

## B

## Bacterial Computing

MARTYN AMOS

Department of Computing and Mathematics,  
Manchester Metropolitan University, Manchester, UK

### Article Outline

[Glossary](#)

[Definition of the Subject](#)

[Introduction](#)

[Motivation for Bacterial Computing](#)

[The Logic of Life](#)

[Rewiring Genetic Circuitry](#)

[Successful Implementations](#)

[Future Directions](#)

[Bibliography](#)

### Glossary

**DNA** Deoxyribonucleic acid. Molecule that encodes the genetic information of cellular organisms.

**Operon** Set of functionally related genes with a common promoter (“on switch”).

**Plasmid** Small circular DNA molecule used to transfer genes from one organism to another.

**RNA** Ribonucleic acid. Molecule similar to DNA, which helps in the conversion of genetic information to proteins.

**Transcription** Conversion of a genetic sequence into RNA.

**Translation** Conversion of an RNA sequence into an amino acid sequence (and, ultimately, a protein).

### Definition of the Subject

Bacterial computing is a conceptual subset of *synthetic biology*, which is itself an emerging scientific discipline largely concerned with the *engineering* of biological systems. The goals of synthetic biology may be loosely partitioned into four sets: (1) To better understand the fun-

damental operation of the biological system being engineered, (2) To extend synthetic chemistry, and create improved systems for the synthesis of molecules, (3) To investigate the “optimization” of existing biological systems for human purposes, (4) To develop and apply rational *engineering* principles to the design and construction of biological systems. It is on these last two goals that we focus in the current article.

The main benefits that may accrue from these studies are both theoretical and practical; the construction and study of synthetic biosystems could improve our quantitative understanding of the fundamental underlying processes, as well as suggesting plausible applications in fields as diverse as pharmaceutical synthesis and delivery, biosensing, tissue engineering, bionanotechnology, biomaterials, energy production and environmental remediation.

### Introduction

Complex natural processes may often be described in terms of networks of computational components, such as Boolean logic gates or artificial neurons. The interaction of biological molecules and the flow of information controlling the development and behavior of organisms is particularly amenable to this approach, and these models are well-established in the biological community. However, only relatively recently have papers appeared proposing the use of such systems to perform useful, *human-defined* tasks. For example, rather than merely using the network analogy as a convenient technique for clarifying our understanding of complex systems, it may now be possible to harness the power of such systems for the purposes of computation.

Despite the relatively recent emergence of biological computing as a distinct research area, the link between biology and computer science is not a new one. Of course, for years biologists have used computers to store and analyze experimental data. Indeed, it is widely accepted that the huge advances of the Human Genome Project (as well

as other genome projects) were only made possible by the powerful computational tools available. Bioinformatics has emerged as “the science of the 21st century”, requiring the contributions of truly interdisciplinary scientists who are equally at home at the lab bench or writing software at the computer.

However, the seeds of the relationship between biology and computer science were sown over fifty years ago, when the latter discipline did not even exist. When, in the 17th century, the French mathematician and philosopher René Descartes declared to Queen Christina of Sweden that animals could be considered a class of machines, she challenged him to demonstrate how a clock could reproduce. Three centuries later in 1951, with the publication of “*The General and Logical Theory of Automata*” [38] John von Neumann showed how a machine could indeed construct a copy of itself. Von Neumann believed that the behavior of *natural* organisms, although orders of magnitude more complex, was similar to that of the most intricate machines of the day. He believed that life was based on *logic*. We now begin to look at how this view of life may be used, not simply as a useful analogy, but as the practical foundation of a whole new engineering discipline.

### Motivation for Bacterial Computing

Here we consider the main motivations behind recent work on bacterial computing (and, more broadly, synthetic biology). Before recombinant DNA technology made it possible to construct new genetic sequences, biologists were restricted to crudely “knocking out” individual genes from an organism’s genome, and then assessing the damage caused (or otherwise). Such knock-outs gradually allowed them to piece together fragments of causality, but the process was very time-consuming and error-prone.

Since the dawn of genetic engineering – with the ability to synthesize novel gene segments – biologists have been in a position to make much more finely-tuned modifications to their organism of choice, this generating much more refined data. Other advances in biochemistry have also contributed, allowing scientists to – for example – investigate new types of genetic systems with, for example, *twelve* bases, rather than the traditional four [14]. Such creations have yielded valuable insights into the mechanics of mutation, adaptation and evolution. Researchers in synthetic biology are now extending their work beyond the synthesis of single genes, and are now introducing whole new gene *complexes* into organisms.

The objectives behind this work are both theoretical and practical. As Benner and Seymour argue [5], “... a synthetic goal forces scientists to cross uncharted ground

to encounter and solve problems that are not easily encountered through [top-down] analysis. This drives the emergence of new paradigms [“world views”] in ways that analysis cannot easily do.” Drew Endy agrees. “Worst-case scenario, it’s a complementary approach to traditional discovery science” [18]. “Best-case scenario, we get systems that are simpler and easier to understand ...” Or, to put it bluntly, “Let’s build new biological systems – systems that are easier to understand because we made them that way” [31]. As well as shedding new light on the underlying biology, these novel systems may well have significant practical utility. Such new creations, according to Endy’s “personal wish list” might include “generating biological machines that could clean up toxic waste, detect chemical weapons, perform simple computations, stalk cancer cells, lay down electronic circuits, synthesize complex compounds and even produce hydrogen from sunlight” [18]. In the next Section we begin to consider how this might be achieved, by first describing the underlying logic of genetic circuitry.

### The Logic of Life

Twenty years after von Neumann’s seminal paper, Francois Jacob and Jacques Monod identified *specific* natural processes that could be viewed as behaving according to logical principles:

“The logic of biological regulatory systems abides not by Hegelian laws but, like the workings of computers, by the propositional algebra of George Boole.” [29]

This conclusion was drawn from earlier work of Jacob and Monod [30]. In addition, Jacob and Monod described the “lactose system” [20], which is one of the archetypal examples of a Boolean biosystem. We describe this system shortly, but first give a brief introduction to the operation of genes in general terms.

### DNA as the Carrier of Genetic Information

The *central dogma* of molecular biology [9] is that DNA produces RNA, which in turn produces proteins. The basic “building blocks” of genetic information are known as *genes*. Each gene codes for one specific *protein* and may be turned on (*expressed*) or off (*repressed*) when required.

### Transcription and Translation

We now describe the processes that determine the structure of a protein, and hence its function. Note that in what follows we assume the processes described occur in bacte-

ria, rather than in higher organisms such as humans. For a full description of the structure of the DNA molecule, see the chapter on *DNA computing*. In order for a DNA sequence to be converted into a protein molecule, it must be read (*transcribed*) and the transcript converted (*translated*) into a protein. Transcription of a gene produces a *messenger RNA* (mRNA) copy, which can then be translated into a protein.

Transcription proceeds as follows. The mRNA copy is synthesized by an enzyme known as *RNA polymerase*. In order to do this, the RNA polymerase must be able to recognize the specific region to be transcribed. This specificity requirement facilitates the regulation of genetic expression, thus preventing the production of unwanted proteins. Transcription begins at specific sites within the DNA sequence, known as *promoters*. These promoters may be thought of as “markers”, or “signs”, in that they are not transcribed into RNA. The regions that *are* transcribed into RNA (and eventually translated into protein) are referred to as *structural* genes. The RNA polymerase recognizes the promoter, and transcription begins. In order for the RNA polymerase to begin transcription, the double helix must be opened so that the sequence of bases may be read. This opening involves the breaking of the hydrogen bonds between bases. The RNA polymerase then moves along the DNA *template* strand in the 3 → 5' direction. As it does so, the polymerase creates an *antiparallel* mRNA chain (that is, the mRNA strand is the equivalent of the Watson-Crick complement of the template). However, there is one significant difference, in that RNA contains uracil instead of thymine. Thus, in mRNA terms, “U binds with A.”

The RNA polymerase moves along the DNA, the DNA re-coiling into its double-helix structure behind it, until it reaches the end of the region to be transcribed. The end of this region is marked by a *terminator* which, like the promoter, is not transcribed.

### Genetic Regulation

Each step of the conversion, from stored information (DNA), through mRNA (messenger), to protein synthesis (effector), is itself catalyzed by other effector molecules. These may be enzymes or other factors that are required for a process to continue (for example, sugars). Consequently, a loop is formed, where products of one gene are required to produce further gene products, and may even influence that gene's own expression. This process was first described by Jacob and Monod in 1961 [20], a discovery that earned them a share of the 1965 Nobel Prize in Physiology or Medicine.

Genes are composed of a number of distinct regions, which control and encode the desired product. These regions are generally of the form promoter–gene–terminator. Transcription may be regulated by effector molecules known as *inducers* and *repressors*, which interact with the promoter and increase or decrease the level of transcription. This allows effective control over the expression of proteins, avoiding the production of unnecessary compounds. It is important to note at this stage that, in reality, genetic regulation does not conform to the digital “on-off” model that is popularly portrayed; rather, it is continuous or analog in nature.

### The *Lac Operon*

One of the most well-studied genetic systems is the *lac operon*. An *operon* is a set of functionally related genes with a common promoter. An example of this is the *lac* operon, which contains three structural genes that allow *E. coli* to utilize the sugar lactose.

When *E. coli* is grown on the sugar glucose, the product of the (separate, and unrelated to the *lac* operon) *lacI* gene *represses* the transcription of the *lacZYA* operon (i. e., the operon is turned off). However, if lactose is supplied *together with* glucose, a lactose by-product is produced which interacts with the repressor molecule, preventing it from repressing the *lacZYA* operon. This de-repression does not itself initiate transcription, since it would be inefficient to utilize lactose if the more common sugar glucose were still available. The operon is *positively* regulated (i. e., “encouraged”) by a different molecule, whose level increases as the amount of available glucose decreases. Therefore, if lactose were present as the sole carbon source, the *lacI* repression would be relaxed and the high “encouraging” levels would activate transcription, leading to the synthesis of the *lacZYA* gene products. Thus, the promoter is under the control of two sugars, and the *lacZYA* operon is only transcribed when lactose is *present* and glucose is *absent*.

In essence, Jacob and Monod showed how a gene may be thought of (in very abstract terms) as a binary switch, and how the state of that switch might be affected by the presence or absence of certain molecules. Monod's point, made in his classic book *Chance and Necessity* and quoted above, was that the workings of biological systems operate not by Hegel's philosophical or metaphysical logic of understanding, but according to the formal, mathematically-grounded logical system of George Boole.

What Jacob and Monod found was that the transcription of a gene may be regulated by molecules known as *inducers* and *repressors*, which either increase or decrease the

“volume” of a gene (corresponding to its level of transcription, which isn’t always as clear cut and binary as Monod’s quote might suggest). These molecules interact with the promoter region of a gene, allowing the gene’s level to be finely “tuned”. The *lac* genes are so-called because, in the *E. coli* bacterium, they combine to produce a variety of proteins that allow the cell to metabolise the sugar lactose (which is most commonly found in milk, hence the derivation from the Latin, *lact*, meaning milk).

For reasons of efficiency, these proteins should only be produced (i.e., the genes be turned on) when lactose is present in the cell’s environment. Making these proteins when lactose is *absent* would be a waste of the cell’s resources, after all. However, a different sugar – glucose – will *always* be preferable to lactose, if the cell can get it, since glucose is an “easier” form of sugar to metabolise. So, the input to and output from the *lac* operon may be expressed as a truth table, with G and L standing for glucose and lactose (1 if present, 0 if absent), and O standing for the output of the operon (1 if on, 0 if off):

G	L	O
0	0	0
0	1	1
1	0	0
1	1	0

The Boolean function that the *lac* operon therefore *physically* computes is  $(L \text{ AND } (\text{NOT } G))$ , since it only outputs 1 if  $L=1$  (lactose present) *and*  $G=0$  (glucose is absent). By showing how one gene could affect the expression of another – just like a transistor feeds into the input of another and affects its state – Jacob and Monod laid the foundations for a new way of thinking about genes; not simply in terms of protein blueprints, but of *circuits* of interacting parts, or dynamic networks of interconnected switches and logic gates. This view of the genome is now well-established [22,23,27], but in the next Section we show how it might be used to guide the *engineering* of biological systems.

### Rewiring Genetic Circuitry

A key difference between the wiring of a computer chip and the circuitry of the cell is that “electronics engineers know exactly how resistors and capacitors are wired to each other because they installed the wiring. But biologists often don’t have a complete picture. They may not know which of thousands of genes and proteins are interacting at a given moment, making it hard to predict how circuits will behave inside cells” [12]. This makes the task of reengineering vastly more complex.

Rather than trying assimilate the huge amounts of data currently being generated by the various genome projects, synthetic biologists are taking a novel route – simplify and build. “They create models of genetic circuits, build the circuits, see if they work, and adjust them if they don’t – learning about biology in the process. ‘I view it as a reductionist approach to systems biology,’ says biomedical engineer Jim Collins of Boston University” [12].

The field of *systems biology* has emerged in recent years as an alternative to the traditional reductionist way of doing science. Rather than simply focussing on a single level of description (such as individual proteins), researchers are now seeking to *integrate* information from many different layers of complexity. By studying how different biological components *interact*, rather than simply looking at their structure, systems biologists are attempting to build models of systems from the bottom up. A model is simply an abstract description of how a system operates – for example, a set of equations that describe how a disease spreads throughout a population. The point of a model is to capture the *essence* of a system’s operation, and it should therefore be as simple as possible. Crucially, a model should also be capable of making *predictions*, which can then be tested against reality using real data (for example, if infected people are placed in quarantine for a week, how does this affect the progress of the disease in question, or what happens if I feed this signal into the chip?). The results obtained from these tests may then feed back in to further refinements of the model, in a continuous cycle of improvement.

When a model suggests a plausible structure for a synthetic genetic circuit, the next stage is to engineer it into the chosen organism, such as a bacterium. Once the structure of the DNA molecule was elucidated and the processes of transcription and translation were understood, molecular biologists were frustrated by the lack of suitable experimental techniques that would facilitate more detailed examination of the genetic material. However, in the early 1970s, several techniques were developed that allowed previously impossible experiments to be carried out (see [8,33]). These techniques quickly led to the first ever successful cloning experiments [19,25]. Cloning is generally defined as “... the production of multiple identical copies of a single gene, cell, virus, or organism.” [35]. This is achieved as follows: a specific sequence (corresponding, perhaps, to a novel gene) is inserted in a circular DNA molecule, known as a *plasmid*, or *vector*, producing a *recombinant DNA molecule*. The vector acts as a *vehicle*, transporting the sequence into a *host* cell (usually a bacterium, such as *E.coli*). Cloning single genes is well-established, but is often done on an ad hoc basis. If biological



computing is to succeed, it requires some degree of *standardization*, in the same way that computer manufacturers build different computers, but using a standard library of components. “Biobricks are the first example of standard biological parts,” explains Drew Endy [21]. “You will be able to use biobricks to program systems that do whatever biological systems do.” He continues. “That way, if in the future, someone asks me to make an organism that, say, counts to 3,000 and then turns left, I can grab the parts I need off the shelf, hook them together and predict how they will perform” [15].

Each biobrick is a simple component, such as an AND gate, or an inverter (NOT). Put them together one after the other, and you have a NAND (NOT-AND) gate, which is all that is needed to build any Boolean circuit (an arbitrary circuit can be translated to an equivalent circuit that uses only NAND gates. It will be much bigger than the original, but it will compute the same function. Such considerations were important in the early stages of integrated circuits, when building *different* logic gates was difficult and expensive). Just as transistors can be used together to build logic gates, and these gates then combined into circuits, there exists a hierarchy of complexity with biobricks. At the bottom are the “parts”, which generally correspond to coding regions for proteins. Then, one level up, we have “devices”, which are built from parts – the oscillator of Elowitz and Leibler, for example, could be constructed from three inverter devices chained together, since all an inverter does is “flip” its signal from 1 to 0 (or vice versa). This circuit would be an example of the biobricks at the top of the conceptual tree – “systems”, which are collections of parts to do a significant task (like oscillating or counting).

Tom Knight at MIT made the first 6 biobricks, each held in a plasmid ready for use. As we have stated, plasmids can be used to insert novel DNA sequences in the genomes of bacteria, which act as the “testbed” for the biobrick circuits. “Just pour the contents of one of these vials into a standard reagent solution, and the DNA will transform itself into a functional component of the bacteria,” he explains [6]. Drew Endy was instrumental in developing this work further, one invaluable resource being the Registry of Standard Biological Parts [32], the definitive catalogue of new biological component. At the start of 2006, it contained 224 “basic” parts and 459 “composite” parts, with 270 parts “under construction”. Biobricks are still at a relatively early stage, but “Eventually we’ll be able to design and build *in silico* and go out and have things synthesized,” says Jay Keasling, head of Lawrence Berkeley National Laboratory’s new synthetic biology department [12].

## Successful Implementations

Although important foundational work had been performed by Arkin and Ross as early as 1994 [2], the year 2000 was a particularly significant one for synthetic biology. In January two foundational papers appeared back-to-back in the same issue of *Nature*. “Networks of interacting biomolecules carry out many essential functions in living cells, but the ‘design principles’ underlying the functioning of such intracellular networks remain poorly understood, despite intensive efforts including quantitative analysis of relatively simple systems. Here we present a complementary approach to this problem: the design and construction of a synthetic network to implement a particular function” [11]. That was the introduction to a paper that Drew Endy would call “the high-water mark of a synthetic genetic circuit that does something” [12]. In the first of the two articles, Michael Elowitz and Stanislaw Leibler (then both at Princeton) showed how to build microscopic “Christmas tree lights” using bacteria.

### Synthetic Oscillator

In physics, an *oscillator* is a system that produces a regular, periodic “output”. Familiar examples include a pendulum, a vibrating string, or a lighthouse. Linking several oscillators together in some way gives rise to *synchrony* – for example, heart cells repeatedly firing in unison, or millions of fireflies blinking on and off, seemingly as one [36].

Leibler actually had *two* articles published in the same high-impact issue of *Nature*. The other was a short communication, co-authored with Naama Barkai – also at Princeton, but in the department of Physics [3]. In their paper, titled “Circadian clocks limited by noise”, Leibler and Barkai showed how a simple model of biochemical networks could oscillate reliably, even in the presence of noise. They argued that such oscillations (which might, for example, control the internal circadian clock that tells us when to wake up and when to be tired) are based on networks of *genetic regulation*. They built a simulation of a simple regulatory network using a *Monte Carlo* algorithm. They found that, however they perturbed the system, it still oscillated reliably, although, at the time, their results existed only *in silico*. The other paper by Leibler was much more applied, in the sense that they had constructed a biological circuit [11]. Elowitz and Leibler had succeeded in constructing an artificial genetic oscillator in cells, using a synthetic network of repressors. They called this construction the *repressilator*.

Rather than investigating *existing* oscillator networks, Elowitz and Leibler decided to build one entirely from first principles. They chose three repressor-promoter pairs that

had already been sequenced and characterized, and first built a mathematical model in software. By running the sets of equations, they identified from their simulation results certain molecular characteristics of the components that gave rise to so-called *limit cycle* oscillations; those that are robust to perturbations. This information from the model's results lead Elowitz and Leibler to select strong promoter molecules and repressor molecules that would rapidly decay. In order to implement the oscillation, they chose three genes, each of which affected one of the others by repressing it, or turning it off. For the sake of illustration, we call the genes A, B and C. The product of gene A turns off (represses) gene B. The absence of B (which represses C) allows C to turn on. C is chosen such that it turns gene A *off* again, and the three genes loop continuously in a “daisy chain” effect, turning on and off in a repetitive cycle. However, some form of *reporting* is necessary in order to confirm that the oscillation is occurring as planned.

Green fluorescent protein (GFP) is a molecule found occurring naturally in the jellyfish *Aequorea victoria*. Biologists find it invaluable because it has one interesting property – when placed under ultraviolet light, it *glows*. Biologists quickly sequenced the gene responsible for producing this protein, as they realized that it could have many applications as a *reporter*. By inserting the gene into an organism, you have a ready-made “status light” – when placed into bacteria, they glow brightly if the gene is turned on, and look normal if it's turned off. We can think of it in terms of a Boolean circuit – if the circuit outputs the value 1, the GFP promoter is produced to turn on the light. If the value is 0, the promoter isn't produced, the GFP gene isn't expressed, and the light stays off.

Elowitz and Leibler set up their gene network so that the GFP gene would be expressed whenever gene C was turned *off* – when it was turned on, the GFP would gradually decay and fade away. They synthesized the appropriate DNA sequences and inserted them into a *plasmid*, eventually yielding a population of bacteria that blinked on and off in a repetitive cycle, like miniature lighthouses. Moreover, and perhaps most significantly, the period between flashes was longer than the time taken for the cells to divide, showing that the state of the system had been passed on during reproduction.

### Synthetic Toggle Switch

Rather than model an existing circuit and then altering it, Elowitz and Leibler had taken a “bottom up” approach to learning about how gene circuits operate. The other notable paper to appear in that issue was written by Timo-

thy Gardner, Charles Cantor and Jim Collins, all of Boston University in the US. In 2000, Gardner, Collins and Cantor observed that genetic switching (such as that observed in the lambda phage [34]) had not yet been “demonstrated in networks of non-specialized regulatory components” [13]. That is to say, at that point nobody had been able to construct a switch out of genes that hadn't already been “designed” by evolution to perform that specific task. The team had a similar philosophy to that of Elowitz and Leibler, in that their main motivation was being able to test theories about the fundamental behaviour of gene regulatory networks. “Owing to the difficulty of testing their predictions,” they explained, “these theories have not, in general, been verified experimentally. Here we have integrated theory and experiment by constructing and testing a synthetic, bistable [two-state] gene circuit based on the predictions of a simple mathematical model.”

“We were looking for the equivalent of a light switch to flip processes on or off in the cell,” explained Gardner [10]. “Then I realized a way to do this with genes instead of with electric circuits.” The team chose two genes that were mutually inhibitory – that is, each produced a molecule that would turn the other off. One important thing to bear in mind is that the system didn't have a single input. Although the team acknowledged that bistability might be possible – in theory – using only a single promoter that regulated *itself*, they anticipated possible problems with robustness and experimental tunability if they used that approach. Instead, they decided to use a system whereby each “side” of the switch could be “pressed” by a different stimulus – the addition of a chemical on one, and a change in temperature on the other. Things were set up so that if the system was in the state induced by the chemical, it would stay in that state until the temperature was changed, and would only change *back* again if the chemical was reintroduced. Importantly, these stimuli did not have to be applied continuously – a “short sharp” burst was enough to cause the switch to flip over. As with the other experiment, Gardner and his colleagues used GFP as the system state reporter, so that the cells glowed in one state, and looked “normal” in the other.

In line with the bottom-up approach, they first created a mathematical model of the system and made some predictions about how it would behave inside a cell. Within a year, Gardner had spliced the appropriate genes into the bacteria, and he was able to flip them –at will – from one state the the other. As McAdams and Arkin observed, synthetic “one way” switches had been created in the mid-1980s, but “this is perhaps the first engineered design exploiting bistability to produce a switch with capability of *reversibly* switching between two ... stable states” [28].

The potential applications of such a bacterial switch were clear. As they state in the conclusion to their article, “As a practical device, the toggle switch ... may find applications in gene therapy and biotechnology.” They also borrowed from the language of computer programming, using an analogy between their construction and the short “applets” written in the Java language, which now allow us to download and run programs in our web browser. “Finally, as a cellular memory unit, the toggle forms the basis for ‘genetic applets’ – self-contained, programmable, synthetic gene circuits for the control of cell function”.

### Engineered Communication

Towards the end of his life, Alan Turing did some foundational work on pattern formation in nature, in an attempt to explain how zebras get their striped coats or leopards their spots. The study of *morphogenesis* (from the Greek, *morphe* – shape, and *genesis* – creation. “Amorphous” therefore means “without shape or structure”) is concerned with how cells split to assume new roles and communicate with another to form very precise shapes, such as tissues and organs. Turing postulated that the diffusion of chemical signals both within and between cells is the main driving force behind such complex pattern formation [37].

Although Turing’s work was mainly concerned with the processes occurring amongst cells inside a developing embryo, it is clear that chemical signalling also goes on between bacteria. Ron Weiss of Princeton University was particularly interested in *Vibrio fischeri*, a bacterium that has a symbiotic relationship with a variety of aquatic creatures, including the Hawaiian squid. This relationship is due mainly to the fact that the bacteria exhibit *bioluminescence* – they generate a chemical known as a *Luciferase* (coded by the *Lux* gene), a version of which is also found in fireflies, and which causes them to glow when gathered together in numbers. Cells within the primitive light organs of the squid draw in bacteria from the seawater and encourage them to grow. Crucially, once enough bacteria are packed into the light organ they produce a signal to tell the squid cells to stop attracting their colleagues, and only then do they begin to glow. The cells get a safe environment in which to grow, protected from competition, and the squid has a light source by which to navigate and catch prey. The mechanism by which the *Vibrio* “know” when to start glowing is known as *quorum sensing*, since there have to be sufficient “members” present for luminiscence to occur.

The bacteria secrete an *autoinducer* molecule, known as VAI (*Vibrio* Auto Inducer), which diffuses through the cell wall. The *Lux* gene (which generates the glowing

chemical) needs to be activated (turned on) by a particular protein – which attracts the attention of the polymerase – but the protein can only do this with help from the VAI. Its particular 3D structure is such that it can’t fit tightly onto the gene unless it’s been slightly bent out of shape. Where there’s enough VAI present, it locks onto the protein and alters its conformation, so that it can turn on the gene. Thus, the concentration of VAI is absolutely crucial; once a critical threshold has been passed, the bacteria “know” that there are enough of them present, and they begin to glow.

Weiss realized that this quorum-based cell-to-cell communication mechanism could provide a powerful framework for the construction of bacterial devices – imagine, for example, a tube of solution containing engineered bacteria that can be added to a sample of seawater, causing it to glow only if the concentration of a particular pollutant exceeds a certain threshold. Crucially, as we will see shortly, it also allows the possibility of generating precise “complex patterned” development.

Weiss set up two colonies of *E. coli*, one containing “sender”, and the other “receivers”. The idea was that the senders would generate a chemical signal made up of VAI, which could diffuse across a gap and then be picked up by the receivers. Once a strong enough signal was being communicated, the receivers would glow using GFP to say that it had been picked up. Weiss cloned the appropriate gene sequences (corresponding to a type of biobrick) into his bacteria, placed colonies of receiver cells on a plate, and the receivers started to glow in acknowledgment.

### Synthetic Circuit Evolution

In late 2002, Weiss and his colleagues published another paper, this time describing how rigorous engineering principles may be brought to bear on the problem of designing and building entirely new genetic circuitry. The motivation was clear – “biological circuit engineers will have to confront their inability to predict the precise behavior of even the most simple synthetic networks, a serious shortcoming and challenge for the design and construction of more sophisticated genetic circuitry in the future” [39].

Together with colleagues Yohei Yokobayashi and Frances Arnold, Weiss proposed a two stage strategy: first, design a circuit from the bottom up, as Elowitz and others had before, and clone it into bacteria. Such circuits are highly unlikely to work first time, “because the behavior of biological components inside living cells is highly context-dependent, the actual circuit performance will likely differ from the design predictions, often resulting in a poorly performing or nonfunctional circuit.” Rather than simply

abandoning their design, Weiss and his team decided to then *tune* the circuit inside the cell itself, by applying the principles of evolution. By inducing mutations in the DNA that they had just introduced, they were able to slightly modify the behaviour of the circuit that it represented. Of course, many of these changes would be catastrophic, giving even worse performance than before, but, occasionally, they observed a minor improvement. In that case, they kept the “winning” bacteria, and subjected it to another round of mutation, in a repeated cycle. In a microcosmic version of Darwinian evolution, mutation followed by selection of the fittest took an initially unpromising pool of broken circuits and transformed them into winners. “Ron is utilizing the power of evolution to design networks in ways so that they perform exactly the way you want them to,” observed Jim Collins [16]. In a commentary article in the same issue of the journal, Jeff Hasty called this approach “design then mutate” [17]. The team showed how a circuit made up of three genetic gates could be fine-tuned *in vivo* to give the correct performance, and they concluded that “the approach we have outlined should serve as a robust and widely applicable route to obtaining circuits, as well as new genetic devices, that function inside living cells.”

### Pattern Formation

The next topic studied by Weiss and his team was the problem of *space* – specifically, how to get a population of bacteria to cover a surface with a specific density. This facility could be useful when designing bacterial biosensors – devices that detect chemicals in the environment and produce a response. By controlling the density of the microbial components, it might be possible to tune the sensitivity of the overall device. More importantly, the ability for cells to control their own density would provide a useful “self-destruct” mechanism were these genetically-modified bugs ever to be released into the environment for “real world” applications.

In “Programmed population control by cell-cell communication and regulated killing” [40], Weiss and his team built on their previous results to demonstrate the ability to keep the density of an *E. coli* population artificially low – that is, below the “natural” density that could be supported by the available nutrients. They designed a genetic circuit that caused the bacteria to generate a different *Vibrio* signalling molecule, only this time, instead of making the cells glow, a sufficient concentration would flip a switch inside the cell, turning on a *killer gene*, encoding a protein that was toxic in sufficient quantities. The system behaved exactly as predicted by their mathematical model.

The culture grew at an exponential rate (that is, doubling every time step) for seven hours, before hitting the defined density threshold. At that point the population dropped sharply, as countless cells expired, until the population settled at a steady density significantly (ten times) lower than an unmodified “control” colony. The team concluded that “The population-control circuit lays the foundations for using cell-cell communication to programme interactions among bacterial colonies, allowing the concept of communication-regulated growth and death to be extended to engineering synthetic ecosystems” [40].

The next stage was to programme cells to form *specific* spatial patterns in the dish. As we have already mentioned briefly, pattern formation is one characteristic of multicellular systems. This is generally achieved using some form of chemical signalling, combined with a differential response – that is, different cells, although genetically identical, may “read” the environmental signals and react in different ways, depending on their internal state. For example, one cell might be “hungry” and choose to move towards a food source, while an identical cell might choose to remain in the same spot, since it has adequate levels of energy.

The team used a variant of the sender-receiver model, only this time adding a “distance detection” component to the receiver circuit. The senders were placed in the center of the dish, and the receivers distributed uniformly across the surface. The receivers constructed so that they could measure the *strength* of the signal being “beamed” from the senders, a signal which decayed over distance (a little like a radio station gradually breaking up as you move out of the reception area). The cells were engineered so that only those that were either “near” to the senders or “far” from the senders would generate a response (those in the middle region were instructed to keep quiet). These cells are genetically identical, and are uniformly distributed over the surface – the differential response comes in the way that they assess the strength of the signal, and make a decision on whether or not to respond. The power of the system was increased further by making the near cells glow green, and those far away glow red (using a different fluorescent protein).

When the team set the system running, they observed the formation of a “dartboard” pattern, with the “bullseye” being the colony of senders (instructed to glow cyan, or light blue), which was surrounded by a green ring, which in turn was surrounded by a red ring. By placing three sender colonies in a triangle, they were also able to obtain a green heart-shaped pattern, formed by the intersection of three green circles, as well as other patterns, determined solely by the initial configuration of senders [4].



## Bacterial Camera

Rather than generating light, a different team decided to use bacteria to *detect* light – in the process, building the world's first microbial camera. By engineering a dense bed of *E. coli*, a team of students led by Chris Voight at Berkeley developed light-sensitive “film” capable of storing images at a resolution of 100 megapixels per square inch. *E. coli* are not normally sensitive to light, so the group took genes coding for *photoreceptors* from blue-green algae, and spliced them into their bugs [24]. When light was shone on the cells, it turned on a genetic switch that cause a chemical inside them to permanently darken, thus generating a black “pixel”. By projecting an image onto a plate of bacteria, the team were able to obtain several monochrome images, including the *Nature* logo and the face of team member Andrew Ellington. Nobel Laureate Sir Harry Kroto, discoverer of “buckballs”, called the team's camera an “extremely exciting advance” [26], going on to say that “I have always thought that the first major nanotechnology advances would involve some sort of chemical modification of biology.”

## Future Directions

Weiss and his team suggest that “the integration of such systems into higher-level organisms and with different cell functions will have practical applications in three-dimensional tissue engineering, biosensing, and biomaterial fabrication.” [4] One possible use for such a system might lie in the detection of bio-weapons – spread a culture of bacteria over a surface, and, with the appropriate control circuit, they will be able to accurately pinpoint the location of any pathogens. Programmed cells could eventually replace artificial tissues, or even organs – current attempts to build such constructions in the laboratory rely on cells arranging themselves around an artificial scaffold. Controlled cellular structure formation could do away with the need for such support – “The way we're doing tissue engineering, right now, ... is very unnatural,” argues Weiss. “Clearly cells make scaffolds themselves. If we're able to program them to do that, we might be able to embed them in the site of injury and have them figure out for themselves what the pattern should be” [7]. In addition to building structures, others are considering engineering cells to act as miniature drug delivery systems – fighting disease or infection from the *inside*. Adam Arkin and Chris Voight are currently investigating the use of modified *E. coli* to battle against cancer tumours, while Jay Keasling and co-workers at Berkeley are looking at engineering circuits into the same bacteria to persuade them to

generate a potent antimalarial drug that is normally found in small amounts in wormwood plants.

Clearly, bacterial computing/synthetic biology is still at a relatively early stage in its development, although the field is growing at a tremendous pace. It could be argued, with some justification, that the dominant science of the new millennium may well prove to be at the intersection of biology and computing. As biologist Roger Brent argues, “I think that synthetic biology ... will be as important to the 21st century as [the] ability to manipulate bits was to the 20th” [1].

## Bibliography

### Primary Literature

1. Anon (2004) Roger Brent and the alpha project. *ACM Ubiquity*, 5(3)
2. Arkin A, Ross J (1994) Computational functions in biochemical reaction networks. *Biophysical J* 67:560–578
3. Barkai N, Leibler S (2000) Circadian clocks limited by noise. *Nature* 403:267–268
4. Basu S, Gerchman Y, Collins CH, Arnold FH, Weiss R (2005) A synthetic multicellular system for programmed pattern formation. *Nature* 434:1130–1134
5. Benner SA, Sismour M (2005) Synthetic biology. *Nature Rev Genet* 6:533–543
6. Brown C (2004) BioBricks to help reverse-engineer life. *EE Times*, June 11
7. Brown S (2005) Command performances. *San Diego Union-Tribune*, December 14
8. Brown TA (1990) Gene cloning: an introduction, 2nd edn. Chapman and Hall, London
9. Crick F (1970) Central dogma of molecular biology. *Nature* 227:561–563
10. Eisenberg A (2000) Unlike viruses, bacteria find a welcome in the world of computing. *New York Times*, June 1
11. Elowitz M, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* 403:335–338
12. Ferber D (2004) Synthetic biology: microbes made to order. *Science* 303(5655):158–161
13. Gardner T, Cantor R, Collins J (2000) Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403:339–342
14. Geyer CR, Battersby TR, Benner SA (2003) Nucleobase pairing in expanded Watson-Crick-like genetic information systems. *Structure* 11:1485–1498
15. Gibbs WW (2004) Synthetic life. *Scientific Am* April 26
16. Gravit L (2004) 10 emerging technologies that will change your world. *MIT Technol Rev* February
17. Hasty J (2002) Design then mutate. *Proc Natl Acad Sci* 99(26):16516–16518
18. Hopkin K (2004) Life: the next generation. *The Scientist* 18(19):56
19. Jackson DA, Symons RH, Berg P (1972) Biochemical method for inserting new genetic information into DNA of simian virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli*. *Proc Natl Acad Sci* 69:2904–2909



20. Jacob F, Monod J (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* 3:318–356
21. Jha A (2005) From the cells up. *The Guardian*, March 10
22. Kauffman S (1993) Gene regulation networks: a theory for their global structure and behaviors. *Current topics in developmental biology* 6:145–182
23. Kauffman SA (1993) *The origins of order: Self-organization and selection in evolution*. Oxford University Press, New York
24. Levskaya A, Chevalier AA, Tabor JJ, Simpson ZB, Lavery LA, Levy M, Davidson EA, Scouras A, Ellington AD, Marcotte EM, Voight CA (2005) Engineering *Escherichia coli* to see light. *Nature* 438:441–442
25. Lobban PE, Sutton CA (1973) Enzymatic end-to-end joining of DNA molecules. *J Mol Biol* 78(3):453–471
26. Marks P (2005) For ultrasharp pictures, use a living camera. *New Scientist*, November 26, p 28
27. McAdams HH, Shapiro L (1995) Circuit simulation of genetic networks. *Science* 269(5224):650–656
28. McAdams HH, Arkin A (2000) Genetic regulatory circuits: Advances toward a genetic circuit engineering discipline. *Current Biol* 10:318–320
29. Monod J (1970) *Chance and Necessity*. Penguin, London
30. Monod J, Changeux JP, Jacob F (1963) Allosteric proteins and cellular control systems. *J Mol Biol* 6:306–329
31. Morton O (2005) Life, Reinvented. *Wired* 13(1), January
32. Registry of Standard Biological Parts. <http://parts.mit.edu/>
33. Old R, Primrose S (1994) *Principles of Gene Manipulation, an Introduction to Genetic Engineering*, 5th edn. Blackwell, Boston
34. Ptashne M (2004) *A Genetic Switch*, 3rd edn. Phage Lambda Revisited. Cold Spring Harbor Laboratory Press, Woodbury
35. Roberts L, Murrell C (eds) (1998) *An introduction to genetic engineering*. Department of Biological Sciences, University of Warwick
36. Strogatz S (2003) *Sync: The Emerging Science of Spontaneous Order*. Penguin, London
37. Turing AM (1952) The chemical basis of morphogenesis. *Phil Trans Roy Soc B* 237:37–72
38. von Neumann J (1941) The general and logical theory of automata. In: *Cerebral Mechanisms in Behavior*. Wiley, New York
39. Yokobayashi Y, Weiss R, Arnold FH (2002) Directed evolution of a genetic circuit. *Proc Natl Acad Sci* 99(26):16587–16591
40. You L, Cox III RS, Weiss R, Arnold FH (2004) Programmed population control by cell-cell communication and regulated killing. *Nature* 428:868–871
- Sayler GS, Simpson ML, Cox CD (2004) Emerging foundations: nano-engineering and bio-microelectronics for environmental biotechnology. *Curr Opin Microbiol* 7:267–273

## Bayesian Games: Games with Incomplete Information

SHMUEL ZAMIR

Center for the Study of Rationality, Hebrew University, Jerusalem, Israel

### Article Outline

[Glossary](#)

[Definition of the Subject](#)

[Introduction](#)

[Harsanyi's Model: The Notion of Type](#)

[Aumann's Model](#)

[Harsanyi's Model and Hierarchies of Beliefs](#)

[The Universal Belief Space](#)

[Belief Subspaces](#)

[Consistent Beliefs and Common Priors](#)

[Bayesian Games and Bayesian Equilibrium](#)

[Bayesian Equilibrium and Correlated Equilibrium](#)

[Concluding Remarks and Future Directions](#)

[Acknowledgments](#)

[Bibliography](#)

### Glossary

**Bayesian game** An interactive decision situation involving several decision makers (players) in which each player has beliefs about (i. e. assigns probability distribution to) the payoff relevant parameters and the beliefs of the other players.

**State of nature** Payoff relevant data of the game such as payoff functions, value of a random variable, etc. It is convenient to think of a state of nature as a full description of a 'game-form' (actions and payoff functions).

**Type** Also known as *state of mind*, is a full description of player's beliefs (about the state of nature), beliefs about beliefs of the other players, beliefs about the beliefs about his beliefs, etc. ad infinitum.

**State of the world** A specification of the state of nature (payoff relevant parameters) and the players' types (belief of all levels). That is, a state of the world is a state of nature and a list of the states of mind of all players.

**Common prior and consistent beliefs** The beliefs of players in a game with incomplete information are said to be consistent if they are derived from the same

### Books and Reviews

- Alon U (2006) *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman and Hall/CRC
- Amos M (ed) (2004) *Cellular Computing*. Series in Systems Biology Oxford University Press
- Amos M (2006) *Genesis Machines: The New Science of Biocomputing*. Atlantic Books, London
- Benner SA (2003) Synthetic biology: Act natural. *Nature* 421:118
- Endy D (2005) Foundations for engineering biology. *Nature* 436:449–453
- Kobayashi H, Kaern M, Araki M, Chung K, Gardner TS, Cantor CR, Collins JJ (2004) Programmable cells: interfacing natural and engineered gene networks. *Proc Natl Acad Sci* 101(22):8414–8419

probability distribution (the common prior) by conditioning on each player's private information. In other words, if the beliefs are consistent, the only source of differences in beliefs is difference in information.

**Bayesian equilibrium** A Nash equilibrium of a Bayesian game: A list of behavior and beliefs such that each player is doing his best to maximize his payoff, according to his beliefs about the behavior of the other players.

**Correlated equilibrium** A Nash equilibrium in an extension of the game in which there is a chance move, and each player has only partial information about its outcome.

### Definition of the Subject

Bayesian games (also known as *Games with Incomplete Information*) are models of interactive decision situations in which the decision makers (players) have only partial information about the data of the game and about the other players. Clearly this is typically the situation we are facing and hence the importance of the subject: The basic underlying assumption of classical game theory according to which the data of the game is *common knowledge* (CK) among the players, is too strong and often implausible in real situations. The importance of Bayesian games is in providing the tools and methodology to relax this implausible assumption, to enable modeling of the overwhelming majority of real-life situations in which players have only partial information about the payoff relevant data. As a result of the interactive nature of the situation, this methodology turns out to be rather deep and sophisticated, both conceptually and mathematically: Adopting the classical Bayesian approach of statistics, we encounter the need to deal with an *infinite hierarchy of beliefs*: what does each player believe that the other player believes about what he believes... is the actual payoff associated with a certain outcome? It is not surprising that this methodological difficulty was a major obstacle in the development of the theory, and this article is largely devoted to explaining and resolving this methodological difficulty.

### Introduction

A game is a mathematical model for an interactive decision situation involving several decision makers (players) whose decisions affect each other. A basic, often implicit, assumption is that the data of the game, which we call the *state of nature*, are *common knowledge* (CK) among the players. In particular the actions available to the players and the payoff functions are CK. This is a rather strong assumption that says that every player knows all actions

and payoff functions of all players, every player knows that all other players know all actions and payoff functions, every player knows that every player knows that every player knows... etc. ad infinitum. *Bayesian games* (also known as *games with incomplete information*), which is the subject of this article, are models of interactive decision situations in which each player has only partial information about the payoff relevant parameters of the given situation.

Adopting the Bayesian approach, we assume that a player who has only partial knowledge about the state of nature has some *beliefs*, namely prior distribution, about the parameters which he does not know or he is uncertain about. However, unlike in a statistical problem which involves a single decision maker, this is not enough in an interactive situation: As the decisions of other players are relevant, so are their beliefs, since they affect their decisions. Thus a player must have beliefs about the beliefs of other players. For the same reason, a player needs beliefs about the beliefs of other players about his beliefs and so on. This interactive reasoning about beliefs leads unavoidably to *infinite hierarchies of beliefs* which looks rather intractable. The natural emergence of hierarchies of beliefs is illustrated in the following example:

*Example 1* Two players, P1 and P2, play a  $2 \times 2$  game whose payoffs depend on an unknown state of nature  $s \in \{1, 2\}$ . Player P1's actions are  $\{T, B\}$ , player P2's actions are  $\{L, R\}$  and the payoffs are given in the following matrices:

		P2	
		L	R
P1	T	0, 1	1, 0
	B	1, 0	0, 1

a Payoffs when  $s = 1$

		P2	
		L	R
P1	T	1, 0	0, 1
	B	0, 1	1, 0

b Payoffs when  $s = 2$

Assume that the belief (prior) of P1 about the event  $\{s = 1\}$  is  $p$  and the belief of P2 about the same event is  $q$ . The best action of P1 depends both on his prior and on the action of P2, and similarly for the best action of P2. This is given in the following tables:

		P2's action		
		$L$	$R$	
$p < 0.5$		$T$	$B$	
		$B$	$T$	
		Best reply of P1		
		$q < 0.5$	$q > 0.5$	
P1's action	$T$	$R$	$L$	
	$B$	$L$	$R$	
		Best reply of P2		

Now, since the optimal action of P1 depends not only on his belief  $p$  but also on the, unknown to him, action of P2, which depends on his belief  $q$ , player P1 must therefore have beliefs about  $q$ . These are his *second-level beliefs*, namely beliefs about beliefs. But then, since this is relevant and unknown to P2, he must have beliefs about that which will be *third-level beliefs* of P2, and so on. The whole infinite hierarchies of beliefs of the two players pop out naturally in the analysis of this simple two-person game of incomplete information.

The objective of this article is to model this kind of situation. Most of the effort will be devoted to the modeling of the mutual beliefs structure and only then we add the underlying game which, together with the beliefs structure, defines a *Bayesian game* for which we define the notion of *Bayesian equilibrium*.

### Harsanyi's Model: The Notion of Type

As suggested by our introductory example, the straightforward way to describe the mutual beliefs structure in a situation of incomplete information is to specify explicitly the whole hierarchies of beliefs of the players, that is, the beliefs of each player about the unknown parameters of the game, each player's beliefs about the other players' beliefs about these parameters, each player's beliefs about the other players' beliefs about his beliefs about the parameters, and so on ad infinitum. This may be called the *explicit* approach and is in fact feasible and was explored and developed at a later stage of the theory (see [18,5,6,7]). We will come back to it when we discuss the universal belief space. However, for obvious reasons, the explicit approach is mathematically rather cumbersome and hardly manageable. Indeed this was a major obstacle to the development of the theory of games with incomplete information at its early stages. The breakthrough was provided by John Harsanyi [11] in a seminal work that earned him the

Nobel Prize some thirty years later. While Harsanyi actually formulated the problem verbally, in an explicit way, he suggested a solution that 'avoided' the difficulty of having to deal with infinite hierarchies of beliefs, by providing a much more workable *implicit*, encapsulated model which we present now.

The key notion in Harsanyi's model is that of *type*. Each player can be of several types where a type is to be thought of as a full description of the player's beliefs about the *state of nature* (the data of the game), beliefs about the beliefs of other players about the state of nature and about his own beliefs, etc. One may think of a player's type as his *state of mind*; a specific configuration of his brain that contains an answer to any question regarding beliefs about the state of nature and about the types of the other players. Note that this implies self-reference (of a type to itself through the types of other players) which is unavoidable in an interactive decision situation. A Harsanyi game of incomplete information consists of the following ingredients (to simplify notations, assume all sets to be finite):

- $I$  – Player's set.
- $S$  – The set of states of nature.
- $T_i$  – The type set of player  $i \in I$ .  
Let  $T = \times_{i \in I} T_i$  – denote the *type set*, that is, the set type profiles.
- $Y \subset S \times T$  – a set of *states of the world*.
- $p \in \Delta(Y)$  – probability distribution on  $Y$ , called the *common prior*.

(For a set  $A$ , we denote the set of probability distributions on  $A$  by  $\Delta(A)$ .)

*Remark* A state of the world  $\omega$  thus consists of a state of nature and a list of the types of the players. We denote it as

$$\omega = (s(\omega); t_1(\omega), \dots, t_n(\omega)) .$$

We think of the state of nature as a full description of the game which we call a *game-form*. So, if it is a game in strategic form, we write the state of nature at state of the world  $\omega$  as:

$$s(\omega) = (I, (A_i(\omega))_{i \in I}, (u_i(\cdot; \omega))_{i \in I}) .$$

The payoff functions  $u_i$  depend only on the state of nature and not on the types. That is, for all  $i \in I$ :

$$s(\omega) = s(\omega') \Rightarrow u_i(\cdot; \omega) = u_i(\cdot; \omega') .$$

The game with incomplete information is played as follows:

- (1) A chance move chooses  $\omega = (s(\omega); t_1(\omega), \dots, t_n(\omega)) \in Y$  using the probability distribution  $p$ .
- (2) Each player is told his chosen type  $t_i(\omega)$  (but not the chosen state of nature  $s(\omega)$  and not the other players' types  $t_{-i}(\omega) = (t_j(\omega))_{j \neq i}$ ).
- (3) The players choose simultaneously an action: player  $i$  chooses  $a_i \in A_i(\omega)$  and receives a payoff  $u_i(a; \omega)$  where  $a = (a_1, \dots, a_n)$  is the vector of chosen actions and  $\omega$  is the state of the world chosen by the chance move.

*Remark* The set  $A_i(\omega)$  of actions available to player  $i$  in state of the world  $\omega$  must be known to him. Since his only information is his type  $t_i(\omega)$ , we must impose that  $A_i(\omega)$  is  $T_i$ -measurable, i. e.,

$$t_i(\omega) = t_i(\omega') \Rightarrow A_i(\omega) = A_i(\omega').$$

Note that if  $s(\omega)$  was commonly known among the players, it would be a regular game in strategic form. We use the term 'game-form' to indicate that the players have only partial information about  $s(\omega)$ . The players do not know which  $s(\omega)$  is being played. In other words, in the extensive form game of Harsanyi, the game-forms  $(s(\omega))_{\omega \in Y}$  are not subgames since they are interconnected by information sets: Player  $i$  does not know which  $s(\omega)$  is being played since he does not know  $\omega$ ; he knows only his own type  $t_i(\omega)$ .

An important application of Harsanyi's model is made in *auction theory*, as an auction is a clear situation of incomplete information. For example, in a closed private-value auction of a single indivisible object, the type of a player is his private-value for the object, which is typically known to him and not to other players. We come back to this in the section entitled "Examples of Bayesian Equilibria".

### Aumann's Model

A frequently used model of incomplete information was given by Aumann [2].

**Definition 2** An Aumann model of incomplete information is  $(I, Y, (\pi_i)_{i \in I}, P)$  where:

- $I$  is the players' set.
- $Y$  is a (finite) set whose elements are called *states of the world*.
- For  $i \in I$ ,  $\pi_i$  is a partition of  $Y$ .
- $P$  is a probability distribution on  $Y$ , also called the *common prior*.

In this model a state of the world  $\omega \in Y$  is chosen according to the probability distribution  $P$ , and each player  $i$  is informed of  $\pi_i(\omega)$ , the element of his partition that contains the chosen state of the world  $\omega$ . This is the informational structure which becomes a game with incomplete information if we add a mapping  $s: Y \rightarrow S$ . The state of nature  $s(\omega)$  is the game-form corresponding to the state of the world  $\omega$  (with the requirement that the action sets  $A_i(\omega)$  are  $\pi_i$ -measurable).

It is readily seen that Aumann's model is a Harsanyi model in which the type set  $T_i$  of player  $i$  is the set of his partition elements, i. e.,  $T_i = \{\pi_i(\omega) | \omega \in Y\}$ , and the common prior on  $Y$  is  $P$ . Conversely, any Harsanyi model is an Aumann model in which the partitions are those defined by the types, i. e.,  $\pi_i(\omega) = \{\omega' \in Y | t_i(\omega') = t_i(\omega)\}$ .

### Harsanyi's Model and Hierarchies of Beliefs

As our starting point in modeling incomplete information situations was the appearance of hierarchies of beliefs, one may ask how is the Harsanyi (or Aumann) model related to hierarchies of beliefs and how does it capture this unavoidable feature of incomplete information situations? The main observation towards answering this question is the following:

**Proposition 3** Any state of the world in Aumann's model or any type profile  $t \in T$  in Harsanyi's model defines (uniquely) a hierarchy of mutual beliefs among the players.

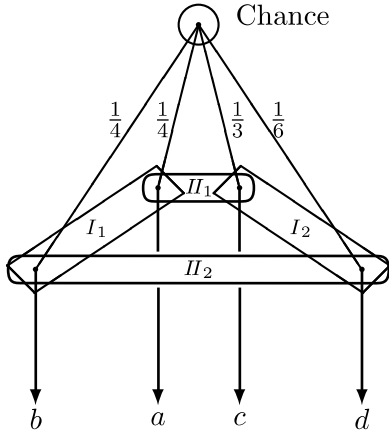
Let us illustrate the idea of the proof by the following example:

*Example* Consider a Harsanyi model with two players,  $I$  and  $II$ , each of which can be of two types:  $T_I = \{I_1, I_2\}$ ,  $T_{II} = \{II_1, II_2\}$  and thus:  $T = \{(I_1, II_1), (I_1, II_2), (I_2, II_1), (I_2, II_2)\}$ . The probability  $p$  on types is given by:

	$II_1$	$II_2$
$I_1$	$\frac{1}{4}$	$\frac{1}{4}$
$I_2$	$\frac{1}{3}$	$\frac{1}{6}$

Denote the corresponding states of nature by  $a = s(I_1 II_1)$ ,  $b = s(I_1 II_2)$ ,  $c = s(I_2 II_1)$  and  $d = s(I_2 II_2)$ . These are the *states of nature* about which there is incomplete information.

The game in extensive form:



Assume that the state of nature is  $a$ . What are the belief hierarchies of the players?

	$II_1$	$II_2$
$I_1$	$\begin{matrix} a \\ \frac{1}{4} \end{matrix}$	$\begin{matrix} b \\ \frac{1}{4} \end{matrix}$
$I_2$	$\begin{matrix} c \\ \frac{1}{3} \end{matrix}$	$\begin{matrix} d \\ \frac{1}{6} \end{matrix}$

First-level beliefs are obtained by each player from  $p$ , by conditioning on his type:

- $I_1$ : With probability  $\frac{1}{2}$  the state is  $a$  and with probability  $\frac{1}{2}$  the state is  $b$ .
- $I_2$ : With probability  $\frac{2}{3}$  the state is  $c$  and with probability  $\frac{1}{3}$  the state is  $d$ .
- $II_1$ : With probability  $\frac{3}{7}$  the state is  $a$  and with probability  $\frac{4}{7}$  the state is  $c$ .
- $II_2$ : With probability  $\frac{3}{5}$  the state is  $b$  and with probability  $\frac{2}{5}$  the state is  $d$ .

Second-level beliefs (using short-hand notation for the above beliefs:  $(\frac{1}{2}a + \frac{1}{2}b)$ , etc.):

- $I_1$ : With probability  $\frac{1}{2}$ , player  $II$  believes  $(\frac{3}{7}a + \frac{4}{7}c)$ , and with probability  $\frac{1}{2}$ , player  $II$  believes  $(\frac{3}{5}b + \frac{2}{5}d)$ .
- $I_2$ : With probability  $\frac{2}{3}$ , player  $II$  believes  $(\frac{3}{7}a + \frac{4}{7}c)$ , and with probability  $\frac{1}{3}$ , player  $II$  believes  $(\frac{3}{5}b + \frac{2}{5}d)$ .
- $II_1$ : With probability  $\frac{3}{7}$ , player  $I$  believes  $(\frac{1}{2}a + \frac{1}{2}b)$ , and with probability  $\frac{4}{7}$ , player  $I$  believes  $(\frac{2}{3}c + \frac{1}{3}d)$ .
- $II_2$ : With probability  $\frac{3}{5}$ , player  $I$  believes  $(\frac{1}{2}a + \frac{1}{2}b)$ , and with probability  $\frac{2}{5}$ , player  $I$  believes  $(\frac{2}{3}c + \frac{1}{3}d)$ .

Third-level beliefs:

- $I_1$ : With probability  $\frac{1}{2}$ , player  $II$  believes that: “With probability  $\frac{3}{7}$ , player  $I$  believes  $(\frac{1}{2}a + \frac{1}{2}b)$  and with probability  $\frac{4}{7}$ , player  $I$  believes  $(\frac{2}{3}c + \frac{1}{3}d)$ ”. And with probability  $\frac{1}{2}$ , player  $II$  believes that: “With probability  $\frac{3}{5}$ , player  $I$  believes  $(\frac{1}{2}a + \frac{1}{2}b)$  and with probability  $\frac{2}{5}$ , player  $I$  believes  $(\frac{2}{3}c + \frac{1}{3}d)$ ”.

and so on and so on. The idea is very simple and powerful; since each player of a given type has a probability distribution (beliefs) both about the types of the other players and about the set  $S$  of states of nature, the hierarchies of beliefs are constructed inductively: If the  $k$ th level beliefs (about  $S$ ) are defined for each type, then the beliefs about types generates the  $(k + 1)$ th level of beliefs.

Thus the compact model of Harsanyi *does* capture the whole hierarchies of beliefs and it is rather tractable. The natural question is whether this model can be used for *all* hierarchies of beliefs. In other words, given any hierarchy of mutual beliefs of a set of players  $I$  about a set  $S$  of states of nature, can it be represented by a Harsanyi game? This was answered by Mertens and Zamir [18], who constructed the universal belief space; that is, given a set  $S$  of states of nature and a finite set  $I$  of players, they looked for the space  $\Omega$  of *all possible* hierarchies of mutual beliefs about  $S$  among the players in  $I$ . This construction is outlined in the next section.

### The Universal Belief Space

Given a finite set of players  $I = \{1, \dots, n\}$  and a set  $S$  of states of nature, which are assumed to be compact, we first identify the mathematical spaces in which lie the hierarchies of beliefs. Recall that  $\Delta(A)$  denotes the set of probability distributions on  $A$  and define inductively the sequence of spaces  $(X_k)_{k=1}^{\infty}$  by

$$X_1 = \Delta(S) \quad (1)$$

$$X_{k+1} = X_k \times \Delta(S \times X_k^{n-1}), \quad \text{for } k = 1, 2, \dots \quad (2)$$

Any probability distribution on  $S$  can be a first-level belief and is thus in  $X_1$ . A second-level belief is a joint probability distribution on  $S$  and the first-level beliefs of the other  $(n - 1)$  players. This is an element in  $\Delta(S \times X_1^{n-1})$  and therefore a two-level hierarchy is an element of the product space  $X_1 \times \Delta(S \times X_1^{n-1})$ , and so on for any level.



Note that at each level belief is a joint probability distribution on  $S$  and the previous level beliefs, allowing for correlation between the two. In dealing with these probability spaces we need to have some mathematical structure. More specifically, we make use of the *weak\** topology:

**Definition 4** A sequence  $(F_n)_{n=1}^{\infty}$  of probability measures (on  $\Omega$ ) converges in the *weak\** topology to the probability  $F$  if and only if  $\lim_{n \rightarrow \infty} \int_{\Omega} g(\omega) dF_n = \int_{\Omega} g(\omega) dF$  for all bounded and continuous functions  $g: \Omega \rightarrow \mathbb{R}$ .

It follows from the compactness of  $S$  that all spaces defined by (1)–(2) are compact in the *weak\** topology. However, for  $k > 1$ , not every element of  $X_k$  represents a coherent hierarchy of beliefs of level  $k$ . For example, if  $(\mu_1, \mu_2) \in X_2$  where  $\mu_1 \in \Delta(S) = X_1$  and  $\mu_2 \in \Delta(S \times X_1^{n-1})$ , then for this to describe meaningful beliefs of a player, the marginal distribution of  $\mu_2$  on  $S$  must coincide with  $\mu_1$ . More generally, any event  $A$  in the space of  $k$ -level beliefs has to have the same (marginal) probability in any higher-level beliefs. Furthermore, not only are each player's beliefs coherent, but he also considers only coherent beliefs of the other players (only those that are in support of his beliefs). Expressing formally this coherency condition yields a selection  $T_k \subseteq X_k$  such that  $T_1 = X_1 = \Delta(S)$ . It is proved that the projection of  $T_{k+1}$  on  $X_k$  is  $T_k$  (that is, any coherent  $k$ -level hierarchy can be extended to a coherent  $k + 1$ -level hierarchy) and that all the sets  $T_k$  are compact. Therefore, the *projective limit*,  $T = \lim_{\leftarrow k} T_k$ , is well defined and nonempty.<sup>1</sup>

**Definition 5** The *universal type space*  $T$  is the projective limit of the spaces  $(T_k)_{k=1}^{\infty}$ .

That is,  $T$  is the set of *all coherent infinite hierarchies of beliefs* regarding  $S$ , of a player in  $I$ . It does not depend on  $i$  since by construction it contains *all possible* hierarchies of beliefs regarding  $S$ , and it is therefore the same for all players. It is determined only by  $S$  and the number of players  $n$ .

**Proposition 6** The *universal type space*  $T$  is compact and satisfies

$$T \approx \Delta(S \times T^{n-1}). \quad (3)$$

The  $\approx$  sign in (3) is to be read as an *isomorphism* and Proposition 6 says that a type of player can be identified with a joint probability distribution on the state of nature and the types of the other players. The implicit equation (3) reflects the self-reference and circularity of the notion of type: The type of a player is his beliefs about the

state of nature and about all the beliefs of the other players, in particular, their beliefs about his own beliefs.

**Definition 7** The *universal belief space* (UBS) is the space  $\Omega$  defined by:

$$\Omega = S \times T^n. \quad (4)$$

An element of  $\Omega$  is called a *state of the world*.

Thus a state of the world is  $\omega = (s(\omega); t_1(\omega), t_2(\omega), \dots, t_n(\omega))$  with  $s(\omega) \in S$  and  $t_i(\omega) \in T$  for all  $i$  in  $I$ . This is the specification of the states of nature and the types of all players. The universal belief space  $\Omega$  is what we looked for: the set of *all* incomplete information and mutual belief configurations of a set of  $n$  players regarding the state of nature. In particular, as we will see later, all Harsanyi and Aumann models are embedded in  $\Omega$ , but it includes also belief configurations that cannot be modeled as Harsanyi games. As we noted before, the UBS is determined only by the set of states of nature  $S$  and the set of players  $I$ , so it should be denoted as  $\Omega(S, I)$ . For the sake of simplicity we shall omit the arguments and write  $\Omega$ , unless we wish to emphasize the underlying sets  $S$  and  $I$ .

The execution of the construction of the UBS according to the outline above involves some non-trivial mathematics, as can be seen in Mertens and Zamir [18]. The reason is that even with a finite number of states of nature, the space of first-level beliefs is a continuum, the second level is the space of probability distributions on a continuum and the third level is the space of probability distributions on the space of probability distributions on a continuum. This requires some structure for these spaces: For a (Borel) measurable event  $E$  let  $B_i^p(E)$  be the event “player  $i$  of type  $t_i$  believes that the probability of  $E$  is at least  $p$ ”, that is,

$$B_i^p(E) = \{\omega \in \Omega \mid t_i(E) \geq p\}$$

Since this is the object of beliefs of players other than  $i$  (beliefs of  $j \neq i$  about the beliefs of  $i$ ), this set must also be measurable. Mertens and Zamir used the *weak\** topology which is the minimal topology with which the event  $B_i^p(E)$  is (Borel) measurable for any (Borel) measurable event  $E$ . In this topology, if  $A$  is a compact set then  $\Delta(A)$ , the space of all probability distributions on  $A$ , is also compact. However, the hierarchic construction can also be made with stronger topologies on  $\Delta(A)$  (see [9,12,17]). Heifetz and Samet [14] worked out the construction of the universal belief space without topology, using only a measurable structure (which is implied by the assumption that the beliefs of the players are measurable). All these explicit constructions of the belief space are within what is

<sup>1</sup> The projective limit (also known as the *inverse limit*) of the sequence  $(T_k)_{k=1}^{\infty}$  is the space  $T$  of all sequences  $(\mu_1, \mu_2, \dots) \in \times_{k=1}^{\infty} T_k$  which satisfy: For any  $k \in \mathbb{N}$ , there is a probability distribution  $\nu_k \in \Delta(S \times T_k^{n-1})$  such that  $\mu_{k+1} = (\mu_k, \nu_k)$ .

called the *semantic* approach. Aumann [6] provided another construction of a belief system using the *syntactic* approach based on *sentences and logical formulas* specifying explicitly what each player believes about the state of nature, about the beliefs of the other players about the state of nature and so on. For a detailed construction see Aumann [6], Heifetz and Mongin [13], and Meier [16]. For a comparison of the syntactic and semantic approaches see Aumann and Heifetz [7].

### Belief Subspaces

In constructing the universal belief space we implicitly assumed that each player knows his own type since we specified only his beliefs about the state of nature and about the beliefs of the *other* players. In view of that, and since by (3) a type of player  $i$  is a probability distribution on  $S \times T^{I \setminus \{i\}}$ , we can view a type  $t_i$  also as a probability distribution on  $\Omega = S \times T^I$  in which the marginal distribution on  $T_i$  is a degenerate delta function at  $t_i$ ; that is, if  $\omega = (s(\omega); t_1(\omega), t_2(\omega), \dots, t_n(\omega))$ , then for all  $i$  in  $I$ ,

$$t_i(\omega) \in \Delta(\Omega) \quad \text{and} \quad t_i(\omega)[t_i = t_i(\omega)] = 1. \quad (5)$$

In particular it follows that if  $\text{Supp}(t_i)$  denotes the support of  $t_i$ , then

$$\omega' \in \text{Supp}(t_i(\omega)) \Rightarrow t_i(\omega') = t_i(\omega). \quad (6)$$

Let  $P_i(\omega) = \text{Supp}(t_i(\omega)) \subseteq \Omega$ . This defines a *possibility correspondence*; at state of the world  $\omega$ , player  $i$  does not consider as possible any point not in  $P_i(\omega)$ . By (6),

$$P_i(\omega) \cap P_i(\omega') \neq \emptyset \Rightarrow P_i(\omega) = P_i(\omega').$$

However, unlike in Aumann's model,  $P_i$  does not define a partition of  $\Omega$  since it is possible that  $\omega \notin P_i(\omega)$ , and hence the union  $\cup_{\omega \in \Omega} P_i(\omega)$  may be strictly smaller than  $\Omega$  (see Example 7). If  $\omega \in P_i(\omega) \subseteq Y$  holds for all  $\omega$  in some subspace  $Y \subset \Omega$ , then  $(P_i(\omega))_{\omega \in Y}$  is a partition of  $Y$ .

As we said, the universal belief space includes all possible beliefs and mutual belief structures over the state of nature. However, in a specific situation of incomplete information, it may well be that only part of  $\Omega$  is relevant for describing the situation. If the state of the world is  $\omega$  then clearly all states of the world in  $\cup_{i \in I} P_i(\omega)$  are relevant, but this is not all, because if  $\omega' \in P_i(\omega)$  then all states in  $P_j(\omega')$ , for  $j \neq i$ , are also relevant in the considerations of player  $i$ . This observation motivates the following definition:

**Definition 8** A *belief subspace* (BL-subspace) is a closed subset  $Y$  of  $\Omega$  which satisfies:

$$P_i(\omega) \subseteq Y \quad \forall i \in I \quad \text{and} \quad \forall \omega \in Y. \quad (7)$$

A belief subspace is *minimal* if it has no proper subset which is also a belief subspace. Given  $\omega \in \Omega$ , the belief subspace at  $\omega$ , denoted by  $Y(\omega)$ , is the minimal subspace containing  $\omega$ .

Since  $\Omega$  is a BL-subspace,  $Y(\omega)$  is well defined for all  $\omega \in \Omega$ . A BL-subspace is a closed subset of  $\Omega$  which is also *closed under beliefs* of the players. In any  $\omega \in Y$ , it contains all states of the world which are relevant to the situation: If  $\omega' \notin Y$ , then no player believes that  $\omega'$  is possible, no player believes that any other player believes that  $\omega'$  is possible, no player believes that any player believes that any player believes... etc.

*Remark 9* The subspace  $Y(\omega)$  is meant to be the minimal subspace which is belief-closed by all players *at the state*  $\omega$ . Thus, a natural definition would be:  $\tilde{Y}(\omega)$  is the minimal BL-subspace containing  $P_i(\omega)$  for all  $i$  in  $I$ . However, if for every player the state  $\omega$  is not in  $P_i(\omega)$  then  $\omega \notin \tilde{Y}(\omega)$ . Yet, even if it is not in the belief closure of the players, the real state  $\omega$  is still relevant (at least for the analyst) because it determines the true state of nature; that is, it determines the true payoffs of the game. This is the reason for adding the true state of the world  $\omega$ , even though "it may not be in the mind of the players".

It follows from (5), (6) and (7) that a BL-subspace  $Y$  has the following structure:

**Proposition 10** A closed subset  $Y$  of the universal belief space  $\Omega$  is a BL-subspace if and only if it satisfies the following conditions:

1. For any  $\omega = (s(\omega); t_1(\omega), t_2(\omega), \dots, t_n(\omega)) \in Y$ , and for all  $i$ , the type  $t_i(\omega)$  is a probability distribution on  $Y$ .
2. For any  $\omega$  and  $\omega'$  in  $Y$ ,

$$\omega' \in \text{Supp}(t_i(\omega)) \Rightarrow t_i(\omega') = t_i(\omega).$$

In fact condition 1 follows directly from Definition 8 while condition 2 follows from the general property of the UBS expressed in (6).

Given a BL-subspace  $Y$  in  $\Omega(S, I)$  we denote by  $T_i$  the *type set of player*  $i$ ,

$$T_i = \{t_i(\omega) | \omega \in Y\},$$

and note that unlike in the UBS, in a specific model  $Y$ , the type sets are typically *not the same* for all  $i$ , and the analogue of (4) is

$$Y \subseteq S \times T_1 \times \dots \times T_n.$$

A BL-subspace is a model of incomplete information about the state of nature. As we saw in Harsanyi's model,

in any model of incomplete information about a fixed set  $S$  of states of nature, involving the same set of players  $I$ , a state of the world  $\omega$  defines (encapsulates) an infinite hierarchy of mutual beliefs of the players  $I$  on  $S$ . By the universality of the belief space  $\Omega(S, I)$ , there is  $\omega' \in \Omega(S, I)$  with the same hierarchy of beliefs as that of  $\omega$ . The mapping of each  $\omega$  to its corresponding  $\omega'$  in  $\Omega(S, I)$  is called a *belief morphism*, as it preserves the belief structure. Mertens and Zamir [18] proved that the space  $\Omega(S, I)$  is universal in the sense that any model  $Y$  of incomplete information of the set of players  $I$  about the state of nature  $s \in S$  can be embedded in  $\Omega(S, I)$  via belief morphism  $\varphi: Y \rightarrow \Omega(S, I)$  so that  $\varphi(Y)$  is a belief subspace in  $\Omega(S, I)$ . In the following examples we give the *BL*-subspaces representing some known models.

### Examples of Belief Subspaces

*Example 1 (A game with complete information)* If the state of nature is  $s_0 \in S$  then in the universal belief space  $\Omega(S, I)$ , the game is described by a *BL*-subspace  $Y$  consisting of a single state of the world:

$$Y = \{\omega\} \quad \text{where } \omega = (s_0; [1\omega], \dots, [1\omega]).$$

Here  $[1\omega]$  is the only possible probability distribution on  $Y$ , namely, the trivial distribution supported by  $\omega$ . In particular, the state of nature  $s_0$  (i. e., the data of the game) is commonly known.

*Example 2 (Commonly known uncertainty about the state of nature)* Assume that the players' set is  $I = \{1, \dots, n\}$  and there are  $k$  states of nature representing, say,  $k$  possible  $n$ -dimensional payoff matrices  $G_1, \dots, G_k$ . At the beginning of the game, the payoff matrix is chosen by a chance move according to the probability vector  $p = (p_1, \dots, p_k)$  which is commonly known by the players but no player receives any information about the outcome of the chance move. The set of states of nature is  $S = \{G_1, \dots, G_k\}$ . The situation described above is embedded in the UBS,  $\Omega(S, I)$ , as the following *BL*-subspace  $Y$  consisting of  $k$  states of the world (denoting  $p \in \Delta(Y)$  by  $[p_1\omega_1, \dots, p_k\omega_k]$ ):

- $Y = \{\omega_1, \dots, \omega_k\}$
- $\omega_1 = (G_1; [p_1\omega_1, \dots, p_k\omega_k], \dots, [p_1\omega_1, \dots, p_k\omega_k])$
- $\omega_2 = (G_2; [p_1\omega_1, \dots, p_k\omega_k], \dots, [p_1\omega_1, \dots, p_k\omega_k])$
- $\dots$
- $\omega_k = (G_k; [p_1\omega_1, \dots, p_k\omega_k], \dots, [p_1\omega_1, \dots, p_k\omega_k])$ .

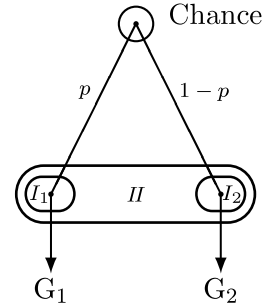
There is a single type,  $[p_1\omega_1, \dots, p_k\omega_k]$ , which is the same for all players. It should be emphasized that the type is a distribution on  $Y$  (and not just on the states of nature),

which implies that the beliefs  $[p_1G_1, \dots, p_kG_k]$  on the state of nature are commonly known by the players.

*Example 3 (Two players with incomplete information on one side)* There are two players,  $I = \{I, II\}$ , and two possible payoff matrices,  $S = \{G_1, G_2\}$ . The payoff matrix is chosen at random with  $P(s = G_1) = p$ , known to both players. The outcome of this chance move is known only to player  $I$ . Aumann and Maschler have studied such situations in which the chosen matrix is played repeatedly and the issue is how the informed player strategically uses his information (see Aumann and Maschler [8] and its references). This situation is presented in the UBS by the following *BL*-subspace:

- $Y = \{\omega_1, \omega_2\}$
- $\omega_1 = (G_1; [1\omega_1], [p\omega_1, (1-p)\omega_2])$
- $\omega_2 = (G_2; [1\omega_2], [p\omega_1, (1-p)\omega_2])$ .

Player  $I$  has two possible types:  $I_1 = [1\omega_1]$  when he is informed of  $G_1$ , and  $I_2 = [1\omega_2]$  when he is informed of  $G_2$ . Player  $II$  has only one type,  $II = [p\omega_1, (1-p)\omega_2]$ . We describe this situation in the following *extensive form-like* figure in which the oval forms describe the types of the players in the various vertices.



*Example 4 (Incomplete information about the other players' information)* In the next example, taken from Sorin and Zamir [23], one of two players always knows the state of nature but he may be uncertain whether the other player knows it. There are two players,  $I = \{I, II\}$ , and two possible payoff matrices,  $S = \{G_1, G_2\}$ . It is commonly known to both players that the payoff matrix is chosen at random by a toss of a fair coin:  $P(s = G_1) = 1/2$ . The outcome of this chance move is told to player  $I$ . In addition, if (and only if) the matrix  $G_1$  was chosen, another fair coin toss determines whether to inform player  $II$  which payoff matrix was chosen. In any case player  $I$  is *not* told the result of the second coin toss. This situation is described by the following belief space with three states of the world:

- $Y = \{\omega_1, \omega_2, \omega_3\}$
- $\omega_1 = (G_1; [\frac{1}{2}\omega_1, \frac{1}{2}\omega_2], [1\omega_1])$

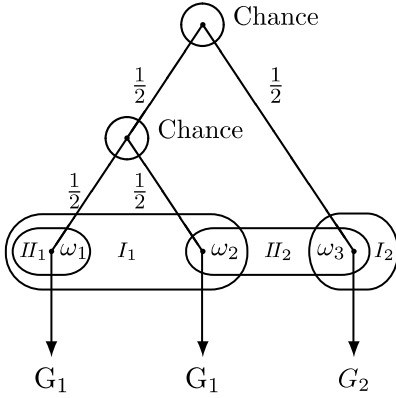
- $\omega_2 = (G_1; [\frac{1}{2}\omega_1, \frac{1}{2}\omega_2], [\frac{1}{3}\omega_2, \frac{2}{3}\omega_3])$
- $\omega_3 = (G_2; [1\omega_3], [\frac{1}{2}\omega_2, \frac{1}{2}\omega_3])$

Each player has two types and the type sets are:

$$T_I = \{I_1, I_2\} = \left\{ \left[ \frac{1}{2}\omega_1, \frac{1}{2}\omega_2 \right], [1\omega_3] \right\}$$

$$T_{II} = \{II_1, II_2\} = \left\{ [1\omega_1], \left[ \frac{1}{3}\omega_2, \frac{2}{3}\omega_3 \right] \right\}.$$

Note that in all our examples of belief subspaces, condition (6) is satisfied; the support of a player's type contains only states of the world in which he has that type. The game with incomplete information is described in the following figure:



**Example 5 (Incomplete information on two sides: A Harsanyi game)** In this example, the set of players is again  $I = \{I, II\}$  and the set of states of nature is  $S = \{s_{11}, s_{12}, s_{21}, s_{22}\}$ . In the universal belief space  $\Omega(S, I)$  consider the following BL-subspace consisting of four states of the world.

- $Y = \{\omega_{11}, \omega_{12}, \omega_{21}, \omega_{22}\}$
- $\omega_{11} = (s_{11}; [\frac{3}{7}\omega_{11}, \frac{4}{7}\omega_{12}], [\frac{2}{3}\omega_{11}, \frac{1}{3}\omega_{21}])$
- $\omega_{12} = (s_{12}; [\frac{3}{7}\omega_{11}, \frac{4}{7}\omega_{12}], [\frac{4}{5}\omega_{12}, \frac{1}{5}\omega_{22}])$
- $\omega_{21} = (s_{21}; [\frac{2}{3}\omega_{21}, \frac{1}{3}\omega_{22}], [\frac{2}{5}\omega_{11}, \frac{3}{5}\omega_{21}])$
- $\omega_{22} = (s_{22}; [\frac{2}{3}\omega_{21}, \frac{1}{3}\omega_{22}], [\frac{4}{5}\omega_{12}, \frac{1}{5}\omega_{22}])$

Again, each player has two types and the type sets are:

$$T_I = \{I_1, I_2\} = \left\{ \left[ \frac{3}{7}\omega_{11}, \frac{4}{7}\omega_{12} \right], \left[ \frac{2}{3}\omega_{21}, \frac{1}{3}\omega_{22} \right] \right\}$$

$$T_{II} = \{II_1, II_2\} = \left\{ \left[ \frac{3}{5}\omega_{11}, \frac{2}{5}\omega_{21} \right], \left[ \frac{4}{5}\omega_{12}, \frac{1}{5}\omega_{22} \right] \right\}.$$

The type of a player determines his beliefs about the type of the other player. For example, player  $I$  of type  $I_1$  assigns probability  $3/7$  to the state of the world  $\omega_{11}$  in which player

$II$  is of type  $II_1$ , and probability  $4/7$  to the state of the world  $\omega_{12}$  in which player  $II$  is of type  $II_2$ . Therefore, the beliefs of type  $I_1$  about the types of player  $II$  are  $P(II_1) = 3/7$ ,  $P(II_2) = 4/7$ . The mutual beliefs about each other's type are given in the following tables:

	$II_1$	$II_2$
$I_1$	$3/7$	$4/7$
$I_2$	$2/3$	$1/3$

Beliefs of player  $I$

	$II_1$	$II_2$
$I_1$	$3/5$	$4/5$
$I_2$	$2/5$	$1/5$

Beliefs of player  $II$

These are precisely the beliefs of Bayesian players if the pair of types  $(t_I, t_{II})$  in  $T = T_I \times T_{II}$  is chosen according to the prior probability distribution  $p$  below, and each player is then informed of his own type:

	$II_1$	$II_2$
$I_1$	0.3	0.4
$I_2$	0.2	0.1

Prior distribution  $p$  on  $T$

In other words, this BL-subspace is a Harsanyi game with type sets  $T_I, T_{II}$  and the prior probability distribution  $p$  on the types. Actually, as there is one-to-one mapping between the type set  $T$  and the set  $S$  of states of nature, the situation is generated by a chance move choosing the state of nature  $s_{ij} \in S$  according to the distribution  $p$  (that is,  $P(s_{ij}) = P(I_i, II_j)$  for  $i$  and  $j$  in  $\{1, 2\}$ ) and then player  $I$  is informed of  $i$  and player  $II$  is informed of  $j$ . As a matter of fact, *all* the BL-subspaces in the previous examples can also be written as Harsanyi games, mostly in a trivial way.

**Example 6 (Inconsistent beliefs)** In the same universal belief space,  $\Omega(S, I)$  of the previous example, consider now another BL-subspace  $\tilde{Y}$  which differs from  $Y$  only by changing the type  $II_1$  of player  $II$  from  $[\frac{3}{5}\omega_{11}, \frac{2}{5}\omega_{21}]$  to  $[\frac{1}{2}\omega_{11}, \frac{1}{2}\omega_{21}]$ , that is,

- $\tilde{Y} = \{\omega_{11}, \omega_{12}, \omega_{21}, \omega_{22}\}$
- $\omega_{11} = (s_{11}; [\frac{3}{7}\omega_{11}, \frac{4}{7}\omega_{12}], [\frac{1}{2}\omega_{11}, \frac{1}{2}\omega_{21}])$
- $\omega_{12} = (s_{12}; [\frac{3}{7}\omega_{11}, \frac{4}{7}\omega_{12}], [\frac{4}{5}\omega_{12}, \frac{1}{5}\omega_{22}])$
- $\omega_{21} = (s_{21}; [\frac{2}{3}\omega_{21}, \frac{1}{3}\omega_{22}], [\frac{1}{2}\omega_{11}, \frac{1}{2}\omega_{21}])$
- $\omega_{22} = (s_{22}; [\frac{2}{3}\omega_{21}, \frac{1}{3}\omega_{22}], [\frac{4}{5}\omega_{12}, \frac{1}{5}\omega_{22}])$

with type sets:

$$T_I = \{I_1, I_2\} = \left\{ \left[ \frac{3}{7}\omega_{11}, \frac{4}{7}\omega_{12} \right], \left[ \frac{2}{3}\omega_{21}, \frac{1}{3}\omega_{22} \right] \right\}$$

$$T_{II} = \{II_1, II_2\} = \left\{ \left[ \frac{1}{2}\omega_{11}, \frac{1}{2}\omega_{21} \right], \left[ \frac{4}{5}\omega_{12}, \frac{1}{5}\omega_{22} \right] \right\}.$$

Now, the mutual beliefs about each other's type are:

	$\Pi_1$	$\Pi_2$		$\Pi_1$	$\Pi_2$
$I_1$	3/7	4/7	$I_1$	1/2	4/5
$I_2$	2/3	1/3	$I_2$	1/2	1/5

Beliefs of player  $I$       Beliefs of player  $II$

Unlike in the previous example, these beliefs *cannot be derived from a prior distribution*  $p$ . According to Harsanyi, these are *inconsistent beliefs*. A  $BL$ -subspace with inconsistent beliefs *cannot be described* as a Harsanyi or Aumann model; it cannot be described as a game in extensive form.

**Example 7** (“Highly inconsistent” beliefs) In the previous example, even though the beliefs of the players were inconsistent in all states of the world, the true state was considered possible by all players (for example in the state  $\omega_{12}$  player  $I$  assigns to this state probability 4/7 and player  $II$  assigns to it probability 4/5). As was emphasized before, the UBS contains *all* belief configurations, including *highly inconsistent* or *wrong* beliefs, as the following example shows. The belief subspace of the two players  $I$  and  $II$  concerning the state of nature which can be  $s_1$  or  $s_2$  is given by:

- $Y = \{\omega_1, \omega_2\}$
- $\omega_1 = (s_1; [\frac{1}{2}\omega_1, \frac{1}{2}\omega_2], [1\omega_2])$
- $\omega_2 = (s_2; [\frac{1}{2}\omega_1, \frac{1}{2}\omega_2], [1\omega_2])$ .

In the state of the world  $\omega_1$ , the state of nature is  $s_1$ , player  $I$  assigns equal probabilities to  $s_1$  and  $s_2$ , but player  $II$  assigns probability 1 to  $s_2$ . In other words, he does not consider as possible the true state of the world (and also the true state of nature):  $\omega_1 \notin P_I(\omega_1)$  and consequently  $\cup_{\omega \in Y} P_I(\omega) = \{\omega_2\}$  which is strictly smaller than  $Y$ . By the definition of belief subspace and condition (6), this also implies that  $\cup_{\omega \in \Omega} P_I(\omega)$  is strictly smaller than  $\Omega$  (as it does not contain  $\omega_1$ ).

### Consistent Beliefs and Common Priors

A  $BL$ -subspace  $Y$  is a semantic belief system presenting, via the notion of types, the hierarchies of belief of a set of players having incomplete information about the state of nature. A state of the world captures the situation at what is called the *interim stage*: Each player knows his own type and has beliefs about the state of nature and the types of the other players. The question “what is the *real state of the world*  $\omega$ ?” is not addressed. In a  $BL$ -subspace, there is no chance move with explicit probability distribution that chooses the state of the world, while such a probability distribution is part of a Harsanyi or an Aumann model. Yet, in the belief space  $Y$  of Example 5 in the previous section,

such a *prior distribution*  $p$  emerged *endogenously* from the structure of  $Y$ . More specifically, if the state  $\omega \in Y$  is chosen by a chance move according to the probability distribution  $p$  and each player  $i$  is told his type  $t_i(\omega)$ , then his beliefs are *precisely* those described by  $t_i(\omega)$ . This is a property of the  $BL$ -subspace that we call *consistency* (which does not hold, for instance, for the  $BL$ -subspace  $\tilde{Y}$  in Example 6) and that we define now: Let  $Y \subseteq \Omega$  be a  $BL$ -subspace.

#### Definition 11

- (i) A probability distribution  $p \in \Delta(Y)$  is said to be *consistent* if for any player  $i \in I$ ,

$$p = \int_Y t_i(\omega) dp. \quad (8)$$

- (ii) A  $BL$ -subspace  $Y$  is said to be *consistent* if there is a consistent probability distribution  $p$  with  $\text{Supp}(p) = Y$ . A consistent  $BL$ -subspace will be called a  $C$ -subspace. A state of the world  $\omega \in \Omega$  is said to be *consistent* if it is a point in a  $C$ -subspace.

The interpretation of (8) is that the probability distribution  $p$  is “the average” of the types  $t_i(\omega)$  of player  $i$  (which are also probability distributions on  $Y$ ), when the average is taken on  $Y$  according to  $p$ . This definition is not transparent; it is not clear how it captures the consistency property we have just explained, in terms of a chance move choosing  $\omega \in Y$  according to  $p$ . However, it turns out to be equivalent.

For  $\omega \in Y$  denote  $\pi_i(\omega) = \{\omega' \in Y | t_i(\omega') = t_i(\omega)\}$ ; then we have:

**Proposition 12** A probability distribution  $p \in \Delta(Y)$  is *consistent* if and only if

$$t_i(\omega)(A) = p(A | \pi_i(\omega)) \quad (9)$$

holds for all  $i \in I$  and for any measurable set  $A \subseteq Y$ .

In particular, a Harsanyi or an Aumann model is represented by a consistent  $BL$ -subspace since, by construction, the beliefs are derived from a common prior distribution which is part of the data of the model. The role of the prior distribution  $p$  in these models is actually not that of an additional parameter of the model but rather that of an additional assumption on the belief system, namely, the *consistency assumption*. In fact, if a minimal belief subspace is consistent, then the common prior  $p$  is uniquely determined by the beliefs, as we saw in Example 5; *there is no need to specify*  $p$  *as additional data of the system*.



**Proposition 13** *If  $\omega \in \Omega$  is a consistent state of the world, and if  $Y(\omega)$  is the smallest consistent BL-subspace containing  $\omega$ , then the consistent probability distribution  $p$  on  $Y(\omega)$  is uniquely determined.*

(The formulation of this proposition requires some technical qualification if  $Y(\omega)$  is a continuum.)

The consistency (or the existence of a common prior), is quite a strong assumption. It assumes that differences in beliefs (i. e., in probability assessments) are due *only* to differences in information; players having precisely the same information will have precisely the same beliefs. It is no surprise that this assumption has strong consequences, the most known of which is due to Aumann [2]: Players with consistent beliefs *cannot agree to disagree*. That is, if at some state of the world it is commonly known that one player assigns probability  $q_1$  to an event  $E$  and another player assigns probability  $q_2$  to the same event, then it must be the case that  $q_1 = q_2$ . Variants of this result appear under the title of “No trade theorems” (see, e. g., [19]): Rational players with consistent beliefs cannot believe that they both can gain from a trade or a bet between them.

The plausibility and the justification of the common prior assumption was extensively discussed in the literature (see, e. g., [4,10,11]). It is sometimes referred to in the literature as the *Harsanyi doctrine*. Here we only make the observation that within the set of BL-subspaces in  $\Omega$ , the set of consistent BL-subspaces is a set of measure zero. To see the idea of the proof, consider the following example:

**Example 8 (Generalization of Examples 5 and 6)** Consider a BL-subspace as in Examples 5 and 6 but with type sets:

$$\begin{aligned} T_I &= \{I_1, I_2\} \\ &= \{[\alpha_1\omega_{11}, (1-\alpha_1)\omega_{12}], [\alpha_2\omega_{21}, (1-\alpha_2)\omega_{22}]\} \\ T_{II} &= \{II_1, II_2\} \\ &= \{[\beta_1\omega_{11}, (1-\beta_1)\omega_{21}], [\beta_2\omega_{12}, (1-\beta_2)\omega_{22}]\}. \end{aligned}$$

For any  $(\alpha_1, \alpha_2, \beta_1, \beta_2) \in [0, 1]^4$  this is a BL-subspace. The mutual beliefs about each other's type are:

	$II_1$	$II_2$
$I_1$	$\alpha_1$	$1 - \alpha_1$
$I_2$	$\alpha_2$	$1 - \alpha_2$

Beliefs of player  $I$

	$II_1$	$II_2$
$I_1$	$\beta_1$	$\beta_2$
$I_2$	$1 - \beta_1$	$1 - \beta_2$

Beliefs of player  $II$

If the subspace is consistent, these beliefs are obtained as conditional distributions from some prior probability distribution  $p$  on  $T = T_I \times T_{II}$ , say, by  $p$  of the following matrix:

	$II_1$	$II_2$
$I_1$	$p_{11}$	$p_{12}$
$I_2$	$p_{21}$	$p_{22}$

Prior distribution  $p$  on  $T$

This implies (assuming  $p_{ij} \neq 0$  for all  $i$  and  $j$ ),

$$\begin{aligned} \frac{p_{11}}{p_{12}} &= \frac{\alpha_1}{1 - \alpha_1}; & \frac{p_{21}}{p_{22}} &= \frac{\alpha_2}{1 - \alpha_2} \\ \text{and hence} & & \frac{p_{11}p_{22}}{p_{12}p_{21}} &= \frac{\alpha_1}{1 - \alpha_1} \frac{1 - \alpha_2}{\alpha_2}. \end{aligned}$$

Similarly,

$$\begin{aligned} \frac{p_{11}}{p_{21}} &= \frac{\beta_1}{1 - \beta_1}; & \frac{p_{12}}{p_{22}} &= \frac{\beta_2}{1 - \beta_2} \\ \text{and hence} & & \frac{p_{11}p_{22}}{p_{12}p_{21}} &= \frac{\beta_1}{1 - \beta_1} \frac{1 - \beta_2}{\beta_2}. \end{aligned}$$

It follows that the types must satisfy:

$$\frac{\alpha_1}{1 - \alpha_1} \frac{1 - \alpha_2}{\alpha_2} = \frac{\beta_1}{1 - \beta_1} \frac{1 - \beta_2}{\beta_2}, \quad (10)$$

which is *generally* not the case. More precisely, the set of  $(\alpha_1, \alpha_2, \beta_1, \beta_2) \in [0, 1]^4$  satisfying the condition (10) is a set of measure zero; it is a three-dimensional set in the four-dimensional set  $[0, 1]^4$ . Nyarko [21] proved that even the ratio of the dimensions of the set of consistent BL-subspaces to the dimension of the set of BL-subspaces goes to zero as the latter goes to infinity. Summing up, *most* BL-subspaces are inconsistent and thus do not satisfy the common prior condition.

## Bayesian Games and Bayesian Equilibrium

As we said, a game with incomplete information played by Bayesian players, often called a *Bayesian game*, is a game in which the players have incomplete information about the data of the game. Being a Bayesian, each player has beliefs (probability distribution) about any relevant data he does not know, including the beliefs of the other players. So far, we have developed the belief structure of such a situation which is a BL-subspace  $Y$  in the universal belief space  $\Omega(S, I)$ . Now we add the action sets and the pay-off functions. These are actually part of the description of the state of nature: The mapping  $s: \Omega \rightarrow S$  assigns to each state of the world  $\omega$  the *game-form*  $s(\omega)$  played at this state. To emphasize this interpretation of  $s(\omega)$  as a game-form, we denote it also as  $\Gamma_\omega$ :

$$\Gamma_\omega = (I, A_i(t_i(\omega))_{i \in I}, (u_i(\omega))_{i \in I}),$$

where  $A_i(t_i(\omega))$  is the actions set (pure strategies) of player  $i$  at  $\omega$  and  $u_i(\omega): A(\omega) \rightarrow \mathbb{R}$  is his payoff function and  $A(\omega) = \times_{i \in I} A_i(t_i(\omega))$  is the set of action profiles at state  $\omega$ . Note that while the actions of a player depend only on his type, his payoff depends on the actions and types of all the players. For a vector of actions  $a \in A(\omega)$ , we write  $u_i(\omega; a)$  for  $u_i(\omega)(a)$ . Given a  $BL$ -subspace  $Y \subseteq \Omega(S, I)$  we define the *Bayesian game on  $Y$*  as follows:

**Definition 14** The Bayesian game on  $Y$  is a vector payoff game in which:

- $I = \{1, \dots, n\}$  – the players' set.
- $\Sigma_i$  – the strategy set of player  $i$ , is the set of mappings

$$\sigma_i: Y \longrightarrow A_i \quad \text{which are } T_i\text{-measurable.}$$

In particular:

$$t_i(\omega_1) = t_i(\omega_2) \implies \sigma_i(\omega_1) = \sigma_i(\omega_2).$$

Let  $\Sigma = \times_{i \in I} \Sigma_i$ .

- The payoff function  $u_i$  for player  $i$  is a *vector-valued function*  $u_i = (u_{t_i})_{t_i \in T_i}$ , where  $u_{t_i}$  (the payoff function of player  $i$  of type  $t_i$ ) is a mapping

$$u_{t_i}: \Sigma \longrightarrow \mathbb{R}$$

defined by

$$u_{t_i}(\sigma) = \int_Y u_i(\omega; \sigma(\omega)) dt_i(\omega). \quad (11)$$

Note that  $u_{t_i}$  is  $T_i$ -measurable, as it should be. When  $Y$  is a finite  $BL$ -subspace, the above-defined Bayesian game is an  $n$ -person “game” in which the payoff for player  $i$  is a vector with a payoff for each one of his types (therefore, a vector of dimension  $|T_i|$ ). It becomes a *regular game-form* for a given state of the world  $\omega$  since then the payoff to player  $i$  is  $u_{t_i(\omega)}$ . However, these game-forms are not regular games since they are *interconnected*; the players do not know which of these “games” they are playing (since they do not know the state of the world  $\omega$ ). Thus, just like a Harsanyi game, a Bayesian game on a  $BL$ -subspace  $Y$  consists of a family of *connected* game-forms, one for each  $\omega \in Y$ . However, unlike a Harsanyi game, a Bayesian game has no chance move that chooses the state of the world (or the vector of types). A way to transform a Bayesian game into a regular game was suggested by R. Selten and was named by Harsanyi as the *Selten game*  $G^{**}$  (see p. 496 in [11]). This is a game with  $|T_1| \cdot |T_2| \cdot \dots \cdot |T_n|$  players (one for each type) in which each player  $t_i \in T_i$  chooses a strategy and then *selects* his  $(n-1)$  partners, one from each  $T_j$ ;  $j \neq i$ , according to his beliefs  $t_i$ .

## Bayesian Equilibrium

Although a Bayesian game is not a regular game, the Nash equilibrium concept based on the notion of *best reply* can be adapted to yield the solution concept of *Bayesian equilibrium* (also called *Nash–Bayes equilibrium*).

**Definition 15** A vector of strategies  $\sigma = (\sigma_1, \dots, \sigma_n)$ , in a Bayesian game, is called a *Bayesian equilibrium* if for all  $i$  in  $I$  and for all  $t_i$  in  $T_i$ ,

$$u_{t_i}(\sigma) \geq u_{t_i}(\sigma_{-i}; \tilde{\sigma}_i), \quad \forall \tilde{\sigma}_i \in \Sigma_i, \quad (12)$$

where, as usual,  $\sigma_{-i} = (\sigma_j)_{j \neq i}$  denotes the vector of strategies of players other than  $i$ .

Thus, a Bayesian equilibrium specifies a behavior for each player which is a *best reply* to what he believes is the behavior of the other players, that is, a best reply to the strategies of the other players given his type. In a game with complete information, which corresponds to a  $BL$ -subspace with one state of the world ( $Y = \{\omega\}$ ), as there is only one type of each player, and the beliefs are all probability one on a singleton, the Bayesian equilibrium is just the well-known Nash equilibrium.

**Remark 16** It is readily seen that when  $Y$  is finite, any Bayesian equilibrium is a Nash equilibrium of the Selten game  $G^{**}$  in which each type is a player who selects the types of his partners according to his beliefs. Similarly, we can transform the Bayesian game into an ordinary game in strategic form by defining the payoff function to player  $i$  to be  $\tilde{u}_i = \sum_{t_i \in T_i} \gamma_{t_i} u_{t_i}$  where  $\gamma_{t_i}$  are strictly positive. Again, independently of the values of the constants  $\gamma_{t_i}$ , any Bayesian equilibrium is a Nash equilibrium of this game and vice versa. In particular, if we choose the constants so that  $\sum_{t_i \in T_i} \gamma_{t_i} = 1$ , we obtain the game suggested by Aumann and Maschler in 1967 (see p. 95 in [8]) and again, the set of Nash equilibria of this game is precisely the set of Bayesian equilibria.

## The Harsanyi Game Revisited

As we observed in Example 5, the belief structure of a consistent  $BL$ -subspace is the same as in a Harsanyi game *after the chance move choosing the types*. That is, the embedding of the Harsanyi game as a  $BL$ -subspace in the universal belief space is only at the *interim stage*, after the moment that each player gets to know his type. The Harsanyi game on the other hand is at the *ex ante stage*, before a player knows his type. Then, what is the relation between the Nash equilibrium in the Harsanyi game at the ex ante stage and the equilibrium at the interim stage, namely, the Bayesian equilibrium of the corresponding  $BL$ -subspace?

This is an important question concerning the embedding of the Harsanyi game in the UBS since, as we said before, the chance move choosing the types *does not appear explicitly in the UBS*. The answer to this question was given by Harsanyi (1967–8) (assuming that each type  $t_i$  has a positive probability):

**Theorem 17 (Harsanyi)** *The set of Nash equilibria of a Harsanyi game is identical to the set of Bayesian equilibria of the equivalent BL-subspace in the UBS.*

In other words, this theorem states that any equilibrium in the ex ante stage is also an equilibrium at the interim stage and vice versa.

In modeling situations of incomplete information, the interim stage is the natural one; if a player *knows his beliefs (type)*, then why should he analyze the situation, as Harsanyi suggests, from the ex ante point of view as if his type was not known to him and he could equally well be of another type? Theorem 17 provides a technical answer to this question: The equilibria are the *same* in both games and the equilibrium strategy of the ex ante game specifies for each type precisely his equilibrium strategy at the interim stage. In that respect, for a player who knows his type, the Harsanyi model is just an *auxiliary game* to compute his equilibrium behavior. Of course the deeper answer to the question above comes from the interactive nature of the situation: Even though player  $i$  knows he is of type  $t_i$ , he knows that his partners do not know that and that they may consider the possibility that he is of type  $\bar{t}_i$ , and since this affects their behavior, the behavior of type  $\bar{t}_i$  is also relevant to player  $i$  who knows he is of type  $t_i$ . Finally, Theorem 17 makes the Bayesian equilibrium the natural extension of the Nash equilibrium concept to games with incomplete information for consistent or inconsistent beliefs, when the Harsanyi ordinary game model is unavailable.

**Examples of Bayesian Equilibria** In Example 6, there are two players of two types each, and with *inconsistent* mutual beliefs given by

	$\Pi_1$	$\Pi_2$
$I_1$	3/7	4/7
$I_2$	2/3	1/3

Beliefs of player  $I$

	$\Pi_1$	$\Pi_2$
$I_1$	1/2	4/5
$I_2$	1/2	1/5

Beliefs of player  $II$

Assume that the payoff matrices for the four type's of profiles are:

		$II$	
		$L$	$R$
$I$	$T$	2, 0	0, 1
	$B$	0, 0	1, 0

$G_{11}$ : Payoffs when  $t = (I_1, II_1)$

		$II$	
		$L$	$R$
$I$	$T$	0, 0	0, 0
	$B$	1, 1	1, 0

$G_{12}$ : Payoffs when  $t = (I_1, II_2)$

		$II$	
		$L$	$R$
$I$	$T$	0, 0	0, 0
	$B$	1, 1	0, 0

$G_{21}$ : Payoffs when  $t = (I_2, II_1)$

		$II$	
		$L$	$R$
$I$	$T$	0, 0	2, 1
	$B$	0, 0	0, 2

$G_{22}$ : Payoffs when  $t = (I_2, II_2)$

As the beliefs are inconsistent they cannot be presented by a Harsanyi game. Yet, we can compute the Bayesian equilibrium of this Bayesian game. Let  $(x, y)$  be the strategy of player  $I$ , which is:

- Play the mixed strategy  $[x(T), (1 - x)(B)]$  when you are of type  $I_1$ .
- Play the mixed strategy  $[y(T), (1 - y)(B)]$  when you are of type  $I_2$ .

and let  $(z, t)$  be the strategy of player  $II$ , which is:

- Play the mixed strategy  $[z(L), (1 - z)(R)]$  when you are of type  $II_1$ .
- Play the mixed strategy  $[t(L), (1 - t)(R)]$  when you are of type  $II_2$ .

For  $0 < x, y, z, t < 1$ , each player of each type must be indifferent between his two pure actions; that yields the values in equilibrium:

$$x = \frac{3}{5}, \quad y = \frac{2}{5}, \quad z = \frac{7}{9}, \quad t = \frac{2}{9}.$$

There is no “expected payoff” since this is a Bayesian game and not a game; the expected payoffs depend on the *actual* state of the world, i. e., the actual types of the players and the actual payoff matrix. For example, the state of the world is  $\omega_{11} = (G_{11}; I_1, II_1)$ ; the expected payoffs are:

$$\pi(\omega_{11}) = \left(\frac{3}{5}, \frac{2}{5}\right) G_{11} \begin{pmatrix} 7/9 \\ 2/9 \end{pmatrix} = \left(\frac{46}{45}, \frac{6}{45}\right).$$

Similarly:

$$\pi(\omega_{12}) = \left(\frac{3}{5}, \frac{2}{5}\right) G_{12} \begin{pmatrix} 2/9 \\ 7/9 \end{pmatrix} = \left(\frac{18}{45}, \frac{4}{45}\right)$$

$$\pi(\omega_{21}) = \left(\frac{2}{5}, \frac{3}{5}\right) G_{21} \begin{pmatrix} 7/9 \\ 2/9 \end{pmatrix} = \left(\frac{21}{45}, \frac{21}{45}\right)$$

$$\pi(\omega_{22}) = \left(\frac{2}{5}, \frac{3}{5}\right) G_{22} \begin{pmatrix} 2/9 \\ 7/9 \end{pmatrix} = \left(\frac{28}{45}, \frac{70}{45}\right).$$

However, these are the *objective* payoffs as viewed by the analyst; they are viewed differently by the players. For

player  $i$  of type  $t_i$  the relevant payoff is his subjective payoff  $u_{t_i}(\sigma)$  defined in (11). For example, at state  $\omega_{11}$  (or  $\omega_{12}$ ) player  $I$  believes that with probability  $3/7$  the state is  $\omega_{11}$  in which case his payoff is  $46/45$  and with probability  $4/7$  the state is  $\omega_{12}$  in which case his payoff is  $18/45$ . Therefore his *subjective* expected payoff at state  $\omega_{11}$  is  $3/7 \times 46/45 + 4/7 \times 18/45 = 2/3$ . Similar computations show that in states  $\omega_{21}$  or  $\omega_{22}$  player  $I$  “expects” a payoff of  $7/15$  while player  $II$  “expects”  $3/10$  at states  $\omega_{11}$  or  $\omega_{21}$  and  $86/225$  in states  $\omega_{12}$  or  $\omega_{22}$ .

Bayesian equilibrium is widely used in Auction Theory, which constitutes an important and successful application of the theory of games with incomplete information. The simplest example is that of two buyers bidding in a first-price auction for an indivisible object. If each buyer  $i$  has a private value  $v_i$  for the object (which is independent of the private value  $v_j$  of the other buyer), and if he further believes that  $v_j$  is random with uniform probability distribution on  $[0, 1]$ , then this is a Bayesian game in which the type of a player is his private valuation; that is, the type sets are  $T_1 = T_2 = [0, 1]$ , which is a continuum. This is a consistent Bayesian game (that is, a Harsanyi game) since the beliefs are derived from the uniform probability distribution on  $T_1 \times T_2 = [0, 1]^2$ . A Bayesian equilibrium of this game is that in which each player bids half of his private value:  $b_i(v_i) = v_i/2$  (see, e.g., Chap. III in [25]). Although auction theory was developed far beyond this simple example, almost all the models studied so far are Bayesian games with consistent beliefs, that is, Harsanyi games. The main reason of course is that consistent Bayesian games are more manageable since they can be described in terms of an equivalent ordinary game in strategic form. However, inconsistent beliefs are rather plausible and exist in the market place in general and even more so in auction situations. An example of that is the case of collusion of bidders: When a bidding ring is formed, it may well be the case that some of the bidders outside the ring are unaware of its existence and behave under the belief that all bidders are competitive. The members of the ring may or may not know whether the other bidders know about the ring, or they may be uncertain about it. This rather plausible mutual belief situation is typically inconsistent and has to be treated as an inconsistent Bayesian game for which a Bayesian equilibrium is to be found.

### Bayesian Equilibrium and Correlated Equilibrium

*Correlated equilibrium* was introduced in Aumann (1974) as the Nash equilibrium of a game extended by adding to it random events about which the players have partial information. Basically, starting from an ordinary game, Au-

mann added a probability space and information structure and obtained a game with incomplete information, the equilibrium of which he called a *correlated equilibrium of the original game*. The fact that the Nash equilibrium of a game with incomplete information is the Bayesian equilibrium suggests that the concept of correlated equilibrium is closely related to that of Bayesian equilibrium. In fact Aumann noticed that and discussed it in a second paper entitled “Correlated equilibrium as an expression of Bayesian rationality” [3]. In this section, we review briefly, by way of an example, the concept of correlated equilibrium, and state formally its relation to the concept of Bayesian equilibrium.

**Example 18** Consider a two-person game with actions  $\{T, B\}$  for player 1 and  $\{L, R\}$  for player 2 with corresponding payoffs given in the following matrix:

		2	
		L	R
1	T	6,6	2,7
	B	7,2	0,0

$G$ : Payoffs of the basic game

This game has three Nash equilibria:  $(T, R)$  with payoff  $(2, 7)$ ,  $(B, L)$  with payoff  $(7, 2)$  and the mixed equilibrium  $([\frac{2}{3}(T), \frac{1}{3}(B)], [\frac{2}{3}(L), \frac{1}{3}(R)])$  with payoff  $(4\frac{2}{3}, 4\frac{2}{3})$ . Suppose that we add to the game a chance move that chooses an element in  $\{T, B\} \times \{L, R\}$  according to the following probability distribution  $\mu$ :

		L	R
	T	1/3	1/3
	B	1/3	0,0

$\mu$ : Probability distribution on  $\{T, B\} \times \{L, R\}$

Let us now extend the game  $G$  to a game with incomplete information  $G^*$  in which a chance move chooses an element in  $\{T, B\} \times \{L, R\}$  according to the probability distribution above. Then player 1 is informed of the first (left) component of the chosen element and player 2 is informed of the second (right) component. Then each player chooses an action in  $G$  and the payoff is made. If we interpret the partial information as a suggestion of which action to choose, then it is readily verified that following the suggestion is a Nash equilibrium of the extended game yielding a payoff  $(5, 5)$ . This was called by Aumann a *correlated equilibrium of the original game  $G$* . In our terminology, the extended game  $G^*$  is a Bayesian game and its Nash equilibrium is its Bayesian equilibrium. Thus what

we have here is that a correlated equilibrium of a game is just the Bayesian equilibrium of its extension to a game with incomplete information. We now make this a general formal statement. For simplicity, we use the Aumann model of a game with incomplete information.

Let  $G = (I, (A_i)_{i \in I}, (u_i)_{i \in I})$  be a game in strategic form where  $I$  is the set of players,  $A_i$  is the set of actions (pure strategies) of player  $i$  and  $u_i$  is his payoff function.

**Definition 19** Given a game in strategic form  $G$ , an incomplete information extension (the  $I$ -extension) of the game  $G$  is the game  $G^*$  given by

$$G^* = (I, (A_i)_{i \in I}, (u_i)_{i \in I}, (Y, p)), (\pi_i)_{i \in I},$$

where  $(Y, p)$  is a finite probability space and  $\pi_i$  is a partition of  $Y$  (the information partition of player  $i$ ).

This is an Aumann model of incomplete information and, as we noted before, it is also a Harsanyi type-based model in which the type of player  $i$  at state  $\omega \in Y$  is  $t_i(\omega) = \pi_i(\omega)$ , and a strategy of player  $i$  is a mapping from his type set to his mixed actions:  $\sigma_i: T_i \rightarrow \Delta(A_i)$ .

We identify a correlated equilibrium in the game  $G$  by the probability distribution  $\mu$  on the vectors of actions  $A = A_1 \times \dots \times A_n$ . Thus  $\mu \in \Delta(A)$  is a correlated equilibrium of the game  $G$  if when  $a \in A$  is chosen according to  $\mu$  and each player  $i$  is suggested to play  $a_i$ , his best reply is in fact to play the action  $a_i$ .

Given a game with incomplete information  $G^*$  as in definition 19, any vector of strategies of the players  $\sigma = (\sigma_1, \dots, \sigma_n)$  induces a probability distribution on the vectors of actions  $a \in A$ . We denote this as  $\mu_\sigma \in \Delta(A)$ .

We can now state the relation between correlated and Bayesian equilibria:

**Theorem 20** Let  $\sigma$  be a Bayesian equilibrium in the game of incomplete information  $G^* = (I, (A_i)_{i \in I}, (u_i)_{i \in I}, (Y, p)), (\pi_i)_{i \in I}$ ; then the induced probability distribution  $\mu_\sigma$  is a correlated equilibrium of the basic game  $G = (I, (A_i)_{i \in I}, (u_i)_{i \in I})$ .

The other direction is:

**Theorem 21** Let  $\mu$  be a correlated equilibrium of the game  $G = (I, (A_i)_{i \in I}, (u_i)_{i \in I})$ ; then  $G$  has an extension to a game with incomplete information  $G^* = (I, (A_i)_{i \in I}, (u_i)_{i \in I}, (Y, p)), (\pi_i)_{i \in I}$  with a Bayesian equilibrium  $\sigma$  for which  $\mu_\sigma = \mu$ .

## Concluding Remarks and Future Directions

### The Consistency Assumption

To the heated discussion of the merits and justification of the consistency assumption in economic and game-theo-

retical models, we would like to add a couple of remarks. In our opinion, the appropriate way of modeling an incomplete information situation is at the *interim stage*, that is, when a player *knows his own beliefs (type)*. The Harsanyi *ex ante* model is just an auxiliary construction for the analysis. Actually this was also the view of Harsanyi, who justified his model by proving that it provides the same equilibria as the interim stage situation it generates (Theorem 17). The *Harsanyi doctrine* says roughly that our models “*should be consistent*” and if we get an inconsistent model it must be the case that it not be a “correct” model of the situation at hand. This becomes less convincing if we agree that the interim stage is what we are interested in: Not only are *most* mutual beliefs inconsistent, as we saw in the section entitled “[Consistent Beliefs and Common Priors](#)” above, but it is hard to argue convincingly that the model in Example 5 describes an *adequate* mutual belief situation while the model in Example 6 does not; the only difference between the two is that in one model, a certain type’s beliefs are  $[\frac{3}{5}\omega_{11}, \frac{2}{5}\omega_{21}]$  while in the other model his beliefs are  $[\frac{1}{2}\omega_{11}, \frac{1}{2}\omega_{21}]$ .

Another related point is the fact that if players’ beliefs are the data of the situation (in the interim stage), then these are typically imprecise and rather hard to measure. Therefore any meaningful result of our analysis should be robust to small changes in the beliefs. This cannot be achieved within the consistent belief systems which are a thin set of measure zero in the universal belief space.

### Knowledge and Beliefs

Our interest in this article was mostly in the notion of *beliefs* of players and less in the notion of *knowledge*. These are two related but different notions. Knowledge is defined through a knowledge operator satisfying some axioms. Beliefs are defined by means of probability distributions. Aumann’s model, discussed in the section entitled “[Aumann’s Model](#)” above, has both elements: The knowledge was generated by the partitions of the players while the beliefs were generated by the probability  $P$  on the space  $Y$  (and the partitions). Being interested in the subjective beliefs of the player we could understand “at state of the world  $\omega \in \Omega$  player  $i$  knows the event  $E \subseteq \Omega$ ” to mean “at state of the world  $\omega \in \Omega$  player  $i$  assigns to the event  $E \subseteq \Omega$  probability 1”. However, in the universal belief space, “belief with probability 1” does not satisfy a central axiom of the knowledge operator. Namely, if at  $\omega \in \Omega$  player  $i$  knows the event  $E \subseteq \Omega$ , then  $\omega \in E$ . That is, if a player *knows* an event, then this event in fact happened. In the universal belief space where all coherent beliefs are possible, in a state  $\omega \in \Omega$  a player may assign probabil-



ity 1 to the event  $\{\omega'\}$  where  $\omega' \neq \omega$ . In fact, if in a *BL*-subspace  $Y$  the condition  $\omega \in P_i(\omega)$  is satisfied for all  $i$  and all  $\omega \in Y$ , then belief with probability 1 is a knowledge operator on  $Y$ . This in fact was the case in Aumann's and in Harsanyi's models where, by construction, the support of the beliefs of a player in the state  $\omega$  always included  $\omega$ . For a detailed discussion of the relationship between knowledge and beliefs in the universal belief space see Vassilakis and Zamir [24].

### Future Directions

We have not said much about the existence of Bayesian equilibrium, mainly because it has not been studied enough and there are no general results, especially in the non-consistent case. We can readily see that a Bayesian game on a finite *BL*-subspace in which each state of nature  $s(\omega)$  is a finite game-form has a Bayesian equilibrium in mixed strategies. This can be proved, for example, by transforming the Bayesian game into an ordinary finite game (see Remark 16) and applying the Nash theorem for finite games. For games with incomplete information with a continuum of strategies and payoff functions not necessarily continuous, there are no general existence results. Even in consistent auction models, existence was proved for specific models separately (see [20,15,22]). Establishing general existence results for large families of Bayesian games is clearly an important future direction of research. Since, as we argued before, most games are Bayesian games, the existence of a Bayesian equilibrium should, and could, reach at least the level of generality available for the existence of a Nash equilibrium.

### Acknowledgments

I am grateful to two anonymous reviewers for their helpful comments.

### Bibliography

- Aumann R (1974) Subjectivity and Correlation in Randomized Strategies. *J Math Econ* 1:67–96
- Aumann R (1976) Agreeing to disagree. *Ann Stat* 4:1236–1239
- Aumann R (1987) Correlated equilibrium as an expression of Bayesian rationality. *Econometrica* 55:1–18
- Aumann R (1998) Common priors: A reply to Gul. *Econometrica* 66:929–938
- Aumann R (1999) Interactive epistemology I: Knowledge. *Intern J Game Theory* 28:263–300
- Aumann R (1999) Interactive epistemology II: Probability. *Intern J Game Theory* 28:301–314
- Aumann R, Heifetz A (2002) Incomplete Information. In: Aumann R, Hart S (eds) *Handbook of Game Theory*, vol 3. Elsevier, pp 1666–1686
- Aumann R, Maschler M (1995) *Repeated Games with Incomplete Information*. MIT Press, Cambridge
- Brandenburger A, Dekel E (1993) Hierarchies of beliefs and common knowledge. *J Econ Theory* 59:189–198
- Gul F (1998) A comment on Aumann's Bayesian view. *Econometrica* 66:923–927
- Harsanyi J (1967–8) Games with incomplete information played by 'Bayesian' players, parts I–III. *Manag Sci* 8:159–182, 320–334, 486–502
- Heifetz A (1993) The Bayesian formulation of incomplete information, the non-compact case. *Intern J Game Theory* 21:329–338
- Heifetz A, Mongin P (2001) Probability logic for type spaces. *Games Econ Behav* 35:31–53
- Heifetz A, Samet D (1998) Topology-free topology of beliefs. *J Econ Theory* 82:324–341
- Maskin E, Riley J (2000) Asymmetric auctions. *Rev Econ Stud* 67:413–438
- Meier M (2001) An infinitary probability logic for type spaces. CORE Discussion Paper 2001/61
- Mertens J-F, Sorin S, Zamir S (1994) Repeated Games, Part A: Background Material. CORE Discussion Paper No. 9420
- Mertens J-F, Zamir S (1985) Foundation of Bayesian analysis for games with incomplete information. *Intern J Game Theory* 14:1–29
- Milgrom PR, Stokey N (1982) Information, trade and common knowledge. *J Eco Theory* 26:17–27
- Milgrom PR, Weber RJ (1982) A Theory of Auctions and Competitive Bidding. *Econometrica* 50:1089–1122
- Nyarko Y (1991) Most games violate the Harsanyi doctrine. C.V. Starr working paper #91–39, NYU
- Reny P, Zamir S (2004) On the existence of pure strategy monotone equilibria in asymmetric first price auctions. *Econometrica* 72:1105–1125
- Sorin S, Zamir S (1985) A 2-person game with lack of information on  $1\frac{1}{2}$  sides. *Math Oper Res* 10:17–23
- Vassilakis S, Zamir S (1993) Common beliefs and common knowledge. *J Math Econ* 22:495–505
- Wolfstetter E (1999) *Topics in Microeconomics*. Cambridge University Press, Cambridge

## Bayesian Methods in Non-linear Time Series

OLEG KORENOK

Department of Economics, VCU School of Business,  
Richmond, USA

### Article Outline

Glossary

Definition of the Subject

Introduction

Threshold Autoregressive Model

Smooth Transition Autoregressive Model

Markov-Switching Model

Future Directions  
Acknowledgments  
Bibliography

## Glossary

**Autoregressive model** describes a stochastic process as a weighted average of its previous values and a stochastic error term.

**Threshold autoregressive model** is an autoregressive model in which parameters change depending on the time index or the previous values of the process.

**Markov-switching autoregressive model** is an autoregressive model in which parameters change over time depending on an unobserved Markov chain.

**Prior distribution** summarizes the information about the parameters of interest after observing the data.

**Posterior distribution** summarizes the information about the parameters of interest after observing the data.

## Definition of the Subject

Economic fluctuations display definite nonlinear features. Recessions, wars, financial panics, and varying government policies change the dynamics of almost all macroeconomic and financial time series. In the time series literature, such events are modeled by modifying the standard linear autoregressive (abbreviated, AR) model

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t,$$

where  $y_t$  is a covariance stationary process,  $\epsilon_t$  is an independent and identically distributed noise process,  $\epsilon_t \sim i.i.d.N(0, \sigma^2)$ , and the parameters  $c$ ,  $\phi_i$ , and  $\sigma^2$  are fixed over time. In particular, the literature assumes that  $y_t$  follows two or more regimes. The three most commonly used nonlinear models differ in their description of the transition between regimes. In the threshold autoregressive (abbreviated, TAR) model, regime changes abruptly; in the smooth threshold autoregressive (abbreviated, STAR) model, regime changes slowly. Nevertheless, in both models the regime change depends on the time index or lagged values of  $y_t$ . In the Markov-switching autoregressive (abbreviated, MAR) model, however, the regime change depends on the past values of an unobserved random variable, the state of the Markov chain, and possibly the lagged values of  $y_t$ .

Arguably, the best-known example of the nonlinear time series model is the model of cyclical fluctuations of the US economy. It was first introduced and estimated by Hamilton [45] for quarterly US real Gross National Product over the 1952(II)–1984(IV) period. The model has two

discrete regimes. The first regime is associated with a positive 1.2% growth rate and the second regime is associated with a negative  $-0.4\%$  growth rate. Against his original motivation to find decade-long changes in growth rate trends for the US economy, Hamilton finds that negative growth regimes occur at the business cycle frequency. Positive growth regimes last, on average, 10 quarters, and negative growth regimes last, on average, 4 quarters. Moreover, he finds that the estimated regimes coincide closely with the official National Bureau of Economic Research (abbreviated, NBER) recession dates.

Figure 1 illustrates Hamilton's results for the extended 1952(II)–2006(IV) sample. Panel (a) shows the quarterly growth rate of the US real Gross Domestic Product, currently the more common measure of output; panel (b) plots the estimated probability that the US economy is in a negative growth regime. The shaded regions represent recessionary periods as determined informally and with some delay by the NBER: It took nine months for the NBER's Business Cycle Dating Committee to determine the latest peak of the US economy, which occurred in March 2001 but was officially announced in November 2001. Even though the NBER dates were not used in the model, the periods with high probability of a negative growth rate coincide almost perfectly with the NBER dates.

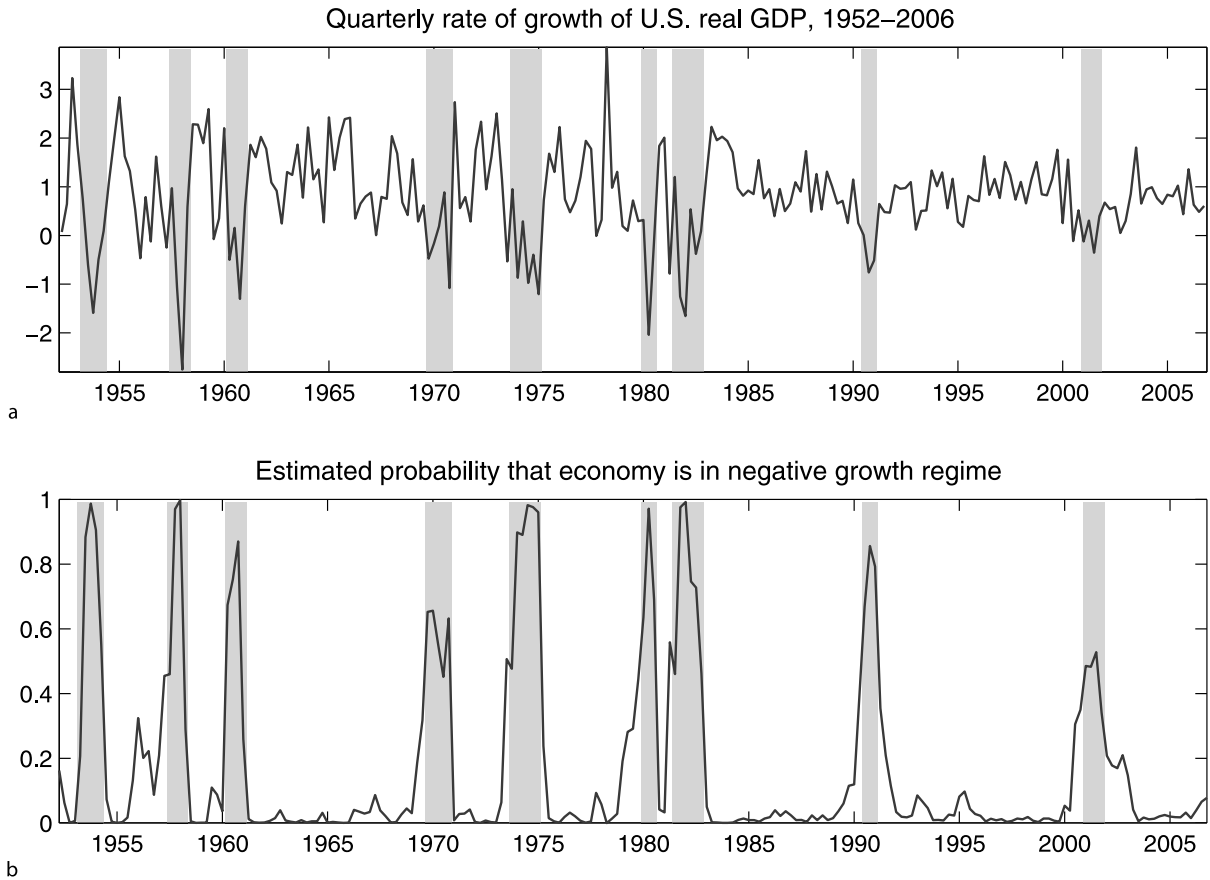
In addition to the formal recession dating methodology, Hamilton [45] presents clear statistical evidence for the proposition that the US business cycle is asymmetric: Behavior of output during normal times, when labor, capital, and technology determine long-run economic growth, is distinct from behavior during recessions, when all these factors are underutilized.

## Introduction

Hamilton's paper triggered an explosion of interest in nonlinear time series. The purpose of this paper is to give a survey of the main developments from the Bayesian perspective. The Bayesian framework treats model parameters as random variables and interprets probability as a degree of belief about particular realizations of a random variable conditional on available information. Given the observed sample, the inference updates prior beliefs, formulated before observing the sample, into posterior beliefs using Bayes' theorem

$$p(\theta|y) = \frac{f(y|\theta)\pi(\theta)}{f(y)},$$

where  $y$  is the sample observations  $y = (y_1, \dots, y_T)$ ,  $\theta$  is the vector of parameters  $\theta = (c, \phi_1, \dots, \phi_p, \sigma^2)$ ,  $\pi(\theta)$  is



**Bayesian Methods in Non-linear Time Series, Figure 1**  
Output growth and recession probabilities

the prior distribution that describes beliefs prior to observing the data,  $f(y|\theta)$  is the distribution of the sample conditional on the parameters,  $f(y)$  is the marginal distribution of the sample, and  $p(\theta|y)$  is the posterior distribution that describes the beliefs after observing the sample. Zellner [100], Bauwens, Lubrano, and Richard [7], Koop [58], Lancaster [61], and Geweke [39] cover Bayesian econometrics extensively and provide excellent introductions to relevant computational techniques.

We review the three most commonly used nonlinear models in three separate sections. We start each section by describing a baseline model and discussing possible extensions and applications (Matlab implementation of baseline models is available at <http://www.people.vcu.edu/~okorenok/share/mlab.zip>). Then we review the choice of prior, inference, tests against the linear hypothesis, and conclude with models selection. A short discussion of recent progress in incorporating regime changes into theoretical macroeconomic models concludes our survey.

Our survey builds on reviews of the TAR and STAR models in Tong [95], Granger and Terasvirta [41], Terasvirta [90], Bauwens, Lubrano, and Richard [7], Lubrano [63], Potter [74], Franses and van Dijk [34], van Dijk, Terasvirta, and Franses [98], and on reviews of the MAR models in Hamilton [46], Potter [74], and Kim and Nelson [51].

We limit our survey of nonlinear models only to the TAR, STAR, and MAR models. For a reader interested in a wider range of time series models from a Bayesian perspective, we recommend Steel's [84] survey: He overviews linear, as well as nonlinear, and parametric, as well as nonparametric, models.

### Threshold Autoregressive Model

A threshold regression was introduced by Quandt [75] and was extended to the threshold autoregressive model by Tong [92,93] and Tong and Lim [94]. Tong [95] had a great impact on popularizing TAR models.

We limit our baseline model to a single switching variable  $z_t$ . The choice of the switching variable depends on the purpose of the investigation. For the analysis of structural breaks at an unknown point in time, Perron and Vogelsang [70], as well as DeJong [24], among many others, use the time index ( $z_t = t$ ). For the purpose of prediction, Geweke and Terui [37], Chen and Lee [15], and others, use a lagged value of the time series ( $z_t = y_{t-d}$ ), the self-exciting threshold autoregressive (abbreviated, SETAR) model.

In our discussion, the number of lags in the model  $p$  and a delay  $d$  is fixed. We also limit the baseline model to the homoscedastic case so that the variance of  $\epsilon_t$  is constant in both regimes.

Introducing a more general notation,  $x'_t = (1, y_{t-1}, \dots, y_{t-p})$ ,  $\beta' = (c, \phi_1, \dots, \phi_p)$ , the two-regime TAR model becomes

$$\begin{aligned} y_t &= x'_t \beta_1 + \epsilon_t & \text{if } z_t < \tau & \text{(first regime),} \\ y_t &= x'_t \beta_2 + \epsilon_t & \text{if } z_t \geq \tau & \text{(second regime),} \end{aligned}$$

or more succinctly

$$y_t = [1 - I_{[\tau, \infty)}(z_t)]x'_t \beta_1 + I_{[\tau, \infty)}(z_t)x'_t \beta_2 + \epsilon_t, \quad (1)$$

where  $I_A(x)$  is an indicator function that is equal to one if  $x \in A$ , in particular  $I_{[\tau, \infty)}(z_t) = 1$  if  $z_t \in [\tau, \infty)$ . The indicator function introduces the abrupt transition between regimes. It is convenient to rewrite the model in a more compact form

$$y_t = x'_t(\tau)\beta + \epsilon_t, \quad (2)$$

where  $x'_t(\tau) = (x'_t, I_{[\tau, \infty)}(z_t)x'_t)$  and  $\beta' = (\beta'_1, \delta')$  with  $\delta = \beta_2 - \beta_1$ .

If the number of observations in regime  $i$  is less than or equal to the number of parameters, we cannot estimate parameters, or the model is not identified. In the Bayesian inference, we resolve the identification problem by restricting the region of possible parameter values to the one where the number of observations per regime is greater than the number of regressors.

The baseline model can be extended in several ways. First, we can allow the variance of the error term to differ in each regime. In this case, we rescale the data and introduce an additional parameter  $\phi = \sigma_2^2/\sigma_1^2$ , as in Lubrano [63]. Second, we can allow the number of lags to differ in each regime. Then  $p$  equals to  $\max\{p_1, p_2\}$ .

A more substantial change is required if we want to increase the number of regimes  $r$ . We can either use a single transition variable

$$y_t = x_t \beta_i(t) + \sigma_i(t) \epsilon_t,$$

where  $i(t) = 1$  if  $z_t < \tau_1$ ,  $i(t) = 2$  if  $\tau_1 \leq z_t < \tau_2$ , ...,  $i(t) = r$  if  $\tau_{r-1} \leq z_t$ ; or we can use a combination of two (or more) transition variables as in Astatkie, Watts, and Watt [5], where first stage transition is nested in the second stage transition

$$\begin{aligned} y_t &= [(1 - I_{[\tau_1, \infty)}(z_{1t}))x'_t \beta_1 + I_{[\tau_1, \infty)}(z_{1t})x'_t \beta_2] \\ &\quad \cdot [1 - I_{[\tau_2, \infty)}(z_{2t})] \\ &\quad + [(1 - I_{[\tau_1, \infty)}(z_{1t}))x'_t \beta_3 + I_{[\tau_1, \infty)}(z_{1t})x'_t \beta_4] \\ &\quad \cdot I_{[\tau_2, \infty)}(z_{2t}) + \epsilon_t, \end{aligned}$$

nested TAR model.

Also, we can treat either the choice of number of lags, the delay, or the number of regimes as an inference problem. Then  $p$ ,  $d$ , and  $r$  are added to the vector of the model parameters, as in Geweke and Terui [37] and Koop and Potter [57].

Finally, the univariate TAR model can be extended to describe a vector of time series as in Tsay [96]. The  $n$  dimensional two-regime TAR model can be specified in a manner similar to Eq. (1) as

$$\begin{aligned} Y_t &= [1 - I_{[\tau, \infty)}(z_t)](C_1 + \Phi_{11}Y_{t-1} + \dots + \Phi_{1p}Y_{t-p}) \\ &\quad + I_{[\tau, \infty)}(z_t)(C_2 + \Phi_{21}Y_{t-1} + \dots + \Phi_{2p}Y_{t-p}) + \epsilon_t, \end{aligned}$$

where  $Y_t = (y_{1t}, \dots, y_{nt})'$  is a  $(n \times 1)$  vector,  $C_1$  is a  $(n \times 1)$  vector,  $\Phi_{ji}$ ,  $j = 1, 2$ ,  $i = 1, \dots, p$  are  $(n \times n)$  matrices, and  $\epsilon_t = (\epsilon_{1t}, \dots, \epsilon_{nt})$  is a vector of error terms with mean zero and positive definite covariance matrix  $\Sigma$ .

The TAR model has a wide range of applications. Tiao and Tsay [91], Potter [73], Pesaran and Potter [71], Rothman [78], and Koop and Potter [54] demonstrate both statistically significant and economically important nonlinearities in the US business cycle. Pfann, Schotman, and Tschernig [72] find strong evidence of high volatility and low volatility regimes in the behavior of US short-term interest rates. Dwyer, Locke, and Yu [26], Martens, Kofman, and Vorst [66], and Forbes, Kalb, and Kofman [33] describe the relationship between spot and futures prices of the S&P 500 index and model financial arbitrage in these markets as a threshold process. Obstfeld and Taylor [68] study the law of one price and purchasing power parity convergences and find strong evidence of two regimes. They demonstrate fast, months rather than years, convergence when price differences are higher than transaction costs, and slow or no convergence otherwise.

To simplify the exposition, our discussion of inference for all models will be conditional on the initial observations in the sample. We assume that  $y_{1-p}, \dots, y_0$  are observable. Two alternative treatments are possible.

One can treat the initial observations as unobserved random variables and include the marginal density of initial observations into the likelihood. Alternatively, in the Bayesian analysis, one can treat the initial observations as any other parameter and augment the parameter space,  $\theta$ , with  $y_{1-p}, \dots, y_0$ .

### Prior

The first step in Bayesian inference is to formalize prior beliefs about the model's parameters by choosing functional forms and parameters of prior distributions.

The prior density for  $\tau$  depends on our choice of  $z_t$ . First, we can limit the prior support by the minimum and the maximum of  $z_t$ . Second, if  $z_t = t$  the threshold is a date, and so the prior density is naturally discrete. If, however,  $z_t = y_{t-d}$ , the threshold  $\tau$  is continuous and so is the prior density.

For a model to be identified, we restrict the support of the prior density to the region where the number of observations per regime is greater than the number of regressors. We assign an equal weight to the entire support to get the 'non-informative' prior for  $\tau$  that is proportional to a constant

$$\pi(\tau) \propto I_{[z_{(k_1)}, z_{(T-k_2)}]}(\tau), \quad (3)$$

where  $k_1$  and  $k_2$  are the number of regressors in the first and second regimes, and the subscript  $(t)$  indicates the order in the sample,  $z_{(1)} \leq z_{(2)} \leq \dots \leq z_{(T)}$ . For example,  $z_{(1)} = 1$  and  $z_{(T)} = T$  if  $z_t$  is a time index since the ordering is natural. For an alternative prior distribution of  $\tau$  see Ferreira [31].

We assume that the prior density for  $\beta$  and  $\sigma^2$  is independent of the prior density for  $\tau$ . Also, because, conditional on  $\tau$ , the model (2) is linear, we use the natural conjugate prior for  $\beta$  and  $\sigma^2$

$$\begin{aligned} \pi(\beta|\sigma^2) &= N(\beta|\beta_0, \sigma^2 M_0^{-1}), \\ \pi(\sigma^2) &= IG_2(\sigma^2|v_0, s_0), \end{aligned}$$

where  $IG_2(\cdot)$  denotes the density of the Inverted Gamma-2 distribution. The functional form of the Inverted Gamma-2 density is given by

$$IG_2(\sigma^2|v, s) = \Gamma\left(\frac{v}{2}\right)^{-1} \left(\frac{s}{2}\right)^{\frac{v}{2}} (\sigma^2)^{-\frac{1}{2}(v+2)} \exp\left(-\frac{s}{2\sigma^2}\right).$$

The natural conjugate prior allows us to use analytical integration that considerably simplifies the inference.

### Estimation

The next step of the Bayesian analysis is to combine sample information with our prior beliefs to form the posterior beliefs. Given prior distributions, we update prior distributions with the sample likelihood into posterior distributions using Bayes' theorem. The posterior distribution can be further summarized for each parameter with its marginal expectation and variance.

Using the assumption of Normal errors, the likelihood function of the model (2) is

$$f(\beta, \sigma^2, \tau|y) \propto \sigma^{-T} \exp\left\{-\frac{1}{2\sigma^2} \sum (y_t - x'_t(\tau)\beta)^2\right\}. \quad (4)$$

The posterior density is a product of the prior and the likelihood

$$p(\beta, \sigma^2, \tau|y) = \pi(\beta|\sigma^2)\pi(\sigma^2)\pi(\tau)f(\beta, \sigma^2, \tau|y). \quad (5)$$

Conditional on the threshold parameter, model (2) is linear. Applying the results from the standard natural conjugate analysis in the linear regression model (for details see Zellner [100]), the posteriors density of  $\beta$ , conditional on threshold and the data, can be obtained by integrating the posterior with respect to  $\sigma^2$

$$\begin{aligned} p(\beta|\tau, y) &= \int p(\beta, \sigma^2|\tau, y) d\sigma^2 \\ &= t(\beta|\tau, s(\tau), M(\tau), v), \end{aligned} \quad (6)$$

where  $t(\cdot)$  denotes the density of the multivariate Student t-distribution with

$$\begin{aligned} M(\tau) &= M_0 + \sum x_t(\tau)'x_t(\tau), \\ \beta(\tau) &= M(\tau)^{-1} \left( \sum x_t(\tau)y_t + M_0\beta_0 \right), \\ s(\tau) &= s_0 + \beta_0' M_0 \beta_0 + \sum y_t^2 - \beta_0'(\tau) M(\tau) \beta_0(\tau), \\ v &= v_0 + T. \end{aligned}$$

Further, by integrating Eq. (6) with respect to  $\beta$ , we obtain the marginal posterior density for  $\tau$ , which is proportional to the inverse of the integrating constant of  $t(\beta|\tau, s(\tau), M(\tau), v)$  times the threshold prior density

$$p(\tau|y) \propto s(\tau)^{-v/2} |M(\tau)|^{-1/2} \pi(\tau). \quad (7)$$

Though analytical integration of this function is not available, the fact that it is a univariate function defined on bounded support greatly simplifies the numerical integration.



By integrating numerically the posterior for  $\beta$  conditional on the threshold and the data, we find marginal posterior density for  $\beta$

$$p(\beta|y) = \int p(\beta|\tau, y)p(\tau|y) d\tau.$$

Finally, using analytical results for the expectation of the conditional density  $\beta$ , we can find the marginal moments of  $\beta$  by integrating only over  $\tau$

$$\begin{aligned} E(\beta|y) &= \int E(\beta|\tau, y)p(\tau|y) d\tau, \\ \text{Var}(\beta|y) &= \int \text{Var}(\beta|\tau, y)p(\tau|y) d\tau \\ &\quad + \int (E(\beta|\tau, y) - E(\beta|y))(E(\beta|\tau, y) \\ &\quad - E(\beta|y))' p(\tau|y) d\tau. \end{aligned}$$

Similarly, applying the results from the standard natural conjugate analysis, we obtain the posterior density of  $\sigma^2$  conditional on the threshold and the data. Then we integrate out  $\tau$  numerically to get the marginal posterior density for  $\sigma^2$

$$p(\sigma^2|y) = \int IG_2(\sigma^2|v, s(\tau))p(\tau|y) d\tau,$$

and the marginal moments  $E(\sigma^2|y)$  and  $\text{Var}(\sigma^2|y)$ .

### Testing for Linearity and Model Selection

After estimating the TAR model, we might ask whether our data are best characterized by two regimes or a single regime? Model (2) becomes linear when both regimes have identical regression coefficients, so that the difference  $\beta_1 - \beta_2 = \delta$  is zero. There are two methods to the null hypothesis test  $H_0 : \delta = 0$ . The first approach is the Bayesian equivalent of the F-test. Taking into account that  $\beta$  conditional on  $\tau$  has a Student t-distribution and that the linear transformation of a Student random vector is also a Student, the quadratic transformation of  $\delta$

$$\xi(\delta|\tau, y) = (\delta - \delta(\tau))' M_{22.1}(\tau) (\delta - \delta(\tau)) \frac{T-k}{k_2 s(\tau)} \quad (8)$$

has a Fisher distribution, where

$$M_{22.1}(\tau) = M_{22}(\tau) - M_{21}(\tau) M_{11}^{-1}(\tau) M_{12},$$

and  $\delta(\tau)$  is our estimate.  $M(\tau)$  is partitioned by dividing  $\beta$  into  $\beta_1$  and  $\delta$ . The posterior ‘p-value’ of the Bayesian

F-test gives the unconditional probability that  $\xi(\delta|y)$  exceeds  $\xi(\delta = 0|y)$ . It can be computed numerically as

$$\begin{aligned} \Pr(\xi(\delta) > \xi(\delta = 0|y)) \\ = \int F(\xi(\delta = 0|y), k_2, T-k) \cdot p(\tau|y) d\tau, \quad (9) \end{aligned}$$

where  $F(\xi(\delta = 0|y), k_2, T-k)$  is the Fisher distribution function with  $k_2$  and  $T-k$  degrees of freedom. The null hypothesis is accepted if, for example,  $\Pr(\xi(\delta) > \xi(\delta = 0|y))$  is larger than 5%.

The second approach, the posterior odds, is more general, and can also be used to select the number of lags  $p$ , the delay parameter  $d$ , or the number of regimes  $r$ . Koop and Potter [55,56] advocate and illustrate this approach in the context of the TAR model. To choose between two competing models,  $m_1$  with  $\theta_1 = (\beta_1, \delta, \tau, \sigma^2)$  and  $m_2$  with  $\theta_2 = (\beta_1, 0, \tau, \sigma^2)$ , we calculate the posterior odds ratio

$$po_{12} = \frac{f(y|m_1)\pi(m_1)}{f(y|m_2)\pi(m_2)},$$

where  $\pi(m_i)$  is the prior probability for the model  $i$ , and  $f(y|m_i)$  is the marginal likelihood or marginal density of the sample. Since  $f(y|m_i)$  is a normalizing constant of the posterior density, it can be calculated as

$$f(y|m_i) = \int f(y|\theta_i, m_i)\pi(\theta_i|m_i) d\theta_i.$$

With a ‘non-informative’ prior that assigns equal weight to each model, the posterior odds reduces to the ratio of marginal likelihoods, or the Bayes factor. Again, applying the standard natural conjugate analysis of the linear regression model, the marginal likelihood for model  $i$  is

$$\begin{aligned} f(y|m_i) &= \int \frac{\Gamma\left(\frac{v(\tau_i|m_i)}{2}\right) s_0^{\frac{v_0}{2}}}{\Gamma\left(\frac{v_0}{2}\right) \pi^{\frac{T}{2}}} s(\tau_i|m_i)^{-\frac{v(\tau_i|m_i)}{2}} \\ &\quad \cdot \left(\frac{|M_0|}{|M(\tau_i|m_i)|}\right)^{\frac{1}{2}} \pi(\tau_i|m_i) d\tau, \quad (10) \end{aligned}$$

which can be calculated numerically. The model with the highest marginal likelihood is preferred.

### Smooth Transition Autoregressive Model

In some applications, imposing an abrupt transition between regimes might be undesirable. For example, if the initial estimate of output is slightly below the threshold, even a small upward revision will result in a substantial change of the forecast in the TAR model. Bacon and

Watts [6], in a regression model context, and Chan and Tong [14], in the TAR model context, propose to make the transition between regimes smooth. Terasvirta [89] develops a modeling cycle for the STAR model that includes specification, estimation, and evaluation stages as in the Box and Jenkins [9] modeling cycle for the linear time series model.

In the STAR model, a smooth transition is imposed by replacing the indicator function in Eq. (1) by the cumulative distribution function

$$y_t = [1 - F(\gamma(z_t - \tau))]x'_t\beta_1 + F(\gamma(z_t - \tau))x'_t\beta_2 + \epsilon_t. \quad (1a)$$

Terasvirta [89] uses the logistic function

$$F(\gamma(z_t - \tau)) = \frac{1}{1 + \exp(-\gamma(z_t - \tau))},$$

where  $\gamma \in [0, \infty)$  determines the degree of smoothness. As  $\gamma$  increases, smoothness decreases. In the limit, as  $\gamma$  approaches infinity,  $F(\cdot)$  becomes an indicator function, with  $F(\gamma(z_t - \tau)) \sim 1$  when  $z_t \geq \tau$ . We can rewrite Eq. (1a) as

$$y_t = x'_t(\gamma, \tau)\beta + \epsilon_t, \quad (2a)$$

where  $x'_t(\gamma, \tau) = (x'_t, F(\gamma(z_t - \tau))x'_t)$ .

Note that the identification problem discussed for the TAR model does not occur in the STAR model. We cannot have fewer observations than regressors because we no longer classify observations into regimes. The new parameter  $\gamma$ , however, introduces a new identification problem. If  $\gamma = 0$ , the logistic function equals  $\frac{1}{2}$  for any value of  $\tau$ , so  $\tau$  is not identified. Also  $x'_t(\gamma, \tau)$  is perfectly collinear unless the two regimes have no common regressors. Perfect collinearity implies that  $\delta$  is also not identified. As in the TAR model, we choose such prior densities that resolve the identification problem.

The baseline model can be extended in several directions. Generally, the transition function  $F(\cdot)$  is not limited to the logistic function. Any continuous, monotonically increasing function  $F(\cdot)$  with  $F(-\infty) = 0$  and  $F(\infty) = 1$  can be used. For example, the popular alternative to the logistic function is the exponential function

$$F(\gamma(z_t - \tau)) = 1 - \exp(-\gamma(z_t - \tau)^2).$$

In the regression model context, Bacon and Watts [6] show that results are not sensitive to the choice of  $F(\cdot)$ . As in the TAR model, we can increase the number of regimes either with a single transition variable

$$y_t = x'_t\beta_1 + F(\gamma_1(z_t - \tau_1))x'_t(\beta_2 - \beta_1) + \dots \\ + F(\gamma_r(z_t - \tau_r))x'_t(\beta_r - \beta_{r-1}) + \epsilon_t,$$

or with a combination of transition variables

$$y_t = [(1 - F(\gamma_1(z_{1t} - \tau_1)))x'_t\beta_1 + F(\gamma_1(z_{1t} - \tau_1))x'_t\beta_2] \\ \cdot [(1 - F(\gamma_2(z_{2t} - \tau_2)))] \\ + [(1 - F(\gamma_1(z_{1t} - \tau_1)))x'_t\beta_3 \\ + F(\gamma_1(z_{1t} - \tau_1))x'_t\beta_4] \cdot [F(\gamma_2(z_{2t} - \tau_2))] + \epsilon_t.$$

See van Dijk and Franses [97] for a discussion of the multiple regime STAR model.

Also, we can treat the choice of number of lags  $p$ , delay  $d$ , or number of regimes  $r$  as an inference problem, adding  $p$ ,  $d$ , and  $r$  to the vector of parameters in the model. In addition, we can allow the variance of the error term to change between regimes, or more generally, use an autoregressive conditional heteroscedasticity form as in Lundbergh and Terasvirta [64], or a stochastic volatility form as in Korenok and Radchenko [59].

Finally, similar to the TAR model, the univariate STAR model can be extended to model a vector of time series as in Granger and Swanson [42]. The  $n$  dimensional two-regime STAR model can be specified as

$$Y_t = [1 - F(\gamma(z_t - \tau))](C_1 + \Phi_{11}Y_{t-1} + \dots + \Phi_{1p}Y_{t-p}) \\ + F(\gamma(z_t - \tau))(C_2 + \Phi_{21}Y_{t-1} + \dots + \Phi_{2p}Y_{t-p}) \\ + \epsilon_t,$$

where we use the same notation as in the multivariate TAR model.

Applications of the STAR model include models of the business cycles, real exchange rates, stock and futures prices, interest rates, and monetary policy. Terasvirta and Anderson [88] and van Dijk and Franses [97] demonstrate nonlinearities in the US business cycles. Skalin and Terasvirta [82] find similar nonlinearities in Swedish business cycles. Michael, Nobay, and Peel [67], Sarantis [80], and Taylor, Peel, and Sarno [87] show that the real exchange rate nonlinearly depends on the size of the deviation from purchasing power parity; Lundbergh and Terasvirta [65] and Korenok and Radchenko [59] use the STAR model to fit the behavior of exchange rates inside a target zone. Taylor, van Dijk, Franses, and Lucas [86] describe the nonlinear relationship between spot and futures prices of the FTSE100 index. Anderson [1] uses the STAR model to study yield movements in the US Treasury Bill Market. Finally, Rothman, van Dijk, and Franses [79] find evidence of a nonlinear relationship between money and output; Weise [99] demonstrates that monetary policy has a stronger effect on output during recessions.

### Prior

As in the TAR model, the natural conjugate priors for  $\beta$  and  $\sigma^2$  facilitate analytical integration. Bauwens, Lubrano, and Richard [7] impose the identification at  $\gamma = 0$  by modifying the prior density of  $\beta$

$$\pi(\beta|\sigma^2, \gamma) = N(\beta|0, \sigma^2 M_0^{-1}(\gamma)),$$

where, assuming prior independence between  $\beta_1$  and  $\delta$ ,  $M_0$  is defined as

$$M_0(\gamma) = \begin{pmatrix} M_{0,11} & 0 \\ 0 & M_{0,22}/\exp(\gamma) \end{pmatrix}.$$

As  $\gamma$  gets closer to zero, the prior variance falls, increasing precision around  $\delta = 0$ . The choice of  $\delta = 0$  is consistent with the linear hypothesis, which can be formulated as either  $\delta = 0$  or  $\gamma = 0$ . When  $\gamma$  is positive, prior precision about  $\delta = 0$  decreases as variance rises, so more weight is given to the information in the sample. We keep the natural conjugate prior of  $\sigma^2$  without modifications.

We do not modify the prior for the threshold parameter  $\tau$ . When  $\gamma$  is large, the smooth transition function is close to the step transition function. Thus, we prefer to limit the prior to the region where the number of observations per regime is greater than the number of regressors to avoid the TAR identification problem.

The prior for the smoothness parameter,  $\gamma$ , cannot be ‘non-informative’ or flat. As  $\gamma \rightarrow \infty$  the smooth transition function becomes a step transition with a strictly positive likelihood. This means that the marginal likelihood function of  $\gamma$  is not integrable. To avoid the integration problem, Bauwens, Lubrano, and Richard [7] use the truncated Cauchy density

$$\pi(\gamma) \propto (1 + \gamma^2)^{-1} I_{[0, \infty)}(\gamma).$$

### Estimation

Inference in the STAR model follows the TAR methodology, taking into account the additional parameter  $\gamma$ , and the new definitions of  $M_0(\gamma)$  and  $x_t(\tau, \gamma)$ .

In particular, the likelihood function of model (2a) is

$$f(\beta, \sigma^2, \tau, \gamma|y) \propto \sigma^{-T} \exp \left\{ -\frac{1}{2\sigma^2} \sum (y_t - x'_t(\tau, \gamma)\beta)^2 \right\}, \quad (4a)$$

the posterior density is

$$p(\beta, \sigma^2, \tau, \gamma|y) = \pi(\beta|\sigma^2)\pi(\sigma^2)\pi(\tau)\pi(\gamma) f(\beta, \sigma^2, \tau, \gamma|y), \quad (5a)$$

and the joint posterior density of  $\tau$  and  $\gamma$  is proportional to the inverse of the integrating constant of the Student  $t$ -density  $t(\beta|\beta(\tau, \gamma), s(\tau, \gamma), M(\tau, \gamma), \nu)$  times the prior densities for  $c$  and  $\gamma$

$$p(\tau, \gamma|y) \propto |s(\tau, \gamma)|^{-(T-k)/2} |M(\tau, \gamma)|^{-1/2} \pi(\tau)\pi(\gamma), \quad (7a)$$

where

$$\begin{aligned} M(\tau, \gamma) &= M_0(\gamma) + \sum x_t(\tau, \gamma)' x_t(\tau, \gamma), \\ \beta(\tau, \gamma) &= M(\tau, \gamma)^{-1} \left( \sum x_t(\tau, \gamma) y_t + M_0(\gamma) \beta_0 \right), \\ s(\tau, \gamma) &= s_0 + \beta_0' M_0(\gamma) \beta_0 \\ &\quad + \sum y_t^2 - \beta'(\tau, \gamma) M(\tau, \gamma) \beta(\tau, \gamma), \\ \nu &= \nu_0 + T. \end{aligned}$$

This function is bivariate and can be integrated numerically with respect to  $\tau$  and  $\gamma$ . Then, as in the TAR model, we use numerical integration to obtain marginal densities and moments for  $\beta$  and  $\sigma^2$ .

Compared to the TAR model,  $\beta_1$  and  $\beta_2$  cannot be interpreted as regression coefficients in regime 1 and regime 2. Smooth transition implies that the effect of change in  $x_t$  on  $y_t$  is a weighted average of two regimes with weights changing from one observation to the other.

### Testing for Linearity and Model Selection

The STAR model becomes linear when either  $\delta = 0$  or  $\gamma = 0$ . The test for  $H_0 : \delta = 0$  is equivalent to the test in the TAR model. The quadratic transformation of  $\delta$

$$\begin{aligned} \xi(\delta|\tau, \gamma, y) &= (\delta - \delta(\tau, \gamma))' M_{22.1}(\tau, \gamma) (\delta - \delta(\tau, \gamma)) \frac{T-k}{k_2 s(\tau, \gamma)}, \end{aligned} \quad (8a)$$

where

$$M_{22.1}(\tau, \gamma) = M_{22}(\tau, \gamma) - M_{21}(\tau, \gamma) M_{11}^{-1}(\tau, \gamma) M_{12}(\tau, \gamma),$$

has a Fisher distribution. We can find the posterior ‘p-value’ of the Bayesian F-test numerically as

$$\begin{aligned} \Pr(\xi(\delta) > \xi(\delta = 0)|y) &= \iint F(\xi(\delta = 0|y), k_2, T-k) p(\tau, \gamma|y) d\tau d\gamma. \end{aligned} \quad (9a)$$

The null hypothesis is accepted, for example, if  $\Pr(\xi(\delta) > \xi(\delta = 0)|y)$  is larger than 5%.

The test for  $H_0 : \gamma = 0$  can be conducted using the 95% highest posterior density interval (abbreviated, HPDI), defined as the smallest interval with 95% probability of  $\gamma$  to be in the interval

$$\max_h \text{PDI}(h) = \left\{ \gamma \mid \int p(\tau, \gamma) \pi(\tau) d\tau \geq h \right\},$$

s.t.  $\text{Pr}(\text{PDI}(h)) \geq 0.95$ .

The null hypothesis is accepted, for example, if  $\gamma = 0$  is inside the 95% HPDI.

As in the TAR model, linearity tests and model selection can be conducted using posterior odds. In the STAR model, the marginal likelihood for model  $i$  is given by

$$f(y|m_i) = \iint \frac{\Gamma\left(\frac{v(\tau_i, \gamma_i|m_i)}{2}\right) s_0^{\frac{v_0}{2}}}{\Gamma\left(\frac{v_0}{2}\right) \pi^{\frac{T}{2}}} s(\tau_i, \gamma_i|m_i)^{-\frac{v(\tau_i, \gamma_i|m_i)}{2}} \cdot \left(\frac{|M_0|}{|M(\tau_i, \gamma_i|m_i)|}\right)^{\frac{1}{2}} \pi(\tau_i|m_i) \pi(\gamma_i|m_i) d\tau_i d\gamma_i, \quad (10a)$$

which can be calculated numerically. The model with the highest marginal likelihood is preferred.

### Markov-Switching Model

Unlike the threshold models, where the regime transition depends on a time index or on lagged values of  $y_t$ , the Markov-switching autoregressive model relies on a random variable,  $s_t$ . A Markov-switching regression was introduced in econometrics by Goldfeld and Quandt [40] and was extended to the Markov-switching autoregressive model by Hamilton [45].

As in the threshold models, we limit our baseline MAR model to two regimes that differ only in mean. The variance of the error term is constant. The number of lags  $p$  is determined by the model choice. The two-regime MAR model becomes

$$(y_t - \mu_{s_t}) = \sum_{i=1}^p \phi_i (y_{t-i} - \mu_{s_{t-i}}) + \epsilon_t, \quad (11)$$

$$\begin{aligned} \mu_{s_t} &= \mu_0 & \text{if } s_t = 0 & \text{ (first regime),} \\ \mu_{s_t} &= \mu_0 + \mu_1 & \text{if } s_t = 1 & \text{ (second regime),} \end{aligned}$$

where  $\mu_{s_t} = \mu_0 + s_t \mu_1$ . An unobserved discrete random variable  $s_t$  takes only integer values of 0 or 1. The transition probability  $\text{Pr}(s_t = j | s_{t-1} = i) = p_{ij}$  that state  $i$  will be followed by state  $j$  depends only on  $s_{t-1}$ , the first order Markov-switching process, with transition probability

matrix

$$P = \begin{pmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{pmatrix}.$$

Since we have only two possible regimes and  $p_{i1} + p_{i2} = 1$ , we estimate only two free parameters, the probabilities of remaining in the same regime  $p_{11}$  and  $p_{22}$ . We also assume that, conditional on previous history of states  $s = (s_1, \dots, s_T)'$ , the transition probabilities are independent of other parameters and the data.

In general, we do not have a clear association between regimes and the state indicator. This introduces an identification problem when we change regime identifiers, 0 and 1, and accordingly change  $\mu_0^* = \mu_0 + \mu_1$  and  $\mu_1^* = -\mu_1$ . For example, if  $s_t = 0$  during recessions, then the long run average during recessions is  $\mu_0$  and the long-run average during expansions is  $\mu_0 + \mu_1$ . On the other hand, if  $s_t = 0$  during expansions, then the long-run average during expansions is  $\mu_0^* = \mu_0 + \mu_1$  and the long-run average during recessions is  $\mu_0^* - \mu_1$  or  $\mu_1^* = -\mu_1$ .

The second identification problem occurs in the MAR model when  $\mu_1 = 0$ ; the model becomes linear. In this case, the conditional mean  $E(y_t | s_t = 0) = E(y_t | s_t = 1) = \mu_0$  is independent of the state realizations,  $s$ , and transition probability matrix,  $P$ . Neither  $s$  nor  $P$  are identified.

The baseline model can be extended in several directions. The Markov-switching component can be modified by increasing the number of regimes as in Calvet and Fisher [11] and Sims and Zha [81] or by increasing the order of the Markov-switching process so that  $s_t$  depends on  $s_{t-1}, \dots, s_{t-r}$ . Both changes can be incorporated by increasing the number of states in the baseline model, as in Hamilton [46].

Diebold, Lee, and Weinbach [22], Filardo [32], and Peria [69] relax the assumption of time invariant Markov-switching by making the transition probabilities depend on lagged values of  $y_t$ . In most applications, however, relatively few transitions between regimes makes it difficult to estimate the transition probabilities and restricts model choice to two or three regimes with time-invariant probabilities.

The error term can be modified by introducing regime-switching for the variance of the error term as in Hamilton and Susmel [47], and Cai [10]; by relaxing the assumption of Gaussian density for the error term as in Dueker [25]; or by specifying a general Markov-switching moving average structure for the error term as in Billio, Monfort, and Robert [8].

Finally, the univariate Markov-switching model can be extended to a multivariate model. Diebold and Rude-

busch [23] propose a model where a number of time series are driven by a common unobserved Markov-switching variable, the dynamic factor model. The dynamic factor model captures the fact that many economic series show similar changes in dynamic behavior during recessions. Krolzig [60] provides a detailed exposition of how the baseline model can be extended to the Markov-switching vector autoregressive model.

The applications of the MAR model include models of business cycles, interest rates, financial crises, portfolio diversification, options pricing, and changes in government policy. Hamilton [45], Filardo [32], Diebold and Rudebusch [23], Kim and Nelson [51], Kim and Piger [53], and Hamilton [48] find statistically significant evidence that expansionary and contractionary phases of the US business cycle are distinct. Hamilton [44], Cai [10], Garcia and Perron [35], Gray [43], Dueker [25], Smith [83], Hamilton [48], and Dai, Singleton, and Yang [18] describe dramatic changes in interest rate volatility associated with the OPEC oil shocks, the changes in the Federal Reserve operating procedures in 1979–1982, and the stock market crash of October 1987. Ang and Bekaert [3] show a similar increase in volatility in Germany during the reunification period. Jeanne and Masson [49] use the MAR model to describe the crisis of the European Monetary System in 1992–1993; Cerra and Saxena [13] find permanent losses in output after the Asian crisis. Ang and Bekaert [2] report that the correlation between international equity returns is higher during bear markets relative to bull markets. Radchenko [76] shows that gasoline prices respond faster to a permanent oil price change compared to a transitory change. Finally, Sims and Zha [81] document abrupt changes of shocks to US monetary policy, and Davig and Leeper [20] document the regime changes in fiscal policy.

### Prior

As in the threshold models, the natural conjugate priors facilitate considerably the integration of the posterior density. Conditional on  $s_t$ ,  $\mu_0$ , and  $\mu_1$ , the MAR model is linear

$$y_t(s_t) = x'_t(s_t)\tilde{\phi} + \epsilon_t, \quad (12)$$

where

$$\begin{aligned} y_t(s_t) &= y_t - \mu_{s_t}, \\ x'_t(s_t) &= (y_{t-1} - \mu_{s_{t-1}}, \dots, y_{t-p} - \mu_{s_{t-p}}), \end{aligned}$$

and  $\tilde{\phi} = (\phi_1, \dots, \phi_p)'$ . For the regression coefficient  $\tilde{\phi}$  and the variance of the error term  $\sigma^2$ , the natural conjugate prior is given by

gate prior is given by

$$\begin{aligned} \pi(\tilde{\phi}|\sigma^2) &= N(\tilde{\phi}|\tilde{\phi}_0, \sigma^2 M_{0,\tilde{\phi}}^{-1}) I_A(\tilde{\phi}), \\ \pi(\sigma^2) &= IG_2(\sigma^2|\nu_0, s_0), \end{aligned}$$

where  $A$  is a region where the roots of polynomial  $1 - \phi_1 L - \dots - \phi_p L^p = 0$  lie outside the complex unit circle. This restriction imposes stationarity on  $y_t(s_t)$ .

Conditional on  $s_t$  and  $\tilde{\phi}$ , the MAR model is also linear

$$y_t(\tilde{\phi}) = x'_t(\tilde{\phi})\tilde{\mu} + \epsilon_t, \quad (13)$$

where

$$\begin{aligned} y_t(\tilde{\phi}) &= y_t - \sum_{i=1}^p \phi_i y_{t-p}, \\ x'_t(\tilde{\phi}) &= \left( 1, s_t - \sum_{i=1}^p \phi_i s_{t-p} \right), \end{aligned}$$

and  $\tilde{\mu} = (\mu_0, \mu_1)'$ . The natural conjugate prior for  $\tilde{\mu}$  is

$$\pi(\tilde{\mu}) = N(\tilde{\mu}|\tilde{\mu}_0, M_{0,\tilde{\mu}}^{-1}) I_{(0,\infty)}(\mu_1),$$

where the indicator function imposes an identification constraint. In particular, we constrain the mean of the second regime to be greater than the mean of the first regime and in this way fix the order of regimes. We also impose  $\mu_1 \neq 0$ .

Kim and Nelson [51] show that the natural conjugate prior for the vector of transition probabilities  $\tilde{p} = (p_{11}, p_{22})'$  is

$$\pi(\tilde{p}) = B(p_{11}|\alpha_1, \beta_1) B(p_{22}|\alpha_2, \beta_2),$$

where  $B(\cdot)$  denotes the density of Beta distribution defined on the interval  $[0, 1]$ .

### Estimation

In the Bayesian approach, we add realizations of the vector of states to the model parameters:  $\theta = (\mu_0, \mu_1, \phi_1, \dots, \phi_p, \sigma, p_{11}, p_{22}, s_1, \dots, s_T)'$ . Analytical or numerical integration of the posterior density  $p(\theta|y)$ , where  $\theta$  is  $p + 5 + T \times 1$ , may be difficult.

Albert and Chib [4] developed inference methodology that overcomes the curse of dimensionality using Gibbs-sampling, a Markov chain Monte Carlo simulation method of integration. The technique was further refined by Kim and Nelson [50]. Monte Carlo integration takes random draws from the posterior density and, by averaging them, produces estimates of moments. In particular, Gibbs-sampling allows us to generate many draws



$\theta^{(g)}$ ,  $g = 1, \dots, G$ , from joint density of  $p(\theta|y)$  using only conditional densities  $p(\theta_i|\theta_{i \neq j}, y)$  either for all  $i$  or for blocks of parameters. The joint and marginal distribution of  $\theta^{(g)}$  converge at an exponential rate to the joint and marginal distribution of  $\theta$  under fairly weak conditions. Casella and George [12], Gelfand and Smith [36], and Geweke [38] provide the details.

To implement the Gibbs-sampling simulation, we have to describe the conditional posterior distributions for all parameters or parameter blocks. It is convenient to separate parameters into five blocks: the state vector  $s$ , the transition probabilities  $\tilde{p}$ , the regression coefficients  $\tilde{\phi}$  in the conditional linear model (12), the regression coefficients  $\tilde{\mu}$  in the conditional linear model (13), and the variance of the error term  $\sigma^2$ .

The state vector  $s$  is a first-order Markov process, which implies that given  $s_{t+1}$  all information, for example  $s_{t+2}, \dots, s_T$  and  $y_{t+1}, \dots, y_T$ , is irrelevant in describing  $s_t$ . Then the posterior density of  $s$  conditional on other parameters becomes

$$\begin{aligned} p(s|\tilde{p}, \tilde{\phi}, \tilde{\mu}, \sigma^2, y) \\ = p(s_T|\tilde{p}, \tilde{\phi}, \tilde{\mu}, \sigma^2, y) \prod_{t=1}^{T-1} p(s_t|s_{t+1}, \tilde{p}, \tilde{\phi}, \tilde{\mu}, \sigma^2, y^t), \end{aligned} \quad (14)$$

where  $y^t = (y_1, \dots, y_t)'$ . The functional form of the posterior density suggests that we can generate draw of the state vector recursively. First we generate the last element  $s_T$ . Then, conditional on  $s_T$ , we generate  $s_{T-1}$ . More generally, conditional on  $s_{t+1}$ , we generate  $s_t$  for  $t = T-1, T-2, \dots, 1$ .

To generate the state vector, Kim and Nelson [50] use the output from Hamilton's [45] filter. To facilitate exposition, we suppress the conditioning on parameters and consider first a model without lags.

Hamilton's filter starts from the observation that, before observing the data, the probability of finding the state in regime  $j$ ,  $\Pr(s_0 = j|y^0)$ , equals the unconditional probability,  $\Pr(s_t = j)$ , which is proportional to the eigenvector of  $P$  associated with unitary eigenvalue.

Using transition probabilities and the probability of observing regime  $j$  conditional on observations obtained through date  $t$ ,  $\Pr(s_t = j|y^t)$ , we predict the next period regime

$$\Pr(s_{t+1} = j|y^t) = \Pr(s_t = 0|y^t)p_{0j} + \Pr(s_t = 1|y^t)p_{1j}. \quad (15)$$

Once  $y_{t+1}$  is observed, we update the prediction using

Bayes rule

$$\begin{aligned} \Pr(s_{t+1} = j|y^{t+1}) &= \Pr(s_{t+1} = j|y_{t+1}, y^t) \\ &= \frac{f(y_{t+1}|s_{t+1} = j, y^t) \Pr(s_{t+1} = j|y^t)}{f(y_{t+1}|y^t)}, \end{aligned} \quad (16)$$

where the numerator is the joint probability of observing  $y_{t+1}$  and  $s_{t+1} = j$ , which is a product of the probability of observing  $y_{t+1}$  given that state  $s_{t+1}$  is in regime  $j$  (for example  $f(y_{t+1}|s_{t+1} = 0, y^t) = N(\mu_0, \sigma^2)$ ) and our prediction from Eq. (15). The denominator is the unconditional density of observing  $y_{t+1}$ , which is a sum of the numerator over all possible regimes

$$f(y_{t+1}|y^t) = \sum_j f(y_{t+1}|s_{t+1} = j, y^t) \Pr(s_{t+1} = j|y^t). \quad (17)$$

Starting from  $\Pr(s_0 = j|y^0)$ , the filter iterates through Eqs. (15)–(17) until we calculate  $\Pr(s_t = j|y^t)$  for every  $t$  and  $j$ . As a by-product of the filter we obtain the likelihood function

$$f(\tilde{\phi}, \tilde{\mu}, \tilde{p}, \sigma^2, s|y) = \prod_t f(y_{t+1}|y^t). \quad (18)$$

For the AR(1) model, the filter should be adjusted. Given  $\Pr(s_t = j|y^t)$ , we forecast the next period regime and the previous period regime jointly, taking one summand in Eq. (15) at a time

$$\Pr(s_{t+1} = j, s_t = i|y^t) = p_{ij} \Pr(s_t = i|y^t), \quad (15a)$$

for  $j = 0, 1$  and  $i = 0, 1$ . After  $y_{t+1}$  is observed, we update our prediction to

$$\begin{aligned} \Pr(s_{t+1} = j, s_t = i|y^{t+1}) \\ = \frac{f(y_{t+1}|s_{t+1} = j, s_t = i, y^t) \Pr(s_{t+1} = j, s_t = i|y^t)}{f(y_{t+1}|y^t)}, \end{aligned} \quad (16a)$$

where  $f(y_{t+1}|s_{t+1} = j, s_t = i, y^t)$  is the density of observing  $y_{t+1}$  given that state  $s_{t+1}$  is in regime  $j$  and state  $s_t$  is in regime  $i$  (for example  $f(y_{t+1}|s_{t+1} = 0, s_t = 0, y^t) = N(\mu_0 + \phi_1(y_t - \mu_0), \sigma^2)$ )

$$f(y_{t+1}|y^t) = \sum_j \sum_i f(y_{t+1}|s_{t+1} = j, s_t = i, y^t) \cdot \Pr(s_{t+1} = j, s_t = i|y^t). \quad (17a)$$

Summing (16a) over  $i$ ,

$$\Pr(s_{t+1} = j|y^{t+1}) = \sum_i \Pr(s_{t+1} = j, s_t = i|y^{t+1}), \quad (19)$$

finishes the iteration. Iterating through Eqs. (15a)–(17a) and (19) we get  $\Pr(s_t = j|y^t)$  for every  $t$  and  $j$ . The extension to a more general AR(p) model is similar.

The output of Hamilton's filter gives only the first term in the product (14), which is sufficient to generate  $s_T$ . To generate the other states  $s_t$  conditional on  $y^t$  and  $s_t + 1$ ,  $t = T - 1, T - 2, \dots, 1$ , we again use Bayes rule

$$\Pr(s_t = j|s_{t+1} = i, y^t) = \frac{p_{ji} \Pr(s_t = j|y^t)}{\sum_j p_{ji} \Pr(s_t = j|y^t)}, \quad (20)$$

where  $\Pr(s_t = j|y^t)$  is the output from Hamilton's filter. Since  $s_t$  is a discrete random variable taking on values 0 and 1, we can generate it by drawing random numbers from uniform distribution between 0 and 1, and comparing them to  $\Pr(s_t = 1|s_{t+1} = i, y^t)$ .

Conditional on other parameters in the model, the likelihood function of transition probabilities reduces to a simple count  $n_{ij}$  of transitions from state  $i$  to state  $j$

$$f(\tilde{p}|\tilde{\mu}, \tilde{\phi}, \sigma^2, s, y) = p_{11}^{n_{11}}(1 - p_{11})^{n_{12}} p_{22}^{n_{22}}(1 - p_{22})^{n_{21}},$$

which is the product of the independent beta distributions. The posterior distribution for the transition probabilities conditional on the other parameters is a product of independent beta distributions

$$\begin{aligned} p(\tilde{p}|\tilde{\phi}, \tilde{\mu}, \sigma^2, s, y) \\ = B(\alpha_1 + n_{11}, \beta_1 + n_{12}) \cdot B(\alpha_2 + n_{22}, \beta_2 + n_{21}). \end{aligned}$$

To derive posterior distributions for  $\tilde{\phi}$ ,  $\tilde{\mu}$ , and  $\sigma^2$  conditional on other parameters, we use standard results for a linear model with the natural conjugate priors. The natural conjugate priors are reviewed, for example, by Geweke [39], Koop [58], or Lancaster [61]. In particular, the conditional distribution of the regression coefficients is Normal

$$\begin{aligned} p(\tilde{\phi}|\tilde{p}, \tilde{\mu}, \sigma^2, s, y) \\ = N\left(\Sigma_{\phi}\left(\sigma^{-2}M_{0,\phi}\tilde{\phi}_0 + \sigma^{-2}\sum x_t(s)'y_t(s)\right), \Sigma_{\phi}\right) \\ \cdot I_A(\tilde{\phi}), \\ p(\tilde{\mu}|\tilde{p}, \tilde{\phi}, \sigma^2, s, y) \\ = N\left(\Sigma_{\mu}\left(M_{0,\mu}\tilde{\mu}_0 + \sigma^{-2}\sum x_t(\tilde{\phi})'y_t(\tilde{\phi})\right), \Sigma_{\mu}\right) \\ \cdot I_{(0,\infty)}(\mu_1), \end{aligned}$$

where

$$\begin{aligned} \Sigma_{\phi} &= \left(\sigma^{-2}M_{0,\phi} + \sigma^{-2}\sum x_t(s)'x_t(s)\right)^{-1}, \\ \Sigma_{\mu} &= \left(M_{0,\mu} + \sigma^{-2}\sum x_t(\tilde{\phi})'x_t(\tilde{\phi})\right)^{-1}. \end{aligned}$$

The conditional distribution for the variance of error term is Inverted Gamma-2

$$\begin{aligned} p(\sigma^2|\tilde{p}, \tilde{\phi}, \tilde{\mu}, s, y) \\ = IG_2\left(s_0 + \sum (y_t(s_t) - x_t'(s_t)\tilde{\phi})^2, \nu_0 + T\right). \end{aligned}$$

### Testing for Linearity and Model Selection

Given our prior, the linear model is not nested in the MAR model. To test against a linear model, we use the Bayes factor. We also use the Bayes factor to select the number of regimes and the number of lags.

The Bayes factor is a ratio of marginal likelihoods of the alternative models. To find the marginal likelihood, we need to integrate the product of the likelihood function and the prior density with respect to all parameters. Chib [16] shows that the marginal likelihood can be computed from the output of the Gibbs sampler requiring only that the integrating constants of the conditional posterior distributions be known. This requirement is satisfied for the natural conjugate priors.

From the Bayes's theorem it follows that the identity

$$f(y) = \frac{f(y|\theta)\pi(\theta)}{p(\theta|y)},$$

holds for any  $\theta$ . The complete functional form of the numerator is given by the product of the likelihood (18) and the prior densities. Chib suggests evaluating the denominator, the posterior density, at the posterior mode  $\theta^*$ . Then the posterior density at the posterior mode can be written as

$$\begin{aligned} p(\theta^*|y) &= p(\tilde{\mu}^*|y) p(\tilde{\phi}^*|\tilde{\mu}^*, y) \\ &\cdot p(\tilde{\sigma}^{2*}|\tilde{\mu}^*, \tilde{\phi}^*, y) p(\tilde{p}^*|y, \mu^*, \tilde{\phi}^*, \sigma^{2*}). \end{aligned}$$

The first term

$$\begin{aligned} p(\tilde{\mu}^*|y) &= \\ \int p(\tilde{\mu}^*|\tilde{\phi}, \sigma^2, \tilde{p}, s, y) p(\tilde{\phi}, \sigma^2, \tilde{p}, s|y) d\tilde{\phi} d\sigma^2 d\tilde{p} ds, \end{aligned}$$

can be estimated by averaging over the full conditional density

$$\hat{p}(\tilde{\mu}^*|y) = G^{-1} \sum_{g=1}^G p(\tilde{\mu}^*|\tilde{\phi}^{(g)}, \sigma^{2(g)}, \tilde{p}^{(g)}, s^{(g)}, y).$$

This estimate converges at an exponential rate to the true marginal distribution of  $\tilde{\mu}$ .

In the second term,

$$p(\tilde{\phi}|\tilde{\mu}^*, y) = \int p(\tilde{\phi}^*|\tilde{\mu}^*, \sigma^2, \tilde{p}, s, y) p(\sigma^2, \tilde{p}, s|\tilde{\mu}^*, y) d\sigma^2 d\tilde{p} ds,$$

the complete conditional density of  $\tilde{\phi}$  cannot be averaged directly because the Gibbs sampler does not provide draws conditional on  $\tilde{\mu}^*$ . We generate necessary draws by additional  $G$  iterations of the original Gibbs sampler, but instead of generating  $\tilde{\mu}$  we set it equal to  $\tilde{\mu}^*$ . Then the estimate of the second term

$$\hat{p}(\tilde{\phi}^*|\tilde{\mu}^*, y) = G^{-1} \sum_{g=G+1}^{2G} p(\tilde{\phi}^*|\tilde{\mu}^*, \sigma^{2(g)}, \tilde{p}^{(g)}, s^{(g)}, y),$$

converges at an exponential rate to the true  $p(\tilde{\phi}|\tilde{\mu}^*, y)$ . Similarly, by generating additional draws from the Gibbs sampler we compute  $\hat{p}(\tilde{\sigma}^{2*}|\tilde{\mu}^*, \tilde{\phi}^*, y)$  and  $\hat{p}(\tilde{p}^*|y, \mu^*, \tilde{\phi}^*, \sigma^{2*})$ .

Substituting our estimate of posterior density into marginal likelihood results in

$$\begin{aligned} \ln f(y) = & \ln f(y|\theta^*) + \ln \pi(\theta^*) - \ln \hat{p}(\tilde{\mu}^*|y) \\ & - \ln \hat{p}(\tilde{\phi}^*|\tilde{\mu}^*, y) - \ln \hat{p}(\tilde{\sigma}^{2*}|\tilde{\mu}^*, \tilde{\phi}^*, y) \\ & - \ln \hat{p}(\tilde{p}^*|y, \mu^*, \tilde{\phi}^*, \sigma^{2*}). \end{aligned}$$

The model with the highest marginal likelihood is preferred.

### Future Directions

Given the large volume of evidence collected in the non-linear time series, incorporating regime-switching policies and disturbances into general equilibrium models may lead to a better understanding of monetary and fiscal policies.

Over the years, the time series literature has collected substantial statistical evidence that output, unemployment, and interest rates in the US exhibit different behavior in recessions and expansions. Contrary to the real business cycle models in which short-run and long-run fluctuations have the same origin, the statistical evidence suggests that the forces that cause output to rise may be quite different from those that cause it to fall.

Also, many studies provide evidence that monetary and fiscal policies have changed substantially throughout US history. Taylor [85], Clarida, Gali, and Gertler [17], Romer and Romer [77], and Lubik and Schorfheide [62] show that, since the mid-1980s, the Fed reacted more forcefully to inflation. Favero and Monacelli [30] and Davig and Leeper [20] demonstrate that US fiscal policy

has fluctuated frequently responding to wars, recessions, and more generally to the level of debt. Sims and Zha [81], after extensive comparison of 17 regime-switching structural VAR models, report that their best-fitting model requires nine regimes to incorporate the large shocks, for example, generated by the OPEC oil embargo or the Vietnam War. They conclude that, "It is time to abandon the idea that policy change is best modelled as a once-and-for-all, nonstochastic regime switch" (p. 56).

The research by Davig and Leeper [19,20,21] and Farmer, Waggoner, and Zha [27,28,29] show considerable promise in introducing nonlinear regime-switching components into dynamic stochastic general equilibrium models. For example, Davig and Leeper [20] estimate regime-switching rules for monetary policy and tax policy and incorporate them into the otherwise standard new-Keynesian model. Unlike expansionary fiscal policy in the fixed-regime model, fiscal expansion in the regime-switching model increases inflation and output.

### Acknowledgments

The author is grateful to Bruce Mizrach, the Finance and Econometrics Section editor, as well as to David Harless, Carol S. Lehr, Ming Lo, Stan Radchenko, Philip Rothman, and Tara Sinclair for many helpful comments.

### Bibliography

1. Anderson HM (1997) Transaction costs and nonlinear adjustment towards equilibrium in the US treasury bill market. *Oxf Bull Econ Stat* 59:465–484
2. Ang A, Bekaert G (2002) International asset allocation with regime shifts. *Rev Financ Stud* 15:1137–1187
3. Ang A, Bekaert G (2002) Regime switches in interest rates. *J Bus Econ Stat* 20:163–197
4. Albert JH, Chib S (1993) Bayes inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts. *J Bus Econ Stat* 11(1):1–15
5. Astatkie T, Watts DG, Watt WE (1997) Nested threshold autoregressive NeTAR models. *Int J Forecast* 13:105–116
6. Bacon DW, Watts DG (1971) Estimating the transition between intersecting straight lines. *Biometrika* 62:525–534
7. Bauwens L, Lubrano M, Richard JF (1999) *Bayesian Inference in Dynamic Econometric Models*. Oxford University Press, New York
8. Billio M, Monfort A, Robert CP (1999) Bayesian estimation of switching ARMA models. *J Econ* 93:229–255
9. Box GEP, Jenkins GM (1970) *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco
10. Cai J (1994) A Markov model of switching-regime ARCH. *J Bus Econ Stat* 12:309–316
11. Calvet L, Fisher A (2004) Regime switching and the estimation of multifractal processes. *J Financ Econ* 2:49–83
12. Casella G, George EI (1992) Explaining the Gibbs sampler. *Am Stat* 46:167–174

13. Cerra V, Saxena SC (2005) Did output recover from the Asian crisis? *IMF Staff Papers* 52:1–23
14. Chan KS, Tong H (1986) On estimating thresholds in autoregressive models. *J Time Ser Anal* 7:178–190
15. Chen CWS, Lee JC (1995) Bayesian inference of threshold autoregressive models. *J Time Ser Anal* 16:483–492
16. Chib S (1995) Marginal likelihood from the Gibbs output. *J Am Stat Assoc* 90:1313–1321
17. Clarida R, Gali J, Gertler M (2000) Monetary policy rules and macroeconomic stability: evidence and some theory. *Q J Econ* 115:147–180
18. Dai Q, Singleton KJ, Yang W (2007) Regime shifts in a dynamic term structure model of US treasury bonds. *Rev Financ Stud* 20(5):1669–1706
19. Davig T, Leeper E (2005) Generalizing the Taylor principle. National Bureau of Economic Research, Working Paper No 11874
20. Davig T, Leeper E (2006) Fluctuating macro policies and the fiscal theory. In: Acemoglu D, Rogoff, Woodford M (eds) *NBER Macroeconomic Annual*. MIT Press, Cambridge
21. Davig T, Leeper E (2007) Generalizing the Taylor principle. *Am Econ Rev* 97(3):607–635
22. Diebold FX, Lee JH, Weinbach GC (1994) Regime switching with time-varying transition probabilities. In: Hargreaves C (ed) *Nonstationary Time Series Analysis and Cointegration*. Oxford University Press, Oxford
23. Diebold FX, Rudebusch GD (1996) Measuring business cycles: a modern perspective. *Rev Econ Stat* 78:67–77
24. DeJong DN (1996) A Bayesian Search for Structural Breaks in US GNP. In: Fomby TB (ed) *Advances in Econometrics: Bayesian Methods Applied to Time Series Data*, vol 11, part B. JAI Press, Greenwich, Connecticut, pp 109–146
25. Dueker M (1997) Markov switching in GARCH processes and mean-reverting stock-market volatility. *J Bus Econ Stat* 15:26–34
26. Dwyer GP, Locke P, Yu W (1996) Index arbitrage and nonlinear dynamics between the S&P 500 futures and cash. *Rev Financ Stud* 9:301–332
27. Farmer RE, Waggoner DF, Zha T (2006) Indeterminacy in a forward looking regime switching model. *NBER Working Paper* No 12540
28. Farmer RE, Waggoner DF, Zha T (2006) Minimal state variable solutions to Markov-switching rational expectations models. (unpublished manuscript)
29. Farmer RE, Waggoner DF, Zha T (2007) Understanding the New-Keynesian model when monetary policy switches regimes. (unpublished manuscript)
30. Favero CA, Monacelli T (2005) Fiscal policy rules and regime (in)stability: evidence from the US Manuscript, IGIER
31. Ferreira PE (1975) A Bayesian analysis of a switching regression model: known number of regimes. *J Am Stat Assoc* 70:370–374
32. Filardo AJ (1994) Business cycle phases and their transitional dynamics. *J Bus Econ Stat* 12:299–308
33. Forbes CS, Kalb GRJ, Kofman P (1999) Bayesian arbitrage threshold analysis. *J Bus Econ Stat* 17:364–372
34. Franses PH, van Dijk D (2000) *Nonlinear Time Series Models in Empirical Finance*. Cambridge University Press, Cambridge
35. Garcia R, Perron P (1996) An analysis of real interest under regime shift. *Rev Econ Stat* 78:111–125
36. Gelfand AE, Smith AFM (1990) Sampling-based approaches to calculating marginal densities. *J Am Stat Assoc* 85(410):398–409
37. Geweke J, Terui N (1993) Bayesian threshold autoregressive models for nonlinear time series. *J Time Ser Anal* 14:441–454
38. Geweke J (1999) Using simulation methods for Bayesian econometric models: inference, development and communication. *Econ Rev* 18:1–127
39. Geweke J (2005) *Contemporary Bayesian Econometrics and Statistics*. Wiley, Hoboken
40. Goldfeld SM, Quandt RE (1973) A Markov model for switching regressions. *J Econ* 1:3–16
41. Granger CWJ, Terasvirta T (1993) *Modeling Nonlinear Economic Relationships*. Oxford University Press, Oxford
42. Granger CWJ, Swanson NR (1996) Future developments in the study of cointegrated variables. *Oxf Bull Econ Stat* 58:537–553
43. Gray SF (1996) Modeling the conditional distribution of interest rates as a regime-switching process. *J Financ Econ* 42:27–62
44. Hamilton JD (1988) Rational-expectations econometric analysis of changes in regime: an investigation of the term structure of interest rates. *J Econ Dyn Control* 12:385–423
45. Hamilton JD (1989) A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57(2):357–384
46. Hamilton JD (1994) *Time Series Analysis*. Princeton University Press, Princeton
47. Hamilton JD, Susmel R (1994) Autoregressive conditional heteroskedasticity and changes in regime. *J Econ* 64:207–333
48. Hamilton JD (2005) What's real about the business cycle? *Fed Reserve Bank St Louis Rev* 87(4):435–452
49. Jeanne O, Masson P (2000) Currency crises, sunspots, and Markov-switching regimes. *J Int Econ* 50:327–350
50. Kim CJ, Nelson CR (1998) Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime-switching. *Rev Econ Stat* 80(2):188–201
51. Kim CJ, Nelson CR (1999) Friedman's plucking model of business fluctuations: tests and estimates of permanent and transitory components. *J Money Credit Bank* 31:317–334
52. Kim CJ, Nelson CR (1999) *State-Space Models with Regime Switching: Classical and Gibbs-sampling Approaches with Applications*. MIT Press, Cambridge
53. Kim CJ, Piger J (2002) Common stochastic trends, common cycles, and asymmetry in economic fluctuations. *J Monet Econ* 49:1189–1211
54. Koop G, Potter SM (1999) Dynamic asymmetries in US unemployment. *J Bus Econ Stat* 17:198–312
55. Koop G, Potter SM (1999) Bayes factors and nonlinearity: evidence from economic time series. *J Econ* 88:251–281
56. Koop G, Potter SM (2000) Nonlinearity, structural breaks or outliers in economic time series? In: Barnett B, Johansen S (eds) *Nonlinear Econometric Modeling in Time Series*. Cambridge University Press, Cambridge
57. Koop G, Potter SM (2003) Bayesian analysis of endogenous delay threshold models. *J Bus Econ Stat* 21(1):93–103
58. Koop G (2003) *Bayesian Econometrics*. Wiley, Chichester
59. Korenok O, Radchenko R (2005) The smooth transition autoregressive target zone model with the Gaussian stochastic

- volatility and TGARCH error terms with applications. VCU Economics Department, No 0505
60. Krolzig HM (1997) Markov-Switching Vector Autoregressions: Modeling, Statistical Inference, and Application to Business Cycle Analysis. Springer, Berlin
  61. Lancaster T (2004) An Introduction to Modern Bayesian Econometrics. Blackwell Publishing, Malden
  62. Lubik TA, Schorfheide F (2004) Testing for indeterminacy: an application to US monetary policy. *Am Econ Rev* 94:190–217
  63. Lubrano M (1999) Bayesian Analysis of Nonlinear Time Series Models with a threshold. In: Barnett WA, Hendry DF, Hylleberg S, Terasvirta T, Tjostheim D, Wurts A (eds) *Nonlinear Econometric Modeling*. Cambridge University Press, Cambridge
  64. Lundbergh S, Terasvirta T (1998) Modelling economic high-frequency time series with STAR-GARCH models. Working Paper Series in Economics and Finance No. 291, Stockholm School of Economics
  65. Lundbergh S, Terasvirta T (2006) A time series model for an exchange rate in a target zone with applications. *J Econ* 131:579–609
  66. Martens M, Kofman P, Vorst ACF (1998) A threshold error correction for intraday futures and index returns. *J Appl Econ* 13:245–263
  67. Michael P, Nobay AR, Peel DA (1997) Transaction costs and nonlinear adjustment in real exchange rates: an empirical investigation. *J Political Econ* 105:862–879
  68. Obstfeld M, Taylor AM (1997) Nonlinear aspects of goods-market arbitrage and adjustment: Heckscher's commodity points revisited. *J Japan Int Econ* 11:441–479
  69. Peria MSM (2002) A regime-switching approach to the study of speculative attacks: a focus on EMS crises. In: Hamilton JD, Raj B (eds) *Advances in Markov-Switching Models*. Physica-Verlag, Heidelberg
  70. Perron P, Vogelsang TJ (1992) Nonstationarity and level shifts with an application to purchasing power parity. *J Bus Econ Stat* 10:301–320
  71. Pesaran MH, Potter S (1997) A floor and ceiling model of US output. *J Econ Dyn Control* 21:661–695
  72. Pfann GA, Schotman PC, Tschernig R (1996) Nonlinear interest rate dynamics and the implications for the term structure. *J Econ* 74:149–176
  73. Poter SM (1995) A nonlinear approach to US GNP. *J Appl Econ* 10:109–125
  74. Potter SM (1999) Nonlinear time series modelling: an introduction. *J Econ Surv* 13:505–528
  75. Quandt RE (1958) The estimation of the parameters of a linear regression system obeying two separate regimes. *J Am Stat Assoc* 53:873–880
  76. Radchenko S (2005) Lags in the response of gasoline prices to changes in crude oil prices: the role of short-term and long-term shocks. *Energy Econ* 27:573–602
  77. Romer CD, Romer DH (2002) A rehabilitation of monetary policy in the 1950s. *Am Econ Rev* 92:121–127
  78. Rothman P (1998) Forecasting asymmetric unemployment rates. *Rev Econ Stat* 80:164–168
  79. Rothman P, van Dijk D, Franses PH (2001) A multivariate STAR analysis of the relationship between money and output. *Macroecon Dyn* 5:506–532
  80. Sarantis N (1999) Modeling non-linearities in real effective exchange rates. *J Int Money Finance* 18:27–45
  81. Sims C, Zha T (2006) Were there switches in US monetary policy? *Am Econ Rev* 96:54–81
  82. Skalin J, Terasvirta T (1999) Another look at Swedish business cycles. *J Appl Econ* 14:359–378
  83. Smith DR (2002) Markov-switching and stochastic volatility diffusion models of short-term interest rates. *J Bus Econ Stat* 20:183–197
  84. Steel M (2008) Bayesian Time Series Analysis. In: Durlauf S, Blume L (eds) *The New Palgrave Dictionary of Economics*, 2nd ed. Palgrave Macmillan, London
  85. Taylor JB (1999) An historical analysis of monetary policy rules. In: Taylor JB (ed) *Monetary Policy Rules*, pp 319–341
  86. Taylor N, van Dijk D, Franses PH, Lucas A (2000) SETS, arbitrage activity, and stock price dynamics. *J Bank Finance* 24: 1289–1306
  87. Taylor MP, Peel DA, Sarno L (2001) Nonlinear mean-reversion in exchange rate rates: towards a solution to the purchasing power parity puzzles. *Int Econ Rev* 42:1015–1042
  88. Terasvirta T, Anderson H (1992) Characterising nonlinearities in business cycles using smooth transition autoregressive models. *J Appl Econ* 75:119–136
  89. Terasvirta T (1994) Specification, estimation and evaluation of smooth transition autoregressive models. *J Am Stat Assoc* 89:208–219
  90. Terasvirta T (1998) Modelling economic relationships with smooth transition regressions. In *Handbook of Appl Economic Statistics*. Marcel Dekker, New York, pp 507–552
  91. Tiao CG, Tsay RS (1994) Some advances in non linear and adaptive modelling in time series. *J Forecast* 13:109–131
  92. Tong H (1978) On a threshold model. In: Chan CH (ed) *Pattern Recognition and Signal Processing*. Sijthoff and Noordhoff, Amsterdam
  93. Tong H (1983) *Threshold Models in Non-Linear Time Series Analysis*. Lecture Notes in Statistics, no 21. Springer, Heidelberg
  94. Tong H, Lim KS (1980) Threshold autoregression, limit cycles and cyclical data. *J Royal Stat Soc B* 42:245–292
  95. Tong H (1990) *Nonlinear Time Series: A Dynamical System Approach*. Oxford University Press, Oxford
  96. Tsay RS (1998) Testing and modeling multivariate threshold models. *J Am Stat Assoc* 93:1188–1202
  97. van Dijk D, Franses PH (1999) Modeling multiple regimes in the business cycle. *Macroecon Dyn* 3:311–340
  98. van Dijk D, Terasvirta T, Franses PH (2002) Smooth transition autoregressive models – a survey of recent developments. *Econ Rev* 21:1–47
  99. Weise CL (1999) The Asymmetric effects of monetary policy. *J Money Credit Bank* 31:85–108
  100. Zellner A (1971) *An Introduction to Bayesian Inference in Econometrics*. Wiley, New York

## Bayesian Statistics

DAVID DRAPER

Department of Applied Mathematics and Statistics,  
Baskin School of Engineering, University of California,  
Santa Cruz, USA



## Article Outline

Glossary

Definition of the Subject and Introduction

The Bayesian Statistical Paradigm

Three Examples

Comparison with the Frequentist Statistical Paradigm

Future Directions

Bibliography

## Glossary

**Bayes' theorem; prior, likelihood and posterior distributions** Given (a)  $\theta$ , something of interest which is unknown to the person making an uncertainty assessment, conveniently referred to as You, (b)  $y$ , an information source which is relevant to decreasing Your uncertainty about  $\theta$ , (c) a desire to learn about  $\theta$  from  $y$  in a way that is both internally and externally logically consistent, and (d)  $\mathcal{B}$ , Your background assumptions and judgments about how the world works, as these assumptions and judgments relate to learning about  $\theta$  from  $y$ , it can be shown that You are compelled in this situation to reason within the standard rules of probability as the basis of Your *inferences* about  $\theta$ , *predictions* of future data  $y^*$ , and *decisions* in the face of uncertainty (see below for contrasts between inference, prediction and decision-making), and to quantify Your uncertainty about any unknown quantities through conditional probability distributions. When inferences about  $\theta$  are the goal, Bayes' Theorem provides a means of combining all relevant information internal and external to  $y$ :

$$p(\theta|y, \mathcal{B}) = c p(\theta|\mathcal{B}) l(\theta|y, \mathcal{B}). \quad (1)$$

Here, for example in the case in which  $\theta$  is a real-valued vector of length  $k$ , (a)  $p(\theta|\mathcal{B})$  is Your *prior distribution* about  $\theta$  given  $\mathcal{B}$  (in the form of a probability density function), which quantifies all relevant information available to You about  $\theta$  external to  $y$ , (b)  $c$  is a positive normalizing constant, chosen to make the density on the left side of the equation integrate to 1, (c)  $l(\theta|y, \mathcal{B})$  is Your *likelihood distribution* for  $\theta$  given  $y$  and  $\mathcal{B}$ , which is defined to be a density-normalized multiple of Your *sampling distribution*  $p(\cdot|\theta, \mathcal{B})$  for future data values  $y^*$  given  $\theta$  and  $\mathcal{B}$ , but re-interpreted as a function of  $\theta$  for fixed  $y$ , and (d)  $p(\theta|y, \mathcal{B})$  is Your *posterior distribution* about  $\theta$  given  $y$  and  $\mathcal{B}$ , which summarizes Your current total information about  $\theta$  and solves the basic inference problem.

## Bayesian parametric and non-parametric modeling (1)

Following de Finetti [23], a Bayesian statistical model is a joint predictive distribution  $p(y_1, \dots, y_n)$  for observable quantities  $y_i$  that have not yet been observed, and about which You are therefore uncertain. When the  $y_i$  are real-valued, often You will not regard them as probabilistically *independent* (informally, the  $y_i$  are independent if information about any of them does not help You to predict the others); but it may be possible to identify a *parameter vector*  $\theta = (\theta_1, \dots, \theta_k)$  such that You would judge the  $y_i$  *conditionally independent* given  $\theta$ , and would therefore be willing to model them via the relation

$$p(y_1, \dots, y_n|\theta) = \prod_{i=1}^n p(y_i|\theta). \quad (2)$$

When combined with a prior distribution  $p(\theta)$  on  $\theta$  that is appropriate to the context, this is *Bayesian parametric modeling*, in which  $p(y_i|\theta)$  will often have a standard distributional form (such as binomial, Poisson or Gaussian). (2) When a (finite) parameter vector that induces conditional independence cannot be found, if You judge your uncertainty about the real-valued  $y_i$  *exchangeable* (see below), then a representation theorem of de Finetti [21] states informally that all internally logically consistent predictive distributions  $p(y_1, \dots, y_n)$  can be expressed in a way that is equivalent to the *hierarchical model* (see below)

$$\begin{aligned} (F|\mathcal{B}) &\sim p(F|\mathcal{B}) \\ (y_i|F, \mathcal{B}) &\stackrel{\text{iid}}{\sim} F, \end{aligned} \quad (3)$$

where (a)  $F$  is the cumulative distribution function (CDF) of the underlying process  $(y_1, y_2, \dots)$  from which You are willing to regard  $p(y_1, \dots, y_n)$  as (in effect) like a random sample and (b)  $p(F|\mathcal{B})$  is Your prior distribution on the space  $\mathcal{F}$  of all CDFs on the real line. This (placing probability distributions on infinite-dimensional spaces such as  $\mathcal{F}$ ) is *Bayesian non-parametric modeling*, in which priors involving *Dirichlet processes* and/or *Pólya trees* (see Sect. “[Inference: Parametric and Non-Parametric Modeling of Count Data](#)”) are often used.

**Exchangeability** A sequence  $y = (y_1, \dots, y_n)$  of random variables (for  $n \geq 1$ ) is (*finitely*) *exchangeable* if the joint probability distribution  $p(y_1, \dots, y_n)$  of the elements of  $y$  is invariant under permutation of the indices  $(1, \dots, n)$ , and a countably infinite sequence  $(y_1, y_2, \dots)$  is (*infinitely*) *exchangeable* if every finite subsequence is finitely exchangeable.

**Hierarchical modeling** Often Your uncertainty about something unknown to You can be seen to have a *nested* or *hierarchical* character. One class of examples arises in *cluster sampling* in fields such as education and medicine, in which students (level 1) are nested within classrooms (level 2) and patients (level 1) within hospitals (level 2); cluster sampling involves random samples (and therefore uncertainty) at two or more levels in such a data hierarchy (examples of this type of hierarchical modeling are given in Sect. “[Strengths and Weaknesses of the Two Approaches](#)”). Another, quite different, class of examples of Bayesian hierarchical modeling is exemplified by equation (3) above, in which it was helpful to decompose Your overall predictive uncertainty about  $(y_1, \dots, y_n)$  into (a) uncertainty about  $F$  and then (b) uncertainty about the  $y_i$  given  $F$  (examples of this type of hierarchical modeling appear in Sect. “[Inference and Prediction: Binary Outcomes with No Covariates](#)” and “[Inference: Parametric and Non-Parametric Modeling of Count Data](#)”).

**Inference, prediction and decision-making; samples and populations** Given a data source  $y$ , *inference* involves drawing probabilistic conclusions about the underlying process that gave rise to  $y$ , *prediction* involves summarizing uncertainty about future observable data values  $y^*$ , and *decision-making* involves looking for optimal behavioral choices in the face of uncertainty (about either the underlying process, or the future, or both). In some cases inference takes the form of reasoning backwards from a *sample* of data values to a *population*: a (larger) universe of possible data values from which You judge that the sample has been drawn in a manner that is *representative* (i. e., so that the sampled and unsampled values in the population are (likely to be) similar in relevant ways).

**Mixture modeling** Given  $y$ , unknown to You, and  $\mathcal{B}$ , Your background assumptions and judgments relevant to  $y$ , You have a choice: You can either model (Your uncertainty about)  $y$  directly, through the probability distribution  $p(y|\mathcal{B})$ , or (if that is not feasible) You can identify a quantity  $x$  upon which You judge  $y$  to depend and model  $y$  hierarchically, in two stages: first by modeling  $x$ , through the probability distribution  $p(x|\mathcal{B})$ , and then by modeling  $y$  given  $x$ , through the probability distribution  $p(y|x, \mathcal{B})$ :

$$p(y|\mathcal{B}) = \int_{\mathcal{X}} p(y|x, \mathcal{B}) p(x|\mathcal{B}) dx, \quad (4)$$

where  $\mathcal{X}$  is the space of possible values of  $x$  over which Your uncertainty is expressed. This is *mixture model-*

*ing*, a special case of hierarchical modeling (see above). In hierarchical notation (4) can be re-expressed as

$$y = \left\{ \begin{array}{c} x \\ (y|x) \end{array} \right\}. \quad (5)$$

Examples of mixture modeling in this article include (a) equation (3) above, with  $F$  playing the role of  $x$ ; (b) the basic equation governing Bayesian prediction, discussed in Sect. “[The Bayesian Statistical Paradigm](#)”; (c) Bayesian model averaging (Sect. “[The Bayesian Statistical Paradigm](#)”); (d) de Finetti’s representation theorem for binary outcomes (Sect. “[Inference and Prediction: Binary Outcomes with No Covariates](#)”); (e) random-effects parametric and non-parametric modeling of count data (Sect. “[Inference: Parametric and Non-Parametric Modeling of Count Data](#)”); and (f) integrated likelihoods in Bayes factors (Sect. “[Decision-Making: Variable Selection in Generalized Linear Models; Bayesian Model Selection](#)”).

**Probability – frequentist and Bayesian** In the *frequentist* probability paradigm, attention is restricted to phenomena that are inherently repeatable under (essentially) identical conditions; then, for an event  $A$  of interest,  $P_f(A)$  is the limiting relative frequency with which  $A$  occurs in the (hypothetical) repetitions, as the number of repetitions  $n \rightarrow \infty$ . By contrast, Your Bayesian probability  $P_B(A|\mathcal{B})$  is the numerical weight of evidence, given Your background information  $\mathcal{B}$  relevant to  $A$ , in favor of a true-false proposition  $A$  whose truth status is uncertain to You, obeying a series of reasonable axioms to ensure that Your Bayesian probabilities are internally logically consistent.

**Utility** To ensure internal logical consistency, optimal decision-making proceeds by (a) specifying a *utility* function  $U(a, \theta_0)$  quantifying the numerical value associated with taking action  $a$  if the unknown is really  $\theta_0$  and (b) *maximizing expected utility*, where the expectation is taken over uncertainty in  $\theta$  as quantified by the posterior distribution  $p(\theta|y, \mathcal{B})$ .

## Definition of the Subject and Introduction

*Statistics* may be defined as the study of uncertainty: how to measure it, and how to make choices in the face of it. Uncertainty is quantified via *probability*, of which there are two leading paradigms, *frequentist* (discussed in Sect. “[Comparison with the Frequentist Statistical Paradigm](#)”) and *Bayesian*. In the Bayesian approach to probability the primitive constructs are true-false *propositions*  $A$  whose truth status is uncertain, and the probability of  $A$  is the numerical weight of evidence in favor of  $A$ ,

constrained to obey a set of axioms to ensure that Bayesian probabilities are *coherent* (internally logically consistent).

The discipline of statistics may be divided broadly into four activities: *description* (graphical and numerical summaries of a data set  $y$ , without attempting to reason outward from it; this activity is almost entirely non-probabilistic and will not be discussed further here), *inference* (drawing probabilistic conclusions about the underlying process that gave rise to  $y$ ), *prediction* (summarizing uncertainty about future observable data values  $y^*$ ), and *decision-making* (looking for optimal behavioral choices in the face of uncertainty). Bayesian statistics is an approach to inference, prediction and decision-making that is based on the Bayesian probability paradigm, in which uncertainty about an unknown  $\theta$  (this is the inference problem) is quantified by means of a conditional probability distribution  $p(\theta|y, \mathcal{B})$ ; here  $y$  is all available relevant data and  $\mathcal{B}$  summarizes the background assumptions and judgments of the person making the uncertainty assessment. Prediction of a future  $y^*$  is similarly based on the conditional probability distribution  $p(y^*|y, \mathcal{B})$ , and optimal decision-making proceeds by (a) specifying a *utility* function  $U(a, \theta_0)$  quantifying the numerical reward associated with taking action  $a$  if the unknown is really  $\theta_0$  and (b) *maximizing expected utility*, where the expectation is taken over uncertainty in  $\theta$  as quantified by  $p(\theta|y, \mathcal{B})$ .

### The Bayesian Statistical Paradigm

*Statistics* is the branch of mathematical and scientific inquiry devoted to the study of uncertainty: its consequences, and how to behave sensibly in its presence. The subject draws heavily on *probability*, a discipline which predates it by about 100 years: basic probability theory can be traced [48] to work of Pascal, Fermat and Huygens in the 1650s, and the beginnings of statistics [34,109] are evident in work of Bayes published in the 1760s.

The Bayesian statistical paradigm consists of three basic ingredients:

- $\theta$ , something of interest which is unknown (or only partially known) to the person making the uncertainty assessment, conveniently referred to, in a convention proposed by Good (1950), as *You*. Often  $\theta$  is a *parameter* vector of real numbers (of finite length  $k$ , say) or a matrix, but it can literally be almost anything: for example, a function (three leading examples are a cumulative distribution function (CDF), a density, or a regression surface), a phylogenetic tree, an image of a region on the surface of Mars at a particular moment in time, . . .
- $y$ , an information source which is relevant to decreasing Your uncertainty about  $\theta$ . Often  $y$  is a vector of real numbers (of finite length  $n$ , say), but it can also literally be almost anything: for instance, a time series, a movie, the text in a book, . . .
- A desire to learn about  $\theta$  from  $y$  in a way that is both *coherent* (internally consistent: in other words, free of internal logical contradictions; Bernardo and Smith [11] give a precise definition of coherence) and *well-calibrated* (externally consistent: for example, capable of making accurate predictions of future data  $y^*$ ).

It turns out [23,53] that You are compelled in this situation to reason within the standard rules of probability (see below) as the basis of Your inferences about  $\theta$ , predictions of future data  $y^*$ , and decisions in the face of uncertainty, and to quantify Your uncertainty about any unknown quantities through conditional probability distributions, as in the following three basic equations of Bayesian statistics:

$$\begin{aligned} p(\theta|y, \mathcal{B}) &= c p(\theta|\mathcal{B}) l(\theta|y, \mathcal{B}) \\ p(y^*|y, \mathcal{B}) &= \int_{\Theta} p(y^*|\theta, \mathcal{B}) p(\theta|y, \mathcal{B}) d\theta \\ a^* &= \operatorname{argmax}_{a \in \mathcal{A}} E_{(\theta|y, \mathcal{B})} [U(a, \theta)] . \end{aligned} \quad (6)$$

(The basic rules of probability [71] are: for any true-false propositions  $A$  and  $B$  and any background assumptions and judgments  $\mathcal{B}$ , (*convexity*)  $0 \leq P(A|\mathcal{B}) \leq 1$ , with equality at 1 iff  $A$  is known to be true under  $\mathcal{B}$ ; (*multiplication*)  $P(A \text{ and } B|\mathcal{B}) = P(A|\mathcal{B}) P(B|A, \mathcal{B}) = P(B|\mathcal{B}) P(A|B, \mathcal{B})$ ; and (*addition*)  $P(A \text{ or } B|\mathcal{B}) = P(A|\mathcal{B}) + P(B|\mathcal{B}) - P(A \text{ and } B|\mathcal{B})$ .)

The meaning of the equations in (6) is as follows.

- $\mathcal{B}$  stands for Your background (often not fully stated) assumptions and judgments about how the world works, as these assumptions and judgments relate to learning about  $\theta$  from  $y$ .  $\mathcal{B}$  is often omitted from the basic equations (sometimes with unfortunate consequences), yielding the simpler-looking forms

$$\begin{aligned} p(\theta|y) &= c p(\theta) l(\theta|y) \\ p(y^*|y) &= \int_{\Theta} p(y^*|\theta) p(\theta|y) d\theta \\ a^* &= \operatorname{argmax}_{a \in \mathcal{A}} E_{(\theta|y)} [U(a, \theta)] . \end{aligned} \quad (7)$$

- $p(\theta|\mathcal{B})$  is Your *prior information* about  $\theta$  given  $\mathcal{B}$ , in the form of a probability density function (PDF) or probability mass function (PMF) if  $\theta$  lives continuously or discretely on  $\mathbb{R}^k$  (this is generically referred to as

Your *prior distribution*), and  $p(\theta|y, \mathcal{B})$  is Your *posterior distribution* about  $\theta$  given  $y$  and  $\mathcal{B}$ , which summarizes Your current total information about  $\theta$  and solves the basic inference problem. These are actually not very good names for  $p(\theta|\mathcal{B})$  and  $p(\theta|y, \mathcal{B})$ , because (for example)  $p(\theta|\mathcal{B})$  really stands for all (relevant) information about  $\theta$  (given  $\mathcal{B}$ ) external to  $y$ , whether that information was obtained before (or after)  $y$  arrives, but (a) they do emphasize the sequential nature of learning and (b) through long usage it would be difficult for more accurate names to be adopted.

- $c$  (here and throughout) is a generic positive normalizing constant, inserted into the first equation in (6) to make the left-hand side integrate (or sum) to 1 (as any coherent distribution must).
- $p(y^*|\theta, \mathcal{B})$  is Your *sampling distribution* for future data values  $y^*$  given  $\theta$  and  $\mathcal{B}$  (and presumably You would use the same sampling distribution  $p(y|\theta, \mathcal{B})$  for (past) data values  $y$ , mentally turning the clock back to a point before the data arrives and thinking about what values of  $y$  You might see). This assumes that You are willing to regard Your data as like random draws from a *population* of possible data values (an heroic assumption in some cases, for instance with observational rather than randomized data; this same assumption arises in the frequentist statistical paradigm, discussed below in Sect. “[Comparison with the Frequentist Statistical Paradigm](#)”).
- $l(\theta|y, \mathcal{B})$  is Your *likelihood function* for  $\theta$  given  $y$  and  $\mathcal{B}$ , which is defined to be any positive constant multiple of the sampling distribution  $p(y|\theta, \mathcal{B})$  but re-interpreted as a function of  $\theta$  for fixed  $y$ :

$$l(\theta|y, \mathcal{B}) = c p(y|\theta, \mathcal{B}). \quad (8)$$

The likelihood function is also central to one of the main approaches to frequentist statistical inference, developed by Fisher [37]; the two approaches are contrasted in Sect. “[Comparison with the Frequentist Statistical Paradigm](#)”.

All of the symbols in the first equation in (6) have now been defined, and this equation can be recognized as *Bayes’ Theorem*, named after Bayes [5] because a special case of it appears prominently in work of his that was published posthumously. It describes how to pass coherently from information about  $\theta$  external to  $y$  (quantified in the prior distribution  $p(\theta|\mathcal{B})$ ) to information both internal and external to  $y$  (quantified in the posterior distribution  $p(\theta|y, \mathcal{B})$ ), via the likelihood function  $l(\theta|y, \mathcal{B})$ : You multiply the prior and likelihood pointwise in  $\theta$  and normalize so that the posterior distribution  $p(\theta|y, \mathcal{B})$  integrates (or sums) to 1.

- According to the second equation in (6),  $p(y^*|y, \mathcal{B})$ , Your (posterior) *predictive distribution* for future data  $y^*$  given (past) data  $y$  and  $\mathcal{B}$ , which solves the basic prediction problem, must be a weighted average of Your sampling distribution  $p(y^*|\theta, \mathcal{B})$  weighted by Your current best information  $p(\theta|y, \mathcal{B})$  about  $\theta$  given  $y$  and  $\mathcal{B}$ ; in this integral  $\Theta$  is the space of possible values of  $\theta$  over which Your uncertainty is expressed. (The second equation in (6) contains a simplifying assumption that should be mentioned: in full generality the first term  $p(y^*|\theta, \mathcal{B})$  inside the integral would be  $p(y^*|y, \theta, \mathcal{B})$ , but it is almost always the case that the information in  $y$  is redundant in the presence of complete knowledge of  $\theta$ , in which case  $p(y^*|y, \theta, \mathcal{B}) = p(y^*|\theta, \mathcal{B})$ ; this state of affairs could be described by saying that *the past and future are conditionally independent given the truth*. A simple example of this phenomenon is provided by coin-tossing: if You are watching a Bernoulli( $\theta$ ) process unfold (see Sect. “[Inference and Prediction: Binary Outcomes with No Covariates](#)”) whose success probability  $\theta$  is unknown to You, the information that 8 of the first 10 tosses have been heads is definitely useful to You in predicting the 11th toss, but if instead You somehow knew that  $\theta$  was 0.7, the outcome of the first 10 tosses would be irrelevant to You in predicting any future tosses.)
- Finally, in the context of making a choice in the face of uncertainty,  $\mathcal{A}$  is Your set of possible actions,  $U(a, \theta_0)$  is the numerical value (*utility*) You attach to taking action  $a$  if the unknown is really  $\theta_0$  (specified, without loss of generality, so that large utility values are preferred by You), and the third equation in (6) says that to make the choice coherently You should find the action  $a^*$  that *maximizes expected utility* (MEU); here the expectation

$$E_{(\theta|y, \mathcal{B})} [U(a, \theta)] = \int_{\Theta} U(a, \theta) p(\theta|y, \mathcal{B}) d\theta \quad (9)$$

is taken over uncertainty in  $\theta$  as quantified by the posterior distribution  $p(\theta|y, \mathcal{B})$ .

This summarizes the entire Bayesian statistical paradigm, which is driven by the three equations in (6). Examples of its use include clinical trial design [56] and analysis [105]; spatio-temporal modeling, with environmental applications [101]; forecasting and dynamic linear models [115]; non-parametric estimation of receiver operating characteristic curves, with applications in medicine and agriculture [49]; finite selection models, with health policy applications [79]; Bayesian CART model search, with applications in breast cancer research [16]; construc-



tion of radiocarbon calibration curves, with archaeological applications [15]; factor regression models, with applications to gene expression data [114]; mixture modeling for high-density genotyping arrays, with bioinformatic applications [100]; the EM algorithm for Bayesian fitting of latent process models [76]; state-space modeling, with applications in particle-filtering [92]; causal inference [42,99]; hierarchical modeling of DNA sequences, with genetic and medical applications [77]; hierarchical Poisson regression modeling, with applications in health care evaluation [17]; multiscale modeling, with engineering and financial applications [33]; expected posterior prior distributions for model selection [91]; nested Dirichlet processes, with applications in the health sciences [96]; Bayesian methods in the study of sustainable fisheries [74,82]; hierarchical non-parametric meta-analysis, with medical and educational applications [81]; and structural equation modeling of multilevel data, with applications to health policy [19].

Challenges to the paradigm include the following:

- **Q:** How do You specify the sampling distribution/likelihood function that quantifies the information about the unknown  $\theta$  internal to Your data set  $y$ ?  
**A:** (1) The solution to this problem, which is common to all approaches to statistical inference, involves imagining future data  $y^*$  from the same process that has yielded or will yield Your data set  $y$ ; often the variability You expect in future data values can be quantified (at least approximately) through a standard parametric family of distributions (such as the Bernoulli/binomial for binary data, the Poisson for count data, and the Gaussian for real-valued outcomes) and the parameter vector of this family becomes the unknown  $\theta$  of interest. (2) Uncertainty in the likelihood function is referred to as *model uncertainty* [67]; a leading approach to quantifying this source of uncertainty is *Bayesian model averaging* [18,25,52], in which uncertainty about the models  $M$  in an ensemble  $\mathcal{M}$  of models (specifying  $\mathcal{M}$  is part of  $\mathcal{B}$ ) is assessed and propagated for a quantity, such as a future data value  $y^*$ , that is common to all models via the expression

$$p(y^*|y, \mathcal{B}) = \int_{\mathcal{M}} p(y^*|y, M, \mathcal{B}) p(M|y, \mathcal{B}) dM. \quad (10)$$

In other words, to make coherent predictions in the presence of model uncertainty You should form a weighted average of the conditional predictive distributions  $p(y^*|y, M, \mathcal{B})$ , weighted by the posterior model probabilities  $p(M|y, \mathcal{B})$ . Other potentially useful approaches to model uncertainty include

*Bayesian non-parametric modeling*, which is examined in Sect. “[Inference: Parametric and Non-Parametric Modeling of Count Data](#)”, and methods based on *cross-validation* [110], in which (in Bayesian language) part of the data is used to specify the prior distribution on  $\mathcal{M}$  (which is an input to calculating the posterior model probabilities) and the rest of the data is employed to update that prior.

- **Q:** How do You quantify information about the unknown  $\theta$  external to Your data set  $y$  in the prior probability distribution  $p(\theta|\mathcal{B})$ ?  
**A:** (1) There is an extensive literature on *elicitation* of prior (and other) probabilities; notable references include O’Hagan et al. [85] and the citations given there. (2) If  $\theta$  is a parameter vector and the likelihood function is a member of the *exponential family* [11], the prior distribution can be chosen in such a way that the prior and posterior distributions for  $\theta$  have the same mathematical form (such a prior is said to be *conjugate* to the given likelihood); this may greatly simplify the computations, and often prior information can (at least approximately) be quantified by choosing a member of the conjugate family (see Sect. “[Inference and Prediction: Binary Outcomes with No Covariates](#)” for an example of both of these phenomena).

In situations where it is not precisely clear how to quantify the available information external to  $y$ , two sets of tools are available:

- *Sensitivity analysis* [30], also known as *pre-posterior analysis* [4]: Before the data have begun to arrive, You can (a) generate data similar to what You expect You will see, (b) choose a plausible prior specification and update it to the posterior on the quantities of greatest interest, (c) repeat (b) across a variety of plausible alternatives, and (d) see if there is substantial stability in conclusions across the variations in prior specification. If so, fine; if not, this approach can be combined with hierarchical modeling [68]: You can collect all of the plausible priors and add a layer hierarchically to the prior specification, with the new layer indexing variation across the prior alternatives.
- *Bayesian robustness* [8,95]: If, for example, the context of the problem implies that You only wish to specify that the prior distribution belongs to an infinite-dimensional class (such as, for priors on  $(0, 1)$ , the class of monotone non-increasing functions) with (for instance) bounds on the first two moments, You can in turn quantify bounds on summaries of the resulting posterior distribution, which may be narrow enough to demonstrate that Your



uncertainty in specifying the prior does not lead to differences that are large in practical terms.

Often context suggests specification of a prior that has relatively little information content in relation to the likelihood information; for reasons that are made clear in Sect. “[Inference and Prediction: Binary Outcomes with No Covariates](#)”, such priors are referred to as relatively *diffuse* or *flat* (the term *non-informative* is sometimes also used, but this seems worth avoiding, because any prior specification takes a particular position regarding the amount of relevant information external to the data). See Bernardo [10] and Kass and Wasserman [60] for a variety of formal methods for generating diffuse prior distributions.

- **Q:** How do You quantify Your utility function  $U(a, \theta)$  for optimal decision-making? **A:** There is a rather less extensive statistical literature on elicitation of utility than probability; notable references include Fishburn [35,36], Schervish et al. [103], and the citations in Bernardo and Smith [11]. There is a parallel (and somewhat richer) economics literature on utility elicitation; see, for instance, Abdellaoui [1] and Blavatsky [12]. Sect. “[Decision-Making: Variable Selection in Generalized Linear Models; Bayesian Model Selection](#)” provides a decision-theoretic example.
- Suppose that  $\theta = (\theta_1, \dots, \theta_k)$  is a parameter vector of length  $k$ . Then (a) computing the normalizing constant in Bayes’ Theorem

$$c = \left( \int \cdots \int p(y|\theta_1, \dots, \theta_k, \mathcal{B}) \cdot p(\theta_1, \dots, \theta_k|\mathcal{B}) d\theta_1 \cdots d\theta_k \right)^{-1} \quad (11)$$

involves evaluating a  $k$ -dimensional integral; (b) the predictive distribution in the second equation in (6) involves another  $k$ -dimensional integral; and (c) the posterior  $p(\theta_1, \dots, \theta_k|y, \mathcal{B})$  is a  $k$ -dimensional probability distribution, which for  $k > 2$  can be difficult to visualize, so that attention often focuses on the *marginal* posterior distributions

$$p(\theta_j|y, \mathcal{B}) = \int \cdots \int p(\theta_1, \dots, \theta_k|y, \mathcal{B}) d\theta_{-j} \quad (12)$$

for  $j = 1, \dots, k$ , where  $\theta_{-j}$  is the  $\theta$  vector with component  $j$  omitted; each of these marginal distributions involves a  $(k - 1)$ -dimensional integral. If  $k$  is large these integrals can be difficult or impossible to evaluate exactly, and a general method for computing accurate approximations to them proved elusive from the time of Bayes in the eighteenth century until recently (in the

late eighteenth century Laplace [63]) developed an analytical method, which today bears his name, for approximating integrals that arise in Bayesian work [11], but his method is not as general as the computationally-intensive techniques in widespread current use). Around 1990 there was a fundamental shift in Bayesian computation, with the belated discovery by the statistics profession of a class of techniques – *Markov chain Monte Carlo* (MCMC) methods [41,44] – for approximating high-dimensional Bayesian integrals in a computationally-intensive manner, which had been published in the chemical physics literature in the 1950s [78]; these methods came into focus for the Bayesian community at a moment when desktop computers had finally become fast enough to make use of such techniques.

MCMC methods approximate integrals associated with the posterior distribution  $p(\theta|y, \mathcal{B})$  by (a) creating a Markov chain whose equilibrium distribution is the desired posterior and (b) sampling from this chain from an initial  $\theta_{(0)}$  (i) until equilibrium has been reached (all draws up to this point are typically discarded) and (ii) for a sufficiently long period thereafter to achieve the desired approximation accuracy. With the advent and refinement of MCMC methods since 1990, the Bayesian integration problem has been solved for a wide variety of models, with more ambitious sampling schemes made possible year after year with increased computing speeds: for instance, in problems in which the dimension of the parameter space is not fixed in advance (an example is *regression change-point* problems [104], where the outcome  $y$  is assumed to depend linearly (apart from stochastic noise) on the predictor(s)  $x$  but with an unknown number of changes of slope and intercept and unknown locations for those changes), ordinary MCMC techniques will not work; in such problems methods such as *reversible-jump MCMC* [47,94] and *Markov birth-death processes* [108], which create Markov chains that permit trans-dimensional jumps, are required.

The main drawback of MCMC methods is that they do not necessarily scale well as  $n$  (the number of data observations) increases; one alternative, popular in the machine learning community, is *variational* methods [55], which convert the integration problem into an optimization problem by (a) approximating the posterior distribution of interest by a family of distributions yielding a closed-form approximation to the integral and (b) finding the member of the family that maximizes the accuracy of the approximation.

- Bayesian decision theory [6], based on maximizing expected utility, is unambiguous in its normative rec-

ommendation for how a single agent (You) should make a choice in the face of uncertainty, and it has had widespread success in fields such as economics (e.g., [2,50]) and medicine (e.g., [87,116]). It is well known, however [3,112], that Bayesian decision theory (or indeed any other formal approach that seeks an optimal behavioral choice) can be problematic when used normatively for group decision-making, because of conflicts in preferences among members of the group. This is an important unsolved problem.

### Three Examples

#### Inference and Prediction:

##### Binary Outcomes with No Covariates

Consider the problem of measuring the quality of care at a particular hospital  $H$ . One way to do this is to examine the outcomes of that care, such as mortality, after adjusting for the burden of illness brought by the patients to  $H$  on admission. As an even simpler version of this problem, consider just the  $n$  binary mortality observables  $y = (y_1, \dots, y_n)$  (with mortality measured within 30 days of admission, say; 1 = died, 0 = lived) that You will see from all of the patients at  $H$  with a particular admission diagnosis (heart attack, say) during some pre-specified future time window. You acknowledge Your uncertainty about which elements in the sequence will be 0s and which 1s, and You wish to quantify this uncertainty using the Bayesian paradigm. As de Finetti [20] noted, in this situation Your fundamental imperative is to construct a predictive distribution  $p(y_1, \dots, y_n | \mathcal{B})$  that expresses Your uncertainty about the future observables, rather than – as is perhaps more common – to reach immediately for a standard family of parametric models for the  $y_i$  (in other words, to posit the existence of a vector  $\theta = (\theta_1, \dots, \theta_k)$  of parameters and to model the observables by appeal to a family  $p(y_i | \theta, \mathcal{B})$  of probability distributions indexed by  $\theta$ ).

Even though the  $y_i$  are binary, with all but the smallest values of  $n$  it still seems a formidable task to elicit from Yourself an  $n$ -dimensional predictive distribution  $p(y_1, \dots, y_n | \mathcal{B})$ . De Finetti [20] showed, however, that the task is easier than it seems. In the absence of any further information about the patients, You notice that Your uncertainty about them is *exchangeable*: if someone (without telling You) were to rearrange the order in which their mortality outcomes become known to You, Your predictive distribution would not change. This still seems to leave  $p(y_1, \dots, y_n | \mathcal{B})$  substantially unspecified (where  $\mathcal{B}$  now includes the judgment of exchangeability of the  $y_i$ ), but

de Finetti [20] proved a remarkable theorem which shows (in effect) that all exchangeable predictive distributions for a vector of binary observables are representable as *mixtures* of Bernoulli sampling distributions: if You're willing to regard  $(y_1, \dots, y_n)$  as the first  $n$  terms in an *infinitely exchangeable* binary sequence  $(y_1, y_2, \dots)$  (which just means that every finite subsequence is exchangeable), then to achieve coherence Your predictive distribution must be expressible as

$$p(y_1, \dots, y_n | \mathcal{B}) = \int_0^1 \theta^{s_n} (1 - \theta)^{n-s_n} p(\theta | \mathcal{B}) d\theta, \quad (13)$$

where  $s_n = \sum_{i=1}^n y_i$ . Here the quantity  $\theta$  on the right side of (13) is more than just an integration variable: the equation says that in Your predictive modeling of the binary  $y_i$  You may as well proceed as if

- There is a quantity called  $\theta$ , interpretable both as the marginal death probability  $p(y_i = 1 | \theta, \mathcal{B})$  for each patient and as the long-run mortality rate in the infinite sequence  $(y_1, y_2, \dots)$  (which serves, in effect, as a population of values to which conclusions from the data can be generalized);
- Conditional on  $\theta$  and  $\mathcal{B}$ , the  $y_i$  are independent identically distributed (IID) Bernoulli ( $\theta$ ); and
- $\theta$  can be viewed as a realization of a random variable with density  $p(\theta | \mathcal{B})$ .

In other words, exchangeability of Your uncertainty about a binary process is functionally equivalent to assuming the simple Bayesian *hierarchical* model [27]

$$\begin{aligned} (\theta | \mathcal{B}) &\sim p(\theta | \mathcal{B}) \\ (y_i | \theta, \mathcal{B}) &\stackrel{\text{iid}}{\sim} \text{Bernoulli}(\theta), \end{aligned} \quad (14)$$

and  $p(\theta | \mathcal{B})$  is recognizable as Your prior distribution for  $\theta$ , the underlying death rate for heart attack patients similar to those You expect will arrive at hospital  $H$  during the relevant time window.

Consider now the problem of quantitatively specifying prior information about  $\theta$ . From (13) and (14) the likelihood function is

$$l(\theta | y, \mathcal{B}) = c \theta^{s_n} (1 - \theta)^{n-s_n}, \quad (15)$$

which (when interpreted in the Bayesian manner as a density in  $\theta$ ) is recognizable as a member of the Beta family of probability distributions: for  $\alpha, \beta > 0$  and  $0 < \theta < 1$ ,

$$\theta \sim \text{Beta}(\alpha, \beta) \text{ iff } p(\theta) = c \theta^{\alpha-1} (1 - \theta)^{\beta-1}. \quad (16)$$

Moreover, this family has the property that the product of two Beta densities is another Beta density, so by Bayes'

Theorem if the prior  $p(\theta|\mathcal{B})$  is chosen to be  $\text{Beta}(\alpha, \beta)$  for some (as-yet unspecified)  $\alpha > 0$  and  $\beta > 0$ , then the posterior will be  $\text{Beta}(\alpha + s_n, \beta + n - s_n)$ : this is conjugacy (Sect. “The Bayesian Statistical Paradigm”) of the Beta family for the Bernoulli/binomial likelihood. In this case the conjugacy leads to a simple interpretation of  $\alpha$  and  $\beta$ : the prior acts like a data set with  $\alpha$  1s and  $\beta$  0s, in the sense that if person 1 does a Bayesian analysis with a  $\text{Beta}(\alpha, \beta)$  prior and sample data  $y = (y_1, \dots, y_n)$  and person 2 instead merges the corresponding “prior data set” with  $y$  and does a maximum-likelihood analysis (Sect. “Comparison with the Frequentist Statistical Paradigm”) on the resulting merged data, the two people will get the same answers. This also shows that the *prior sample size*  $n_0$  in the Beta-Bernoulli/binomial model is  $(\alpha + \beta)$ . Given that the mean of a  $\text{Beta}(\alpha, \beta)$  distribution is  $\alpha / (\alpha + \beta)$ , calculation reveals that the posterior mean  $(\alpha + s_n) / (\alpha + \beta + n)$  of  $\theta$  is a weighted average of the prior mean and the data mean  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ , with prior and data weights  $n_0$  and  $n$ , respectively:

$$\frac{\alpha + s_n}{\alpha + \beta + n} = \frac{n_0 \left( \frac{\alpha}{\alpha + \beta} \right) + n\bar{y}}{n_0 + n}. \quad (17)$$

These facts shed intuitive light on how Bayes’ Theorem combines information internal and external to a given data source: thinking of prior information as equivalent to a data set is a valuable intuition, even in non-conjugate settings.

The choice of  $\alpha$  and  $\beta$  naturally depends on the available information external to  $y$ . Consider for illustration two such specifications:

- Analyst 1 does a web search and finds that the 30-day mortality rate for heart attack (given average quality of care and average patient sickness at admission) in her country is 15%. The information she has about hospital  $H$  is that its care and patient sickness are not likely to be wildly different from the country averages but that a mortality deviation from the mean, if present, would be more likely to occur on the high side than the low. Having lived in the community served by  $H$  for some time and having not heard anything either outstanding or deplorable about the hospital, she would be surprised to find that the underlying heart attack death rate at  $H$  was less than (say) 5% or greater than (say) 30%. One way to quantify this information is to set the prior mean to 15% and to place (say) 95% of the prior mass between 5% and 30%.
- Analyst 2 has little information external to  $y$  and thus wishes to specify a relatively diffuse prior distribution

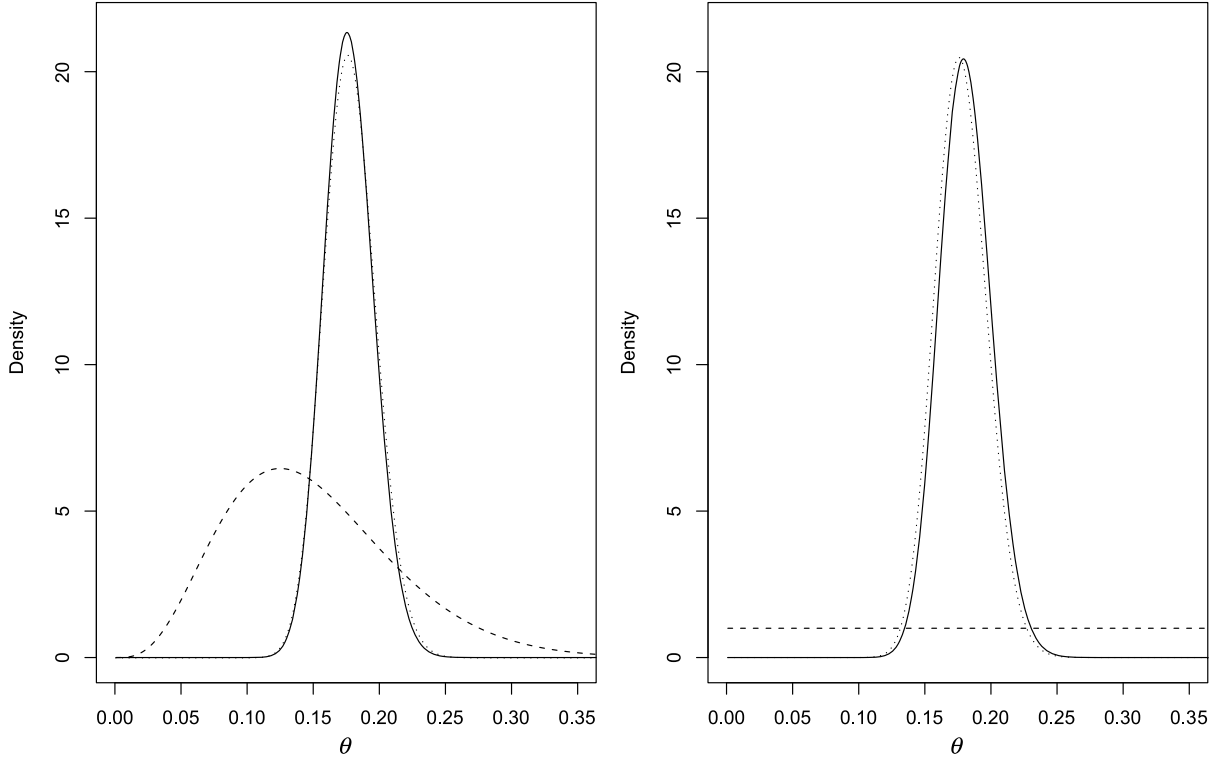
that does not dramatically favor any part of the unit interval.

Numerical integration reveals that  $(\alpha, \beta) = (4.5, 25.5)$ , with a prior sample size of 30.0, satisfies Analyst 1’s constraints. Analyst 2’s diffuse prior evidently corresponds to a rather small prior sample size; a variety of positive values of  $\alpha$  and  $\beta$  near 0 are possible, all of which will lead to a relatively flat prior.

Suppose for illustration that the time period in question is about four years in length and  $H$  is a medium-size US hospital; then there will be about  $n = 385$  heart attack patients in the data set  $y$ . Suppose further that the observed mortality rate at  $H$  comes out  $\bar{y} = s_n / n = 69 / 385 \doteq 18\%$ . Figure 1 summarizes the prior-to-posterior updating with this data set and the two priors for Analysts 1 (left panel) and 2 (right panel), with  $\alpha = \beta = 1$  (the Uniform distribution) for Analyst 2. Even though the two priors are rather different – Analyst 1’s prior is skewed, with a prior mean of 0.15 and  $n_0 \doteq 30$ ; Analyst 2’s prior is flat, with a prior mean of 0.5 and  $n_0 = 2$  – it is evident that the posterior distributions are nearly the same in both cases; this is because the data sample size  $n = 385$  is so much larger than either of the prior sample sizes, so that the likelihood information dominates. With both priors the likelihood and posterior distributions are nearly the same, another consequence of  $n_0 \ll n$ . For Analyst 1 the posterior mean, standard deviation, and 95% central posterior interval for  $\theta$  are (0.177, 0.00241, 0.142, 0.215), and the corresponding numerical results for Analyst 2 are (0.181, 0.00258, 0.144, 0.221); again it is clear that the two sets of results are almost identical. With a large sample size, careful elicitation – like that undertaken by Analyst 1 – will often yield results similar to those with a diffuse prior.

The posterior predictive distribution  $p(y_{n+1}|y_1, \dots, y_n, \mathcal{B})$  for the next observation, having observed the first  $n$ , is also straightforward to calculate in closed form with the conjugate prior in this model. It is clear that  $p(y_{n+1}|y, \mathcal{B})$  has to be a Bernoulli( $\theta^*$ ) distribution for some  $\theta^*$ , and intuition says that  $\theta^*$  should just be the mean  $\alpha^* / (\alpha^* + \beta^*)$  of the posterior distribution for  $\theta$  given  $y$ , in which  $\alpha^* = \alpha + s_n$  and  $\beta^* = \beta + n - s_n$  are the parameters of the Beta posterior. To check this, making use of the fact that the normalizing constant in the  $\text{Beta}(\alpha, \beta)$  family is  $\Gamma(\alpha + \beta) / \Gamma(\alpha) \Gamma(\beta)$ , the second equation in (6) gives

$$\begin{aligned} p(y_{n+1}|y_1, \dots, y_n, \mathcal{B}) &= \int_0^1 \theta^{y_{n+1}} (1 - \theta)^{1-y_{n+1}} \frac{\Gamma(\alpha^* + \beta^*)}{\Gamma(\alpha^*) \Gamma(\beta^*)} \theta^{\alpha^*-1} \\ &\quad \cdot (1 - \theta)^{\beta^*-1} d\theta \end{aligned}$$



Bayesian Statistics, Figure 1

Prior-to-posterior updating with two prior specifications in the mortality data set (in both panels, prior: long dotted lines; likelihood: short dotted lines; posterior: solid lines). The left and right panels give the updating with the priors for Analysts 1 and 2, respectively

$$\begin{aligned}
 &= \frac{\Gamma(\alpha^* + \beta^*)}{\Gamma(\alpha^*) \Gamma(\beta^*)} \int_0^1 \theta^{\alpha^* + y_{n+1} - 1} \\
 &\quad \cdot (1 - \theta)^{(\beta^* - y_{n+1} + 1) - 1} d\theta \\
 &= \left[ \frac{\Gamma(\alpha^* + y_{n+1})}{\Gamma(\alpha^*)} \right] \left[ \frac{\Gamma(\beta^* - y_{n+1} + 1)}{\Gamma(\beta^*)} \right] \\
 &\quad \cdot \left[ \frac{\Gamma(\alpha^* + \beta^*)}{\Gamma(\alpha^* + \beta^* + 1)} \right]; \quad (18)
 \end{aligned}$$

this, combined with the fact that  $\Gamma(x + 1) / \Gamma(x) = x$  for any real  $x$ , yields, for example in the case  $y_{n+1} = 1$ ,

$$\begin{aligned}
 p(y_{n+1} = 1 | y, \mathcal{B}) &= \left[ \frac{\Gamma(\alpha^* + 1)}{\Gamma(\alpha^*)} \right] \left[ \frac{\Gamma(\alpha^* + \beta^*)}{\Gamma(\alpha^* + \beta^* + 1)} \right] \\
 &= \frac{\alpha^*}{\alpha^* + \beta^*}, \quad (19)
 \end{aligned}$$

confirming intuition.

### Inference: Parametric and Non-Parametric Modeling of Count Data

Most elderly people in the Western world say they would prefer to spend the end of their lives at home, but many instead finish their lives in an institution (a nursing home

or hospital). How can elderly people living in their communities be offered health and social services that would help to prevent institutionalization? Hendriksen et al. [51] conducted an experiment in the 1980s in Denmark to test the effectiveness of *in-home geriatric assessment* (IHGA), a form of preventive medicine in which each person's medical and social needs are assessed and acted upon individually. A total of  $n = 572$  elderly people living in non-institutional settings in a number of villages were randomized,  $n_C = 287$  to a control group, who received standard health care, and  $n_T = 285$  to a treatment group, who received standard care plus IHGA. The number of hospitalizations during the two-year life of the study was an outcome of particular interest.

The data are presented and summarized in Table 1. Evidently IHGA lowered the mean hospitalization rate per two years (for the elderly Danish people in the study, at least) by  $(0.944 - 0.768) \doteq 0.176$ , which is about an 18% reduction from the control level, a clinically large difference. The question then becomes, in Bayesian inferential language: what is the posterior distribution for the treatment effect in the entire population  $\mathcal{P}$  of patients judged exchangeable with those in the study?

Bayesian Statistics, Table 1

Distribution of number of hospitalizations in the IHGA study

Group	Number of Hospitalizations								n	Sample	
	0	1	2	3	4	5	6	7		Mean	Variance
Control	138	77	46	12	8	4	0	2	287	0.944	1.54
Treatment	147	83	37	13	3	1	1	0	285	0.768	1.02

Continuing to refer to the relevant analyst as You, with a binary outcome variable and no covariates in Sect. “Inference and Prediction: Binary Outcomes with No Covariates” the model arose naturally from a judgment of exchangeability of Your uncertainty about all  $n$  outcomes, but such a judgment of *unconditional* exchangeability would not be appropriate initially here; to make such a judgment would be to assert that the treatment and control interventions have the same effect on hospitalization, and it was the point of the study to see if this is true. Here, at least initially, it would be more scientifically appropriate to assert exchangeability separately and in parallel within the two experimental groups, a judgment de Finetti [22] called *partial exchangeability* and which has more recently been referred to as *conditional exchangeability* [28,72] given the treatment/control status covariate.

Considering for the moment just the control group outcome values  $C_i, i = 1, \dots, n_C$ , and seeking as in Sect. “Inference and Prediction: Binary Outcomes with No Covariates” to model them via a predictive distribution  $p(C_1, \dots, C_{n_C} | \mathcal{B})$ , de Finetti’s previous representation theorem is not available because the outcomes are real-valued rather than binary, but he proved [21] another theorem for this situation as well: if You’re willing to regard  $(C_1, \dots, C_{n_C})$  as the first  $n_C$  terms in an infinitely exchangeable sequence  $(C_1, C_2, \dots)$  of values on  $\mathbb{R}$  (which plays the role of the population  $\mathcal{P}$ , under the control condition, in this problem), then to achieve coherence Your predictive distribution must be expressible as

$$p(C_1, \dots, C_{n_C} | \mathcal{B}) = \int_{\mathcal{F}} \prod_{i=1}^{n_C} F(C_i) dG(F | \mathcal{B}); \quad (20)$$

here (a)  $F$  has an interpretation as  $F(t) = \lim_{n_C \rightarrow \infty} F_{n_C}(t)$ , where  $F_{n_C}$  is the empirical CDF based on  $(C_1, \dots, C_{n_C})$ ; (b)  $G(F | \mathcal{B}) = \lim_{n_C \rightarrow \infty} p(F_{n_C} | \mathcal{B})$ , where  $p(\cdot | \mathcal{B})$  is Your joint probability distribution on  $(C_1, C_2, \dots)$ ; and (c)  $\mathcal{F}$  is the space of all possible CDFs on  $\mathbb{R}$ . Equation (20) says informally that exchangeability of Your uncertainty about an observable process unfolding on the real line is functionally equivalent to assuming the Bayesian hierarchical

model

$$\begin{aligned} (F | \mathcal{B}) &\sim p(F | \mathcal{B}) \\ (y_i | F, \mathcal{B}) &\stackrel{\text{iid}}{\sim} F, \end{aligned} \quad (21)$$

where  $p(F | \mathcal{B})$  is a prior distribution on  $\mathcal{F}$ . Placing distributions on functions, such as CDFs and regression surfaces, is the topic addressed by the field of *Bayesian non-parametric* (BNP) modeling [24,80], an area of statistics that has recently moved completely into the realm of day-to-day implementation and relevance through advances in MCMC computational methods. Two rich families of prior distributions on CDFs about which a wealth of practical experience has recently accumulated include (mixtures of) *Dirichlet processes* [32] and *Pólya trees* [66].

Parametric modeling is of course also possible with the IHGA data: as noted by Krnjajić et al. [62], who explore both parametric and BNP models for data of this kind, Poisson modeling is a natural choice, since the outcome consists of counts of relatively rare events. The first Poisson model to which one would generally turn is a *fixed-effects* model, in which  $(C_i | \lambda_C)$  are IID  $\text{Poisson}(\lambda_C)$  ( $i = 1, \dots, n_C = 287$ ) and  $(T_j | \lambda_T)$  are IID  $\text{Poisson}(\lambda_T)$  ( $j = 1, \dots, n_T = 285$ ), with a diffuse prior on  $(\lambda_C, \lambda_T)$  if little is known, external to the data set, about the underlying hospitalization rates in the control and treatment groups. However, the last two columns of Table 1 reveal that the sample variance is noticeably larger than the sample mean in both groups, indicating substantial Poisson over-dispersion. For a second, improved, parametric model this suggests a *random-effects* Poisson model of the form

$$\begin{aligned} (C_i | \lambda_{iC}) &\stackrel{\text{indep}}{\sim} \text{Poisson}(\lambda_{iC}) \\ [\log(\lambda_{iC}) | \beta_{0C}, \sigma_C^2] &\stackrel{\text{iid}}{\sim} N(\beta_{0C}, \sigma_C^2), \end{aligned} \quad (22)$$

and similarly for the treatment group, with diffuse priors for  $(\beta_{0C}, \sigma_C^2, \beta_{0T}, \sigma_T^2)$ . As Krnjajić et al. [62] note, from a medical point of view this model is more plausible than the fixed-effects formulation: each patient in the control group has his/her own *latent* (unobserved) underlying rate of hospitalization  $\lambda_{iC}$ , which may well differ from the underlying rates of the other control patients because of unmeasured differences in factors such as health status at the beginning of the experiment (and similarly for the treatment group).

Model (22), when complemented by its analogue in the treatment group, specifies a Lognormal mixture of Poisson distributions for each group and is straightforward to fit by MCMC, but the Gaussian assumption for the mixing distribution is conventional, not motivated by the underlying science of the problem, and if the distribution of the



latent variables is not Gaussian – for example, if it is multimodal or skewed – model (22) may well lead to incorrect inferences. Krnjajić et al. [62] therefore also examine several BNP models that are centered on the random-effects Poisson model but which permit learning about the true underlying distribution of the latent variables instead of assuming it is Gaussian. One of their models, when applied (for example) to the control group, was

$$\begin{aligned} (C_i | \lambda_{iC}) &\stackrel{\text{indep}}{\sim} \text{Poisson}(\lambda_{iC}) \\ [\log(\lambda_{iC}) | G] &\stackrel{\text{iid}}{\sim} G \\ (G | \alpha, \mu, \sigma^2) &\sim \text{DP}[\alpha N(\mu, \sigma^2)]. \end{aligned} \quad (23)$$

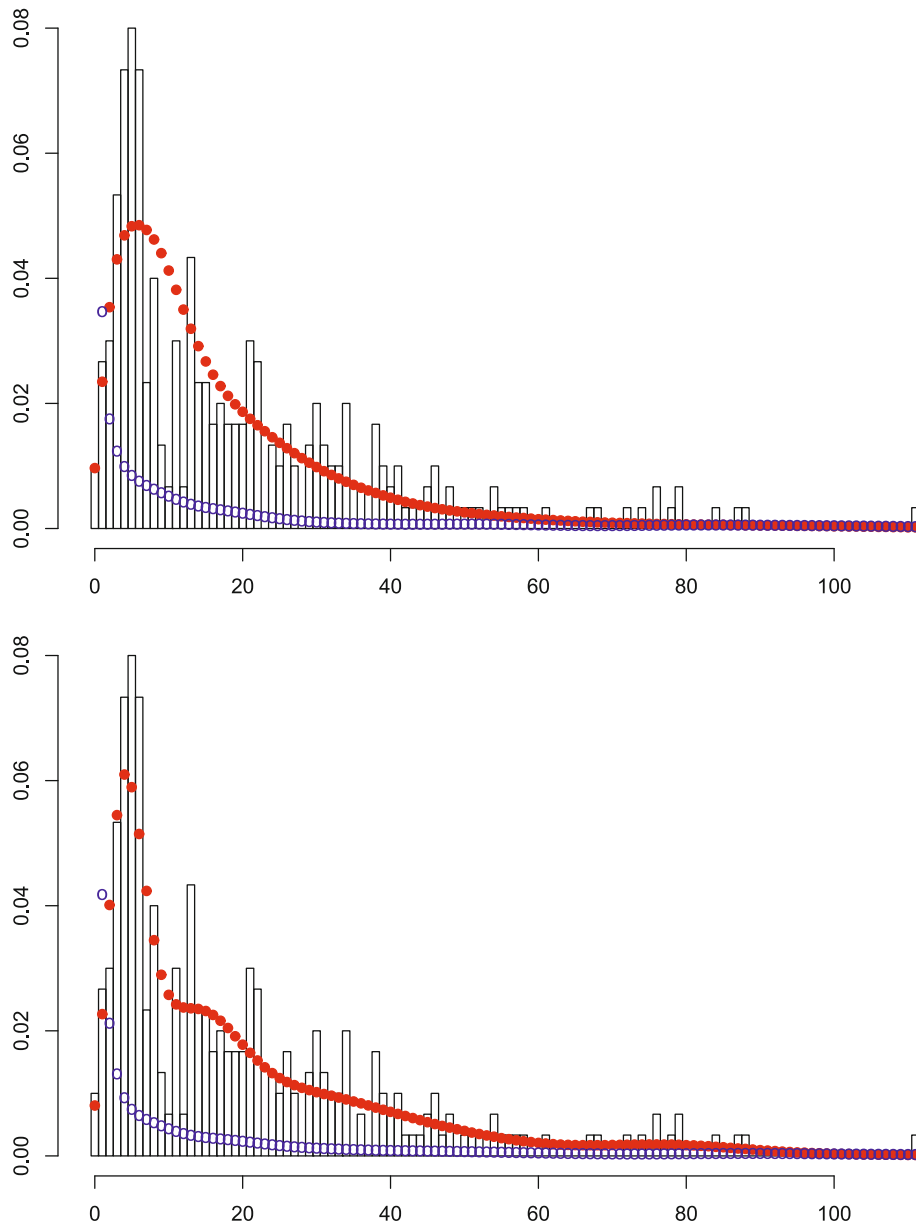
Here  $\text{DP}[\alpha N(\mu, \sigma^2)]$  refers to a Dirichlet process prior distribution, on the CDF  $G$  of the latent variables, which is centered at the  $N(\mu, \sigma^2)$  model with *precision parameter*  $\alpha$ . Model (23) is an *expansion* of the random-effects Poisson model (22) in that the latter is a special case of the former (obtained by letting  $\alpha \rightarrow \infty$ ). *Model expansion* is a common Bayesian analytic tool which helps to assess and propagate model uncertainty: if You are uncertain about a particular modeling detail, instead of fitting a model that assumes this detail is correct with probability 1, embed it in a richer model class of which it is a special case, and let the data tell You about its plausibility.

With the IHGA data, models (22) and (23) turned out to arrive at similar inferential conclusions – in both cases point estimates of the ratio of the treatment mean to the control mean were about 0.82 with a posterior standard deviation of about 0.09, and a posterior probability that the (population) mean ratio was less than 1 of about 0.95, so that evidence is strong that IHGA lowers mean hospitalizations not just in the sample but in the collection  $\mathcal{P}$  of elderly people to whom it is appropriate to generalize. But the two modeling approaches need not yield similar results: if the latent variable distribution is far from Gaussian, model (22) will not be able to adjust to this violation of one of its basic assumptions. Krnjajić et al. [62] performed a simulation study in which data sets with 300 observations were generated from various Gaussian and non-Gaussian latent variable distributions and a variety of parametric and BNP models were fit to the resulting count data; Fig. 2 summarizes the prior and posterior predictive distributions from models (22; top panel) and (23; bottom panel) with a bimodal latent variable distribution. The parametric Gaussian random-effects model cannot fit the bimodality on the data scale, but the BNP model – even though centered on the Gaussian as the random-effects distribution – adapts smoothly to the underlying bimodal reality.

### Decision-Making: Variable Selection in Generalized Linear Models; Bayesian Model Selection

Variable selection (choosing the “best” subset of predictors) in generalized linear models is an old problem, dating back at least to the 1960s, and many methods [113] have been proposed to try to solve it; but virtually all of them ignore an aspect of the problem that can be important: the cost of data collection of the predictors. An example, studied by Fouskakis and Draper [39], which is an elaboration of the problem examined in Sect. “[Inference and Prediction: Binary Outcomes with No Covariates](#)”, arises in the field of quality of health care measurement, where patient sickness at admission is often assessed by using logistic regression of an outcome, such as mortality within 30 days of admission, on a fairly large number of sickness indicators (on the order of 100) to construct a sickness scale, employing standard variable selection methods (for instance, backward selection from a model with all predictors) to find an “optimal” subset of 10–20 indicators that predict mortality well. The problem with such *benefit-only* methods is that they ignore the considerable differences among the sickness indicators in the *cost* of data collection; this issue is crucial when admission sickness is used to drive programs (now implemented or under consideration in several countries, including the US and UK) that attempt to identify substandard hospitals by comparing observed and expected mortality rates (given admission sickness), because such quality of care investigations are typically conducted under cost constraints. When both data-collection cost and accuracy of prediction of 30-day mortality are considered, a large variable-selection problem arises in which the only variables that make it into the final scale should be those that achieve a cost-benefit tradeoff.

Variable selection is an example of the broader process of *model selection*, in which questions such as “Is model  $M_1$  better than  $M_2$ ?” and “Is  $M_1$  good enough?” arise. These inquiries cannot be addressed, however, without first answering a new set of questions: good enough (better than) for what purpose? Specifying this purpose [26,57,61,70] identifies model selection as a decision problem that should be approached by constructing a contextually relevant utility function and maximizing expected utility. Fouskakis and Draper [39] create a utility function, for variable selection in their severity of illness problem, with two components that are combined additively: a data-collection component (in monetary units, such as US\$), which is simply the negative of the total amount of money required to collect data on a given set of patients with a given subset of the sickness indicators; and a predictive-accuracy component, in which a method

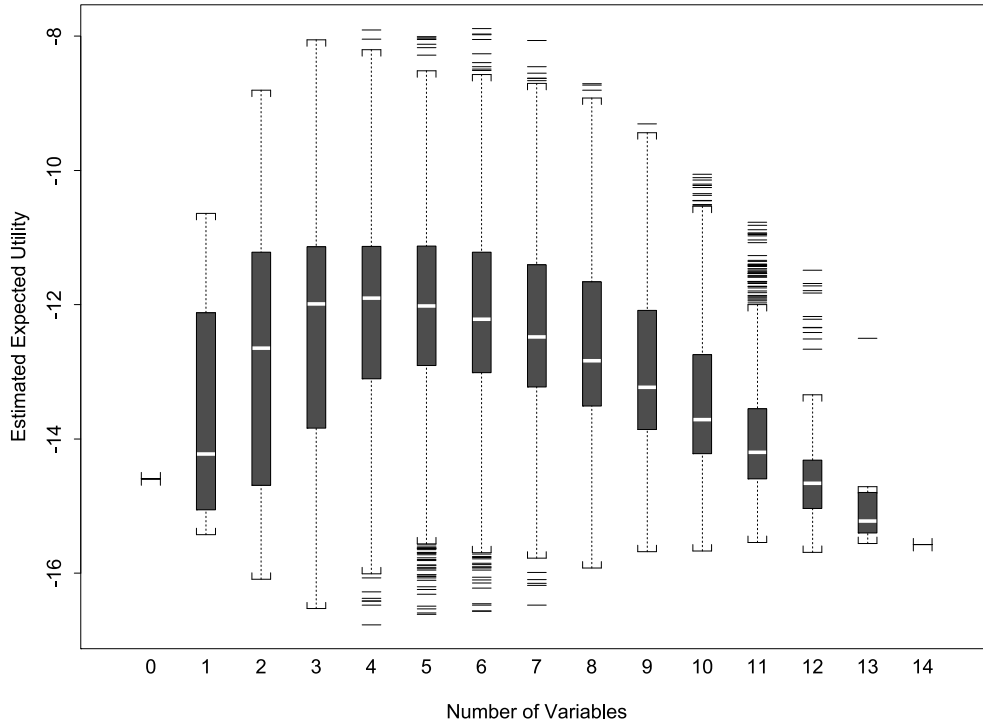


Bayesian Statistics, Figure 2

Prior (open circles) and posterior (solid circles) predictive distributions under models (22) and (23) (top and bottom panels, respectively) based on a data set generated from a bimodal latent variable distribution. In each panel, the histogram plots the simulated counts

is devised to convert increased predictive accuracy into decreased monetary cost by thinking about the consequences of labeling a hospital with bad quality of care “good” and vice versa. One aspect of their work, with a data set (from a RAND study: [58]) involving  $p = 83$  sickness indicators gathered on a representative sample of  $n = 2,532$  elderly American patients hospitalized in the period 1980–86 with pneumonia, focused only on the  $p = 14$  variables

in the original RAND sickness scale; this was chosen because  $2^{14} = 16,384$  was a small enough number of possible models to do brute-force enumeration of the estimated expected utility (EEU) of all the models. Figure 3 is a parallel boxplot of the EEUs of all 16,384 variable subsets, with the boxplots sorted by the number of variables in each model. The model with no predictors does poorly, with an EEU of about US\$–14.5, but from a cost-benefit point of



Bayesian Statistics, Figure 3

Estimated expected utility of all 16 384 variable subsets in the quality of care study based on RAND data

view the RAND sickness scale with all 14 variables is even worse (US\$−15.7), because it includes expensive variables that do not add much to the predictive power in relation to cheaper variables that predict almost as well. The best subsets have 4–6 variables and would save about US\$8 per patient when compared with the entire 14–variable scale; this would amount to significant savings if the observed-versus-expected assessment method were applied widely.

Returning to the general problem of Bayesian model selection, two cases can be distinguished: situations in which the precise purpose to which the model will be put can be specified (as in the variable-selection problem above), and settings in which at least some of the end uses to which the modeling will be put are not yet known. In this second situation it is still helpful to reason in a decision-theoretic way: the hallmark of a good (bad) model is that it makes good (bad) predictions, so a utility function based on predictive accuracy can be a good general-purpose choice. With (a) a single sample of data  $y$ , (b) a future data value  $y^*$ , and (c) two models  $M_j$  ( $j = 1, 2$ ) for illustration, what is needed is a *scoring rule* that measures the discrepancy between  $y^*$  and its predictive distribution  $p(y^*|y, M_j, \mathcal{B})$  under model  $M_j$ . It turns out [46,86] that the optimal (*impartial, symmetric, proper*) scoring

rules are linear functions of  $\log p(y^*|y, M_j, \mathcal{B})$ , which has a simple intuitive motivation: if the predictive distribution is Gaussian, for example, then values of  $y^*$  close to the center (in other words, those for which the prediction has been good) will receive a greater reward than those in the tails. An example [65], in this one-sample setting, of a model selection criterion (a) based on prediction, (b) motivated by utility considerations and (c) with good model discrimination properties [29] is the *full-sample log score*

$$LS_{FS}(M_j|y, \mathcal{B}) = \frac{1}{n} \sum_{i=1}^n \log p^*(y_i|y, M_j, \mathcal{B}), \quad (24)$$

which is related to the *conditional predictive ordinate* criterion [90]. Other Bayesian model selection criteria in current use include the following:

- *Bayes factors* [60]: Bayes' Theorem, written in odds form for discriminating between models  $M_1$  and  $M_2$ , says that

$$\frac{p(M_1|y, \mathcal{B})}{p(M_2|y, \mathcal{B})} = \left[ \frac{p(M_1|\mathcal{B})}{p(M_2|\mathcal{B})} \right] \cdot \left[ \frac{p(y|M_1, \mathcal{B})}{p(y|M_2, \mathcal{B})} \right]; \quad (25)$$

here the prior odds in favor of  $M_1$ ,

$$\frac{p(M_1|\mathcal{B})}{p(M_2|\mathcal{B})},$$

are multiplied by the *Bayes factor*

$$\frac{p(y|M_1, \mathcal{B})}{p(y|M_2, \mathcal{B})}$$

to produce the posterior odds

$$\frac{p(M_1|y, \mathcal{B})}{p(M_2|y, \mathcal{B})}.$$

According to the logic of this criterion, models with high posterior probability are to be preferred, and if all the models under consideration are equally plausible a priori this reduces to preferring models with larger Bayes factors in their favor. One problem with this approach is that – in parametric models in which model  $M_j$  has parameter vector  $\theta_j$  defined on parameter space  $\Theta_j$  – the *integrated likelihoods*  $p(y|M_j, \mathcal{B})$  appearing in the Bayes factor can be expressed as

$$\begin{aligned} p(y|M_j, \mathcal{B}) &= \int_{\Theta_j} p(y|\theta_j, M_j, \mathcal{B}) p(\theta_j|M_j, \mathcal{B}) d\theta_j \\ &= E_{(\theta_j|M_j, \mathcal{B})} [p(y|\theta_j, M_j, \mathcal{B})]. \end{aligned} \quad (26)$$

In other words, the numerator and denominator ingredients in the Bayes factor are each expressible as expectations of likelihood functions with respect to the *prior* distributions on the model parameters, and if context suggests that these priors should be specified diffusely the resulting Bayes factor can be unstable as a function of precisely how the diffuseness is specified. Various attempts have been made to remedy this instability of Bayes factors (for example, {partial, intrinsic, fractional} Bayes factors, well calibrated priors, conventional priors, intrinsic priors, expected posterior priors, ...; [9]); all of these methods appear to require an appeal to ad-hockery which is absent from the log score approach.

- *Deviance Information Criterion* (DIC): Given a parametric model  $p(y|\theta_j, M_j, \mathcal{B})$ , Spiegelhalter et al. [106] define the *deviance information criterion* (DIC) (by analogy with other information criteria) to be a trade-off between (a) an estimate of the model lack of fit, as measured by the *deviance*  $D(\bar{\theta}_j)$  (where  $\bar{\theta}_j$  is the posterior mean of  $\theta_j$  under  $M_j$ ; for the purpose of DIC, the deviance of a model [75] is minus twice the logarithm

of the likelihood for that model), and (b) a penalty for model complexity equal to twice the effective number of parameters  $p_{Dj}$  of the model:

$$DIC(M_j|y, \mathcal{B}) = D(\bar{\theta}_j) + 2 \hat{p}_{Dj}. \quad (27)$$

When  $p_{Dj}$  is difficult to read directly from the model (for example, in complex hierarchical models, especially those with random effects), Spiegelhalter et al. motivate the following estimate, which is easy to compute from standard MCMC output:

$$\hat{p}_{Dj} = \overline{D(\theta_j)} - D(\bar{\theta}_j); \quad (28)$$

in other words,  $\hat{p}_{Dj}$  is the difference between the posterior mean of the deviance and the deviance evaluated at the posterior mean of the parameters.

DIC is available as an option in several MCMC packages, including WinBUGS [107] and MLwiN [93]. One difficulty with DIC is that the MCMC estimate of  $p_{Dj}$  can be poor if the marginal posteriors for one or more parameters (using the parameterization that defines the deviance) are far from Gaussian; reparameterization (onto parameter scales where the posteriors are approximately Normal) helps but can still lead to mediocre estimates of  $p_{Dj}$ .

Other notable recent references on the subject of Bayesian variable selection include Brown et al. [13], who examine multivariate regression in the context of compositional data, and George and Foster [43], who use *empirical Bayes* methods in the Gaussian linear model.

## Comparison with the Frequentist Statistical Paradigm

### Strengths and Weaknesses of the Two Approaches

Frequentist statistics, which has concentrated mainly on inference, proceeds by (i) thinking of the values in a data set  $y$  as like a random sample from a *population*  $\mathcal{P}$  (a set to which it is hoped that conclusions based on the data can validly be generalized), (ii) specifying a summary  $\theta$  of interest in  $\mathcal{P}$  (such as the population mean of the outcome variable), (iii) identifying a function  $\hat{\theta}$  of  $y$  that can serve as a reasonable estimate of  $\theta$ , (iv) imagining repeating the random sampling from  $\mathcal{P}$  to get other data sets  $y$  and therefore other values of  $\hat{\theta}$ , and (v) using the random behavior of  $\hat{\theta}$  across these repetitions to make inferential probability statements involving  $\theta$ . A leading implementation of the frequentist paradigm [37] is based on using the value  $\hat{\theta}_{MLE}$  that maximizes the likelihood function as the estimate of  $\theta$  and obtaining a measure of uncertainty

for  $\hat{\theta}_{MLE}$  from the curvature of the logarithm of the likelihood function at its maximum; this is *maximum likelihood inference*.

Each of the frequentist and Bayesian approaches to statistics has strengths and weaknesses.

- The frequentist paradigm has the advantage that repeated-sampling calculations are often more tractable than manipulations with conditional probability distributions, and it has the clear strength that it focuses attention on the scientifically important issue of *calibration*: in settings where the true data-generating process is known (e. g., in simulations of random sampling from a known population  $\mathcal{P}$ ), how often does a particular method of statistical inference recover known truth? The frequentist approach has the disadvantage that it only applies to inherently repeatable phenomena, and therefore cannot be used to quantify uncertainty about many true-false propositions of real-world interest (for example, if You are a doctor to whom a new patient (male, say) has just come, strictly speaking You cannot talk about the frequentist probability that *this* patient is HIV positive; he either is or he is not, and his arriving at Your office is not the outcome of any repeatable process that is straightforward to identify). In practice the frequentist approach also has the weaknesses that (a) model uncertainty is more difficult to assess and propagate in this paradigm, (b) predictive uncertainty assessments are not always straightforward to create from the frequentist point of view (the *bootstrap* [31] is one possible solution) and (c) inferential calibration may not be easy to achieve when the sample size  $n$  is small.

An example of several of these drawbacks arises in the construction of *confidence intervals* [83], in which repeated-sampling statements such as

$$P_f(\hat{\theta}_{low} < \theta < \hat{\theta}_{high}) = 0.95 \quad (29)$$

(where  $P_f$  quantifies the frequentist variability in  $\hat{\theta}_{low}$  and  $\hat{\theta}_{high}$  across repeated samples from  $\mathcal{P}$ ) are interpreted in the frequentist paradigm as suggesting that the unknown  $\theta$  lies between  $\hat{\theta}_{low}$  and  $\hat{\theta}_{high}$  with 95% “confidence.” Two difficulties with this are that (a) equation (29) looks like a probability statement about  $\theta$  but is not, because in the frequentist approach  $\theta$  is a fixed unknown constant that cannot be described probabilistically, and (b) with small sample sizes *nominal* 95% confidence intervals based on maximum likelihood estimation can have actual *coverage* (the percentage of time in repeated sampling that the interval includes the true  $\theta$ ) substantially less than 95%.

- The Bayesian approach has the following clear advantages: (a) It applies (at least in principle) to uncertainty about anything, whether associated with a repeatable process or not; (b) inference is unambiguously based on the first equation in (6), without the need to face questions such as what constitutes a “reasonable” estimate of  $\theta$  (step (iii) in the frequentist inferential paradigm above); (c) prediction is straightforwardly and unambiguously based on the second equation in (6); and (d) in the problem of decision analysis a celebrated theorem of Wald [111] says informally that all good decisions can be interpreted as having been arrived at by maximizing expected utility as in the third equation of (6), so the Bayesian approach appears to be the way forward in decision problems rather broadly (but note the final challenge at the end of Sect. “[The Bayesian Statistical Paradigm](#)”). The principal disadvantage of the Bayesian approach is that coherence (internal logical consistency) by itself does not guarantee good calibration: You are free in the Bayesian paradigm to insert strong prior information in the modeling process (without violating coherence), and – if this information is seen after the fact to have been out of step with the world – Your inferences, predictions and/or decisions may also be off-target (of course, the same is true in the both the frequentist and Bayesian paradigms with regard to Your modeling of the likelihood information).

Two examples of frequentist inferences having poor calibration properties in small samples were given by Browne and Draper [14]. Their first example again concerns the measurement of quality of care, which is often studied with *cluster samples*: a random sample of  $J$  hospitals (indexed by  $j$ ) and a random sample of  $N$  total patients (indexed by  $i$ ) nested in the chosen hospitals is taken, and quality of care for the chosen patients and various hospital- and patient-level predictors are measured. With  $y_{ij}$  as the quality of care score for patient  $i$  in hospital  $j$ , a first step would often be to fit a *variance-components model* with random effects at both the hospital and patient levels, to assess the relative magnitudes of within- and between-hospital variability in quality of care:

$$y_{ij} = \beta_0 + u_j + e_{ij}, \quad i = 1, \dots, n_j, \quad j = 1, \dots, J; \\ \sum_{j=1}^J n_j = N, \quad (u_j | \sigma_u^2) \stackrel{\text{iid}}{\sim} N(0, \sigma_u^2), \quad (e_{ij} | \sigma_e^2) \stackrel{\text{iid}}{\sim} N(0, \sigma_e^2). \quad (30)$$

Browne and Draper [14] used a simulation study to show that, with a variety of maximum-likelihood-based meth-



ods for creating confidence intervals for  $\sigma_u^2$ , the actual coverage of nominal 95% intervals ranged from 72–94% across realistic sample sizes and true parameter values in the fields of education and medicine, versus 89–94% for Bayesian methods based on diffuse priors.

Their second example involved a re-analysis of a Guatemalan National Survey of Maternal and Child Health [89,97], with three-level data (births nested within mothers within communities), working with the random-effects logistic regression model

$$(y_{ijk} | p_{ijk}) \overset{\text{indep}}{\sim} \text{Bernoulli}(p_{ijk}) \quad \text{with} \\ \text{logit}(p_{ijk}) = \beta_0 + \beta_1 x_{1ijk} + \beta_2 x_{2jk} + \beta_3 x_{3k} + u_{jk} + v_k, \quad (31)$$

where  $y_{ijk}$  is a binary indicator of modern prenatal care or not and where  $u_{jk} \sim N(0, \sigma_u^2)$  and  $v_k \sim N(0, \sigma_v^2)$  were random effects at the mother and community levels (respectively). Simulating data sets with 2 449 births by 1 558 women living in 161 communities (as in the Rodríguez and Goldman study [97]), Browne and Draper [14] showed that things can be even worse for likelihood-based methods in this model, with actual coverages (at nominal 95%) as low as 0–2% for intervals for  $\sigma_u^2$  and  $\sigma_v^2$ , whereas Bayesian methods with diffuse priors again produced actual coverages from 89–96%. The technical problem is that the marginal likelihood functions for random-effects variances are often heavily skewed, with maxima at or near 0 even when the true variance is positive; Bayesian methods, which integrate over the likelihood function rather than maximizing it, can have (much) better small-sample calibration performance as a result.

### Some Historical Perspective

The earliest published formal example of an attempt to do statistical inference – to reason backwards from effects to causes – seems to have been Bayes [5], who defined conditional probability for the first time and noted that the result we now call Bayes’ Theorem was a trivial consequence of the definition. From the 1760s til the 1920s, all (or almost all) statistical inference was Bayesian, using the paradigm that Fisher and others referred to as *inverse probability*; prominent Bayesians of this period included Gauss [40], Laplace [64] and Pearson [88]. This Bayesian consensus changed with the publication of Fisher [37], which laid out a user-friendly program for maximum-likelihood estimation and inference in a wide variety of problems. Fisher railed against Bayesian inference; his principal objection was that in settings where little was known about a parameter (vector)  $\theta$  external to the data, a number of prior distributions could be put forward to quantify

this relative ignorance. He believed passionately in the late Victorian–Edwardian goal of scientific *objectivity*, and it bothered him greatly that two analysts with somewhat different diffuse priors might obtain somewhat different posteriors. (There is a Bayesian account of objectivity: a probability is objective if many different people more or less agree on its value. An example would be the probability of drawing a red ball from an urn known to contain 20 red and 80 white balls, if a sincere attempt is made to thoroughly mix the balls without looking at them and to draw the ball in a way that does not tend to favor one ball over another.)

There are two problems with Fisher’s argument, which he never addressed:

1. He would be perfectly correct to raise this objection to Bayesian analysis if investigators were often forced to do inference based solely on prior information with no data, but in practice with even modest sample sizes the posterior is relatively insensitive to the precise manner in which diffuseness is specified in the prior, because the likelihood information in such situations is relatively so much stronger than the prior information; Sect. “[Inference and Prediction: Binary Outcomes with No Covariates](#)” provides an example of this phenomenon.
2. If Fisher had looked at the entire process of inference with an engineering eye to sensitivity and stability, he would have been forced to admit that uncertainty in how to specify the likelihood function has inferential consequences that are often an order of magnitude larger than those arising from uncertainty in how to specify the prior. It is an inescapable fact that subjectivity, through assumptions and judgments (such as the form of the likelihood function), is an integral part of any statistical analysis in problems of realistic complexity.

In spite of these unrebutted flaws in Fisher’s objections to Bayesian inference, two schools of frequentist inference – one based on Fisher’s maximum-likelihood estimation and *significance tests* [38], the other based on the confidence intervals and *hypothesis tests* of Neyman [83] and Neyman and Pearson [84] – came to dominate statistical practice from the 1920s at least through the 1980s. One major reason for this was practical: the Bayesian paradigm is based on integrating over the posterior distribution, and accurate approximations to high-dimensional integrals were not available during the period in question. Fisher’s technology, based on differentiation (to find the maximum and curvature of the logarithm of the likelihood function) rather than integration, was a much more

tractable approach for its time. Jeffreys [54], working in the field of astronomy, and Savage [102] and Lindley [69], building on de Finetti's results, advocated forcefully for the adoption of Bayesian methods, but prior to the advent of MCMC techniques (in the late 1980s) Bayesians were often in the position of saying that they knew the best way to solve statistical problems but the computations were beyond them. MCMC has removed this practical objection to the Bayesian paradigm for a wide class of problems.

The increased availability of affordable computers with decent CPU throughput in the 1980s also helped to overcome one objection raised in Sect. “[Strengths and Weaknesses of the Two Approaches](#)” against likelihood methods, that they can produce poorly-calibrated inferences with small samples, through the introduction of the bootstrap by Efron [31] in 1979. At this writing (a) both the frequentist and Bayesian paradigms are in vigorous inferential use, with the proportion of Bayesian articles in leading journals continuing an increase that began in the 1980s; (b) Bayesian MCMC analyses are often employed to produce meaningful predictive conclusions, with the use of the bootstrap increasing for frequentist predictive calibration; and (c) the Bayesian paradigm dominates decision analysis.

### A Bayesian-Frequentist Fusion

During the 20th century the debate over which paradigm to use was often framed in such a way that it seemed it was necessary to choose one approach and defend it against attacks from people who had chosen the other, but there is nothing that forces an analyst to choose a single paradigm. Since both approaches have strengths and weaknesses, it seems worthwhile instead to seek a *fusion* of the two that makes best use of the strengths. Because (a) the Bayesian paradigm appears to be the most flexible way so far developed for quantifying all sources of uncertainty and (b) its main weakness is that coherence does not guarantee good calibration, a number of statisticians, including Rubin [98], Draper [26], and Little [73], have suggested a fusion in which inferences, predictions and decisions are formulated using Bayesian methods and then evaluated for their calibration properties using frequentist methods, for example by using Bayesian models to create 95% predictive intervals for observables not used in the modeling process and seeing if approximately 95% of these intervals include the actual observed values. Analysts more accustomed to the purely frequentist (likelihood) paradigm who prefer not to explicitly make use of prior distributions may still find it useful to reason in a Bayesian way, by integrating over the parameter uncer-

tainty in their likelihood functions rather than maximizing over it, in order to enjoy the superior calibration properties that integration has been demonstrated to provide.

### Future Directions

Since the mid- to late-1980s the Bayesian statistical paradigm has made significant advances in many fields of inquiry, including agriculture, archaeology, astronomy, bioinformatics, biology, economics, education, environmetrics, finance, health policy, and medicine (see Sect. “[The Bayesian Statistical Paradigm](#)” for recent citations of work in many of these disciplines). Three areas of methodological and theoretical research appear particularly promising for extending the useful scope of Bayesian work, as follows:

- *Elicitation of prior distributions and utility functions:* It is arguable that too much use is made in Bayesian analysis of diffuse prior distributions, because (a) accurate elicitation of non-diffuse priors is hard work and (b) lingering traces still remain of a desire to at least appear to achieve the unattainable Victorian-Edwardian goal of objectivity, the (false) argument being that the use of diffuse priors somehow equates to an absence of subjectivity (see, e.g., the papers by Berger [7] and Goldstein [45] and the ensuing discussion for a vigorous debate on this issue). It is also arguable that too much emphasis was placed in the 20th century on inference at the expense of decision-making, with inferential tools such as the Neyman-Pearson hypothesis testing machinery (Sect. “[Some Historical Perspective](#)”) used incorrectly to make decisions for which they are not optimal; the main reason for this, as noted in Sect. “[Strengths and Weaknesses of the Two Approaches](#)” and “[Some Historical Perspective](#)”, is that (a) the frequentist paradigm was dominant from the 1920s through the 1980s and (b) the high ground in decision theory is dominated by the Bayesian approach. Relevant citations of excellent recent work on elicitation of prior distributions and utility functions were given in Sect. “[The Bayesian Statistical Paradigm](#)”; it is natural to expect that there will be a greater emphasis on decision theory and non-diffuse prior modeling in the future, and elicitation in those fields of Bayesian methodology is an important area of continuing research.
- *Group decision-making:* As noted in Sect. “[The Bayesian Statistical Paradigm](#)”, maximizing expected utility is an effective method for decision-making by a single agent, but when two or more agents are involved in the decision process this approach cannot be guaranteed to

yield a satisfying solution: there may be conflicts in the agents' preferences, particularly if their relationship is at least partly adversarial. With three or more possible actions, *transitivity* of preference – if You prefer action  $a_1$  to  $a_2$  and  $a_2$  to  $a_3$ , then You should prefer  $a_1$  to  $a_3$  – is a criterion that any reasonable decision-making process should obey; informally, a well-known theorem by Arrow [3] states that even if all of the agents' utility functions obey transitivity, there is no way to combine their utility functions into a single decision-making process that is guaranteed to respect transitivity. However, Arrow's theorem is temporally static, in the sense that the agents do not share their utility functions with each other and iterate after doing so, and it assumes that all agents have the same set  $\mathcal{A}$  of feasible actions. If agents  $A_1$  and  $A_2$  have action spaces  $\mathcal{A}_1$  and  $\mathcal{A}_2$  that are not identical and they share the details of their utility specification with each other, it is possible that  $A_1$  may realize that one of the actions in  $\mathcal{A}_2$  that (s)he had not considered is better than any of the actions in  $\mathcal{A}_1$  or vice versa; thus a temporally dynamic solution to the problem posed by Arrow's theorem may be possible, even if  $A_1$  and  $A_2$  are partially adversarial. This is another important area for new research.

- *Bayesian computation*: Since the late 1980s, simulation-based computation based on Markov chain Monte Carlo (MCMC) methods has made useful Bayesian analyses possible in an increasingly broad range of application areas, and (as noted in Sect. “[The Bayesian Statistical Paradigm](#)”) increases in computing speed and sophistication of MCMC algorithms have enhanced this trend significantly. However, if a regression-style data set is visualized as a matrix with  $n$  rows (one for each subject of inquiry) and  $k$  columns (one for each variable measured on the subjects), MCMC methods do not necessarily scale well in either  $n$  or  $k$ , with the result that they can be too slow to be of practical use with large data sets (e.g. at current desktop computing speeds, with  $n$  and/or  $k$  on the order of  $10^5$  or greater). Improving the scaling of MCMC methods, or finding a new approach to Bayesian computation that scales better, is thus a third important area for continuing study.

## Bibliography

1. Abdellaoui M (2000) Parameter-free elicitation of utility and probability weighting functions. *Manag Sci* 46:1497–1512
2. Aleskerov F, Bouyssou D, Monjardet B (2007) *Utility Maximization, Choice and Preference*, 2nd edn. Springer, New York
3. Arrow KJ (1963) *Social Choice and Individual Values*, 2nd edn. Yale University Press, New Haven CT
4. Barlow RE, Wu AS (1981) Preposterior analysis of Bayes estimators of mean life. *Biometrika* 68:403–410
5. Bayes T (1764) An essay towards solving a problem in the doctrine of chances. *Philos Trans Royal Soc Lond* 53:370–418
6. Berger JO (1985) *Statistical Decision Theory and Bayesian Analysis*. Springer, New York
7. Berger JO (2006) The case for objective Bayesian analysis (with discussion). *Bayesian Anal* 1:385–472
8. Berger JO, Betro B, Moreno E, Pericchi LR, Ruggeri F, Salinetti G, Wasserman L (eds) (1995) *Bayesian Robustness*. Institute of Mathematical Statistics Lecture Notes-Monograph Series, vol 29. IMS, Hayward CA
9. Berger JO, Pericchi LR (2001) Objective Bayesian methods for model selection: introduction and comparison. In: Lahiri P (ed) *Model Selection*. Monograph Series, vol 38. Institute of Mathematical Statistics Lecture Notes Series, Beachwood, pp 135–207
10. Bernardo JM (1979) Reference posterior distributions for Bayesian inference (with discussion). *J Royal Stat Soc, Series B* 41:113–147
11. Bernardo JM, Smith AFM (1994) *Bayesian Theory*. Wiley, New York
12. Blavatskyy P (2006) Error propagation in the elicitation of utility and probability weighting functions. *Theory Decis* 60:315–334
13. Brown PJ, Vannucci M, Fearn T (1998) Multivariate Bayesian variable selection and prediction. *J Royal Stat Soc, Series B* 60:627–641
14. Browne WJ, Draper D (2006) A comparison of Bayesian and likelihood methods for fitting multilevel models (with discussion). *Bayesian Anal* 1:473–550
15. Buck C, Blackwell P (2008) Bayesian construction of radiocarbon calibration curves (with discussion). In: *Case Studies in Bayesian Statistics*, vol 9. Springer, New York
16. Chipman H, George EI, McCulloch RE (1998) Bayesian CART model search (with discussion). *J Am Stat Assoc* 93:935–960
17. Christiansen CL, Morris CN (1997) Hierarchical Poisson regression modeling. *J Am Stat Assoc* 92:618–632
18. Clyde M, George EI (2004) Model uncertainty. *Stat Sci* 19:81–94
19. Das S, Chen MH, Kim S, Warren N (2008) A Bayesian structural equations model for multilevel data with missing responses and missing covariates. *Bayesian Anal* 3:197–224
20. de Finetti B (1930) Funzione caratteristica di un fenomeno aleatorio. *Mem R Accad Lincei* 4:86–133
21. de Finetti B (1937). La prévision: ses lois logiques, ses sources subjectives. *Ann Inst H Poincaré* 7:1–68 (reprinted in translation as de Finetti B (1980) Foresight: its logical laws, its subjective sources. In: Kyburg HE, Smokler HE (eds) *Studies in Subjective Probability*. Dover, New York, pp 93–158)
22. de Finetti B (1938/1980). Sur la condition d'équivalence partielle. *Actual Sci Ind* 739 (reprinted in translation as de Finetti B (1980) On the condition of partial exchangeability. In: Jeffrey R (ed) *Studies in Inductive Logic and Probability*. University of California Press, Berkeley, pp 193–206)
23. de Finetti B (1970) *Teoria delle Probabilità*, vol 1 and 2. Einaudi, Torino (reprinted in translation as de Finetti B (1974–75) *Theory of probability*, vol 1 and 2. Wiley, Chichester)
24. Dey D, Müller P, Sinha D (eds) (1998) *Practical Nonparametric and Semiparametric Bayesian Statistics*. Springer, New York

25. Draper D (1995) Assessment and propagation of model uncertainty (with discussion). *J Royal Stat Soc, Series B* 57:45–97
26. Draper D (1999) Model uncertainty yes, discrete model averaging maybe. Comment on: Hoeting JA, Madigan D, Raftery AE, Volinsky CT (eds) Bayesian model averaging: a tutorial. *Stat Sci* 14:405–409
27. Draper D (2007) Bayesian multilevel analysis and MCMC. In: de Leeuw J, Meijer E (eds) *Handbook of Multilevel Analysis*. Springer, New York, pp 31–94
28. Draper D, Hodges J, Mallows C, Pregibon D (1993) Exchangeability and data analysis (with discussion). *J Royal Stat Soc, Series A* 156:9–37
29. Draper D, Krnjajić M (2008) Bayesian model specification. Submitted
30. Duran BS, Booker JM (1988) A Bayes sensitivity analysis when using the Beta distribution as a prior. *IEEE Trans Reliab* 37:239–247
31. Efron B (1979) Bootstrap methods. *Ann Stat* 7:1–26
32. Ferguson T (1973) A Bayesian analysis of some nonparametric problems. *Ann Stat* 1:209–230
33. Ferreira MAR, Lee HKH (2007) *Multiscale Modeling*. Springer, New York
34. Fienberg SE (2006) When did Bayesian inference become “Bayesian”? *Bayesian Anal* 1:1–40
35. Fishburn PC (1970) *Utility Theory for Decision Making*. Wiley, New York
36. Fishburn PC (1981) Subjective expected utility: a review of normative theories. *Theory Decis* 13:139–199
37. Fisher RA (1922) On the mathematical foundations of theoretical statistics. *Philos Trans Royal Soc Lond, Series A* 222:309–368
38. Fisher RA (1925) *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh
39. Fouskakis D, Draper D (2008) Comparing stochastic optimization methods for variable selection in binary outcome prediction, with application to health policy. *J Am Stat Assoc*, forthcoming
40. Gauss CF (1809) *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium*, vol 2. Perthes and Besser, Hamburg
41. Gelfand AE, Smith AFM (1990) Sampling-based approaches to calculating marginal densities. *J Am Stat Assoc* 85:398–409
42. Gelman A, Meng X-L (2004) *Applied Bayesian Modeling and Causal Inference From Incomplete-Data Perspectives*. Wiley, New York
43. George EI, Foster DP (2000) Calibration and empirical Bayes variable selection. *Biometrika* 87:731–747
44. Gilks WR, Richardson S, Spiegelhalter DJ (eds) (1996) *Markov Chain Monte Carlo in Practice*. Chapman, New York
45. Goldstein M (2006) Subjective Bayesian analysis: principles and practice (with discussion). *Bayesian Anal* 1:385–472
46. Good IJ (1950) *Probability and the Weighing of Evidence*. Charles Griffin, London
47. Green P (1995) Reversible jump Markov chain Monte carlo computation and Bayesian model determination. *Biometrika* 82:711–713
48. Hacking I (1984) *The Emergence of Probability*. University Press, Cambridge
49. Hanson TE, Kottas A, Branscum AJ (2008) Modelling stochastic order in the analysis of receiver operating characteristic data: Bayesian non-parametric approaches. *J Royal Stat Soc, Series C (Applied Statistics)* 57:207–226
50. Hellwig K, Speckbacher G, Weniges P (2000) Utility maximization under capital growth constraints. *J Math Econ* 33:1–12
51. Hendriksen C, Lund E, Stromgard E (1984) Consequences of assessment and intervention among elderly people: a three year randomized controlled trial. *Br Med J* 289:1522–1524
52. Hoeting JA, Madigan D, Raftery AE, Volinsky CT (1999) Bayesian model averaging: a tutorial. *Stat Sci* 14:382–417
53. Jaynes ET (2003) *Probability Theory: The Logic of Science*. Cambridge University Press, Cambridge
54. Jeffreys H (1931) *Scientific Inference*. Cambridge University Press, Cambridge
55. Jordan MI, Ghahramani Z, Jaakkola TS, Saul L (1999) An introduction to variational methods for graphical models. *Mach Learn* 37:183–233
56. Kadane JB (ed) (1996) *Bayesian Methods and Ethics in a Clinical Trial Design*. Wiley, New York
57. Kadane JB, Dickey JM (1980) Bayesian decision theory and the simplification of models. In: Kmenta J, Ramsey J (eds) *Evaluation of Econometric Models*. Academic Press, New York
58. Kahn K, Rubenstein L, Draper D, Koseoff J, Rogers W, Keeler E, Brook R (1990) The effects of the DRG-based Prospective Payment System on quality of care for hospitalized Medicare patients: An introduction to the series (with discussion). *J Am Med Assoc* 264:1953–1955
59. Kass RE, Raftery AE (1995) Bayes factors. *J Am Stat Assoc* 90:773–795
60. Kass RE, Wasserman L (1996) The selection of prior distributions by formal rules. *J Am Stat Assoc* 91:1343–1370
61. Key J, Pericchi LR, Smith AFM (1999) Bayesian model choice: what and why? (with discussion). In: Bernardo JM, Berger JO, Dawid AP, Smith AFM (eds) *Bayesian Statistics 6*. Clarendon Press, Oxford, pp 343–370
62. Krnjajić M, Kottas A, Draper D (2008) Parametric and non-parametric Bayesian model specification: a case study involving models for count data. *Comput Stat Data Anal* 52: 2110–2128
63. Laplace PS (1774) Mémoire sur la probabilité des causes par les événements. *Mém Acad Sci Paris* 6:621–656
64. Laplace PS (1812) *Théorie Analytique des Probabilités*. Courcier, Paris
65. Laud PW, Ibrahim JG (1995) Predictive model selection. *J Royal Stat Soc, Series B* 57:247–262
66. Lavine M (1992) Some aspects of Pólya tree distributions for statistical modelling. *Ann Stat* 20:1222–1235
67. Leamer EE (1978) *Specification searches: Ad hoc inference with non-experimental data*. Wiley, New York
68. Leonard T, Hsu JSJ (1999) *Bayesian Methods: An Analysis for Statisticians and Interdisciplinary Researchers*. Cambridge University Press, Cambridge
69. Lindley DV (1965) *Introduction to Probability and Statistics*. Cambridge University Press, Cambridge
70. Lindley DV (1968) The choice of variables in multiple regression (with discussion). *J Royal Stat Soc, Series B* 30:31–66
71. Lindley DV (2006) *Understanding Uncertainty*. Wiley, New York
72. Lindley DV, Novick MR (1981) The role of exchangeability in inference. *Ann Stat* 9:45–58
73. Little RJA (2006) Calibrated Bayes: A Bayes/frequentist roadmap. *Am Stat* 60:213–223



74. Mangel M, Munch SB (2003) Opportunities for Bayesian analysis in the search for sustainable fisheries. *ISBA Bulletin* 10:3–5
75. McCullagh P, Nelder JA (1989) *Generalized Linear Models*, 2nd edn. Chapman, New York
76. Meng XL, van Dyk DA (1997) The EM Algorithm: an old folk song sung to a fast new tune (with discussion). *J Royal Stat Soc, Series B* 59:511–567
77. Merl D, Prado R (2007) Detecting selection in DNA sequences: Bayesian modelling and inference (with discussion). In: Bernardo JM, Bayarri MJ, Berger JO, Dawid AP, Heckerman D, Smith AFM, West M (eds) *Bayesian Statistics 8*. University Press, Oxford, pp 1–22
78. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equations of state calculations by fast computing machine. *J Chem Phys* 21:1087–1091
79. Morris CN, Hill J (2000) The Health Insurance Experiment: design using the Finite Selection Model. In: Morton SC, Rolph JE (eds) *Public Policy and Statistics: Case Studies from RAND*. Springer, New York, pp 29–53
80. Müller P, Quintana F (2004) Nonparametric Bayesian data analysis. *Stat Sci* 19:95–110
81. Müller P, Quintana F, Rosner G (2004) Hierarchical meta-analysis over related non-parametric Bayesian models. *J Royal Stat Soc, Series B* 66:735–749
82. Munch SB, Kottas A, Mangel M (2005) Bayesian nonparametric analysis of stock-recruitment relationships. *Can J Fish Aquat Sci* 62:1808–1821
83. Neyman J (1937) Outline of a theory of statistical estimation based on the classical theory of probability. *Philos Trans Royal Soc Lond A* 236:333–380
84. Neyman J, Pearson ES (1928) On the use and interpretation of certain test criteria for purposes of statistical inference. *Biometrika* 20:175–240
85. O'Hagan A, Buck CE, Daneshkhah A, Eiser JR, Garthwaite PH, Jenkinson DJ, Oakley JE, Rakow T (2006) *Uncertain Judgements. Eliciting Experts' Probabilities*. Wiley, New York
86. O'Hagan A, Forster J (2004) Bayesian Inference, 2nd edn. In: Kendall's *Advanced Theory of Statistics*, vol 2B. Arnold, London
87. Parmigiani G (2002) *Modeling in medical decision-making: A Bayesian approach*. Wiley, New York
88. Pearson KP (1895) Mathematical contributions to the theory of evolution, II. Skew variation in homogeneous material. *Proc Royal Soc Lond* 57:257–260
89. Pebley AR, Goldman N (1992) Family, community, ethnic identity, and the use of formal health care services in Guatemala. Working Paper 92-12. Office of Population Research, Princeton
90. Pettit LI (1990) The conditional predictive ordinate for the Normal distribution. *J Royal Stat Soc, Series B* 52:175–184
91. Pérez JM, Berger JO (2002) Expected posterior prior distributions for model selection. *Biometrika* 89:491–512
92. Polson NG, Stroud JR, Müller P (2008) Practical filtering with sequential parameter learning. *J Royal Stat Soc, Series B* 70:413–428
93. Rashbash J, Steele F, Browne WJ, Prosser B (2005) *A User's Guide to MLwiN, Version 2.0*. Centre for Multilevel Modelling, University of Bristol, Bristol UK; available at [www.cmm.bristol.ac.uk](http://www.cmm.bristol.ac.uk) Accessed 15 Aug 2008
94. Richardson S, Green PJ (1997) On Bayesian analysis of mixtures with an unknown number of components (with discussion). *J Royal Stat Soc, Series B* 59:731–792
95. Rios Insua D, Ruggeri F (eds) (2000) *Robust Bayesian Analysis*. Springer, New York
96. Rodríguez A, Dunston DB, Gelfand AE (2008) The nested Dirichlet process. *J Am Stat Assoc*, 103, forthcoming
97. Rodríguez G, Goldman N (1995) An assessment of estimation procedures for multilevel models with binary responses. *J Royal Stat Soc, Series A* 158:73–89
98. Rubin DB (1984) Bayesianly justifiable and relevant frequency calculations for the applied statistician. *Ann Stat* 12:1151–1172
99. Rubin DB (2005) Bayesian inference for causal effects. In: Rao CR, Dey DK (eds) *Handbook of Statistics: Bayesian Thinking, Modeling and Computation*, vol 25. Elsevier, Amsterdam, pp 1–16
100. Sabatti C, Lange K (2008) Bayesian Gaussian mixture models for high-density genotyping arrays. *J Am Stat Assoc* 103:89–100
101. Sansó B, Forest CE, Zantedeschi D (2008) Inferring climate system properties using a computer model (with discussion). *Bayesian Anal* 3:1–62
102. Savage LJ (1954) *The Foundations of Statistics*. Wiley, New York
103. Schervish MJ, Seidenfeld T, Kadane JB (1990) State-dependent utilities. *J Am Stat Assoc* 85:840–847
104. Seidou O, Asselin JJ, Ouarda TMBJ (2007) Bayesian multivariate linear regression with application to change point models in hydrometeorological variables. In: *Water Resources Research* 43, W08401, doi:10.1029/2005WR004835.
105. Spiegelhalter DJ, Abrams KR, Myles JP (2004) *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. Wiley, New York
106. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A (2002) Bayesian measures of model complexity and fit (with discussion). *J Royal Stat Soc, Series B* 64:583–640
107. Spiegelhalter DJ, Thomas A, Best NG (1999) *WinBUGS Version 1.2 User Manual*. MRC Biostatistics Unit, Cambridge
108. Stephens M (2000) Bayesian analysis of mixture models with an unknown number of components – an alternative to reversible-jump methods. *Ann Stat* 28:40–74
109. Stigler SM (1986) *The History of Statistics: The Measurement of Uncertainty Before 1900*. Harvard University Press, Cambridge
110. Stone M (1974) Cross-validation choice and assessment of statistical predictions (with discussion). *J Royal Stat Soc, Series B* 36:111–147
111. Wald A (1950) *Statistical Decision Functions*. Wiley, New York
112. Weerahandi S, Zidek JV (1981) Multi-Bayesian statistical decision theory. *J Royal Stat Soc, Series A* 144:85–93
113. Weisberg S (2005) *Applied Linear Regression*, 3rd edn. Wiley, New York
114. West M (2003) Bayesian factor regression models in the “large p, small n paradigm.” *Bayesian Statistics* 7:723–732
115. West M, Harrison PJ (1997) *Bayesian Forecasting and Dynamic Models*. Springer, New York
116. Whitehead J (2006) Using Bayesian decision theory in dose-escalation studies. In: Chevret S (ed) *Statistical Methods for Dose-Finding Experiments*. Wiley, New York, pp 149–171



## Biochemistry, Chaotic Dynamics, Noise, and Fractal Space in

MIGUEL ANTONIO AON, SONIA CORTASSA

Institute of Molecular Cardiobiology, The Johns Hopkins University, Baltimore, USA

### Article Outline

Glossary  
 Definition of the Subject  
 Introduction  
 Cellular Self-Organization and Nonlinear Dynamics  
 Fractals as Geometric, Statistic,  
 and Dynamic Objects  
 Fractals in Biochemistry and Physiology  
 Chaos in Biochemistry and Physiology  
 Chaos and Colored Noise  
 Noise in Biochemical and Cellular Systems  
 Noise and Network Behavior  
 Future Directions  
 Bibliography

### Glossary

**Scaling** a quantity scales if it changes according to a power law whose exponent is insensitive to the details of the processes involved, a feature usually referred to as universality.

**Scale free** refers to the geometry or dynamic of structures or processes with no single characteristic scale (space) or frequency (temporal).

**Critical** applies to states or phenomena characterized by their extreme sensitivity to small changes or perturbations.

**Fractal** is an object of geometric, statistical, or dynamical nature that obeys power laws of the form  $M(L) \propto L^D$ , with  $D$  as the non-integer fractal dimension. Mandelbrot introduced “fractal geometry” and defined a fractal “as a shape made of parts similar to the whole in some way”.

**Chaotic** is a dynamic, non-periodic, behavior that describes ‘turbulent’ motions in time that exhibiting sensitive dependence on initial conditions.

**Deterministic** dynamic behavior specified by difference or differential equations.

**Noise** refer to random time series originating from stochastic fluctuations of events of different nature (e.g. molecular, electric, energetic as in the biological realm) that in a certain context may represent a signal.

**Mitochondria** sub-cellular organelles considered to be the main energy-producers of the cell as well as the source of important signaling functions, among them apoptosis or cell death.

**Network** the collective organization of an ensemble of objects, or groups of them, in space (structural and topological) and time (dynamics).

### Definition of the Subject

The description of fractals, noise, and chaotic dynamics, is deeply associated with the concepts of *scaling* and *critical* phenomena. The recent discovery of the occurrence of scaling, criticality, and colored noise behavior in mitochondrial networks – at a fundamental level of cellular organization and dynamics as it relates to energetics, life and death – further emphasizes the role of fractals and chaos in biochemistry.

The concept of fractals introduced by Mandelbrot was initially applied to the geometric description of irregular objects but quickly spread to the field of dynamics where it joined chaos. Initially described by Lorenz in the context of meteorology, chaos extended his influence from fluid dynamics to biology and biochemistry.

Analysis of time series of biological variables with techniques inspired by fractal and chaos theories are providing a more clear understanding of the relationship between scaling in space and time exhibited by cells and organs in health and disease. These studies have also contributed to differentiate between chaos and noise, while revealing the increasing role of the power law description in networks’ topology and dynamics.

The example of the cardiac mitochondrial network clearly states that the major elements of a theory of organized complexity based on the integration of Systems Biology and the Complex Systems Approach are at hand.

### Introduction

In the undergoing shift between the *analytical* and *integrative* periods of biology, the importance of the concept of complex systems and systemic approaches has become dominant [1]. Cells, organisms and ecosystems are complex because they consist of a large number of mutually interacting parts. By exchanging energy, matter or information with their environment complex systems exhibit emergence, that is they self-organize their internal structure and their dynamics with novel and sometimes surprising macroscopic properties [2,3,4,5]. The appearance of emergent properties – that do not result from the existence of a central controller – is the single most distin-

guishing feature of complex systems, and self-organization the main mechanism of generating complexity [2,5,6,7,8].

The *Complex Systems Approach* involves analyzing inter-connections and relationships rather than component parts. It is nowadays pervasive, playing an increasing role in most of the scientific disciplines, including biology (cellular biochemistry, molecular biology, cellular and developmental biology). Its main foundation is the concept of self-organization based on non-equilibrium thermodynamics of nonlinear open systems, i. e. those that exchange energy and matter like cells, organisms, and ecosystems. Three modern developments, fractals, chaos, and networks' topology and dynamics, have further extended the complex systems approach.

The advent of high throughput technologies, and the feasibility to mathematically model complicated systems, have made possible the quantitative integration of information to gain insights into the nature of health and disease, and their control. This resumes the approach of *Systems Biology* as understood at present. It is timely to integrate *Systems Biology* to the group of important theoretical concepts as well as experimental and mathematical techniques loosely grouped here as the *Complex Systems Approach* (Fig. 1). This may prove fruitful and insightful for biochemistry and physiology.

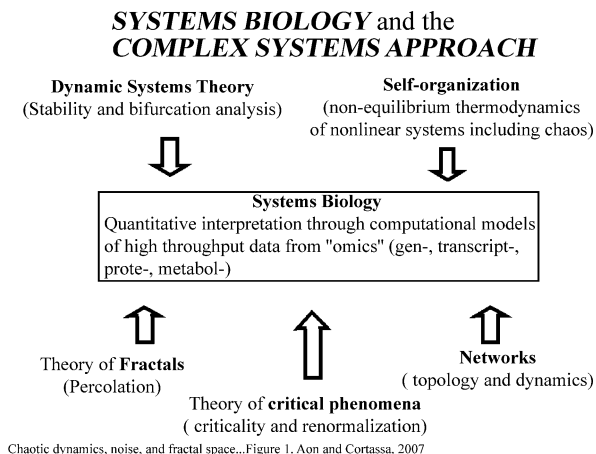
Originally, the concept of fractals arose to explain the geometry of irregular objects that is the great majority either from the macroscopic or microscopic world that surrounds us [9,10]. The discovery of chaos in dynamics [11]

(see [12] for a historical review) and criticality in phase transitions [5,13] allowed us to clearly see that *scaling* is a common, foundational, concept of chaos and criticality as it is with fractals. Thus, scaling appears in the description of fractals, in the theory of critical phenomena and in chaotic dynamics [14].

Scaling appeals to the interaction between the multiple levels of organization exhibited by cells and organisms, thus linking the spatial and temporal aspects of their organization [15,16,17]. Scaling is a feature displayed by an unexpectedly large selection of biological systems. Scaling means that a quantity depends through a power law on its argument. Thus, if a variable value changes according to a *power law* when the parameter it depends on is growing linearly, we say it *scales*, and the corresponding exponent is called a *critical exponent*. The concept of scaling has been extended to the fractal description and characterization of the dynamic performance of systems in general (e. g. biological, economical, geophysical) exhibiting self-organized (chaotic) or (colored) noise behavior (see below). A great insight there has been to show that in time series from, e. g. market prices or mitochondrial membrane potential, the short-term fluctuations are intrinsically related to the long-term trends, appealing to the notion of networks.

Three main directions have characterized the study of cellular networks [18]: (i) architectural (structural morphology); (ii) topological (connectivity properties); and (iii) dynamical (time-dependent behavior). Network organization is at the basis of the collective and large-scale spreading consequences exhibited by complex phenomena [19,20]. Networks exhibit scale-free topologies, i. e. most of the nodes in a network will have only a few links and these will be held together by a small number of nodes exhibiting high connectivity [21]. This is a manifestation of the scaling and inverse power law features exhibited by network organization in the topological and dynamical sense. As a result, inverse power laws characterize the normal, physiological, as well as the pathophysiological behavior of the cardiac mitochondrial network [18,22,23].

Biologists have the notion of scaling from long ago, especially associated with the relationship between the size of organisms and their metabolic rates, or the so-called allometric relationship [2,24]. The concept of organisms' design in terms of fractal geometry as an underlying explanation of the link between size and metabolism has reinvigorated this research field (see [25] and Refs. therein). Why is the concept of scaling so pervading and fundamental for biology and biochemistry? Because it shows the intrinsic interdependence among the different levels of organization exhibited by living systems, which is expressed by the correlation among the different scales of space and



#### Biochemistry, Chaotic Dynamics, Noise, and Fractal Space in, Figure 1

This diagram shows the main theoretical concepts from the *Complex Systems Approach* that may contribute to a more integrated and realistic view of *Systems Biology* as presently understood

time involved [2,15,16,25,26,27]. Geometric and dynamic fractals capture the essence of this crucial feature through scaling and self-similarity [26].

Scaling also expresses the essential organizational traits of different systems independent of their nature. Some relevant examples are the complex set of dynamic motions or chaos detected at the onset of turbulence, as well as critical phenomena. Chaos and critical behavior display *universality* since universal numbers [28,29] or a few critical exponents [5,10] characterize them. In this context, *universality* means the insensitivity of the exponent of power laws to the details of the processes considered. Critical exponents also characterize the transition of the cardiac mitochondrial network from physiological to pathophysiological behavior [30], as another recent example described later. If one day we may presume of having a theory of complexity the concept of scaling will certainly be in its foundations.

A central insight derived from the Complex Systems Approach has been to show that networks of social or economic nature, or complex systems – such as human-engineered machines or biological ones tinkered by evolution – when stressed, can collapse, crash, or rupture. These events that happen under extreme conditions will trigger fundamental biological and societal changes and as such qualify as crises ranging from disease and epidemics to large natural catastrophes (e.g. earthquakes, hurricanes) [5,17] or stock markets crash. The cascade of failures shown to scale from the subcellular (mitochondrial network) to the organ (heart) resulting in reperfusion-related arrhythmias after ischemic injury, and eventually the death of the organism, is another thrilling example [31]. From our innermost biochemistry and physiology to the large scale of natural and human phenomena, the same laws are at play.

### Cellular Self-Organization and Nonlinear Dynamics

A cell is the autonomous unit of life. Its autonomy is based on two fundamental concepts: “operational closure” [32] and self-organization [4]. Operational closure refers to the capacity of living systems for self-making of their organization (autopoietic networks) through self-organization by continuous exchange of matter and energy [32]. According to the emerging view of living systems, cells consist of self-making networks of mass, energy and information continually flowing, but the overall system maintains a stable form autonomously through self-organization, i.e. dynamic organization [1,2,7,18]. Chaos along with multiple stationary states and oscillations belong to the range of dynamic self-organized behaviors exhibited by cells.

Thermodynamically, cells and organisms behave as open systems continuously exchanging matter and energy. Under non-equilibrium steady state conditions, free energy is continuously dissipated thus the entropy production from irreversible processes inside the system,  $d_i S$ , always increases whereas the flow of entropy exchange with the surroundings,  $d_e S$ , may be negative. It is in the non-equilibrium state that self-organized structures can emerge spontaneously implying a decrease in entropy maintained by energy and matter dissipation and as such called *dissipative structures* [4]. Although living organisms continually increase their entropy (positive  $d_i S$ ) and tend to approach the dangerous state of maximum entropy (or death), they stay away from it by continually drawing from its environment negative entropy ( $d_e S$ ) or, more rigorously, free energy to keep its organization [33,34].

In many cases, the stationary state represents a delicate balance among the units of the system but avalanche-like series of changes may take place when it is close to a critical state. Such large events express more than anything else the underlying “forces” usually hidden under almost perfect balance providing a potential for a better understanding of complex systems [5,35]. Under those conditions, the system becomes very sensitive to perturbations. Many biological systems can be considered both, self-organized and critical.

### Fractals as Geometric, Statistic, and Dynamic Objects

Fractals are of geometrical, statistical and dynamical nature [14]. Geometrically, fractals can be looked at as structures exhibiting *scaling* in *space* since their mass as a function of size, or their density as a function of distance, behave as a power law [9,35]. Additionally, the scaling is self-similar meaning that the object remains invariant at several length scales. The fractal dimension,  $D_f$ , characterizes the invariant scaling of objects exhibiting this type of geometry.

Mandelbrot (see [36] for a review), and West (1990) among others, expanded the concept of fractals from geometric objects to dynamic processes exhibiting fractal statistics. Fractal dynamics correspond to processes that are scale free, i.e. they do not exhibit a fundamental scale, and are represented by inverse power laws.

The double nature of fractals as geometric (spatial) objects and dynamic processes has allowed the use of common techniques for their characterization and quantification. For instance, lacunarity used as a measure of the non-uniformity (heterogeneity) of structures, is quantified through mass related distributions. These distribu-

tions can be characterized by the coefficient of variation or relative dispersion, RD (= standard deviation, SD/mean) that is a strong function of scale [9,26,37]. Relative dispersion as applied to the temporal domain suggests that events in different time scales are tied together through fractal statistics whose span will be as large as the scaling holds for at least two orders of magnitude. For self-similar time series, Relative Dispersional Analysis (RDA) shows that RD is constant with scale (i. e., the object looks the same at all scales) [17,26]. Since RDA implies the binning of the data set at successively larger time scales, for a system exhibiting completely random fluctuations in time, its RD drops off more rapidly as compared with a system that shows long-term memory. As a result, an inverse power law with a slope corresponding to a  $D_f$  close to 1.0 suggests high correlation among the components of a system exhibiting dynamics processes with long-term memory (see next section below).

In the temporal domain, the existence of fractals has also been investigated using Power Spectral Analysis (PSA). As in geometrically fractal objects, the dynamics of processes with long-term memory correspond to self-similar scaling in frequency. Thus, correlations in time series are also revealed using PSA; the power spectrum of such data usually follows an inverse power law proportional to  $1/f^\beta$ . This power law, more generally known as  $1/f$  noise, describes colored noise depending on the value of  $\beta$  or the spectral exponent (e. g.  $\beta = 1, 2$  or  $3$ , for pink, brown, or black noise, respectively). The spectral exponent equals zero for white noise; thus, the spectrum of white noise is independent of frequency. If we integrate a white noise over time, we get “brown” noise that has a power spectrum that is proportional to  $f^{-2}$  over an extended frequency.

## Fractals in Biochemistry and Physiology

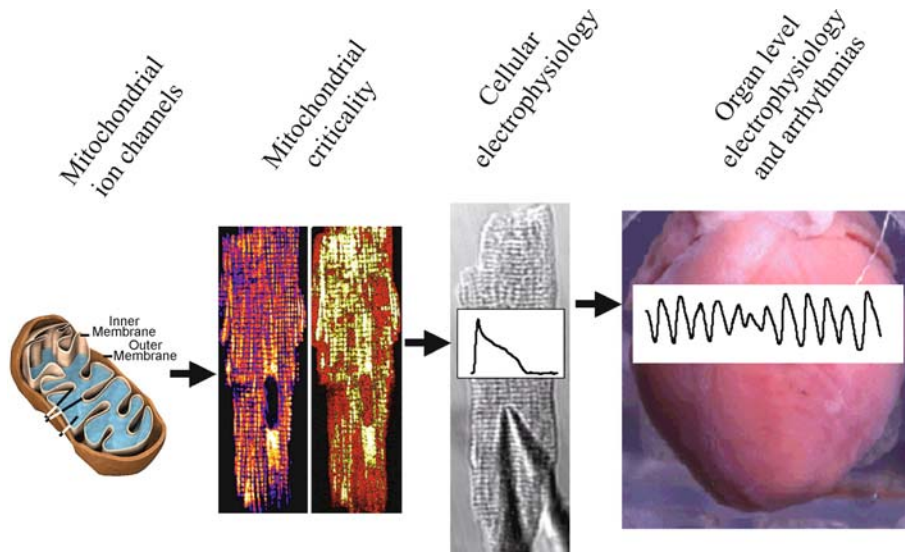
In cellular biochemistry, the fractal approach has been directed to understanding the organization and behavior of (macro)molecules in cells [16,38,39,40] whereas in physiology it has been used as a theoretical and analytical tool of signals [26,41,42]. The former field is still in its infancy whereas the latter has already shown a relevant progress.

The recent finding of fractal organization and scaling during transitions in mitochondrial physiology [30] is significant because it links cellular biochemistry with organ (patho)physiology [31,43,44] (Fig. 2). An early proposal [2,45] suggested that in the cellular cytoplasm fractal-like kinetics could arise as a result of cytoskeletal proteins’ fractal organization and anomalous diffusion [46]. Recently, this proposal has been further sub-

stantiated [16,40,47] as primary steps towards understanding the dynamics of biochemical reactions in the intracellular environment [48,49,50]. The approach to fractal kinetics in cells differs between these authors. All of them depart from Kopelman (1988) whose main observation was that in chemical reactions occurring in heterogeneous media the rate constants become time-dependent scaling with time as  $t^{-h}$ , the  $h$  parameter reflecting the characteristics of the medium, e. g. obstacle density. The dependence of the rate constant upon  $h$  has been modeled according to a fractal Zipf–Mandelbrot distribution [40], assuming Michaelis–Menten kinetics in 2D lattices using Monte Carlo simulations with time-dependent rate constant [47], or in terms of the dependence of the parameter  $h$  on the spectral dimension,  $D_s$ , for Michaelis–Menten or sigmoidal kinetics [16]. The main findings show that: (i) spatial segregation of substrate and products increase with the degree of obstruction in the medium making stronger the deviation of the rate constants at longer times and, consequently, the fractal kinetic description as compared with the classical approach [40,47]; (ii) at short times the reaction rate becomes much larger, at low substrate levels, in fractal than in Euclidean space; this behavior being explained by the time-dependence of the Michaelis–Menten constant ( $K_M$ ) or an increase in cooperativity and reaction amplification in allosteric kinetics [16]. The latter behavior is of utmost importance for cells when challenged by sudden changes in environmental conditions since the quick relaxation of molecular mechanisms will allow fast and precise adaptation.

Heart physiology is one of the fields that benefited the most from applying fractal ideas. Pioneering efforts to apply fractal concepts in cardiac physiology uncovered that healthiness, as judged from electrocardiographic (EKG) data, display a broadband frequency spectrum of the QT interval [51]. The finding of an inverse power-law spectrum in the QRS complex arising from the EKG [51] suggested that heart dynamics exhibits no fundamental time scale (period) [41,52,53] unlike in cardiac disease where there is “loss of spectral reserve”. The latter was suggested by spectrum simplification and the appearance of characteristic frequencies [14,26,54]. The loss of “spectral reserve” in cardiac interbeat interval spectra [54] has been determined in a number of disease processes including development of heart failure or postinfarction arrhythmias [55,56,57].

Inverse power law behavior is a manifestation of self-similar, fractal, behavior in the dynamics of a process. Fractal time series are associated with power spectra in which the higher the frequency component, the lower its power [14]. On the basis of the fractal organization of the



#### Biochemistry, Chaotic Dynamics, Noise, and Fractal Space in, Figure 2

Scaling mitochondrial dynamics to arrhythmias in the heart. Opening of inner mitochondrial membrane ion channels results in loss of membrane potential ( $\Delta\Psi_m$ ) (left panel) that under criticality (a mitochondrial cluster with threshold levels of reactive oxygen species, ROS) can scale to the mitochondrial network through neighbor-neighbor interaction (second panel). Whole-cell electrical and mechanical dysfunction occurs due to the coupling of metabolism with energy-sensitive ion channels in the sarcolemma and the  $\text{Ca}^{2+}$ -handling subsystem (third panel). Dispersion of action potential repolarization contributes to impaired propagation in the myocardial syncytium, resulting in reentrant cardiac arrhythmias in the postischemic heart (fourth panel)

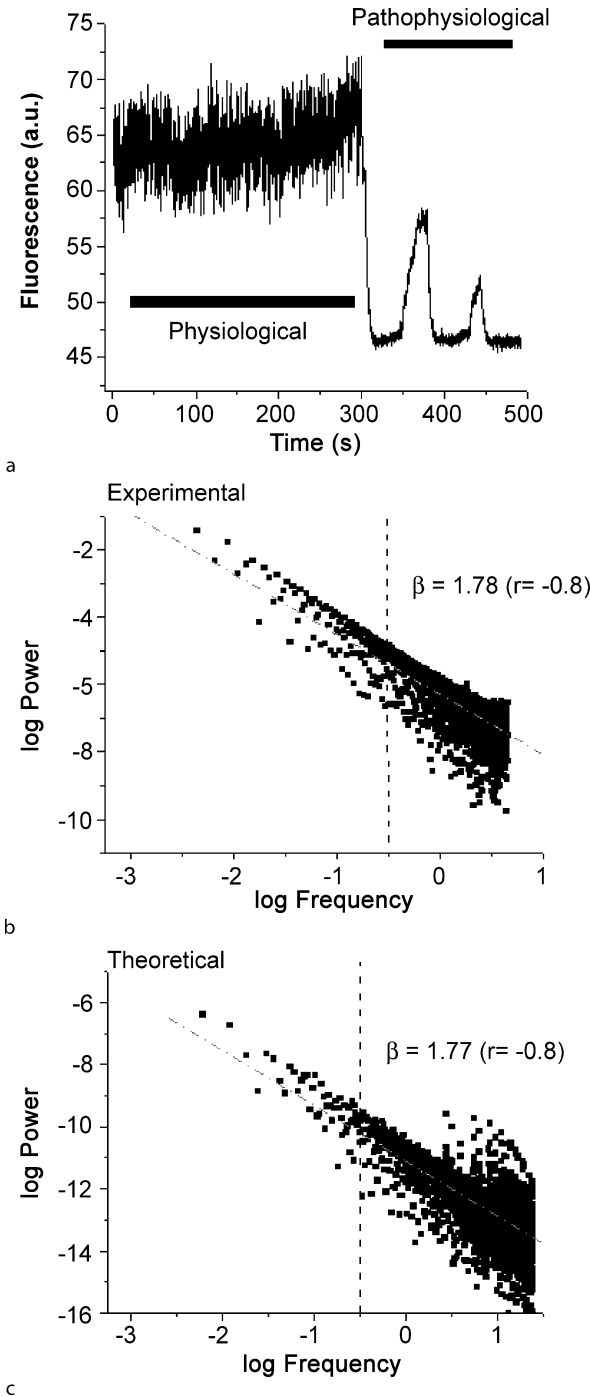
His–Purkinje conduction system, an interpretation of the inverse power law spectrum of the QRS-complex has been offered [14,51]. It was proposed that the effect of the finely branching His–Purkinje conduction network might be to subtly “decorrelate” the individual pulses that superpose to form the QRS-complex. A fractal decorrelation cascade of  $N$  voltage-time pulses results whose time trace exhibits a spectrum obeying an inverse power-law behavior. What is the mechanism of such inverse power law spectra? The inverse power law spectrum can be viewed as the resultant of many processes interacting over a myriad of interdependent scales [14]. Multiple control processes including neurohumoral regulation and local electrochemical factors regulate fluctuations in heart rate. Competition between sympathetic and parasympathetic stimulation on pacemaker cells of the sinoatrial node represent the autonomic component of the heart rate variability. A minor role of the sympathetic nervous system on the  $1/f$  spectrum following administration of propranolol, a  $\beta$ -adrenergic blocker, has been reported [58]. The contribution of non-autonomic factors such as mitochondrial energetics to the long-range correlations in the human heartbeat time series has not been investigated.

Inverse power laws also characterize the physiological and pathophysiological behaviors of the cardiac mitochondrial network [22] (Fig. 3a). At the center stage of

cell energetics, mitochondrial oxidative phosphorylation is a prime example of a biological system regulated so as to rapidly and robustly match changes in energy supply to meet cellular demand. A profound implication of the power law behavior exhibited by the mitochondrial network is that beyond a critical level, the loss of a few mitochondria may cause the entire network to fail [30] because of their multiplicative interdependency [14] (Fig. 4). In heart, the loss of mitochondrial function through collapse of  $\Delta\Psi_m$  following metabolic stress directly affects the sarcolemmal  $K_{\text{ATP}}$  channel and alters the cellular action potential [AP] [59,60]. The rapid uncoupling of oxidative phosphorylation during depolarization of  $\Delta\Psi_m$  is closely linked to the activation of sarcolemmal  $K_{\text{ATP}}$  currents, consequently shortening the cellular AP [60] and rendering the myocyte electrically inexcitable during the nadir of  $\Delta\Psi_m$  oscillation [59]. This mechanism has been recently shown to contribute to destabilize the AP repolarization during reperfusion after ischemic injury in the whole heart, leading to arrhythmias [31] (Fig. 2). This confirms the prediction derived from the inherent organization of mitochondria as a network, i.e. that failures can scale to higher levels of organization [43,44].

In cardiac cells, the collective dynamics of the mitochondrial network was characterized from time series using RDA and PSA analysis (see previous section). The





collective dynamics of the mitochondrial membrane potential,  $\Delta\Psi_m$ , belongs to a statistically fractal, self-similar process characterized by a large number of frequencies in multiple timescales, rather than an inherent “characteristic” frequency [22]. The dynamic behavior of a spatially

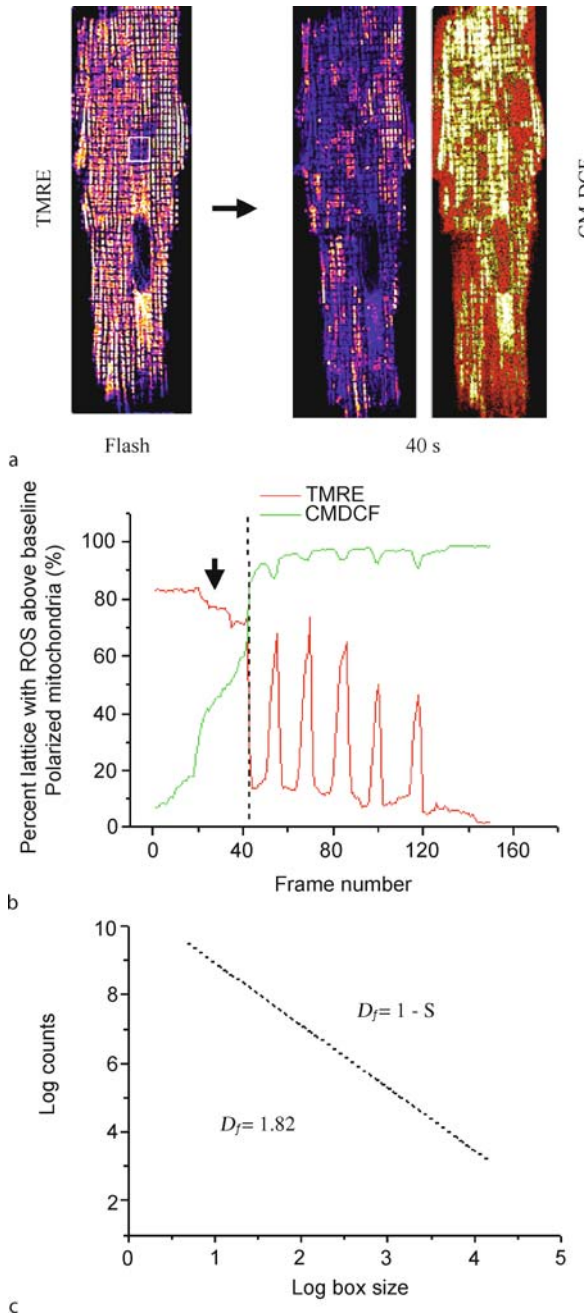
#### ◀ Biochemistry, Chaotic Dynamics, Noise, and Fractal Space in, Figure 3

Physiological and pathophysiological behaviors of the mitochondrial network in heart cells. **a** Freshly isolated ventricular cardiomyocytes were loaded with 100 nM of the cationic potentiometric fluorescent dye TMRM, and imaged by two photon microscopy (150 ms time resolution). The results obtained from a stack of 3,720 images are shown. Before the mitochondrial network reaches criticality, the  $\Delta\Psi_m$  (as measured by TMRM fluorescence) oscillates at high frequencies and small amplitudes. After criticality, the network dynamics evolve into “pathophysiological” behavior characterized by low frequency high-amplitude oscillations. **b** Statistical analysis of experimentally obtained  $\Delta\Psi_m$  time series of the mitochondrial network of living cardiomyocytes loaded with TMRM. Power Spectral Analysis (PSA) of  $\Delta\Psi_m$  time series after Fast Fourier Transform reveals a broad spectrum of oscillation in normally polarized mitochondria with a spectral exponent,  $\beta = 1.78$  (a random process gives a  $\beta = 0$  meaning that there is no relationship between the amplitude and the frequency in a random signal). The observed relationship between the amplitude squared (power) and frequency obeys a homogeneous power law ( $1/f^\beta$ ; with  $f$ , frequency, and  $\beta$ , the spectral exponent), and is statistically self-similar which means that there is no dominant frequency. **c** After simulation of oscillations with our computational model of a mitochondrial oscillator, we selected five oscillatory periods in the high frequency domain (between 70 and 300 ms) and one from the low frequency (1 min period) domain and attributed each one of them proportionally to a network composed by 500 mitochondria [22]. A matrix containing a total of 500 columns (mitochondria) and 6,000 rows was constructed. The time steps represented by the rows correspond to a fixed integration step of 20 ms for the numerical integration of the system of ordinary differential equations. Shown are the results after applying PSA to the average value of each row of the matrix at e.g. time 1,  $T_1$ , that represents the experimental average value of fluorescence intensity of the  $\Delta\Psi_m$  probe (corresponding to mV) obtained every 150 ms from 500 mitochondria (on average as in panel a) from each image of our stack. In panels b and c, the vertical dashed line at 0.3 Hz delimits the high frequency region that contains 93% of the data points (to the right of the line corresponds to the high frequency low-amplitude domain of the spectrum). Notice the close parallel between the experimental (b) and theoretical data (c)

heterogeneous system can be interpreted from the correlation between nearest neighbors in terms of covariance through the spatial correlation coefficient,  $r_1$  [62]:

$$r_1 = 2^{3-2D} - 1. \quad (1)$$

The RD of aggregates of nearest neighbors obeys a fractal relationship [26]. Equation (1) can also be applied to fractal random time series, in addition to spatially inhomogeneous data sets, where  $D$  (or  $D_f$ ) is the fractal dimension [17]. The derivation of the single-time-lag correlation coefficient,  $r_1$ , of a time series [17] is based on similar arguments as those used for Eq. (1). With perfect correla-



tion,  $r_1 = 1$ , the mitochondrial behavior is correlated everywhere [26]. Processes without memory such as white or brown noise show completely random behavior which is characterized by an exponential law with slopes corresponding to  $D_f = 1.5$  [17,26] whereas pink noise exhibits  $D_f$  close to 1.0 [22].

In the physiological regime, the power spectrum of mitochondria corresponds to a network of coupled os-

#### ◀ Biochemistry, Chaotic Dynamics, Noise, and Fractal Space in, Figure 4

Mitochondrial criticality and fractal organization at the percolation threshold. **a** The mitochondrial lattice of cardiomyocytes, as imaged by two-photon laser scanning microscopy, was labeled with the membrane potential marker TMRE (left and middle panels; TMRE labeled cell, intensity scaled from blue to white;) and a reactive oxygen species (ROS)-sensitive fluorescent probe (right panel; CM-H<sub>2</sub>DCFDA signal, scaled from red [reduced] to yellow [oxidized]). **b** The time course of  $\Delta\Psi_m$  depolarization (the laser flash is indicated by the arrow in this panel and corresponds to the white square on the left image of panel a). After the laser flash a mitochondrial spanning cluster develops (within 40 s on average) as mitochondrial ROS accumulates. Criticality is reached when about 60% of the mitochondria have CM-DCF fluorescence levels  $\sim 20\%$  above baseline (right image on panel a), and after which sustained oscillations in  $\Delta\Psi_m$  happen as depicted. This coincides with the expected theoretical percolation threshold according to standard 2D percolation theory [16,30]. **c** The fractal dimension,  $D_f$ , of the mitochondrial cluster, as revealed by the ROS probe (right image on panel a), was also consistent with percolation theory

cillators obeying an inverse power law proportional to  $1/f^\beta$  ( $\beta = 1.7$ ) (Fig. 3b). This is indicative of a broad range of frequencies and amplitudes spanning at least three orders of magnitude in time (from ms to a few min). The broad frequency distribution, and the long-term temporal correlations exhibited by the mitochondrial network, allows a change in one time scale to be felt across the frequency range. This is reminiscent of heart rate regulation in a normal subject, suggesting the intriguing possibility that a change in mitochondrial network dynamics could entrain pathological consequences for the cell and the organism.

The inverse power law behavior exhibited by the mitochondrial network dynamics obeys a precise mechanism. Amplitude and oscillatory frequency exhibit an inverse relationship that, mechanistically, reflects the balance between the rate of ROS production and scavenging [22]. This relationship demonstrates two essential features of the oscillatory dynamics displayed by mitochondria: (i) the link between period and amplitude, (i. e., the longer the period, the higher the amplitude), and (ii) at periods  $> 300$  ms, the transition to a single dominant frequency with large  $\Delta\Psi_m$  depolarization that define the limit of the physiological domain.

#### Chaos in Biochemistry and Physiology

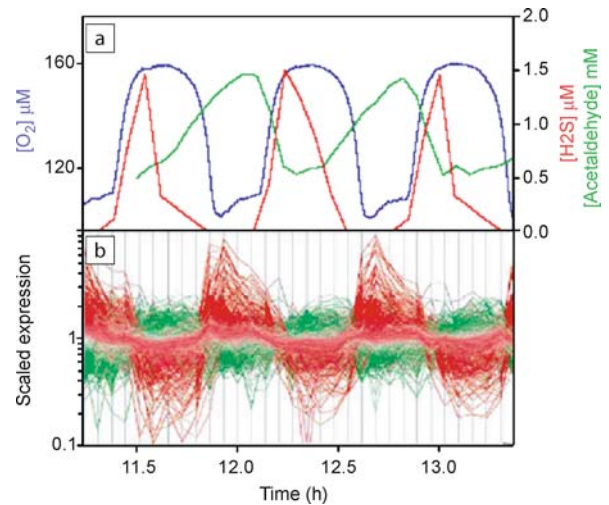
Homeodynamics rather than homeostasis characterize the dynamic motion of biological systems [7]. In a biological system, homeodynamics can be better visualized in the *state space*, which is  $n$ -dimensional and represents the set

of all possible states of the system with  $n$  as the number of variables. In the case of the *phase space* (i.e. a two-dimensional state space) the system's motion is given by the trajectories (whose ensemble is the phase portrait) of the system like a fluid flowing around itself.

An *attractor* is a restricted region of the phase space to which all the trajectories settle or approach over time. *Strange attractors* characterize chaotic dynamics when represented in phase space. It is well known that strange attractors can be quantitatively described as fractals, its dimension being a non-integer that is less than the dimensionality,  $n$ , of the phase space (see below). In contrast, time series of colored noise will have an integer dimension equal to the dimensionality of the phase space.

A great number of unstable limit cycles (with different periods) are embedded in a strange attractor that may be selected (at least theoretically) to a stable periodic output by small perturbations of their parameters [63,64]. Mechanistically, chaos can be controlled by a combination of periodic perturbations in a feedback control circuit with delay [65,66]. A controlled chaotic attractor can, potentially, be a multi-oscillator that can be tuned into different time domains. Evidence has been provided that in the heart, multifrequency oscillations in rhythm and amplitude create a highly unstable prebrillatory condition that presages fibrillation [67,68].

In recent years, an ultradian clock that functions as an intracellular timekeeper coordinating biochemical activity has been proposed [3,69]. This ultradian (circadian) clock with  $\sim 40$  min period has been described in the yeast *Saccharomyces cerevisiae* growing aerobically in continuous culture [70,71,72]. Respiratory oscillations are (among many) a main physiological manifestation of the  $\sim 40$  min temperature-compensated periodicity (see [73] for a review). Genome-wide expression (transcriptome) during the time frame period provided by the ultradian clock revealed the existence of two blocks of redox superclusters [74]. A genome-wide activity of ca. 5300 transcripts expressed in two phases represented by two blocks of  $\sim 600$  and  $\sim 4700$  maximally expressed genes during active respiration (oxidative) and low respiration (reductive), respectively [27] (Fig. 5). Within the 40 min time frame, a 10–15 min period is conducive to DNA synthesis and replication, a time window that opens during the reductive phase of the clock cycle [74], suggesting that this might be an evolutionary strategy to avoid oxidative damage [27,75]. These gene expression clusters extend beyond central metabolism and mitochondrial energetics to other highly conserved cellular processes such as peroxisome function and repair and turnover of protein, RNA and DNA synthesis [27,75]. By separating into different

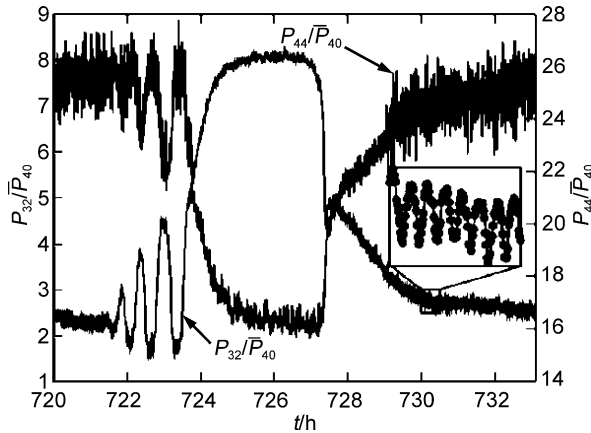


**Biochemistry, Chaotic Dynamics, Noise, and Fractal Space in, Figure 5**

The transcriptome of *Saccharomyces cerevisiae* (yeast) in continuous culture: activity with respect to the respiratory oscillations. **a** Dissolved O<sub>2</sub> in the culture, acetaldehyde and H<sub>2</sub>S. **b** Transcripts in phase with the oxidative phase of the culture (red), and those in phase with the reductive phase (green). The scale for the transcripts was obtained by dividing the intensity of expression of that gene at that point by the median intensity for the 32 points obtained over the duration of the experiment. The genome-wide transcriptional activity during respiratory oscillations has shown that 5329 transcripts are expressed and, of these, two blocks of redox superclusters could be assigned: (i) 650 that give maximal expression during the phase of active respiration (the oxidative phase of the cycle), and (ii) 4679 that are maximally observed during the "reductive" phase of the cycle (Reproduced from [27])

temporal windows contemporary incompatible biochemical reactions, the time keeping function of the ultradian clock uncovers a powerful biological rationale. Because of the sensitivity to reactive oxygen species and oxidative stress exhibited by biochemical processes such as DNA synthesis, their occurrence manifests at a time when oxygen uptake is minimal, and antioxidant species (e.g. reduced glutathione) are highest [76].

Recent contributions show that mitochondria oscillate as coupled oscillators either during respiratory oscillations in yeast [77] or in cardiac mitochondria [22]. It is now clear that a yeast population functions like a multicellular system, a network of coupled oscillators, through chemical communication by metabolites like acetaldehyde [78] or hydrogen sulfide, H<sub>2</sub>S [79]. The very recent experimental finding of a chaotic attractor of metabolic origin in continuous cultures of yeast could, potentially, give support to the idea that "controlled chaos" may be at the origin of ultradian (even circadian) rhythms [73,80]. Several periods



**Biochemistry, Chaotic Dynamics, Noise, and Fractal Space in, Figure 6**

Self-organized multioscillatory behavior of yeast respiration in continuous culture. One oscillatory period of the cell cycle-dependent, circadian, and fast oscillations exhibited by yeast in chemostat cultures corresponding to the relative membrane-inlet mass spectrometry signals of the  $m/z$  32 and 44 components (i.e. relative  $O_2$  and  $CO_2$  signals, respectively) vs. time. The inset shows the individual data points for  $m/z > 1/4$  32 ( $O_2$ ) for a 30-min span (Reproduced from [80])

of 13 h, 40 min and 4 min are embedded in the strange attractor of yeast metabolism and likely more may be found with greater time resolution in data collection (Fig. 6).

### Chaos and Colored Noise

A number of investigators have shown that while it is true that a low-dimensional chaotic attractor will generate a fractal time series, so too will processes other than white noise exhibiting power-law spectra like colored noise [81]; see also [14]. These colored noises are clearly not necessarily related to any deterministic dynamical process. However, it has been demonstrated that colored noise may emerge from deterministic (oscillatory) dynamics in a mitochondrial network (Fig. 3c and see below) [22].

Random noises exhibit a well defined, monotonically decreasing function of the correlation dimension,  $\nu$ , and the spectral exponent  $\alpha$ ,  $\nu(\alpha)$  (see Fig. 3b and 3c where  $\beta = \alpha$ ). The correlation integral  $C(r)$ , that for a chaotic or strange attractor is based on the correlation method proposed by Grassberger and Procaccia (1983) (see also [14]), has the power law form:

$$C(r) \sim r^\nu \quad (2)$$

where  $\nu$ , the “correlation dimension”, is closely related to the fractal dimension,  $D$ , and the “information dimension” of the attractor [82]. The  $D$  of a stochastic trajectory generated by a colored noise process can be related to the slope

of the inverse power law spectrum,  $\alpha$ , through the following equation:

$$D = 2/(\alpha - 1). \quad (3)$$

Osborne and Provenzale (1988) found that a finite value for the correlation dimension  $\nu$  may be found even for non-deterministic, random signals. For  $\alpha = 1$  (corresponds to a white noise spectrum) they found that scaling is present, and by repeating this analysis for a number of values of  $\alpha$  they found a quantitative relation between  $\nu$  and  $\alpha$ . Additionally, systems dominated by some kind of stochastic process are thought to provide a non-saturating value of  $\nu$ .

These results have implications for experimental studies of deterministic chaos. Dynamic processes characterized by a finite and small  $D$  correspond to low dimensional strange attractors whereas a stochastic or random time series is believed to completely fill the available 3D space [26].

### Noise in Biochemical and Cellular Systems

The intrinsic autonomous dynamics of biochemical processes may generate noise through stochastic fluctuations of molecular events involved in cellular physiology and biochemistry [83]. These dynamics generate time series that are characterized as colored noise when exhibiting an inverse power law spectrum. For this class of processes there is a one-to-one correspondence between the power-law dependence of the spectrum and fractal behavior, i.e. the power law index is related to the fractal dimension [17,81]. Moreover, these fractal processes are scale free, i.e. they do not exhibit a characteristic time scale, and, unlike white or brown noise, they exhibit long-term memory. The latter is a property of the autonomous system's dynamics as opposed to the so-called “extrinsic noise” or short-term random fluctuations produced by the interaction of the system with the environment [17]. Although in principle, colored noises are not related to any deterministic dynamic process as in the case of “deterministic chaos” (see above), noise-like behavior may arise from deterministic oscillatory dynamics.

Despite the stochastic function of regulatory and signaling circuits within cells, most cellular events show a precise spatio-temporal regulation. What is the role of noise, then? Several cellular processes are either noise-driven or are influenced by noise such as ion channel gating [38] and gene expression among others [83]. However, intracellular signaling networks have the ability to filter noise that in the case of periodic signals may function as *frequency filters*. A filter that passes only low frequencies is



a *low-pass* filter whereas one that passes only high frequencies is a *high-pass* filter [84]. In terms of signal processing, cascades and relays such as two-component systems and the mitogen-activated protein kinase pathway function as low-pass filters as they transduce low-frequency signals whereas high frequencies are attenuated [83]. A chemical low-pass filter can at low frequencies amplify a signal [84]. Considering nonlinear biochemical reactions, band-pass filtering is also possible. This is based on *integral feedback*, a form of negative feedback that uses an internal memory state to amplify intermediate frequencies. Bacterial chemotaxis uses integral feedback that by measuring temporal changes in chemical concentrations rather than steady state values, results in biased motion towards attractants [83].

In a cascade-type signaling pathway there seems to be a trade off between noise attenuation (arising from the input signal) and inherent noise at each step of the pathway (by molecular fluctuations). As a consequence, from the point of view of noise attenuation, there seems to be an optimal cascade length [85]. The biological ubiquitous negative feedback is the most common noise-attenuating device that in terms of signal processing functions as a low-pass filter. This has been experimentally demonstrated using the constitutive expression of GFP that is highly variable in *E. coli*. The addition of a negative feedback using the tetracycline repressor significantly reduced the measured variability [86]. Other noise-attenuating mechanisms are gene dosage, parallel cascades and regulatory checkpoints that either increase the likelihood of gene expression, establish the consensus from multiple signals, or ensure that each step in a pathway is successfully completed before proceeding to the next step.

There are mechanisms that amplify or exploit noise as well. Among these we count mechanisms that give rise to population heterogeneity or mechanisms that use noise to attenuate noise. Autocatalysis or positive feedback amplify the effects of noise giving rise to heterogeneity in a cell population as has been demonstrated in yeast using the tetracycline transactivator and a GFP marker [61]. Noise may actually enhance a signal when certain nonlinear effects are present as in stochastic resonance [87]. Concerning the latter, several biological examples have been reported [83].

Biochemical switches may be differentially sensitive to noise. Switches are based on ultrasensitive or bistable mechanisms depending upon their dynamic behavior based on a series of continuous or discontinuous steady values as a function of a control parameter [48]. Bistable mechanisms that may function in futile cycles of phosphorylation-dephosphorylation cascades, exhibit

hysteresis or “memory” effects since depending on the direction of change in the control parameter (i.e. increasing or decreasing) two steady state values are possible for the same value of the control parameter [2]. Ultrasensitivity is a type of signal amplification by which the percentage change in the response of a biological system increases several-fold in a narrow variation range of the stimulus [48,49,88,89]. One main advantage of ultrasensitivity is to allow increases in flux by several-fold after fluctuations in substrate or effector concentration slightly change over the background. Bistable mechanisms are more robust to noisy signals as compared with ultrasensitive ones that in turn may exploit noise such as in slightly changing illumination conditions in cyanobacteria. These relatively small fluctuations over the background allow the triggering of the mechanism of light harvesting and polysaccharide storage even under poor stimulus conditions and quickly deactivating it hindering ATP wastage [50].

Several modeling approaches are available to address the problem of how cells control, exploit and tolerate noise. Among them are the Langevin equation and the continuous or discrete descriptions by the Fokker–Planck or master equations, respectively [83].

## Noise and Network Behavior

Complex networks involve multiple feedback loops and the coexistence of mechanisms of noise attenuation, amplification and exploitation. A network keeping a reliable functionality is robust if tolerating a wide range of parametric variation and inherent noise due to stochastic molecular fluctuations. A complex network such as the chemotaxis pathway in *E. coli* is robust since it keeps its functionality despite a wide range of variation in enzymatic activities and protein concentrations [19,90,91]. The intrinsic stochasticity of biochemical reactions can be bypassed then bestowing noise resistance when perturbations affect variables that are not directly responsible for the dynamics or when molecular species are present in low numbers for short periods of time, as shown in a simple model of circadian rhythms [92].

Timing exerted by oscillatory mechanisms are found throughout the biological world and their periods span a wide range from milliseconds, as in the action potential of neurons and the heart, to the slow evolutionary changes that require thousands of generations [7,93]. Rhythms with temperature compensation are biological clocks. In virtually all light-sensitive organisms from cyanobacteria to humans, biological clocks adapt cyclic physiology to geophysical time with time-keeping properties in the circadian (24 h), ultradian (< 24 h) and infradian (> 24 h)



domains [94]. Within this framework, it becomes clear why it is important to understand the synchronization of a population of coupled oscillators as an important problem for the dynamics of physiology in living systems (see Sect. “Chaos in Biochemistry and Physiology”). Through frequency and amplitude modulation, oscillatory dynamics may function as a temporal-encoding signaling mechanism. Recent experimental findings are in agreement with the hypothesis that mitochondrial oscillatory behavior may function as a frequency- and/or amplitude-encoded signaling mechanism under physiological conditions and over a wide range of time scales [22,95]. This indicates that each mitochondrion affects the temporal behavior of its neighbors such that the collective ensemble of mitochondria in a cardiac cell acts as a network. The colored noise dynamics exhibited by mitochondria could be simulated from deterministic (oscillatory) dynamics by mixing a certain number of frequencies (Fig. 3c). This colored noise does not necessarily belong to chaotic dynamics.

### Future Directions

Is it possible to have a theory of organized complexity in biology? The apparent problem in formulating such a theory is that *contingency* and *specificity* are attributes of biological systems. These two features emphasize the importance of detailed observations in apparent contradiction with the theoretical aim of condensing specific system details into a small number of mathematical equations, similar to the fundamental laws of physics. What makes the situation different for biology as compared to physics is that the details are important because the outcome (sometimes, life or death) is important. Thus, we are aimed at finding universal behavior taking into account most of the biologically relevant details [35].

The existence of self-organization, scaling, criticality, percolation, fractals in the cardiac mitochondrial network show that there exists a clear crossroad between the universality of physical concepts and a defined biological system such as the heart [22,30]. The cardiac mitochondrial network exhibits the transition from physiological to pathophysiological behavior with all the characteristics of systems at critical state, i.e. a mitochondrial percolation cluster attains criticality at a certain threshold level of ROS. The transition that happens at the percolation threshold is self-organized and occurs with all the traits of universality, i.e. similar critical exponents as predicted by percolation theory, and the mitochondrial cluster exhibiting fractal organization. Moreover, the time series of mitochondrial membrane potential (a crucial cellular variable) exhibit fractal organization and long-term memory. Inverse

power laws statistically describe this behavior, and, in this case, the mechanisms underlying its origin are known.

Another major breakthrough comes from respiratory oscillations in yeast and their relationship to an ultradian clock. The genome-wide activity of yeast during respiratory oscillations [27,74] has now been linked to a chaotic attractor embedding multiple oscillatory periods [80].

The multiple time scales exhibited by the heart and yeast systems suggest that intracellular network dynamic organization manifests itself as scale-free in the form of dynamic fractals. Scale-free architecture has been shown in network connectivity but not in dynamics. The scale free behavior exhibited by mitochondrial network dynamics would allow simultaneous modulation of intracellular timekeeping in several time scales [18]. Two main features of mitochondria as metabolic “hubs” and as producers of ROS as signaling molecules with scale-free dynamics are in agreement with: (i) the cardiac mitochondrial network having both properties of constancy and flexibility, i.e. providing a steady supply of ATP to fuel contraction, and adapting the rate of energy production to meet changing metabolic demand [96,97]; and (ii) achieving the right ROS balance both in amount and dynamics, compatible with intracellular signaling.

We witness exciting times since several fundamental discoveries done in unrelated areas of research start to reveal that scaling and critical states are fundamental in normal physiology and pathology at the (sub)cellular, tissue and organ levels. Under steady conditions, biochemical and physiological functions exhibit a delicate balance that under stress unleashes as crises. When stressed, systems of disparate nature (e.g. markets, shuttles, cells) become extremely sensitive to small changes or perturbations exhibiting an exceptional large susceptibility to external factors and strong correlation between different parts of the system. Overall, the collective, repetitive actions of interactive nonlinear influences on many scales lead to large-scale correlations, determining the astonishing autonomy and interdependence of the living.

### Bibliography

#### Primary Literature

1. Aon MA, Cortassa S (2006) Metabolic dynamics in cells viewed as multilayered, distributed, mass-energy-information networks. In: Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics. Wiley Interscience, New York
2. Aon MA, Cortassa S (1997) Dynamic biological organization. Fundamentals as applied to cellular systems. Chapman & Hall, London
3. Lloyd D (1992) Intracellular time keeping: epigenetic oscillations reveal the functions of an ultradian clock. In: Lloyd D,

- Rossi EL (eds) *Ultradian rhythms in life processes*. Springer, London, pp 5–22
4. Nicolis G, Prigogine I (1977) *Self-organization in nonequilibrium systems: from dissipative structures to order through fluctuations*. Wiley, New York
5. Sornette D (2000) *Critical Phenomena in Natural Sciences. Chaos, fractals, selforganization and disorder: Concepts and tools*. Springer, Berlin
6. Boccaro N (2004) *Modeling Complex Systems*. Springer, New York
7. Lloyd D, Aon MA, Cortassa S (2001) Why homeodynamics, not homeostasis? *Sci World J* 1:133–145
8. Mainzer K (2005) *Symmetry and Complexity. The spirit and beauty of nonlinear science*. World Scientific, Singapore
9. Mandelbrot BB (1983) *The fractal geometry of nature*. WH Freeman, New York
10. Schroeder M (1991) *Fractals, chaos, power laws. Minutes from an infinite paradise*. WH Freeman, New York
11. Lorenz EN (1963) Deterministic nonperiodic flow. *J Atmos Sci* 20:130–141
12. Gleick J (1988) *Chaos: making a new science*. Penguin Books, New York
13. Wilson KG (1979) Problems in physics with many scales of length. *Sci Am* 241:158–179
14. West BJ (1990) *Fractal Physiology and Chaos in Medicine*. World Scientific, Singapore
15. Aon MA, Cortassa S (1993) An allometric interpretation of the spatio-temporal organization of molecular and cellular processes. *Mol Cell Biochem* 120:1–13
16. Aon MA, O'Rourke B, Cortassa S (2004) The fractal architecture of cytoplasmic organization: scaling, kinetics and emergence in metabolic networks. *Mol Cell Biochem*:256–257:169–184
17. West BJ (1999) *Physiology, promiscuity and prophecy at The Millennium: A tale of tails*. World Scientific, Singapore
18. Aon MA, Cortassa S, O'Rourke B (2007) On the network properties of mitochondria. In: Saks V (ed) *Molecular system bioenergetics: energy for life*. Wiley, Weinheim, pp 111–135
19. Alon U (2007) *An introduction to systems biology. Design principles of biological circuits*. Chapman & Hall/CRC, London
20. Barabasi AL, Oltvai ZN (2004) Network biology: understanding the cell's functional organization. *Nat Rev Genet* 5:101–113
21. Wagner A, Fell DA (2001) The small world inside large metabolic networks. *Proc Biol Sci* 268:1803–1810
22. Aon MA, Cortassa S, O'Rourke B (2006) The fundamental organization of cardiac mitochondria as a network of coupled oscillators. *Biophys J* 91:4317–4327
23. Aon MA, Cortassa S, O'Rourke B (2008) Mitochondrial oscillations in physiology and pathophysiology. *Adv Exp Med Biol* 641:98–117
24. Enquist BJ, West GB, Brown JH (2000) Quarter-power allometric scaling in vascular plants: Functional basis and ecological consequences. Oxford University Press, New York
25. Brown JH, West GB, Enquist BJ (2000) *Scaling in biology*. Oxford University Press, New York
26. Bassingthwaite JB, Liebovitch LS, West BJ (1994) *Fractal Physiology*. Oxford University Press for the American Physiological Society, New York
27. Lloyd D, Murray DB (2005) Ultradian metronome: timekeeper for orchestration of cellular coherence. *Trends Biochem Sci* 30:373–377
28. Cvitanovic P (1989) *Universality in chaos*. Adam Hilger, Bristol
29. Feigenbaum MJ (1980) Universal behavior in nonlinear systems. *Los Alamos Sci* 1:4–27
30. Aon MA, Cortassa S, O'Rourke B (2004) Percolation and criticality in a mitochondrial network. *Proc Natl Acad Sci USA* 101:4447–4452
31. Akar FG, Aon MA, Tomaselli GF, O'Rourke B (2005) The mitochondrial origin of postischemic arrhythmias. *J Clin Invest* 115:3527–3535
32. Varela F, Maturana H, Uribe R (1974) Autopoiesis: the organization of living systems, its characterization and a model. *Biosystems* 5:187–196
33. Schneider ED, Sagan D (2005) *Into the cool. Energy flow, thermodynamics and life*. The University of Chicago Press, Chicago
34. Schrodinger E (1944) *What is life? The physical aspects of the living cell*. Cambridge University Press, Cambridge
35. Vicsek T (2001) *Fluctuations and Scaling in Biology*. Oxford University Press, New York
36. Mandelbrot BB, Hudson RL (2004) *The (mis)behavior of market. A fractal view of risk, ruin, and reward*. Basic Books, New York
37. Smith TGJ, Lange GD (1996) *Fractal studies of neuronal and glial cellular morphology*. CRC Press, Boca Raton
38. Liebovitch LS, Todorov AT (1996) Using fractals and nonlinear dynamics to determine the physical properties of ion channel proteins. *Crit Rev Neurobiol* 10:169–187
39. Savageau MA (1995) Michaelis-Menten mechanism reconsidered: implications of fractal kinetics. *J Theor Biol* 176:115–124
40. Schnell S, Turner TE (2004) Reaction kinetics in intracellular environments with macromolecular crowding: simulations and rate laws. *Prog Biophys Mol Biol* 85:235–260
41. Goldberger AL, Amaral LA, Hausdorff JM, Ivanov P, Peng CK, Stanley HE (2002) Fractal dynamics in physiology: alterations with disease and aging. *Proc Natl Acad Sci USA* 99 Suppl 1:2466–2472
42. Ivanov PC, Amaral LA, Goldberger AL, Havlin S, Rosenblum MG, Struzik ZR, Stanley HE (1999) Multifractality in human heart-beat dynamics. *Nature* 399:461–465
43. Aon MA, Cortassa S, Akar FG, O'Rourke B (2006) Mitochondrial criticality: a new concept at the turning point of life or death. *Biochim Biophys Acta* 1762:232–240
44. O'Rourke B, Cortassa S, Aon MA (2005) Mitochondrial ion channels: gatekeepers of life and death. *Physiol (Bethesda)* 20:303–315
45. Rabouille C, Cortassa S, Aon MA (1992) Fractal organisation in biological macromolecular lattices. *J Biomol Struct Dyn* 9:1013–1024
46. Kopelman R (1988) Fractal reaction kinetics. *Science* 241:1620–1626
47. Berry H (2002) Monte carlo simulations of enzyme reactions in two dimensions: fractal kinetics and spatial segregation. *Biophys J* 83:1891–1901
48. Aon MA, Gomez-Casati DF, Iglesias AA, Cortassa S (2001) Ultrasensitivity in (supra)molecularly organized and crowded environments. *Cell Biol Int* 25:1091–1099
49. Gomez-Casati DF, Aon MA, Iglesias AA (1999) Ultrasensitive glycogen synthesis in Cyanobacteria. *FEBS Lett* 446:117–121
50. Gomez-Casati DF, Cortassa S, Aon MA, Iglesias AA (2003) Ultrasensitive behavior in the synthesis of storage polysaccharides in cyanobacteria. *Planta* 216:969–975
51. Goldberger AL, West BJ (1987) Applications of nonlinear dynamics to clinical cardiology. *Ann N Y Acad Sci* 504:195–213

52. Ivanov PC, Nunes Amaral LA, Goldberger AL, Havlin S, Rosenblum MG, Stanley HE, Struzik ZR (2001) From 1/f noise to multifractal cascades in heartbeat dynamics. *Chaos* 11:641–652
53. Peng CK, Mietus J, Hausdorff JM, Havlin S, Stanley HE, Goldberger AL (1993) Long-range anticorrelations and non-gaussian behavior of the heartbeat. *Phys Rev Lett* 70:1343–1346
54. Goldberger AL, Bhargava V, West BJ, Mandell AJ (1985) On a mechanism of cardiac electrical stability. The fractal hypothesis. *Biophys J* 48:525–528
55. Casolo G, Balli E, Taddei T, Amuhasi J, Gori C (1989) Decreased spontaneous heart rate variability in congestive heart failure. *Am J Cardiol* 64:1162–1167
56. Ewing DJ (1991) Heart rate variability: an important new risk factor in patients following myocardial infarction. *Clin Cardiol* 14:683–685
57. Skinner JE, Pratt CM, Vybiral T (1993) A reduction in the correlation dimension of heartbeat intervals precedes imminent ventricular fibrillation in human subjects. *Am Heart J* 125:731–743
58. Yamamoto Y, Hughson RL (1994) On the fractal nature of heart rate variability in humans: effects of data length and beta-adrenergic blockade. *Am J Physiol* 266:R40–49
59. Aon MA, Cortassa S, Marban E, O'Rourke B (2003) Synchronized whole cell oscillations in mitochondrial metabolism triggered by a local release of reactive oxygen species in cardiac myocytes. *J Biol Chem* 278:44735–44744
60. O'Rourke B, Ramza BM, Marban E (1994) Oscillations of membrane current and excitability driven by metabolic oscillations in heart cells. *Science* 265:962–966
61. Becskei A, Seraphin B, Serrano L (2001) Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *Embo J* 20:2528–2535
62. van Beek JHGM, Roger SA, Bassingthwaighite JB (1989) Regional myocardial flow heterogeneity explained with fractal networks. *Am J Physiol* 257:H1670–H1680
63. Garfinkel AM, Spano L, Ditto WL, Weiss JN (1992) Controlling cardiac chaos. *Science* 257:1230–1235
64. Shinbrot T, Ditto W, Grebogi C, Ott E, Spano M, Yorke JA (1992) Using the sensitive dependence of chaos (the “butterfly effect”) to direct trajectories in an experimental chaotic system. *Phys Rev Lett* 68:2863–2866
65. Meucci R, Labate A, Ciofini C (1998) Experimental techniques for controlling chaos in lasers. *Int J Bifurcat Chaos* 8:1759–1768
66. Moss F (1994) Neurobiology: Chaos under control. *Nature* 370:596–597
67. Garfinkel A, Chen PS, Walter DO, Karagueuzian HS, Kogan B, Evans SJ, Karpoukhin M, Hwang C, Uchida T, Gotoh M, Nwasonkwa O, Sager P, Weiss JN (1997) Quasiperiodicity and chaos in cardiac fibrillation. *J Clin Invest* 99:305–314
68. Verrier RL, Nearing BD, Lovett EG (1997) Complex oscillatory heart rhythm: a dance macabre. *J Clin Invest* 99:156–157
69. Lloyd AL, Lloyd D (1993) Hypothesis: the central oscillator of the circadian clock is a controlled chaotic attractor. *Biosystems* 29:77–85
70. Keulers M, Suzuki T, Satroutdinov AD, Kuriyama H (1996) Autonomous metabolic oscillation in continuous culture of *Saccharomyces cerevisiae* grown on ethanol. *FEMS Microbiol Lett* 142:253–258
71. Lloyd D, Salgado LE, Turner MP, Suller MT, Murray D (2002) Cycles of mitochondrial energization driven by the ultradian clock in a continuous culture of *Saccharomyces cerevisiae*. *Microbiology* 148:3715–3724
72. Satroutdinov AD, Kuriyama H, Kobayashi H (1992) Oscillatory metabolism of *Saccharomyces cerevisiae* in continuous culture. *FEMS Microbiol Lett* 77:261–267
73. Lloyd D (2008) Respiratory oscillations in yeast. In: Maroto M, Monk N (eds) *Cellular oscillatory mechanisms*. Landes Bioscience (in press)
74. Klevecz RR, Bolen J, Forrest G, Murray DB (2004) A genomewide oscillation in transcription gates DNA replication and cell cycle. *Proc Natl Acad Sci USA* 101:1200–1205
75. Lloyd D, Murray DB (2006) The temporal architecture of eukaryotic growth. *FEBS Lett* 580:2830–2835
76. Tu BP, Kudlicki A, Rowicka M, McKnight SL (2005) Logic of the yeast metabolic cycle: temporal compartmentalization of cellular processes. *Science* 310:1152–1158
77. Aon MA, Cortassa S, Lemar KM, Hayes AJ, Lloyd D (2007) Single and cell population respiratory oscillations in yeast: a 2-photon scanning laser microscopy study. *FEBS Lett* 581:8–14
78. Murray DB, Klevecz RR, Lloyd D (2003) Generation and maintenance of synchrony in *Saccharomyces cerevisiae* continuous culture. *Exp Cell Res* 287:10–15
79. Lloyd D (2006) Hydrogen sulfide: clandestine microbial messenger? *Trends Microbiol* 14:456–462
80. Roussel MR, Lloyd D (2007) Observation of a chaotic multioscillatory metabolic attractor by real-time monitoring of a yeast continuous culture. *FEBS J* 274:1011–1018
81. Osborne AR, Provenzale A (1989) Finite correlation dimension for stochastic systems with power-law spectra. *Physica D* 35:357–381
82. Grassberger P, Procaccia I (1983) Measuring the strangeness of strange attractors. *Physica D* 9:189–208
83. Rao CV, Wolf DM, Arkin AP (2002) Control, exploitation and tolerance of intracellular noise. *Nature* 420:231–237
84. Arkin AP (2000) *Signal processing by biochemical reaction networks*. Cambridge University Press, Cambridge
85. Thattai M, van Oudenaarden A (2002) Attenuation of noise in ultrasensitive signaling cascades. *Biophys J* 82:2943–2950
86. Becskei A, Serrano L (2000) Engineering stability in gene networks by autoregulation. *Nature* 405:590–593
87. Kosko B (2006) *Noise*. Viking, New York
88. Goldbeter A, Koshland DE Jr (1982) Sensitivity amplification in biochemical systems. *Q Rev Biophys* 15:555–591
89. Koshland DE Jr, Goldbeter A, Stock JB (1982) Amplification and adaptation in regulatory and sensory systems. *Science* 217:220–225
90. Alon U, Surette MG, Barkai N, Leibler S (1999) Robustness in bacterial chemotaxis. *Nature* 397:168–171
91. Barkai N, Leibler S (1997) Robustness in simple biochemical networks. *Nature* 387:913–917
92. Vilar JM, Kueh HY, Barkai N, Leibler S (2002) Mechanisms of noise-resistance in genetic oscillators. *Proc Natl Acad Sci USA* 99:5988–5992
93. Lloyd D (1998) Circadian and ultradian clock-controlled rhythms in unicellular microorganisms. *Adv Microb Physiol* 39:291–338
94. Schibler U, Naef F (2005) Cellular oscillators: rhythmic gene expression and metabolism. *Curr Opin Cell Biol* 17:223–229
95. Cortassa S, Aon MA, Winslow RL, O'Rourke B (2004) A mitochondrial oscillator dependent on reactive oxygen species. *Biophys J* 87:2060–2073

96. Cortassa SC, Aon MA, O'Rourke B, Jacques R, Tseng HJ, Marban E, Winslow RL (2006) A computational model integrating electrophysiology, contraction and mitochondrial bioenergetics in the ventricular myocyte. *Biophys J* 91:1564–1589
97. Saks V, Dzeja P, Schlattner U, Vendelin M, Terzic A, Wallimann T (2006) Cardiac system bioenergetics: metabolic basis of the Frank-Starling law. *J Physiol* 571:253–273

## Books and Reviews

- Bak P (1996) *How nature works: the science of self-organized criticality*. Copernicus, New York
- Barabasi AL (2003) *Linked*. Plume, New York
- Capra F (1996) *The web of life*. Anchor books Doubleday, New York
- Dewey GT (1997) *Fractals in molecular biophysics*. Oxford University Press, New York
- Cortassa S, Aon MA, Iglesias AA, Lloyd D (2002) *An Introduction to Metabolic and Cellular Engineering*. World Scientific, Singapore
- Feder J (1988) *Fractals*. Plenum Press, New York
- Glass L, Mackey MC (1988) *From clocks to chaos. The rhythms of life*. Princeton University Press, Princeton
- Haken H (1978) *Synergetics*. Springer, Heidelberg
- Jantsch E (1989) *The self-organizing universe. Scientific and human implications of the emerging paradigm of evolution*. Pergamon Press, Oxford
- Kauffman SA (1993) *Origins of order: Self-organization and selection in evolution*. Oxford University Press, New York
- Lane N (2005) *Power, sex, suicide. Mitochondria and the meaning of life*. Oxford University Press, Oxford
- Noble D (2006) *The music of life*. Oxford University Press, New York
- Sornette D (2003) *Why stock markets crash. Critical events in complex financial systems*. Princeton University Press, New Jersey
- Varela FJ (1989) *Autonomie et connaissance. Seuil*, Paris
- West BJ, Deering B (1995) *The Lure of Modern Science. Fractal Thinking*, vol 3. World Scientific, Singapore
- Yates EF (1987) *Self-organizing systems. The emergence of order*. Plenum Press, New York

## Biological Complexity and Biochemical Information

CHRISTOPH ADAMI

Keck Graduate Institute of Applied Life Sciences,  
State University of New York, Claremont, USA

## Article Outline

[Glossary](#)

[Definition of the Subject](#)

[Introduction](#)

[Measures of Biological Complexity](#)

[Biochemical Information](#)

[Network Complexity](#)

[Future Directions](#)

[Acknowledgments](#)

[Bibliography](#)

## Glossary

**C-value** The haploid genome size of an organism, measured either in picograms (pg) or base pairs (bp).

**Degree distribution** The probability distribution  $P(d)$  to find a node with  $d$  edges in a network.

**Entropic profile** A graph of the per-site entropy along the sites of a biomolecular sequence, such as a DNA, RNA, or protein sequence.

**Epistasis** Generally, an interaction between genes, where the fitness effect of the modification of one gene influences the fitness effect of the modification of another gene. More specifically, an interaction between *mutations* that can be either positive (reinforcing or *synergistic*), or negative (mitigating or *antagonistic*).

**Erdős–Rényi network** A random graph with a binomial degree distribution.

**Fitness** A numerical measure predicting the long-term success of a lineage.

**Jensen–Shannon divergence** In probability and statistics, a measure for the similarity of probability distributions, given by the symmetrized relative entropy of the distributions.

**Module** In network theory, a group of nodes that is closely associated in connections or function, but only weakly associated to other such groups.

**Motif** In network theory, a subgraph of small size.

**Network diameter** For networks, the average geodesic distance between nodes, defined as  $D = 1/m \sum_{i=1}^n \sum_{j=1}^n d(i, j)$ , where  $m$  is the number of edges of the graph,  $n$  is the number of nodes, and  $d(i, j)$  is the shortest path distance between nodes  $i$  and  $j$ .

**Phylogenetic depth** A measure of the genetic distance between a genome and its ancestor on the same line of descent, given by the number of genetically different genomes on the line between the genomes plus one.

**Random variable** In probability and statistics, a mathematical object with discrete or continuous states that the object takes on with probabilities drawn from a probability distribution associated to the random variable.

**Source entropy** The entropy of a sequence generated by a process that generates symbols with a given probability distribution.

**Wright–Fisher process** In population genetics, a stochastic process that describes how genes are transmitted from one generation to the next.

**Turing machine** In mathematics, an abstract automaton that manipulates symbols on a tape directed by a finite set of rules.

**Watson–Crick pairing** In biochemistry, the pairing be-



tween nucleotides adenine and thymine (A-T), and guanine and cytosine (G-C).

**Zipf's law** A relationship between the frequency  $f$  and the rank  $k$  of words in a text, of the form  $f(k) \sim k^s$ , where  $s$  is the exponent of the distribution.

### Definition of the Subject

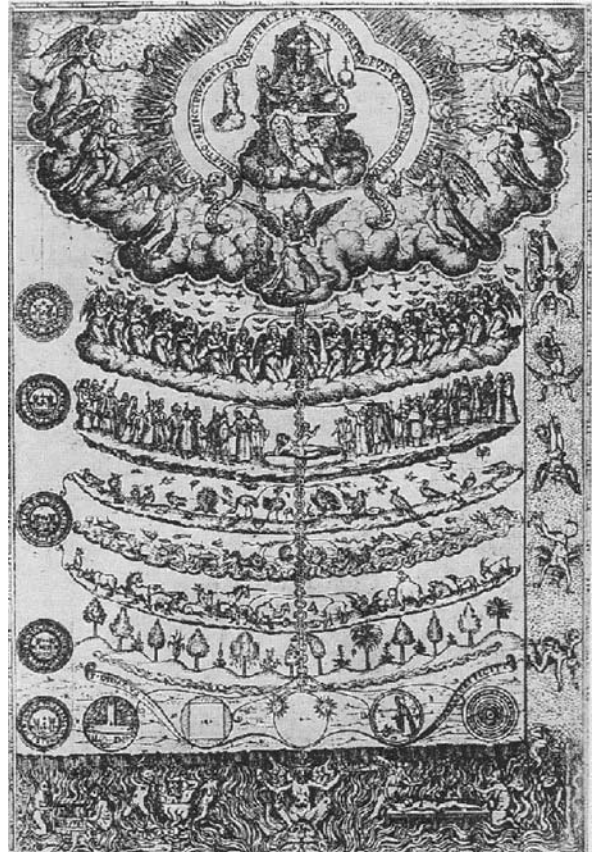
Biological complexity refers to a measure of the intricateness, or complication, of a biological organism that is directly related to that organism's ability to successfully function in a complex environment. Because organismal complexity is difficult to define, several different measures of complexity are often used as proxies for biological complexity, such as structural, functional, or sequence complexity. While the complexity of single proteins can be estimated using tools from information theory, a whole organism's biological complexity is reflected in its set of expressed proteins and its interactions, whereas the complexity of an ecosystem is summarized by the network of interacting species and their interaction with the environment.

### Introduction

Mankind's need to classify the world around him is perhaps nowhere more apparent than in our zeal to attach to each and every living organism a tag that reveals its relationship to ourselves. The idea that all forms of life (and even inanimate matter) belong to a "Great Chain of Being" goes back to medieval times [1] and usually depicts rocks at the bottom of the chain, with plants next, followed by fish, birds, mammals, humans, angels, and ultimately god (see Fig. 1).

In more modern times, biologists and lay people alike have, sometimes unconsciously, given in to the same need by classifying organisms according to their perceived complexity. So, for example, viruses are often perceived as the least complex organisms, followed by bacteria, unicellular eukaryotes, fungi, plants, invertebrates, vertebrates, mammals, and, ultimately, ourselves. Such a hierarchical view of biological complexity with *Homo sapiens* at its top has repeatedly been decried (see, e.g., [2] and recently [3]). However, in the absence of a quantitative measure of biological complexity, neither a position in favor nor against a hierarchical ordering of species can authoritatively be maintained. Indeed, even though it has become less and less accepted to view *Homo* as the "crown jewel" of evolution, there is no reason a priori why this cannot in fact be true.

Many different measures for biological complexity have been introduced in the past literature, some of them



**Biological Complexity and Biochemical Information, Figure 1**  
A depiction of the Great Chain of Being by Fra Diego de Valdes, in *Rhetorica Christiana*, from 1579

with obvious biological roots (such as number of cells, number of different tissue types, genome length, etc.), some of them inspired by the physics of dynamical systems. Neither of these measures has managed to convince a majority that it represents the best proxy for a biological complexity measure, and most have very obvious shortcomings. As we learn more about the genome and proteome of different organisms, it seems as if we are well on our way to accumulate a sufficient amount of detailed knowledge concerning the function of organisms that a universal measure of organism complexity is at last imaginable. For example, the set of all expressed proteins and their interactions (including their regulation and post-translational modification) could conceivably enter into an equation whose result is a number: the biological (functional) complexity of an organism that allows for comparisons across the tree of life.

Such a fantasy, however, is almost surely misguided unless we can enumerate along with the list of proteins and their interactions the possible environments within which



the organism can make a living. After all, the complexity of an organism must depend crucially on the environment within which it functions. Consider as an example a 150 pound rock and an average human, transplanted to the surface of the moon. Functionally, they are very comparable there. Likewise, a human is much less complex than even an algal bloom, if forced under water for an extended period of time. Given that we are unable to characterize even a single environment for most bacteria (environmental microbiologists estimate that fewer than 2% of all bacteria can be cultured in the laboratory) it is highly unlikely that the full functional complexity of most organisms can be ascertained.

Still, for organisms that make a living in comparable environments, a classification in terms of their biological complexity given their bioinformatic pedigree is not unreasonable, and the concepts and issues discussed below are steps toward such a goal. In the following section, I review standard types of complexity measures, how they compare to each other and how they fail to capture the essence of biological complexity. I then discuss a type of sequence complexity that can be shown to reduce to the information encoded into the biochemistry of DNA and proteins, and discuss its advantage over other measures, as well as its practical problems. I discuss complexity measures for biological and biochemical networks in the following section, and close with some thoughts on future directions.

## Measures of Biological Complexity

The different measures put forward in the literature to quantify an organism's complexity can be grouped into three main categories: structural, functional, and sequence complexities. A fourth category—network complexity—is very recent, and I will consider that last.

### Structural Complexity

Perhaps the most obvious way to define complexity is with reference to the structural complication of an organism. Clearly, structural complexity is not limited to biological organisms alone, and a good measure of this kind might allow us to compare biological organisms to human feats of engineering. Two problems are immediately apparent for any measure of structural complexity. First, to generate a scalar (that is, a single number) that will rank all possible structures appears to be a formidable task. Second, there is no guarantee that structural complexity is always a good predictor for an organism's success in the biosphere. On the one hand, something that looks like a complicated contraption could, in principle, be an evolution-

ary artifact: a non-adaptive feature that is either a necessity or a consequence of another feature (called a "spandrel" by Gould and Lewontin [4]). On the other hand, a complicated device could conceivably be functional only in a very different environment, perhaps one in which the organism can no longer survive. In other words, complex structure is not necessarily predictive of complex function, although we expect this to be true in the majority of cases.

The second difficulty, however, pales compared to the first. Commonly, a measure of structural complexity attempts to count the number of parts and their connection. Several such measures are reviewed by McShea [5]. A typical example for a structural complexity measure is the number of *different* cell types within an organism [6], perhaps normalized by the total number of cells in order to counter the bias of organism size. Bell and Mooers analyzed such a measure [7], and found that it robustly classifies animals as more complex than plants, and plants more complex than algae. However, the measure remains a very crude estimate of complexity, unable to shed light on finer gradations of the tree of life.

*Hierarchical* measures of complexity represent a different type of structural complexity. Generally speaking, a measure of hierarchical complexity seeks to quantify the number of levels needed to build a biological system, for example as the minimum amount of hierarchical structuring needed to build an understanding of the system [8]. The problem with a hierarchical scale of complexity for biological systems is that there are only four clear hierarchies: the prokaryote cell, the eukaryotic cell viewed as a symbiotic assembly of prokaryotic cells, the multicellular organism, and colonial individuals or integrated societies [9]. However, it is possible to introduce a higher-resolution scale by decomposing each hierarchy into levels and sublevels, for example by differentiating a monomorphic aggregate of elements of the lower hierarchy from a differentiated aggregate, and an aggregate of nested differentiated elements [9]. Even though a hierarchical measure of complexity necessarily represents a fairly coarse scale, it is one of only few measures of structural complexity that shows an unambiguous increase in complexity throughout the fossil record.

### Functional Complexity

Ideally, any measure of biological complexity should be functional, that is, reflecting how the organism functions in a complex world. Because of the obvious difficulty in relating either form to function, or sequence to function, function-based measures are currently even less well-formed than structural ones. Because function

is understood to be a product of natural selection, it is often implicitly assumed that a measure of functional complexity should increase in evolution. However, arguments advanced earlier clearly destroy this notion: it is certainly possible that following a drastic change in the environment the functional complexity of an organism decreases, for example because a preferred metabolic substrate is eliminated from the environment, abolishing the metabolic pathway—and thus the function associated with it—for the unlucky species. A functional complexity measure for single molecules that is based on information theory has recently been introduced by Szostak [10]. I will discuss this measure in more detail in Sect. “[Biochemical Information and Functional Complexity](#)”.

Attempts to count the number of different functions that an organism can perform are (and perhaps will remain) hopeless. This problem is directly related to our inability to characterize the necessary, or even sufficient, elements of an organism’s environment. McShea proposed to catalogue the different behaviors of an organism as proxy for its different functions [11], as behaviors are more amenable to observation than functions. In turn, he suggested that the number of *behavioral parts* of an organism could be used as a proxy for the number of behaviors, if it is true that parts usually play a role only in one function or behavior. He gives the example of the Swiss Army knife as a device more complex than a screwdriver, on account of the different parts that represent different functions. McShea’s analysis of functional parts leads him to consider networks of interacting parts that display varying degrees of modularity. I will return to such network-based measures further below.

### Sequence Complexity

Given that all forms of life on Earth contain a genetic code that is responsible for generating their form and function, we might naively assume that the amount of haploid DNA (measured either in picograms (pg) as was done before the advent of whole genome sequencing or in millions of base pairs (mbp) as is more common today) would reflect—even if only roughly—the complexity of the organism. This hope was quashed relatively early on: Britten and Davidson showed conclusively [12] that no correlation between genome size and perceived complexity exists. This disconnect has been termed the “C-value paradox” [13] (reviewed in [14]) and is perhaps best exemplified by the giant free living amoeba *Amoeba dubia*, whose total DNA content was estimated at 700 pg, which would correspond to about 675,000 mpb if it was all haploid. This would correspond to about 200 times the C-value of hu-

mans. However, the haploidy of the *A. dubia* genome is now in doubt [15]. The variation in genome size and the absence of a correlation to a complexity scale such as that given by the classical chain of being is depicted in Fig. 2.

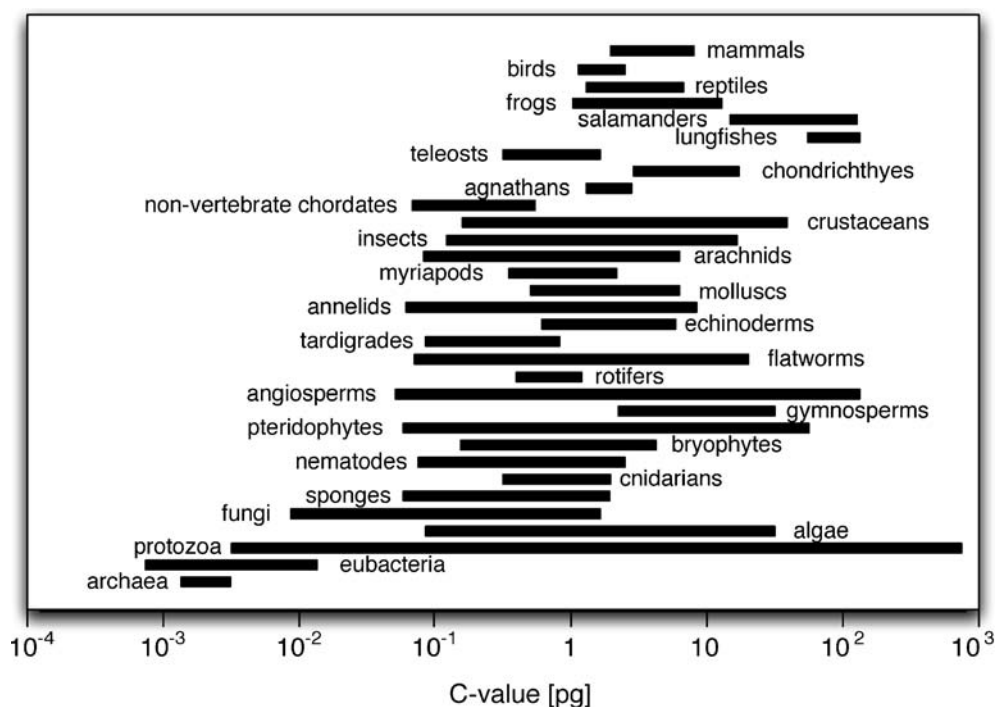
Because sequence length cannot be used, the focus for measures of sequence complexity instead has been on mathematical measures. The literature on mathematical sequence complexity, or more specifically, the complexity of symbolic strings, is far richer than that concerning functional or structural and morphological complexity. Among sequence complexities there are many different types, such as Kolmogorov, compositional, or information-theoretic ones. I shall give brief expositions of examples of each of these types here, without attempting to be even nearly exhaustive. For a good review of sequence complexities used in physics, see Ref. [16].

The most well-known sequence complexity is that introduced by the Russian mathematician Andrey Kolmogorov. He proposed to assign to each symbolic sequence a scalar that represents the *regularity* of the sequence [17]. So, for example, a string consisting of the repetition of a symbol or pattern is classified as regular, whereas a string with no discernable pattern would be irregular, and therefore complex. Note however that this algorithm does not classify a string as complex just because no pattern is readily identifiable. The Kolmogorov measure assumes that *all possible* computer programs (of a universal Turing machine) are tested so as to find the shortest one that produces the sequence in question. Mathematically, the Kolmogorov complexity of a string  $s$  is given by the length of the shortest program  $p$  (denoted as  $|p|$ ) that produces  $s$  when executed on a universal Turing machine  $T$ :

$$K(s) = \min \{|p| : s = C_T(p)\} , \quad (1)$$

where  $C_T(p)$  denotes the result of running program  $p$  on Turing machine  $T$ . So, for example, the binary equivalent of the irrational number  $\pi$  is random *prima facie*, however a concise algorithm (a short program  $p_\pi$ ) exists to produce it, leading to a fairly low complexity for  $s_\pi$ . There is a vast amount of mathematical and information-theoretical literature concerning the Kolmogorov complexity (see, e.g., [18]), but when applied to biological sequences, this measure has two obvious flaws.

First, the procedure to generate the Kolmogorov complexity is uncomputable because the search for the smallest program may never end (the computation may not halt). Second, truly random strings, that is, those that cannot be generated by any computation, are assigned maximal complexity. But in physics (and in particular in biology!) truly random sequences are meaningless, and there-



Biological Complexity and Biochemical Information, Figure 2

Ranges in haploid genome size measured in picograms (the C-value) in different organisms [14]

fore should not be assigned a large complexity. In fact, the Kolmogorov complexity is even logically inconsistent. Remember that the procedure to calculate the Kolmogorov complexity is one where an automaton returns the smallest program that produces the sequence in question. For a random sequence, the procedure is to return the sequence itself (and thus the complexity estimate of a random sequence is the length of that sequence, the largest possible result). But, logically, a random sequence can never be the result of a computation, because a computation is a deterministic process. These flaws can be partly fixed by focusing on conditional and mutual Kolmogorov complexities, as I outline further below. In summary, while the Kolmogorov complexity is a mathematically sound measure of *mathematical* sequence regularity, it is not a measure of the complexity of a sequence that describes a physical object.

The key concept to account for objects in the physical—rather than mathematical—world is the *conditional Kolmogorov complexity* (first introduced in [17]) of a string  $s$  given another string  $t$ , defined as the length of the smallest program  $|p|$  that produces  $s$  using that program and an external string  $t$ :

$$K(s|t) = \min \{|p| : s = C_T(p|t)\} , \quad (2)$$

with the notation  $(p|t)$  pronounced as “ $p$  given  $t$ ”. Now,  $t$  can be a sequence from the physical world, or describing an object in the physical world. The program  $p$  is small if the sequence  $s$  can be obtained from  $t$  using a simple computation. Using this construction, the conditional complexity of a random sequence  $r$  can be defined rigorously, because the shortest program to produce a random sequence  $r$  involves both the sequence  $r$  as input to the machine, and the vanishingly small (in the limit of infinite strings) program  $p$ =“print”. In other words, the conditional Kolmogorov complexity of random strings *vanishes* in the limit of long sequences, rather than being maximal.

To define a physical complexity of a symbolic string (physical, because it refers to a physical world, rather than to abstract mathematics), we can imagine a sequence  $e$  that represents everything that can be measured in that world. We now ask of our Turing machine to compute string  $s$  given everything that is knowable about the physical world, that is, given  $e$ . In that case, the conditional complexity  $K(s|e)$  represents everything that *cannot* be obtained from knowing the physical world. In other words, it represents the “remaining randomness”: the unmeasurable. Naturally, the physical complexity of the string is then just the length of the sequence *minus* the remaining

randomness [19]:

$$C_P(s) = K_0(s) - K(s|e), \quad (3)$$

where  $K_0(s)$ , the Kolmogorov complexity in the absence of both a “world” string  $e$  and the rules of mathematics, is just the length of  $s$ :  $K_0(s) = |s|$ . This notation is chosen in anticipation of results from information theory introduced in the following section. In particular, it allows us to see that  $C_P(s)$  is simply a *mutual Kolmogorov complexity* [17], between string  $s$  and the world  $e$ . Put in another way, the physical complexity of a sequence is that part of the sequence that can be obtained from the world string  $e$  using a concise (and therefore short in the limit of very long sequences) computation using program  $p$ . Only *those* sequences can be obtained by computation from the world  $e$  that *mean* something in world  $e$ , or refer to something there. Note that by its construction, the physical complexity represents a special case of the “effective complexity” measure introduced earlier by Gell-Mann and Lloyd [20].

As an example, consider a world of machines, whose blueprints are stored in such a way that they can be represented as sequences of symbols. In this world,  $e$  represents all the blueprints of all the possible machines that exist there. A sequence  $s$  is complex in this world if a part—or all—of the sequence can be obtained by manipulating, or translating, the world tape  $e$ . (It is conceivable that  $s$  contains part of a blueprint from  $e$  in encrypted form, in which case the program  $p$  must try to compare  $e$  to encrypted forms of  $s$ .) Of course, it would be an accident bordering on the unreasonable to find that a string that is complex in  $e$  is also complex mathematically. Instead, from a mathematical point of view, such a sequence most likely would be classified as random, or rather, the search for a shortest program would not halt. Similarly, it is extremely unlikely that sequence  $s$  would be classified as complex in a world in which  $e$  represents, say, all the literature produced on Earth (unless there are a few books on the blueprints of certain machines!). Thus, the complexity of a sequence  $s$ , by this construction, is never absolute (like in mathematics), but always conditional with respect to the world within which the sequence is to be *interpreted*.

This is precisely what we need in order to quantify the complexity of biological sequences, because it is immediately clear that a biological sequence only means something in a very specific environment, given very specific rules of chemistry. So, according to this argument, the sequence describing a particular organism is complex only with respect to the environment within which that organism “makes its living”, that is, its niche. Take the organism out of its niche, and it is unlikely to function as well as in its native niche; some of its structural complexity may turn

into a useless appendage or, worse, a liability. In the following section, we will see under what circumstances this physical complexity can be understood in terms of information theory.

## Biochemical Information

The inception of information theory [21] created the widely accepted expectation that biology would ultimately become a subdiscipline of cybernetics, information theory, and control theory [22]. That this expectation has not come to pass lies partly in a gross underestimate of the complexity of molecular and cellular biology by engineers and physicists, and partly in a fundamentally misguided application of information theory to genomes and molecular sequences in general. Most of the early applications of information theory to biomolecular sequences focused on an estimation of the entropy of sequences [23], rather than the information content. For example, a significant amount of work was expended on estimating the compressibility of DNA and protein sequences by studying long range correlations [24,25,26], all the while mistaking the average per-site entropy for information content [27]. These studies concluded, for the most part, that coding DNA sequences and proteins are *essentially random*, or uncompressible (non-coding DNA was found to be less random, due to repeats). From the point of view of coding theory, such a finding should not have been surprising, as Shannon’s coding theorem [28] implies that the length of a message conveyed by a symbolic sequence is limited by the per-site (source) entropy [29]. In other words, we expect evolution to try to maximize the source entropy. However, source entropy does not equal information content. Let us briefly review the basics of information theory as applied to molecular sequences [30].

## Entropy

For a random variable  $X$  that can take on states  $x_1, \dots, x_D$  with probabilities  $p(x_i)$ , the entropy of  $X$  is given by

$$H(X) = - \sum_i^D p(x_i) \log p(x_i), \quad (4)$$

where the logarithm is taken to a convenient base, and determines the units of the entropy. If the base is chosen to be 2, for example, the entropy is given in bits. Often, it is convenient to take the number of possible states  $D$  as the basis. In that case, the entropy is bounded from above by 1, and a sequence of  $N$  such random variables (a polymer) has a maximal entropy of  $N$ .

Suppose we are given a DNA sequence

$$\begin{array}{c} \text{AGAGCGCCTGCTTTGCACGCAGGAGGTCTGCGGTTCTGA} \\ \text{TCCCGCATAGCTCCACCA} \end{array} \quad (5)$$

We can either take the stance that this is one of  $4^{56}$  possible DNA sequences of length  $N = 56$ , or we can imagine that this is the record of 56 consecutive independent trials of a single DNA random variable that can take on the states  $x = A, T, G, C$  only. If we take the latter view, then we can try to estimate the *source entropy* by using the results of the 56 measurements to calculate the *plug-in probabilities*:

$$\begin{aligned} p_A &\approx \frac{10}{56} \approx 0.18, & p_T &\approx \frac{11}{56} \approx 0.20, \\ p_G &\approx \frac{16}{56} \approx 0.29, & p_C &\approx \frac{19}{56} \approx 0.34. \end{aligned} \quad (6)$$

These probabilities are quite uncertain due to the lack of statistics, and only give us limited knowledge about the nature of the gene. The deviation of the calculated probabilities from the unbiased prediction  $p_i = 1/4$  (an independent identically distributed, or i.i.d., random variable) is not significant. Indeed, using the four probabilities above, we obtain as an estimate of the source entropy

$$H(X) = - \sum_{i=A,C,G,T} p(x_i) \log_4 p(x_i) \approx 0.978, \quad (7)$$

which is almost maximal and confirms that, whatever message is encoded in the sequence, it is encoded in a nearly maximally compressed manner. For longer sequences, an analysis of the usage probabilities can reveal important information such as a bias in nucleotide usage, which will reduce the source entropy.

A typical application of sequence entropy is the calculation of “n-gram” entropies for blocks of sequences of length  $n$ :

$$H_n = - \sum_i^{4^n} P_i^{(n)} \log P_i^{(n)}, \quad (8)$$

where  $P_i^{(n)}$  represents the probability to find the  $i$ th of the  $4^n$  blocks of length  $n$  in the sequence.  $H_n$  then measures the average uncertainty within an average n-gram. We can write down the difference [31]

$$h_n = H_{n+1} - H_n, \quad (9)$$

which is sometimes called the amount of information necessary to predict the next symbol on a sequence

given the  $n$  previous symbols. For example, for the sequence Eq. (5),  $H_2 = 1.877$  (based on 55 2-grams) so that  $h_1 = H_2 - H_1 = 0.899$ , using the per-site entropy Eq. (7) for  $H_1$ . Note, however, that the difference Eq. (9) is, strictly speaking, not a measure of information but rather the conditional entropy (see below) to find a symbol  $A$  given the sequence  $S_n$  of length  $n$ :  $h_n = H(A|S_n)$ .

Using the entropy per letter  $h = \lim_{n \rightarrow \infty} h_n$ , we can calculate Grassberger’s *effective measure complexity* [31]

$$EMC = \sum_{n=0}^{\infty} (h_n - h). \quad (10)$$

This measure sums up the “memory” effects within the string, that is, it sums up the correlations at all scales. This measure vanishes if all sites are independent (because then  $H_n = nh$  and  $h_n = h$ ) and is maximal if there are strong long-range correlations within a sequence. The utility of this measure to capture the complexity of a gene is much in doubt, however, because the ability to predict the next symbol on a sequence, as discussed before, is unlikely to shed light on the function and utility of the gene if the information contained in the sequence is *encoded*.

The *compositional complexity* is another usage of n-gram entropies to capture sequence complexity, but this measure attempts to find the optimal segmentation of the sequence into  $m$  partitions based on a significance criterion. The more heterogeneous a sequence is, the higher the compositional complexity  $C_{\text{comp}}$  of a sequence  $s$  with length  $L$  [32]

$$C_{\text{comp}} = \max_{\pi} J(s_m), \quad (11)$$

which is the maximum over all possible partitions  $\pi$  of the Jensen–Shannon divergence

$$J(s_m) = H(s) - \sum_{i=1}^m \frac{l_i}{L} H(s_i), \quad (12)$$

where  $H(s)$  is again the source entropy, and  $H(s_i)$  is the entropy of the  $i$ th segment of length  $l_i$ . As with most of the sequence measures discussed up to now, this measure addresses coding-style more than function, and is unlikely to capture the essence of functional information.

### Conditional Entropy

Before we can talk about information, we have to introduce the concept of *conditional probabilities*, that is, the probability  $p(x_i|y_j)$  that a random variable  $X$  takes on one of its states  $x_i$  given that another variable  $Y$  is in one of its



states  $y_j$ . Then, the conditional entropy of  $X$  given that  $Y$  is in state  $y_j$  can be written as

$$H(X|Y = y_j) = - \sum_i^D p(x_i|y_j) \log p(x_i|y_j). \quad (13)$$

This concept allows us to discover the *relative state* of two different random variables, or in this case two different nucleotide positions. Suppose we are given 32 more sequences like the one in Eq. (5), and suppose we are told that these sequences represent the *same information*, that is, same gene, but from different samples. After aligning these 33 sequences, the result looks like Table 1.

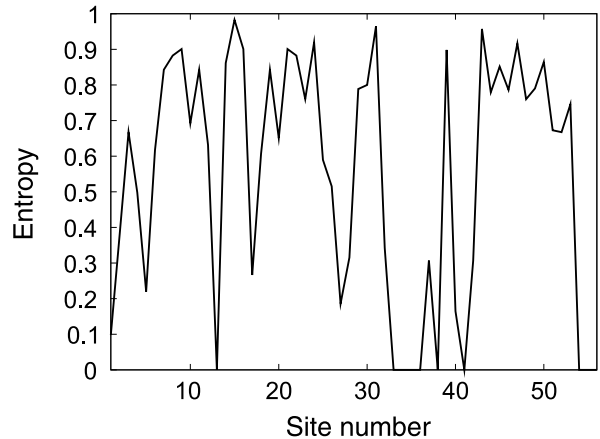
The extra symbols ‘-’ and ‘.’ in this alignment need further explanation. The symbol ‘-’ denotes a “missing” symbol, that is, a deletion in that sequence with respect to the other sequences. The symbol ‘.’, on the other hand, signals that the identity of the nucleotide at that position is not known. Here, we will treat each of these symbols as if there was no data for the random variable at that position, reducing the statistics at that site. Clearly, we now are viewing the set of sequences in Table 1 as representing independent samples of a joint random variable

$$Z = X_1 X_2 \cdots X_{56} \quad (14)$$

Using the conditional entropy, we can now study *correlations* between nucleotide positions. For example, we can study position 8, or more precisely, the random variable  $X_8$ . We can gather a probability estimate for this position by gathering statistics *vertically*, rather than horizontally as before. Here, we find

$$\begin{aligned} p_A(8) &= 5/33, & p_C(8) &= 17/33, \\ p_G(8) &= 5/33, & p_T(8) &= 6/33. \end{aligned} \quad (15)$$

which, even though 33 sequences does not represent a large sample, are significantly different from the random assumption  $p_i = 1/4$ . Even more important deviations from the i.i.d case can be found for  $X_{33}$  and  $X_{36}$ : only one particular nucleotide can be found at these positions, all others are absent. The reasons for this *abnormal* distribution is immediately clear when we reveal the origin of these sequences: they are the last 56 nucleotides of the tRNA gene of the *E. coli* bacterium [33], and the function of this molecule requires positions 33 and 36 to have that particular nucleotide. A mutation leading to an alternative nucleotide at these positions implies the death of the organism carrying the mutant, which explains why we do not find it in a sample of living organisms. We can now



**Biological Complexity and Biochemical Information, Figure 3**  
Per-site entropy of the last 56 positions of *E. coli* tRNA from the alignment in Table 1

check how uneven the probabilities per-site are, by plotting the per-site entropy (normalized to lie between zero and one by choosing base  $D = 4$  for the logarithm) against site-number in Fig. 3.

We notice that there are sites that are almost random, and some sites that are absolutely conserved. While this pattern, conserved by evolution over billions of years, clearly reflects the function of the organism, it does not tell us whether the states of any two sites are correlated.

The relative state of two random variables will tell us whether knowing the state of one variable will reveal to us information about another variable. For example, we can ask whether knowing, say,  $X_8$ , allows us to say something about other sites that we would not already know in the absence of that knowledge. Let us see whether the probabilities characterizing  $X_{21}$ , say, depend on knowing  $X_8$ . Collecting frequencies for  $X_{21}$  gives approximately

$$\begin{aligned} p_{21}(A) &= 0.24, & p_{21}(C) &= 0.46, \\ p_{21}(G) &= 0.21, & p_{21}(T) &= 0.09, \end{aligned} \quad (16)$$

while  $p(X_{21}|X_8) =$

$$\begin{pmatrix} p(A|A) & p(A|C) & p(A|G) & p(A|T) \\ p(C|A) & p(C|C) & p(C|G) & p(C|T) \\ p(G|A) & p(G|C) & p(G|G) & p(G|T) \\ p(T|A) & p(T|C) & p(T|G) & p(T|T) \end{pmatrix} = \begin{pmatrix} 0.2 & 0.235 & 0 & 0.5 \\ 0 & 0.706 & 0.2 & 0.333 \\ 0.8 & 0 & 0.4 & 0.167 \\ 0 & 0.059 & 0.4 & 0 \end{pmatrix}. \quad (17)$$

## Biological Complexity and Biochemical Information, Table 1

Alignment of the last 56 nucleotides of sequences of tRNA genes of *E. coli*

AGAGCGCTGCTTGCACGCAGGAGGTCTGCGGTTTCGATCCCGCATAGCTCCACCA
AGAGCGCTTGCATGGCATGCAAGAGGTCAGCGGTTTCGATCCCGCTTAGCTCCACCA
TATGTAGCGGATTGCAAAATCCGTCTA-GTCCGGTTCGACTCCGGAACGCGCCTCCA
AGAATACCTGCCTGTACACGCAGGGGTCGCGGGTTCGAGTCCCGTCCGTTCGCGCA
AGGACACCGCCCTTTCACGGCGGTAA-CAGGGGTTTCAATCCCTAGGGGACGCCA
AGAGCAGGGGATTGAAAATCCCGTGTCTTGGTTTCGATTCCGAGTCCGGGACCA
ATTACCTCAGCCTTCCAAGCTGATGA-TGCGGGTTCGATTCCCGCTGCCGCTCCA
AGAGCACGACCTTGCCAAGGTCGGGGTCGCGAGTTCGAGTCTCGTTTCCCGCTCCA
AGAACGAGAGCTTCCCAAGCTCTATA-CGAGGGTTCGATTCCCTTCGCCCGCTCCA
AGAGCCCTGGATTGTGATTCCAGTTGTCTGTGGGTTTCAATCCCATAGCCACCCCA
AGAGCGCACCCCTGATAAGGGTGAGGTCGGTGGTTCAAGTCCACTCAGGCCATACCA
AGAGCAGCGACTCATAATCGCTTGGTTCGCTGGTTCAAGTCCAGCAGGGGACCA
AGAGCAGTTGACTTTTAATCAATTGGTCGCAGGTTTCAATCCTGCACGACCCACCA
AGAGCACATCACTCATAATGATGGGGTCACAGGTTTCAATCCCGTTCGTAGCCACCA
AGAACGCGGACTGTTAATCCGTATGTCACTGGTTCGAGTCCAGTCAGAGGAGCCA
AGCGCAACTGGTTTGGGACAGTGGGTCGGAGGTTTCAATCCTCTCTCGCCGACCA
AGCGCACTTCGTTCGGGACGAAGGGTCGGAGGTTTCAATCCTCTATACCCGACCA
AGCGCACCGCTCATGGGGTGTCTGGGGTCGGAGGTTCAAATCCTCTCTGCCGACCA
AAGGCACCGGTTTTTGATACCGGCATTCCTGGTTTCAATCCAGGTACCCAGCCA
AAGGCACCGGATTCTGATTCCCGCATTCGAGGTTTCAATCCTCTGATCCCGACCA
AGAGCGCTGCCCTCCCGAGGCAGAGGTTTCAGGTTTCAATCCTGTCTGGGCGCGCCA
AGAGCAACGACCTTCTAAGTCGTGGGCGCAGGTTTCAATCCTGCAGGGCGCGCCA
AGAGCAACGACCTTCTAAGTCGTGGGCGCAGGTTTCAATCCTGCAGGGCGCGCCA
AGAGTACTCGGCTACGAACCGAGCGGTTCGAGGTTTCAATCCTCCCGGATGCACCA
ATAACGAGCCCCCTCTAAGGGCTAAT-TGCAGGTTTCGATTCTGCAGGGGACACCA
AGAGCGCACCCCTTGGTAGGGGTGGGGTCCCCAGTTCGACTCTGGGTATCAGACCA
AGAGCGCACCCCTTGGTAAGGGTGAGGTCGGCAGTTCGAATCTGCCTATCAGACCA
AGAGCAACTGACTTGTAAATCAGTAGGTCACAGTTCGATTCCGGTA.TCGGCACCA
AGAGCAGCGCATTCGTAATGCGAAGGTCGTAGGTTTCGACTCCTATTATCGGCACCA
AGAGCGCACCCCTTGGTAAGGGTGAGGTCGCCAGTTCGACTCTGGGTATCAGACCA
AGAGCACCGGTCTCCAAAACCGGGTGTGGGAGTTCGAGTCTCTCCGCCCTGCCA
AGCTCGTCGGGTCATAACCCGAAGATCGTCGGTTCAAATCCGGCCCCCGCAACCA
AGCTCGTCGGGTCATAACCCGAAGGTCGTTCGGTTCAAATCCGGCCCCCGCAACCA

Knowing  $X_8$  would help us in predicting the state of  $X_{21}$  if any of the numbers in any of the columns of Eq. (17) is significantly larger than 0.25, which is the probability to guess  $X_{21}$  right by random chance. In fact, however, only a very few of the probabilities are even above 0.5. Clearly, the knowledge of position 8 does not help much in predicting position 21. Given  $X_8 = G$ , for example, the probabilities to correctly predict A, C, G or T at position 21 are given by the third column of Eq. (17), and none of those probabilities exceed 0.5. The only case where knowing  $X_8$  helps us significantly in predicting  $X_{21}$  is if  $X_8 = A$ , since then  $X_{21} = G$  with probability 0.8. But let us look instead at another pair of sites, this time  $X_8$  and  $X_{22}$ . We can obtain the *unconditional* probabilities for  $X_{22}$  from the alignment (as before, we are rounding the probabilities)

$$\begin{aligned} p_{22}(A) &= 0.18, \quad p_{22}(C) = 0.15, \\ p_{22}(G) &= 0.52, \quad p_{22}(T) = 0.15, \end{aligned} \quad (18)$$

but the conditional probability matrix is very different

from Eq. (17):

$$p(X_{22}|X_8) = \begin{matrix} & \begin{matrix} A & C & G & T \end{matrix} \\ \begin{matrix} A \\ C \\ G \\ T \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \end{matrix}. \quad (19)$$

This matrix immediately tells us that, based on the alignment in Table 1, if we see a T at position 8, then we can be certain ( $p = 1$ ) to find A at position 22, while if we see a G at position 8, we are sure to encounter C at 22, and so forth. Obviously, these are the associations implied by Watson-Crick pairing, so this conditional probability matrix suggests to us that position 8 and 22 are in fact *paired* within the molecule: this is how biochemistry stores information in molecules.

Calculating the conditional entropy  $H(X_{22}|X_8 = T)$  makes the correlation apparent in another way. As the entropy of a random variable measures the amount of uncertainty that we have about it, we immediately see that

knowing  $X_8$  tells us everything about the state of  $X_{22}$  that there is to know: for every one of  $X_8$ 's possible states,  $H(X_{22}|X_8 = A, C, G, T) = 0$ . We can now introduce the *average conditional entropy*:

$$H(X|Y) = \sum_{j=1}^D p(y_j) H(X|Y = y_j), \quad (20)$$

which gives us the average value of the entropy of  $X$  given any of the states of  $Y$ . For the pair  $(X_{22}, X_8)$  this gives of course  $H(X_{22}|X_8) = 0$ , while for the pair  $(X_{21}, X_8)$  we find

$$H(X_{21}|X_8) = 0.58. \quad (21)$$

We saw earlier that knowing site 8 does not significantly affect our probability to predict site 21, but still the conditional entropy Eq. (21) is significantly smaller than the unconditional one,  $H(X_{21}) = 0.9$ .

### Information

From the example we just discussed, the definition of biochemical information is clear: it is just the reduction in entropy (or uncertainty) of one variable using the knowledge of the state of another. So, if  $Y$  is again a variable whose entropy we seek to decrease using our knowledge of  $X$ , then the information  $X$  conveys about  $Y$  is just

$$I(X : Y) = H(Y) - H(Y|X). \quad (22)$$

This definition of information (also sometimes called “shared” or “mutual” entropy) is symmetric:  $I(X : Y) = I(Y : X)$ , and we can therefore equivalently write

$$I(X : Y) = H(X) - H(X|Y). \quad (23)$$

Thus, “what  $X$  knows about  $Y$ ,  $Y$  also knows about  $X$ ”. This is particularly clear for the example of biochemical information through Watson-Crick pairing that we looked at above. We can calculate the information position 8 has about position 22 as

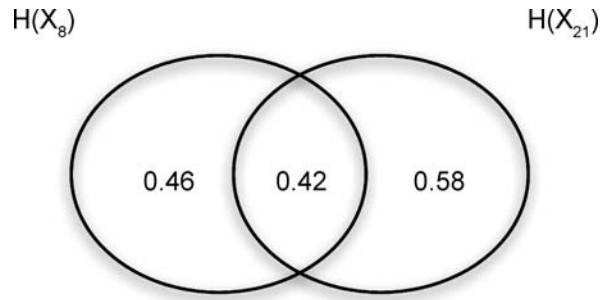
$$I(X_{22} : X_8) = H(X_{22}) - H(X_{22}|X_8) = H(X_{22}) \quad (24)$$

because  $H(X_{22}|X_8) = 0$  as we found above. At the same time,

$$I(X_8 : X_{22}) = H(X_8) - H(X_8|X_{22}) = H(X_8) \quad (25)$$

because  $H(X_8|X_{22})$  also vanishes. Then,  $H(X_8)$  and  $H(X_{22})$  have to be equal. We can repeat this analysis for the other pair we looked at:

$$I(X_8 : X_{21}) = H(X_{21}) - H(X_{21}|X_8) = 0.9 - 0.58 = 0.42. \quad (26)$$



Biological Complexity and Biochemical Information, Figure 4  
Venn diagram of entropies between sites 8 and 21 of the sequence alignment (1)

A simple diagram (Fig. 4) helps to see how entropies are distributed among shared and conditional entropies.

As we discussed earlier, if two sites share most of their entropy, we can conclude that they bind to each other in a Watson-Crick pair. Because this binding is responsible for the structure of the molecule, information theory can help us to determine the molecule's *secondary structure*, that is, how the molecule is arranged as a chain in two dimensions [34].

### Molecular Complexity

While it is important to understand how biochemistry encodes information in chemical bonds, calculating information-theoretical correlations between nucleotides (or between residues in proteins [35,36,37,38]) does not reveal how much information a molecule stores about the environment within which it functions. We can achieve this by imagining that the physical and chemical world has a *description* in terms of a symbolic sequence. Then, given this particular sequence (and thus, given the identity of the world within which we study the entropy of biomolecular sequences), the set of information-rich sequences can be determined, and the information content calculated. The mathematical form of the measure of a molecule's information content will be shown to be closely related to the physical complexity measure discussed earlier in Eq. (3).

We start by considering a molecule that is guaranteed *not* to encode information about its environment: a random molecule. Let us imagine a polypeptide  $P$  of  $L$  residues, written as a joint random variable  $P = P_1 P_2 \dots P_L$ . If  $P$  is truly random, then each of its  $20^L$  states are equally likely, and the entropy of  $P$  is

$$H(P) = \log 20^L. \quad (27)$$

If we choose to take 20 as the base of the logarithm, we see that the entropy of the random  $L$ -mer is  $H(P) = L$ , that is, one unit of entropy for each random monomer. Of course, functional proteins are nowhere near random. Instead, the probabilities to find a particular residue  $x$  at position  $n$ ,  $p_n(x)$ , are strongly constrained by evolution. The opposite extreme of a protein would be the case where one and only one particular residue is allowed at each position in a protein that has a specific function. Then, only one state of the polymer is consistent with that function, and  $H(P) = 0$ . In general, the entropy of  $P$  is given by

$$H(P) = - \sum_i^{20^L} p_i \log p_i, \quad (28)$$

where  $p_i$  is the probability to find any of the possible  $20^L$  states of  $P$  in an infinitely large ensemble of polymers of length  $L$ . In practice, however, ensembles large enough to estimate these  $p_i$  cannot exist (note that  $20^{100}$  is approximately  $10^{130}$ ), so we need to find approximations to calculate  $H(P)$  if we would like to measure a polymer's entropy. A common approximation is to neglect interactions between the different sites  $n$ , so that the entropy of  $P$  can be written in terms of a sum over the entropies of each monomer:

$$H(P) \approx \sum_{n=1}^L H(P_n). \quad (29)$$

In practice this is not a bad approximation [39], but we should keep in mind that interactions between residues, known as *epistasis*, are extremely important in evolution and are often studied in detail [40,41,42]. Typically, while many pairs of residues interact epistatically, some do so positively and some negatively, so that on average the approximation Eq. (29) often holds.

If we can consider each residue separately, we can now focus on the probabilities  $p_n(x)$  introduced above. Typically, we expect those sites that are very important for the function of the protein to be strongly conserved (all the  $p_n(x)$  are either one or zero at position  $n$ ), leading to a vanishing entropy at that site, while those that are less important are less constrained, leading to a larger entropy at those sites. We can now define the amount of information a sequence  $P$  stores about its environment as the mutual entropy (information) between the random variable describing protein  $P$  and a random variable  $E$  describing all possible environments (we imagine, as before, that all possible environments can be described in terms of sequences, and listed)

$$I(P : E) = H(P) - H(P|E), \quad (30)$$

which is the same formula as Eq. (46). The first term on the right hand side of Eq. (30) is the *unconditional* entropy of the polymer  $P$ . An entropy that is not conditional on any environment is an entropy that is unspecified. This is the same as the entropy of a random protein, and thus  $H(P) = H_{\max}(P) = L$ . The second term in that equation is the average conditional entropy of the protein, averaged over all the possible environments described by the random variable  $E$ . Our world, of course, is just one particular such environment  $E = e$ , and therefore the amount of information stored in a polymer  $P$  about environment  $E = e$  is given by

$$I(P : e) = L - H(P|e), \quad (31)$$

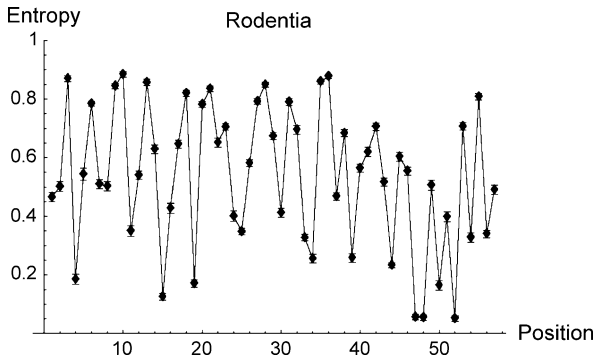
where

$$H(P|e) = \sum_{i=1}^L H(P_i|e) = - \sum_{n=1}^L \sum_{x=1}^{20} p_n(x) \log_{20} p_n(x), \quad (32)$$

and the sum over  $x$  goes over the possible 20 amino acids at each site.

The probabilities  $p_n(x)$  can be obtained by a sequence alignment of structurally identical proteins of different species, such as those listed in the *Pfam* database [43]. As an example, we can align the sequences of the homeobox proteins of the rodent family *Rodentia*. Aligning 703 of the sequences of 57 residues in the database (as of July 2006) allows us to estimate the  $p_n(x)$  necessary for calculating the rodent homeodomain information content. Note that for this analysis to be significant, we have to ensure that the sequences of all the aligned proteins code for a protein with the *same* functionality. Indeed, any time two different residues are allowed at a particular position, we must be able to imply that these are neutral substitutions in the protein. Adaptive changes that influence protein function should not appear in the alignment. The average sequence identity of the set of 703 proteins is about 39%, which gives confidence that the set is composed of sequences coding for proteins that are at least structurally identical. (Sequences with more than 30% identity have more than a 90% chance of coding for structurally identical proteins [44].)

When estimating entropies from finite ensembles, care must be taken to correct the estimates for a bias that arises when the ensemble is small. The method for correcting this bias is well-known [45,46], and applied to all data shown here. We start by calculating the per-site entropy of all 57 aligned sites, giving the *entropy profile* of the sequence shown in Fig. 5. This view reveals a curious alterna-



**Biological Complexity and Biochemical Information, Figure 5**

Entropic profile of the 57 amino acid rodent homeodomain, obtained from 703 sequences in Pfam (St. Louis mirror, accessed July 20, 2006). Per-site entropy between zero and one by taking logs to base 20

tion between high and low entropy sites in the homeobox protein.

The information content is obtained by summing up the per-site entropies and subtracting this value from the length of the sequence, as implied by Eq. (31). If we call the unit of information obtained by using logarithms to the base of the size of the alphabet the “mer” (such that an  $L$ -mer has a maximum information content of  $L$  mers) then the information content of rodent homeobox proteins is

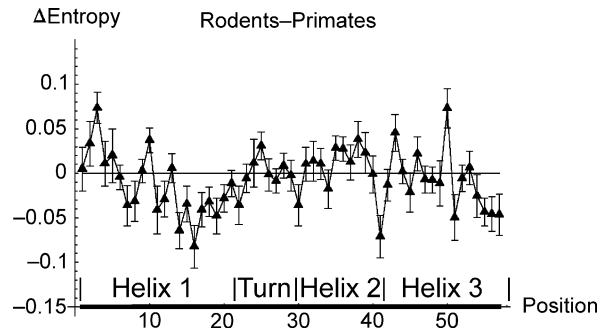
$$I(P_{\text{rodents}}) = 26.83 \pm 0.13 \text{ mers}, \quad (33)$$

where the quoted error reflects the statistical nature of the probability estimate, quantified in [45,46]. Note that for amino acids, 1 mer equals about 4.32 bits, while for nucleic acids 1 mer equals 2 bits.

From the point of view of evolution, we may ask whether this information content has changed as species evolved. For example, we might ask whether animals that, according to our intuitive understanding of the “great chain”, are considered “higher” than rodents show a different information content for this protein. The homeobox protein sequences of many other families of organisms are available in *Pfam* to examine this. As an example, the alignment of 504 sequences of primate homeobox proteins results in an entropic profile remarkably similar to that of Fig. 5. The information content for these proteins can be calculated to be

$$I(P_{\text{primates}}) = 26.71 \pm 0.14 \text{ mers}, \quad (34)$$

in other words, identical to that of the rodents within statistical error. But just because the total information content is the same does not imply that information is coded



**Biological Complexity and Biochemical Information, Figure 6**

Difference between entropic profile of the homeobox protein of rodents and primates (the latter from 504 sequences in Pfam St. Louis, accessed July 20, 2006)

in a similar manner. We can study this by subtracting one entropic profile from the other, to see if the entropy of some sites has increased, and some others decreased in the evolution from rodents to primates. We can see this difference plot in Fig. 6, which suggest that some recoding did indeed take place, in particular in the first helix of the protein, but the significance of this result is not strong.

Thus, information theory can track some aspects of evolutionary changes between proteins, even if the total amount of information is unchanged.

### Biochemical Information and Functional Complexity

It is tempting to speculate that the information content of a biomolecule is related to the functional complexity of an organism. After all, possessing information about an ensemble enables the *prediction* of the possible states of an ensemble with accuracy better than random, something that is highly valuable for a biological organism whose ensemble is an uncertain environment. Let us recall the measure of sequence complexity (“physical complexity”) introduced in the previous section:

$$C_P(s) = K_0(s) - K(s|e), \quad (35)$$

for any sequence  $s$ , given an environment sequence  $e$ . If we take the average of this quantity over an infinite ensemble of sequences  $s_i$  drawn from an ensemble  $S$ , we obtain

$$\langle C_P(s) \rangle_S = \sum_{s_i} p(s_i) (K_0(s_i) - K(s_i|e)), \quad s_i \in S. \quad (36)$$

It is easy to prove the following inequality between average Kolmogorov complexities and the Shannon entropy [47]:

$$\sum_{s_i} p(s_i) K(s_i) \geq \sum_{s_i} p(s_i) \log \frac{1}{p(s_i)} = H(S). \quad (37)$$



The inequality in Eq. (37) reflects the possibility that the complexities  $K(s_i)$ , which can be viewed as compressed *encodings* of the sequences  $s_i$ , do not necessarily form a *perfect code* that saturates the Kraft inequality [48]. However, because the  $K(s_i)$  do represent the *smallest* program encoding  $s_i$ , it is reasonable to assume that the average Kolmogorov complexity is given by the Shannon entropy of the ensemble up to an additive constant, which represents the length of a program that tells one computer how to simulate another (see, e.g., [48]):

$$\langle K(s_i) \rangle_S \approx H(S) + c. \quad (38)$$

In that case (and assuming that the overall constant cancels from the difference), the average physical complexity becomes

$$\langle C_P(s) \rangle_S \approx H(S) - H(S|e), \quad (39)$$

where  $H(S)$  is the unconditional Shannon entropy of the ensemble of sequences, and  $H(S|e)$  is the conditional entropy. If the ensemble  $S$  consists of sequences of fixed length  $L$ , then the unconditional entropy is  $H(S) = L$ , and  $H(S|e)$  is the conditional entropy of the sequences as in Eq. (31). (Note that, technically, the Kolmogorov complexity for fixed length sequences  $K(s|L)$  is related to the arbitrary length complexity  $K(s)$  via  $K(s) \leq K(s|L) + 2 \log L + c$ , where  $c$  is again the “simulation constant”.) To summarize, the average physical complexity is (assuming perfect coding) equal to the Shannon information that the ensemble has about the environment, that is, a sequence’s information content.

This interpretation of complexity is particularly satisfying from an evolutionary point of view. The value of information lies in the ability of the observer who is in possession of it to make *predictions* about the system that the information is *about*. Organisms, armed with the functionality bestowed upon them by their genetic code, do precisely that to survive. An organism’s metabolism is a chemical machine making predictions about the availability and concentrations of the surrounding chemicals. A cell’s surface proteins make predictions about the type of cells it might interact with, and so on. Viewed in this way, informational complexity should be a near perfect proxy for functional complexity, because information must be used for function: if it is not so used, a sequence represents entropy, not information. An investigation of the informational complexity of evolving computer programs (an instance of “digital life” [49,50]) has shown that the complexity increases in evolution [51] and correlates well with the functional complexity of the programs [52]. A good

example for how the equivalence of function and information is achieved in biochemistry is the evolution of functionality in ribozymes by in-vitro evolution, achieved by Jack Szostak’s group at Massachusetts General Hospital.

This group evolved short GTP-binding RNAs (aptamers) in vitro, and found eleven distinct structures with different binding affinities [53]. By measuring the information content as outlined here (within each pool of sequences that evolved for each structure, there were sufficient mutants that could be aligned in order to determine the substitution probabilities  $p_n(x)$ ), it was possible to show that increased functional activity went hand-in-hand with increased information content, so much so that the group was able to derive a simple law that predicts, within this GTP-binding set of ribozymes, that a ten-fold higher binding affinity is achieved by about 10 bits of extra information. In other words, the informational complexity is linearly proportional to the functional complexity. Even more, the structural complexity, as measured by the number of different stems (ladders) within the secondary structure of the enzyme, also seemed to increase with functional activity.

Based on this type of observation, Szostak has proposed a new measure of functional complexity [10,54] that is based both on function and on information. For a particular function  $x$ , let  $E_x$  represent the degree of that function achieved by a system. Then the *functional information* is defined as [54]

$$I(E_x) = -\log(F(E_x)), \quad (40)$$

where  $F(E_x)$  is the fraction of all possible configurations of the system that possess a degree larger or equal to  $E_x$ . For sequences, the function could represent a binding affinity, or the number of ATPs produced by a pathway within which the enzyme is the bottleneck factor, or any other real-valued attribute that characterizes the performance of the sequence. This measure introduces a clear link between information and function, but fundamentally turns out to be a coarse-grained version of the information content Eq. (31), as can be seen as follows.

Suppose we are interested in measuring the information content of a sequence  $s$  that performs function  $x$  to the degree  $E_x$ . We can obtain the functional information of  $s$  by creating all possible mutants of  $s$ , and measuring the fraction  $F(E_x)$  of sequences that have the same function as  $s$ , given by  $\nu(s)/N$ , where  $\nu(s)$  is the number of neutral mutants of  $s$  within  $S$ , and  $N$  is the total number of possible sequences. Thus,

$$I(E_x) = \log N - \log \nu(s). \quad (41)$$

The conditional probability to find a sequence  $s_i$  given environment  $e$  in an evolving populations of sequences of the type  $s$  is given by  $p(s_i|e)$ . If  $e$  specifies the function  $x$  at level  $E_x$  for  $s_i$ , then  $p(s_i|e) = 1$  if  $s_i$  performs the function at the required level, and zero otherwise (coarse-graining of the entropy). There are  $v(s)$  such sequences in the ensemble, and thus

$$\begin{aligned} H(S|e) &= - \sum_i p(s_i|e) \log p(s_i|e) \\ &= - \sum_{v(s)} \frac{1}{v(s)} \log \frac{1}{v(s)} = \log v(s). \end{aligned} \quad (42)$$

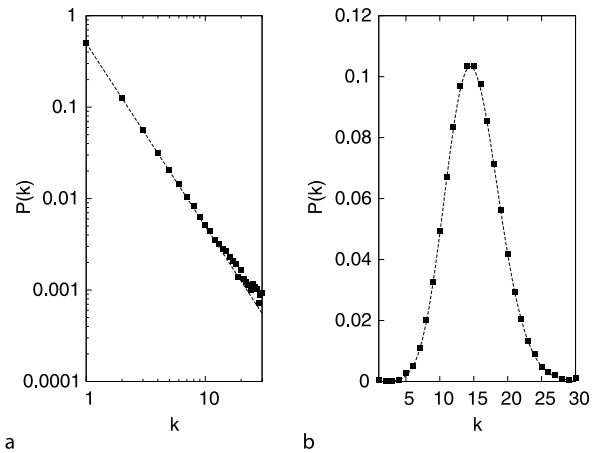
As  $\log N = \log D^L = L$ , Eq. (41) recovers Eq. (31), that is, functional information is a coarse-grained version of the Shannon information content of a sequence.

### Network Complexity

If we are given the informational complexity of every single protein in an organism, is the sum of these complexities equal to the complexity of the cell, or the organism? Certainly not, because much of the complexity of an organism lies in how the proteins interact, and in particular in the complicated temporal sequence of events dictated by the regulation of the expression of proteins as a function of environmental influences.

It is well-known that functional biological networks have properties that distinguish them from a randomly connected set of nodes (a random graph, or Erdős-Rényi network). In fact, many biological networks have degree distributions that are approximately scale-free (Fig. 7a), in stark contrast to the binomial distribution of the random graph (Fig. 7b). Also, the diameter of biological networks, defined as the average of all internode distances  $d(i,j)$ , is small (“small-world-network”, see [55]) and depends only weakly on the number of nodes in the network, again unlike what we find for random graphs.

Assessing the complexity of a network usually entails measuring the structural complexity of the network, as compared to a random graph. Often, the *modularity* of a network is used as a proxy for the structural or the functional complexity, even though the concept of a module for networks is not universally defined [56,57]. Usually, a module is defined as a discrete entity whose function is separable from those of other modules, but in biology modules can have significant overlap. Within protein-protein interaction networks, putative modules can be obtained by clustering, so that modules are sets of proteins that are strongly interconnected, but only weakly connected to other such sets. In this section I discuss



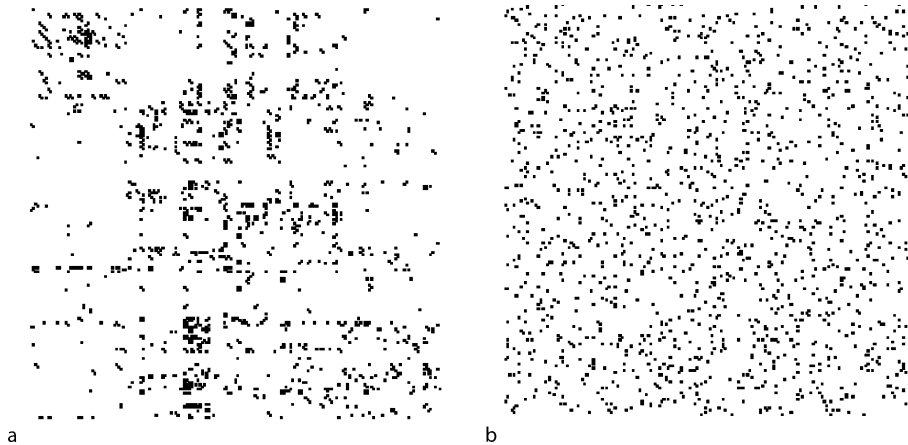
**Biological Complexity and Biochemical Information, Figure 7**  
**a** Degree distribution  $P(k)$  to find a node with  $k$  edges in a scale free network (squares) with degree distribution  $P(k) \sim 1/k^2$  (dashed line), **b** degree distribution for a random graph with 1,000 nodes and a connectivity  $p = 0.015$  (squares) and a Poisson distribution with mean  $\langle k \rangle = 15$  (line)

several ways to understand network complexity, starting with an estimate of the information content of a functional network by assessing the information content of the genome that gave rise to it, using the methods discussed in Sect. “Molecular Complexity”. I then review an information-theoretic approach to network modularity, followed by an exposition of a method to assess network structure and information by measuring subnetwork (sub-graph) abundances.

### Evolution of Information in Networks

The difficulty in assessing network complexity from network topology is clear to anyone who has studied the multitudes of networks arising in engineering and biology. Biological networks usually have thousands of nodes and several thousand edges, and often appear to be unstructured. For example, the network summarizing the connectivity of neurons in the brain of the nematode *C. elegans* shows very little modularity or structure at first sight, but is markedly different from a random network [58] (see Fig. 8).

Because the functionality of a network is not necessarily reflected in its topological structure, the best hope for assessing the complexity of a network is to measure the complexity of the set of rules used to construct it. In biology, this set of rules is encoded in the genome, so a first-order estimate of the complexity of a network should be given by the complexity of the genome that produced it. Of course, this is difficult for all the reasons given in the



**Biological Complexity and Biochemical Information, Figure 8**

Adjacency matrix of 179 of the 302-neuron neural network of a *C. elegans* brain (left), and a random network of the same size and connectivity (right)

previous section, but even more difficult in this case because a network of proteins, for example, is specified not just by the open reading frames coding for the proteins, but also all the untranslated regulatory regions as well as the transcription factors affecting them.

We can test the evolution of network complexity in computational models where a genome represents the functionality of a cellular network, as was done recently in Ref. [59]. In this work, an artificial chemistry and genetics was encoded in a simple linear (circular) code based on the monomers 0, 1, 2, 3, where enzymatic proteins with variable specificity act on 53 precursor molecules to form up to 555 metabolites. The metabolic reactions involving transport and enzymatic proteins are obtained by a translation of the genetic code into reactions, and implementing chemostat physics and reaction kinetics. Evolution proceeded from a simple ancestral genome with only 3 genes to large and complex metabolic networks of thousands of nodes and several thousand edges, in a completely asexual Wright–Fisher process acting on two chromosomes.

In order to be considered fit, an artificial cell has to import precursor molecules that are available outside the cell walls and convert them into metabolites. The fitness of an organism was determined by calculating the produced biomass of metabolites (see [59]). Evolution was carried out in three different environments that differ in their predictability. In the simplest environment, the location of precursor sources and their abundance is constant during evolution (the “static” environment), while in the quasistatic environment one randomly selected precursor source location is changed per update. In the dynamic environment, the source location of all precursors is changed

randomly, and 25% of all precursors are made unavailable, giving rise to a highly unpredictable environment.

The information content of the genomes was measured as outlined above, that is

$$I = L - H(s), \quad (43)$$

where  $L$  is the total length of the sequence and  $H(s)$  is the sum of per-site entropies

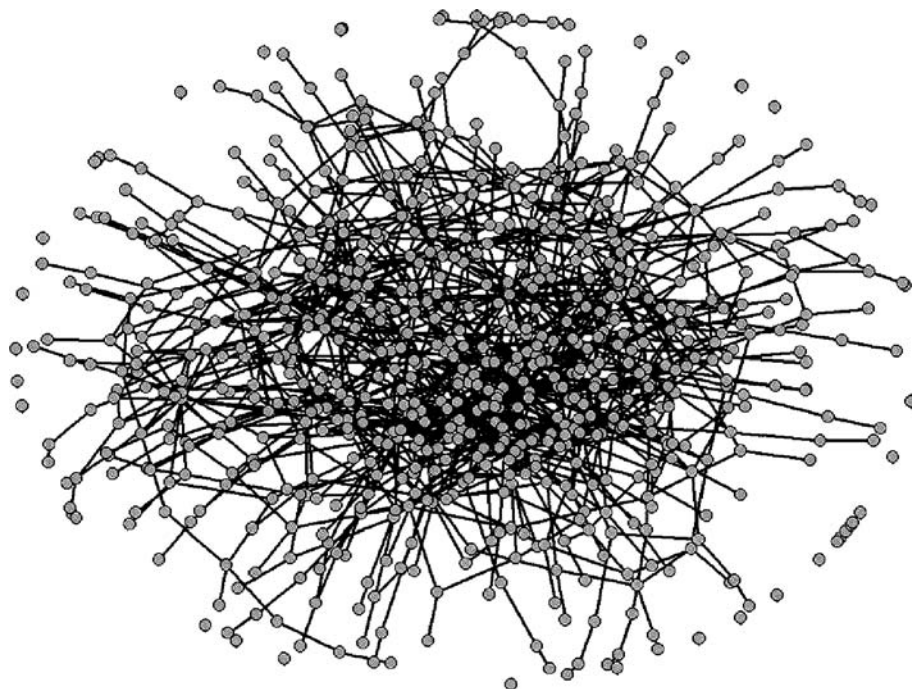
$$H(s) = \sum_{x=1}^L H(x) \quad (44)$$

and the per-site entropy  $H(x)$  is obtained by summing over the substitution probabilities  $p_i$  at that site:

$$H(x) = - \sum_{i=0}^3 p_i \log_4 p_i. \quad (45)$$

Because we only have 4 possible symbols per site, taking the logarithm to base 4 again ensures that the per-site entropy lies between 0 and 1.

As the evolutionary mechanism allows for insertions and deletion of entire genes or genetic regions along with point mutations, genomes can change length during evolution. As a consequence, an alignment of genomes in a population to ascertain substitution probabilities is problematic. Instead, an approach can be used that determines the substitution probabilities  $p_i$  from the fitness effect of the substitution on organism fitness, along with an application of population genetics theory. If a substitution of allele  $i$  has fitness effect  $w_i$ , then the probability to find this allele in an equilibrated population evolving at mutation



Biological Complexity and Biochemical Information, Figure 9

Evolved metabolic network with 969 nodes and 1,698 edges, rendered with PAJEK [60]

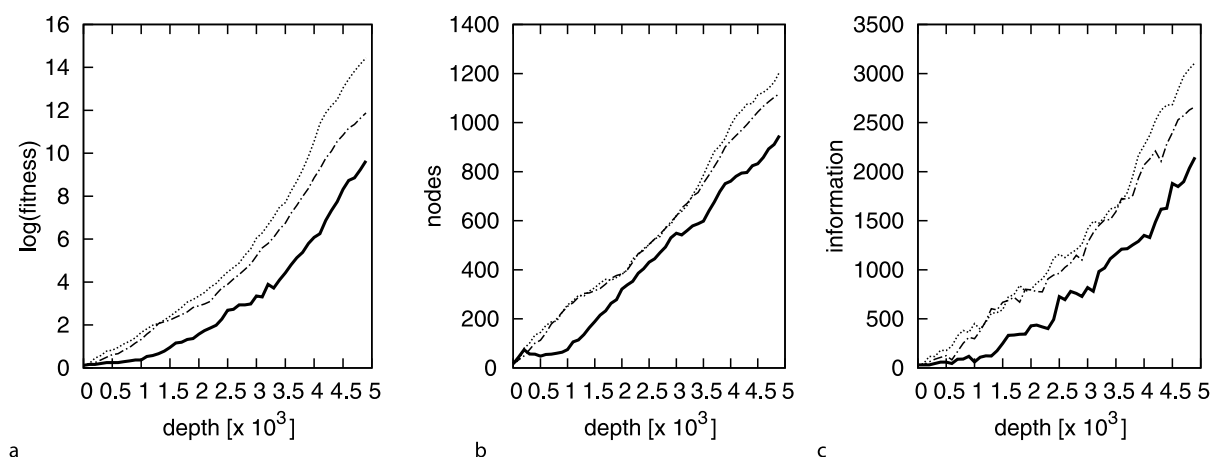
rate  $\mu$  is given by [61]

$$p_i = \frac{p_i w_i}{\bar{w}} (1 - \mu) + \frac{\mu}{4} \sum_{j=0}^3 \frac{p_j w_j}{\bar{w}}, \quad (46)$$

where  $\bar{w} = \sum_{i=0}^3 p_i w_i$  is the mean fitness of the 4 possible alleles at that position.

Figure 10 shows the evolution of network fitness, size and complexity (measured in units of monomer entropy or “mer”, where one mer equals 2 bits) for the three different environments discussed above, as a function of the phylogenetic depth of the organism. The phylogenetic depth of an organism is its position on the line of descent of the particular evolutionary run: The organism with the highest fitness at the end of the run is used to reconstruct the line of descent by following its direct ancestry and increasing the depth counter whenever an organism’s genome differs from that of its direct parent. When arriving at the initial organism, this counter is set to zero, so that the phylogenetic depth counter increases up until the last organism on the line. Because on average we find about one new organism on the line of descent per generation in these runs, the phylogenetic depth is a good proxy for evolutionary time, even though in principle many generations could pass without an advance on the line of descent.

Because the fitness of an organism is multiplicative in the biomass (discovering how to produce a new metabolite multiplies the previous fitness by a number greater than one), the log of the fitness grows about linearly for all three environments (Fig. 10a). The fitness grows fastest for the static environment that is the easiest to predict (dotted line), while it takes longer for complexity to emerge in the dynamic environment (solid line). The same trend is reflected in the growth of the number of nodes and edges in these environments (Fig. 10b). Finally, the information content as calculated by Eq. (43) using the substitution probabilities Eq. (46) follows the same trend: the informational complexity grows the fastest for the static and quasi-static environments, and lags behind for evolution in a dynamic environment. The reason for the slower growth of complexity for networks evolving in dynamic environments is clear: because the availability of precursors necessary for the production of complex metabolites cannot be relied upon in such environments, the cells end up manufacturing the precursor molecules within the cells (rather than importing them from the outside). This machinery is complex in itself, but takes time to evolve. Ultimately, we expect networks evolving in dynamic environments to be more complex than those evolving in static environments because of the added flexibility of producing precursor molecules within the cells. However, such net-



Biological Complexity and Biochemical Information, Figure 10

Evolution of complexity in artificial metabolic networks. **a** log fitness for networks evolving in a static (dotted line), quasistatic (dash-dotted), and dynamic environment (solid line). **b** Evolution of the number of nodes (number of edges follows a similar trend). **c** Evolution of informational complexity, lines as in **a**

works lag behind slightly during the time this complexity is generated.

Note that the informational complexity of the networks used to seed these evolutionary experiments is rather low: the network with three genes (two for importing precursors and one for metabolizing those) is specified with a genome of informational complexity of just 36 mers (72 bits), even though the starting genome has 1,000 “nucleotide” positions in each of the two chromosomes. The non-coding part of these initial genomes thus does not contribute to the informational complexity, because changing any of these positions to any other allele cannot change the fitness of the organism (we do not take beneficial mutations into account in the fitness tests). For these non-coding nucleotides, the  $p_i$  calculated by Eq. (46) all are exactly equal ( $p_i = 1/4$ ), guaranteeing that they do not contribute to  $I$ , as is easily checked. However, the informational complexity grows rather quickly once more and more metabolic reactions are discovered and optimized, at a pace of about 0.5 mers (1 bit) per depth step (roughly one bit per generation).

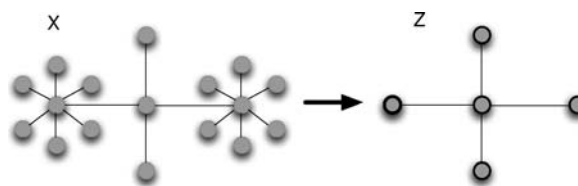
### Modules from Information Theory

The path towards an understanding of the functional organization of a network in the absence of genomic information usually involves the *decomplexification* of the network, either by clustering nodes that are related in function [62], removing those nodes that are immaterial to (or redundant in) function, or analyzing the subgraph decomposition [63] as discussed further below. A par-

ticularly insightful method to decompose networks into modules—both overlapping and non-overlapping—uses information theory to estimate how much information about the original network is present in the abstract—that is, decomplexified—version, while maximizing a variable that measures the *relevance* of the abstraction. This method, sometimes called the “information-bottleneck” approach [64] was applied to biological and engineering networks by Ziv et al. [65].

The *Network Information Bottleneck* (NIB) approach attempts to replace a complex network by a simpler one while still retaining the essential aspects of the network. For example, a highly connected star topology could be replaced by a single node that represents the modular function of the star, as in Fig. 11.

The main idea of the method is that while there are many different ways in which one can collapse a topology, the optimal mapping is one where the new topology retains as much information as possible about the original



Biological Complexity and Biochemical Information, Figure 11

Collapse of a topology described by the nodes  $X$  to a more succinct one described by  $Z$  in which clusters are replaced by a cluster assignment variable  $Z$

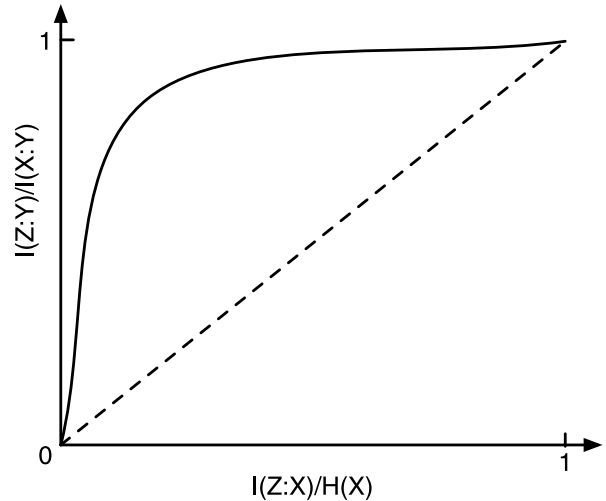


one, while maintaining as much *relevance* of the description as possible. Say, for example, that a random variable  $X$  stands for nodes  $x \in X$  in a network that occur with probability  $p(x)$ , and total number of states  $|X| = N$ , where  $N$  is the size of the network. A *model* of this network can then be made by a random variable  $Z$  with *fewer* states  $|Z| < N$ , and where a cluster assignment is given by a set of probabilities  $p(z|x)$ : the probability that  $z$  is assigned to a particular cluster given the input node  $x$ . Ideally, we would like to maximize the mutual entropy (information) between the random variables  $X$  and  $Z$ , but we shall do this with a constraint given by a relevance variable mentioned earlier. This relevance variable will distinguish different ways in which clusters are assigned. In this application of the NIB to networks, the relevance variable involves diffusion on the network: those nodes that are close to each other are preferentially visited by a diffusion algorithm, and are more likely to be clustered together. In this algorithm, the relevance is represented by a random variable  $Y$  that is defined such that a diffusive process determines the probability to arrive at node  $y$ . The relation to the network variable  $X$  is given by the joint probability  $p(x, y) = p(y|x)p(x)$ , the probability to arrive at node  $y$  via a diffusive process given the process started at node  $x$ , times the probability that we started at node  $x$ . The latter probability is always assumed to be uniform, that is,  $p(x) = 1/N$ .

The NIB algorithm allows optimal solutions where several different nodes  $z$  could be assigned to any given input node  $x$ . This flexibility allows for the possibility of *overlapping* clusters or modules (soft clustering), but in the following we will follow only the algorithm for hard clustering, so that for any choice of  $x$  and  $z$ ,  $p(z|x)$  is either one or zero.

One version of the algorithm (agglomerative clustering) begins with a random variable  $Z$  that has exactly one fewer nodes than  $X$ , and attempts to find the optimal pair of  $x$  nodes to join to produce a model of the network with one fewer nodes. For each possible cluster assignment  $p(z|x)$  we can execute a diffusion process to determine the matrix  $p(y|z)$ , that is, the probability to arrive at node  $y$  given a node  $z$  as starting point. We choose to merge those nodes that maximize  $I(Y : Z)$ , and the algorithm repeats with a set  $Z$  smaller by one node until all nodes have been merged. At each step, we can calculate the normalized variables  $0 < I(Z : X)/H(X) < 1$  and  $0 < I(Z : Y)/I(X : Y) < 1$  and plot them against each other, giving rise to the *information curve* [65], as in Fig. 12.

A completely random network gives rise to the diagonal in Fig. 12, and represents the least modular network. We can define a modularity score, the *network modularity*,



**Biological Complexity and Biochemical Information, Figure 12**  
The information curve for a modular network (solid line), obtained by starting with a model network  $Z$  of the same size as  $X$  (maximal information  $I(Z : X)/H(X) = 1$  and  $I(Z : Y)/I(X : Y) = 1$ , upper right corner), and merging nodes while maximizing  $I(Z : Y)$ . This process generates the information curve from the upper right corner all the way down to the lower left corner, where  $|Z| = 1$  and the mutual entropy vanishes. The dashed line represents the information curve for a random network. The modularity score is given by the area under the information curve

as the area under the information curve. Perfectly modular networks then have a maximal modularity score of 1, whereas random networks have a score of  $1/2$ . Several different networks have been analyzed using this modularity measure in Ref. [65], such as the network of co-authors for papers presented at a meeting of the American Physical Society. This network has 5,604 nodes and 19,761 edges, and yielded a modularity score of 0.9775. The regulatory network of the bacterium *E. coli* (328 nodes and 456 edges), also analyzed by these authors, yielded a score of 0.9709. Thus, both of these networks are highly modular according to this measure.

Interestingly, the network of connections of the *C. elegans* brain (see Fig. 8) has a network modularity score 0.9027, whereas a randomized version retaining the same number of nodes and edges scores 0.4984 on average, as predicted for a random network. The evolved metabolic networks discussed in Sect. “[Evolution of Information in Networks](#)” also score high on this scale. For the largest connected component of a 453 node network, we find a modularity score of 0.8486 for the 5,000th organism on the line of descent (about one organism per generation). While a score for small networks is fairly meaningless, the score increases slightly as the networks become more complex.

### Information in Motifs

We have seen that networks that are functional are built from modules, or at least can be understood in terms of strongly connected sets of nodes that are only weakly connected to other such clusters. This clustering—carried out using information theory in the previous section—can also be performed on the basis of topology alone. For example, every network can be analyzed in terms of its *subgraph composition* [63], that is, the frequency with which particular subgraphs or motifs appear within the entire network. The degree with which certain motifs are overutilized—and some others underutilized—compared to a uniform distribution or one obtained from a random network, reflects the *local structure* of the network and can be used to classify networks from very different realms into similar categories [66].

Subgraph abundances can also be used to study the modular composition of networks, in analogy to the modular composition of sentences in written language. Meaning can be conveyed in text only because the utilization frequency of letters in words, and words in sentences, is different from uniform. For example, the letters e,t,a,i,o, and n appear in decreasing frequency in an average text written in English, while the rank-abundance distribution of the words follows a scale-free distribution (Zipf's Law [67,68]). If we assume that a random sequence of letters contains no information, then the deviation from the uniform distribution could be used to distinguish and perhaps classify functional (that is, meaningful) text from gibberish. In the same vein, it is possible that functional networks differ significantly from random networks in the subgraph utilization, and we can study this difference by estimating the *subgraph information content* as follows.

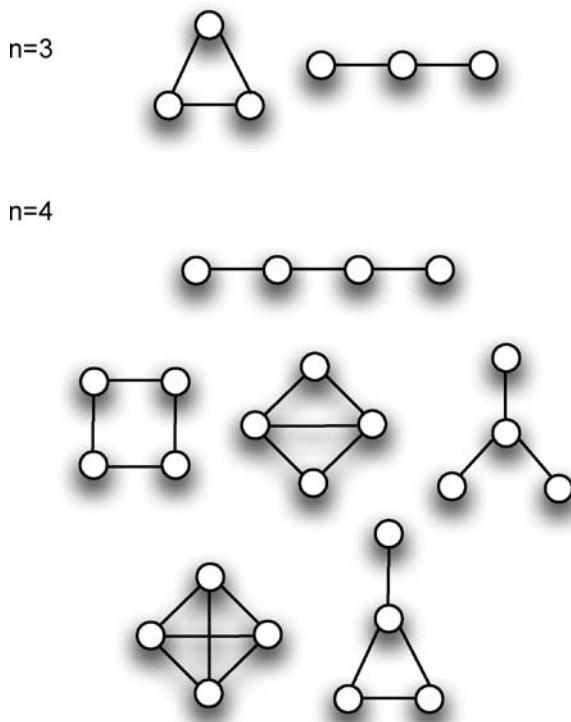
Suppose we compute the probability to find any of the two possible motifs that can be made of three nodes (see Fig. 13). For simplicity, we are considering here only undirected graphs, and do not allow self-interactions, that is, nodes that link to themselves. We can then compare these empirical probabilities to the probabilities with which these subgraphs appear in a random network.

A priori, we might think that any of the two motifs of size  $n = 3$  should appear with equal probability in a random network, giving rise to a motif entropy that is maximal:

$$H_3\left(\frac{1}{2}, \frac{1}{2}\right) = -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} = 1. \quad (47)$$

Here, we defined the size- $n$  motif entropy

$$H_n(p_1, \dots, p_m) = -\sum_{i=1}^m p_i \log_m p_i, \quad (48)$$



**Biological Complexity and Biochemical Information, Figure 13**  
Undirected motifs of size  $n = 3$  and  $n = 4$ , without self-interaction

where  $m$  is the number of possible connected motifs of size  $n$ , and the  $p_i$  are the probabilities to find the  $i$ th motif in the network. (Because the base of the logarithm is also  $m$ , this entropy is normalized to lie between zero and one.) The information stored within  $n = 3$ -motifs would then be (the superscript (u) refers to the uniform baseline distribution)

$$I_3^{(u)} = H_3\left(\frac{1}{2}, \frac{1}{2}\right) - H_3(p_1, p_2) = 1 - H_3(p_1, p_2), \quad (49)$$

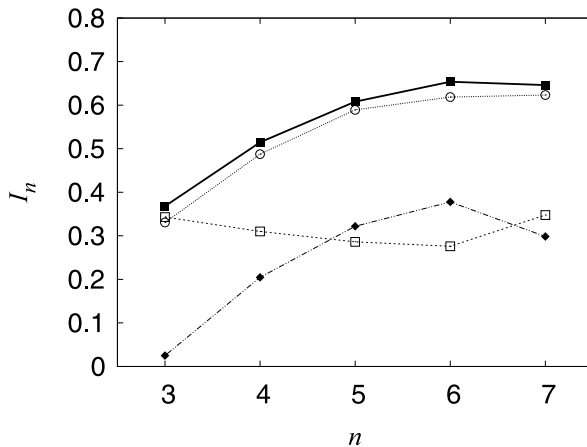
while the information stored in  $n$ -motifs is naturally

$$\begin{aligned} I_n^{(u)} &= H_n\left(\frac{1}{m}, \dots, \frac{1}{m}\right) - H_n(p_1, \dots, p_m) \\ &= 1 - H_n(p_1, \dots, p_m). \end{aligned} \quad (50)$$

However, even random networks do not have a uniform distribution of motifs, and we can instead consider the information stored in motifs compared to a random network as baseline, as (for  $n = 3$ )

$$I_3^{(r)} = H_3(p_1^{(r)}, p_2^{(r)}) - H_3(p_1, p_2). \quad (51)$$

where the  $p_i^{(r)}$  refer to the probability of finding motif  $i$  in a random network.



**Biological Complexity and Biochemical Information, Figure 14**

Motif information  $I_n$  for motifs of size  $n$  for evolved networks with a uniform distribution as a baseline (solid line, filled squares), an Erdős-Rényi network of the same size and number of edges and connectivity  $p \approx 0.005$  (dashed line, open squares), and a randomized network of the same size with a scale-free edge distribution (dotted line, open circles). The motif information using the probability distribution of a random network as the baseline Eq. ( $I_n^{(r)}$ ) is the difference between the solid and dashed lines (dash-dotted line, filled diamonds)

In Fig. 14, we see the information  $I_n^{(u)}$  stored in motifs of size  $n$  for  $n = 3 - 7$  (filled squares, solid line), for a single, evolved, functional, simulated metabolic network discussed in Sect. “Evolution of Information in Networks”, of 598 nodes. The network information increases as more complex motifs (larger  $n$ ) are used for encoding, but appears to stabilize. This behavior mirrors the statistics of  $n$ -gram entropies in English text, as noted early on by Shannon [68]. Note that because the shortest path between any two nodes is on average of the order of 4-5 in these networks [59], motifs of size 7 or larger are not well-sampled.

We can study whether the network information is dictated by functionality, edge distribution, or both, by constructing analogous networks that have the functionality removed by randomizing connections but keeping the scale-free edge distribution, and by randomizing the network but destroying also the edge distribution. If, for example, we randomize the connections in our functional evolved network while keeping the scale-free edge distribution, we find that the network information is only slightly lowered (open circles and dotted line in Fig. 14). On the other hand, if we randomize the network in such a way that the degree distribution is that of a random graph (but still keeping the same number of nodes and edges), the dependence of the network information as a function of the subgraph size is markedly different (open squares, dashed line in Fig. 14). This suggests that the network in-

formation is significantly dictated by the biological scale-free distribution of edges per nodes, and only weakly by the actual function of the network.

Because random graphs do not utilize subgraphs with equal probability (as witnessed by the non-zero network information in Erdős-Rényi networks), it is more appropriate to use the random network probabilities as the baseline to calculate the network information. In this case, we see the network information  $I_n^{(r)}$  increase from small values at  $n = 3$  up to  $n = 6$  (dash-dotted line, filled diamonds in Fig. 14). The decrease noted for  $n = 7$  is due to incomplete sampling of size  $n = 7$  networks for this small graph (598 nodes), and is not significant.

In summary, network information as measured by subgraph “ $n$ -gram” entropies behaves similar to the dependence of  $n$ -gram entropies in written language, and can be used to distinguish functional networks from random ones. However, the network information appears to be controlled mostly by the form of the edge distribution. Insight into the modular structure of networks will likely depend on understanding how the subgraphs of networks are assembled into modules, in analogy to how letters are assembled into words in written text.

## Future Directions

Several billions of years of evolution have shaped our biosphere to become the complicated, interdependent, hierarchical complex system we witness today. Among the parts and members of this system, there are certainly differences in complexity—some obvious to any observer—some less so. Structural complexity, the most intuitive of all measures, is notoriously difficult to define because there is no universal system to rank all possible physical structures. We have seen that automata theory and information theory allow us to quantify the complexity of a sequence in terms of its information content about the environment within which it has evolved, and that this information is a good proxy for the functional complexity of the organism precisely because the information is used by the organism to function in a complex environment.

But the measure is limited because so far it is only practical for short stretches of DNA or single proteins. Thus, a quantitative measure for the complexity of a whole genome using information theory will only be possible when a large number of complete genomes of closely related species is available. Another shortcoming of the informational complexity is that it refers to a particular niche only, and furthermore cannot quantify the complexity of genes that are adapted to varying environments. So, for example, so far we cannot use the informational complex-

ity to estimate the complexity of an ecosystem, nor of an organism that spends part of its lifecycle in one environment, and another part in a completely different one (simple forms of life such as arboviruses are notorious for such a cycle). However, a natural extension of the informational complexity exists that may cover multiple environments both in time and space. Recall that the informational complexity of an ensemble of sequences  $S$  refers to a single environment description  $E = e$ :

$$I(S : e) = L - H(S|e), \quad (52)$$

where  $H(S|e)$  is the ensemble entropy of the sequences. We can generalize this expression by promoting the environment to a true random variable  $E$  that can take on states  $e_i$  with probability  $p(e_i)$ . This formalism can describe environments that are composed of different (spatially separated or overlapping) niches  $e_i$ , as well as environments that take on the different states  $e_i$  periodically or even randomly, in time. The informational complexity then becomes

$$\begin{aligned} I(S : e) &\rightarrow \sum_e p(e) I(S : e) \\ &= I(S : E) = L - H(S|E). \end{aligned} \quad (53)$$

Here,  $H(S|E)$  is the average conditional entropy of the sequence ensemble  $S$  given the environment  $E$ . Whether this construction will turn out to be useful for characterizing the complexity of ecosystems or variable environments remains to be seen, as the practical obstacles are only amplified by having to measure the informational complexity in multiple environments. But at the very least this construction addresses one of the fundamental problems in assessing functional complexity that we encountered in the introduction, namely that for organisms that have adapted to be functional in a variety of environments, we can find genes that appear to show no phenotype upon knockout in the laboratory. Such genes, however, may very well show a phenotype in a particular environment that the organism encounters in the wild, and this functional capacity of an organism needs to be taken into account when assessing functional complexity. If the random variable  $E$  accounts for the multitude of environments with the correct probabilities  $p(e)$ , then the full functional complexity of the organism may be characterized using Eq. (53). But to apply this measure, more efforts need to be expended towards understanding the modes in which an organism functions in its native environment(s) (as opposed to the unnatural laboratory conditions that are the norm today). If such an effort is made, then as we can expect an exponential increase in sequence data in the coming years, the

prospects for a general understanding of biological complexity in terms of sequence complexity are good.

## Acknowledgments

I am grateful to Arend Hintze for the collaborative work in Sect. “Network Complexity”, as well as numerous discussions. I am also indebted to Matthew Rupp for the analysis shown in Figs. 5 and 6. This work was supported in part by the National Science Foundations Frontiers in Integrative Biological Research grant FIBR-0527023, a Templeton Foundation research grant, and DARPA’s FunBio initiative.

## Bibliography

1. Lovejoy AO (1936) The great chain of being: A study of the history of the idea. Harvard University Press, Cambridge
2. Gould S (1996) Full house: The spread of excellence from Plato to Darwin. Harmony Books, New York
3. Nee S (2005) The great chain of being. *Nature* 435:429
4. Gould S, Lewontin R (1979) The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proc R Soc London B* 205:581–598
5. McShea DW (1996) Metazoan complexity and evolution: Is there a trend? *Evolution* 50:477–492
6. Valentine J, Collins A, Meyer C (1994) Morphological complexity increase in metazoans. *Paleobiology*, 20:131–142
7. Bell G, Mooers A (1997) Size and complexity among multicellular organisms. *Biol J Linnean Soc* 60:345–363
8. Nehaniv CL, Rhodes JL (2000) The evolution and understanding of hierarchical complexity in biology from an algebraic perspective. *Artif Life* 6:45–67
9. McShea D (2001) The hierarchical structure of organisms: A scale and documentation of a trend in the maximum. *Paleobiology* 27:405–423
10. Szostak JW (2003) Functional information: Molecular messages. *Nature* 423:689
11. McShea DW (2000) Functional complexity in organisms: Parts as proxies. *Biol Philosoph* 15:641–668
12. Britten RJ, Davidson EH (1971) Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. *Q Rev Biol* 46:111–138
13. Cavalier-Smith T (1985) Eukaryotic gene numbers, non-coding DNA and genome size. In: Cavalier-Smith T (ed) The evolution of genome size. Wiley, New York, pp. 69–103
14. Gregory TR (2004) Macroevolution, hierarchy theory, and the c-value enigma. *Paleobiology* 30:179–202
15. Gregory TR (2005) Genome size evolution in animals. In: Gregory TR (ed) The evolution of the genome. Elsevier, San Diego, pp. 3–87
16. Badii R, Politi A (1997) Complexity: Hierarchical structures and scaling in physics, Cambridge Nonlinear Science Series, vol. 6. Cambridge University Press, Cambridge (UK)
17. Kolmogorov A (1965) Three approaches to the quantitative definition of information. *Probl Inf Transm* 1:4
18. Li M, Vitanyi P (1997) An introduction to Kolmogorov complexity and its applications. Springer, New York
19. Adami C, Cerf NJ (2000) Physical complexity of symbolic sequences. *Physica D* 137:62–69



20. Gell-Mann M, Lloyd S (1996) Information measures, effective complexity, and total information. *Complexity* 2:44–52
21. Shannon C, Weaver W (1949) *The mathematical theory of communication*. University of Illinois Press, Urbana
22. Quastler H (ed) (1953) *Information theory in biology*. University of Illinois Press, Urbana
23. Gatlin L (1972) *Information theory and the living system*. Columbia University Press, New York
24. Mantegna RN, Buldyrev SV, Goldberger AL, Havlin S, Peng CK, et al (1994) Linguistic features of noncoding DNA sequences. *Phys Rev Lett* 73:3169–3172
25. Schmitt AO, Herzel H (1997) Estimating the entropy of DNA sequences. *J Theor Biol* 188:369–377
26. Weiss O, Jimenez-Montaña MA, Herzel H (2000) Information content of protein sequences. *J theor Biol*, 206:379–386
27. Herzel H, Ebeling W, Schmitt AO (1994) Entropy of biosequences: The role of repeats. *Phys Rev E* 50:5061–5071
28. Shannon C (1948) A mathematical theory of communication. *Bell Syst Tech J* 27:379–423, 623–656
29. MacKay DJC (2002) *Information theory, inference and learning algorithms*. Cambridge University Press, Cambridge
30. Adami C (2004) Information theory in molecular biology. *Phys Life Rev* 1:3–22
31. Grassberger P (1986) Toward a quantitative theory of self-generated complexity. *Int J Theor Phys* 25:907–938
32. Bernaola-Galvan P, Roman-Roldan R, Oliver J (1996) Compositional segmentation and long-range fractal correlations in DNA sequences. *Phys Rev E* 53:5181–5189
33. Sprinzl M, Horn C, Brown M, Ioudovitch A, Steinberg S (1998) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res*, 26:148–153
34. Eddy SR, Durbin R (1994) RNA sequence analysis using covariance models. *Nucl Acids Res*, 22:2079–2088
35. Korber BT, Farber RM, Wolpert DH, Lapedes AS (1993) Covariation of mutations in the V3 loop of human immunodeficiency virus type 1 envelope protein: an information theoretic analysis. *Proc Natl Acad Sci USA* 90:7176–7180
36. Clarke ND (1995) Covariation of residues in the homeodomain sequence family. *Protein Sci*, 4:2269–2278
37. Atchley WR, Wollenberg KR, Fitch WM, Terhalle W, Dress AW (2000) Correlations among amino acid sites in bhlh protein domains: an information theoretic analysis. *Mol Biol Evol* 17:164–178
38. Wang LY (2005) Covariation analysis of local amino acid sequences in recurrent protein local structures. *J Bioinform Comput Biol* 3:1391–1409
39. Swofford DL, Olsen GJ, Waddell PJ, Hillis DM (1996) Phylogenetic inference. In: Hillis DM, Moritz C, Mable BK (eds) *Molecular systematics*, 2nd edn, Sinauer, Sunderland, pp. 407–514
40. Wolf JB, Brodie III ED, Wade MJ (eds) (2000) *Epistasis and the evolutionary process*. Oxford University Press, Oxford
41. Bridgham JT, Carroll SM, Thornton JW (2006) Evolution of hormone-receptor complexity by molecular exploitation. *Science* 312:97–101
42. Cowperthwaite MC, Bull JJ, Ancel Meyers L (2006) From bad to good: Fitness reversals and the ascent of deleterious mutations. *PLoS Comput Biol* 2:e141
43. Finn RD et al (2006) Pfam: Clans, web tools and services. *Nucleic Acids Res* 34:D247–D251
44. Brenner SE, Chothia C, Hubbard TJP (1998) Assessing sequence comparison methods with reliable structurally identified distant evolutionary relationships. *Proc Natl Acad Sci USA*, 95:6073–6078
45. Miller GA, Madow WG (1954) On the maximum likelihood estimate of the Shannon-Wiener measure of information. Technical Report 54–75, Air Force Cambridge Research Center, Bedford
46. Basharin GP (1959) On a statistical estimate for the entropy of a sequence of independent random variables. *Theory Probab Appl* 4:333–337
47. Zurek WH (1990) Algorithmic information content, Church-Turing thesis, physical entropy, and Maxwell's demon. In: Zurek WH (ed) *Complexity, entropy, and the physics of information*. SFI Studies in the Sciences of Complexity, vol. 8 Addison-Wesley, Redwood City pp. 73–89
48. Cover TM, Thomas JA (1991) *Elements of Information Theory*. John Wiley, New York
49. Adami C (1998) *Introduction to Artificial Life*. Springer, New York
50. Adami C (2006) Digital genetics: Unravelling the genetic basis of evolution. *Nat Rev Genet* 7:109–118
51. Adami C, Ofria C, Collier T (1999) Evolution of biological complexity. *Proc Natl Acad Sci USA* 97:4463–4468
52. Ofria C, Huang W, Torng E (2008) On the gradual evolution of complexity and the sudden emergence of complex features. *Artif Life* 14, to appear
53. Carothers JM, Oestreich SC, Davis JH, Szostak JW (2004) Informational complexity and functional activity of RNA structures. *J Amer Chem Soc*, 126:5130–5137
54. Hazen RM, Griffin PL, Carothers JM, Szostak JW (2007) Functional information and the emergence of biocomplexity. *Proc Natl Acad Sci USA* 104:8574–8581
55. Barabasi AL, Oltvai ZN (2004) Network biology: understanding the cell's functional organization. *Nat Rev Genet* 5:101–113
56. Schlosser G, Wagner GP (eds) (2004) *Modularity in development and evolution*. University of Chicago Press, Chicago, IL
57. Callebaut W, Rasskin-Gutman D (eds) (2005) *Modularity: Understanding the development and evolution of natural complex systems*. MIT Press, Cambridge, Mass
58. Reigl M, Alon U, Chklovskii DB (2004) Search for computational modules in the *C. elegans* brain. *BMC Biol* 2:25
59. Hintze A, Adami C (2008) Evolution of complex modular biological networks. *PLoS Comput Biol* 4:e23
60. Batagelj V, Mrvar A (2003) Pajek: Analysis and visualization of large networks. In: M Jünger PM (ed) *Graph Drawing Software*. Springer, Berlin, pp. 77–103
61. Huang W, Ofria C, Torng E (2004) Measuring biological complexity in digital organisms. In: Pollack J, Bedau MA, Husbands P, Ikegami T, Watson R (eds) *Proceedings of Artificial Life IX*, MIT Press, Cambridge, pp. 315–321
62. Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. *Nature* 393:440–442
63. Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D et al. (2002) Network motifs: simple building blocks of complex networks. *Science* 298:824–827
64. Tishby N, Pereira F, Bialek W (1999) The information bottleneck method. In: Hajek B, Sreenivas RS (eds) *Proceedings of the 37th Annual Allerton Conference on Communication, Control and Computing*, University of Illinois Press, Champaign, IL, pp. 368–377
65. Ziv E, Mitterdorf M, Wiggins CH (2005) Information-theoretic approach to network modularity. *Phys Rev E* 71:046117



66. Milo R, Itzkovitz S, Kashtan N, Levitt R, Shen-Orr S et al. (2004) Superfamilies of evolved and designed networks. *Science*, 303:1538–1542
67. Zipf GK (1935) *The psycho-biology of languages*. Houghton-Mifflin, Boston
68. Shannon CE (1951) Prediction and entropy of printed English. *Bell System Tech J* 30:50–64

## Biological Data Integration and Model Building

JAMES A. EDDY<sup>1</sup>, NATHAN D. PRICE<sup>1,2,3</sup>

<sup>1</sup> Department of Bioengineering, University of Illinois, Urbana-Champaign, USA

<sup>2</sup> Department of Chemical and Biomolecular Engineering, University of Illinois, Urbana-Champaign, USA

<sup>3</sup> Institute for Genomic Biology, University of Illinois, Urbana-Champaign, USA

### Article Outline

[Glossary](#)

[Definition of the Subject](#)

[Introduction](#)

[The Challenge of Heterogeneous Data Types in Biology](#)

[Integration Through Interaction Network Representations](#)

[Model Building as Informative Data Integration](#)

[Data Integration Using Influence Network Models](#)

[Using Biochemical Network Models](#)

[Comparison of Biochemical and Statistical Network Models for Data Integration](#)

[Future Directions](#)

[Bibliography](#)

### Glossary

**Constraint-based analysis** A modeling framework based on excluding infeasible network states via environmental, physicochemical, and regulatory constraints to improve predictions of achievable cellular states and behavior.

**Data space** Multidimensional space containing all possible states of a system; this space can be reduced using defined constraints.

**Interaction network** A graph where the nodes represent biomolecules (e.g. genes) and the edges represent defined interactions between the nodes, whether they be

direct physical interactions (e.g. protein–protein binding, protein–DNA binding) or functional relationships (e.g. synthetic lethality).

**Biochemical reaction network** Collection of metabolic, signaling, or regulatory chemical reactions described in stoichiometric detail.

**Statistical inference network** A network model designed from statistical inference from large-scale biological data sets to be quantitatively predictive for novel perturbations and/or environmental conditions.

**Genome** The complete DNA nucleotide sequence in all chromosomes of an organism.

**Transcriptome** The complete set of RNA transcripts produced from an organism's genome under a particular set of conditions.

**Proteome** The complete set of expressed proteins produced by the genome.

**Metabolome** The complete set of small molecules which are the intermediates and products of an organism's metabolism.

**Boolean network** A set of  $N$  discrete-valued variables,  $\sigma_1, \sigma_2, \dots, \sigma_N$  where  $\sigma_n \in \{0, 1\}$ . To each node a set of  $k_n$  nodes,  $\sigma_{n_1}, \sigma_{n_2}, \dots, \sigma_{n_{k_n}}$  is assigned, which controls the value of  $\sigma_n$  through the equation  $\sigma_n(t+1) = f_n(\sigma_{n_1}(t), \dots, \sigma_{n_{k_n}}(t))$ . In the case of Boolean networks, the functions  $f_n$  can be chosen from the ensemble of all possible Boolean functions.

### Definition of the Subject

Data integration and model building have become essential activities in biological research as technological advancements continue to empower the measurement of biological data of increasing diversity and scale. High-throughput technologies provide a wealth of global data sets (e.g. genomics, transcriptomics, proteomics, metabolomics), and the challenge becomes how to integrate this data to maximize the amount of useful biological information that can be extracted. Integrating biological data is important and challenging because of the nature of biology. Biological systems have evolved over the course of billions of years, and in that time biological mechanisms have become very diverse, with molecular machines of intricate detail. Thus, while there are certainly great general scientific principles to be distilled – such as the foundational evolutionary theory – much of biology is found in the details of these evolved systems. This emphasis on the details of systems and the history by which they came into being (i.e. evolution) are distinct features of biology as a science, and influence the need for large-scale data integration. Also, biological systems are responsive to vary-

ing environments, with potential system states influenced by the combinatorics of all possible molecular and environmental perturbations. Thus, data space in this realm is extraordinarily large. There is no shortage of possibilities to explore, and vast amounts of data will be needed to fully reengineer biological systems and understand their workings at a (near) complete level of detail or at a level where accurate prediction can be achieved.

Another reason data integration is essential, is that biology arises through the complex interworking of many components. Thus, to understand biology we must integrate biological data in a way that will allow us, not only to access all our acquired data efficiently, but even more importantly allow us to study, predict, modify, and even engineer the emergent properties of biological systems, where the whole is more than the sum of the parts. As Linus Pauling stated, “Life is a relationship among molecules and not a property of any molecule.” Thus data integration and model building to study the emergent properties of biomolecules operating in concert is a pursuit fundamental to the study of biology.

## Introduction

With the emergence of full genome sequencing capabilities in the mid-nineties - marked by the first sequencing of a free-living organism, *Haemophilus influenzae* [77] – the stage was set to greatly expand the scope of data integration through computational models in biology. While modeling in biology certainly predates this era (for review, see e.g. [86]), the availability of known genome sequences has enabled the enumeration of the components of an entire cell to become a potentially definable goal. While it is still true that no cell’s component ‘parts list’ is fully defined, the existence of genome sequences allows a much more complete view, and, importantly, an indication of the amount of genomic and gene product information that is yet to be explored. In this way, genome sequencing has catalyzed the reconstruction of genome-scale models, with an eye towards the eventual creation of whole-cell simulators and beyond.

The creation of models of the integrated functions of genes and proteins in cells is of fundamental and immediate importance to the emerging field of systems biology, and represents a powerful means for integrating biological data to aid in understanding emergent properties of biological systems. Some of the most successful attempts at cell-scale modeling to date have been based on reconstructing networks that represent hundreds to thousands of experimentally-determined biochemical interactions, while others have been very successful at inferring

statistical networks from large amounts of high-throughput data. These biological networks (e.g. metabolic, regulatory, or signaling) can be analyzed, and predictions about cellular behavior under a variety of conditions and perturbations can be made and tested. Many types of models have been built and applied to study cellular behavior. In this review we focus on two broad types: biochemical network models and statistical inference models. Through iterative model prediction, experimentation, and network refinement, the molecular circuitry and functions of biological networks can be elucidated. The construction of genome-scale models that integrate data about the myriad components that produce cellular behavior is a fundamental goal of systems biology research today.

## The Challenge of Heterogeneous Data Types in Biology

Data-space is highly varied in biology for a number of reasons deriving from the nature of biological systems and the heterogeneity of methods that are used to probe different aspects of these systems. To begin thinking about the heterogeneity of biological data, we may first consider the numerous classes of biological molecules in the cell, such as the genome, messenger RNAs, proteins, protein complexes, metabolites and so forth. The global measurement of each of these types of data leads to the study of genomics, transcriptomics, proteomics, metabolomics and so forth. Beyond these biomolecule classifications, heterogeneity in biological data is increased further because of the many forms in which each of the biomolecules in these global data can exist. These modifications give rise to difficulties in characterizing precisely what constitutes a gene or a protein for analysis purposes. For example, a protein can take on many different forms, it may be phosphorylated or not, sites on the protein may be glycosylated or not, it may have different shapes into which it folds depending on its microenvironment, and so forth. In one sense then, one can imagine that integration databases require data on a protein (or a gene) to consist of a packet of information defining both the range of chemical entities defined under the umbrella of the protein (or gene) name, as well as a description of the states in which it is known to be able to exist. Also, the environmental context can greatly affect the values measured for (mostly) all of these data types. Thus, there seem to be confounding factors, uncertainties, and combinatorics at virtually every level of attempting to describe biology – a daunting challenge for data integration moving forward.

Another key factor in defining the heterogeneity of biological data is that the different chemistries of the vari-

ous molecules means that our ability to measure them is often based on very different technologies, with very different capacities for accuracy and global coverage. Also, the history of the development of the technology can also play a role, in that the typical pattern for successful emerging technologies is to increase their capacities logarithmically over time, as has been seen for genomic and transcriptomic data as well as, famously, for computing power (for review of this phenomenon in many fields, see [45]). Thus, factors such as the rate at which these technologies are improving as well as the technical challenges involved in their measurement means that for different data types, we are often at very different stages of our ability to measure them accurately and globally. For example, because of the complementary nature of RNA molecules, we can use template chemistry, as well as amplification of signal using PCR, to enable global measurements. Because of this chemistry, and a history of development, these technologies are now quite mature – i.e. they are available commercially and their accuracy and global coverage are benefiting from the economies of scale and a competitive economic environment. Thus, researchers in the field can currently generate quite accurate data for transcriptomes compared to what is currently possible for proteomics or metabolomics. Genome sequencing has reached such an advanced state that it is becoming possible to sequence individual copies of genomes. Indeed, this progress can be seen in the notable sequencing of the individual genomes of both James Watson and J. Craig Venter [48]. Emerging technologies will enable sequencing capabilities not just from single individuals, but even single cells.

That data space in biology is varied not only because of the nature of biological systems, but also because of variations in our ability to measure different components of biological systems, is a critical issue because what we are really interested in studying is variation in biology. In a sense, the interaction between the biological system and the measurement technology leads to a type of biological ‘uncertainty principle’ whereby the process of measurement typically involves significant perturbation to the system. Thus, it is important to try to distinguish between variation in biological data that occurs because of measurement technologies and that which occurs because of underlying biology. This task is a fundamental challenge to biological data integration. With any given measurement technology, repeated measurements give rise to estimates of the ‘technical variance’, while biological replicates are used to study differences due to subtle changes affecting the biology of the organisms. It is thus a more challenging endeavor to develop methods that account for the different accuracies of measuring heterogeneous data while

attempting to integrate genomics, proteomics, transcriptomics and their interactions into cohesive networks. One example of a methods arising to meet this challenge is the *Pointillist* software package, which enables the reconstruction of interaction networks using p-value-based methods for integrating various types of biological data [33,34].

Another factor to consider is that, even within each of these different data types, many hierarchies of information are contained. For example, the core information of the genome is typically represented as a linear sequence of base pairs (A,T,C,G), since this sequence does contain a tremendous amount of biological information. Even just within the linear sequence of genomic information are different levels of information. For example, many analyses focus on quantizing the genome into genes and focusing on the properties of the resulting gene products. However, the position in the linear sequence of the genome effects genes – their functioning is not independent of their position. Of course, all of these molecules exist in three dimensional space, and thus their three dimensional structure and the ways in which those three dimensional structures interact with each other is at the heart of biology. However, most network techniques in use today do not account for three dimensional structure. While there is much work in computational biology focused on three dimensional structures in, for example, the problem of protein folding based on genomic sequence, these endeavors are typically limited in the number of proteins that they can consider. Thus, three-dimensional interaction modeling is not feasible today for the simulation of cells or large networks, leaving such detailed whole cell simulators out of reach, based both on availability of data and on computational power. Simulations of large-scale networks using the three dimensional properties of molecules, including their chemical interactions as well as their biomechanical properties, remains a very large challenge yet to be met in systems biology.

Finally, another critical aspect to biological data integration is that most biological data types do not translate easily from one form to the next. For example, although transcribed genes yield messenger RNAs which are then converted into proteins, the correlation between the measured transcriptome and the measured proteome is typically weak. This fact means that the inferences that can accurately be made about protein concentrations from transcriptome measurements are very limited (although done quite often in practice). This lack of correlation can be understood, given that there is a time delay in the translation of protein from mRNA, that proteins have different degradation rates, and that protein synthesis and degradation are regulated processes. One could in principle model the

relationship between gene expression and protein concentration accurately if one had access to sufficient data (e. g. protein degradation rates, chemical conversion pathways, tRNA binding rates, etc.), but generating the data at the level of detail required for such approaches is currently prohibitive at the genome-scale. Thus, researchers in the field seek to identify shorter and less data-intensive paths to gain the ability to correlate and predict across such heterogeneous data sets.

### Integration Through Interaction Network Representations

Technological advancements in the post-genome era have enabled high-throughput screens for protein-protein, protein-DNA, and genetic interactions. The enormous amount of data resulting from these experimental studies has encouraged the development of methods to functionally characterize and visualize the interactions between components using interaction networks. Interaction networks are graphs where the nodes represent biomolecules and the edges represent defined relationships between the nodes. Methods have been developed to integrate interaction network data with complementary high-throughput data, such as from gene expression microarrays, providing insight into the relation of network interactions to cellular states under different conditions. Interaction networks can serve as databases storing a large amount of information about the components in a system and the interactions between them. Interaction network maps and databases generally serve to integrate biological data in a graph, as DNA, transcripts, proteins, and metabolites are organized into networks of direct and indirect interactions. These interaction networks are also useful for visually integrating information from genomics, proteomics, and metabolomics studies. They can be very broadly applied, and are thus used for functional relationships as well as specific binding events. For example, one can define a synthetic lethality network [82], where each node is a gene and an edge indicates that the knockout of the two genes is lethal to the cell whereas each individual knockout is not. In addition to serving as databases and means to allow for visualization of a large amount of data, these interaction networks can be studied to reveal the global topological properties of these networks.

Analysis of large and complex systems biology datasets requires automated principled methods based on robust statistical procedures. One important method that has been developed recently is an integrated framework, named *Pointillist* [33,34], for combining diverse datasets and inferential algorithms, which are incorporated into

*Cytoscape* [73] for visualization and simulation. *Cytoscape* is a powerful tool for visualizing and integrating interaction network data with gene expression and other types of state data [38,73]. In addition to its strong visualization capabilities, *Cytoscape* has the potential to be expanded using plug-ins to facilitate a number of analyses on the topological features of the network, as well as on how high-throughput data maps onto the network. For example, *Cytoscape* implements methods of searching interaction networks for connected sets of differentially expressed genes. These methods also examine which conditions significantly affect gene expression in each set (also referred to as active subnetworks). A statistical scoring system is used to quantify the gene expression change in each active subnetwork. The integration methodology of *Pointillist* is able to handle data of different types and sizes, yielding a combined p-value from multiple data sources for the existence of each link in an interaction network. A novel aspect of this methodology is that it does not require a 'gold standard' set of data to be used for training nor does it make assumptions about the underlying statistical distributions in the form of parametric models. This involves designing an efficient deterministic optimization algorithm to minimize the numbers of misses and false positives via an iterative element by element procedure. The methodology is general-purpose so that it can be applied to integrate data from any existing and future technologies. Such approaches are critical to the improved reconstruction of high-quality interaction networks from high-throughput data sources.

Protein interaction networks have also been studied from an evolutionary standpoint, comparing networks of different species to identify conserved components and to yield insight into the functional importance of selected proteins, complexes, and pathways. *PathBLAST* is a network alignment and search tool used to compare protein interaction networks across species [41], providing a framework for addressing some of the challenges in studying interaction networks. For example, common signals present in both networks are reinforced while independent noise from false-positive reactions is reduced. Additionally, these comparisons can systematically catalog conserved network regions arising from shared homologous components. Similar to using *BLAST* to align protein sequences, *PathBLAST* is used to rapidly align protein interaction networks by searching for high-scoring pathway alignments. Each alignment score is a product of independent probabilities, where the probability of each protein pair is based on the *BLAST* *E*-value of sequence alignment and the protein interaction probability is based on false-positive rates associated with interactions in the target net-

work. Score distributions cannot currently be calculated within acceptable time frames, but it is expected that as more organisms become available this approach will become increasingly powerful. Other methods have focused on using the existing protein interaction network in one organism to predict the network in a different organism.

Interaction networks are an informative and powerful way to visualize large amounts of biological data. They also serve as a basis from which to study the topological properties of biological networks [25,29,36,38,70,78]. However, they fall short in terms of explaining or predicting quantitatively the emergent dynamic behaviors of these component interactions. Thus, while interaction networks are important inputs for building models, they are not in themselves quantitative systems models. For example, they are generally insufficient to predict how a perturbation to a particular node (e.g. gene) in the network will affect other nodes, as will the statistical and biochemical network models that will be the focus of the rest of this article. While interaction networks by themselves are not predictive of states and functions, they do provide insight into evolutionary relationships as well as provide a basis for further experimental investigation. Interaction networks are highly useful for differentiating direct and indirect effects that are found from statistical inference, and also provide essential input information for reconstructing statistical inference networks of any large scale because they help to constrain the number of combinations that need to be evaluated (as will be discussed more below).

### Model Building as Informative Data Integration

The challenge of biological data integration begins with high-quality relational databases to house this information in an accessible manner [13,85]. However, it does not end there. While extremely useful, building these large databases does not by itself provide the ability to quantitatively integrate data to predict the emergent properties of interacting systems of components. Such highly-informative data integration, however, is possible, through the use of quantitative computational models.

As discussed above, high throughput technologies have initiated a phase transition in biology, making possible the study of large scale biological systems as a whole. Biology has become an information rich science, where the limiting step in biological discovery is now often not data generation but data analysis. Many in the scientific community have responded to this transformation by suggesting that the fundamental way in which much biology is done must change in order to best harness the large amount of biological data becoming avail-

able [1,31,37,43,44,]. Harnessing this information is critical to furthering our ultimate goal of improving human health and understanding the mechanisms underlying life. In response to these opportunities, three priorities have arisen in recent years from discussions at the National Institutes of Health: (1) the integration of heterogeneous biological data sets; (2) the integration of experiment and computations; and (3) the development of methods for systems analysis in biology. All of these high priority objectives for biological data integration can be met through the creation of quantitative models of biological systems – many of which can be expanded to the genome-scale.

Quantitative models provide for the integration of heterogeneous data sets in mathematically precise terms. Within the context of a cell-scale model, various data types interact in precisely defined ways such that predictions of cellular behavior can be made. These precise interactions are represented through defined mathematical relationships that mirror the underlying physical mechanisms to varying degrees of detail. Thus, quantitative models provide a most compact and powerful representation of diverse data sets.

Quantitative models themselves also serve as detailed hypotheses of how cells function, and thus constructing these models leads naturally to an integral connection between computational and experimental programs. When a quantitative model is constructed, the relationships between various aspects of a cell are related through mathematics in such a way as to reflect a physical process. Failure modes of these models point towards gaps in understanding the basic physical processes that underlie how cellular processes work. Thus, predictions from quantitative models, when found to be incorrect, lead to directed lines of questioning from which understanding of the cell's workings can be increased in a systematic fashion. Constructing and revising these models based on data forms a powerful iterative model building approach that can drive robust experimental programs.

Finally, the quest to create reliable cell-scale models provides a fertile ground for the development of systems analysis methods in biology. Developing the capacity to correctly model each new facet of cellular function provides its own set of challenges and requires the development of methods for its incorporation into the existing theoretical framework of a developing whole cell model. In this way, many modeling approaches are developed based upon the specific questions being asked and the problem being studied. Historically, mathematics have been developed primarily from the necessity to solve physics problems. Now, as biology becomes more and more amenable to mathematical analysis, it is likely that new mathematics



and new computational tools will be developed in order to meet the specific needs of modeling biology [10]. One example of this is the constraint-based modeling philosophy that will be discussed below, which has arisen to model biological systems because of the lack of complete information about the systems being studied and the inherent adaptability of living systems [55,67]. Creating increasingly accurate and detailed whole-cell models promises to be a rewarding effort that will result in the development of novel modeling methods.

The creation of cell-scale models is thus a fundamental and immediate goal of systems biology today and a powerful means for integrating large-scale data sets into an overarching cohesive framework. Such models are of great importance because they allow for the quantitative statement of a hypothesis of how a cell functions. Some disagreement may exist about what exactly constitutes a whole-cell model, but ideally one would like to be able to: 1) represent the multiple ways cells can accomplish the same task (redundancy); 2) model their dynamic and steady state behavior; 3) elucidate differences in pathological and normal states; 4) lead directly towards treatments for disease (where applicable); 5) provide mechanistic description (not just statistical outcome); 6) be specific for a personal genotype to phenotype relationship; 7) remain valid under a wide range of environmental conditions; 8) scale well such that models of relevant size are computable in reasonable length of time; 9) predict cellular behavior resulting from changes in gene content and the environment (preferably should not require fitting of parameters for each condition); 10) represent cells with models that are simpler than the cell itself. Additional and/or modified criteria could no doubt be stated. Suffice it to say, however, that models that satisfy all of these criteria do not yet exist, although models that meet some of these criteria are becoming available for various cellular processes.

As described above, quantitative models integrate heterogeneous data sets, combine experiment and computation, and drive the development of methods of systems analysis in biology. Thus, as is widely recognized, the pursuit of the goal of creating quantitative models is fundamental to the development of biology as a whole. Reliable cell-scale simulations hold the long-term promise of significantly increasing our ability to rationally design and modify biological systems. This power is fundamental to such applications as treating disease, developing new drugs, and using biosystems to improve our environment and create useful products.

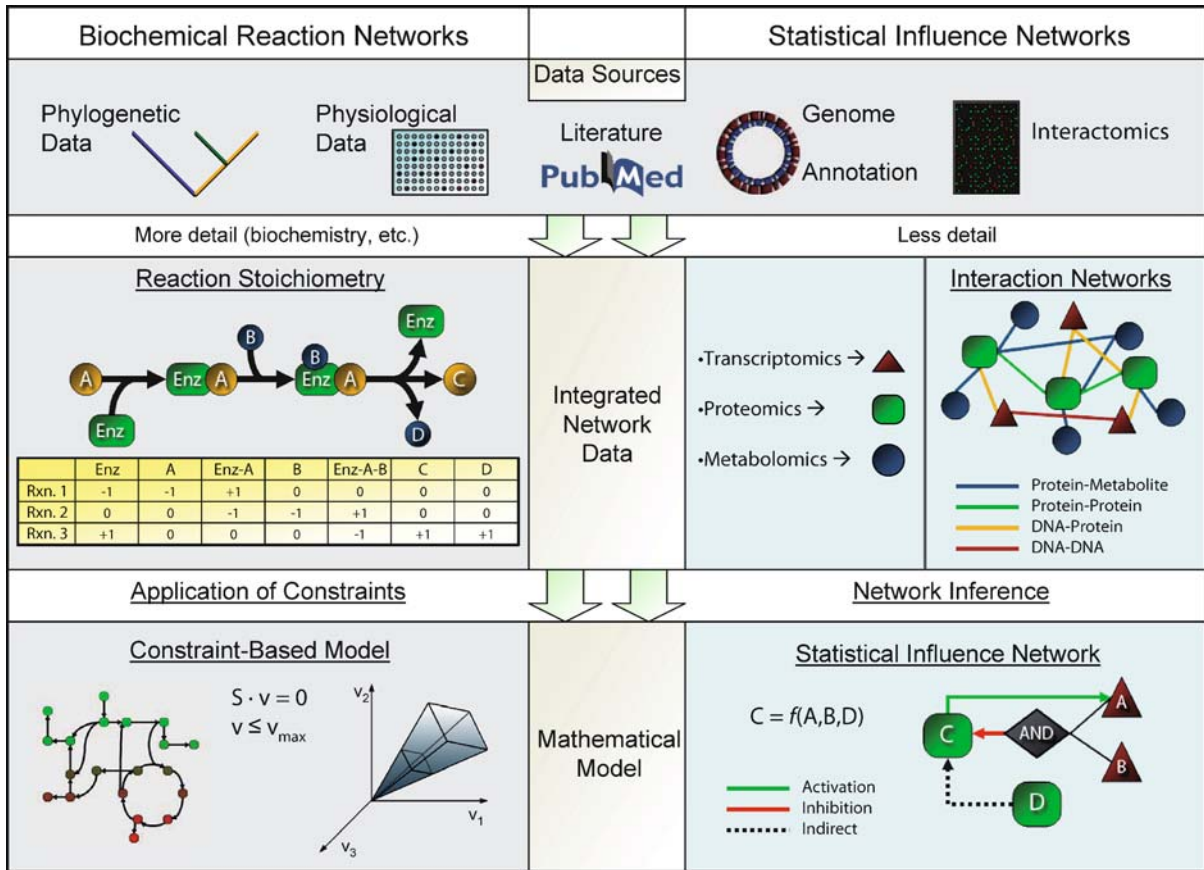
For the purposes of this article, we will discuss two general modeling approaches, based on network models of two different kinds, namely, biochemical reaction net-

works, and statistical inference networks (Fig. 1). These two types of quantitative models represent fundamentally different approaches to network reconstruction and typically use quite different mathematical techniques. However, they share the same fundamental goal: namely, the quantitative prediction of the behavior of a biological system under a wide range of conditions and perturbations. Each of these modeling approaches will be discussed in detail in the following sections.

### Data Integration Using Influence Network Models

One approach to modeling biological networks that holds tremendous potential for advancing knowledge of biology is the use of statistical associations in high-throughput data types to infer predictive network models. These statistical inference networks differ from interaction networks in that the connections between different network components is accompanied by mathematical functions precisely relating how perturbations to one (or a set of) component(s) effects the others. In essence, statistical inference networks are predictive and computable, while interaction networks as a rule are not. Interaction networks can, however, be an important – even essential – information source for the reconstruction of predictive statistical inference networks. The growing amounts of high-quality microarray gene expression data, coupled with input from gene-protein interaction networks, are now making it possible to make robust statistical inferences about the structure and dynamics of biomolecular control systems, such as transcriptional regulatory networks. Many approaches attempt to infer relationships between gene expression measurements using deterministic or stochastic formalisms. The fundamental idea behind these approaches is that models that faithfully capture such relationships have predictive capacity as regards system behavior and can be used to gain insight about system-wide properties, such as steady-state behavior or responses to perturbations or specific stimuli. There are a number of ways in which such relationships can be represented, both in the discrete and continuous domains.

One popular modeling approach that captures the nonlinear multivariate relationships exhibited by biological control circuits, such as gene regulatory networks, is the class of Boolean networks, which owes its inception to the work of Stuart Kauffman in the late 1960s [40]. In the Boolean network model, the variables (e. g., genes or proteins) are binary-valued, meaning that their states can be either on or off, and the relationships between the variables are captured by Boolean functions. Each (target) gene is assigned a Boolean rule that determines its value as a func-



Biological Data Integration and Model Building, Figure 1

Data integration and model building through biochemical reaction networks and statistical influence networks

tion of the values of a set of other (predictor) genes, possibly including the target gene itself. System dynamics are generated by updating the Boolean functions, either synchronously or asynchronously, causing the system to transition from state to state in accordance with its Boolean update rules, where a state is a binary representation of the activities of all of the variables in the system (i. e., a binary vector representing the genes that are on or off at any given time).

Boolean network models have been constructed and analyzed for a number of developmental and physiological processes. For example, Albert et al. constructed a Boolean network model for a subset of genes of the fruit-fly *Drosophila melanogaster*, which describes different stable gene expression patterns in the segmentation process of the developing embryo [1]. The steady-state behavior of this model was in excellent agreement with experimentally observed expression patterns under wild type and several gene mutation conditions. This study highlighted the importance of the network topology in determining bi-

ologically correct asymptotic states of the system. Indeed, when the segment polarity gene control network was modeled with more detailed kinetic models, such as systems of nonlinear differential equations, exceptional robustness to changes in the kinetic parameters was observed [83].

Boolean networks have also been used to model the yeast and mammalian cell cycle [20,49]. Li et al. demonstrated that the cell cycle sequence of protein states, which is a globally attracting trajectory of the dynamics, is extremely robust with respect to small perturbations to the network. The Boolean network formalism was also recently used to model systems-level regulation of the host immune response, which resulted in experimentally validated predictions regarding cytokine regulation and the effects of perturbations [79]. Boolean rules can be learned from gene expression data using methods from computational learning theory [47] and statistical signal processing [74,75,76].

A limitation of the Boolean network approach is its inherent determinism. Because of the inherent stochasticity

of gene expression and the uncertainty associated with the measurement process due to experimental noise and possible interacting latent variables (e.g. protein concentrations or activation states that are not measured), the inference of a single deterministic function may result in poor predictive accuracy, particularly in the context of small sample sizes (e.g., number of microarrays) relative to the number of genes.

One approach to “absorb” this uncertainty is to infer a number of simple functions (having few variables), each of which performs relatively well, and probabilistically synthesize them into a stochastic model, called a probabilistic Boolean network (PBN) [75]. The contribution of each function is proportional to its determinative potential as captured by statistical measures such as the coefficient of determination, which are estimated from the data [76]. The dynamical behavior of PBNs can be studied using the theory of Markov chains, which allows the determination of steady-state behavior as well as systematic intervention and control strategies designed to alter system behavior in a specified manner [77]. The PBN formalism has been used to construct networks in the context of several cancer studies, including glioma [26], melanoma [54], and leukemia [50]. PBNs, which are stochastic rule-based models, bear a close relationship to dynamic Bayesian networks [46] – a popular model class for representing the dynamics of gene expression.

Bayesian networks are graphical models that have been used to represent conditional dependencies and independencies among the variables corresponding to gene expression measurements [23]. One limitation of Bayesian networks for modeling genetic networks is that these models must be in the form of directed acyclic graphs and, as such, are not able to represent feedback control mechanisms. Dynamic Bayesian networks, on the other hand, are Bayesian networks that are capable of representing temporal processes [42,87] that may include such feedback loops. Since not all causal relationships can be inferred from correlation data, meaning that there can be different directed graphs that explain the data equally well, intervention experiments where genes are manipulated by overexpression or deletion have been proposed to learn networks [62]. The Bayesian network formalism has also been used to infer signaling networks from multicolor flow cytometry data [71].

There exist a number of other approaches for inferring large-scale molecular regulatory networks from high-throughput data sets. One example is a method, called the Inferelator, that selects the most likely regulators of a given gene using a nonlinear model that can incorporate combinatorial nonlinear influences of a regulator on target gene

expression, coupled with a sparse regression approach to avoid overfitting [69]. In order to constrain the network inference, the Inferelator performs a preprocessing step of biclustering using the cMonkey algorithm [69], which results in a reduction of dimensionality and places the inferred interactions into experiment-specific contexts. The authors used this approach to construct a model of transcriptional regulation in *Halobacterium* that relates 80 transcription factors to 500 predicted gene targets.

Another method that predicts functional associations among genes by extracting statistical dependencies between gene expression measurements is the ARACNe algorithm [52]. This information-theoretic method uses a pairwise mutual information criterion across gene expression profiles to determine significant interactions. A key step in the method is the use of the so-called data processing inequality, which is intended to eliminate indirect relationships in which two genes are co-regulated through one or more intermediaries. Thus, the relationships in the final reconstructed network are more likely to represent the direct regulatory interactions. The ARACNe algorithm was applied to 336 genome-wide expression profiles of human B cells, resulting in the identification of MYC as a major regulatory hub along with newly identified and validated MYC targets [4].

A method related to the ARACNe algorithm, called the context likelihood of relatedness (CLR), also uses the mutual information measure but applies an adaptive background correction step to eliminate false correlations and indirect influences [18]. CLR was applied to a compendium of 445 *E. coli* microarray experiments collected under various conditions and compared to other inference algorithms on the same data set. The CLR algorithm had superior performance as compared to the other algorithms, which included Bayesian networks and ARACNe, when tested against experimentally determined interactions curated in the RegulonDB database. It also identified many novel interactions, a number of which were verified with chromatin immunoprecipitation [18].

### Data Integration Using Biochemical Network Models

When the necessary biochemical detail is known, we can reconstruct biochemical reaction networks. These networks represent the underlying chemistry of the system, and thus at a minimum represent stoichiometric relationships between inter-converted biomolecules. Biochemical reaction networks are thus directly based on chemistry, rather than a reflection of statistical associations, as are statistical inference networks. These stoichiomet-

ric reconstructions have been most commonly applied to small-molecule interconversions in metabolic networks, but this formalism can easily incorporate biological transformations of all types, including for metabolic, signaling, protein–protein interactions, and gene regulatory networks [55]. The stoichiometry of biochemical reaction networks can now be reconstructed at the genome-scale, and at smaller scale with sufficient detail to generate kinetic models. These biochemical reaction networks represent many years of accumulated experimental data and can be interrogated *in silico* to determine their functional states. Genome-scale models based on biochemical networks provide a comprehensive, yet concise, description of cellular functions.

For metabolism, the reconstruction of the biochemical reaction network is a well-established procedure [14,22,27,29,66,68,81], while methods for the reconstruction of the associated regulatory [11,24] and signaling networks [56,57,58] with stoichiometric detail are developing. In any case, the reconstruction process is typically labor-intensive, involving massive amounts of data collection, biochemical characterization, and concurrent validation. A metabolic network reconstruction begins by collecting the list of enzymatic genes for the organism of interest. Using online databases and published literature, the predicted products and functions of each gene are recorded. It is often advantageous to model an organism whose annotated genome is publicly available. *BLAST* searches and other bioinformatics approaches seeking specifically to uncover functions the cell must have the capacity to carry out, however, can be used to identify the functions of previously uncharacterized genes. This is a simple example of how the process of iterative network model building guides discovery of new biological hypotheses. Along with the gene and peptide properties for each enzyme in the network, the stoichiometric description for each reaction is also recorded, with care taken to ensure mass and charge balance. Each enzyme may be the product of one or multiple genes, whereas a given enzyme may catalyze one or multiple reactions. The association between genes, proteins, and reactions (GPRs) are described using Boolean language. Examples of “OR” relationships are 1) isozymes, where either one protein or the other is required for the reaction to occur; or 2) transcript variants, where at least one in a set of alternatively spliced transcripts is needed to produce the enzyme corresponding to a particular reaction. In contrast, an “AND” relationship might describe a multi-meric protein, where multiple genes are required for production, or a complex of proteins, all necessary to catalyze a given reaction. Finally, for organisms with multiple cellular compartments, the cellular localization is determined

for all enzymes and reactions. Metabolic networks can be reconstructed in a pathway-by-pathway manner, allowing a team of researchers to work in parallel to complete the model more quickly. Computational methods can be used to interrogate the properties of each pathway, or subnetwork, to evaluate the completeness of each of the submodels given the proper ‘exchange’ molecules that will be trafficked back and forth between subsystems.

Signaling network reconstructions follow the general process of metabolic networks, but are based on a slightly different set of data. These networks can again be reconstructed in a pathway-by-pathway manner, and in some cases, only one pathway may be the focus of study. While stand-alone signaling network models typically do not rely on GPR associations, the Boolean rules described above become necessary when integrating signaling and regulatory networks. Published literature is most often used, with signaling databases more scarce than metabolic, to identify the relevant signaling interactions (e. g. phosphorylation, dephosphorylation, receptor binding, poly-ubiquitination, etc.). These interactions are then represented as mass-balanced biochemical reactions. Reconstructing both metabolic and signaling pathways in the form of biochemical reactions allows for each type of network to be represented as a stoichiometric matrix, serving as an integrated database to synthesize large amounts of data about signaling networks. In the stoichiometric matrix (*S*), each row represents a molecular compound and each column represents a reaction. For metabolism, these networks are often focused on just the metabolites, where the existence of a protein that catalyzes this reaction is used to allow that reaction to be present in the network. It is also possible (and truer to the realities in the system) to represent the proteins themselves as compounds in the network, which enables the integration of proteomics, metabolomics, and flux data (Fig. 1). For signaling, as well as regulatory networks, the inclusion of the proteins as compounds is essential. Transcriptional regulatory networks are generally constructed based on experimental gene expression data, high-throughput protein–DNA interaction data from, for example, CHIP–chip experiments, as well as compilations of data from specific transcription factor binding experiments. Regulatory reactions are represented as Boolean switches, where active or inactive genes correspond to both metabolic and signaling activity. Regulatory reconstructions are most useful when coupled with either metabolic, signaling networks, or both – thus fully capturing cellular behavior under different states. The collective process of reconstructing biochemical reaction networks has been referred to as the “two-dimensional annotation” of genomes [55].



Genome-scale models can be generated using a constraint-based approach, as has been reviewed in detail elsewhere [67]. This modeling process involves a three-step procedure. First, the biochemical network is reconstructed, as described above. Second, the physico-chemical and environmental constraints under which the reconstructed network operates are represented mathematically. Physico-chemical constraints are based upon properties such as enzyme capacity, reaction stoichiometry and thermodynamics associated with reaction directionality and biochemical loops.

Constraints are represented in two fundamental forms: balances and bounds. Balance constraints dictate the conservation of quantities such as mass, energy, and momentum, as well as electroneutrality, solvent capacity, and osmotic pressure. The conservation of mass, for example, is commonly used to describe fluxes through biochemical networks at steady state (i.e. flux balance analysis). At steady-state, there is no accumulation or loss of metabolites within the network. Therefore, the rate of production of each metabolite must equal its rate of consumption. The balance of fluxes is typically represented mathematically as  $S \bullet v = 0$ , where  $v$  is a vector of fluxes through the metabolic network and  $S$  is the stoichiometric matrix containing all the reactions in the reconstructed network (Fig. 1). Bounds limit numerical ranges of individual variables and parameters such as fluxes, concentrations, and kinetic constants. For example, upper and lower limits can be applied to individual fluxes ( $v_{\min} \leq v \leq v_{\max}$ ).

Environmental constraints include nutrient availability, pH, temperature, osmolarity, and the availability of electron acceptors. These constraints are important for the quantitative analysis of biochemical network models. For example, defined media and well-documented environmental conditions are used to integrate data from various laboratories into quantitative models that are accurately descriptive and predictive. Environmental constraints can also be represented with bounds, defining the range of metabolite fluxes into and out of the model system to agree with a desired medium or condition. For example, to simulate cellular growth on a glucose-free medium, the upper and lower bounds on glucose flux into the cell would be set equal to zero ( $v_{\text{glucose}} = 0$ ).

Regulatory constraints are enforced via the Boolean rules describing GPR associations. In the case of an isozyme or multimeric protein (both “AND” relationships), the absence of one gene will prevent the production of its protein product, effectively turning off the corresponding reaction in the system. Gene knockout or deletion simulations can therefore provide further insight to the relation between sets of genes and network states. Reg-

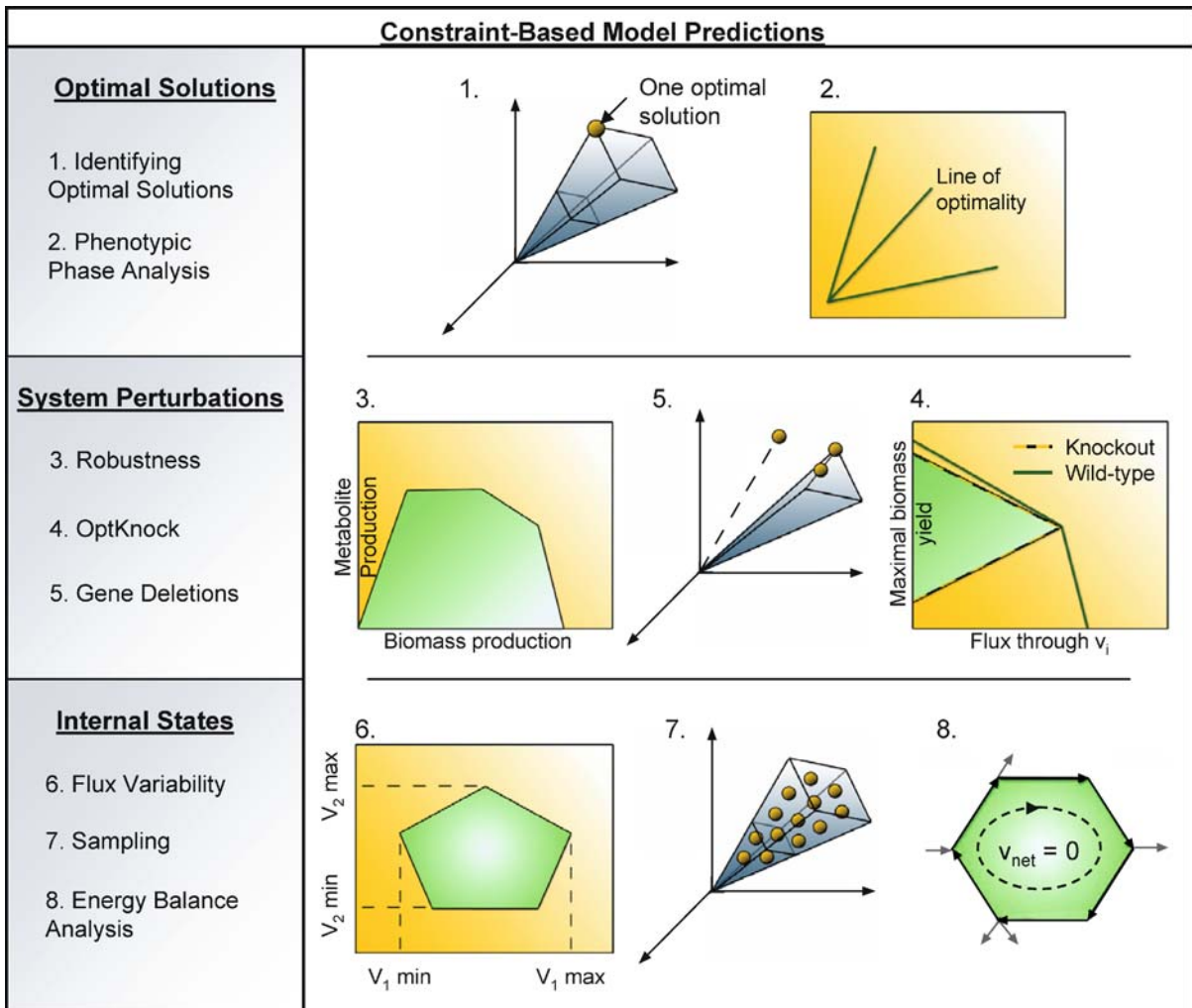
ulatory constraints also provide a link between environmental conditions and the reconstructed network; deletion tests can also be used to determine gene essentiality under various conditions. Finally, regulatory constraints are essential for network reconstructions aiming to integrate signaling and metabolic pathways.

The statement of constraints leads to the definition of a solution space that contains all non-excluded network states, describing the possible functions of the reconstructed network or all the allowable phenotypes. The third step is the determination of the possible solutions in this space that correspond to physiologically meaningful states (Fig. 2). Numerous *in silico* techniques have been developed to determine optimal network states (flux balance analysis, extreme pathways analysis, etc.) under different biological objectives (biomass, ATP production) [68]. This three-step constraint-based modeling procedure has been successfully utilized to study phenotypes in various model [3,16,21,35] and infectious [59,63] microorganisms. Recently, this approach has passed a significant milestone, namely the reconstruction of the human metabolic network [14]. Constraint-based analysis has been applied to selected human systems in the past [65,80], but with this reconstruction in place, this approach is poised to have a much more significant impact on modeling human systems.

Stoichiometric reaction networks can be expanded into dynamic models using standard differential equation models and the addition of information in the form of kinetic rate constants for network reactions. These systems, unlike those generally used in the constraint-based approach, do not make the steady-state assumption for all compounds but rather can simulate detailed dynamic behavior. Thus, they can model a greater degree of complexity for a given network and simultaneously account for both concentrations of compounds and fluxes through reactions. The disadvantage of these systems compared with constraint-based analysis is that they require many more parameters and are thus more data intensive to create and/or can be more prone to overfitting when large numbers of parameters are unknown. These models are not typically created at the whole cell scale, with a notable exception being cell-scale models of red blood cell metabolism [16,53].

In recent years, these types of dynamic models have been used very successfully to provide mechanistic models of, for example, key signaling pathways involved in critical physiological and pathophysiological processes. For example, a differential equation model of NF- $\kappa$ B signaling [30] was updated and used to elucidate the role of tumor necrosis factor in controlling the sustained phase of lipopolysac-





Biological Data Integration and Model Building, Figure 2

Representative computational predictions that can be made using constraint-based modeling of biochemical reaction networks. Example computational predictions include: 1) identifying optimal solutions, such as predicting an optimal growth rate [16], 2) defining phenotypic phase planes that can be used to interpret and predict the outcome of adaptive evolution [35], 3) robustness analyzes evaluating the effect of changing fluxes on optimal growth [17], 4) OptKnock and related methods for predicting the effect of genetic modifications on the relationship between cell growth and the production of a desired product [8,32,60,61], 5) the lethality of gene deletions [19,29], 6) the analysis of the range of allowable values for each flux in a biochemical network under a set of specified constraints [51,67], 7) monte-carlo methods for characterizing the space of potential network steady states [3,65,80], and 8) the introduction of thermodynamic constraints to more realistically constrain the range of internal biochemical fluxes [5,6].

charide-induced IKK activation [12,84], demonstrating the usefulness of these models for deciphering functions of biological networks. Another example was the use of a model to study the integrated dynamic behavior of a network of epidermal growth factor receptor family members [28]. Christopher et al. have constructed a dynamic simulator of gene expression and signaling networks in a human cancer cell [9]. In the long run, such dynamic models set the stage for personalized medicine by offering

the promise of rational drug design and control of the outcome of induced molecular perturbations.

### Comparison of Biochemical and Statistical Network Models for Data Integration

There are fundamental differences between the biochemical and statistical classes of network modeling described herein. One clear difference is the manner in which these

underlying networks are reconstructed. For biochemical networks, reconstruction is typically a work-intensive process requiring significant biochemical characterization with little network inference done (other than inferences of single gene function for catalyzing a reaction based on sequence homology). Thus, the ability to rapidly generate networks for organisms that are relatively uncharacterized from high-throughput data is an inherent advantage of the inferred statistical networks. One advantage of the biochemical network models is that, once reconstructed, the networks are not as subject to change (other than addition) since many of the links are based directly on biochemical evidence. Inferred networks, on the other hand, can undergo substantial changes in light of additional data. Another common difference, although not fundamental, is that constraint-based biochemical network models have mostly been used to model flux, whereas inference networks have mostly been used to predict substance amounts (e. g. mRNA expression). One way this can be thought of is that the biochemical network models currently link more closely to functional phenotype (i. e. fluxes) [72], while the inferred networks relate more directly to available high-throughput data (i. e. transcriptomes). The kinetic biochemical network models, of course, have the capacity to account for both flux and abundance, but suffer from the limitation that they are by far the most data intensive to reconstruct. Another key advantage of biochemical reaction networks, stemming from their basis in chemistry, is that physico-chemical laws apply, such as mass-energy balance, while such laws are not generally applicable to the inferred networks. Of course, the advantage of the inferred networks is that, since they do not need to be mechanistic or require biochemical detail, they can be applied very broadly to systems that are not yet well characterized and can link very disparate data types as long as underlying correlations exist. In summary, both modeling types are essential to contemporary computational systems biology and provide advantages over each other in different settings.

### Future Directions

The future for biological data integration through model building is bright. High throughput experimental technologies keep improving, yielding exponentially increasing sources of data. Computing power continues to increase exponentially as well, meaning that both the necessary computing power and data streams are expected to keep the need for data integration through modeling at a high and increasing premium in biological research. However, very significant challenges clearly remain in in-

tegrating biological data going forward. Even fundamental concepts such as what exactly constitutes a gene or a protein remain somewhat ambiguous, and so the definition of data sets and the rigorous data storage itself are challenges. New and improved mathematical and computational approaches will also need to be developed for accurate biological modeling, adaptive over time along with the emergence of new measurement technologies. One interesting challenge for model building going forward is whether hybrid models that take advantage of the strengths of both biochemical and statistical inference modeling approaches can be constructed to move us further towards the goal of predictive whole-cell models and beyond. Early attempts have been done to link Boolean regulatory networks with constraint-based flux models [11], but the extent to which these approaches can be married to provide significant advances in our ability to model biological networks remains an open question.

### Bibliography

1. Albert R, Othmer HG (2003) The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*. *J Theor Biol* 223(1):1–18
2. Alm E, Arkin AP (2003) Biological networks. *Curr Opin Struct Biol* 13(2):193–202
3. Almaas E, Kovacs B et al (2004) Global organization of metabolic fluxes in the bacterium *Escherichia coli*. *Nature* 427(6977):839–43
4. Basso K, Margolin AA et al (2005) Reverse engineering of regulatory networks in human B cells. *Nat Genet* 37(4):382–90
5. Beard DA, Babson E et al (2004) Thermodynamic constraints for biochemical networks. *J Theor Biol* 228(3):327–33
6. Beard DA, Liang SD et al (2002) Energy balance for analysis of complex metabolic networks. *Biophys J* 83(1):79–86
7. Bonneau R, Reiss DJ et al (2006) The Inferelator: an algorithm for learning parsimonious regulatory networks from systems-biology data sets de novo. *Genome Biol* 7(5):R36
8. Burgard AP, Pharkya P et al (2003) Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol Bioeng* 84(6):647–57
9. Christopher R, Dhiman A et al (2004) Data-driven computer simulation of human cancer cell. *Ann NY Acad Sci* 1020:132–53
10. Cohen JE (2004) Mathematics is biology's next microscope, only better; biology is mathematics' next physics, only better. *PLoS Biol* 2(12):e439
11. Covert MW, Knight EM et al (2004) Integrating high-throughput and computational data elucidates bacterial networks. *Nature* 429(6987):92–6
12. Covert MW, Leung TH et al (2005) Achieving stability of lipopolysaccharide-induced NF- $\kappa$ B activation. *Science* 309(5742):1854–1857
13. Deshpande N, Address KJ et al (2005) The RCSB Protein Data Bank: a redesigned query system and relational database

- based on the mmCIF schema. *Nucleic Acids Res (Database issue)* 33:D233–7
14. Duarte NC, Becker SA et al (2007) Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc Natl Acad Sci USA* 104(6):1777–82
  15. Duarte NC, Herrgard MJ et al (2004) Reconstruction and validation of *Saccharomyces cerevisiae* iND750, a fully compartmentalized genome-scale metabolic model. *Genome Res* 14(7):1298–309
  16. Edwards JS, Ibarra RU et al (2001) In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat Biotechnol* 19(2):125–30
  17. Edwards JS, Palsson BO (2000) Robustness analysis of the *Escherichia coli* metabolic network. *Biotechnol Prog* 16(6):927–39
  18. Faith JJ, Hayete B et al (2007) Large-scale mapping and validation of *Escherichia coli* transcriptional regulation from a compendium of expression profiles. *PLoS Biol* 5(1):e8
  19. Famili I, Forster J et al (2003) *Saccharomyces cerevisiae* phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network. *Proc Natl Acad Sci USA* 100(23):13134–13139
  20. Faure A, Naldi A et al (2006) Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. *Bioinformatics* 22(14):e124–e131
  21. Forster J, Famili I et al (2003) Large-scale evaluation of in silico gene deletions in *Saccharomyces cerevisiae*. *Omics* 7(2):193–202
  22. Francke C, Siezen RJ et al (2005) Reconstructing the metabolic network of a bacterium from its genome. *Trends Microbiol* 13(11):550–558
  23. Friedman N (2004) Inferring cellular networks using probabilistic graphical models. *Science* 303(5659):799–805
  24. Gianchandani EP, Papin JA et al (2006) Matrix formalism to describe functional states of transcriptional regulatory systems. *PLoS Comput Biol* 2(8):e101
  25. Han JD, Bertin N et al (2004) Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* 430(6995):88–93
  26. Hashimoto RF, Kim S et al (2004) Growing genetic regulatory networks from seed genes. *Bioinformatics* 20(8):1241–7
  27. Heinemann M, Kummel A et al (2005) In silico genome-scale reconstruction and validation of the *Staphylococcus aureus* metabolic network. *Biotechnol Bioeng* 92(7):850–864
  28. Hendriks BS, Wiley HS et al (2003) HER2-mediated effects on EGFR endosomal sorting: analysis of biophysical mechanisms. *Biophys J* 85(4):2732–2745
  29. Herrgard MJ, Palsson BO (2005) Untangling the web of functional and physical interactions in yeast. *J Biol* 4(2):5
  30. Hoffmann A, Levchenko A et al (2002) The I $\kappa$ B-NF- $\kappa$ B signaling module: temporal control and selective gene activation. *Science* 298(5596):1241–1245
  31. Hood L, Heath JR et al (2004) Systems biology and new technologies enable predictive and preventative medicine. *Science* 306(5696):640–643
  32. Hua Q, Joyce AR et al (2006) Metabolic analysis of adaptive evolution for in silico-designed lactate-producing strains. *Biotechnol Bioeng* 95(5):992–1002
  33. Hwang D, Rust AG et al (2005) A data integration methodology for systems biology. *Proc Natl Acad Sci USA* 102(48):17296–17301
  34. Hwang D, Smith JJ et al (2005) A data integration methodology for systems biology: experimental verification. *Proc Natl Acad Sci USA* 102(48):17302–17307
  35. Ibarra RU, Edwards JS et al (2002) *Escherichia coli* K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature* 420(6912):186–189
  36. Ideker T (2004) A systems approach to discovering signaling and regulatory pathways—or, how to digest large interaction networks into relevant pieces. *Adv Exp Med Biol* 547:21–30
  37. Ideker T, Galitski T et al (2001) A new approach to decoding life: systems biology. *Annu Rev Genomics Hum Genet* 2:343–372
  38. Ideker T, Ozier O et al (2002) Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics (Suppl)* 18(1):S233–S2340
  39. Jamshidi N, Edwards JS et al (2001) Dynamic simulation of the human red blood cell metabolic network. *Bioinformatics* 17(3):286–287
  40. Kauffman SA (1993) The origins of order : self organization and selection in evolution. Oxford University Press, New York
  41. Kelley BP, Yuan B et al (2004) PathBLAST: a tool for alignment of protein interaction networks. *Nucleic Acids Res (Web Server issue)* 32:W83–W88
  42. Kim SY, Imoto S et al (2003) Inferring gene networks from time series microarray data using dynamic Bayesian networks. *Brief Bioinform* 4(3):228–235
  43. Kirschner MW (2005) The meaning of systems biology. *Cell* 121(4):503–504
  44. Kitano H (2002) Computational systems biology. *Nature* 420(6912):206–210
  45. Kurzweil R (2005) The singularity is near: when humans transcend biology. Penguin, London
  46. Lahdesmaki H, Hautaniemi S et al (2006) Relationships between probabilistic Boolean networks and dynamic Bayesian networks as models of gene regulatory networks. *Signal Processing* 86(4):814–834
  47. Lahdesmaki H, Shmulevich I et al (2003) On Learning Gene Regulatory Networks Under the Boolean Network Model. *Machine Learning* 52(1–2):147–167
  48. Levy S, Sutton G et al (2007) The Diploid Genome Sequence of an Individual Human. *PLoS Biol* 5(10):e254
  49. Li F, Long T et al (2004) The yeast cell-cycle network is robustly designed. *Proc Natl Acad Sci USA* 101(14):4781–4786
  50. Li H, Zhan M (2006) Systematic intervention of transcription for identifying network response to disease and cellular phenotypes. *Bioinformatics* 22(1):96–102
  51. Mahadevan R, Schilling CH (2003) The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab Eng* 5(4):264–276
  52. Margolin AA, Wang K et al (2006) Reverse engineering cellular networks. *Nat Protoc* 1(2):662–671
  53. Mulquiney PJ, Kuchel PW (2003) Modelling metabolism with Mathematica, detailed examples including erythrocyte metabolism. CRC Press, Boca Raton
  54. Pal R, Datta A et al (2005) Intervention in context-sensitive probabilistic Boolean networks. *Bioinformatics* 21(7):1211–1218
  55. Palsson B (2004) Two-dimensional annotation of genomes. *Nat Biotechnol* 22(10):1218–1219
  56. Papin JA, Hunter T et al (2005) Reconstruction of cellular sig-

- nalling networks and analysis of their properties. *Nat Rev Mol Cell Biol* 6(2):99–111
57. Papin JA, Palsson BO (2004) The JAK-STAT signaling network in the human B-cell: an extreme signaling pathway analysis. *Biophys J* 87(1):37–46
  58. Papin JA, Palsson BO (2004) Topological analysis of mass-balanced signaling networks: a framework to obtain network properties including crosstalk. *J Theor Biol* 227(2):283–297
  59. Papin JA, Price ND et al (2002) The genome-scale metabolic extreme pathway structure in *Haemophilus influenzae* shows significant network redundancy. *J Theor Biol* 215(1):67–82
  60. Pharkya P, Burgard AP et al (2003) Exploring the overproduction of amino acids using the bilevel optimization framework OptKnock. *Biotechnol Bioeng* 84(7):887–899
  61. Pharkya P, Burgard AP et al (2004) OptStrain: a computational framework for redesign of microbial production systems. *Genome Res* 14(11):2367–76
  62. Pournara I, Wernisch L (2004) Reconstruction of gene networks using Bayesian learning and manipulation experiments. *Bioinformatics* 20(17):2934–2942
  63. Price ND, Papin JA et al (2002) Determination of redundancy and systems properties of the metabolic network of *Helicobacter pylori* using genome-scale extreme pathway analysis. *Genome Res* 12(5):760–769
  64. Price ND, Reed JL et al (2004) Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat Rev Microbiol* 2(11):886–897
  65. Price ND, Schellenberger J et al (2004) Uniform sampling of steady-state flux spaces: means to design experiments and to interpret enzymopathies. *Biophys J* 87(4):2172–2186
  66. Reed JL, Palsson BO (2003) Thirteen years of building constraint-based in silico models of *Escherichia coli*. *J Bacteriol* 185(9):2692–2699
  67. Reed JL, Palsson BO (2004) Genome-scale in silico models of *E. coli* have multiple equivalent phenotypic states: assessment of correlated reaction subsets that comprise network states. *Genome Res* 14(9):1797–1805
  68. Reed JL, Vo TD et al (2003) An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). *Genome Biol* 4(9):R54
  69. Reiss DJ, Baliga NS et al (2006) Integrated biclustering of heterogeneous genome-wide datasets for the inference of global regulatory networks. *BMC Bioinformatics* 7:280
  70. Rual JF, Venkatesan K et al (2005) Towards a proteome-scale map of the human protein-protein interaction network. *Nature* 437(7062):1173–1178
  71. Sachs K, Perez O et al (2005) Causal protein-signaling networks derived from multiparameter single-cell data. *Science* 308(5721):523–529
  72. Sauer U (2004) High-throughput phenomics: experimental methods for mapping fluxomes. *Curr Opin Biotechnol* 15(1):58–63
  73. Shannon P, Markiel A et al (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13(11):2498–504
  74. Shmulevich I, Dougherty ER et al (2002) Probabilistic Boolean Networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics* 18(2):261–274
  75. Shmulevich I, Dougherty ER et al (2002) From Boolean to probabilistic Boolean networks as models of genetic regulatory networks. *Proceedings of the IEEE* 90(11):1778–1792
  76. Shmulevich I, Dougherty ER et al (2002) Gene perturbation and intervention in probabilistic Boolean networks. *Bioinformatics* 18(10):1319–1331
  77. Smith HO, Tomb JF et al (1995) Frequency and distribution of DNA uptake signal sequences in the *Haemophilus influenzae* Rd genome. *Science* 269(5223):538–540
  78. Stelzl U, Worm U et al (2005) A human protein-protein interaction network: a resource for annotating the proteome. *Cell* 122(6):957–968
  79. Thakar J, Pillione M et al (2007) Modelling Systems-Level Regulation of Host Immune Responses. *PLoS Comput Biol* 3(6):e109
  80. Thiele I, Price ND et al (2005) Candidate metabolic network states in human mitochondria. Impact of diabetes, ischemia, and diet. *J Biol Chem* 280(12):11683–11695
  81. Thiele I, Vo TD et al (2005) Expanded metabolic reconstruction of *Helicobacter pylori* (iIT341 GSM/GPR): an in silico genome-scale characterization of single- and double-deletion mutants. *J Bacteriol* 187(16):5818–5830
  82. Tong AH, Lesage G et al (2004) Global mapping of the yeast genetic interaction network. *Science* 303(5659):808–813
  83. von Dassow G, Meir E et al (2000) The segment polarity network is a robust developmental module. *Nature* 406(6792):188–192
  84. Werner SL, Barken D et al (2005) Stimulus specificity of gene expression programs determined by temporal control of IKK activity. *Science* 309(5742):1857–1861
  85. Westbrook J, Feng Z et al (2002) The Protein Data Bank: unifying the archive. *Nucleic Acids Res* 30(1):245–248
  86. Westerhoff HV, Palsson BO (2004) The evolution of molecular biology into systems biology. *Nat Biotechnol* 22(10):1249–1252
  87. Zou M, Conzen SD (2005) A new dynamic Bayesian network (DBN) approach for identifying gene regulatory networks from time course microarray data. *Bioinformatics* 21(1):71–79

---

## Biological Development and Evolution, Complexity and Self-organization in

STUART A. NEWMAN<sup>1</sup>, GABOR FORGACS<sup>2</sup>

<sup>1</sup> New York Medical College, Valhalla, USA

<sup>2</sup> University of Missouri, Columbia, USA

### Article Outline

[Glossary](#)

[Definition of the Subject](#)

[Introduction](#)

[Dynamic Multistability: Cell Differentiation](#)

[Differential Adhesion: Gastrulation](#)

[Biochemical Oscillations: Segmentation](#)

[Reaction-Diffusion Mechanisms: Body Axis Formation](#)

[Evolution of Developmental Mechanisms](#)

[Future Directions](#)

[Bibliography](#)



## Glossary

**Differential adhesion** The capacity of cells to adhere to each other in a cell type-dependent manner. The strength of adhesion between two cells of type *A* typically differs from that between cells of type *B*. This may be due to differences either in the number or type of cell adhesion molecules.

**Differential gene expression** The main regulatory basis of cell differentiation. Cells of different type in a given organism carry the same set of genes (the genome) but vary in which of these genes are active, that is, expressed.

**Biochemical oscillation** The variation in the concentration of a given molecule in principle either in space or time, although typically the term is reserved for the latter.

**Reaction-diffusion mechanism** A conceptual framework for describing spatiotemporal pattern formation in a system of several interacting and diffusing chemical species.

**Canalization** An evolved property of developmental pathways that permits the robust generation of a phenotype in the face of perturbations. The perturbations can be those internal to the organism, in the form of gene mutation or developmental noise, or external to it, in the form of environmental variability.

## Definition of the Subject

Much work over the past half-century in developmental and evolutionary biology has focused on a subset of an organism's components, its genes. The hierarchical regulatory relationships among genes have been a major emphasis in studies in development, while the variation of genes has played a corresponding role in evolutionary studies. In the past decade, however, investigators have increasingly considered the part played by physical and dynamical properties of cells and tissues, and their molecular components, in producing biological characteristics over the course of ontogeny and phylogeny. Living cells and tissues encompass numerous interpenetrating multicomponent systems in which dynamical interactions among intracellular gene products, metabolites, ions, etc., and interactions between cells, directly via their adhesive surfaces, or indirectly via secreted extracellular matrix (ECM) molecules or diffusible signaling molecules ("morphogens"), generate topologically complex, three-dimensional self-perpetuating entities consisting of up to several hundred different cell types in elaborate arrangements. In this description, the systems properties that give rise to organisms and their substructures,

however, are nowhere specified in any "genetic program".

While these "generic" (i. e., common to living and non-living systems) physical and chemical-dynamic processes contribute to the production of form and pattern in all modern-day organisms, it is likely that early in the evolution of multicellular organisms such processes were even more efficacious in determining organismal forms. In particular, morphologies originally generated by physics and chemical dynamics could have provided templates for genetic evolution that stabilized and reinforced (rather than innovated) multicellular body plans and organ forms. The hierarchical genetic control of development seen in modern organisms can thus be considered an outcome of this evolutionary interplay between genetic change and generic physicochemical processes.

## Introduction

The field of developmental biology has as its major concern *embryogenesis*: the generation of fully-formed organisms from a fertilized egg, the *zygote*. Other issues in this field, organ *regeneration* and tissue *repair* in organisms that have already passed through the embryonic stages, have in common with embryogenesis three interrelated phenomena: *cell differentiation*, the production of distinct cell types, *cell pattern formation*, the generation of specific spatial arrangements of cells of different types, and *morphogenesis*, the molding and shaping of tissues [25]. The cells involved in these developmental processes and outcomes generally have the same genetic information encoded in their DNA, the *genome* of the organism, so that the different cell behaviors are largely associated with *differential gene expression*.

Because of the ubiquity and importance of differential gene expression during development, and the fact that each type of organism has its own unique genome, a highly gene-centered view of development prevailed for several decades after the discovery of DNA's capacity to encode information. In particular, development was held to be the unfolding of a "genetic program" specific to each kind of organism and organ, based on differential gene expression. This view became standard despite the fact that no convincing models had ever been presented for how genes or their products (proteins and RNA molecules) could alone build three dimensional shapes and forms, or even generate populations of cells that utilized the common pool of genetic information in different ways. By "alone" is meant without the help of physics and chemical dynamics, the scientific disciplines traditionally invoked to ex-



plain changes in shape, form, and chemical composition in nonliving material systems.

It has long been recognized, however, that biological systems, in addition to being repositories of genes, are also physicochemical systems, and that phenomena first identified in the nonliving world can also provide models for biological processes. Indeed, at the level of structure and function of biomolecules and macromolecular assemblages such as membranes and molecular motors, biology has historically drawn on ideas from chemistry and physics. More recently there has been intensified cooperation between biological and physical scientists at the level of complex biological systems, including developmental systems. The rise of systems developmental, and evolutionary-developmental, biology required the generation by physical and information scientists of theoretical tools and computational power sufficient to model systems of great complexity. But also, and equally important, has been the growing recognition by modern biologists that organisms are more than programmed expressions of their genes, and that the behavior of systems of many interacting components is neither obvious nor predictable on the basis of the behavior of their parts.

Since biology is increasingly studied at the systems level, there has also been new attention to the origination of these complex systems. In many instances, it is reasonable to assume that complexity and integration of subsystems in living organisms has evolved in the context of forms and functions that originally emerged (in evolutionary history) by simpler physicochemical means. Thus the elaborate system of balanced, antagonistic signaling interactions that keep cell metabolism homeostatic, and embryogenesis on-track, can be seen as the result of accretion, by natural selection, of stabilizing mechanisms for simpler physicochemical generative processes that would otherwise be less reliable.

In what follows we will give a fairly wide-ranging, though hardly comprehensive, survey of the physicochemical processes organizing, and presumed to have promoted the emergence of the complex systems underlying the multicellular development of animals.

### Dynamic Multistability: Cell Differentiation

The early embryos of multicellular organisms are referred to as *blastulae*. These are typically hollow clusters of several dozen to several hundred cells that arise through the cleavage of the zygote (cell divisions without volume increase of the embryo). While the zygote is “totipotent” – it has the potential to give rise to any of the more than 200 specialized cell types (e.g., bone, cartilage, the vari-

ous kinds of muscle, blood, and nerve cells) of the mature human body – its first few divisions generate cells that are “pluripotent” – capable of giving rise to only a limited range of cell types. These cells, in turn, diversify into ones with progressively limited potency, ultimately generating all the (generally unipotent) specialized cells of the body [25].

The transition from wider to narrower developmental potency is referred to as *determination*. This stage of cell specialization generally occurs with no overt change in the appearance of cells. Instead, subtle modifications, only discernable at the molecular level, set the altered cells on new and restricted developmental pathways. A later stage of cell specialization, referred to as *differentiation*, results in cells with vastly different appearances and functional modifications – electrically excitable neurons with extended processes up to a meter long, bone and cartilage cells surrounded by solid matrices, red blood cells capable of soaking up and disbursing oxygen, and so forth.

Since each cell of the organism contains an identical set of genes (except for the egg and sperm and their immediate precursors, and some cells of the immune system), a fundamental question of development is how the same genetic instructions can produce different types of cells. This question pertains to both determination and differentiation. Multicellular organisms solve the problem of specialization by activating a type-specific subset of genes in each cell type.

The *biochemical state* of a cell can be defined as the list of all the different types of molecules contained within it, along with their concentrations. The dynamical state of a cell, like that of any dynamical system, resides in a multidimensional space, the “state space”, with dimensionality equal to the number of system variables (e.g., chemical components) [96]. During the cell division cycle (i.e., the sequence of changes that produces two cells from one), also referred to as the cell cycle, the biochemical state changes periodically with time. (This, of course, assumes that cells are not undergoing differentiation). If two cells have the same complement of molecules at corresponding stages of the cell cycle, then, they can be considered to be of the same differentiated state. The cell’s biochemical state also has a spatial aspect – the concentration of a given molecule might not be uniform throughout the cell. We will discuss an example of this related to *Drosophila* early development in Sect. “[Evolution of Development Mechanisms](#)”. The state of differentiation of the cell (its *type*) can be identified with the collection of proteins it is capable of producing.

Of the estimated 25,000 human genes [76], a large proportion constitutes the “housekeeping genes,” involved in

functions common to all or most cells types. In contrast, a relatively small number of genes – possibly fewer than a thousand – specify the type of determined or differentiated cells; these genes are *developmentally regulated* (i. e., turned on and off in an embryonic stage- and embryonic position-dependent fashion) during embryogenesis. Cell type diversity is based to a great extent on sharp transitions in the dynamical state of embryonic cells, particularly with respect to their developmentally-regulated genes.

When a cell divides it inherits not just a set of genes and a particular mixture of molecular components, but also a dynamical system at a particular *dynamical state*. Dynamical states of many-component systems can be transient, stable, unstable, oscillatory, or chaotic [96]. The cell division cycle in the early stages of frog embryogenesis, for example, is thought to be controlled by a *limit cycle* oscillator [5]. A limit cycle is a continuum of dynamical states that define a stable orbit in the state space surrounding an unstable *node*, a node being a stationary (i. e., time-independent) point, or steady state, of the dynamical system [102]. The fact that cells can inherit dynamical states was demonstrated experimentally by Elowitz and Leibler [20]. These investigators used genetic engineering techniques to provide the bacterium *Escherichia coli* with a set of feedback circuits involving transcriptional repressor proteins such that a biochemical oscillator not previously found in this organism was produced. Individual cells displayed a chemical oscillation with a period longer than the cell cycle. This implied that the dynamical state of the artificial oscillator was inherited across cell generations (if it had not, no periodicity distinct from that of the cell cycle would have been observed). Because the biochemical oscillation was not tied to the cell cycle oscillation, newly divided cells in successive generations found themselves at different phases of the engineered oscillation.

The ability of cells to pass on dynamical states (and not just “informational” macromolecules such as DNA) to their progeny has important implications for developmental regulation, since continuity and stability of a cell’s biochemical identity is key to the performance of its role in a fully developed organism. Inheritance which does not depend directly on genes is called “epigenetic inheritance” [68] and the biological or biochemical states inherited in this fashion are called “epigenetic states”. Although epigenetic states can be determined by reversible chemical modifications of DNA [41], they can also represent alternative steady states of a cell’s network of active or expressed genes [46]. The ability of cells to undergo transitions among a limited number of discrete, stable epigenetic states and to propagate such decisions from one cell gener-

ation to the next is essential to the capacity of the embryo to generate diverse cell types.

All dividing cells exhibit oscillatory dynamical behavior in the subspace of the full state space whose coordinates are defined by cell cycle-related molecules. In contrast, cells exhibit alternative stable steady states in the subspace defined by molecules related to states of cell determination and differentiation. During development, a dividing cell might transmit its particular system state to each of its daughter cells, but it is also possible that some internal or external event accompanying cell division could push one or both daughter cells out of the “basin of attraction” in which the precursor cell resided and into an alternative state. (The basin of attraction of a stable node is the region of state space surrounding the node, in which all system trajectories, present in this region, terminate at that node [96].

It follows that not every molecular species needs to be considered simultaneously in modeling a cell’s transitions between alternative biochemical states. Changes in the concentrations of the small molecules involved in the cell’s housekeeping functions such as energy metabolism and amino acid, nucleotide, and lipid synthesis, occur much more rapidly than changes in the pools of macromolecules such as RNAs and proteins. The latter class of molecules is indicative of the cell’s gene expression profile and can therefore be considered against a metabolic background defined by the average concentrations of molecules of the former class. Most of a cell’s active genes are kept in the “on” state during the cell’s lifetime, since like the metabolites they are also mainly involved in housekeeping functions. The pools of these “constitutively active” gene products can often be considered constant, with their concentrations entering into the dynamic description of the developing embryo as fixed parameters rather than variables. (See Goodwin [27] for an early discussion of separation of time scales in cell activities).

As mentioned above, it is primarily the regulated genes that are important to consider in analyzing determination and differentiation. And of these regulated genes, the most important ones for understanding cell type specification are those whose products control the activity of other genes.

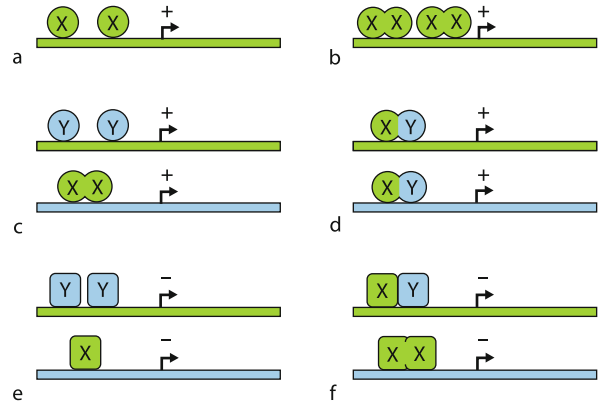
Genes are regulated by a set of proteins called *transcription factors*. Like all proteins, these factors are themselves gene products, in this case having the specific function of turning genes on and off. They do this by binding to specific sequences of DNA usually “upstream” of the gene’s transcription start site, called the “promoter”. Transitions between cell types during development are frequently controlled by the relative levels of transcription

factors. Because the control of most developmentally-regulated genes is a consequence of the synthesis of the factors that regulate their transcription, cell-type transitions can be driven by changes in the relative levels of a fairly small number of transcription factors [14]. We can thus gain insight into the dynamical basis of cell type switching (i. e., determination and differentiation) by focusing on molecular circuits, or networks, consisting solely of transcription factors and the genes that specify them. Networks in which the components mutually regulate one another's expression are termed "autoregulatory". Although a variety of autoregulatory transcription factor circuits appear during development, it is clear that such circuits will have different properties depending on their particular "wiring diagrams", that is the interaction patterns among their components.

Transcription factors can be classified either as *activators*, which bind to a site on a gene's promoter, and enhance the rate of that gene's transcription over its basal rate, or *repressors*, which decrease the rate of a gene's transcription when bound to a site on its promoter. Repression can be competitive or noncompetitive. In the first case, the repressor will interfere with activator binding and can only depress the gene's transcription rate to the basal level. In the second case, the repressor acts independently of any activator and can therefore potentially depress the transcription rate below basal levels. The basal rate of transcription depends on constitutive transcription factors, housekeeping proteins which are distinct from those in the autoregulatory circuits that we consider below.

Networks of autoregulatory transcription factors constitute dynamical systems. Like such systems encountered in physics and mathematics they can productively be modeled as systems of ordinary differential equations (ODE). A classic approach, exemplified by the work of Keller [47], and more recently by Cinquin and Demongeot [10], is to set up systems of ODEs representing experimentally motivated topologies and kinetics of transcription factor interactions. The solutions of these systems of ODEs are then examined with respect to stationary states, limit cycles, bifurcations, etc., under variations in initial conditions, external inputs and kinetic parameters. When considered a part of a multicellular aggregate, this single-cell dynamical system can switch between its intrinsic alternative states depending on signals it experiences from its microenvironment – other such cells, the ECM, or in certain cases the external world.

In his analysis of developmental switching, Keller [47] used simulation methods to investigate the behavior of several autoregulatory transcription factor networks with a range of wiring diagrams (Fig. 1). Each network was rep-



**Biological Development and Evolution, Complexity and Self-organization in, Figure 1**

Six model genetic circuits discussed by Keller [47]. **a** Autoactivation by monomer  $X$ ; **b** autoactivation by dimer  $X_2$ ; **c** mutual activation by monomer  $Y$  and dimer  $X_2$ ; **d** autoactivation by heterodimer  $XY$ ; **e** mutual repression by monomers  $X$  and  $Y$ ; **f** mutual repression by dimer  $X_2$  and heterodimer  $XY$ . Activated and repressed states of gene expression are represented respectively by + and -. The various transcription factors (circles for activators and squares for repressors) bind to the promoters of the genes, the colors of which correspond to the transcription factors they specify. Based on Keller [47]; see original paper for details. Figure modified and reprinted, with permission, from [21]

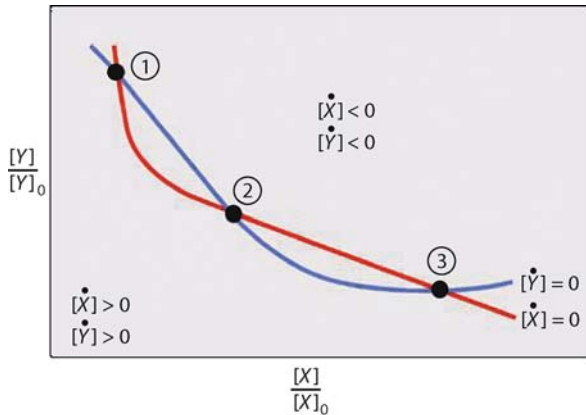
resented by a set of  $n$  coupled ordinary differential equations – one for the concentration of each factor in the network – and the steady-state behaviors of the systems were explored. The questions asked were: how many stationary states exist; are they stable or unstable?

One such network, designated as the "Mutual repression by dimer and heterodimer" (MRDH) (Fig. 1f), comprises the genes encoding the transcriptional repressor,  $X$ , and the gene encoding the protein,  $Y$ , and thus represents a two component network. The stationary values of the two components are governed by the following equations:

$$\frac{d[X]}{dt} = S_{XB} - \{d_X[X] + 2d_{X_2}K_{X_2}[X]^2 + d_{XY}K_{XY}[X][Y]\} = 0 \quad (1)$$

$$\frac{d[Y]}{dt} = \frac{1 + \rho K_X K_{X_2}[X]^2}{1 + K_X K_{X_2}[X]^2} S_{YB} - \{d_Y[Y] + d_{XY}K_{XY}[X][Y]\} = 0, \quad (2)$$

where  $[\dots]$  represents concentrations of the components, the  $S$ 's synthesis rates, and the  $d$ 's decay constants for the indicated monomeric ( $[X]$ ,  $[Y]$ ) and dimeric ( $[X]^2$ ) species. (See Keller [47] for details).



### Biological Development and Evolution, Complexity and Self-organization in, Figure 2

The solutions of the steady-state Eqs. (1) and (2), given in terms of the solutions  $[X_0]$  and  $[Y_0]$ . Here  $[X_0]$  is defined as the steady state cellular level of monomer  $X$  produced in the presence of steady state cellular level  $[Y_0]$  of monomer  $Y$  by the rate of transcription  $S_{X_0}$ . By definition (see Eq. (1)),  $S_{X_0} = d_X[X_0] + 2d_{X_2}K_{X_2}[X_0]^2 + d_{XY}K_{XY}[X_0][Y_0]$ . Since along the blue and red lines, respectively,  $d[X]/dt \equiv [\dot{X}] = 0$  and  $d[Y]/dt \equiv [\dot{Y}] = 0$ , the intersections of these curves correspond to the steady state solutions of the system of equations, Eqs. (1) and (2). Steady states 1 and 3 are stable, while steady state 2 is unstable. Based on Keller [47]; see original paper for details. Figure modified and reprinted, with permission, from [21]

It was found that if in the absence of the repressor  $X$  the rate of synthesis of protein  $Y$  is high, in its presence, the system described by Eqs. (1) and (2) exhibits three steady states, as shown in Fig. 2. Steady states 1 and 3 are stable, thus could be considered as defining two distinct cell types, while steady state 2 is unstable. In an example using realistic kinetic constants, the steady-state values of  $[X]$  and  $[Y]$  at the two stable steady states differ substantially from one another, showing that the dynamical properties of these autoregulatory networks of transcription factors can provide the basis for generating stable alternative cell fates during early development.

The biological validity of Keller's model depends on whether switching between these alternative states is predicted to occur under realistic conditions. By this criterion the model works: the microenvironment of a cell, containing an autoregulatory network of transcription factors, could readily induce changes in the rate of synthesis of one or more of the factors via signal transduction pathways that originate outside the cell ("outside-in" signaling [24]). Moreover, the microenvironment can also affect the activity of transcription factors in an autoregulatory network by indirectly interfering with their localization in the cell's nucleus, where transcription takes place [59]. In addition,

cell division may perturb the cellular levels of autoregulatory transcription factors, particularly if they or their mRNAs are unequally partitioned between the daughter cells. Any jump in the concentration of one or more factors in the autoregulatory system can bring it into a new basin of attraction and thereby lead to a new stable cell state.

The Keller model shows that the existence of multiple steady states in an embryonic cell's state space makes it possible, in principle, for more than one cell type to arise among its descendants. However this capability does not, by itself, provide the conditions under which such a potentially divergent cell population would actually be produced and persist as long as it is required.

Keller's is a single cell analysis. Another, less conventional, approach, acknowledges from the start the multicellular nature of developing organisms. Here the dynamical attractors sought are ones that pertain to the multicellular system as a whole, which consists of multiple copies of initially identical model cells in interaction with one another [43,44].

The need for an inherently multicellular approach to the cell type switching problem is indicated by experimental observations that suggest that cell differentiation depends on properties of cell aggregates rather than simply those of individual cells. For example, during muscle differentiation in the early frog embryo, the muscle precursor cells must be in contact with one another throughout gastrulation (the set of rearrangements that establish the body's main tissue layers referred to as the "germ layers", the ectoderm, mesoderm and endoderm) in order to develop into terminally differentiated muscle [29].

The need for cells to act in groups in order to acquire new identities during development has been termed the "community effect" [29]. This phenomenon is a developmental manifestation of the general property of cells and other dynamical systems of assuming one or another of their possible internal states in a fashion that is dependent on inputs from their external environment. In the case noted above the external environment consists of other cells of the same genotype.

Kaneko, Yomo and co-workers [43,44] have described a previously unknown chemical-dynamic process, termed "isologous diversification", by which replicate copies of the same dynamical system (i. e., cells of the same initial type) can undergo stable differentiation simply by virtue of exchanging chemical substances with one another. This differs from Keller's model described above in that the final state achieved exists only in the phase space of the collective "multicellular" system. Whereas the distinct local states of each cell within the collectivity are mutually re-

inforcing, these local states are not necessarily attractors of the dynamical system representing the individual cell, as they are in Keller's model. The Kaneko–Yomo system thus provides a model for the community effect.

The following simple version of the model, based on Kaneko and Yomo [44], was initially presented in Forgacs and Newman [21]. Improvements and generalizations of the model presented in subsequent publications [43] do not change its qualitative features.

Consider a system of originally identical cells with intra- and inter-cell dynamics, which incorporate cell growth, cell division and cell death (for the general scheme of the model, see Fig. 3a). The dynamical variables are the concentrations of molecular species (“chemicals”) inside and outside the cells. The criterion by which differentiated cells are distinguished is the average of the intracellular concentrations of these chemicals (over the cell cycle). Cells are assumed to interact with each other through their effect on the intercellular concentrations of the chemicals *A* and *B* (only two are considered in this simplified model). Chemicals enter the cells by either active transport or diffusion. It is further assumed that a source chemical *S* is supplied by a flow from an external tank to the chamber containing the cells.

Kaneko and Yomo [44] consider cell division to be equivalent to the accumulation of a threshold quantity of DNA. To avoid infinite growth in cell number, a condition for cell death has to also be imposed. It is assumed that a cell will die if the amount of chemicals *A* and *B* in its interior is below the “starvation” threshold.

The system described can be expressed as the following set of equations (see also Fig. 3b):

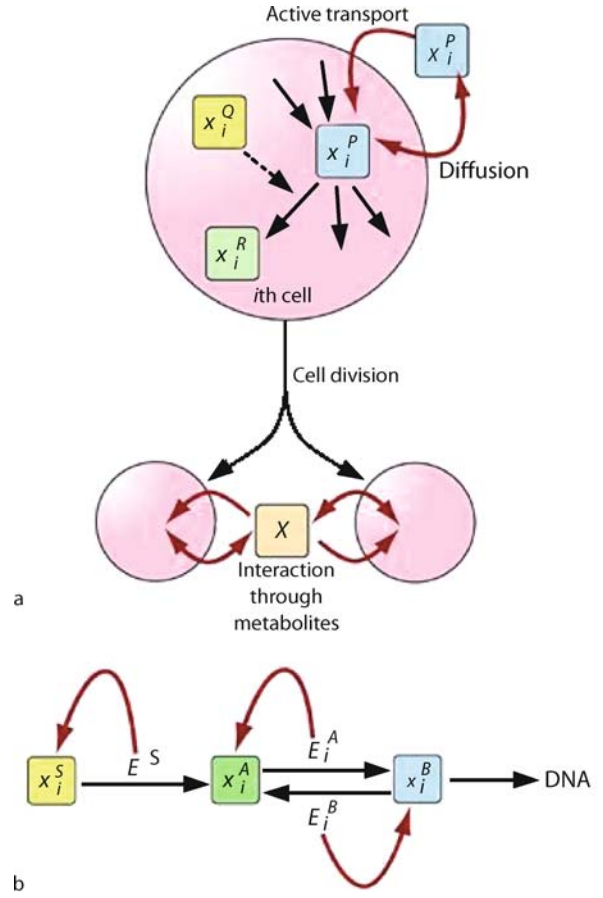
$$\frac{dx_i^S}{dt} = -E^S x_i^S + \text{Transp}_i^S + \text{Diff}_i^S \quad (3)$$

$$\frac{dx_i^A}{dt} = (e_B x_i^B) x_i^B - (e_A x_i^A) x_i^A + E^S x_i^S + \text{Transp}_i^A + \text{Diff}_i^A \quad (4)$$

$$\frac{dx_i^B}{dt} = (e_A x_i^A) x_i^A - (e_B x_i^B) x_i^B - k x_i^B + \text{Transp}_i^B + \text{Diff}_i^B \quad (5)$$

$$\frac{dX^S}{dt} = f(\bar{X}^S - X^S) - \sum_{i=1}^N (\text{Transp}_i^S + \text{Diff}_i^S), \quad (6)$$

Here the  $x_i$  are concentrations of the various chemicals in the  $i$ th cell,  $X^S$  the external concentration of the source chemical.  $E^S$  is the (constant) concentration in each cell of an enzyme that converts *S* into *A*, which in turn is catalyzed by the enzyme (with concentration  $E_i^A = e_A x_i^B$ ) to produce *B*. *B* is catalyzed by an enzyme (with concentra-



**Biological Development and Evolution, Complexity and Self-organization in, Figure 3**

**a** General scheme of the model of Kaneko and Yomo. The concentrations of each chemical may differ in the various cells (cell  $i$  is shown in the center) but some ( $P$  in the example) may have a fixed value in the extracellular environment. External chemicals can enter the cells by passive diffusion or active transport. The concentration of any chemical inside a given cell depends on chemical reactions in which other chemicals are precursors or products (solid arrows) or cofactors (dashed arrow). Once cell division has taken place by the synthesis of DNA (one of the “chemicals”) exceeding a threshold, cells communicate with one another by exchange of transportable chemicals. **b** Schematic representation of the intracellular dynamics of a simple version of the Kaneko–Yomo model. Red arrows symbolize catalysis. The variables  $x_i^A(t)$ ,  $x_i^B(t)$ ,  $x_i^S(t)$  and  $E_i^A$ ,  $E_i^B$ ,  $E^S$  denote respectively the concentrations of chemicals *A*, *B* and *S* and their enzymes in the  $i$ th cell, as described in the text. Reprinted, with permission, from [21]

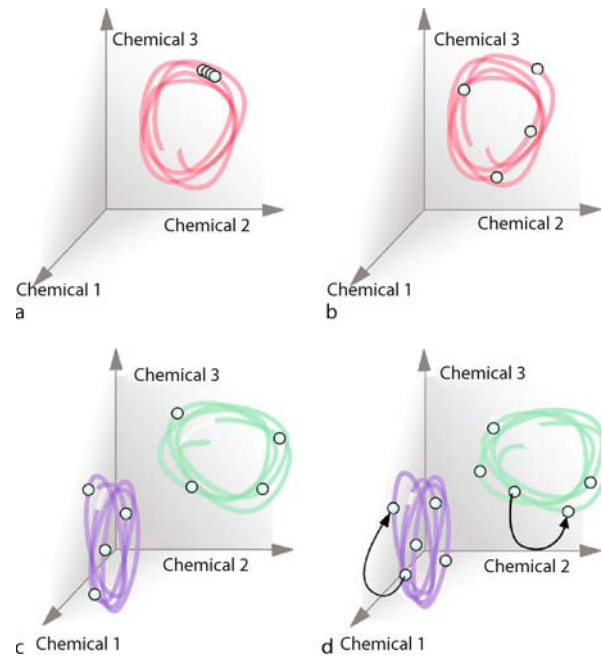
tion  $E_i^B = e_B x_i^A$ ) to produce *A* and it also produces DNA (at a rate  $k$ ). *Transp* and *Diff* are respectively active transport and diffusion functions, which depend on both the intracellular ( $x_i$ ) and extracellular ( $X_i$ ) concentrations. (See Forgacs and Newman [21] for details).



Simulations based on the above model and its generalizations, using a larger number of chemicals [43], led to the following general features, which are likely to pertain to real, interacting cells as well:

1. As the model cells replicate (by division) and interact with one another, eventually multiple biochemical states corresponding to distinct cell types appear. The different types are related to each other by a hierarchical structure in which one cell type stands at the apex, cell types derived from it stand at subnodes, and so on. Such pathways of generation of cell type, which are seen in real embryonic systems, are referred to as developmental lineages.
2. The hierarchical structure appears gradually. Up to a certain number of cells (which depends on the model parameters), all cells have the same biochemical state (Fig. 4a). When the total number of cells rises above a certain threshold value, the state with identical cells is no longer stable. Small differences between cells first introduced by random fluctuations in chemical concentrations start to be amplified. For example, synchrony of biochemical oscillations in different cells of the cluster may break down (Fig. 4b). Ultimately, the population splits into a few groups (“dynamical clusters”), with the phase of the oscillator in each group being offset from that in other groups, like groups of identical clocks in different time zones.
3. When the ratio of the number of cells in the distinct clusters falls within some range (depending on model parameters), the differences in intracellular biochemical dynamics are mutually stabilized by cell-cell interactions.
4. With further increase of cell number, the average concentrations of the chemicals over the cell cycle become different (Fig. 4c). That is to say, groups of cells come to differ not only in the phases of the same biochemical oscillations, but also in their average chemical composition integrated over the entire lifetimes of the cells. After the formation of cell types, the chemical compositions of each group are inherited by their daughter cells (Fig. 4d).

In contrast to the Keller model in which different cell types represent a choice among basins of attraction for a multi-attractor system, with external influences having the potential to bias such preset alternatives, in the Kaneko–Yomo model interactions between cells can give rise to stable intracellular states, which would not exist without such interactions. Isologous diversification thus provides a plausible model for the community effect [29]. It is reasonable to expect that both intrinsic multistability of a dy-



**Biological Development and Evolution, Complexity and Self-organization in, Figure 4**

Schematic representation of the differentiation scenario in the isologous diversification model of Kaneko and Yomo. When there are  $N$  cells and  $C$  chemicals ( $C = 3$  in the figure), the state space of the multicellular system is  $N \times C$  dimensional. A point in this space corresponds to the instantaneous value of all the chemicals and in each cell the orbits represent the evolution of these concentrations in the multicellular system. As long as the biochemical states of the replicating cells are identical, points along the orbit could characterize the synchronous states of the cells. This is illustrated in panel a, where the four circles, representing cells with the same phase and magnitude of their chemicals, overlap. With further replication, cells with differing biochemical states appear. First, chemicals in different cells differ only in their phases, thus the circles in panel b still fall on the same orbit, albeit are well separated in space. With further increase in cell number differentiation takes place: not only the phases but also the magnitudes (i. e., the averages over the cell cycle) of the chemicals in different cells will differ. The two orbits in panel c represent two distinct cell types, each different from the original cell type shown in panels a and b. Panel d illustrates the “breeding true” of the differentiated cells. After the formation of distinct cell types, the chemical compositions of each group are inherited by their daughter cells. That is, chemical compositions of cells are recursive over subsequent divisions as the result of stabilizing interactions. Cell division is represented here by an arrow from a progenitor cell to its progeny. Adapted, with changes, from Kaneko [42]. Reprinted, with permission, from [21]

namical system of the sort analyzed by Keller, and interaction-dependent multistability, as described by Kaneko, Yomo, and coworkers, based as they are on generic properties of complex dynamical systems, are utilized in initiating developmental decisions in various contexts in different organisms.

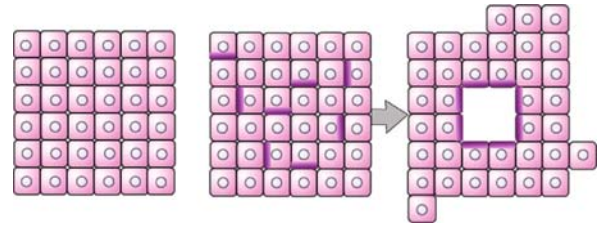
### Differential Adhesion: Gastrulation

As cells differentiate they become biochemically and structurally specialized and capable of forming multicellular structures with characteristic shapes, such as those of the earliest developmental stages, *blastulae* and multilayered *gastrulae*. Later during development tubular vessels, ducts and crypts emerge. The appearance and function of these specialized structures reflect, among other things, differences in the ability of cells to adhere to each other and the distinct mechanisms by which they do so. The distinct adhesion molecules expressed by differentiating cells thus mobilize physical forces to produce stereotypical multicellular structures.

The organization of tissues is determined not only by the chemical nature of the cells' adhesion molecules but also the distribution of these molecules on the cell surface. *Epithelioid* cells express cell adhesion molecules (CAMs), such as cadherins, uniformly over their surfaces and tend to form solid cell masses. *Epithelial* cells, which form two dimensional sheets must have CAMs predominantly along their lateral surfaces, whereas substrate adhesion molecules (SAMs), such as integrins, must populate their basal surfaces, along which interaction with the planar ECM, the basal lamina, is carried out. Such distribution of CAMs and SAMs renders epithelial cells *polarized*. Finally, *mesenchymal* cells express SAMs over their surfaces by which they interact with interstitial ECM.

Adhesion enters into the phenomena of embryonic development in several distinct ways. The first, and most straightforward way, is in simply holding tissues together. Whereas mature tissues contain the definitive, relatively long-lived forms of CAM-containing junctions, during early development the CAM-containing junctions are present in apparently immature forms [17], consistent with the provisional arrangement of cells and their capacity to rearrange during these stages.

The other roles for adhesion during development are based on its modulation – the phenomenon of *differential adhesion*. The regulated spatiotemporal modulation of adhesion is an important driving force for major morphogenetic transitions during embryogenesis. The simplest form of this is the detachment of cell populations from existing tissues. This is usually followed by their relocation.



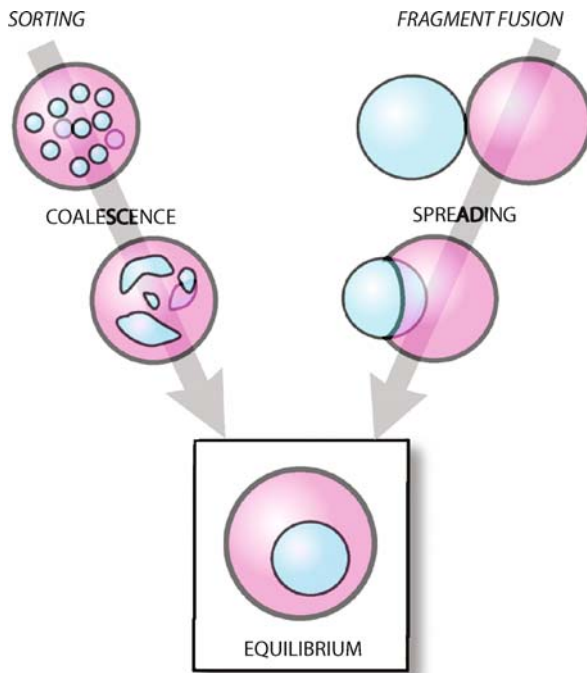
Biological Development and Evolution, Complexity and Self-organization in, Figure 5

Schematic illustration of lumen formation by polarized cells expressing cell adhesion molecules only on restricted parts of their surface. The shading in the eight cells in the *middle* panel represents the lack of adhesion molecules on corresponding regions of the cells. As a consequence, if minimization of configurational energy is the driving force in cell rearrangement, a lumen, shown on the *right* is bound to appear. (Based on Newman [62]) Reprinted, with permission, from [21]

But modulation of adhesive strength without complete detachment also has morphogenetic consequences, whether it occurs locally, on the scale of the single cell surface, or more globally, on the scale of whole cells within a common tissue.

Polar expression of CAMs can lead rather directly to morphogenetic change, as illustrated in Fig. 5. In the process of differentiation some of the originally non-polarized cells may lose their CAMs along part of their surface, thus preventing them from adhering to one another at those sites [98]. If the cells move around randomly, maximizing their contacts with one another in accordance with this constraint, a hollow region (a lumen) will naturally arise.

One of the most dramatic morphogenetic processes exhibited by embryonic cells is sorting, the ability of cells of distinct types to segregate into distinct tissues which do not intermix at their common boundary. Steinberg postulated that cells of different origin adhere to each other with different strengths and, in analogy with immiscible liquids such as oil and water, undergo a process of phase separation in which the final configuration corresponds to the minimum of interfacial and surface free energies [93]. (Here the different tissues play a role analogous to liquids and the constituent cells mimic the molecules of the liquid). This “differential adhesion hypothesis” (DAH) was expressed in quantitative terms by Phillips [77] based on geometric analysis of surface and interfacial tensions of immiscible droplets of liquids. According to the DAH, the final “phase-separated” state of two adjacent tissues is an equilibrium configuration not dependent on the pathway by which it was achieved. That is, it will be the same whether arrived at by fusion of two intact fragments of tissue or by the sorting out of their respective cells from a binary mixture (Fig. 6). Another implication of the DAH is



#### Biological Development and Evolution, Complexity and Self-organization in, Figure 6

Different paths by which two immiscible liquids or cell aggregates (composed of cells with differing adhesive properties) may arrive at the same equilibrium state. The path on the left shows sorting or nucleation, which proceeds through the gradual coalescence of groups of cells. The path on the right corresponds to engulfment, which occurs through spreading. Reprinted, with permission, from [21]

that tissue engulfment relationships should form a hierarchy: if tissue *A* engulfs tissue *B* and *B* engulfs *C* in separate experiments, it follows that *A* will engulf *C* if that experiment is performed. Finally, the DAH predicts that the values of tissue surface tensions should fall in a quantitative order that corresponds to the engulfment hierarchy. Each of these predictions of the DAH has been amply confirmed experimentally [19,21,22,93].

By the time the blastula has formed, the embryo already contains, or begins to generate, a number of differentiated cell types. Insofar as these cell types have or acquire distinct adhesive properties, compartmentalization or other forms of regional segregation start taking place. This regionalization, accompanied by the collective movement of the resulting cell masses – that is common to all forms of gastrulation – gives rise to embryos consisting of two or three germ layers along with some subsidiary populations of cells.

To understand how cells can rearrange by differential adhesion it is useful to describe the notion of the “work

of adhesion”. Consider a unit interfacial area between two materials *A* and *B* immersed in a medium, denoted by *M* (which in particular could be vacuum or, in the case of tissues, the extracellular medium or tissue culture medium). We define the work of adhesion  $w_{AB}$  as the energy input required to separate *A* and *B* across the unit area in medium *M*. We can imagine such a unit area to be formed in the following way. First, we separate a rectangular column of *A* and *B* to produce two free unit surfaces of each substance (Fig. 7). This requires  $w_{AA}$  and  $w_{BB}$  amounts of work, respectively. These quantities are called the works of cohesion. (Note that the magnitudes of the work of adhesion and cohesion depend on the medium). We then combine these pieces to end up with two unit interfacial areas between *A* and *B* as shown in Fig. 7. Thus the total work,  $\Delta_{AB}$ , needed to produce a unit interfacial area between *A* and *B* is given by

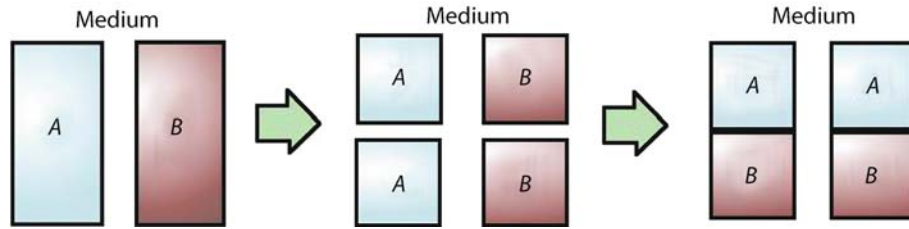
$$\Delta_{AB} = \frac{1}{2}(w_{AA} + w_{BB}) - w_{AB} \quad (7)$$

The quantity  $\Delta_{AB}$  is called the interfacial energy and Eq. (7) is known as the Dupré equation [39]. If the interface is formed by two immiscible liquids then Eq. (7) can readily be expressed in terms of liquid surface and interfacial tensions. By its definition the surface tension is the energy required to increase the surface of the liquid by one unit of area. Since the works  $w_{AA}$  and  $w_{BB}$  create two units of area, we obtain

$$\sigma_{AB} = \sigma_{AM} + \sigma_{BM} - w_{AB}, \quad (8)$$

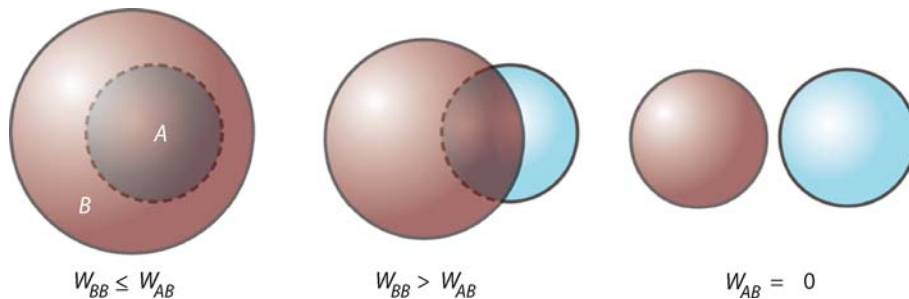
where  $\sigma_{AM}$ ,  $\sigma_{BM}$  and  $\sigma_{AB}$  are respectively the surface tensions of liquids *A* and *B* and their mutual interfacial tension [39]. (Whereas for solids  $\Delta_{AB}$  depends on the amount of interfacial area between *A* and *B* already formed, for liquids,  $\sigma_{AB}$  does not). Note that immiscibility of *A* and *B* implies  $\sigma_{AB} > 0$ . If, on the contrary,  $\sigma_{AB} \leq 0$ , it is energetically more beneficial for the molecules of liquid *A* to be surrounded by molecules of liquid *B*, and vice versa; that is, *A* and *B* are miscible.

If we now invoke the liquid-like behavior of tissues, we can apply Eq. (7) or Eq. (8) to obtain the conditions for sorting in terms of the  $w$ 's or  $\sigma$ 's. We imagine the cells of tissues *A* and *B* to be initially randomly intermixed and surrounded by tissue culture medium and allow them to sort, as discussed above. Let us assume that tissue *A* is the more cohesive one. This implies that  $w_{AA}$  is the largest of the three quantities on the right hand side of Eq. (7). In the energetically most favorable configuration at the end of the sorting process cells of tissue *A* form a sphere, the configuration in which they have minimal contact with their environment and maximal contact with each other. Then, depending on the relative magnitudes of the  $w$ 's,



Biological Development and Evolution, Complexity and Self-organization in, Figure 7

Schematic illustration of creation of an interface between two materials A and B, immersed in a common medium. In the two-step process shown, first free interfaces of materials A or B are produced (*middle panel*), which requires the separation of their corresponding subunits (molecules, cells) from one another, and thus involves the works of cohesion,  $w_{AA}$  and  $w_{BB}$ . In the second step the free interfaces are combined to form the AB interface (by rearranging the A and B blocks; *right panel*). The separation of A and B requires the work of adhesion,  $w_{AB}$ . If the cross-sectional area of columns A and B is of unit magnitude, then the operation shown results in two units of interfacial area between A and B. Reprinted, with permission, from [21]



Biological Development and Evolution, Complexity and Self-organization in, Figure 8

Geometric configurations of immiscible liquids A (represented by the interior smaller sphere in the *left panel*) and B (represented by the larger sphere in the *left panel*), and the corresponding relations between the works of cohesion and adhesion. It is assumed that A is more cohesive than B:  $w_{AA} > w_{BB}$  and  $\sigma_{AM} > \sigma_{BM}$  ( $M$  denotes the surrounding medium). The *left* and *middle* panels correspond respectively to complete and partial envelopment. Figure modified and reprinted, with permission, from [21]

the sphere of tissue B may completely or partially envelop the sphere of tissue A or the two spheres may separate (Fig. 8) [21,93].

### Biochemical Oscillations: Segmentation

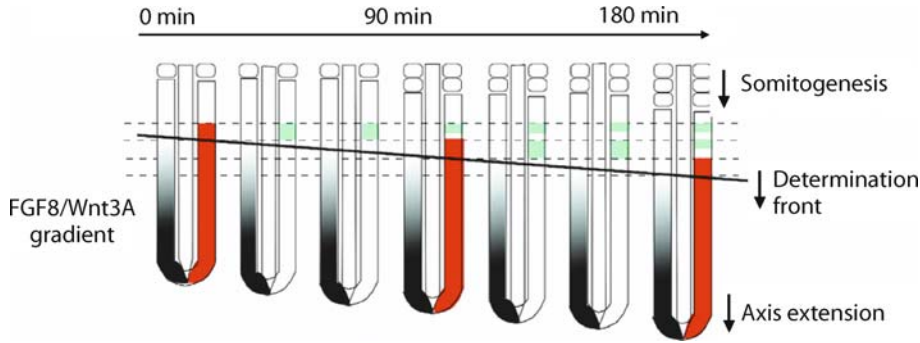
A wide variety of animal types, ranging across groups as diverse as insects, annelids (e.g., earthworms), and vertebrates, undergo *segmentation* early in development, whereby the embryo, or a major portion of it, becomes subdivided into a series of tissue modules [25]. These modules typically appear similar to each other when initially formed; later they may follow distinct developmental fates and the original segmental organization may be all but obscured in the adult form. Somite formation (or somitogenesis) is a segmentation process in vertebrate embryos in which the tissue to either side of the central axis of the embryo (where the backbone will eventually form) becomes organized into parallel blocks of tissue.

Somitogenesis takes place in a sequential fashion. The first somite begins forming as a distinct cluster of cells

in the anterior region (towards the head) of the embryo's body. Each new somite forms just posterior (towards the tail) to the previous one, budding off from the anterior portion of the unsegmented presomitic mesoderm (PSM) (Fig. 9). Eventually, 50 (chick), 65 (mouse), or as many as 500 (certain snakes) of these segments will form.

In the late nineteenth century the biologist William Bateson speculated that the formation of repetitive blocks of tissue, such as the somites of vertebrates or the segments of earthworms might be produced by an oscillatory process inherent to developing tissues [2,65]. More recently, Pourquié and coworkers made the significant observation that the gene *c-hairy1*, which specifies a transcription factor, is expressed in the PSM of avian embryos in cyclic waves whose temporal periodicity corresponds to the formation time of one somite [72,78] (Fig. 9). The *c-hairy1* mRNA and protein product are expressed in a temporally-periodic fashion in individual cells, but since the phase of the oscillator is different at different points along the embryo's axis, the areas of maximal expression sweep along the axis in a periodic fashion.





**Biological Development and Evolution, Complexity and Self-organization in, Figure 9**

Model proposed by Pourquié for segment formation in vertebrates, based on mouse and chick data. A gradient of *FGF8* (see text), shown in *black*, regresses posteriorly during somitogenesis. The anterior boundary of the gradient defines the determination front, which corresponds to the position of the wavefront (*thick black line*). (A coordinately-expressed gradient of *Wnt3A* plays a similar role; Aulehla et al. [2]). Oscillatory (i. e., waxing and waning with developmental stage) expression of *c-hairy1* and related genes is shown in *red*. Expression of genes of the *Mesp* family, which encode transcription factors involved in somite boundary formation, is shown in *green*. (Reprinted, with permission, from Pourquié [78]; see original paper for details)

Experimental evidence suggests that somite boundaries form when cells which have left a posterior growth zone move sufficiently far away from a source of a diffusible protein known as fibroblast growth factor 8 (FGF8) in the tailbud at the posterior end of the embryo [18]. The FGF gradient thus acts as a “gate” that, when its low end coincides with a particular phase of the segmentation clock, results in formation of a boundary [18,78]. The general features of this mechanism (called the “clock and wavefront” model) were predicted on the basis of dynamical principles two decades before there was any direct evidence for a somitic oscillator [13].

During somitogenesis in the zebrafish a pair of transcription factors known as *her1* and *her7*, which are related to chicken *hairyl* (see above), oscillate in the PSM in a similar fashion, as does the cell surface signaling ligand delta C. Lewis [52] has suggested that *her1* and *her7* constitute an autoregulatory transcription factor gene circuit of the sort treated by Keller [47] (see Sect. “Dynamic Multistability: Cell Differentiation”, above), and are the core components of the somitic oscillator in zebrafish. He also hypothesizes that deltaC, whose signaling function is realized by activating the Notch receptors on adjacent cells, is a downstream effector of this oscillation. The two *her* genes negatively regulate their own expression [33,71] and are positively regulated by signaling via Notch [71]. Certain additional experimental results [34] led Lewis to the conclusion that signaling by the Notch pathway, usually considered to act in the determination of cell fate [1] in this case acts to keep cells in the segment-generating growth zone in synchrony [52]. Such synchrony has been experimentally confirmed in chicken embryos [79,94].

Lewis [52] and Monk [58] have independently provided a simple mechanism for the oscillatory expression of the *her1* and *her7* genes, which we briefly summarize here. The model is based on the assumption, mentioned above, that there exists a feedback loop in which the Her1 and Her7 proteins directly bind to the regulatory DNA of their own genes to inhibit transcription. Also incorporated into the model is the recognition that there is always a delay between the initiation of transcription and the initiation of translation,  $T_m$  (since it takes time for the mRNA molecule to translocate into the cytoplasm), as well as between the initiation of translation and the emergence of a complete functional protein molecule,  $T_p$  (see Fig. 10a).

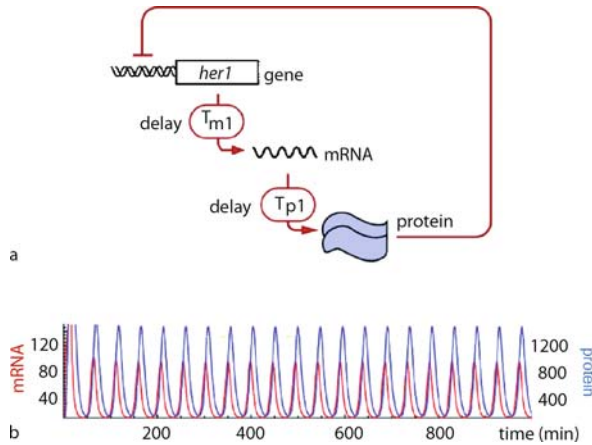
These ingredients are put into mathematical language in the following way. For a given autoregulatory gene, let  $m(t)$  be the number of mRNA molecules in a cell at time  $t$  and let  $p(t)$  be the number of the corresponding protein molecules. The rate of change of  $m$  and  $p$ , are then assumed to obey the following equations

$$\frac{dp(t)}{dt} = am(t - T_p) - bp(t) \quad (9)$$

$$\frac{dm(t)}{dt} = f[p(t - T_m)] - cm(t). \quad (10)$$

Here the constants  $b$  and  $c$  are the decay rates of the protein and its mRNA, respectively,  $a$  is the rate of production of new protein molecules and  $f(p)$  is the rate of production of new mRNA molecules. The function  $f(p)$  is assumed to be a decreasing function of the amount of protein (for its form see Lewis [52]). The results of simulations turned out to be quite insensitive to the specific form of  $f(p)$ .





**Biological Development and Evolution, Complexity and Self-organization in, Figure 10**

Cell autonomous gene expression oscillator for zebrafish somitogenesis. **a** Molecular control circuitry for a single gene, *her1*, whose protein product acts as a homodimer to inhibit *her1* expression. **b** Computed behavior for the system in **a** (defined by Eqs. (9) and (10)) in terms of the number of mRNA molecules per cell in red and protein molecules in blue. Parameter values were chosen appropriate for the *her1* homodimer oscillator based on experimental results. (Reprinted from Lewis [52] with permission from Elsevier; see original paper for details)

The above delay differential equations were numerically solved for *her1* and *her7* (for which Lewis was able to estimate the values of all the model parameters in Eqs. (9) and (10) using experimental results). The solutions indeed exhibit sustained oscillations in the concentration of Her1 and Her7, with the predicted periods close to the observed ones (Fig. 10b). In a subsequent study Lewis and his coworkers have experimentally confirmed the main features of this oscillatory model for segmentation [26].

### Reaction-Diffusion Mechanisms: Body Axis Formation

The nonuniform distribution of any chemical substance, whatever the mechanism of its formation, can clearly provide spatial information to cells. For at least a century embryologists have considered models for pattern formation and its regulation that employ diffusion gradients [25]. Only in the last decade, however, has convincing evidence been produced that this mechanism is utilized in early development. The prime evidence comes from studies of mesoderm induction, a key event preceding gastrulation in the frog *Xenopus*. Nieuwkoop [69] originally showed that mesoderm (the middle of the three germ layers in the three-layered gastrula – the one that gives rise to muscle, skeletal tissue, and blood) only appeared when tissue

from the upper half of an early embryo (“animal cap”) was juxtaposed with tissue from the lower half of the embryo (“vegetal pole”). By themselves, animal cap and vegetal pole cells, respectively, only produce ectoderm, which gives rise to skin and nervous tissue, and endoderm, which gives rise to the intestinal lining. Later it was found that several released, soluble, factors of the TGF- $\beta$  protein superfamily and the FGF protein family could substitute for the inducing vegetal pole cells [28]. Both TGF- $\beta$  [55] and FGFs [9] can diffuse over several cell diameters.

While none of this proves beyond question that simple diffusion of such released signal molecules (called “morphogens”) between and among cells, rather than some other, cell-dependent mechanism, actually establishes the gradients in question, work by Lander and co-workers, among others [49], has shown that diffusion-like processes, mediated by secreted molecules interacting with cell surfaces and ECMs, are likely to be involved in much developmental patterning. Analysis of such effects can best be performed by considering generalized reaction-diffusion systems, which we now describe (see [21] for additional details).

The rate of change in the concentrations of  $n$  interacting molecular species ( $c_i$ ,  $i = 1, 2, \dots, n$ ) is determined by their reaction kinetics and can be expressed in terms of ordinary differential equations

$$\frac{dc_i}{dt} = F_i(c_1, c_2, \dots, c_n). \quad (11)$$

The explicit form of the functions  $F_i$  in Eq. (11) depends on the details of the reactions. Spatial inhomogeneities also cause time variations in the concentrations even in the absence of chemical reactions. If these inhomogeneities are governed by diffusion, then in one spatial dimension,

$$\frac{\partial c_i}{\partial t} = D_i \frac{\partial^2 c_i}{\partial x^2}. \quad (12)$$

Here  $D_i$  is the diffusion coefficient of the  $i$ th species. In general, both diffusion and reactions contribute to the change in concentration and the time dependence of the  $c_i$ 's is governed by reaction-diffusion equations

$$\frac{\partial c_i}{\partial t} = D_i \frac{\partial^2 c_i}{\partial x^2} + F_i(c_1, c_2, \dots, c_n). \quad (13)$$

Reaction-diffusion systems exhibit characteristic parameter-dependent bifurcations (the “Turing instability”), which are thought to serve as the basis for pattern formation in several embryonic systems, including butterfly wing spots, stripes on fish skin, distribution of feathers on the skin of birds, the skeleton of the vertebrate limb, and the primary axis of the developing vertebrate embryo [21].

Gastrulation in the frog embryo is initiated by the formation of an indentation, the “blastopore”, through which the surrounding cells invaginate, or tuck into, to the hollow blastula. Spemann and Mangold [91] discovered that the anterior blastopore lip constitutes an *organizer*: a population of cells that directs the movement of other cells. The action of the Spemann–Mangold organizer ultimately leads to the formation of the notochord, the rod of connective tissue that first defines the anteroposterior body axis, and to either side of which the somites later appear (see Sect. “Biochemical Oscillations: Segmentation”, above). These investigators also found that an embryo with an organizer from another embryo at the same stage transplanted at some distance from its own organizer would form two axes, and conjoined twins would result. Other classes of vertebrates have similarly acting organizers.

A property of this tissue is that if it is removed, adjacent cells differentiate into organizer cells and take up its role. This indicates that one of the functions of the organizer is to suppress nearby cells with similar potential from exercising it. This makes the body axis a partly self-organizing system. The formation of the body axis in vertebrates also exhibits another unusual feature: while it takes place in an apparently symmetrical fashion, with the left and right sides of the embryo seemingly equivalent to one another, at some point the symmetry is broken. Genes such as *nodal* and *lefty* start being expressed differently on the two sides of the embryo [90], and the whole body eventually assumes a partly asymmetric morphology, particularly with respect to internal organs, such as the heart.

Turing [99] first demonstrated that reaction-diffusion systems like that represented in Eq. (12) will, with appropriate choice of parameters and boundary conditions, generate self-organizing patterns, with a particular propensity to exhibit symmetry breaking across more than one axis. Using this class of models, Meinhardt [56] has presented an analysis of axis formation in vertebrates and the breaking of symmetry around these axes.

The first goal a model of axis formation has to accomplish is to generate an organizer *de novo*. For this high local concentrations and graded distributions of signaling molecules are needed. This can be accomplished by the coupling of a self-enhancing feedback loop acting over a short range with a competing inhibitory reaction acting over a longer range. The simplest system that can produce such a molecular pattern in the  $x - y$  plane consists of a positively autoregulatory activator (with concentration  $A(x, y; t)$ ) and an inhibitor (with concentration  $I(x, y; t)$ ). The activator controls the production of the inhibitor, which in turn limits the production of the activator. This process can be described by the following reac-

tion-diffusion system [56]

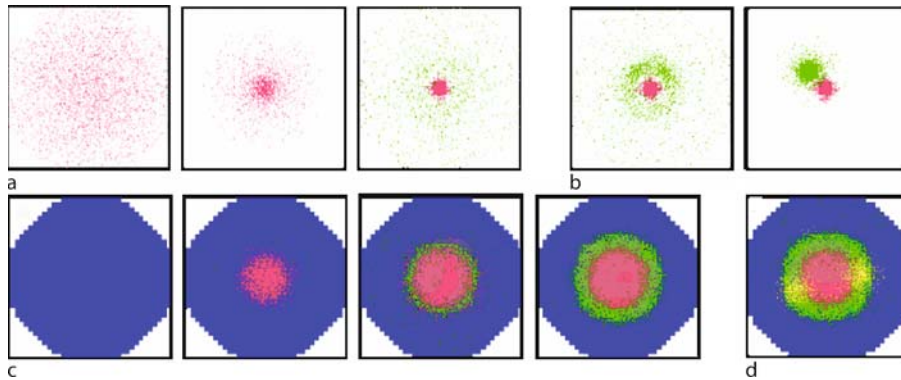
$$\frac{\partial A}{\partial t} = D_A \left( \frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2} \right) + s \frac{A^2 + I_A}{I(1 + s_A A^2)} - k_A A \quad (14)$$

$$\frac{\partial I}{\partial t} = D_I \left( \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \right) + s A^2 - k_I I + I_I \quad (15)$$

The  $A^2$  terms specify that the feedback of the activator on its own production and that of the inhibitor in both cases is non-linear. The factor  $s > 0$  describes positive autoregulation, the capability of a factor to induce positive feedback on its own synthesis. This may occur by purely chemical means (“autocatalysis”), which is the mechanism assumed by Turing [99] when he first considered systems of this type. More generally, in living tissues, positive autoregulation occurs if a cell’s exposure to a factor it has secreted causes it to make more of the same factor [100]. The inhibitor slows down the production of the activator (i. e., the  $1/I$  factor in the second term in Eq. (14)). Both activator and inhibitor diffuse (i. e., spread) and decay with respective diffusion ( $D_A, D_I$ ) and rate constants ( $k_A, k_I$ ). The small baseline inhibitor concentrations,  $I_A$  and  $I_I$  can initiate activator self-enhancement or suppress its onset, respectively, at low values of  $A$ . The factor  $s_A$ , when present, leads to saturation of positive autoregulation. Once the positive autoregulatory reaction is under way, it leads to a stable, self-regulating pattern in which the activator is in dynamic equilibrium with the surrounding cloud of the inhibitor.

The various organizers and subsequent inductions leading to symmetry breaking, axis formation and the appearance of the three germ layers in amphibians during gastrulation, can all, in principle, be modeled by the reaction-diffusion system in Eqs. (14) and (15), or by the coupling of several such systems. The biological relevance of such reaction-diffusion models depends on whether there exist molecules that can be identified as activator-inhibitor pairs. Meinhardt’s model starts with a default state, which consists of ectoderm. Patch-like activation generates the first “hot spot”, the *vegetal pole organizer*, which induces endoderm formation (simulation in Fig. 11a). A candidate for the diffusible activator in the corresponding self-enhancing loop for endoderm specification is the TGF- $\beta$ -like factor *Derriere*, which activates the *VegT* transcription factor [97].

*VegT* expression remains localized to the vegetal pole, but not because of lack of competence of the surrounding cells to produce *VegT* [11]. These findings provide circumstantial evidence for the existence of the inhibitor required by the reaction-diffusion model. Subsequently, a second feedback loop forms a second hot spot in the



**Biological Development and Evolution, Complexity and Self-organization in, Figure 11**

Pattern formation in the reaction-diffusion model of Meinhardt. **a** Induction of the vegetal pole organizer. *Left*: The interaction of an autocatalytic activator, the TGF- $\beta$ -like factor *Derriere* (red), with a long-ranging inhibitor (whose production it controls, and which in turn, limits the activator's production), creates an unstable state in an initially near-uniform distribution of the substances (inhibitor not shown). *Middle and right*: Small local elevation in the activator concentration above steady-state levels triggers a cascade of events governed by Eqs. (14)–(15): further increase of the activator due to autocatalysis, spread of the concomitantly produced surplus of inhibitor into the surrounding, where it suppresses activator production (*middle*), and the establishment of a new stable state, in which the activator maximum (hot spot) is in a dynamic equilibrium with the surrounding cloud of inhibitor (*right*). The *right* panel also shows the near-uniform distribution of the activator (green) in the second reaction-diffusion system discussed in **b**. **b** Induction of the Nieuwkoop center. Once the first hot spot has formed, it activates a second self-enhancing feedback loop. The production of the activator (green) in this reaction is inhibited by the vegetal pole organizer itself. As a consequence, the Nieuwkoop center is displaced from the pole. **c** Zonal separation of ectoderm, endoderm and mesoderm. The competition of several positive feedback loops assures that in one cell only one of these loops is active. As the results indicate, reaction-diffusion systems can produce not only spot-like organizers, but also zones. Endoderm (red) forms as shown from a default ectodermal state (blue). The mesodermal zone (green; which forms by the involvement of the FGF/*brachyury* feedback loop) develops a characteristic width by an additional self-inhibitory influence in Eqs. (14) and (15). **d** Induction of the Spemann–Mangold organizer. The activation of an organizing region (yellow) within the still symmetric mesodermal zone would imply a competition over relatively long distances. In such a case, Eqs. (14) and (15) lead to the occurrence of two organizing regions, a biologically unacceptable result. The strong asymmetry shown in **b** prevents this and suggests a reason for the existence of the Nieuwkoop center. (Reprinted from Meinhardt [56] with permission from the University of the Basque Country Press; see original paper for details)

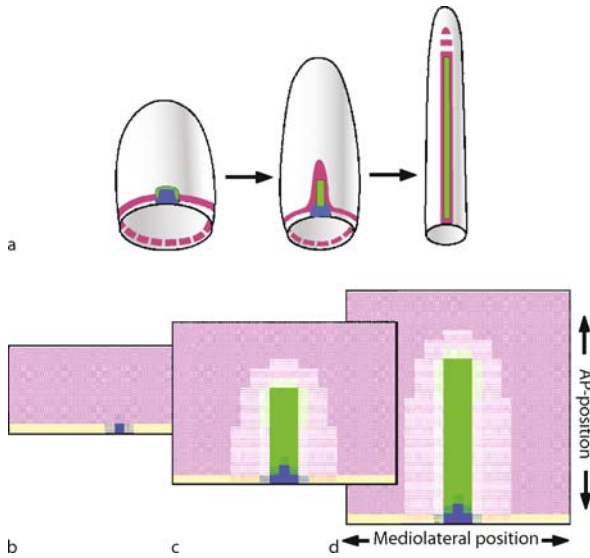
vicinity of the first, in the endoderm. This is identified with the “Nieuwkoop center”, a second organizing region, which appears in a specific quadrant of the blastula. A candidate for the second self-enhancing loop is FGF together with *Brachyury* [87]. Interestingly, the inhibitor for this loop is hypothesized to be the first loop itself (i. e., the vegetal pole organizer), which acts as local repressor for the second. As a result of this local inhibitory effect, the Nieuwkoop center is displaced from the pole (simulation in Fig. 11b). With the formation of the Nieuwkoop center the spherical symmetry of the embryo is broken. In Meinhardt's model this symmetry breaking “propagates” and thus forms the basis of further symmetry breakings, in particular the left-right asymmetry.

By secreting several diffusible factors, the Nieuwkoop center induces the formation of the Spemann–Mangold organizer [31] (If the second feedback loop, responsible for the Nieuwkoop center is not included in the model, two Spemann–Mangold organizers appear symmetrically with respect to the animal-vegetal axis and no symmetry breaking occurs). The organizer triggers gastrulation, in

the course of which the germ layers acquire their relative positions and the notochord forms. This long thin structure marks the midline of the embryo, which itself inherits organizer function and eventually establishes the primary (AP) embryonic axis. A simulation of midline formation, based on Meinhardt's model is shown in Fig. 12.

### Evolution of Developmental Mechanisms

Many key gene products that participate in and regulate multicellular development emerged over several billions of years of evolution in a world containing only single-celled organisms. Less than a billion years ago multicellular organisms appeared, and the gene products that had evolved in the earlier period were now raw material to be acted upon by physical and chemical-dynamic mechanisms on a more macroscopic scale [62]. The mechanisms described in the previous sections – chemical multistability, chemical oscillation, changes in cell adhesive differentials (both between cell types and across the surface of individual cells), and reaction-diffusion-based symmetry



#### Biological Development and Evolution, Complexity and Self-organization in, Figure 12

**Formation of the midline and enfolding of the anteroposterior (AP) axis according to the reaction-diffusion model of Meinhardt.** **a** Schematics of the hypothesized processes involved in axis formation. Cells close to the blastopore (red annulus) move towards the Spemann–Mangold organizer (blue). The spot-like organizer redirects the cells in its vicinity: initially they are drawn to the organizer, but then lose their attraction to it, so that they leave as a unified stripe-like band. Stem cells in the organizer region may contribute to this band to form the most central element of the midline (green). Such highly coordinated cellular motion and strong positional specification along the enfolding AP axis requires the simultaneous action of several feedback loops. **b–d** Simplified simulation of the scenario described in **a**. A reaction-diffusion system tuned to make stripes (green; compare with zone formation in Fig. 11) is triggered by the organizer (blue). The organizer itself is the result of a self-enhancing system, activated in a spot-like manner. (In **b–d** the blastopore is shown in yellow.) Repulsion between the spot system (the organizer) and the stripe system (the notochord) causes the elongation of the latter. Saturation in self-enhancement (due to the factor  $s_A$  in Eq. (14)) ensures that the stripe system does not disintegrate into individual patches and thus establishes the midline. This, in turn, acts as a sink for a ubiquitously produced substance (pink), which could be the product of *BMP-4*. The local concentration of this substance (shades of pink) is a measure of the distance from the midline. This simulation is simplified in that the actual movement of cells toward the organizer is not considered; instead, the midline elongates by the addition of lines of new cells next to the blastopore. (a redrawn, and **b–d** reprinted, from Meinhardt [56] with permission from the University of the Basque Country Press; see original paper for details)

breaking, would have caused these ancient multicellular aggregates to take on a wide range of biological forms. Any set of physicochemical activities that generated a new form in a reliable fashion within a genetically uniform popula-

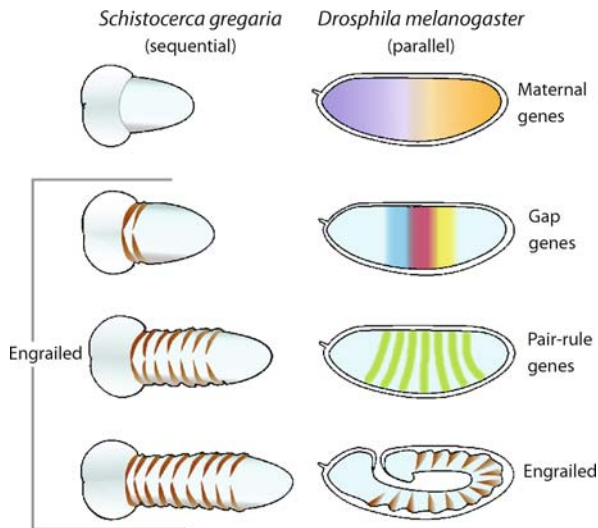
tion of cell aggregates would have constituted a primitive developmental mechanism.

But any developmental process that depends solely on physical mechanisms of morphogenesis would have been variable in its outcome – subject to inconstant external parameters such as temperature and the chemical microenvironment [67]. Modern-day embryos develop in a more rigidly programmed fashion: the operation of cell type- and pattern-generating physical and chemical-dynamic mechanisms is constrained and focused by hierarchical systems of coordinated gene activities. These developmental-genetic programs are the result of eons of molecular evolution that occurred mainly *after* the origination of multicellularity.

The relation between physical mechanisms of morphogenesis and hierarchical developmental programs can be appreciated by considering a major puzzle in the field of evolutionary developmental biology (“EvoDevo”): the disparity between the mechanisms of segment formation in “short germ-band” insects such as beetles and “long germ-band” insects such as fruit flies. As we will see, one solution to this conundrum touches on all categories of multicomponent dynamical behaviors discussed in earlier sections.

Similarly to somitogenesis in vertebrates (see Sect. “Biochemical Oscillations: Segmentation”), in short germ-band insects [75], as well as in other arthropods, such as the horseshoe crab [40], segmental primordia are added in sequence from a zone of cell proliferation (“growth zone”) (Fig. 13). In contrast, in long germ-band insects, such as the fruit fly *Drosophila*, a series of chemical stripes (i. e., parallel bands of high concentration of a molecule) forms in the embryo, which at this stage is a *syncytium*, a large cell with single cytoplasmic compartment containing about 6000 nuclei arranged in a single layer on the inner surface of the plasma membrane [50]. These stripes are actually alternating, evenly-spaced bands of transcription factors of the “pair-rule” class. The pair-rule genes include *even-skipped*, *fushi tarazu*, and *hairy*, which is the insect homolog of the *c-hairy1* gene expressed in a periodic fashion during vertebrate somitogenesis (see Sect. “Biochemical Oscillations: Segmentation”). When cellularization (the enclosure of each nucleus and nearby cytoplasm in their own complete plasma membrane) takes place shortly thereafter, the cells of the resulting blastoderm will have periodically-distributed identities, determined by the particular mix of transcription factors they have incorporated. The different cell states are later transformed into states of differential adhesivity [36], and morphological segments form as a consequence.





**Biological Development and Evolution, Complexity and Self-organization in, Figure 13**

Comparison of sequential and parallel segmentation modes in short and long germ-band insects. Left: In short germ-band insects (the embryo of the grasshopper, *Schistocerca gregaria*, is shown schematically) one or groups of a few segments appear in succession. Brown stripes indicate expression of a segment polarity gene such as *engrailed*. With further development additional segments appear sequentially from a zone of proliferation that remains posterior to the most recently-added segment. Right: In long germ-band insects (the embryo of the fruit-fly, *Drosophila melanogaster*, is shown) gradients of maternal gene products (e. g., *bicoid* and *nanos*) are present in the egg before cellularization of the blastoderm (see text). As development proceeds, the maternal gene products induce the expression of gap genes (e. g., *hunchback*, *Krüppel*), the products of which, in turn, induce the expression of pair-rule genes in a total of 14 stripes (e. g., *eve*, *fushi tarazu*, *hairy*). Each pair-rule gene is expressed in a subset of the stripes: *eve*, for example, is expressed in seven alternating stripes, shown in green. The pair-rule gene products provide a prepattern for the alternating bands of nuclear expression of segment polarity genes. Once cellularization has occurred (bottom panel), these are distributed similarly to that in short germ-band embryos. (*Schistocerca* series after Patel [73]; *Drosophila* series after Ingham [36]. Reprinted, with permission, from [21])

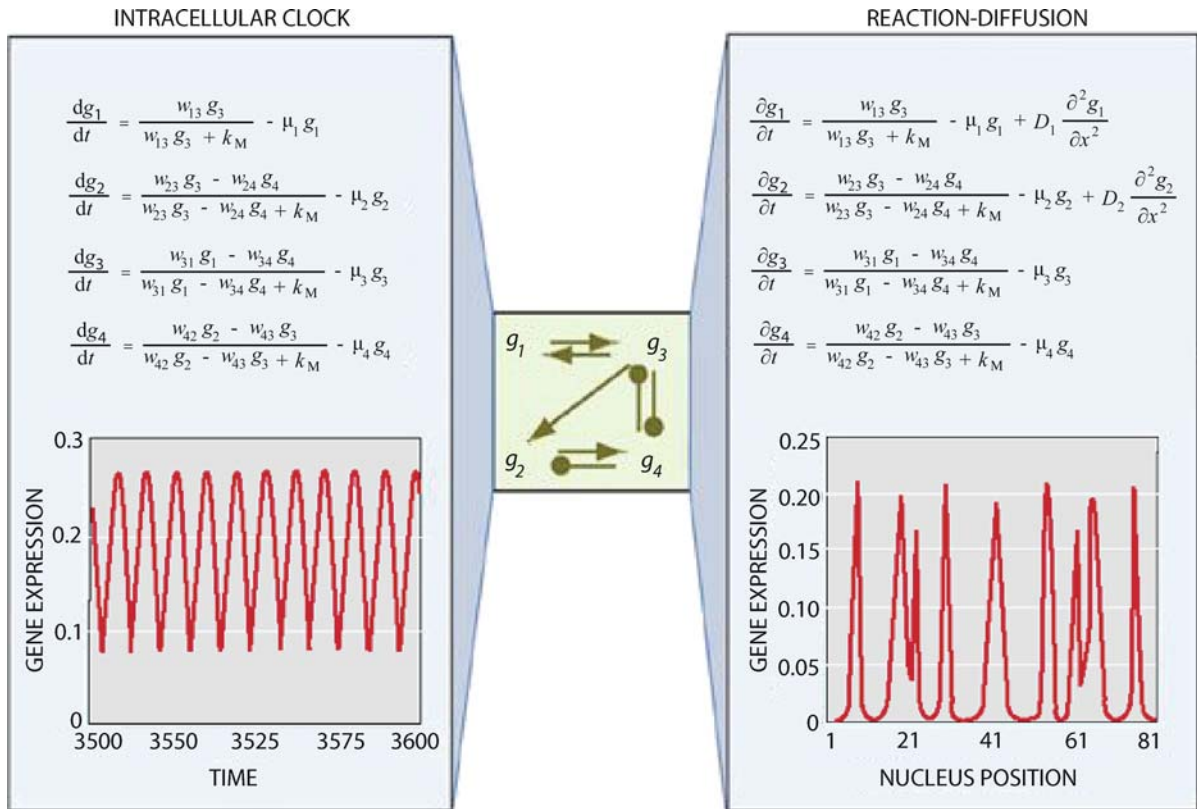
The formation of overt segments in both short (e. g., grasshoppers [75], Patel [74] and beetles [6]) and long germ-band insects [45] requires the prior expression of a stripe of the product of the *engrailed* (*en*) gene, a transcription factor, in the cells of the posterior border of each of the presumptive segments. In *Drosophila*, the positions of the engrailed stripes are largely determined by the activity of the pair-rule genes *even-skipped* (*eve*) and *fushi tarazu* (*ftz*), which, as mentioned above, exhibit alternating, complementary seven stripe patterns prior to the formation of the blastoderm [23,35].

On theoretical [4] and experimental [51] grounds it has long been recognized that the kinetic properties that give rise to a chemical oscillation (Sect. “[Biochemical Oscillations: Segmentation](#)”), can, when one or more of the components is diffusible, also give rise to standing or traveling spatial periodicities of chemical concentration (Sect. “[Reaction-Diffusion Mechanisms: Body Axis Formation](#)”). (See Fig. 14 for a simple dynamical system that exhibits temporal oscillation or standing waves, depending on whether or not diffusion is permitted). This connection between oscillatory and reaction-diffusion mechanisms can potentially unify the different segmentation mechanisms found in short and long germ-band insects. This would be quite straightforward if the *Drosophila* embryo were patterned by a reaction-diffusion system, which can readily give rise to a series of chemical standing waves (“stripes”).

While the stripe patterns of the pair-rule genes in *Drosophila* indeed have the distinct appearance of being produced by a reaction-diffusion system (Fig. 13), at least one of the stripes is generated instead by a complex set of interactions among transcription factors in the syncytial embryo. The formation of *eve* stripe number 2 requires the existence of sequences in the *eve* promoter that switch on the *eve* gene in response to a set of spatially-distributed morphogens that under normal circumstances have the requisite values only at the stripe 2 position (Fig. 15) [88,89]. In particular, these promoter sequences respond to specific combinations of products of the “gap” genes (e. g., *giant*, *knirps*, the embryonically-produced version of *hunchback*). These proteins are transcription factors that are expressed in a spatially-nonuniform fashion and act as activators and competitive repressors of the pair-rule gene promoters (also see discussion of the Keller model in Sect. “[Dynamic Multistability: Cell Differentiation](#)”). The patterned expression of the gap genes, in turn, is controlled by the responses of their own promoters to particular combinations of products of “maternal” genes (e. g., *bicoid*, *staufer*), which are distributed as gradients along the embryo at even earlier stages (Fig. 13). As the category name suggests, the maternal gene products are deposited in the egg during oogenesis.

While the expression of *engrailed* along the posterior margin of each developing segment is a constant theme during development of arthropods, the expression patterns of pair-rule genes is less well-conserved over evolution [15]. The accepted view is that the short germ-band “sequential” mode is the more ancient way of making segments, and that the long germ-band “simultaneous” mode seen in *Drosophila*, which employs pair-rule stripes, is more recently evolved [53].





Biological Development and Evolution, Complexity and Self-organization in, Figure 14

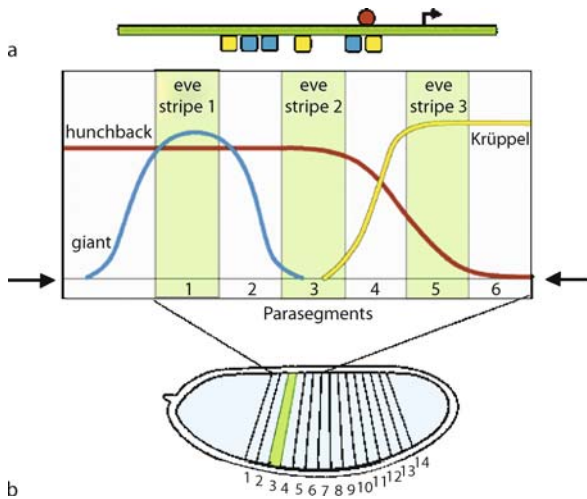
Example of a network that can produce (for the same parameter values) sequential stripes when acting as an intracellular biochemical clock in a one-dimensional cellularized blastoderm with a posterior proliferative zone, and simultaneously forming stripes when acting in a one-dimensional diffusion-permissive syncytium. The network is shown in the central box. Arrows indicate positive regulation; lines terminating in circles, negative regulation. In the upper boxes, the equations governing each of the two behaviors are shown. The four genes involved in the central network diagram, as well as their levels of expression, are denoted by  $g_1, g_2, g_3$ , and  $g_4$ . In the reaction-diffusion case,  $g_1$  and  $g_2$  can diffuse between nuclei (note that the two sets of equations differ only in the presence of a diffusion term for the products of genes 1 and 2). The lower boxes indicate the levels of expression of gene 2 for the two systems. For the intracellular clock the x-axis represents time, and the indicated time variation of the gene's expression level refers to a single nucleus. In the reaction-diffusion system this axis represents space, the distance along a chain of 81 nuclei. In the pattern shown on the right, the initial condition consisted of all gene product levels set to zero except gene 1 in the central of 81 nuclei, which was assigned a small value (the exact quantity did not affect the pattern). All parameters (except for  $D_1$  and  $D_2$ ) had the same value when generating the two patterns (for details see Salazar-Ciudad et al. [85], from which the figure was redrawn). Figure modified and reprinted, with permission, from [21]

No intracellular oscillations have thus far been identified during segmentation of invertebrates, unlike the case in vertebrates such as the mouse, chicken, and zebrafish. However, the sequential appearance of gene expression stripes from the posterior proliferative zone of short germ-band insects and other arthropods such as spiders, has led to the suggestion that these patterns in fact arise from a segmentation clock like that found to control vertebrate somitogenesis [95].

Based on the experimental findings, theoretical considerations and evolutionary inferences described above, it has been hypothesized that the ancestor of *Drosophila* gen-

erated its segments by a reaction-diffusion system [60,85], built upon the presumed chemical oscillator underlying short germ-band segmentation. Modern-day *Drosophila* contains (retains, according to the hypothesis) the ingredients for this type of mechanism. Specifically, in the syncytial embryo several of the pair-rule proteins (e.g., *eve*, *ftz*) diffuse over short distances among the cell nuclei that synthesize their mRNAs, and positively regulate their own synthesis [30,38].

The hypothesized evolutionary scenario can be summarized as follows: the appearance of the syncytial mode of embryogenesis converted the chemical oscillation-de-



### Biological Development and Evolution, Complexity and Self-organization in, Figure 15

**a** Schematic representation of a portion of the *even-skipped* gene, including the promoter and transcription start site. Contained within the promoter is a subregion ("stripe 2 enhancer") responsible for the expression of the second eve stripe. This sequence contains binding sites for gap gene-class transcription factors that positively (hunchback, red) or negatively (giant, blue; Krüppel, orange) regulate eve expression. **b** Illustration of *Drosophila* syncytial blastula, showing distribution of giant, hunchback and Krüppel proteins in region of first six prospective parasegments. At the position of prospective parasegment 3 (the posterior half of the last head segment plus the anterior half of the first thoracic – upper body – segment) the levels of the activator hunchback is high and those of the inhibitors giant and Krüppel are low. This induces eve stripe 2. To either side of this position giant and Krüppel are high, restricting expression of eve to a narrow band. (a based on Small et al. [88]; b after Wolpert [104], 2002. Reprinted, with permission, from [21])

pendent temporal mechanism found in the more ancient short germ-band insects into the spatial standing wave mechanism seen in the more recently evolved long germ-band forms. The pair-rule stripes formed by the proposed reaction-diffusion mechanism would have been equivalent to one another. That is, they would have been generated by a single mechanism acting in a spatially-periodic fashion, not by stripe-specific molecular machinery (see above). Thus despite suggesting an underlying physical connection between modern short germ-band segmentation and segmentation in the presumed ancestor of long germ-band forms, this hypothesis introduces a new puzzle of its own: Why does modern-day *Drosophila* not use a reaction-diffusion mechanism to produce its segments?

Genes are always undergoing random mutation, but morphological change does not always track genetic change. Particularly interesting are those cases in which

the outward form of a body plan or organ does not change, but its genetic "underpinning" does. One example of this is seen in the role of the transcription factor sloppy-paired (slp) in the beetle *Tribolium* and *Drosophila*. In relation to, but in contrast with the paired (prd) protein, which activates *engrailed* and *wingless* in the same odd-even parasegmental register in both species, slp has an opposite parasegmental register in *Tribolium* and *Drosophila* [8].

This situation can result from a mechanism that is initially "plastic", i. e., having variable morphological outcomes despite utilizing the same genes (as physically based mechanisms would typically be), followed by a particular kind of natural selection, termed "canalizing selection" by Waddington [101] (see also [86]). Canalizing selection will preserve those randomly acquired genetic alterations that happen to enhance the reliability of a developmental process. Development would thereby become more complex at the molecular level, but correspondingly more resistant ("robust") to external perturbations or internal noise that could disrupt non-reinforced physical mechanisms of determination.

If the striped expression of pair-rule genes in the ancestor of modern *Drosophila* was generated by a reaction-diffusion mechanism, this inherently variable developmental system would have been a prime candidate for canalizing evolution. The elaborate systems of multiple promoter elements responsive to pre-existing, nonuniformly distributed molecular cues (e. g., maternal and gap gene products), seen in *Drosophila* is therefore not inconsistent with this pattern having originated as a reaction-diffusion process.

In light of the discussion in the previous paragraphs, a tentative answer to why modern-day *Drosophila* does not use a reaction-diffusion mechanism to produce its segments is that such pattern-forming systems are inherently unstable to environmental and random genetic changes, and would therefore, under pressure of natural selection, have been replaced, or at least reinforced, by more hierarchically-organized genetic control systems. We do not know, at present, whether the short germ-band mode of segmentation is similarly hierarchically reinforced.

This evolutionary hypothesis has been examined computationally in a simple physical model by Salazar-Ciudad and coworkers [84]. The model consists of a fixed number of nuclei arranged in a row within a syncytium. Each nucleus has the same genome (i. e., the same set of genes), and the same *epigenetic system* (i. e., the same activating and inhibitory relationships among these genes). The genes in these networks specify receptors or transcription factors that act within the cells that produce them, or paracrine factors that diffuse between cells (Fig. 16). These genes in-

teract with each other according to a set of simple rules that embody unidirectional interactions in which an upstream gene activates a downstream one, as well as reciprocal interactions, in which genes feed back (via their products) on each other's activities. This formalism was based on a similar one devised by Reinitz and Sharp [80], who considered the specific problem of segmentation in the *Drosophila* embryo.

One may ask whether a system of this sort, with particular values of the gene-gene coupling and diffusion constants, can form a spatial pattern of differentiated cells. Salazar-Ciudad and coworkers performed simulations on systems containing 25 nuclei, and a pattern was considered to arise if after some time different nuclei stably expressed one or more of the genes at different levels. The system was isolated from external influences ("zero-flux boundary conditions" were used, that is the boundaries of the domain were impermeable to diffusible morphogens), and initial conditions were set such that at  $t = 0$  the levels of all gene products had zero value except for that of an arbitrarily chosen gene, which had a non-zero value in the nucleus at the middle position.

Isolated single nuclei, or isolated patches of contiguous nuclei expressing a given gene, are the one-dimensional analogue of isolated stripes of gene expression in a two-dimensional sheet of nuclei, such as those in the *Drosophila* embryo prior to cellularization [84]. Regardless of whether a particular system initially gives rise to a pattern, if within the model it evolves by random changes in the values of the gene-gene coupling constants it may acquire or lose pattern-forming ability over subsequent "generations". Salazar-Ciudad and coworkers therefore used such systems to perform a large number of computational "evolutionary experiments". Here we will describe a few of these that are relevant to the questions raised by the evolution of segmentation in insects, discussed above.

It had earlier been determined that the core mechanisms responsible for all stable patterns fell into two non-overlapping topological categories [83]. These were referred to as "emergent" and "hierarchical". Mechanisms in which reciprocal positive and negative feedback interactions give rise to the pattern, are emergent. In contrast, those that form patterns by virtue of the unidirectional influence of one gene on the next, in an ordered succession ("feed-forward" loops [54]), as in the "maternal gene induces gap gene induces pair-rule gene" scheme described above for early steps of *Drosophila* segmentation, are hierarchical. Emergent systems are equivalent to excitable dynamical systems, like the transcription factor networks discussed in Sect. "Dynamic Multistability: Cell Differentiation" and the reaction-diffusion systems dis-

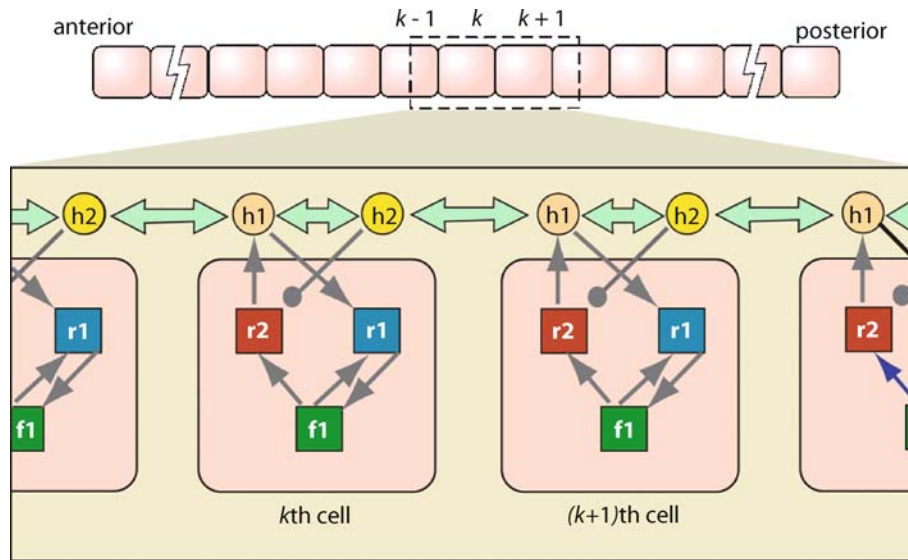
cussed in Sect. "Reaction-Diffusion Mechanisms: Body Axis Formation".

In their computational studies of the evolution of developmental mechanisms the investigators found that randomly-chosen emergent networks were much more likely than randomly-chosen hierarchical networks to generate complex patterns, i. e., patterns with three or more (one-dimensional) stripes. This was taken to suggest that the evolutionary origination of complex forms, such as segmented body plans, would have been more readily achieved in a world of organisms in which dynamical (e. g., multistable, reaction-diffusion, oscillatory) mechanisms were also active [84] and not just mechanisms based on genetic hierarchies.

Following up on these observations, Salazar-Ciudad and co-workers performed a set of computational studies on the evolution of developmental mechanisms after a pattern had been originated. First they identified emergent and hierarchical networks that produced a particular pattern – e. g., three "stripes" [84]. (Note that patterns themselves are neither emergent nor hierarchical – these terms apply to the mechanisms that generate them). They next asked if given networks would "breed true" phenotypically, despite changes to their underlying circuitry. That is, would their genetically-altered "progeny" exhibit the same pattern as the unaltered version? Genetic alterations in these model systems consisted of point mutations (i. e., changes in the value of a gene-gene coupling constants), duplications, recombinations (i. e., interchange of coupling constant values between pairs of genes), and the acquisition of new interactions (i. e., a coupling constant that was initially equal to zero was randomly assigned a small positive or negative value).

It was found that hierarchical networks were much less likely to diverge from the original pattern (after undergoing simulated evolution as described) than emergent networks [84]. That is to say, a given pattern would be more robust (and thus evolutionarily more stable) under genetic mutation if it were generated by a hierarchical, than an emergent network. Occasionally it was observed that networks that started out as emergent were converted into hierarchical networks with the same number of stripes. The results of Salazar-Ciudad and coworkers on how network topology influences evolutionary stability imply that these "converted" networks would produce the original pattern in the face of further genetic evolution. Recall that this is precisely the scenario that was hypothesized above to have occurred during the evolution of *Drosophila* segmentation [84].

Subject to the caveats to what is obviously a highly schematic analysis, the possible implications of these com-



**Biological Development and Evolution, Complexity and Self-organization in, Figure 16**

Schematic of the gene-gene interactions in the model of Salazar-Ciudad et al. [84]. A line of cells is represented at the top. Below, types of genes are illustrated. Genes whose products act solely on or within the cells that produce them (receptors:  $r$ , transcription factors:  $f$ ) are represented by squares; diffusible factors (paracrine factors or hormones:  $h$ ) that pass between cells and enable genes to affect one another's activities, are represented by circles. Activating and inhibitory interactions are denoted by small arrows and lines terminating in circles, respectively. Double-headed green arrows denote diffusion. (After Salazar-Ciudad et al. [84]) Reprinted, with permission, from [21]

putational experiments for the evolution of segmentation in long germ-band insects are the following: (i) if the ancestral embryo indeed generated its seven-stripe pair-rule protein patterns by a reaction-diffusion mechanism, and (ii) if this pattern was sufficiently well-adapted so as to provide a premium on breeding true, then (iii) genetic changes that preserved the pattern but converted the underlying network from an emergent one to a hierarchic one (as seen in present-day *Drosophila*) would have been favored.

As this discussion of the short germ-band/long germ-band puzzle suggests, developmental mechanisms can change over the course of evolution while their morphological outcomes remain constant. While forms (few tissue layers, segments, tubes, and so forth) remain simple and reminiscent of structures that can be generated by generic physical processes, the associated gene-expression networks can become increasingly more complex. This is because the latter tend to evolve so as to bring about a developmental result in a fashion protected against perturbation from external and internal variability, including, in the latter case, mutation and noise.

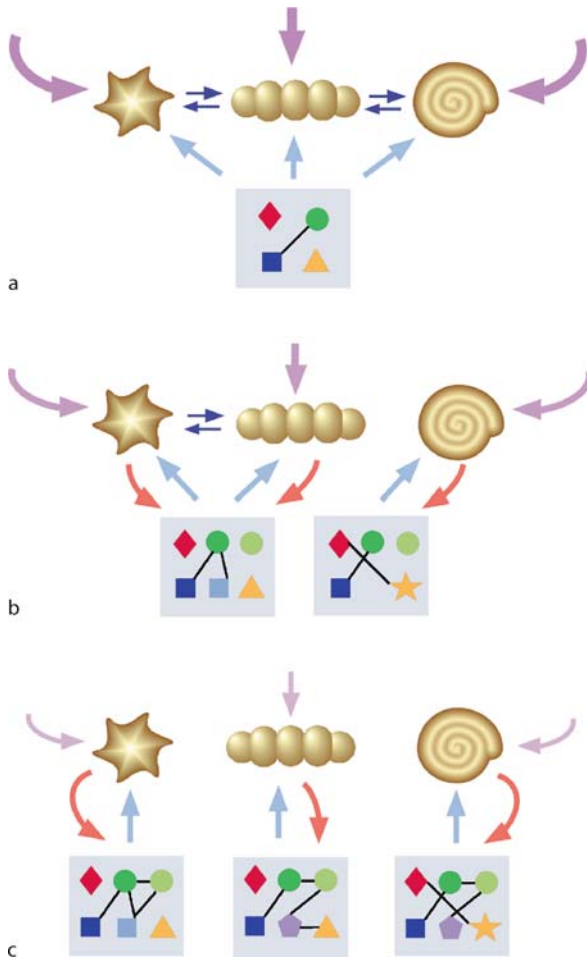
The evolutionary changes that will lead to increasingly canalized or robust outcomes can include genetic redundancies [70,103] such as duplication of developmental control genes [32] and multiplication of their regulatory

elements [88,89], as well as the utilization of chaperone proteins as “phenotypic capacitors” [81]. The effect of this mode of evolution is to convert organisms into more stable types that are less morphologically plastic than ones molded by relatively unconstrained physical mechanisms (Fig. 17).

### Future Directions

While all living systems are part of the physical world and thus must obey its physical laws, it is not self-evident that all biological problems can be adequately addressed by the concepts of physics as we know them today. On a fundamental level this is because whereas the laws of physics have not changed during the lifetime of the Earth, living systems have. Evolution is the hallmark of life, but has little relevance for physics. As living systems evolved, some lineages transformed from single cells into multicellular organisms with a myriad of different interacting components organized on multiple levels. Such systems are considerably more complex than even the most elaborate materials studied by physicists. Thus, on the practical level, the tools of present day physics do not make it possible to study a fully developed adult organism with all its parts or even the subcellular economy of intracellular biochemical networks.





The developing embryo, however, considered at the appropriate level of resolution, is not beyond physical comprehension of its main episodes and transitions. Each of the processes discussed in the previous sections, multistability of dynamical systems, differential adhesion and interfacial tension, chemical oscillation and symmetry-breaking by reaction-diffusion interaction, have their counterparts in the world of nonliving chemically and mechanically excitable [57] “soft matter” [16]. These descriptions of the behavior of matter at an intermediate scale (i. e., between molecules and bulk materials), represent elegant simplifications of highly complex many-body interactions. They are based on models developed by physicists using sophisticated mathematical techniques and systems-level approaches (nonlinear dynamics, and scaling and renormalization group theory, to name a few), which, for the most part, were unavailable before the mid-20th century.

Embryos are approachable by these physical principles because they are, to a major extent, chemically and me-

#### Biological Development and Evolution, Complexity and Self-organization in, Figure 17

Schematic representation of evolutionary partitioning of a morphologically plastic ancestral organism into distinct morphotypes associated with unique genotypes. **a** A hypothetical primitive metazoan is shown with a schematic representation of its genome in the box below it. Specific genes are shown as colored geometric objects; interactions between them by lines. Determinants of the organism’s form include the materials provided by expression of its genes (light blue arrows) and the external environment, including physical causes (purple arrows) acting on its inherent physical properties. At this stage of evolution the organism is highly plastic, exhibiting several condition-dependent forms that are mutually interconvertible (dark blue arrows). **b** Descendants of organism in **a** after some stabilizing evolution. Gene duplication, mutation, etc. have led to non-interbreeding populations that are biased toward subsets of the original morphological phenotypes. Determinants of form are still gene products and the physical environment, but the effect of the latter has become attenuated (smaller, fainter purple arrows) as development has become more programmatic. There is also causal influence of the form on the genotype (orange arrows), exerted over evolutionary time, as ecological establishment of forms filters out those variant genotypes that are not compatible with the established form. Some morphotypes remain interconvertible at this stage of evolution. **c** Modern organisms descended from those in **b**. Further stabilizing evolution has now led to each morphotype being uniquely associated with its own genotype. Physical causation is even more attenuated. Note that in this idealized example the forms have remained unchanged while the genes and mechanisms for generating the forms have undergone extensive evolution. (From Newman et al. [67]; with permission from the University of the Basque Country Press)

chanically excitable soft matter. Developing tissues consist of cells containing complex biochemical networks which are capable of switching between a limited number of stationary or oscillatory states. Their cells are mobile with respect to one another, giving them liquid-like properties, and they secrete diffusible signal molecules, permitting them to become spatially heterogeneous by straightforward physical means.

As we have seen, the forms generated by self-organizing physical processes of soft matter will become increasingly regulated by genetic hierarchies over the course of evolution. But the major structural motifs of the animal body and its organs (hollow, multilayered cell masses, segments, tubes) had their origin more than 500 million years ago [12,61,68]. During this period “before programs” [67], certain molecules of the “developmental-genetic toolkit” [7] – cadherins, Notch and its ligands, secreted molecules such as Wnt, BMP, hedgehog, FGF – though having evolved in the world of single cells, took on, when present in a multicellular context, the role of mediating intermediate-scale physical effects [66]. Once this new set (for organisms) of physical processes was thus mobi-



lized, morphological evolution, unconstrained in this early period by highly integrated programs, could proceed very rapidly. The otherwise puzzling high degree of conservation, in disparate animal forms, of the toolkit genes for morphological development is to be expected under these circumstances [64].

Although multicellular forms are relatively simple, and the physical processes that likely first brought them about comprehensible, the molecular mechanisms of development in modern-day organisms are enormously complex. The challenge for both developmental and evolutionary studies in the future, then, will be to experimentally disentangle and conceptually reintegrate the roles physics and developmental genetics in modern organisms and their ancestors.

## Bibliography

### Primary Literature

- Artavanis-Tsakonas S, Rand MD, Lake RJ (1999) Notch signaling: cell fate control and signal integration in development. *Science* 284:770–6
- Aulehla A, Wehrle C, Brand-Saberi B, Kemler R, Gossler A, Kanzler B, Herrmann BG (2003) Wnt3a plays a major role in the segmentation clock controlling somitogenesis. *Dev Cell* 4:395–406
- Bateson W (1894) *Materials for the study of variation*. Macmillan, London
- Boissonade J, Dulos E, DeKepper P (1994) Turing patterns: From myth to reality. In: Kapral R, Showalter K (eds) *Chemical waves and patterns*. Boston, Kluwer
- Borisuk MT, Tyson JJ (1998) Bifurcation analysis of a model of mitotic control in frog eggs. *J Theor Biol* 195:69–85
- Brown SJ, Patel NH, Denell RE (1994) Embryonic expression of the single *Tribolium* engrailed homolog. *Dev Genet* 15:7–18
- Carroll SB, Grenier JK, Weatherbee SD (2001) *From DNA to diversity: molecular genetics and the evolution of animal design*. Blackwell Science, Malden
- Choe CP, Brown SJ (2007) Evolutionary flexibility of pair-rule patterning revealed by functional analysis of secondary pair-rule genes, paired and sloppy-paired in the short-germ insect, *Tribolium castaneum*. *Dev Biol* 302:281–94
- Christen B, Slack J (1999) Spatial response to fibroblast growth factor signalling in *Xenopus* embryos. *Development* 126:119–125
- Cinquin O, Demongeot J (2005) High-dimensional switches and the modelling of cellular differentiation. *J Theor Biol* 233:391–411
- Clements D, Friday RV, Woodland HR (1999) Mode of action of VegT in mesoderm and endoderm formation. *Development* 126:4903–11
- Conway Morris S (2006) Darwin's dilemma: The realities of the cambrian 'explosion'. *Philos Trans Royal Soc Lond B Biol Sci* 361:1069–83
- Cooke J, Zeeman EC (1976) A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. *J Theor Biol* 58:455–76
- Davidson EH (2006) *The regulatory genome: gene regulatory networks in development and evolution*. Elsevier Academic Press, Amsterdam
- Dawes R, Dawson I, Falciani F et al (1994) Dax, a locust Hox gene related to fushi tarazu but showing no pair-rule expression. *Development* 120:1561–72
- de Gennes PG (1992) Soft matter. *Science* 256:495–497
- DeMarais AA, Moon RT (1992) The armadillo homologs beta-catenin and plakoglobin are differentially expressed during early development of *Xenopus laevis*. *Dev Biol* 153:337–46
- Dubrulle J, McGrew MJ, Pourquié O (2001) FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation. *Cell* 106:219–32
- Duguay D, Foty RA, Steinberg MS (2003) Cadherin-mediated cell adhesion and tissue segregation: qualitative and quantitative determinants. *Dev Biol* 253:309–23
- Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* 403:335–8
- Forgacs G, Newman SA (2005) *Biological physics of the developing embryo*. Cambridge University Press, Cambridge
- Foty RA, Pflieger CM, Forgacs G, Steinberg MS (1996) Surface tensions of embryonic tissues predict their mutual envelopment behavior. *Development* 122:1611–1620
- Frasch M, Levine M (1987) Complementary patterns of even-skipped and fushi tarazu expression involve their differential regulation by a common set of segmentation genes in *Drosophila*. *Genes Dev* 1:981–95
- Giancotti FG, Ruoslahti E (1999) Integrin signaling. *Science* 285:1028–32
- Gilbert SF (2006) *Developmental biology*, 8th edn. Sinauer Associates, Sunderland
- Giudicelli F, Ozbudak EM, Wright GJ, Lewis J (2007) Setting the tempo in development: an investigation of the zebrafish somite clock mechanism. *PLoS Biol* 5:e150
- Goodwin BC (1963) *Temporal organization in cells; a dynamic theory of cellular control processes*. Academic Press, London
- Green J (2002) Morphogen gradients, positional information, and *Xenopus*: Interplay of theory and experiment. *Dev Dyn* 225:392–408
- Gurdon JB (1988) A community effect in animal development. *Nature* 336:772–4
- Harding K, Hoey T, Warrior R et al (1989) Autoregulatory and gap gene response elements of the even-skipped promoter of *Drosophila*. *EMBO J* 8:1205–12
- Harland R, Gerhart J (1997) Formation and function of Spemann's organizer. *Annu Rev Cell Dev Biol* 13:611–67
- Holland PW (1999) Gene duplication: past, present and future. *Semin Cell Dev Biol* 10:541–7
- Holley SA, Geisler R, Nusslein-Volhard C (2000) Control of her1 expression during zebrafish somitogenesis by a Delta-dependent oscillator and an independent wave-front activity. *Genes Dev* 14:1678–1690
- Holley SA, Julich D, Rauch GJ et al (2002) her1 and the notch pathway function within the oscillator mechanism that regulates zebrafish somitogenesis. *Development* 129:1175–83
- Howard K, Ingham P (1986) Regulatory interactions between the segmentation genes fushi tarazu, hairy, and engrailed in the *Drosophila* blastoderm. *Cell* 44:949–57
- Ingham PW (1988) The molecular genetics of embryonic pattern formation in *Drosophila*. *Nature*. 335:25–34

37. Irvine KD, Wieschaus E (1994) Cell intercalation during *Drosophila* germband extension and its regulation by pair-rule segmentation genes. *Development* 120:827–41
38. Ish-Horowicz D, Pinchin SM, Ingham PW et al (1989) Autocatalytic ftz activation and instability induced by ectopic ftz expression. *Cell* 57:223–232
39. Israelachvili JN (1991) Intermolecular and surface forces. Academic Press, London
40. Itow T (1986) Inhibitors of DNA synthesis change the differentiation of body segments and increase the segment number in horseshoe crab embryos. *Roux's Arch Dev Biol* 195:323–333
41. Jablonka E, Lamb MJ (1995) Epigenetic inheritance and evolution. Oxford University Press, Oxford
42. Kaneko K (2003) Organization through intra-inter dynamics. In: Müller GB, Newman SA (eds) *Origination of Organismal Form: Beyond the Gene in Developmental and Evolutionary Biology*. MIT Press, Cambridge, pp 195–220
43. Kaneko K (2006) *Life: an introduction to complex systems biology*. Springer, New York
44. Kaneko K, Yomo T (1994) Cell division, differentiation and dynamic clustering. *Physica D* 75:89–102
45. Karr TL, Weir MP, Ali Z et al (1989) Patterns of engrailed protein in early *Drosophila* embryos. *Development* 105:605–612
46. Kauffman SA (1969) Metabolic stability and epigenesis in randomly constructed genetic nets. *J Theor Biol* 22:437–67
47. Keller AD (1995) Model genetic circuits encoding autoregulatory transcription factors. *J Theor Biol* 172:169–85
48. Kerszberg M, Wolpert L (1998) Mechanisms for positional signaling by morphogen transport: a theoretical study. *J Theor Biol* 191:103–14
49. Lander AD (2007) *Morpheus unbound: reimagining the morphogen gradient*. *Cell* 128:245–56
50. Lawrence PA (1992) *The making of a fly: the genetics of animal design*. Blackwell Scientific Publications, Oxford
51. Lengyel I, Epstein IR (1992) A chemical approach to designing Turing patterns in reaction-diffusion systems. *Proc Natl Acad Sci USA* 89:3977–9
52. Lewis J (2003) Autoinhibition with transcriptional delay: a simple mechanism for the zebrafish somitogenesis oscillator. *Curr Biol* 13:1398–408
53. Liu PZ, Kaufman TC (2005) Short and long germ segmentation: unanswered questions in the evolution of a developmental mode. *Evol Dev* 7:629–46
54. Mangan S, Zaslaver A, Alon U (2003) The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. *J Mol Biol* 334:197–204
55. McDowell N, Gurdon JB, Grainger DJ (2001) Formation of a functional morphogen gradient by a passive process in tissue from the early *Xenopus* embryo. *Int J Dev Biol* 45 (1 Spec No):199–207
56. Meinhardt H (2001) Organizer and axes formation as a self-organizing process. *Int J Dev Biol* 45(1 Spec No):177–88
57. Mikhailov AS (1990) *Foundations of synergetics I*. Springer, Berlin
58. Monk NA (2003) Oscillatory expression of *Hes1*, *p53*, and *NF-kappaB* driven by transcriptional time delays. *Curr Biol* 13:1409–13
59. Morisco C, Seta K, Hardt SE et al (2001) Glycogen synthase kinase 3 $\beta$  regulates GATA4 in cardiac myocytes. *J Biol Chem* 276:28586–97
60. Newman SA (1993) Is segmentation generic? *BioEssays* 15:277–283
61. Newman SA (1994) Generic physical mechanisms of tissue morphogenesis: A common basis for development and evolution. *J Evol Bio* 7:467–488
62. Newman SA (1998) Epithelial morphogenesis: a physico-evolutionary interpretation. In: Chuong C-M (ed) *Molecular Basis of Epithelial Appendage Morphogenesis*. Landes, Austin, pp 341–358
63. Newman SA (2003) From physics to development: the evolution of morphogenetic mechanisms. In: Müller GB, Newman SA (eds) *Origination of Organismal Form: Beyond the Gene in Developmental and Evolutionary Biology*. MIT Press, Cambridge
64. Newman SA (2006) The developmental-genetic toolkit and the molecular homology-analogy paradox. *Biol Theory* 1:12–16
65. Newman SA (2007) William Bateson's physicalist ideas. In: Laubichler M, Maienschein J (eds) *From Embryology to Evo-Devo: a History of Evolutionary Development*. Cambridge, MA, MIT Press, pp 83–107
66. Newman SA, Bhat R (2008) Dynamical patterning modules: physico-genetic determinants of morphological development and evolution. *Phys Biol* 5:15008
67. Newman SA, Forgacs G, Müller GB (2006) Before programs: The physical origination of multicellular forms. *Int J Dev Biol* 50:289–99
68. Newman SA, Müller GB (2000) Epigenetic mechanisms of character origination. *J Exp Zool* 288:304–17
69. Nieuwkoop PD (1969) The formation of mesoderm in Urodelean amphibians. I. Induction by the endoderm. *Wilhelm Roux' Arch Entw Mech Org* 162:341–373
70. Nowak MA, Boerlijst MC, Cooke J, Smith JM (1997) Evolution of genetic redundancy. *Nature* 388:167–71
71. Oates AC, Ho RK (2002) *Hairy/E(spl)*-related (*Her*) genes are central components of the segmentation oscillator and display redundancy with the *Delta/Notch* signaling pathway in the formation of anterior segmental boundaries in the zebrafish. *Development* 129:2929–46
72. Palmeirim I, Henrique D, Ish-Horowicz D et al (1997) Avian *hairy* gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. *Cell* 91:639–48
73. Patel NH (1994) Developmental evolution: insights from studies of insect segmentation. *Science* 266:581–590
74. Patel NH, Ball EE, Goodman CS (1992) Changing role of even-skipped during the evolution of insect pattern formation. *Nature* 357:339–342
75. Patel NH, Kornberg TB, Goodman CS (1989) Expression of engrailed during segmentation in grasshopper and crayfish. *Development* 107:201–212
76. Pennisi E (2003) A low number wins the GeneSweep pool. *Science* 300:1484
77. Phillips HM (1969) Equilibrium measurements of embryonic cell adhesiveness: Physical formulation and testing of the differential adhesion hypothesis. Ph D thesis. Johns Hopkins University
78. Pourquié O (2003) The segmentation clock: converting embryonic time into spatial pattern. *Science* 301:328–30
79. Primmatt DR, Norris WE, Carlson GJ et al (1989) Periodic segmental anomalies induced by heat shock in the chick embryo are associated with the cell cycle. *Development* 105:119–30

80. Reinitz J, Mjolsness E, Sharp DH (1995) Model for cooperative control of positional information in *Drosophila* by bicoid and maternal hunchback. *J Exp Zool* 271:47–56
81. Rutherford SL, Lindquist S (1998) Hsp90 as a capacitor for morphological evolution. *Nature* 396:336–42
82. Sakuma R, Ohnishi Yi Y, Meno C et al (2002) Inhibition of Nodal signalling by Lefty mediated through interaction with common receptors and efficient diffusion. *Genes Cells* 7:401–12
83. Salazar-Ciudad I, Garcia-Fernandez J, Sole RV (2000) Gene networks capable of pattern formation: from induction to reaction-diffusion. *J Theor Biol* 205:587–603
84. Salazar-Ciudad I, Newman SA, Solé R (2001) Phenotypic and dynamical transitions in model genetic networks. I. Emergence of patterns and genotype-phenotype relationships. *Evol Dev* 3:84–94
85. Salazar-Ciudad I, Solé R, Newman SA (2001) Phenotypic and dynamical transitions in model genetic networks. II. Application to the evolution of segmentation mechanisms. *Evol Dev* 3:95–103
86. Schmalhausen II (1949) Factors of evolution. Blakiston, Philadelphia
87. Schulte-Merker S, Smith JC (1995) Mesoderm formation in response to Brachyury requires FGF signalling. *Curr Biol* 5:62–7
88. Small S, Blair A, Levine M (1992) Regulation of even-skipped stripe 2 in the *Drosophila* embryo. *EMBO J* 11:4047–4057
89. Small S, Kraut R, Hoey T et al (1991) Transcriptional regulation of a pair-rule stripe in *Drosophila*. *Genes Dev* 5:827–39
90. Solnica-Krezel L (2003) Vertebrate development: taming the nodal waves. *Curr Biol* 13:R7–9
91. Spemann H, Mangold H (1924) Über Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren. *Wilhelm Roux' Arch Entw Mech Org* 100:599–638
92. St Johnston D, Nusslein-Volhard C (1992) The origin of pattern and polarity in the *Drosophila* embryo. *Cell* 68:201–19
93. Steinberg MS (1963) Reconstruction of tissues by dissociated cells. Some morphogenetic tissue movements and the sorting out of embryonic cells may have a common explanation. *Science* 141:401–8
94. Stern CD, Bellairs R (1984) Mitotic activity during somite segmentation in the early chick embryo. *Anat Embryol (Berl)* 169:97–102
95. Stollwerk A, Schoppmeier M, Damen WG (2003) Involvement of Notch and Delta genes in spider segmentation. *Nature* 423:863–5
96. Strogatz SH (1994) Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering. Perseus Pub, Cambridge
97. Sun B, Bush S, Collins-Racie L et al (1999) derriere: a TGF-beta family member required for posterior development in *Xenopus*. *Development* 126:1467–1482
98. Tsarfaty I, Resau JH, Rulong S, Keydar I, Faletto DL, Vande Woude GF (1992) The met proto-oncogene receptor and lumen formation. *Science* 257:1258–61
99. Turing AM (1952) The chemical basis of morphogenesis. *Phil Trans Royal Soc Lond B* 237:37–72
100. Van Obberghen-Schilling E, Roche NS, Flanders KC et al (1988) Transforming growth factor beta-1 positively regulates its own expression in normal and transformed cells. *J Biol Chem* 263:7741–7746
101. Waddington CH (1957) The Strategy of the Genes. Allen and Unwin, London
102. Winfree AT (1980) The geometry of biological time. Springer, New York
103. Wilkins AS (1997) Canalization: a molecular genetic perspective. *BioEssays* 19:257–262
104. Wolpert L (2002) Principles of development. Oxford University Press, Oxford New York

## Books and Reviews

- Meinhardt H (1982) Models of biological pattern formation. Academic, New York
- Müller GB, Newman SA (2003) Origination of organismal form: beyond the gene in developmental and evolutionary biology. MIT Press, Cambridge, pp 221–239
- Newman SA, Comper WD (1990) 'Generic' physical mechanisms of morphogenesis and pattern formation. *Development* 110: 1–18

## Biological Fluid Dynamics, Non-linear Partial Differential Equations

ANTONIO DESIMONE<sup>1</sup>, FRANÇOIS ALOUGES<sup>2</sup>,  
ALINE LEFEBVRE<sup>2</sup>

<sup>1</sup> SISSA-International School for Advanced Studies,  
Trieste, Italy

<sup>2</sup> Laboratoire de Mathématiques, Université Paris-Sud,  
Orsay cedex, France

## Article Outline

[Glossary](#)  
[Definition of the Subject](#)  
[Introduction](#)  
[The Mathematics of Swimming](#)  
[The Scallop Theorem Proved](#)  
[Optimal Swimming](#)  
[The Three-Sphere Swimmer](#)  
[Future Directions](#)  
[Bibliography](#)

## Glossary

**Swimming** The ability to advance in a fluid in the absence of external propulsive forces by performing cyclic shape changes.

**Navier-Stokes equations** A system of partial differential equations describing the motion of a simple viscous incompressible fluid (a Newtonian fluid)

$$\rho \left( \frac{\partial v}{\partial t} + (v \cdot \nabla) v \right) = -\nabla p + \eta \Delta v$$

$$\operatorname{div} v = 0$$

where  $v$  and  $p$  are the velocity and the pressure in the fluid,  $\rho$  is the fluid density, and  $\eta$  its viscosity. For simplicity external forces, such as gravity, have been dropped from the right hand side of the first equation, which expresses the balance between forces and rate of change of linear momentum. The second equation constrains the flow to be volume preserving, in view of incompressibility.

**Reynolds number** A dimensionless number arising naturally when writing Navier–Stokes equations in non-dimensional form. This is done by rescaling position and velocity with  $x^* = x/L$  and  $v^* = v/V$ , where  $L$  and  $V$  are characteristic length scale and velocity associated with the flow. Reynolds number (Re) is defined by

$$\text{Re} = \frac{VL\rho}{\eta} = \frac{VL}{\nu}$$

where  $\nu = \eta/\rho$  is the kinematic viscosity of the fluid, and it quantifies the relative importance of inertial versus viscous effects in the flow.

**Steady Stokes equations** A system of partial differential equations arising as a formal limit of Navier–Stokes equations when  $\text{Re} \rightarrow 0$  and the rate of change of the data driving the flow (in the case of interest here, the velocity of the points on the outer surface of a swimmer) is slow

$$\begin{aligned} -\eta\Delta v + \nabla p &= 0 \\ \text{div } v &= 0. \end{aligned}$$

Flows governed by Stokes equations are also called creeping flows.

**Microscopic swimmers** Swimmers of size  $L = 1\text{ }\mu\text{m}$  moving in water ( $\nu \sim 1\text{ mm}^2/\text{s}$  at room temperature) at one body length per second give rise to  $\text{Re} \sim 10^{-6}$ . By contrast, a 1 m swimmer moving in water at  $V = 1\text{ m/s}$  gives rise to a Re of the order  $10^6$ .

**Biological swimmers** Bacteria or unicellular organisms are microscopic swimmers; hence their swimming strategies cannot rely on inertia. The devices used for swimming include rotating helical flagella, flexible tails traversed by flexural waves, and flexible cilia covering the outer surface of large cells, executing oar-like rowing motion, and beating in coordination. Self propulsion is achieved by cyclic shape changes described by time periodic functions (*swimming strokes*). A notable exception is given by the rotating flagella of bacteria, which rely on a submicron-size rotary motor capable of turning the axis of an helix without alternating between clockwise and anticlockwise directions.

**Swimming microrobots** Prototypes of artificial microswimmers have already been realized, and it is hoped that they can evolve into working tools in biomedicine. They should consist of minimally invasive, small-scale self-propelled devices engineered for drug delivery, diagnostic, or therapeutic purposes.

## Definition of the Subject

Swimming, i. e., being able to advance in a fluid in the absence of external propulsive forces by performing cyclic shape changes, is particularly demanding at low Reynolds numbers (Re). This is the regime of interest for micro-organisms and micro-robots or nano-robots, where hydrodynamics is governed by Stokes equations. Thus, besides the rich mathematics it generates, low Re propulsion is of great interest in biology (How do microorganism swim? Are their strokes optimal and, if so, in which sense? Have these optimal swimming strategies been selected by evolutionary pressure?) and biomedicine (can small-scale self-propelled devices be engineered for drug delivery, diagnostic, or therapeutic purposes?).

For a microscopic swimmer, moving and changing shape at realistically low speeds, the effects of inertia are negligible. This is true for both the inertia of the fluid and the inertia of the swimmer. As pointed out by Taylor [10], this implies that the swimming strategies employed by bacteria and unicellular organism must be radically different from those adopted by macroscopic swimmers such as fish or humans. As a consequence, the design of artificial microswimmers can draw little inspiration from intuition based on our own daily experience.

Taylor's observation has deep implications. Based on a profound understanding of low Re hydrodynamics, and on a plausibility argument on which actuation mechanisms are physically realizable at small length scales, Berg postulated the existence of a sub-micron scale rotary motor propelling bacteria [5]. This was later confirmed by experiment.

## Introduction

In his seminal paper *Life at low Reynolds numbers* [8], Purcell uses a very effective example to illustrate the subtleties involved in microswimming, as compared to the swimming strategies observable in our mundane experience. He argues that at low Re, any organism trying to swim adopting the reciprocal stroke of a scallop, which moves by opening and closing its valves, is condemned to the frustrating experience of not having advanced at all at the end of one cycle.

This observation, which became known as the *scallop theorem*, started a stream of research aiming at finding the simplest mechanism by which cyclic shape changes may lead to effective self-propulsion at small length scales. Purcell's proposal was made of a chain of three rigid links moving in a plane; two adjacent links swivel around joints and are free to change the angle between them. Thus, shape is described by two scalar parameters (the angles between adjacent links), and one can show that, by changing them independently, it is possible to swim.

It turns out that the mechanics of swimming of Purcell's three-link creature are quite subtle, and a detailed understanding has started to emerge only recently [4,9]. In particular, the direction of the average motion of the center of mass depends on the geometry of both the swimmer and of the stroke, and it is hard to predict by simple inspection of the shape of the swimmer and of the sequence of movements composing the swimming stroke. A radical simplification is obtained by looking at axisymmetric swimmers which, when advancing, will do so by moving along the axis of symmetry. Two such examples are the three-sphere-swimmer in [7], and the push-me-pull-you in [3]. In fact, in the axisymmetric case, a simple and complete mathematical picture of low Re swimming is now available, see [1,2].

### The Mathematics of Swimming

This article focuses, for simplicity, on swimmers having an axisymmetric shape  $\Omega$  and swimming along the axis of symmetry, with unit vector  $\vec{l}$ . The configuration, or state  $s$  of the system is described by  $N + 1$  scalar parameters:  $s = \{x^{(1)}, \dots, x^{(N+1)}\}$ . Alternatively,  $s$  can be specified by a position  $c$  (the coordinate of the center of mass along the symmetry axis) and by  $N$  shape parameters  $\xi = \{\xi^{(1)}, \dots, \xi^{(N)}\}$ . Since this change of coordinates is invertible, the generalized velocities  $u^{(i)} = \dot{x}^{(i)}$  can be represented as linear functions of the time derivatives of position and shape:

$$(u^{(1)}, \dots, u^{(N+1)})^t = A(\xi^{(1)}, \dots, \xi^{(N)})(\dot{\xi}^{(1)}, \dots, \dot{\xi}^{(N)}, \dot{c})^t \quad (1)$$

where the entries of the  $N + 1 \times N + 1$  matrix  $A$  are independent of  $c$  by translational invariance.

Swimming describes the ability to change position in the absence of external propulsive forces by executing a cyclic shape change. Since inertia is being neglected, the total drag force exerted by the fluid on the swimmer must also vanish. Thus, since all the components of the total force in directions perpendicular to  $\vec{l}$  vanish by symmetry,

self-propulsion is expressed by

$$0 = \int_{\partial\Omega} \sigma n \cdot \vec{l} \quad (2)$$

where  $\sigma$  is the stress in the fluid surrounding  $\Omega$ , and  $n$  is the outward unit normal to  $\partial\Omega$ . The stress  $\sigma = \eta(\nabla v + (\nabla v)^t) - p\text{Id}$  is obtained by solving Stokes equation outside  $\Omega$  with prescribed boundary data  $v = \bar{v}$  on  $\partial\Omega$ . In turn,  $\bar{v}$  is the velocity of the points on the boundary  $\partial\Omega$  of the swimmer, which moves according to (1).

By linearity of Stokes equations, (2) can be written as

$$\begin{aligned} 0 &= \sum_{i=1}^{N+1} \varphi^{(i)}(\xi^{(1)}, \dots, \xi^{(N)}) u^{(i)} \\ &= A^t \Phi \cdot (\dot{\xi}^{(1)}, \dots, \dot{\xi}^{(N)}, \dot{c})^t \end{aligned} \quad (3)$$

where  $\Phi = (\varphi^{(1)}, \dots, \varphi^{(N)})^t$ , and we have used (1). Notice that the coefficients  $\varphi^{(i)}$  relating drag force to velocities are independent of  $c$  because of translational invariance. The coefficient of  $\dot{c}$  in (3) represents the drag force corresponding to a rigid translation along the symmetry axis at unit speed, and it never vanishes. Thus (3) can be solved for  $\dot{c}$ , and we obtain

$$\dot{c} = \sum_{i=1}^N V_i(\xi^{(1)}, \dots, \xi^{(N)}) \dot{\xi}^{(i)} = V(\xi) \cdot \dot{\xi}. \quad (4)$$

Equation (4) links positional changes to shape changes through shape-dependent coefficients. These coefficients encode all hydrodynamic interactions between  $\Omega$  and the surrounding fluid due to shape changes with rates  $\dot{\xi}^{(1)}, \dots, \dot{\xi}^{(N)}$ .

A stroke is a closed path  $\gamma$  in the space  $S$  of admissible shapes given by  $[0, T] \ni t \mapsto (\xi^{(1)}, \dots, \xi^{(N-1)})$ . Swimming requires that

$$0 \neq \Delta c = \int_0^T \sum_{i=1}^N V_i \dot{\xi}^{(i)} dt \quad (5)$$

i. e., that the differential form  $\sum_{i=1}^N V_i d\xi^{(i)}$  is not exact.

### The Scallop Theorem Proved

Consider a swimmer whose motion is described by a parametrized curve in two dimensions ( $N = 1$ ), so that (4) becomes

$$\dot{c}(t) = V(\xi(t)) \dot{\xi}(t), \quad t \in \mathbb{R}, \quad (6)$$

and assume that  $V \in L^1(S)$  is an integrable function in the space of admissible shapes and  $\xi \in W^{1,\infty}(\mathbb{R}; S)$  is a Lipschitz-continuous and  $T$ -periodic function for some  $T > 0$ , with values in  $S$ .



Figure 1 is a sketch representing concrete examples compatible with these hypotheses. The axisymmetric case consists of a three-dimensional cone with axis along  $\vec{i}$  and opening angle  $\xi \in [0, 2\pi]$  (an *axisymmetric octopus*). A non-axisymmetric example is also allowed in this discussion, consisting of two rigid parts (*valves*), always maintaining mirror symmetry with respect to a plane (containing  $\vec{i}$  and perpendicular to it) while swiveling around a joint contained in the symmetry plane and perpendicular to  $\vec{i}$  (a *mirror-symmetric scallop*), and swimming parallel to  $\vec{i}$ .

Among the systems that are *not* compatible with the assumptions above are those containing helical elements with axis of rotation  $\vec{i}$ , and capable of rotating around  $\vec{i}$  always in the same direction (call  $\theta$  the rotation angle). Indeed, a monotone function  $t \mapsto \theta(t)$  is not periodic.

The celebrated “scallop theorem” [8] states that, for a system like the one depicted in Fig. 1, the net displacement of the center of mass at the end of a periodic stroke will always vanish. This is due to the linearity of Stokes equation (which leads to symmetry under time reversals), and to the low dimensionality of the system (a one-dimensional periodic stroke is necessarily reciprocal). Thus, whatever forward motion is achieved by the scallop by closing its valves, it will be exactly compensated by a backward motion upon reopening them. Since the low Re world is unaware of inertia, it will not help to close the valves quickly and reopen them slowly. A precise statement and a rigorous short proof of the scallop theorem are given below.

**Theorem 1** Consider a swimmer whose motion is described by

$$\dot{c}(t) = V(\xi(t))\dot{\xi}(t), \quad t \in \mathbb{R}, \quad (7)$$

with  $V \in L^1(S)$ . Then for every  $T$ -periodic stroke  $\xi \in W^{1,\infty}(\mathbb{R}; S)$ , one has

$$\Delta c = \int_0^T \dot{c}(t) dt = 0. \quad (8)$$

*Proof* Define the primitive of  $V$  by

$$\Psi(s) = \int_0^s V(\sigma) d\sigma \quad (9)$$

so that  $\Psi'(\xi) = V(\xi)$ . Then, using (7),

$$\begin{aligned} \Delta c &= \int_0^T V(\xi(t))\dot{\xi}(t) dt \\ &= \int_0^T \frac{d}{dt} \Psi(\xi(t)) dt \\ &= \Psi(\xi(T)) - \Psi(\xi(0)) = 0 \end{aligned}$$

by the  $T$ -periodicity of  $t \mapsto \xi(t)$ .

### Optimal Swimming

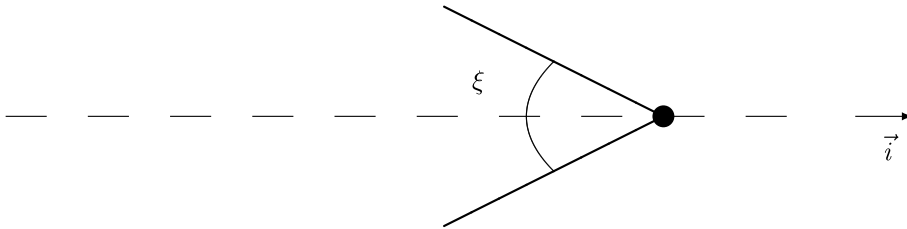
A classical notion of swimming efficiency is due to Lighthill [6]. It is defined as the inverse of the ratio between the average power expended by the swimmer during a stroke starting and ending at the shape  $\xi_0 = (\xi_0^{(1)}, \dots, \xi_0^{(N)})$  and the power that an external force would spend to translate the system rigidly at the same average speed  $\bar{c} = \Delta c/T$ :

$$\text{Eff}^{-1} = \frac{\frac{1}{T} \int_0^T \int_{\partial\Omega} \sigma n \cdot v}{6\pi\eta L \bar{c}^2} = \frac{\int_0^1 \int_{\partial\Omega} \sigma n \cdot v}{6\pi\eta L (\Delta c)^2} \quad (10)$$

where  $\eta$  is the viscosity of the fluid,  $L = L(\xi_0)$  is the effective radius of the swimmer, and time has been rescaled to a unit interval to obtain the second identity. The expression in the denominator in (10) comes from a generalized version of Stokes formula giving the drag on a sphere of radius  $L$  moving at velocity  $\bar{c}$  as  $6\pi\eta L \bar{c}$ .

Let  $DN: H^{1/2}(\partial\Omega) \rightarrow H^{-1/2}(\partial\Omega)$  be the Dirichlet to Neumann map of the outer Stokes problem, i. e., the map such that  $\sigma n = DNv$ , where  $\sigma$  is the stress in the fluid, evaluated on  $\partial\Omega$ , arising in response to the prescribed velocity  $v$  on  $\partial\Omega$ , and obtained by solving the Stokes problem outside  $\Omega$ . The expended power in (10) can be written as

$$\int_{\partial\Omega} \sigma n \cdot v = \int_{\partial\Omega} DN(v) \cdot v. \quad (11)$$



Biological Fluid Dynamics, Non-linear Partial Differential Equations, Figure 1  
A mirror-symmetric scallop or an axisymmetric octopus

At a point  $p \in \partial\Omega$ , the velocity  $v(p)$  accompanying a change of state of the swimmer can be written as a linear combination of the  $u^{(i)}$

$$v(p) = \sum_{i=1}^{N+1} \mathcal{V}_i(p, \xi) u^{(i)} \quad (12)$$

$$= \sum_{i=1}^N \mathcal{W}_i(p, \xi) \dot{\xi}^{(i)}. \quad (13)$$

Indeed, the functions  $\mathcal{V}_i$  are independent of  $c$  by translational invariance, and (4) has been used to get (13) from the line above.

Substituting (13) in (11), the expended power becomes a quadratic form in  $\dot{\xi}$

$$\int_{\partial\Omega} \sigma n \cdot v = (G(\xi) \dot{\xi}, \dot{\xi}) \quad (14)$$

where the symmetric and positive definite matrix  $G(\xi)$  is given by

$$G_{ij}(\xi) = \int_{\partial\Omega} DN(\mathcal{W}_i(p, \xi)) \cdot \mathcal{W}_j(p, \xi) dp. \quad (15)$$

Strokes of maximal efficiency may be defined as those producing a given displacement  $\Delta c$  of the center of mass with minimal expended power. Thus, from (10), maximal efficiency is obtained by minimizing

$$\int_0^1 \int_{\partial\Omega} \sigma n \cdot v = \int_0^1 (G(\xi) \dot{\xi}, \dot{\xi}) \quad (16)$$

subject to the constraint

$$\Delta c = \int_0^1 V(\xi) \cdot \dot{\xi} \quad (17)$$

among all closed curves  $\xi: [0, 1] \rightarrow S$  in the set  $S$  of admissible shapes such that  $\xi(0) = \xi(1) = \xi_0$ .

The Euler-Lagrange equations for this optimization problem are

$$-\frac{d}{dt}(G\dot{\xi}) + \frac{1}{2} \begin{pmatrix} \left( \frac{\partial G}{\partial \xi^{(1)}} \dot{\xi}, \dot{\xi} \right) \\ \vdots \\ \left( \frac{\partial G}{\partial \xi^{(N)}} \dot{\xi}, \dot{\xi} \right) \end{pmatrix} + \lambda \left( \nabla_{\xi} V - \nabla_{\xi}^t V \right) \dot{\xi} = 0 \quad (18)$$

where  $\nabla_{\xi} V$  is the matrix  $(\nabla_{\xi} V)_{ij} = \partial V_i / \partial \xi_j$ ,  $\nabla_{\xi}^t V$  is its transpose, and  $\lambda$  is the Lagrange multiplier associated with the constraint (17).

Given an initial shape  $\xi_0$  and an initial position  $c_0$ , the solutions of (18) are in fact sub-Riemannian geodesics joining the states parametrized by  $(\xi_0, c_0)$  and  $(\xi_0, c_0 + \Delta c)$  in the space of admissible states  $\mathcal{X}$ , see [1]. It is well known, and easy to prove using (18), that along such geodesics  $(G(\gamma)\dot{\gamma}, \dot{\gamma})$  is constant. This has interesting consequences, because swimming strokes are often divided into a power phase, where  $|G(\gamma)|$  is large, and a recovery phase, where  $|G(\gamma)|$  is smaller. Thus, along optimal strokes, the recovery phase is executed quickly while the power phase is executed slowly.

### The Three-Sphere Swimmer

For the three-sphere-swimmer of Najafi and Golestanian [7], see Fig. 2,  $\Omega$  is the union of three rigid disjoint balls  $B^{(i)}$  of radius  $a$ , shape is described by the distances  $x$  and  $y$ , the space of admissible shapes is  $S = (2a, +\infty)^2$ , and the kinematic relation (1) takes the form

$$\begin{aligned} u^{(1)} &= \dot{c} - \frac{1}{3}(2\dot{x} + \dot{y}) \\ u^{(2)} &= \dot{c} + \frac{1}{3}(\dot{x} - \dot{y}) \\ u^{(3)} &= \dot{c} + \frac{1}{3}(2\dot{y} + \dot{x}). \end{aligned} \quad (19)$$

Consider, for definiteness, a system with  $a = 0.05$  mm, swimming in water. Calling  $f^{(i)}$  the total propulsive force on ball  $B^{(i)}$ , we find that the following relation among forces and ball velocities holds

$$\begin{pmatrix} f^{(1)} \\ f^{(2)} \\ f^{(3)} \end{pmatrix} = R(x, y) \begin{pmatrix} u^{(1)} \\ u^{(2)} \\ u^{(3)} \end{pmatrix} \quad (20)$$

where the symmetric and positive definite matrix  $R$  is known as the resistance matrix. From this last equation, using also (19), the condition for self-propulsion  $f^{(1)} + f^{(2)} + f^{(3)} = 0$  is equivalent to

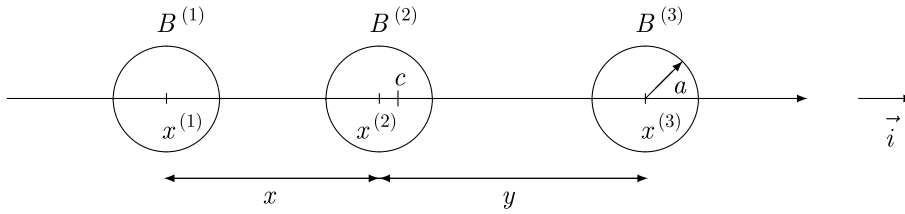
$$\dot{c} = V_x(x, y)\dot{x} + V_y(x, y)\dot{y}, \quad (21)$$

where

$$V_x(x, y) = \frac{Re_c \cdot (e_c \times e_y)}{Re_c \cdot (e_x \times e_y)} \quad (22)$$

$$V_y(x, y) = -\frac{Re_c \cdot (e_c \times e_x)}{Re_c \cdot (e_x \times e_y)}. \quad (23)$$

Moreover,  $e_x = (-1, 1, 0)^t$ ,  $e_y = (0, -1, 1)^t$ ,  $e_c = (1/3, 1/3, 1/3)^t$ .



**Biological Fluid Dynamics, Non-linear Partial Differential Equations, Figure 2**  
Swimmer's geometry and notation

**Biological Fluid Dynamics, Non-linear Partial Differential Equations, Table 1**

Energy consumption ( $10^{-12}$  J) for the three strokes of Fig. 3 inducing the same displacement  $\Delta c = 0.01$  mm in  $T = 1$  s

Optimal stroke	Small square stroke	Large square stroke
0.229	0.278	0.914

Given a stroke  $\gamma = \partial\omega$  in the space of admissible shapes, condition (5) for swimming reads

$$0 \neq \Delta c = \int_0^T (V_x \dot{x} + V_y \dot{y}) dt = \int_\omega \text{curl} V(x, y) dx dy \quad (24)$$

which is guaranteed, in particular, if  $\text{curl} V$  is bounded away from zero. Strokes of maximal efficiency for a given initial shape  $(x_0, y_0)$  and given displacement  $\Delta c$  are obtained by solving Eq. (18). For  $N = 2$ , this becomes

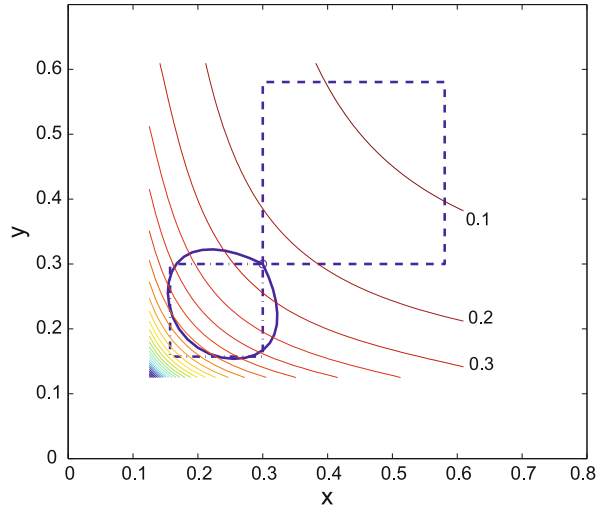
$$-\frac{d}{dt}(G\dot{\gamma}) + \frac{1}{2} \left( \left( \frac{\partial_x G \dot{\gamma}, \dot{\gamma}}{\partial_y G \dot{\gamma}, \dot{\gamma}} \right) \right) + \lambda \text{curl} V(\gamma) \dot{\gamma}^\perp = 0 \quad (25)$$

where  $\partial_x G$  and  $\partial_y G$  stand for the  $x$  and  $y$  derivatives of the  $2 \times 2$  matrix  $G(x, y)$ .

It is important to observe that, for the three-sphere swimmer, all hydrodynamic interactions are encoded in the shape dependent functions  $V(x, y)$  and  $G(x, y)$ . These can be found by solving a two-parameter family of outer Stokes problems, where the parameters are the distances  $x$  and  $y$  between the three spheres. In [1], this has been done numerically via the finite element method: a representative example of an optimal stroke, compared to two more naive proposals, is shown in Fig. 3.

## Future Directions

The techniques discussed in this article provide a head start for the mathematical modeling of microscopic swimmers, and for the quantitative optimization of their strokes. A complete theory for axisymmetric swimmers is already available, see [2], and further generalizations to



**Biological Fluid Dynamics, Non-linear Partial Differential Equations, Figure 3**

Optimal stroke and square strokes which induce the same displacement  $\Delta c = 0.01$  mm in  $T = 1$  s, and equally spaced level curves of  $\text{curl} V$ . The small circle locates the initial shape  $\xi_0 = (0.3 \text{ mm}, 0.3 \text{ mm})$

arbitrary shapes are relatively straightforward. The combination of numerical simulations with the use of tools from sub-Riemannian geometry proposed here may prove extremely valuable for both the question of adjusting the stroke to global optimality criteria, and of optimizing the stroke of complex swimmers. Useful inspiration can come from the sizable literature on the related field dealing with control of swimmers in a perfect fluid.

## Bibliography

### Primary Literature

1. Alouges F, DeSimone A, Lefebvre A (2008) Optimal strokes for low Reynolds number swimmers: an example. *J Nonlinear Sci* 18:277–302
2. Alouges F, DeSimone A, Lefebvre A (2008) Optimal strokes for low Reynolds number axisymmetric swimmers. Preprint SISSA 61/2008/M
3. Avron JE, Kenneth O, Oakmin DH (2005) Pushmepullyou: an efficient micro-swimmer. *New J Phys* 7:234–1–8

4. Becker LE, Koehler SA, Stone HA (2003) On self-propulsion of micro-machines at low Reynolds numbers: Purcell's three-link swimmer. *J Fluid Mechanics* 490:15–35
5. Berg HC, Anderson R (1973) Bacteria swim by rotating their flagellar filaments. *Nature* 245:380–382
6. Lighthill MJ (1952) On the Squirming Motion of Nearly Spherical Deformable Bodies through Liquids at Very Small Reynolds Numbers. *Comm Pure Appl Math* 5:109–118
7. Najafi A, Golestanian R (2004) Simple swimmer at low Reynolds numbers: Three linked spheres. *Phys Rev E* 69:062901-1–4
8. Purcell EM (1977) Life at low Reynolds numbers. *Am J Phys* 45:3–11
9. Tan D, Hosoi AE (2007) Optimal stroke patterns for Purcell's three-link swimmer. *Phys Rev Lett* 98:068105-1–4
10. Taylor GI (1951) Analysis of the swimming of microscopic organisms. *Proc Roy Soc Lond A* 209:447–461

### Books and Reviews

- Agrachev A, Sachkov Y (2004) Control Theory from the Geometric Viewpoint. In: *Encyclopaedia of Mathematical Sciences*, vol 87, Control Theory and Optimization. Springer, Berlin
- Childress S (1981) *Mechanics of swimming and flying*. Cambridge University Press, Cambridge
- Happel J, Brenner H (1983) *Low Reynolds number hydrodynamics*. Nijhoff, The Hague
- Kanso E, Marsden JE, Rowley CW, Melli-Huber JB (2005) Locomotion of Articulated Bodies in a Perfect Fluid. *J Nonlinear Sci* 15: 255–289
- Koiller J, Ehlers K, Montgomery R (1996) Problems and Progress in Microswimming. *J Nonlinear Sci* 6:507–541
- Montgomery R (2002) *A Tour of Subriemannian Geometries, Their Geodesics and Applications*. AMS Mathematical Surveys and Monographs, vol 91. American Mathematical Society, Providence

## Biological Models of Molecular Network Dynamics

HERBERT M. SAURO  
Department of Bioengineering,  
University of Washington, Seattle, USA

### Article Outline

[Glossary](#)  
[Definition of the Subject](#)  
[Introduction](#)  
[Modeling Approaches](#)  
[Deterministic Modeling](#)  
[Stoichiometry Matrix](#)  
[System Equation](#)  
[Theoretical Approaches to Modeling](#)  
[Negative Feedback Systems](#)  
[FeedForward Systems](#)  
[Positive Feedback](#)

[Future Prospects](#)  
[Bibliography](#)

### Glossary

**Deterministic continuous model** A mathematical model where the variables of the model can take any real value and where the time evolution of the model is set by the initial conditions.

**Stochastic discrete model** A mathematical model where the variables of the model take on discrete values and where the time evolution of the model is described by a set of probability distributions.

### Definition of the Subject

Understanding the operation cellular networks is probably one of the most challenging and intellectually exciting scientific fields today. With the availability of new experimental and theoretical techniques our understanding of the operation of cellular networks has made great strides in the last few decades. An important outcome of this work is the development of predictive quantitative models. Such models of cellular function will have a profound impact on our ability of manipulate living systems which will lead to new opportunities for generating energy, mitigating our impact on the biosphere and last but not least, opening up new approaches and understanding of important disease states such as cancer and aging.

### Introduction

Cellular networks are some of the most complex natural systems we know. Even in a “simple” organism such as *E. coli*, there are at least four thousand genes with many thousands of interactions between molecules of many different sizes [11]. In a human cell the number of interactions is probably orders of magnitude larger. Why all this complexity? Presumably the earliest living organisms were much simpler than what we find today but competition for resources and the need to adapt in unfavorable conditions must have led to the development of sensory and decision-making capabilities above and beyond the basic requirements for life. What we see today in almost all living organisms are complex signaling and genetic networks whose complexity and subtlety is beyond most man-made technological systems [87].

Over the last sixty or so years, biochemists and molecular biologists have identified many of the components in living cells and have traced out many of the interactions that delineate cellular networks. What emerges is a picture that would be familiar to many control engineers

where one finds feedback loops, switches, logic gates, oscillators, filters and a myriad of other control and decision-making devices. This knowledge allows us to rationalize the complexity we find in cellular networks by functionalizing a large network into smaller well defined modules. This work is also spawning a new engineering field called synthetic biology which uses knowledge of modular organization and function to engineer new biological networks [3].

One of the important techniques that has permitted us to gain a deeper understanding of network structure and function is mathematical modeling and new theoretical approaches. These, coupled with considerable advances in measurement technology, has revolutionized our picture of how cellular networks operate. Instead of being relatively static structures as portrayed in many textbooks, cellular networks are, unsurprisingly, highly dynamic systems [42,90].

In this chapter the basic approach to mathematically modeling cellular networks will be introduced followed by a brief review of some of the properties of the commonest network modular designs that are known to exist. It should be noted that computational and theoretical models that help us understand the operational characteristics of different network configurations is now an extremely large field and it would be impossible in one short chapter to cover even a small fraction of the advances that have been made in the last few decades. There is a growing list of excellent reviews and some text books that the interested reader can use to expand on some of the topics covered here, these include excellent works by Alon [1], Arkin [6], Fell [26], Ferrell [29], Goldbetter [32], Heinrich [36], Hofmeyr [38], [40], Kacser [44], Kholodenko [46], Klipp [47], Lauffenburger [50], Sauro [67], Savageau [72], [35] and Tyson [79,80] to name but a few.

Much like human engineered systems, the organization of cellular networks is extremely hierarchical, both in time and space. At the most basic level we have a remarkably simple picture of elemental reactions such as unimolecular and bimolecular reactions. All cellular behavior is built on these elemental reactions. On their own, uni and bimolecular reactions show very simple behavior, it is only when combined into more complex structures that more interesting behavior emerges. For example, the simplest systems we can construct from elementary reactions includes catalysis by enzymes where we already see the emergence of more complex responses in the form of hyperbolic and sigmoid behavior. These more complex responses originate through mass conservation and nonlinearities in the underlying bimolecular reactions. The hierarchical range of organization, however, goes much fur-

ther, ranging all the way from simple enzymatic kinetics to complex signal processing encompassing hundreds of components [56]. Any understanding of network dynamics has to appreciate the different levels of organization and the characteristic dynamic properties that they bring to bear. Cellular networks offer an extremely rich source of signal processing capabilities, much of which is largely unexplored. With the rise of systems biology, this picture is changing.

## Modeling Approaches

There are two primary methods for modeling cellular networks, one uses a continuous deterministic approach mainly centered on using ordinary differential equations [88] and the other is based on a discrete stochastic approach, mainly centered on using the chemical master equation [85]. Each approach has its benefits and drawbacks. For systems with large numbers of molecules ( $> 100$ ), the use of differential equations is sufficient and recommended. This approach assumes that concentrations change continuously in time and space and that stochastic events do not occur.

The use of differential equations in modeling physical phenomena has a long history dating back to at least the time of Newton. As a result, a very large body of knowledge, including analytical as well as computational approaches are available to the modeling community. In consequence, differential equation models in systems biology are very popular with many software tools and analytical techniques available for their study [10]. Unfortunately, some cellular systems, particularly bacterial genetic networks, operate at very low molecule numbers. This makes the use of differential equations models for these systems problematic. At very low molecule numbers the stochastic nature of chemical kinetics comes into play and it is necessary to use a stochastic modeling approach [31]. In some cases, stochastic behavior can be markedly different from that shown in an equivalent deterministic model [60,65]. Unfortunately at this point in time, simulation tools and particular analytical tools are not as well developed which impedes progress in understanding the dynamics at the stochastic level although progress has been more rapid in recent years. In this chapter the focus will be on modeling by differential equations even though some of the interesting complexity we see in biological networks emerges as a result of stochastic behavior.

In later sections, models of genetic networks will be described using a deterministic approach. One could argue that in these cases a stochastic approach is crucial because the number of genes and particular promoter sites

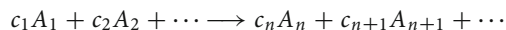


is extremely low ( $< 4$  per gene). The simulations of these models could have been carried out using a stochastic approach, however, the general results obtained would have been the same (the analytical analysis of the stochastic equivalent models could not have been repeated). The question then is when should a stochastic approach be used? There are at least three clear cases when a stochastic approach is critical to modeling biological networks because of differences in predicted behavior: 1) When single events can make a qualitative difference such as shown in the classic work by Arkin et al. on the developmental pathway of phage lambda when infecting *E. coli* [6] or when modeling free radical chemistry [54]; 2) Nonlinearities in the rate laws which can lead to stochastic focusing and thus changes in the mean reaction rates. Such focusing in turn can cause shifts in the steady state solutions or even behavior [60,65]; or 3) When noise can be measured experimentally and an appropriate stochastic model is required to explain the measurements. In many instances both stochastic and deterministic models will yield the same solutions (at least when averaged) but the availability of a mature theory of deterministic dynamics makes deterministic models in such situations much more appealing. When building models, one should always consider the purpose of the model. A model is always a simplification, the reduction of a complicated and messy reality to its essential features. A model is therefore never intended to represent reality in all its detail and efforts for example to build replica *in silico cells* miss the point of modeling. If noise is not of interest or deemed unimportant than a simpler model will suffice and if in doubt one should follow Occam's razor [4].

### Deterministic Modeling

Chemical kinetics is at the heart of all dynamic behavior in biological cells. Almost all chemical transformations are mediated by either enzyme or protein factors. Any understanding of network dynamics must therefore start with a foundation in chemical kinetics. To begin with, one must define what is meant by *the rate of reaction* [55].

Consider a general reaction of the form:



where  $c_i$  represents the stoichiometric mole numbers and  $A_i$  a chemical species. The notation indicates that species  $A_1, A_2$  etc. are transformed with mole numbers  $c_1$  etc. to product species  $A_n$  etc. with mole numbers,  $c_n$  etc. Normally when expressing a reaction in this form signs are not explicitly given to the stoichiometries, however it is

implied that the stoichiometric mole numbers  $c_i$  are **positive** for products and **negative** for reactants.

Let us introduce the notion of the *extent of reaction*, indicated by the symbol  $\xi$ . We define a change from  $\xi$  to  $\xi + d\xi$  in time  $dt$  to mean that  $c_1 d\xi$  moles of  $A_1$ ,  $c_2 d\xi$  moles of  $A_2$  etc., react to form  $c_n d\xi$  moles of  $A_n$  etc. By this definition we can state that for any component  $i$  the following is true in the time interval  $dt$ :

$$\frac{dn_i}{dt} = c_i \frac{d\xi}{dt}$$

where  $n_i$  equals the amount in moles of species  $i$ . From this relation we *define* the rate of reaction,  $v_r$ , to be:

$$v_r \equiv \frac{d\xi}{dt}.$$

In other words

$$\frac{dn_i}{dt} = c_i v_r.$$

The advantage of employing the extent of reaction is that it allows us to define the rate of reaction independently of the species we use to measure the rate. This convenient property can be expressed as:

$$\begin{aligned} v_r \equiv \frac{d\xi}{dt} &= -\frac{1}{c_1} \frac{dn_1}{dt} = -\frac{1}{c_2} \frac{dn_2}{dt} = \cdots \\ &= \frac{1}{c_n} \frac{dn_n}{dt} = \frac{1}{c_{n+1}} \frac{dn_{n+1}}{dt} = \cdots \end{aligned}$$

Often models are expressed in unit volume so that more commonly modelers will use the intensive form for the reaction rate:

$$\frac{dC_i}{dt} = c_i v_j \quad (1)$$

where  $C_i$  is the concentration of species  $i$  and  $v_j$  is the intensive rate of reaction. When building network models, this formalism allows us to separate the connectivity or structure of the network from the rates of reactions.

### Mass-Action Kinetics

One of the basic concepts in chemical kinetics is the Law of Mass Action [7]. This states that the rate of a chemical reaction is proportional to the product of the reactant concentrations raised to a given power. For a simple monomolecular reaction such as:



The rate of reaction,  $v$ , is given by:

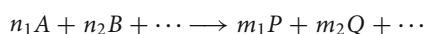
$$v = k_1 A$$

where  $A$  represents the concentration of reactant.  $v$  is the intensive rate of reaction. The rate of change of  $A$  is given by (note that the stoichiometry for  $A = -1$ , see Eq. (1)):

$$\frac{dA}{dt} = -k_1 A. \quad (2)$$

The units for  $k_1$  are  $t^{-1}$  and for the concentration of  $A$ , moles  $l^{-1}$ . The rate of change of  $A$  therefore has units of moles  $l^{-1} t^{-1}$ . By convention, a rate that is proportional to a reactant raised to a first power is called first-order with respect to the reactant. Similarly, reactants raised to the zeroth power are called zero-order and reactants raised to the power of two are called second-order.

For the general reaction such as:



the rate law has been found through empirical observation to often have the form:

$$v = k_1 A^{n_1} B^{n_2} \dots$$

where each reactant is raised to the power of its stoichiometry. For example the reaction catalyzed by adenylate kinase that equilibrates ATP, ADP and AMP:



could have the rate law:

$$v = kADP^2.$$

If the reaction is reversible then the rate law is appended with the reverse rate:

$$v = k_1 A^{n_1} B^{n_2} \dots - k_2 P^{m_1} Q^{m_2} \dots$$

$$v = k_1 ADP^2 - k_2 ATP AMP.$$

The units for the reactant and product terms must be in concentration. The units for the rate constants,  $k$  will depend on the exact form of the rate law but must be set to ensure that the rate of reaction is expressed in the units moles  $l^{-1} t^{-1}$  if the intensive form is used.

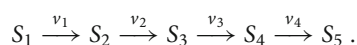
Mass-action kinetics forms the most fundamental and simplest level of description in the cellular hierarchy. Beyond simple mass-action kinetics there is a wide range of aggregate rate laws, probably the most celebrated being the Michaelis–Menten enzyme rate law [20]. These rate

laws make assumptions about the time scale of some reaction processes, such as the binding of substrate to enzyme, so that simplifications can be made that allow one to derive simpler kinetic descriptions. Aggregate rate laws display more complex behavior than simple mass-action kinetics, the most important of these are hyperbolic (or saturable) and sigmoid kinetics [75]. These rate laws can be used across the range of cellular processes, including gene expression, protein signaling and metabolic pathways. A variety of generalized rate laws also exist that make the choice of which rate law to use much easier [35].

## Stoichiometry Matrix

When describing multiple reactions in a network, it is convenient to represent the stoichiometries in a compact form called the stoichiometry matrix,  $\mathbf{N}$  [5,19,25,41]. This matrix is a  $m$  row by  $n$  column matrix where  $m$  is the number species and  $n$  the number reactions. The columns of the stoichiometry matrix correspond to the distinct chemical reactions in the network, the rows to the molecular species, one row per species. Thus the intersection of a row and column in the matrix indicates whether a certain species takes part in a particular reaction or not and, according to the sign of the element, whether there is a net loss or gain of substance, and by the magnitude, the relative quantity of substance that takes part in that reaction. The elements of the stoichiometry matrix thus concerns the relative mole amounts of chemical species that react in a particular reaction; it does not concern itself with the rate of reaction. The use of  $\mathbf{N}$  has a long tradition and is the preferred notation to use [61].

For example, consider the simple chain of reactions which has five molecular species and four reactions. The four reactions are labeled,  $v_1$  to  $v_4$ .

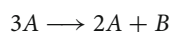


The stoichiometry matrix for this simple system is given by:

$$\mathbf{N} = \begin{bmatrix} & v_1 & v_2 & v_3 & v_4 \\ S_1 & -1 & 0 & 0 & 0 \\ S_2 & 1 & -1 & 0 & 0 \\ S_3 & 0 & 1 & -1 & 0 \\ S_4 & 0 & 0 & 1 & -1 \\ S_5 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

Entries in the stoichiometry matrix are computed as follows. Given a species  $S_i$  and reaction  $v_j$ , the corresponding entry in the stoichiometry matrix at  $ij$  is given by the total stoichiometry for  $S_i$  on the product side minus the total

stoichiometry of  $S_i$  on the reactant side. Thus, considering species  $S_1$  in reaction  $v_1$ , we note that the total stoichiometry on the product side is zero (no  $S_1$  molecules are formed on the product side) and the total stoichiometry of  $S_1$  on the reactant side is +1. Subtracting one from the other ( $0 - 1$ ) we obtain  $-1$ , which is entered into the stoichiometry matrix. This rather long-winded approach to computing stoichiometries avoids errors that arise when a species occurs as both reactant and product. For example, for the more complex reaction such as:



the stoichiometry entry for  $A$  is  $(2 - 3)$  that is  $-1$ , since the stoichiometry for  $A$  on the product side is two and on the reactant side is three.

The stoichiometry matrix is an important element in modeling cellular networks. It represents the connectivity of the network without reference to kinetics or regulation. Even without kinetic information, the stoichiometry matrix contains much interesting information and certain modeling approaches, such as flux balance analysis, make extensive use of the stoichiometry matrix to determine flux conservation relationships from which one can estimate fluxes through a given network [28,59,76].

### System Equation

The central equation in deterministic modeling is the system equation (4). The system equation takes into account both the kinetics of the individual reactions as well as the structure of the network. When models of the time evolution of the molecular species by differential equations are built, the stoichiometry matrix is combined with the vector of reaction rates to obtain the *system equation*:

$$\frac{dS}{dt} = Nv(S) \quad (4)$$

where  $N$  is the  $m \times n$  stoichiometry matrix and  $v$  is the  $n$  dimensional rate vector, whose  $i$ th component gives the rate of reaction  $i$  as a function of the species concentrations.

If we look again at the model depicting the simple chain of reactions (Eq. (3)), the system equation can be written as:

$$\frac{dS}{dt} = Nv = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}.$$

Two particular solutions of the system equation are of interest, time course evolution of molecular species concentrations and the steady state. The steady state is defined as the solution(s) to the system equation when the rates of change of species is zero, Eq. (5)

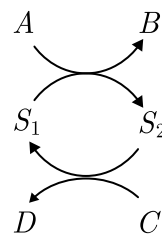
$$Nv(S) = 0. \quad (5)$$

### Redundancies

The system equation (4) will often contain redundant equations. These redundant equations arise from conserved moieties in the network. Conserved moieties are molecular subgroups conserved during the evolution of a network [62]. The total amount of a particular moiety in a network is time invariant and is determined solely by the initial conditions imposed on the system. A typical example of a conserved moiety in a computational model is the conservation of adenine nucleotide, that is when the total amount of ATP, ADP and AMP is constant during the evolution of the system. Other examples include NAD/NADH, phosphorylation/dephosphorylation cycles, and so on. It should be noted that conserved cycles are a model time-scale abstraction because in reality all cycles turn over. However, over the relatively short-time scales used in modeling studies conserved cycles have important biological effects.

Figure 1 illustrates the simplest possible network which displays a conserved moiety, in this case the total mass,  $S_1 + S_2$  is constant during the evolution of the network. Such conservation laws can have dramatic effects on the dynamical properties of models. For example, the hyperbolic behavior in the simple Michaelis–Menten rate law can be attributed to the conservation law between free enzyme and enzyme-substrate complex but more subtle effects, such as the emergence of new regulatory properties [53], can be attributed to moiety conserved cycles.

The effect of conserved moieties on the system equation 4 is to reduce the number of differential equations by the number of conservation laws in the model. In a model-



Biological Models of Molecular Network Dynamics, Figure 1  
Simple Conserved cycle

ing environment, it is therefore customary to divide the set of molecular species into two groups, the dependent ( $S_d$ ) and independent set ( $S_i$ ) [37]. This division is a function of the number and kind of conserved cycles in the network. If the model has no conserved cycles then the dependent set is empty and the size of the independent set equals the number of molecular species in the model. For details on how to compute  $S_d$  and  $S_i$  the reader can consult Sauro and Ingalls [66] and Vallabhajosyula [82]. In many cases it is critical that this separation is made. For simple time course simulations the separation is not so important, but for most other analyses it is very important and for stiff integration methods highly desirable. The reason is that many numerical methods, including the stiff integrators, use the Jacobian matrix [77] as part of the numerical calculation. If the separation is not carried out, the Jacobian becomes singular which renders many analyses numerically unstable if not impossible to carry out. In addition, the separation is also very useful because it enables the number of differential equations to be reduced in number and thereby improve computational efficiency.

Many simulators and mathematical analyses do not actually solve the system equation (4) directly but instead solve the following modified system:

$$\begin{aligned} S_d &= L_0 S_i + T \\ \frac{dS_i}{dt} &= N_R v(S_i, S_d) . \end{aligned} \quad (6)$$

In these equations,  $S_i$  is the vector of independent species,  $S_d$ , the vector of dependent species,  $L_0$  the link matrix,  $T$  the total mass vector,  $N_R$  the reduced stoichiometry matrix and  $v$  the rate vector. This modified equation constitutes the most general expression for a differential equation-based model [35,39,61]. The symbolism used in equation 6 is the standard notation used by many in the Systems Biology community.

Finally, it should be noted that conservation analysis is not confined to deterministic models but is equally applicable to stochastic models and could be used to improve the computational efficiencies of stochastic model solvers.

## Theoretical Approaches to Modeling

One way to study the properties of molecular networks is to build computer simulation models. Simulation is often the only recourse when the system equations that govern the dynamics are nonlinear and no analytical solution is possible. However, certain theoretical approaches allow models to be studied analytically if we confine ourselves to small perturbations and their effects. Such types of analyses are very common in the physical sci-

ences and engineering [14]. In the late 1960s there developed a small perturbation approach for molecular networks, called Biochemical Systems Theory, metabolic control analysis or more simply, control analysis. This approach was developed independently by three different groups [26,34,36,44,69,72] and has since been expanded to deal with arbitrary network structures, time dependent changes and more recently it has been extended into the frequency domain which showed that the approach was identical to classical control analysis used by engineers. A brief description of the theory will be given here.

Control analysis is a mathematical technique that allows us to describe some property of a molecular network, such as a steady state concentration or flux, in terms of the properties of the individual parts. In control analysis, system properties are described using control coefficients, of particular interest are the concentration and flux control coefficients. These are defined as the fractional change in concentration or flux relative to the fractional change in some system parameter, usually an enzyme or gene activity:

$$\begin{aligned} C_E^J &= \frac{dJ}{dE} \frac{E}{J} \\ C_E^S &= \frac{dS}{dE} \frac{E}{S} . \end{aligned}$$

High control coefficients indicate that the system parameter ( $E$ ) has a significant effect on the concentration or flux. To make the analysis simpler, control coefficients are also usually measured at steady state. Therefore the control coefficient,  $C_E^S$ , measures the effect of  $E$  on the steady state concentration of  $S$ . Concentration control coefficients are of particular interest in understanding the dynamics of gene and protein regulatory networks where the notion of a flux is not so important. Flux control coefficients on the other hand are of particular interest when studying the dynamics of metabolic pathways where the flux might represent the consumption of an important nutrient or the export of some valuable commodity such as a drug [63].

Control analysis also defines a property of a network part, the elasticity, a local property. Elasticities describe how a reaction rate is affected by changes in the immediate environment of the enzyme or gene that is catalyzing the reaction. If there are multiple factors that can affect a process then there will be multiple elasticities, one for each factor. Often these factors are simply the concentrations of substrates and products, but they could also be regulators, pH, ion concentrations, etc. Elasticities are defined as follows:

$$\varepsilon_S^v = \frac{\partial v}{\partial S} \frac{S}{v} .$$

Note that the elasticities are defined using a partial derivative whereas the control coefficients are defined using total derivatives, this reflects the local and system characteristics of each of these coefficients.

The real power of control analysis is its ability to relate these two coefficients together. There are various ways to accomplish this but here the approach using the control theorems will be described [27]. The theorems of control analysis describe various relationships among the coefficients, two in particular are significant for this discussion, namely the summation and connectivity theorems.

Two summation theorems of interest include the concentration and flux summation theorems, they can be stated as follows:

$$\sum_{i=1}^n C_{E_i}^J = 1$$

$$\sum_{i=1}^n C_{E_i}^S = 0$$

where the summation is over all steps in the pathway. Operationally the concentration summation theorem states that increasing the activity of all enzymes and/or genes in a network will result in no change to the metabolite or protein concentrations. The flux summation makes a similar statement except that it indicates that the steady state flux increases in proportion to the activity changes in the enzymes and/or genes. The theorems, and particularly generalizations of them [35], have important implications for metabolic engineers. The generalized summation theorems can be used to indicate which steps in a pathway should be changed in order to increase or decrease the output of a particular pathway commodity [21,73,74].

The connectivity theorem is more subtle and relates control coefficients to elasticities:

$$\sum C_{E_i}^J \varepsilon_S^{v_i} = -1$$

$$\sum C_{E_i}^S \varepsilon_S^{v_i} = 0.$$

The sum is over all reactions and/or genes that the molecular species,  $S$  can influence [15,44]. From these theorems (next section) it is possible to write down equations that express control coefficients in terms of elasticities [27,68]. This allows one to understand how properties related to the parts of a network contribute to the properties of the whole network. In the next section a simple example will be given that illustrates this process.

### Negative Feedback Systems

Negative feedback in molecular networks is probably one of the most celebrated regulatory mechanisms in biology.

There have been numerous discussions on the role of feedback in biochemical networks ever since the discovery by Umbarger [81] of feedback inhibition in the isoleucine biosynthesis pathway and the feedback inhibition of aspartate transcarbamylase in *E. coli* by Yates and Pardee [89].

Probably the most extensive mathematical analysis of biochemical feedback was conducted by Savageau [69,70,72] and Burns and Kacser [13,44] and Othmer and Tyson [57,78] in the 1970s and Dibrov et al. in the early 1980s [22]. More recently, Cinquin and Demongeot have published an interesting review on the roles of feedback in biological systems [18].

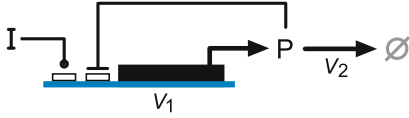
To begin the discussion let us consider the simplest possible system, a linear chain of two reactions. These reactions could be two metabolic enzymes in sequence or a gene expressing a protein and the protein undergoing degradation. Figure 2 illustrates a model of simple gene expression where the gene product,  $P$ , negatively influences its own production. If our reaction steps were two metabolic enzymes in sequence, the negative inhibition would be generated by simple product inhibition. In either case what properties might such a simple system have? First one needs to construct a mathematical model of our system. Since there is only one floating species in the model, protein  $P$ , there will only be one differential equation. It will be assumed that the input concentration,  $I$  is fixed and that the protein product degrades into an infinite volume, both constraints will ensure that the system will have a steady state. This is a common feature of all models of molecular networks. Without these constraints the network would simply evolve to thermodynamic equilibrium which is neither interesting nor relevant to the study of biological networks. For the negative feedback one can assume a simple Hill-like inhibition, the model equation can therefore be written as:

$$\frac{dP}{dt} = \frac{IV_{\max}}{(1 + (P/K)^h)} - k_1 P.$$

$V_{\max}$  is the maximal rate of expression,  $h$  is the Hill coefficient,  $k_1$  the degradation rate and  $K$  the inhibition constant. There are two properties that can be investigated, how long does it take for the network to reach steady state and once at steady state how robust is the steady state to parameter variation?

In order to examine the properties of this network two strategies can be used, the network can be compared with and without negative feedback or small changes can be made to the strength of the feedback and the effect observed.





**Biological Models of Molecular Network Dynamics, Figure 2**

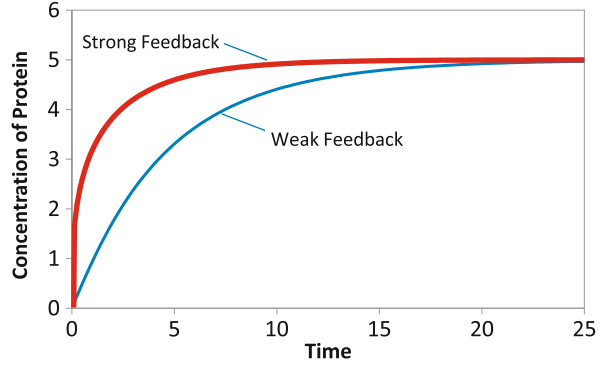
Simple Negative feedback in a genetic circuit.  $P$  represents the protein product from gene expression  $v_1$ .  $v_2$  represents protein degradation.  $I$  represents the input signal that can activate gene expression. Product,  $P$ , binds to an operator on the gene repressing gene expression. Bars represent inhibition; round ends, activation. Similar conventions apply to all subsequent figures

### Time to Reach Steady State

The time taken to reach steady state can be done by carrying out a simulation and observing the progress curves. If one simply took the network with weak feedback and compared it with the network with strong feedback one immediate problem arises. Negative feedback reduces the steady state concentration of  $P$ . This means that a comparison between the two systems will be unfair because of the difference in steady state concentrations in  $P$ . In order to make a fair comparison the maximal rate of gene expression in the strongly regulated system must be raised until the network has the same steady state as the unregulated network [1]. This has been done in the simulations shown in Fig. 3. Note that the rise time for the circuit with negative feedback is much shorter. This can also be easily seen in the rate of change of  $P$  at time zero for both circuits. With weak feedback, the rate of change of  $P$  is 1 unit per unit time. With strong feedback, this rises to 5000 units per unit time. We can therefore conclude from this simple study that for comparable systems, negative feedback reduces the time it takes to reach steady state. This effect has been observed in synthetic networks built in *E. coli* cells. In these experiments the rise time was observed using GFP (green fluorescence protein) tagged proteins in autoregulated and non-regulated transcription units [64]. The researchers observed that negative autoregulation reduced the rise time to one fifth of the time taken for the non-regulated circuit.

### Properties of the Steady State

In the previous section simple mathematical models showed that the time taken to reach steady state in the presence of negative feedback could be significantly reduced although at the expense of increasing the maximal expression rate for the gene. What about the steady state itself, what effect does negative feedback have on the steady state? This question can be answered by using control analysis. The following equations represent the sum-



**Biological Models of Molecular Network Dynamics, Figure 3**

Rise times for a simple negative feedback circuit with weak and strong negative feedback. For a fair comparison, both systems have been adjusted so that they both reach the same steady state. Note that the rise time for the circuit with strong negative feedback is much faster.  $k_1 = 0.1$ ,  $I = 1$ ,  $P = 0$  at  $t = 0$ ,  $n = 2$ . Weak feedback:  $K = 5$ ,  $V_{\max} = 1$ . Strong feedback:  $K = 0.05$ ,  $V_{\max} = 5000$

mation and connectivity theorems for the simple regulated gene circuit in Fig. 2.

$$C_{V_{\max}}^P + C_{k_1}^P = 0$$

$$C_{V_{\max}}^P \varepsilon_P^{v_1} + C_{k_1}^P \varepsilon_P^{v_2} = 0.$$

These two equations can be used to solve for the control coefficients in terms of the elasticities:

$$C_{V_{\max}}^P = \frac{1}{\varepsilon_P^{v_1} - \varepsilon_P^{v_2}}$$

$$C_{k_1}^P = -\frac{1}{\varepsilon_P^{v_1} - \varepsilon_P^{v_2}}.$$

These equations can be used to study how the steady state concentration of  $P$  is influenced by the strength of the feedback. To answer this one need only look at  $C_{V_{\max}}^P$  and  $C_{k_1}^P$ . The feedback strength is given by the elasticity coefficient,  $\varepsilon_P^{v_1}$ , since this measures the influence that the protein product,  $P$ , has on gene expression. Note that the negative feedback elasticity,  $\varepsilon_P^{v_1}$ , is negative, that is increasing  $P$  decreases the rate of gene expression. If the absolute value of  $\varepsilon_P^{v_1}$  is small then the strength of the negative feedback is weak, whereas if the absolute value is large then the strength of the negative feedback is strong. Consider the first case when the feedback is weak, i.e.  $\varepsilon_P^{v_1} \approx 0$ . The degradation rate can be assumed to be linear in  $P$ , so that its elasticity,  $\varepsilon_P^{v_2}$  will equal one. Taking both into consideration means that the control coefficient on the gene will tend to unity and the control coefficient on protein degradation will tend to minus unity. This result means that

variation in either gene expression or protein degradation will have significant effects on the steady state concentration of protein. Consider now the alternative where there is strong feedback, in this case  $\varepsilon_P^{V_1}$  will have a large absolute value and  $\varepsilon_P^{V_2} = 1$  as before. If the feedback elasticity is sufficiently large, which it can be if the Hill coefficient is large, both control coefficients will tend to small values. This means that both variation in gene expression and protein degradation will have little effect on the steady state concentration of protein.

We conclude that negative feedback stabilizes the concentration of the protein by reducing the effect of parameter variation. Intuitively, this is a well known idea and engineers frequently use negative feedback to stabilize systems. However the approach that uses control analysis enables us to quantify the effect and understand the properties of negative feedback in more depth.

There are many examples in metabolic networks where feedback appears to be used to stabilize a concentration, for example the negative feedbacks that go to the enzyme phosphofructokinase can stabilize ATP concentrations in the face of varying ATP demand [48]. In addition, synthetic networks have also been constructed [9] where a comparison is made between a gene circuit with and without negative feedback. They show that with feedback the effect of noise on the system, as indicated by the expression of green fluorescence protein (GFP), is greatly reduced, thus supporting the theoretical predication that negative feedback suppresses the influence of variation in the network.

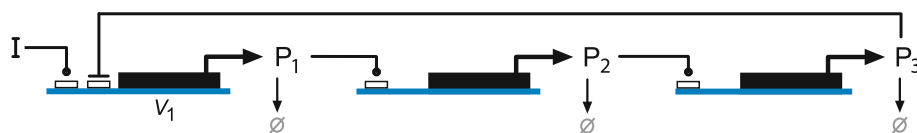
### Long Range Negative Feedback

Figure 4 illustrates a long range negative feedback where instead of the product of a gene inhibiting its own production, a product further downstream of a cascade of genes feeds back a signal to the first gene. The steady state properties of this network can be quite different from a single stage network. What is critical in a multi stage system is the delay in signal transmission caused by the time it takes each genetic unit to respond to changes upstream. Consider a pulse of  $P_1$  added to the pathway in Fig. 4. The pulse will increase the rate of degradation of  $P_1$  but it will also

stimulate the second stage to express more  $P_2$ , this in turn will cause  $P_3$  to increase its production rate. The propagation of the pulse of  $P_1$  down the chain will however take a finite amount of time during which the concentration of  $P_1$  is likely to have recovered to its original steady state level due to degradation. At the same time that  $P_1$  is recovering  $P_3$  will be increasing and via the negative feedback will suppress the production of  $P_1$ . Therefore rather than being restored  $P_1$  will in fact be forced to go under its nominal steady state level. The response to the signal is therefore out of phase, that is whatever changes occur upstream, the time lags in the network cause the negative feedback to restore the system at the wrong time. Given a sufficient delay and enough strength in the feedback, the inability to correctly restore perturbations can develop into an unstable situation leading to sustained oscillations in the network. Such oscillatory systems are well known in engineering and are often called phase-shift oscillators [43].

A critical factor that determines the onset of oscillations, apart from the number of stages in the cascade is the strength of the feedback. Savageau [70,71,72] showed, in a detailed analysis, that for a metabolic pathway with negative feedback, if the substrate elasticities are equal (e.g. all first-order kinetics) then the ratio of the feedback elasticity ( $\varepsilon_{inh}$ ) to the output elasticity ( $\varepsilon_{sub}, \varepsilon_3^4$ ) determined the onset of oscillations (Table 1). Similar criterion would hold for a genetic network although the detailed analysis has not yet been done. Table 1 shows that as the pathway becomes longer, less feedback inhibition is required to destabilize the pathway. This highlights the other factor that contributes to instability, the delay in routing the signal around the network.

All feedback oscillators require some device to provide amplification of a signal combined with a suitable time delay so that the signal response can go out of phase. In metabolic pathways amplification is often provided by a cooperative enzyme while the delay is provided by the intermediate steps in the pathway. In signaling pathways, amplification can be generated by covalent modification cycles [33,45] while in genetic networks, amplification can be provided by cooperative binding of the transcription factor to the promoter site. The criterion for instability is the ratio of the inhibition elasticity to the substrate elas-



Biological Models of Molecular Network Dynamics, Figure 4

Long range negative feedback comprising of three gene stages with the product of the last stage inhibiting the first



Biological Models of Molecular Network Dynamics, Fig. 5  
Ring oscillator, also called the repressilator

Biological Models of Molecular Network Dynamics, Table 1  
Relationship between the pathway length and the degree of feedback inhibition on the threshold for stability.  $\epsilon_{inh}$  is the elasticity of the feedback inhibition and  $\epsilon_{sub}$  is the elasticity of the distal step with respect to the signal. See [71] for details of calculations

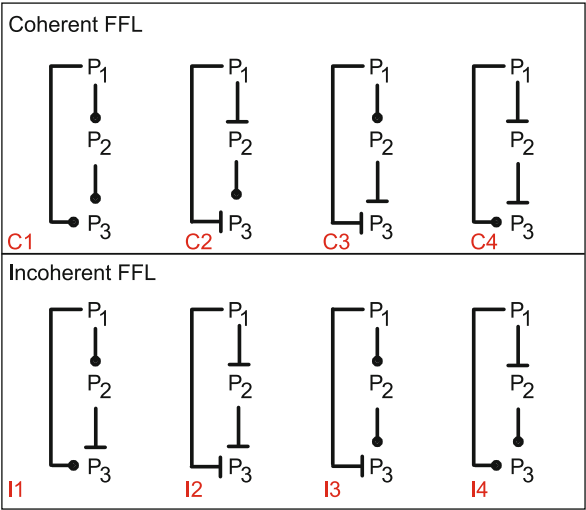
Length of Pathway	Instability Threshold $-\epsilon_{inh}/\epsilon_{sub}$
1	stable
2	stable
3	8.0
4	4.0
5	2.9
6	2.4
7	2.1
$\vdots$	$\vdots$
$\infty$	1.0

ticity. If the output reaction of the pathway is governed by a saturable enzyme then it is possible to have  $\epsilon_{sub}$  less than unity. This means that it is possible to trade cooperativity at the inhibition site with saturation at the output reaction. The modified phase-shift Goodwin model of Bliss [12] illustrates a model with no cooperativity at the inhibition site but with a degree of saturation at the output reaction by using a simple Michaelis–Menten rate law.

The ring oscillator, Fig. 5, also called the repressilator, is a type of phase-shift oscillator. This device is composed of an odd number of gene inverters connected in a closed loop. A ring oscillator has been built experimentally in *E. coli* [23]. Ring oscillators with an even number of inverters can also be used to form memory units or toggle switches. The even number of units means that the signal latches to either on or off, the final state depending on the initial conditions. Toggle circuits have also been implemented experimentally in *E. coli* [30].

FeedForward Systems

Many efforts in modeling cellular networks have been to understand the effects of negative feedback. However in recent years there has been a growing awareness of feed-forward loops (FFL). Recent work in identifying motifs in genetic networks of *E. coli* and yeast has revealed the oc-

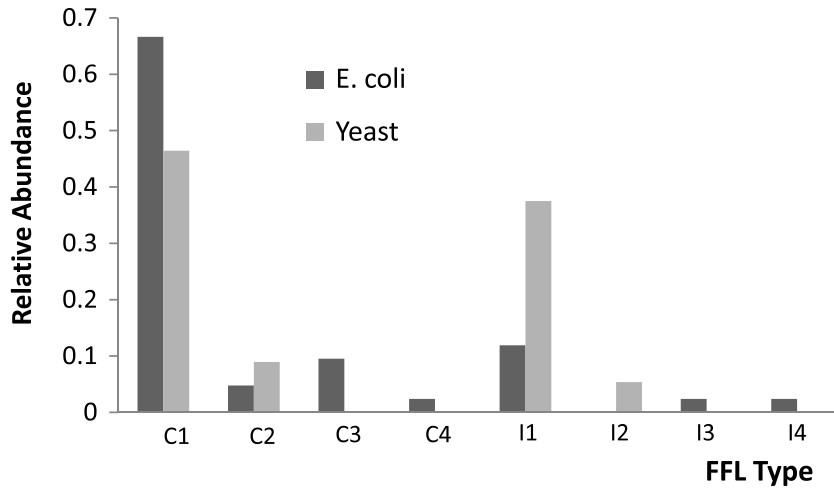


Biological Models of Molecular Network Dynamics, Fig. 6  
Full complement of feedforward motifs, classified into coherent and incoherent types

currence of significant numbers of feedforward loops [52] in these organisms, Fig. 7. These motifs have been classified into two broad groups, coherent and incoherent networks. Within each group four combinations of regulation can be discerned. Figure 6 illustrates all the possible combinations for the FFL motif. Figure 7 shows the relative abundance of the different kinds of feedforward motifs found in *E. coli* and yeast. This data indicates that two types predominate in both organisms, Coherent Type 1 (C1) and Incoherent Type 1 (I1). What is particularly interesting is that these two types have distinct behavioral properties.

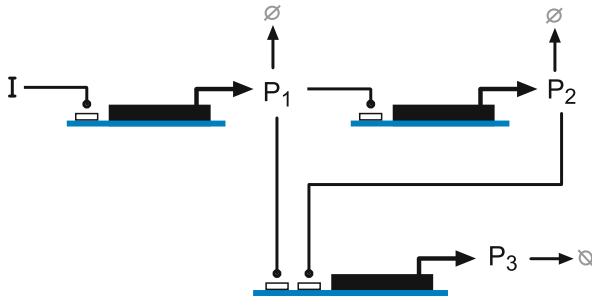
Properties of the Coherent Type I Motif

The best way to understand the properties of the Coherent Type I feedforward network is to build a computer model and run a simulation. Figure 8 illustrates a possible genetic feedforward Coherent Type I network. Figure 9 illustrates one of the key properties of the Coherent Type I feedforward network. The output signal,  $P_3$ , is controlled by  $P_2$  and  $P_1$ . If a pulse signal is applied to  $P_1$ , the signal travels two routes to get to  $P_3$ . If the pulse is too short,  $P_2$  does



Biological Models of Molecular Network Dynamics, Fig. 7

Relative abundance of different FFL types in yeast and *E. coli*. Labels on the X-axis refer to the particular FFL motifs seen in Figure 6. Data taken from Mangan et al. [52]



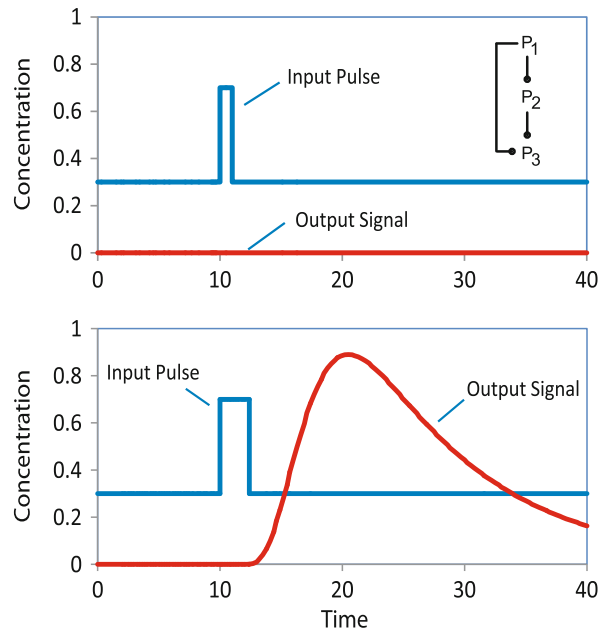
Biological Models of Molecular Network Dynamics, Fig. 8  
Example of a feedforward Coherent Type I genetic network

not have sufficient time to increase in order to reach the threshold level at  $P_3$ , as a result, short pulses do not turn  $P_3$  on. This effect is shown in the upper panel of Fig. 9. In sharp contrast, if the  $P_1$  pulse width is wide enough,  $P_2$  has sufficient time to increase so that it reaches the threshold level and activates  $P_3$ . The network therefore acts as a transient signal filter. Only signals of sufficient duration will result in activation.

### Properties of the Incoherent Type I Motif

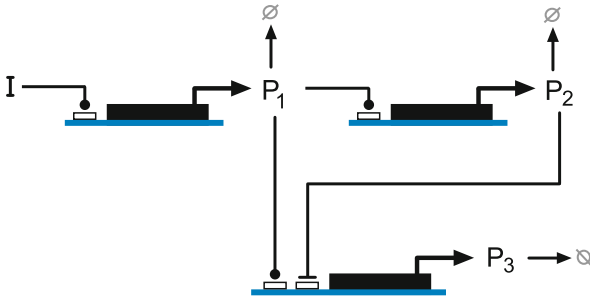
The Incoherent Type I feedforward network (Fig. 10) has a completely different response compared to the Coherent Type I network.

The Incoherent Type I network has a number of interesting behaviors, including pulse generator, concentration band detector and frequency band pass filter. The pulse generator simulation is shown in Fig. 11 where an input step function is applied to the input,  $P_1$  and the effect on

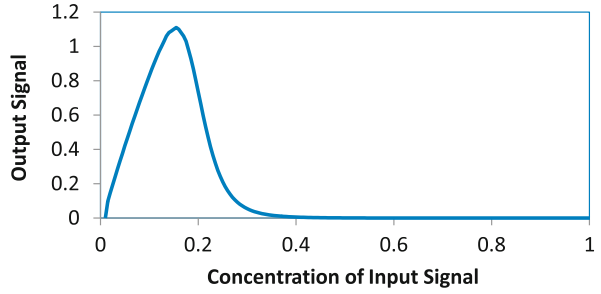


Biological Models of Molecular Network Dynamics, Fig. 9

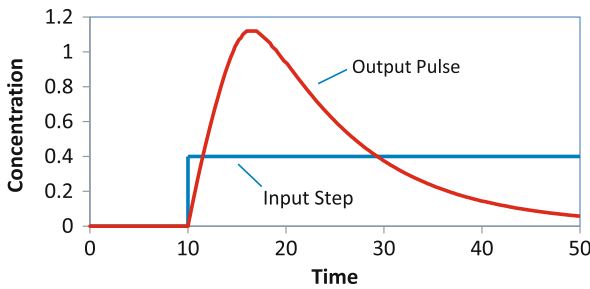
Simulation of a Coherent Type I feedforward network. The top panel shows the effect of a narrow pulse in the concentration of  $P_1$ . The output signal,  $P_3$ , shows no response. In the lower panel, the input signal,  $P_1$ , is wider, this gives time for the feedforward loop to activate as shown by the large change in output signal,  $P_3$ . The Coherent Type I FFL will only fire if the input pulse is of sufficient width. This means that the network can filter out transient input pulses and only respond to input pulses of sufficient duration



**Biological Models of Molecular Network Dynamics, Fig. 10**  
Example of a feedforward Incoherent Type I genetic network. Note that the regulation on  $P_3$  is antagonistic compared to the Coherent Type I network



**Biological Models of Molecular Network Dynamics, Fig. 12**  
Simulation of the concentration band detector characteristics of an Incoherent Type I feedforward network. The X-axis shows the change in the input signal while the Y-axis shows the steady state concentration in the output signal. The network only responds in a particular range of input concentration



**Biological Models of Molecular Network Dynamics, Fig. 11**  
Simulation of the pulse generation characteristics of an Incoherent Type I feedforward network. The X-axis shows the change in the input signal over time. At 10 time units, a step signal is applied, this causes the output to rise then fall in a pulse like manner. The width of the pulse can be adjusted by changing the degree of cooperativity on  $P_3$

$P_3$  observed. In this case,  $P_3$  rises rapidly then falls off in a pulse like manner even though the input signal remains on. This is due to the delay in the inhibition route, initially the inhibition is weak and the output raises, but eventually  $P_2$  increases and begins to repress the production of  $P_3$ . The second type of behavior that the network can display is to act as a concentration band detector. The simulation shown in Fig. 12 shows the network turning on in a specific range of input signal. The position of the peak in the response can be adjusted by changing the strength of inhibition and the input threshold from  $P_1$  to  $P_3$ .

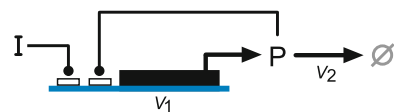
Finally, an Incoherent Type I network can act as a band pass filter. That is the network will respond more strongly if the input signal varies at a specific range of frequencies. That is at high and low frequencies the network will not respond but at mid range frequencies it will.

A series of synthetic Incoherent Type I networks have been built in *E. coli* that illustrate the band pass behavior [24]. Further details on the properties of feedforward networks can be found in Uri Alon's book [1].

## Positive Feedback

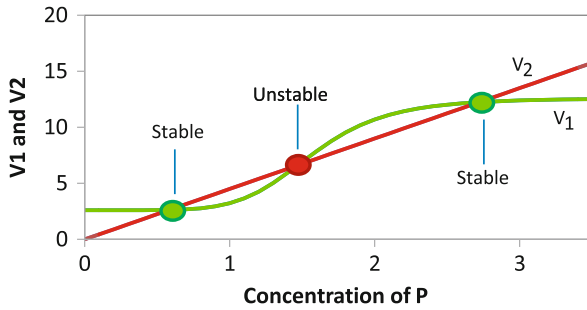
The alternative to negative feedback is positive feedback; that is where the product of a process activates its own production. Ten percent of known transcription factors in *E. coli* are thought to act as positive feedback regulators [1]. Positive feedback shows many of the opposite properties to those of negative feedback. For example, systems with positive feedback have slow responses compared to systems without feedback and unlike negative feedback which tends to stabilize a system, positive feedback tends to destabilize. One may wonder why evolution has favored a destabilizing force since homeostasis and robustness are presumably desirable attributes to have inside a cell. What positive feedback brings is in fact robustness but via a behavior called bistability. Figure 13 shows a simple gene circuit where the gene product,  $P$ , positively influences its own production. We can easily model positive feedback using Hill like equations. A mathematical model for the system in Fig. 13 is given below:

$$\begin{aligned} v_1 &= k_1 + \frac{V_{\max} P^n}{15 + P^n} \\ v_2 &= k_2 P \\ \frac{dP}{dt} &= v_1 - v_2. \end{aligned} \quad (7)$$



**Biological Models of Molecular Network Dynamics, Fig. 13**  
Simple positive feedback



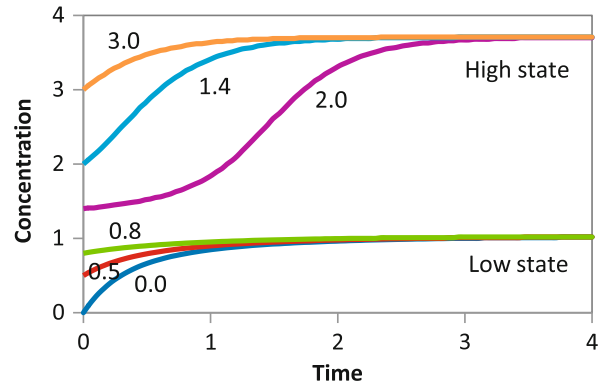


Biological Models of Molecular Network Dynamics, Fig. 14

Simple positive feedback. The plot shows  $v_1$  and  $v_2$  plotted against the concentration of  $P$ . Where the curves intersect, the system is at steady state. By examining perturbations around each steady state it is possible to determine that two of the states are stable and one is unstable

The equation that describes gene expression,  $v_1$ , has two terms,  $k_1$  which represents a basal level of gene expression and a second Hill like term that represents the positive feedback. The second rate,  $v_2$ , is the protein degradation rate. The net rate of change of  $P$  is the difference between the two. To illustrate the key behavior that the simple genetic circuit has, the rates  $v_1$  and  $v_2$  are plotted against the concentration of  $P$ , Fig. 14 shows this plot. The first thing to note is that where the two lines intersect the system is at steady state since at these points  $v_1$  must equal  $v_2$ , which is the definition of steady state. What is interesting is that there are three possible steady states, where each steady state displays a different steady state concentration of  $P$ . If we focus on each steady state and examine the relative rates,  $v_1$  and  $v_2$ , as we move away from the steady state by changing  $P$ , it becomes clear that the lower and upper steady states are stable, that is any changes in  $P$  at these points will restore  $P$ . However, if we were to change  $P$  at the middle steady state, for example increase  $P$ , it should be clear from the graph that  $v_1$  will be higher than  $v_2$ , this means that more  $P$  is being produced which will move  $P$  further away from the steady state. This tells us the middle steady state is in fact unstable.

What we have here is something akin to simple light toggle switch. That is, a switch that has two stable states, which we might nominally call high (on) and low (off) and one middle state which is unstable. In a toggle switch one could imagine, very carefully, balancing the switch so that it is halfway between on and off, knowing that a slight nudge will flip the switch to either the on or off state. Positive feedback can lead to systems that display bistability. To illustrate bistability in a different way, Fig. 15 shows time course simulations of the genetic switch in Fig. 13. Six different simulations are shown, each simulation starts at a different initial concentration of  $P$ . At low concentra-



Biological Models of Molecular Network Dynamics, Fig. 15

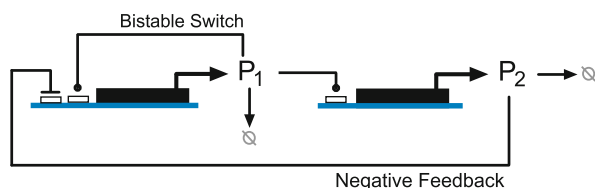
Simulation of model in Fig. 13. Each curve displays the change in concentration for the gene product,  $P$  at different starting concentrations for  $P$ . Low concentrations of  $P$  settle to the low steady state, while higher starting concentrations evolve to the high state. Model given by Eqs. (7). Parameters:  $k_1 = 2.6$ ,  $k_2 = 3.2$ ,  $V_{\max} = 10$  and  $n = 4$

tions, the time course behavior of  $P$  converges to a low state, at around 1.0 unit of  $P$ . High concentrations of  $P$ , on the other hand, converge to a different final state, a high state at around 3.5 units of  $P$ . What evolutionary selective advantage might bistability have? One clear benefit is that bistable states are very stable to perturbations in  $P$  such as thermal noise or transient signals that should be ignored. Once a bistable switch has settled to either the high or low state it is difficult to persuade the system to switch to the alternative state. In a noisy cellular environment, or where a definite decision needs to be made, for example, replication, sporulation or apoptosis, bistability is a useful behavior to have.

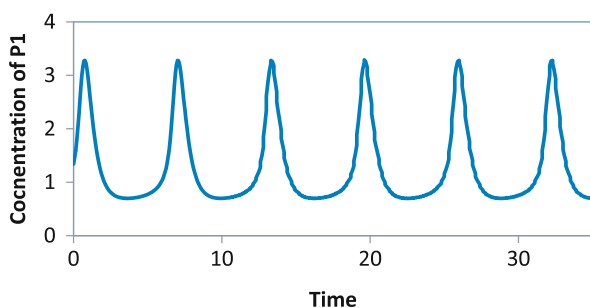
### Relaxation Oscillator

In addition to providing switch like behavior, bistability can also be coupled with negative feedback to generate robust oscillators called relaxation oscillators [80]. Relaxation oscillators are commonly used in electronics where non-sinusoidal waveforms are required, for example, sawtooth or square waveforms [43]. Electronic relaxation oscillators repeatedly alternate between two states at a frequency that depends on a charging device such as a capacitor or an inductor. As a result, relaxation oscillators are characterized by two time-scales which is why they can be used to generate non-sinusoidal waveforms.

In biological networks, a relaxation oscillator operates by charging a species concentration which, upon reaching a threshold, changes the state of a bistable switch. When the switch changes state it causes the species to discharge. Once the species has discharged the bistable switch returns



**Biological Models of Molecular Network Dynamics, Fig. 16**  
Simple genetic circuit that behaves as a relaxation oscillator



**Biological Models of Molecular Network Dynamics, Fig. 17**  
Simulation of the relaxation model depicted in Fig. 16. Relaxation oscillators often have a characteristic *spiky* pattern that originates from the different times scales in the bistable switch and the negative feedback component

to the original state and the sequence begins again. Figure 16 shows a simple genetic relaxation oscillator with the bistable switch on the left (formed from  $P_1$ ) and the negative feedback originating from the gene product  $P_2$ . One of the characteristics of a relaxation oscillator is the *spiky* appearance of the oscillations. This is due to the rapid switching of the bistable circuit which can be faster compared to the operation of the negative feedback. A simulation of the relaxation oscillator shown in Fig. 16 is given in Fig. 17.

The previous sections have covered a few of the common design patterns we find in biological networks but many others exist and most likely many more remain to be discovered. Although the focus has been on genetic networks, the exact same design patterns can be found in metabolic as well as signaling pathways [16,17,46]. This means that once an understanding of how these patterns operate it is relatively easy to translate this understanding to many different biological situations. Crucial to understanding how these networks operate and uncovering the basic design principles of how molecular network operate has required a concerted effort between experimentalists, modelers and theorists. Modeling in particular makes it easy, for both experimentalists and computationalists, to explore the dynamic behavior of complex network configurations. The availability of easy to use software such as the Systems Biology Workbench [10] makes modeling

and model construction accessible to a wide range of researchers in biology.

### Future Prospects

Like all good science, the development of model-based systems biology depends heavily on the ability to make measurements on real systems. In the case of model development the ideal measurements would be high resolution, the time course behavior of molecular states. In recent years there has been significant progress in this area, in particular, researchers now have the ability, through the use of fluorescent proteins for example, to measure the detailed dynamics that occur in the levels of protein in single cells. This kind of data allows model builders and experimentalists to combine their expertise in understanding molecular networks. A particularly striking example of this synergy is found in the work of Bar-Or et al. [51] and Lahav et al. [49] where the authors coupled high resolution measurements of Mdm2 and P53 protein levels in single cells with mathematical modeling. Further progress is needed however, to make more detailed measurements, in particular experimentalists still find it difficult to measure at high time resolution particular protein states (e. g. phosphorylated states) in live single cells. In addition, real-time measurements of metabolic states other than naturally fluorescent molecules such as NAD is still not possible. However recent progress in aptamer technology [8,86] may allow, for the first time, detailed measurements in real-time of both protein and metabolic states.

In addition to making measurements, validating, or more precisely, falsifying models, is still not a well developed area. One of the few researches who has taken this seriously is Tyson and Novak who have developed sophisticated computer tools to match models against experimental data to ensure the validity of models as more data is forthcoming [83].

Perhaps one of the most exiting and promising avenues for matching experimental measurements with model building is to through synthetic biology. The approach that synthetic biologists take is to build new networks in vivo, in organisms such as *E. coli*. The great advantage of synthetic biology is that networks can be studied under controlled conditions [24]. Many insights into network dynamics have been provided by using synthetic biology techniques, far more probably than any other approach [2,58,84].

### Bibliography

1. Alon U (2006) An Introduction to Systems Biology: Design Principles of Biological Circuits. Chapman & Hall/Crc Mathematical

- and Computational Biology Series. Chapman & Hall/CRC, Boca Raton
2. Alon U (2007) Network motifs: theory and experimental approaches. *Nat Rev Genet* 8(6):450–461
  3. Andrianantoandro E, Basu S, Karig DK, Weiss R (2006) Synthetic biology: new engineering rules for an emerging discipline. *Mol Syst Biol* 2:2006–2006
  4. Arieu R (1976) Ockham's Razor: A Historical and Philosophical Analysis of Ockham's Principle of Parsimony. Champaign-Urbana, University of Illinois
  5. Aris R (1965) Prolegomena to the Rational Analysis of Systems of Chemical Reactions. *Arch Rational Mech Anal* 19:81–99
  6. Arkin A, Ross J, McAdams HH (1998) Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. *Genetics* 149:1633–48
  7. Atkins P (2001) *Physical Chemistry*, 7th edn. W. H. Freeman
  8. Bayer TS, Smolke CD (2005) Programmable ligand-controlled riboregulators of eukaryotic gene expression. *Nat Biotechnol* 23(3):337–343
  9. Becskei A, Serrano L (2000) Engineering stability in gene networks by autoregulation. *Nature* 405:590–593
  10. Bergmann FT, Vallabhajosyula RR, Sauro HM (2006) Computational Tools for Modeling Protein Networks. *Current Proteomics* 3(3):181–197
  11. Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, et al (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* 277(5331):1453–1474
  12. Bliss RD, Painter PR, Marr AG (1982) Role of feedback inhibition in stabilizing the classical operon. *J Theor Biol* 97(2):177–193
  13. Burns JA (1971) *Studies on Complex Enzyme Systems*. Dissertation, University of Edinburgh. <http://www.sys-bio.org/BurnsThesis>
  14. Cannon RH (1967) *Dynamics of Physical Systems*. McGraw-Hill College, New York
  15. Chen YD, Westerhoff HV (1986) How do Inhibitors and Modifiers of Individual Enzymes Affect Steady-state Fluxes and Concentrations in Metabolic Systems? *Math Model* 7:1173–1180
  16. Chickarmane V, Kholodenko BN, Sauro HM (2007) Oscillatory dynamics arising from competitive inhibition and multisite phosphorylation. *J Theor Biol* 244(1):68–76
  17. Chickarmane V, Ray A, Sauro HM, Nadim A (2007) A Model for p53 Dynamics Triggered by DNA Damage. *SIAM J Appl Dynamical Syst* 6(1):61–78
  18. Cinquin O, Demongeot J (2002) Roles of positive and negative feedback in biological systems. *C R Biol* 325:1085–1095
  19. Clarke BL (1980) *Stability of complex reaction networks*. Adv. Chem. Phys., vol. 42. Wiley, New York
  20. Cornish-Bowden A (1995) *Fundamentals of Enzyme Kinetics*. Portland Press, London
  21. Cornish-Bowden A, Cárdenas ML (2002) Metabolic balance sheets. *Nature* 420(6912):129–130
  22. Dibrov BF, Zhabotinsky AM, Kholodenko BN (1982) Dynamic stability of steady states and static stabilization in unbranched metabolic pathways. *J Math Biol* 15:51–63
  23. Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* 403:335–338
  24. Entus R, Aufderheide B, Sauro HM (2007) Design and implementation of three incoherent feed-forward motif based biological concentration sensors. *Systems and Synthetic Biology*. doi:10.1007/s11693-007-9008-6
  25. Erdi P, Toth J (1989) *Mathematical Models of Chemical Reactions*. Theory and Applications of Deterministic and Stochastic Models. Manchester University Press, Manchester, Princeton University Press, Princeton
  26. Fell D (1997) *Understanding the Control of Metabolism*. Portland Press, London
  27. Fell DA, Sauro HM (1985) Metabolic Control Analysis: Additional relationships between elasticities and control coefficients. *Eur J Biochem* 148:555–561
  28. Fell DA, Small JR (1986) Fat synthesis in adipose tissue: an examination of stoichiometric constraints. *Biochem J* 238:781–786
  29. Ferrell JE (1996) Tripping the switch fantastic: how a protein kinase cascade can convert graded inputs into switch-like outputs. *Trends in Biochemical Sciences* 21:460–466
  30. Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403:339–342
  31. Gillespie DT (2007) Stochastic simulation of chemical kinetics. *Annu Rev Phys Chem* 58:35–55
  32. Goldbeter A (1997) *Biochemical Oscillations and Cellular Rhythms: The Molecular Bases of Periodic and Chaotic Behaviour*. Cambridge University Press, Cambridge
  33. Goldbeter A, Koshland DE (1984) Ultrasensitivity in biochemical systems controlled by covalent modification. Interplay between zero-order and multistep effects. *J Biol Chem* 259:14441–7
  34. Heinrich R, Rapoport TA (1974) *A Linear Steady-state Treatment of Enzymatic Chains; General Properties, Control and Effector Strength*. *Eur J Biochem* 42:89–95
  35. Heinrich R, Schuster S (1996) *The Regulation of Cellular Systems*. Chapman and Hall, New York
  36. Heinrich R, Rapoport SM, Rapoport TA (1977) Metabolic regulation and mathematical models. *Prog Biophys Molec Biol* 32:1–82
  37. Hofmeyr JHS (1986) Steady state modelling of metabolic pathways: a guide for the prespective simulator. *Comp Appl Biosci* 2:5–11
  38. Hofmeyr JHS (1986) *Studies in steady state modelling and control analysis of metabolic systems*. Dissertation, University of Stellenbosch
  39. Hofmeyr JHS (2001) *Metabolic Control Analysis in a Nutshell*. In: *Proceedings of the Second International Conference on Systems Biology*, Caltech
  40. Hood L, Heath JR, Phelps ME, Lin B (2004) Systems biology and new technologies enable predictive and preventative medicine. *Science* 306(5696):640–643
  41. Horn F, Jackson R (1972) General Mass Action Kinetics. *Arch Rational Mech Anal* 47:81–116
  42. Ideker T, Galitski T, Hood L (2001) A new approach to decoding life: systems biology. *Annu Rev Genomics Hum Genet* 2:343–372
  43. Jones MH (1977) *A practical introduction to electronic circuits*. Cambridge University Press, Cambridge
  44. Kacser H, Burns JA (1973) The Control of Flux. In D. D. Davies (eds) *Rate Control of Biological Processes*, Symp Soc Exp Biol, vol 27, Cambridge University Press, Cambridge, pp 65–104
  45. Kholodenko BN (2000) Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. *Eur J Biochem* 267:1583–1588
  46. Kholodenko BN (2006) Cell-signalling dynamics in time and space. *Nat Rev Mol Cell Biol* 7(3):165–176

47. Klipp E, Herwig R, Kowald A, Wierling C, Lehrach H (2005) *Systems Biology in Practice*. Wiley-VCH, Weinheim
48. Koebmann BJ, Westerhoff HV, Snoep JL, Nilsson D, Jensen PR (2002) The Glycolytic Flux in *Escherichia coli* Is Controlled by the Demand for ATP. *J Bacteriol* 184(14):3909–3916
49. Lahav G, Rosenfeld N, Sigal A, Geva-Zatorsky N, Levine AJ, Elowitz MB, Alon U (2004) Dynamics of the p53-Mdm2 feedback loop in individual cells. *Nature, Genetics* 36(2):147–150
50. Lauffenburger DA (2000) Cell signaling pathways as control modules: complexity for simplicity? *Proc Natl Acad Sci USA* 97:5031–3
51. Lev Bar-Or R, Maya R, Segel LA, Alon U, Levine AJ, Oren M (2000) Generation of oscillations by the p53-Mdm2 feedback loop: a theoretical and experimental study. *Proc Natl Acad Sci USA* 97(21):11250–11255
52. Mangan S, Itzkovitz S, Zaslaver A, Alon U (2006) The incoherent feed-forward loop accelerates the response-time of the gal system of *Escherichia coli*. *J Mol Biol* 356(5):1073–1081
53. Markevich NI, Hoek JB, Kholodenko BN (2004) Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. *J Cell Biol* 164:353–9
54. Moniz-Barreto P, Fell DA (1993) Simulation of dioxygen free radical reactions. *Biochem Soc Trans* 21(3):256–256
55. Moore WJ (1972) *Physical Chemistry*. 5th edn. Longman, London, Prentice Hall, NJ
56. Oda K, Kitano H (2006) A comprehensive map of the toll-like receptor signaling network. *Mol Syst Biol* 2:2006–2006
57. Othmer HH (1976) The quantitative dynamics of a class of biochemical control circuits. *J Math Biol* 37:53–78
58. Ozbudak EM, Thattai M, Lim HN, Shraiman BI, Van Oudenaarden A (2004) Multistability in the lactose utilization network of *Escherichia coli*. *Nature* 427(6976):737–740
59. Papin JA, Stelling J, Price ND, Klamt S, Schuster S, Palsson BO (2004) Comparison of network-based pathway analysis methods. *Trends Biotechnol* 22(8):400–405
60. Paulsson J, Berg OG, Ehrenberg M (2000) Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation. *Proc Natl Acad Sci USA* 97(13):7148–7153
61. Reder C (1988) *Metabolic Control Theory: A Structural Approach*. *J Theor Biol* 135:175–201
62. Reich JG, Selkov EE (1981) *Energy metabolism of the cell*. Academic Press, London
63. Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, et al (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440(7086):940–943
64. Rosenfeld N, Elowitz MB, Alon U (2002) Negative autoregulation speeds the response times of transcription networks. *J Mol Biol* 323(5):785–793
65. Samoilov M, Pilyasnov S, Arkin AP (2005) Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proc Natl Acad Sci USA* 102(7):2310–2315
66. Sauro HM, Ingalls B (2004) Conservation analysis in biochemical networks: computational issues for software writers. *Biophys Chem* 109:1–15
67. Sauro HM, Kholodenko BN (2004) Quantitative analysis of signaling networks. *Prog Biophys Mol Biol*. 86:5–43
68. Sauro HM, Small JR, Fell DA (1987) *Metabolic Control and its Analysis: Extensions to the theory and matrix method*. *Eur J Biochem* 165:215–221
69. Savageau MA (1972) The behaviour of intact biochemical control systems. *Curr Topics Cell Reg* 6:63–130
70. Savageau MA (1974) Optimal design of feedback control by inhibition: Steady-state considerations. *J Mol Evol* 4:139–156
71. Savageau MA (1975) Optimal design of feedback control by inhibition: dynamic considerations. *J Mol Evol* 5(3):199–222
72. Savageau MA (1976) *Biochemical systems analysis: a study of function and design in molecular biology*. Addison-Wesley, Reading
73. Schuster S, Dandekar T, Fell DA (1999) Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends Biotechnol* 17(2):53–60
74. Schuster S, Fell DA, Dandekar T (2000) A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nature Biotechnology* 18:326–332
75. Segel IH (1975) *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*. Wiley-Interscience, New York
76. Stephanopoulos GN, Aristidou AA, Nielsen J (1998) *Metabolic Engineering: Principles and Methodologies*. Academic Press, San Diego
77. Strogatz S (2001) *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. Perseus Books Group, Reading
78. Tyson J, Othmer HG (1978) The dynamics of feedback control circuits in biochemical pathways. In: Rosen R, Snell FM (eds) *Progress in Theoretical Biology*, vol 5. Academic press, New York, pp 1–62
79. Tyson JJ, Chen K, Novak B (2001) *Network Dynamics And Cell Physiology*. *Nat Rev Mol Cell Biol* 2:908–916
80. Tyson JJ, Chen KC, Novak B (2003) Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr Opin Cell Biol* 15:221–231
81. Umbarger HE (1956) Evidence for a Negative-Feedback Mechanism in the Biosynthesis of Leucine. *Science* 123:848
82. Vallabhajosyula RR, Chickarmane V, Sauro HM (2006) Conservation analysis of large biochemical networks. *Bioinformatics* 22(3):346–353
83. Vass M, Allen N, Shaffer CA, Ramakrishnan N, Watson LT, Tyson JJ (2004) the JigCell model builder and run manager. *Bioinformatics* 20(18):3680–3681
84. Voigt CA (2006) Genetic parts to program bacteria. *Curr Opin Biotechnol* 17(5):548–557
85. Wilkinson DJ (2006) *Stochastic Modelling for Systems Biology*. Chapman and Hall, New York
86. Win MN, Smolke CD (2007) A modular and extensible RNA-based gene-regulatory platform for engineering cellular function. *Proc Natl Acad Sci U S A* 104(36):14283–14288
87. Wolf DM, Arkin AP (2003) Motifs, modules and games in bacteria. *Current Opinion in Microbiology* 6:125–34
88. Wolkenhauer O, Ullah M, Wellstead P, Cho KH (2005) The dynamic systems approach to control and regulation of intracellular networks. *FEBS Lett* 579(8):1846–1853
89. Yates RA, Pardee AB (1956) Control of Pyrimidine Biosynthesis in *Escherichia coli* by a Feed-Back Mechanism. *J Biol Chem* 221:757–770
90. Zaslaver A, Mayo AE, Rosenberg R, Bashkin P, Sberro H, Tsalyuk M, Surette MG, et al (2004) Just-in-time transcription program in metabolic pathways. *Nat Genet* 36(5):486–491



## Biomolecular Network Structure and Function

LAN V. ZHANG<sup>1</sup>, FREDERICK P. ROTH<sup>1,2</sup>

<sup>1</sup> Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, USA

<sup>2</sup> Center for Cancer Systems Biology (CCSB), Dana-Farber Cancer Institute, Boston, USA

### Article Outline

Glossary

Definition of the Subject

Introduction

Single-Color Networks

Multi-Color Networks

Rich Network Models

Future Directions

Bibliography

### Glossary

**Biomolecule** Any organic molecule that is produced by or essential to a living organism, sometimes specifically referring to macromolecules such as a protein or nucleic acid.

**Biomolecular network** a graph representation of relationships amongst a group of biomolecules. Nodes or vertices represent biomolecules. An edge or link between two vertices indicates a relationship between the corresponding biomolecules, for example, physical interaction, genetic interaction, or regulatory relationship.

**Protein-protein interaction** the physical association of two protein molecules with each other. A pair of proteins can interact directly with physical contact, or indirectly through other biomolecule(s), often other protein(s).

**Yeast two-hybrid** an experimental method to examine protein-protein interaction, in which one protein is fused to a transcriptional activation domain (the GAL4 activation domain) and the other to a DNA-binding domain (the GAL4 DNA-binding domain), and both fusion proteins are introduced into yeast. Expression of a GAL4-regulated reporter gene with the appropriate DNA-binding sites upstream of its promoter indicates that the two proteins physically interact.

**Genetic interaction (epistasis)** functional interaction between genes, in which the action of one gene is modified by the other gene, sometimes called the modifier

gene. The gene whose phenotype is expressed is said to be *epistatic*, while the one whose phenotype is altered or suppressed is said to be *hypostatic*. *Epistasis* can either refer to this phenomenon, or more broadly to any case in which two mutations together cause a phenotype that is surprising given the phenotypes of each single mutation alone.

**“Single-color” network** a network with edges defined by only one type of interaction or relationship.

**“Multi-color” network** a network with edges defined by more than one type of interaction or relationship, with each type corresponding to a different ‘color’.

**Scale-free network** See **Power-law network**

**Power-law network** a network defined by a degree distribution which follows  $P(k) \sim k^{-\gamma}$ , where the probability  $P(k)$  that a vertex in the network connects with  $k$  other vertices is roughly proportional to  $k^{-\gamma}$ . Sometimes networks that exhibit this behavior only at high degree are also called power-law. The coefficient  $\gamma$  seems to vary approximately between 2 and 3 for most real networks. In a power-law network, majority of the vertices have low degree (connectivity), while a small fraction of the vertices have very high degree. Highly-connected vertices are referred to as *hubs*.

**Small-world network** a special type of network with: 1) short characteristic path length, such that most vertex pairs are connected to one another via only a small number of edges) and 2) high clustering coefficient, such that neighbors of a given vertex tend to be connected to one another.

**Network motif** a specific pattern of connected vertices and edges that occurs frequently within a given network.

### Definition of the Subject

Biological research over the past century or so has been dominated by reductionism – identifying and characterizing individual biomolecules – and has enjoyed enormous success. Throughout this history, however, it has become increasingly clear that an individual biomolecule can rarely account for a discrete biological function on its own. A biological process is almost always the result of a complex interplay of relationships amongst biomolecules [5,19,50,51,67,68,90,128], and the treatment of these relationships as a graph is a natural and useful abstraction.

Broadly speaking, a **biomolecular network** is a graph representation of relationships (of which there are many types) amongst a group of biomolecules. Vertices or nodes represent biomolecules, including macromolecules such



as genes, proteins, and RNAs, or small biomolecules like amino acids, sugars, and nucleic acids. In the next few sections, we focus mostly on the macromolecules. An edge or link between two vertices indicates a relationship between the corresponding biomolecules, which could include physical interaction, genetic interaction, or a regulatory relationship (e. g., the protein product of gene A regulates the expression of gene B). This abstraction, although simplifying, converts a complex web of biological relationships into a mathematical graph, from which we can study its structural features as well as their implications on biological functions.

Aside from biomolecular networks, other biological networks also bear significance in systems biology. Examples include networks of protein domains and neuronal networks [29,86]. Although many of the features described in this chapter also apply, these networks are largely out of the scope of this chapter.

## Introduction

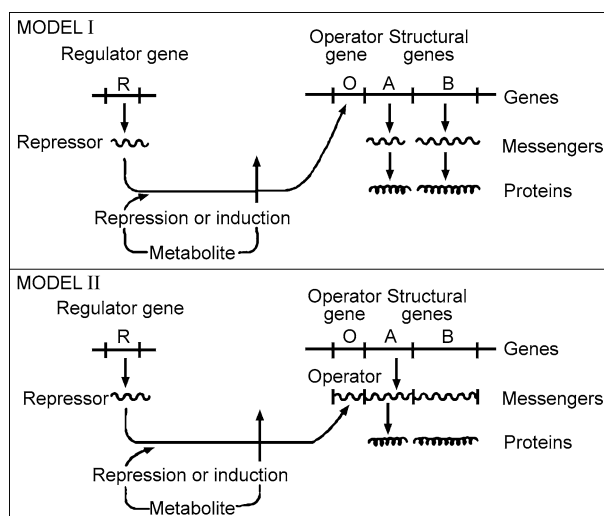
Study of small-scale biomolecular networks has a long history, stemming from “chemical graphs” introduced in the 1700s [98], with knowledge about networks of metabolic reactions reaching a large scale by the 1970s.

A prominent early example may be found in the context of Monod and Jacob’s Nobel-winning research in the 1950s, which first illuminated a transcriptional regulation system, namely control of gene expression in the *Lac* operon [58,84,85]. In their landmark work, Monod and Jacob described the operon model, in which the expression

of structural genes is suppressed by the binding of the repressor to the operator gene, which in turn can be regulated by certain metabolites [58]. Regulatory relationships amongst the repressor, the metabolite, the operator gene, and the structural genes in the operon model present one of the first examples of biomolecular networks (Fig. 1). In addition, Monod and Jacob also proposed several theoretical models, in the form of biochemical or enzymatic networks, as potential mechanisms of biochemical differentiation of cells with identical genomes [85].

Pioneering work on the *lac* operon also includes studying functional implications of bionetwork structure. Building upon Jacob and Monod’s work, the elegant experiments of Novick and Weiner [89] showed interesting features of beta-galactosidase induction – in a single cell, the production of lactose degrading enzymes is either switched on or off, and the cell’s on or off state is inherited by subsequent generations. A mixture of the two types of cells has the intermediate overall level of enzyme production observed in a normal cell culture. This all-or-none phenomenon can be explained by positive feedback in the regulatory network of beta-galactosidase synthesis, where the intracellular concentration of inducer is boosted by an intermediate mechanism, which is also induced by the inducer (Fig. 1). As we can see from Fig. 1, the very first studies on gene regulation also show how the structure of biomolecular networks, even the small-scale pathways known at that time, can illuminate regulation of protein production.

Molecular biology has entered an era in which the nature, function and relationships of macromolecules can be



**Biomolecular Network Structure and Function, Figure 1**  
**Lac operon model**

studied at a much larger scale. High-throughput technology platforms such as genome sequencing, microarrays, and yeast two-hybrid screening have enabled biologists to examine how nearly all genes or proteins in a genome or proteome interact or relate to one another [40,76,100,107,108]. Collections of these genomic or proteomic data have led to networks of various relationship types – including protein-protein interaction, transcriptional regulation, sequence homology – involving thousands of proteins or genes [22,35,52,57,69,73,99,125]. This has offered an unprecedented ‘bird’s-eye view’ of how parts of the network, each studied intensively on a local scale, connect to one another. This perspective may be useful in uncovering unknown pieces of the puzzle that together form the intricate machinery of life responsible for the behavior of the cell or the organism. The extremely large scale of these networks also presents a tremendous challenge as it is practically impossible to study each vertex or each interaction individually. A simplified abstract representation of these networks and mathematical methods permits automated study of the structural features, to identify local and global trends that relate biological macromolecules to their functional implications. Early work on large-scale bionetworks focused on **single-color networks** – networks describing a single type of relationship, while **multi-color networks** – networks combining multiple relationship types have gained increasing attention in recent years.

At an abstract level, a network of protein-protein interactions and a network of computers connected through the internet have much in common. Both have simplified representations as mathematical graphs: vertices represent proteins in the former case and computers in the latter case, and edges between vertices denote physical interactions between proteins or communication wires between computers. Concurrent with the increasing availability of large-scale biological networks, theoretical research in complex networks is advancing rapidly, with significant progress in uncovering the structural features and organizing principles of complex technological and social networks [3,18,27,88]. In a truly interdisciplinary effort, these principles were soon applied to biological networks, leading to the observation that biological networks share many global architectural features with networks from other complex systems [33,43,110,126,129], including power-law connectivity, small-world topology, and hierarchy, each discussed further below. Such resemblance suggested that many biological networks might be governed by similar ‘design’ principles, allowing methods proven in other large, non-biological complex systems to be used to characterize the intricate relationships between biomolecules.

While some reported global features of biomolecular networks, others sought local topological patterns recurrent in the network [5,73,83,103]. **Network motifs** are simple patterns of interconnection in networks that occur more frequently than expected in randomized networks [83,103], and have been proposed to represent basic building blocks of complex networks [73,83,103]. Different types of networks exhibit different motif profiles, providing a means for network classification [82]. Dobrin et al. later showed for the *E. coli* transcriptional network that subgraphs matching two types of transcriptional regulatory circuit motif – feed-forward and bi-fan motifs – overlap with one another and form large clusters which further coalesce into ‘superclusters’ [26]. Although some circuits corresponding to motifs may be capable of functioning independently [79], observation of Debrin et al., among many others, leads to the idea that instances of a network motif are often interdependent on one another as opposed to being building blocks with separable functions [138].

Cellular function is carried out through multiple types of interaction, resulting in a complex interplay of biological relationships amongst the myriad of cellular components. To have a truly global picture of the cell as a complex system, one needs to relate multiple types of interactions to one another. **Multi-color networks** are integrated networks of many relationship types, with frequent overlap between relationships of different types. For example, physically interacting proteins have expression patterns that are slightly more similar than would be expected by chance [36,60]; genes with correlated expression are more likely to be controlled by a common transcription factor [137]; and some genetic interaction types have a weak tendency to occur between homologous genes [122] (see Sect. “[Single-Color Networks](#)” for definitions of the various interaction types).

Topological study of single-color networks is paralleled in multi-color networks. Multi-color network motifs characterize relationships between different biological interaction types within local network neighborhoods [120,134,138]. Furthermore, multi-color networks also exhibit network themes – recurring higher-order interconnection patterns that encompass multiple occurrences of network motifs and reflect a common organizational principle [138]. Network themes can be tied to specific biological phenomena that represent more fundamental ‘design principles’, for example, the tendency for proteins within the same physical complex to be regulated by a common transcription factor [105,138]. This suggests that instead of representing network ‘building blocks’, network motifs should in some cases be viewed as signatures of more fundamental higher-order structures [138]. Net-

work themes also provide a more natural simplification of the otherwise complex set of relationships in an integrated network [138].

Studying the architectural features of biological networks conveys useful biological intuition and also allows us to make biological predictions. Even simple methods exploiting the predictive power of local topology have proven useful in predicting physical and genetic interactions [12,41,43,132]. Richer network models are able to incorporate more information and hence promise better predictions.

A multi-color network can be viewed as a graph in which each edge carries labels denoting the relationships between the corresponding pair of vertices. Similarly, each vertex may carry multiple labels (e.g., denoting that a particular protein is expressed in the pancreas; or that a gene is conserved in chimpanzee). A **rich network** is a network in which a vector describing potentially many vertex properties is associated with each vertex, and a vector describing potentially many edge properties is associated with each pair of vertices. By incorporating a more comprehensive set of information, these rich network models provide additional power in predicting features of genes or proteins, as well the relationships between them. In the following sections, we will survey methodology, findings, and applications for biological network models of increasing complexity, starting with single-color biomolecular networks.

As described previously, study of large-scale biomolecular networks has been enabled by the development of a variety of genomic and proteomic technologies, yielding a variety of relationship types. *Saccharomyces cerevisiae*, the budding yeast, a relatively simple single-celled eukaryote, has been an important model system, and many network studies originate from this yeast. Many core functions in yeast are conserved in humans, and many network trends observed in yeast have proven true for other organisms [99,116], leading us to believe that relationships in different organisms share similar organizational principles.

Before delving into the details of biomolecular network modeling, we add a few notes of caution. First, even the most complex network models remain a vast oversimplification of the much richer and more complex reality of cellular networks. A single ‘vertex’ can actually correspond to hundreds or even millions of physical protein molecules carrying a variety of post-translation modifications, and each diffusing independently throughout the cell. As important nuances relating to spatial and temporal dynamics are often lost, the network abstraction is only relevant to the extent that it proves useful. As the famous statisti-

cian George Box put it, “all models are wrong, some are useful” [71]. Second, the method of network abstraction and the data source/experimental may introduce systematic errors. For example, the so-called ‘matrix’ method of inferring protein interactions (which assumes that proteins interact if each binds directly or indirectly to a common ‘bait’ protein) will show ‘over-clustering’ by definition. Alternatively the ‘spoke’ method, in which interactions are only inferred between a bait protein and each prey protein that it ‘pulls down’, will result in ‘under-clustering’ if only a small subset of proteins are used as baits [9,35,52]. Therefore, one should take extra caution in interpreting observed network features to separate artifacts from real biological phenomena by using appropriate controls. Last, in searching network structure for trends, one often needs to compare the actual biomolecular network with appropriate random networks to ensure that the features identified are not a trivial consequence of a simpler phenomenon. For example, in searching for network motifs, randomized networks with the same **degree distribution** (see Sect. “Single-Color Networks”) can be used as a control to avoid discovering motifs that are a trivial consequence of the underlying degree distribution [66,83,138].

## Single-Color Networks

### Definition and Categorization

The most basic biomolecular network is a single-color network, which contains only one type of biological interaction or relationship. The nature of the biological interaction or relationship defines the type of network, as it also specifies the nature of biomolecules involved. For interactions with clearly defined directionality, the corresponding network can be represented mathematically as a **directed** graph, with vertices denoting biomolecules, and edges connecting pairs of vertices whose corresponding biomolecules have directional interactions. Edge direction can represent material flow, exemplified by a **metabolite network** in which vertices are metabolites, and edges are enzyme-catalyzed reactions. As many metabolic reactions are practically irreversible under physiological conditions, we can use directed edges to represent the direction of a reaction, pointing from substrate to product. Alternatively, one can represent biochemical reactions using a directed edge to connect two enzymes such that the first enzyme’s metabolic product is the second one’s substrate. We refer to networks of the latter type as **metabolic enzyme networks**, to distinguish them from metabolite networks. Direction of interaction can also be derived from the inferred direction of information flow.

For example, in a **signaling pathway**, a directed edge points from one protein to another immediately downstream in signal transduction. Transcriptional regulation is another type of directed biomolecular interaction, and we can construct a **network of transcriptional regulation** from directed edges connecting transcription factors to the genes they regulate. In addition to transcriptional regulation, gene regulation in the cell also happens post-transcriptionally, for example through microRNAs (or small RNAs) which base pair with mRNAs and influence their stability or translation. Researchers have constructed **microRNA networks** by connecting microRNAs with their target mRNAs based on experimental identification at a limited scale [30,111,115] or based on computational predictions [94,124]. Post-translational regulation is yet another important part of the cellular regulatory process. One example of a **post-translational regulatory network** is a **protein modification network** such as a **protein phosphorylation network**, in which each directed edge connects a kinase with its protein substrate or phosphorylation target [81,92].

Single-color networks that capture symmetric relationships can be represented by **undirected** graphs, with each vertex denoting a biomolecule, and an edge connecting vertices whose corresponding biomolecules have a symmetric relationship. An important example is a **protein network of physical interactions**, in which edges connect proteins that physically associate with each other with sufficiently small dissociation constants. Such association may be via direct physical contact, or indirect via one or more intermediates.

Genetic interaction is another important type of biomolecular relationship in studying cellular functions. It describes pairs of genes for which the phenotype of the double mutant departs from expected under some null model of independence, e.g., gene pairs for which the fitness of the double mutant deviates from the product of the fitness values for each of the two single mutants. **Synthetic sick genetic interactions** (**synthetic lethality** in the extreme case), correspond to gene pairs for which double mutant fitness is lower than expected. These synthetic sick or lethal (SSL) interactions can be represented as an undirected network. SSL interactions of ten correspond to genes which have redundant or compensatory function. They are also slightly more likely to encode proteins that interact physically [122]. If the double mutant phenotype for a pair of genes is less extreme than expected, the corresponding gene pair is said to have an antagonistic (or alleviating) genetic interaction. Although we can use undirected edges for alleviating interactions, in some cases additional information can be conveyed using di-

**rected edges**. For example, if phenotypes arising from mutations in genes A and B differ and the double mutant phenotype most closely approximates gene A, we may draw a directed edge from gene A to gene B. With the aid of other prior biological knowledge, such directed edges can be useful in assigning order of action in biological pathways [8,65]. There exist many other types of genetic interaction [28,114].

Another relationship type is sequence homology (describing two genes with similar sequences that each descended from a common ancestor). Other biomolecular relationships may be derived from shared biomolecular properties, for example the relationship of **similar mRNA expression profile** [130], **similar phenotype** [39,45], or **similar gene function** [72]. Because there are many different properties of biomolecules, many ‘similar-property’ networks are possible. Above, we sampled types of biomolecular interactions or relationships used in network studies to date. However, this is not meant to be an exhaustive list, as there are many more relationship types, some unknown to us, and some for which large-scale data is not yet available. As more large-scale experimental methods are developed and applied to biological research, more biomolecular relationships can be represented as networks.

A few special types of graphs have particular relevance to the study of large-scale biomolecular networks. These include **bipartite** and **quasi-bipartite** graphs. A **bipartite** graph is a graph with two disjoint vertex sets, such that no edges connect pairs of vertices within the same set. For example, a bipartite graph might consist of a set of gene vertices and a set of protein vertices, together with edges that each correspond to a relationship described by “gene X encodes protein Y”. If one of the two vertex sets contains within-set edges, while the other does not, the graph is called **quasi-bipartite**. Certain biomolecular networks are intrinsically quasi-bipartite. For example, transcriptional regulatory networks contain edges pointing from genes encoding transcription factors to the regulated genes. Thus, vertices may be divided into transcription factors and non-transcription factors. While the latter vertex set, by definition, does not contain any within-set links, a transcription factor may regulate the expression of another transcription factor, so that the network is quasi-bipartite. Quasi-bipartite graphs can also arise as a simple consequence of experimental design. Protein interaction networks derived from yeast two-hybrid or genetic interactions derived from systematic genetic array (SGA) analysis are two examples, where only a subset of genes have been used as ‘baits’ or ‘queries’. In network analyses, care must be taken to avoid discoveries that are trivial conse-

quences of bipartite or quasi-bipartite properties arising from the experimental design.

## Tools and Methodology

Studying global structural features of a network involves many different measures of graph properties [25]. Here we survey a few frequently used measures below.

**Degree** (or **connectivity**) is the most basic characteristic of a vertex. The degree of vertex  $V$ ,  $k(V)$  is defined as the number of edges incident on vertex  $V$ , and tells us how connected  $V$  is with other vertices in the network. In the case of directed networks, there is an **outgoing degree** (or **out-degree**)  $k_{\text{out}}(V)$ , denoting the number of edges that start from  $V$ , and an **incoming degree** (or **in-degree**)  $k_{\text{in}}(V)$ , denoting the number of edges that end at  $V$ .

The **clustering coefficient**,  $C(V)$ , measures the connectivity amongst the neighbors of a vertex  $V$ , and can be defined as the number of edges among  $V$ 's neighbors, divided by the total possible number of edges if all the neighbors were connected to one another. The average clustering coefficient of a network characterizes the general tendency of vertices to form clusters [129].

In a network, **distance** between two vertices  $U$  and  $V$ ,  $d(U, V)$ , is measured as the number of edges in the shortest **path** from  $U$  to  $V$ . In an undirected graph,  $d(U, V)$  satisfies non-negativity, symmetry and the triangle inequality of a distance metric. In the case of directed graph, however, all edges in a path must be in the same direction. While non-negativity still holds, the distance  $d(U, V)$  from  $U$  to  $V$  could be different from the distance  $d(V, U)$  from  $V$  to  $U$  so that symmetry is not necessarily satisfied, nor is the triangle inequality. Hence, strictly speaking,  $d(U, V)$  is not a distance metric in directed graphs. The overall connectedness of a network can be quantified by its **characteristic path length**, defined as the average distance between pairs of vertices in the network.

A **subgraph** of a graph  $G$  is a graph whose vertex and edge sets are subsets of those of  $G$ . A **component** is a subgraph such that all vertices are connected to one another by some path, and there are no connections to vertices outside of that subgraph. If a component contains a majority of a graph's vertices it is referred to as the **giant component**. Other subgraphs of interest are **cliques**, a set of completely connected vertices. Such dense subgraphs bear particular relevance to protein complexes, which often present themselves as cliques, or dense but incompletely connected subgraphs, in a protein network of physical interaction. One can 'seed' the search for such dense subgraphs in the network (as a proxy search for protein complexes) using **k-cores** – subgraphs such

that each vertex has at least  $k$  neighbors within that subgraph [9].

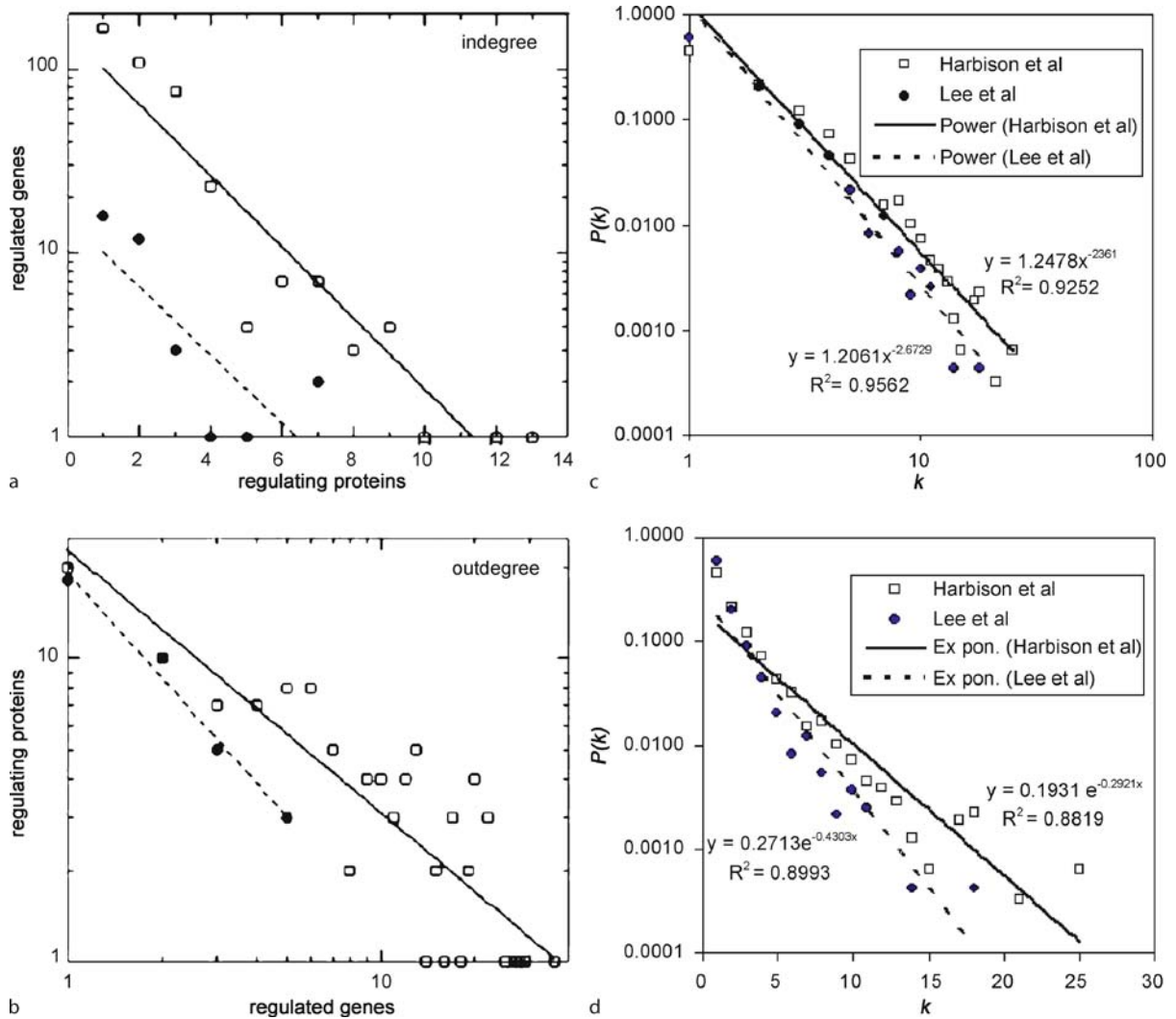
## Global Structural Features

Topological features of biomolecular networks have been subjected to extensive study (reviewed in [16]). Below we describe some features of particular relevance to network structure and function.

**Power-Law Networks** In a biomolecular network, some vertices are better connected than others. To characterize how connectivity varies from vertex to vertex, one can calculate the degree distribution – i.e. the distribution of  $P(k)$ , the fraction of vertices with degree  $k$ . The Erdős–Rényi (ER) model describes random networks in which each vertex pair is assigned an edge randomly with a probability that is fixed for all vertex pairs. In other words, each edge is assigned or not assigned according to the outcome of a 'weighted coin flip' (also known as a Bernoulli trial). For such ER networks,  $P(k)$  follows a binomial distribution, but can often be approximated well and conveniently by a Poisson distribution. The degree distributions of biomolecular networks (and many other complex networks such as food webs and social networks) tend to fall off at high degree much more slowly than a Poisson, and often follow a **power law** distribution  $P(k) \sim k^{-\lambda}$  (Fig. 2) [15]. The degree exponent  $\lambda$  in most biomolecular networks seem to have a value between 2 and 3 [15]. The most intuitive (and perhaps most useful) feature of this power law distribution is that most vertices in the network have only a few edges, while a few vertices have many edges. These highly connected vertices are called **hubs**, which are sometimes described to 'hold the network together' since removal of a hub has a particularly high tendency to break the network up into separated components. Networks with a power-law degree distribution (or sometimes networks that only follow a power-law in the high-degree tail of the distribution) are called **scale-free** networks. The term scale-free is meant to indicate that there is not a typical vertex in the network that can characterize the rest of the vertices. For some random networks, many vertices have degrees equal or close to the average degree; however, it should be noted that ER networks with a low edge probability can be easily confused with power-law networks [47]. Arguments have also been made that **scale-rich** is a more appropriate description of graphs with a power-law degree distribution [119]. For clarity, we prefer the term **power-law network**.

Networks of metabolic reactions were the first biomolecular networks to be characterized as power-





**Biomolecular Network Structure and Function, Figure 2**

Degree distribution of transcriptional regulatory networks. **a/b** incoming and outgoing degree distributions of a manually curated network; **c/d** incoming degree distribution of two large-scale transcriptional regulatory maps, fitted with power-law (**c**) and exponential (**d**)

law [64,127]. Jeong et al. analyzed the metabolite networks of 43 organisms across all three domains of life – bacteria, archaea, and eukaryotes, and discovered that despite significant variation in the components and pathways, all metabolite networks have similar topological scaling properties [64]. As many biochemical reactions are irreversible under physiological conditions, metabolite networks should be represented as directed graphs. The distribution of both in-degrees and out-degrees follow power-law distributions, i.e.  $P(k_{in}) \sim k_{in}^{-\lambda_{in}}$ ,  $P(k_{out}) \sim k_{out}^{-\lambda_{out}}$ , and the values of both  $\lambda_{in}$  and  $\lambda_{out}$  lie within a narrow band of 2.0 to 2.4 [64]. This demonstrated that a power-law degree distribution

is a common feature of cellular metabolism, with most substrates involved in one or two reactions, while a few molecules, such as coenzyme A, function as metabolic hubs by participating in several dozen reactions.

Other biomolecular networks have also been described as power-law, including networks of physical interactions between proteins, transcriptional regulatory networks, and networks of SSL genetic interactions. The protein interaction map of the yeast *Saccharomyces cerevisiae* as measured by systematic two-hybrid screens [56,125] exhibits a power-law degree distribution, as do physical interaction maps of multi-cellular organisms – the worm *C. elegans*, the fly *D. melanogaster*, and hu-

man [41,63,74,126,135]. The degree distribution of correlated expression networks based on microarray data has been shown to follow power-law [1,32]. The outgoing connections from transcription factors to target genes also exhibit power law behavior (Fig. 2b), [44,73,121]. The incoming degree distributions in these networks have been found instead to follow an exponential decay according to manually curated data (Fig. 2a), [44,121]. However, analysis of more recent large-scale data [49,73,78] suggests that power-law distribution may provide a slightly better fit (Fig. 2c,d). The resolution of this awaits further refined data and analysis, but it is clear that the distribution of in-degrees and out-degrees in a directed graph may behave differently. Similar to directed networks, the degree distribution of quasi-bipartite networks, should also be examined separately for the different vertex sets.

Closer examination of several biomolecular networks showed that their degree distributions exhibit better fit to a truncated power law than a power-law model [6,99]. A truncated power-law distribution follows a power-law distribution with an ‘exponential envelope’, such that power-law behavior dominates the distribution at low degrees, while the exponential function dominates the distribution at the high degree tail of the distribution. Truncated power-law behavior may arise from constraints on the number of interaction partners for high degree vertices [6]. However, biological implications of the differences between these degree distributions remain to be demonstrated in most cases. It is nevertheless clear that biomolecular networks share the common feature that a few vertices have many edges, while most vertices have only a few edges.

The ubiquity of power-law and presence of hubs in biomolecular networks prompts us to ask what advantage such network architecture might offer. Some argue that the answer lies in robustness. Intuitively, random removal of a considerable number of vertices could disintegrate a network, breaking it down into small, disconnected groups of vertices. This is true for Erdős–Rényi random networks [16,117]. Power-law networks, such as the electrical power grid, are extremely resilient to accidental vertex failures. Even if 80% of the vertices are randomly removed, the rest often still form a connected component. This is confirmed by both numerical and theoretical analyses examining how removing vertices of varying degrees impacts the average path length and size of the giant component [4,20,21,23]. As most vertices in power-law networks have small degrees, any vertex that fails is most likely one of them and hence the damage will be negligible. However, such robustness is accompanied by the weakness that deliberate attacks on hubs are much more

likely to create disconnected components and hence particularly devastating. Some believe this notion may apply beyond power grids to biological networks, and we discuss this possibility further in the Applications section.

A related question is the evolutionary origin of power-law networks. Building upon Simon’s early work [17,104], Barabasi and Jeong proposed a growth and preferential attachment model [15]. Under this model, the network must emerge from gradual addition of vertices and edges. Preferential attachment specifies that the newly added vertices connect to the rest of the network by preferentially attaching themselves to highly connected vertices (e.g., with a probability proportional to the degree of the target vertex). This rich-get-richer mechanism offers an explanation for hubs. In protein networks, growth and preferential attachment may result from gene duplication, which produces initially-identical proteins with the same interaction partners. The more interactions possessed by a protein, the more likely it is to acquire a new interaction through duplication of an interaction partner, resulting in the rich-get-richer phenomenon associated with preferential attachment.

### Neighborhood Clustering Neighborhood clustering

describes a network characteristic that resembles the phenomenon of neighbors of my neighbors tend to be neighbors of one another. Above we defined a clustering coefficient to quantify this local property of a graph. The average clustering coefficient of all vertices in the graph characterizes the global tendency towards **clustering** in the network [129]. Several biomolecular networks, for example, the metabolic network [96] and physical interaction networks, have a significantly higher average clustering coefficient than random networks of equivalent size and degree distribution, suggesting modular organization in addition to power-law degree distribution. A partially-determined human interactome network, however, was found to be less clustered than random networks with the same degree distribution [99]. This could be because only a subset of edges in the complete network was sampled, or it might suggest ‘clustering’ of protein-protein interaction networks in other organisms have been artificially boosted by previous experimental methods [99]. It is also possible that the human interactome network indeed is less clustered than other organisms. Further experimentation is required to resolve this apparent discrepancy.

**Small World** Another important feature of biomolecular networks is that they are frequently small-world. A small-world network is defined by two properties: 1) short characteristic path length, meaning that the length

of the shortest path between any pair of vertices tends to be very small (this is sometimes confusingly called “the small world property,” despite there being two properties of a small-world network); 2) high clustering coefficient, i.e. densely connected local neighborhoods as described above [129]. Random networks following the ER model satisfy the first requirement, but not the second, whereas some regular lattices possess the second property, but not the first. Both metabolite networks and physical interaction networks have been shown to be small-world [33,110,127]. Short characteristic path length in small world communication networks allows efficient signaling between vertices, in that local perturbations can propagate rapidly to the entire network. Small-worldness seems to be evolutionarily conserved, as the metabolite network of a parasitic bacterium exhibits roughly the same average path length as observed for a large multi-cellular organism [64,127].

**Disassortativity** In some networks, highly connected vertices tend to link to vertices with few edges, avoiding connections to other hubs. This is the disassortative property. Disassortativity is shared by both metabolic networks and protein-protein interaction networks, as well as technological networks such as the world wide web, but not in social networks which are assortative [80,87,91]. The avoidance of hub-hub connections may come from the compartmentalization of many cellular processes [50], although other explanations have been proposed. For example, disassortativity may provide some protection against deliberate attacks to the hub vertices. As the robustness of power-law networks comes with an increased vulnerability to deliberate attacks against hubs, suppression of hub-hub connection has been suggested to alleviate or slow down propagation of such deleterious perturbations [80]. While it is difficult to know how applicable the communication network analogy is to biological networks, these ideas are certainly thought-provoking.

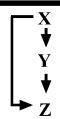

**Network Motifs** The average clustering coefficient discussed above can be expressed in terms of the number of triangles in a network relative to three times the number of paths of length two [88]. It is then interesting to ask whether there are other local patterns of connectivity that are over-represented within biological networks. Such patterns are called **network motifs**. Demonstrating enrichment of a given connectivity pattern requires counting the number of subgraphs in the network matching the pattern and comparing the count to that observed in many randomized versions of the same network [83].

Milo et al. initially proposed that network motifs represent elementary building blocks of cellular networks. The transcription regulatory network in *E. coli* and *S. cerevisiae* both exhibit enrichment for the feed-forward loop and the bi-fan (loop Fig. 3), which can be attributed to specific tasks in information processing [83,103]. In particular, the feed-forward motif (Fig. 3) involves two transcription factors, one regulating the other and both controlling the expression of a third target gene. The feed-forward motif can serve as an AND-gate that corresponds to a low-pass filter which suppresses out transient signals while preserving persistent signals, as is the case in the *ara*-BAN operon in the arabinose feed-forward loop [101].

Different types of networks exhibit distinct sets of motifs. For example, the food web exhibits a ‘three-chain’ motif and a bi-parallel motif, in contrast to the feed-forward motif and bi-fan motif in transcriptional regulatory networks (Fig. 3), [83]. Thus network motifs provide insight into common mechanisms of interaction within a network. The abundance profile of network motifs for a given network serves as a fingerprint for classifying networks into different superfamilies [82].

A given network motif may have many instances within a network. The relationship of network motif instances to one another poses an interesting question. Observations made in the *E. coli* transcriptional regulatory network show that instances of both the feed-forward motif and the bi-fan motif do not exist in isolation. Instead, they tend to overlap with one another and form large clusters, which then coalesce into superclusters [26]. This suggests hierarchical organization of biomolecular networks, and that motifs should not necessarily be viewed as elementary building blocks of complex networks, but might instead be thought of as signatures of higher order structures corresponding to distinct cellular processes. The fact that many motif clusters overlap with known biological functions [26] interference between motif instances. We further discuss this notion below.

**Modularity and Hierarchy** Clustering of network motifs suggested presence of higher order structures in biomolecular networks [26]. The search for enrichment of larger patterns (involving more than four vertices) in large networks is computationally challenging due to the exponential computational complexity associated with enumerating increasingly large subgraphs in biomolecular networks. An alternative approach to brute force computation or sampling methods is to identify densely connected clusters, or network modules. Various clustering methods have been used to characterize modules using network topology [42,53,96,97,102,109,113]. Note that the

Network	Nodes	Edges	$N_{\text{real}}$	$N_{\text{rand}} \pm \text{SD}$	Z score	$N_{\text{real}}$	$N_{\text{rand}} \pm \text{SD}$	Z score
<b>Gene regulation (transcription)</b>			 <b>Feed-forward loop</b>			 <b>Bi-fan</b>		
<i>E. coli</i>	424	519	40	$7 \pm 3$	10	203	$47 \pm 12$	13
<i>S. cerevisiae</i> *	685	1,052	70	$11 \pm 4$	14	1812	$300 \pm 40$	41

Biomolecular Network Structure and Function, Figure 3

## Single-color network motifs

ideal module is a collection of gene products that work together as a unit, which may plug in to perform its molecular function in different contexts. Densely connected clusters may be useful in identifying modules, but we should separate the ideal concept from its surrogate measure [50].

The concept of relatively isolated modules may seem contradictory to the power-law characteristic of biomolecular networks. Both naturally exist [95,96], however, indicating network modules do not exist in isolation, but combine at various levels to form a hierarchical network. The hierarchical structure of biomolecular networks poses a challenge in identifying modules, as there are no clear boundaries between them. However, understanding networks' hierarchical organization can also aid the module identification, at least in the transcriptional regulation network. For example, the transcriptional regulatory networks of *E. coli* and *S. cerevisiae* both exhibit a multi-layer hierarchical structure [77,136], and removing global regulators in the top layers of the hierarchy decomposes the network into functionally relevant modules [77]. Recent work also identified 'origins', or 'organizer' structures within the transcriptional regulatory network that originate at a distinct class of sensor transcription factors and serve as functional modules for various environmental perturbations [13,31]. Combining topological information with other genomic or proteomic data can lead to more biologically relevant modules [9,14,55,62,65,118,123]. We describe this further in the next section.

## Applications

Structural features of biomolecular networks can be used to infer organizing principles of biomolecular interactions and relationships. The power-law nature suggests robustness of biomolecular networks against random attacks, and models of network designation evolution – e.g. growth and preferential attachment – serve as additional support to the gene duplication model of protein network growth. The small-world designation suggests the potential for rapid communication within biomolecular networks, with local perturbations reaching the entire net-

work via a small number of links, while also reflecting clustering due to modularity, e.g. as a result of protein complexes. Network motifs are local interconnection patterns that are often footprints of higher order structure (i.e., functional units composed of more vertices than are considered in network motifs). Different types of biomolecular network have distinctive profiles of the relative abundance of network motifs, so that network motifs can also be used to characterize and compare complex networks.

More practically, structural characteristics of biomolecular networks can be used to make biological inferences. Experimentally derived networks are inevitably susceptible to random as well as systematic errors, and characteristics of network structure can be used to assess confidence of edges in biomolecular networks and also make novel predictions. Goldberg et al. examined a list of *S. cerevisiae* protein-protein interactions and were able to stratify this list into many subsets of varying accuracy by exploiting the neighborhood clustering feature of small-world networks [43]. Other methods using network topology also proved useful in predicting protein interaction networks of other organisms [41], for example, by exploiting conservation of network motifs [2].

High clustering in biomolecular networks often represents physical organization of biomolecules, particularly in the case of physical protein interaction maps. Therefore, identifying densely connected subgraphs (clusters) can help detect protein complexes [10,66,97,113]. In addition, topological modules are believed to represent specific biological processes or functions [9,14,55,59,118,123]. Hence, identifying modules can also help assign biological processes or functions to genes in a network.

Features of vertices can also be inferred from network topology. As discussed above, power-law networks are robust with regard to random vertex failures, but highly sensitive to targeted attacks against hubs. Whether this notion of robustness applies beyond electrical power grids to biological networks is subject to debate. In protein networks that include co-complexed relationships, for example, many proteins are hubs by virtue of their membership in protein complexes (appearing as cliques). Re-

removal of such hubs is not highly disruptive in terms of generating disconnected components, and yet there are many essential complexes for which the removal of a single protein member disrupts the physical assembly of the whole complex and results in death for the organism. Nevertheless, genome-wide deletion analysis in *S. cerevisiae* demonstrated a correlation between degree and essentiality: only ~10% of the genes encoding proteins with degrees less than 5 are essential, as compared to ~60% for genes encoding proteins with 15 or more interactions [39,63,131]. Thus, protein connectivity can to some extent predict the phenotype of its corresponding gene. Recent work on metabolic networks also demonstrated that certain topology-based vertex characteristics can accurately predict knock-out phenotypes of the corresponding enzyme [133].

In the case of SSL genetic interaction networks, highly connected genes are often more critical to fitness than genes of less connectivity as they ‘buffer’ a large number of genes to maintain the cell’s survival [122]. Upon disruption of a hub gene in the SSL network, mutations in an increased number of genes will result in fitness defect. It has been hypothesized that hub genes could be useful targets for anti-cancer drugs to specifically kill cancer cells as they often carry many additional mutations, while normal cells can be protected due to genetic buffering by the hub gene [122].

Examples of structural features and applications discussed thus far have focused on networks composed of a single type of relationship. However, as there are many biological interaction or relationship types, it is worthwhile to combine information from various biomolecular networks. In the next section we discuss multi-color networks, that represent more than one biological interaction/relationship simultaneously.

## Multi-Color Networks

### Definition and Categorization

Previously we described various types of single-color networks, each containing edges that correspond to a single type of relationship between biomolecules. Different types of interactions, however, do not exist independently. Instead, they are inter-related with one another and offer different perspectives on the underlying biological mechanisms. It is therefore useful to consider multiple biomolecular relationships simultaneously by integrating them into one **multi-color network** (Fig. 4). Each edge type can be represented by a color that indicates the type of biological interaction or relationship. Here we focus on multi-color networks where each vertex corresponds to a gene and its

corresponding protein(s). As a multi-color network can be viewed as an overlay of multiple single-color networks, the structural features of single-color networks discussed previously still hold true for the individual component single-color networks, but additional properties of multi-color networks emerge.

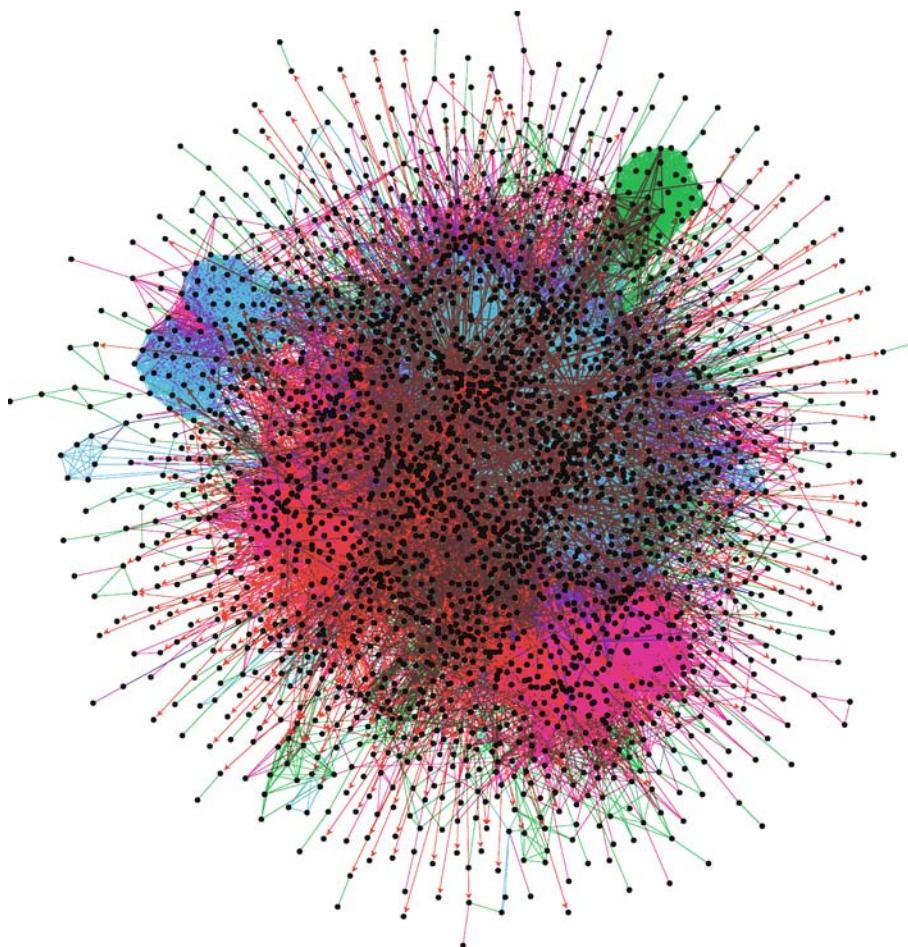
### Features and Implications

**Correlation of Edge ‘Colors’** The first step in studying multi-color networks is to examine how various edge types correlate to one another. Various statistical methods can be used to test significance of correlation. Without considering network topology, one can treat pairs of vertices in the network as independent data points, and assess significance of overlap between two edge types analytically under the null hypothesis of the hypergeometric distribution. Network topology, however, can have significant impact on edge correlation. For example, as most single-color biomolecular networks contain highly-connected hubs, if many of the hubs in one network are hubs in another network, the overlay of the two single-color networks will, as a result, have considerable overlap of the two interaction type. To test edge correlation more rigorously in the context of network topology, we can generate randomized versions of the multi-color network by rewiring edges in each component single-color network independently, while preserving certain topological features, for example, degree distribution [134,138]. We can then determine the significance of correlation by calculating the fraction of random networks that contain more two-color edges than the real network. This approach can also be generalized to assess significance of overlap of more than two edge colors.

Several studies have noted the correlation between different types of biological interactions. For example, interacting proteins are more likely to have similar expression pattern [36,60]; genes with correlated expression are more likely to be controlled by a common transcription factor [137]; and SSL genetic interactions are more likely to occur between homologous genes [122]. Although each of these correlations are statistically significant, no single pair of interaction types shows perfect overlap (and indeed the level of correlation is low in each example). Thus there is dependence amongst many of the observed types of biomolecular interactions, and yet each type offers a distinct view of the same complex biological system.

**‘Multi-Color’ Network Motifs and Themes** The network motif concept, described previously for single-color networks, is extensible to a multi-color network. Multi-color network motifs characterize relationships between

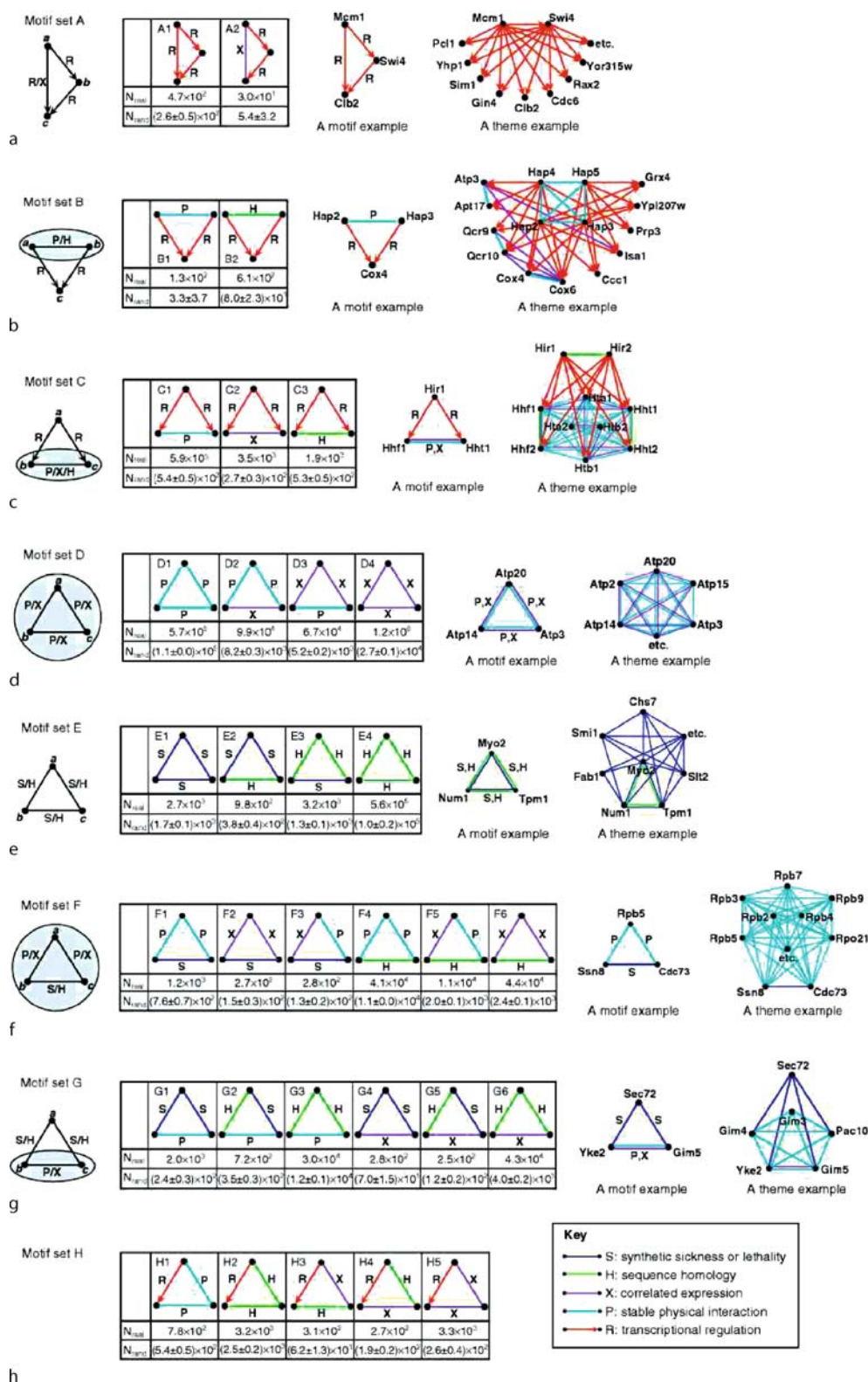




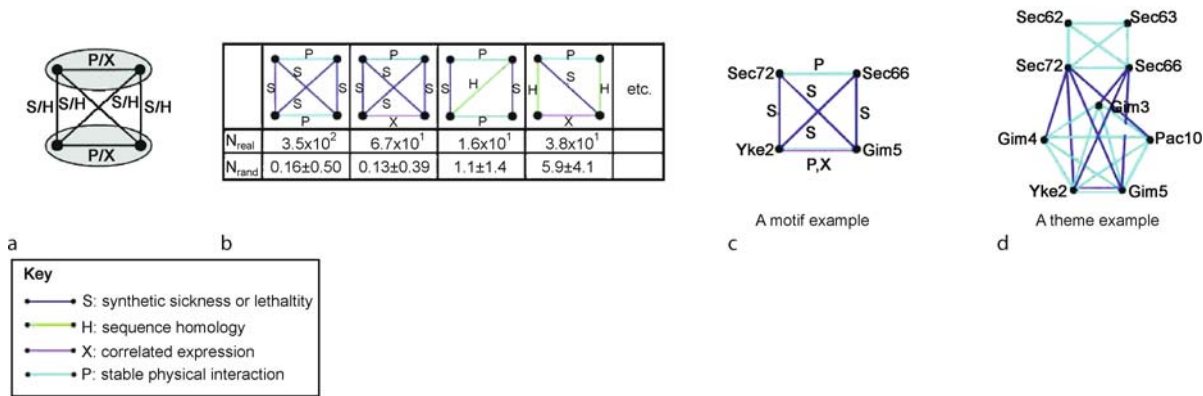
**Biomolecular Network Structure and Function, Figure 4**  
**A multi-color network**

different biological interaction types within local network neighborhoods. Yeager-Lotem et al. examined network motifs in combined cellular networks of two interaction types – transcriptional regulation and protein interaction [134]. They identified three multi-color motifs of three genes – two interacting transcription factors that co-regulate a third gene, two interacting proteins co-regulated by the same transcription factor, and a ‘mixed feedback loop’, where a given transcription factor and the product of its target gene form a protein interaction with each other. More multi-color motifs were characterized in an integrated network of five interaction types – transcription regulation, physical interaction, SSL genetic interaction, sequence homology, and correlated mRNA expression Fig. 5 [138]. These motifs can be explained in terms of high-order network structures, or network themes (Fig. 5), which are representative of the underlying biological phenomena. One interesting ex-

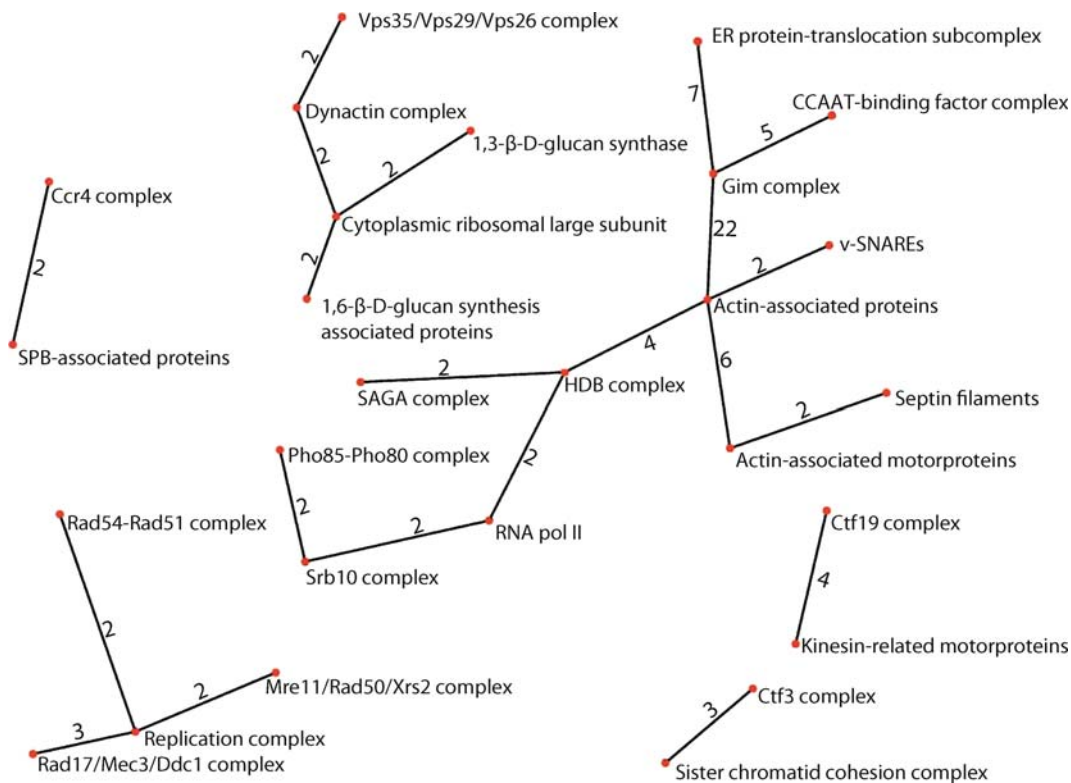
ample is motifs containing two vertices linked by protein interaction or correlated expression, such that both vertices connect to a third vertex either by SSL interaction, or by sequence homology. This suggests the theme that a protein and a distinct protein complex/biological process have compensatory functions, resulting in SSL relationships between the protein and any component essential to the compensatory complex/process. Many four-vertex motifs that correspond to the compensatory complexes theme (Fig. 5) further support this notion. These four-vertex motifs also suggested that some instances of the ‘compensatory protein-complex/process’ theme could simply be signatures of compensatory pairs of complexes or processes, rather than a single protein acting alone to compensate for loss of another process. An excellent example of the compensatory complexes theme is the pairing of endoplasmic reticulum (ER) protein-translocation subcomplex [48] and the Gim complex [37], with each



Biomolecular Network Structure and Function, Figure 5  
Multi-color network motifs and themes



Biomolecular Network Structure and Function, Figure 5 continued



Biomolecular Network Structure and Function, Figure 6 Network themes

complex connected to the other by many SSL interactions [122,138].

**Thematic Maps** In the previous section, we described the idea that network motifs may in general be footprints of larger network structures. Organization of some

multi-color network motifs into larger structures of network themes further supported this notion, and also confirmed that the concept of modularity discovered in single-color networks is extensible to networks of multiple interaction types. As such modules in a multi-color network can exist in various sizes, and exhibit a combina-

torial number of different connection patterns, systematically identifying them by prevalence of network motifs of a fixed size can be impracticable. Network theme representations, in which a simplified network is drawn in terms of specific network themes [138], is a convenient way of examining the system. In the example of the compensatory complex theme, we can collapse members of the same complex into one single vertex, and draw links between complexes with significantly enriched inter-complex SSL interactions [138]. The resulting thematic map of compensatory complexes (Fig. 6) serves as a guide to the ‘redundant systems’ within the integrated *S. cerevisiae* network [138]. Similarly, one can create maps of ‘regulonic complexes’ – complexes for which the corresponding genes are co-regulated – by linking each transcription factor to protein complexes that are significantly enriched for genes targeted by that transcription factor, providing a schematic view of the transcriptional circuitry in the cell [105,106,138].

### Applications

Multi-color networks integrate several biomolecular interaction types and contain information that is richer than the component single-color networks taken individually. Hence they can be used to make more accurate biological inference. Correlation between biological interaction types alone can be useful, especially when a large number of interaction types are combined together. A Bayesian approach achieved considerable success in predicting physical interactions between proteins [59]. To capture the additional information from interdependency of interaction types, one can examine the conditional probability of a physical interaction given all possible ‘multi-color relationships’ (each corresponding to a particular configuration of single-color relationships) [59,83]. However, such an approach scales poorly as the number of interaction types grows (the number of multicolor relationships grows as  $2^N$ , where  $N$  is the number of interaction types, so that there will be increasingly many multicolor relationships with no training examples). More sophisticated methods, such as logistic regression [12,41], decision trees, or random forests, can model the interdependency between interaction types in a more scalable manner, allowing the combination of a large number of biomolecular relationships. Studies using such methods can significantly improve the quality of predictions [59,61,93,132,139].

Multi-color network motifs put correlations of interaction types in the context of local topology, and provide potential to further improve biological inference. Specifically, one can predict a certain type of link between a given

pair of vertices if its addition completes a structure matching an enriched network motif. For example, two genes with a common SSL interaction partner may have an increased probability of protein interaction, as addition of a physical interaction edge results in a match of a motif depicted in Fig. 5 (G1). Similarly, a SSL edge between two genes can complete a match to the same motif, if the two genes are indirectly connected by physical interaction followed by SSL. This ‘two-hop’ relationship of physical interaction and SSL proves to be a strong predictor of SSL interactions [132].

More broadly, we can also predict an interaction if its addition fits into a recurring network theme. For instance, there are many more SSL interactions between the ER-translocation complex and the Gim complex than would be expected by chance given the out-degrees of both complexes in the SSL network. However, no SSL interaction has been observed between Sec62 or Sec63, two members of the ER protein-translocation subcomplex, and any protein in the Gim complex because Sec62 and Sec63 were not used as queries in the SGA analysis [122]. We can therefore hypothesize that Sec62 or Sec63 are SSL with many members of the Gim complex [138].

Network themes represent modular organization of the network at the functional level, and can be used to predict the function or biological process in which a gene is involved. The rationale is that if a gene is part of a module where a particular function is enriched, then there is increased probability that this gene also holds the same function. For example, in the feed-forward theme depicted in Fig. 5, most of the genes regulated by both Mcm1 and Swi4 are involved in the control or execution of the cell cycle. Therefore, we can hypothesize that Yor315w, a protein about which little is known, is also involved in cell cycle regulation. Indeed, this protein is a target of cyclin-dependent kinase Cdk1, a primary regulator of the cell cycle. Predictions based on network themes may be robust to errors in the input data, since they depend on connectivity patterns in extended network neighborhoods instead of one or very few links.

It is worth emphasizing that predictions based on network motifs and themes are more suggestive than definitive. Above we focused on organizational principles that can be utilized for biological inference. Such inferences should be combined with prior information from as many sources as possible, and should be tested experimentally. In the next sections, we will touch upon rich network models that can be used for graphical data integration. We also refer readers to the chapter [► Biological Data Integration and Model Building](#) in this book for more general approaches to data integration.



## Rich Network Models

### Networks of Labeled Vertices and Edges

Just as the relationships between biomolecules come in many colors, the biomolecules themselves (the vertices) also exhibit a range of properties. These include intracellular localization, mRNA expression level, cellular function, and protein complex membership, to name just a few. A network labeled vertices and edges captures multiple properties of biomolecules as well as relationships among them. Extending the concept of multi-color networks, one can build a **rich network model** using a network with vector-labeled vertices and edges. Each vertex or edge can be labeled with a vector, such that each variable in the vector represents some vertex property or edge type. The rich information captured by such a model can be used to infer types of interactions or characteristics of vertices [38].

**Hypergraphs** Previously we mentioned that sharing of the same or similar vertex property can be viewed as a type of biomolecular relationship, and hence represented using an edge. A **hypergraph** can be used to more explicitly represent a shared property of multiple vertices. A hypergraph is a generalization of a graph where edges can connect any number of vertices, instead of only two as in conventional graphs, and a hypergraph edge (or ‘hyperedge’) can be used to connect all vertices sharing the same property. When further generalized to contain multiple types of edges, hypergraphs provide an alternative to networks of labeled vertices and edges described earlier. Hypergraphs have been used, for example, to model single-color networks of protein domains [34]. However, it remains to be seen whether this provides a more useful abstraction than a rich network representation in which vertices and edges are each labeled with vectors.

**Weighted Graphs and Probabilistic Networks** Network models discussed thus far have mostly been discrete. For example, an edge in each single-color network either exists or does not exist. In the rich network models we discussed, variables used to describe a vertex or edge characteristic can be binary or categorical variables. However, some properties of a vertex may have continuous values, for example, a gene product may be described by its corresponding mRNA expression level. An edge property may also be continuous, e. g., affinity of a physical interaction or coefficient of correlation between mRNA abundance profiles. Such quantitative information can also be incorporated in a rich network model, simply by including continuously-valued variables within the vertex or edge vectors.

One special case of a weighted graph is a probabilistic network, in which the weight on each edge represents a posterior probability that the edge is real, given the evidence supporting that edge. Probabilistic networks are particularly useful when knowledge of relationships is uncertain. This is often the case for biological data, sometimes due to inaccuracy of measurement or varying levels of confidence in a computationally inferred relationship [7,139]. Probabilistic networks have been used to predict novel members of partially known protein complexes, for example, by sampling from networks of physical interaction from multiple data sources based on their corresponding confidence levels [7,11]. Algorithms have also been developed to identify protein complexes by finding dense subgraphs within a weighted graph [75].

### Future Directions

Research on the structure and function of biomolecular networks has covered a lot of ground in the past decade. So far, we have studied global topological features such as power-law degree distributions, small-worldness, disassortativity and associated biological implications such as robustness and dense neighborhood clustering. We have also discussed the hierarchical organization of networks, with network ‘structures’ ranging from network motifs to dense clusters, to pairs of clusters, densely connected within each cluster by one ‘color’ of interaction, and connected to one another by interactions of a different color. We have also explored the roles of these structures in cellular functions; for example, the involvement of certain network motifs in information processing and the correspondence of dense clusters to protein complexes and biological processes. We also recognized the interplay of various interaction types, for example, the correspondence of pairs of clusters connected by SSL genetic interaction to protein complexes with mutually compensatory functions. Deeper understandings in the above areas certainly warrant future study. However, many currently under-explored territories remain.

The cell is a constantly changing environment, so are the molecules and processes within that environment. Network studies have been mostly limited to the union of interactions in all locations, times and environments. Cellular state, such as position in the cell cycle, is a key determinant of concentrations and status of biomolecules, and hence the interactions between them. Studies have begun to reveal the dynamic nature of biomolecular networks. For example, by overlaying mRNA expression patterns with the physical interaction network, Han et al. described two types of highly connected hubs – ‘party’



hubs, which interact with most partners simultaneously; and ‘date’ hubs, which connect to different partners at different time [46]. Expression data have also been combined with protein interaction data to reveal cell cycle-dependent protein complexes or arrangements of protein complexes [24]. Genome-scale molecular data across a wide variety of cellular states are so far only available for mRNA expression levels [22,112]. Availability of more biomolecular interaction data under various cellular conditions will allow us to study the additional complexity of cellular processes. Such research will require adding another dimension when modeling biomolecular networks, and will likely also require new mathematical tools.

In addition to overall cellular state, intracellular environment also bears significance on cellular interactions. For example, proteins may be restricted to particular compartments, either defined by a membrane as are the nuclei or mitochondria, or in less-clearly delimited regions such as nucleoli. Such features can be viewed as vertex properties in the network, and can therefore be modeled either using a hypergraph or by labeling vertices. Cellular localization data are already available at a genomic scale [54,70], and should yield additional insights.

One major limitation of network analysis is data collection. Even with the rapid development of high-throughput genomic and proteomic experimentation, we are still far from being able to systematically characterize all biomolecular features and interactions. Quantifying them at high resolution in both space (across subcellular locations and across cell and tissue types) and time (e.g., through the course of organismal development) appears a more distant goal. More comprehensive network data will undoubtedly advance the study of the structure and function of biomolecular networks, and the knowledge can in turn aid experimental discovery. One intuitive example is to use inferences made from network study to apply limited resources to experiments that are more likely to yield discoveries. We are optimistic that experimental discovery and network analysis can form a positive-feedback loop, with each reinforcing the other, and together advancing our understanding of biomolecular networks.

## Bibliography

### Primary Literature

1. Agrawal H (2002) Extreme self-organization in networks constructed from gene expression data. *Phys Rev Lett* 89:268702–268702
2. Albert I, Albert R (2004) Conserved network motifs allow protein-protein interaction prediction. *Bioinformatics* 20(18):3346–3352
3. Albert R, Barabasi AL (2002) Statistical mechanics of complex networks. *Rev Mod Phys* 74(47)
4. Albert R, Jeong H et al (2000) Error and attack tolerance of complex networks. *Nature* 406:378–382
5. Alon U (2003) Biological networks: the tinkerer as an engineer. *Science* 301:1866–1867
6. Amaral LA, Scala A et al (2000) Classes of small-world networks. *Proc Natl Acad Sci USA* 97(21):11149–52
7. Asthana S, King OD et al (2004) Predicting protein complex membership using probabilistic network reliability. *Genome Res* 14(6):1170–5
8. Avery L, Wasserman S (1992) Ordering gene function: the interpretation of epistasis in regulatory hierarchies. *Trends Genet* 8(9):312–6
9. Bader GD, Hogue CW (2002) Analyzing yeast protein-protein interaction data obtained from different sources. *Nat Biotechnol* 20(10):991–7
10. Bader GD, Hogue CW (2003) An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 4(1):2
11. Bader JS (2003) Greedily building protein networks with confidence. *Bioinformatics* 19(15):1869–74
12. Bader JS, Chaudhuri A et al (2004) Gaining confidence in high-throughput protein interaction networks. *Nat Biotechnol* 22(1):78–85
13. Balazsi G, Barabasi AL et al (2005) Topological units of environmental signal processing in the transcriptional regulatory network of *Escherichia coli*. *Proc Natl Acad Sci USA* 102(22):7841–6
14. Bar-Joseph Z (2003) Computational discovery of gene modules and regulatory networks. *Nature Biotechnol* 21:1337–1342
15. Barabasi AL, Albert R (1999) Emergence of scaling in random networks. *Science* 286(5439):509–12
16. Barabasi AL, Oltvai ZN (2004) Network biology: understanding the cell’s functional organization. *Nat Rev Genet* 5(2):101–13
17. Bornholdt S, Ebel H (2001) World Wide Web scaling exponent from Simon’s 1955 model. *Phys Rev E* 64(3):35104
18. Bornholdt S, Schuster HG (2003) Handbook of Graphs and Networks: from the Genome to the Internet
19. Bray D (2003) Molecular networks: the top-down view. *Science* 301:1864–1865
20. Broder A (2000) Graph structure in the web. *Comput Netw* 33:309–320
21. Callaway DS, Newman MEJ et al (2000) Network robustness and fragility: percolation on random graphs. *Phys Rev Lett* 85:5468–5471
22. Cho RJ, Campbell MJ et al (1998) A genome-wide transcriptional analysis of the mitotic cell cycle. *Molecular Cell* 2(1):65–73
23. Cohen R, Erez K et al (2000) Resilience of the Internet to random breakdowns. *Phys Rev Lett* 85:4626–4628
24. de Lichtenberg U, Jensen LJ et al (2005) Dynamic complex formation during the yeast cell cycle. *Science* 307(5710):724–7
25. Diestel R (2005) Graph Theory, 3rd edn. Springer, Heidelberg
26. Dobrin R, Beg QK et al (2004) Aggregation of topological motifs in the *Escherichia coli* transcriptional regulatory network. *BMC Bioinformatics* 5(1):10

27. Dorogovtsev SN, Mendes JF (2003). Evolution of Networks: from Biological Nets to the Internet and WWW. Oxford University Press
28. Drees BL, Thorsson V et al (2005) Derivation of genetic interaction networks from quantitative phenotype data. *Genome Biol* 6(4):R38
29. Fanning AS, Anderson JM (1996) Protein-protein interactions: PDZ domain networks. *Curr Biol* 6(11):1385–8
30. Farh KK, Grimson A et al (2005) The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. *Science* 310(5755):1817–21
31. Farkas IJ, Wu C et al (2006) Topological basis of signal integration in the transcriptional-regulatory network of the yeast, *Saccharomyces cerevisiae*. *BMC Bioinformatics* 7:478
32. Featherstone DE, Broadie K (2002) Wrestling with pleiotropy: genomic and topological analysis of the yeast gene expression network. *Bioessays* 24:267–274
33. Fell DA, Wagner A (2000) The small world of metabolism. *Nat Biotechnol* 18(11):1121–2
34. Freudenberg J, Zimmer R et al (2002) A hypergraph-based method for unification of existing protein structure- and sequence-families. *In Silico Biol* 2(3):339–49
35. Gavin AC, Bosche M et al (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415(6868):141–7
36. Ge H, Liu Z et al (2001) Correlation between transcriptome and interactome mapping data from *Saccharomyces cerevisiae*. *Nat Genet* 29(4):482–6
37. Geissler S, Siegers K et al (1998) A novel protein complex promoting formation of functional alpha- and gamma-tubulin. *Embo J* 17(4):952–66
38. Getoor L, Rhee JT et al (2004) Understanding tuberculosis epidemiology using structured statistical models. *Artif Intell Med* 30(3):233–56
39. Giaever G (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418:387–391
40. Gietz RD, B Triggs-Raine et al (1997) Identification of proteins that interact with a protein of interest: applications of the yeast two-hybrid system. *Mol Cell Biochem* 172(1–2):67–79
41. Giot L (2003) A protein interaction map of *Drosophila melanogaster*. *Science* 302:1727–1736
42. Girvan M, Newman ME (2002) J Community structure in social and biological networks. *Proc Natl Acad Sci USA* 99:7821–7826
43. Goldberg DS, Roth FP (2003) Assessing experimentally derived interactions in a small world. *Proc Natl Acad Sci USA* 3:3
44. Guelzim N, Bottani S et al (2002) Topological and causal structure of the yeast transcriptional regulatory network. *Nat Genet* 31(1):60–3
45. Gunsalus KC, Ge H et al (2005) Predictive models of molecular machines involved in *Caenorhabditis elegans* early embryogenesis. *Nature* 436(7052):861–5
46. Han JD, Bertin N et al (2004) Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* 430(6995):88–93
47. Han JD, Dupuy D et al (2005) Effect of sampling on topology predictions of protein-protein interaction networks. *Nat Biotechnol* 23(7):839–44
48. Hanein D, Matlack KE et al (1996) Oligomeric rings of the Sec61p complex induced by ligands required for protein translocation. *Cell* 87(4):721–32
49. Harbison CT, Gordon DB et al (2004) Transcriptional regulatory code of a eukaryotic genome. *Nature* 431(7004):99–104
50. Hartwell LH, Hopfield JJ et al (1999) From molecular to modular cell biology. *Nature* 402:C47–C52
51. Hasty J, McMillen D et al (2002) Engineered gene circuits. *Nature* 420:224–230
52. Ho Y, Gruhler A et al (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* 415(6868):180–3
53. Holme P, Huss M et al (2003) Subnetwork hierarchies of biochemical pathways. *Bioinformatics* 19:532–538
54. Huh WK, Falvo JV et al (2003) Global analysis of protein localization in budding yeast. *Nature* 425(6959):686–91
55. Ihmels J (2002) Revealing modular organization in the yeast transcriptional network. *Nat Genet* 31:370–377
56. Ito T (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci USA* 98:4569–4574
57. Ito T, Tashiro K et al (2000) Toward a protein-protein interaction map of the budding yeast: A comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins. *Proc Natl Acad Sci USA* 97(3):1143–7
58. Jacob F, Monod J (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* 3:318–56
59. Jansen R (2003) A Bayesian networks approach for predicting protein-protein interactions from genomic data. *Science* 302:449–453
60. Jansen R, Greenbaum D et al (2002) Relating whole-genome expression data with protein-protein interactions. *Genome Res* 12(1): 37–46
61. Jansen R, Lan N et al (2002) Integration of genomic datasets to predict protein complexes in yeast. *J Struct Funct Genomics* 2:71–81
62. Jansen R, Yu H et al (2003) A Bayesian networks approach for predicting protein-protein interactions from genomic data. *Science* 302(5644):449–53
63. Jeong H, Mason SP et al (2001) Lethality and centrality in protein networks. *Nature* 411(6833):41–2
64. Jeong H, Tombor B et al (2000) The large-scale organization of metabolic networks. *Nature* 407:651–654
65. Juven P, Demsar J et al (2005) GenePath: from mutations to genetic networks and back. *Nucleic Acids Res* 33(Web Server issue):W749–52
66. King OD (2004) Comment on Subgraphs in random networks. *Phys Rev E Stat Nonlin Soft Matter Phys* 70(5 Pt 2):058101. author reply 058102
67. Kitano H (2002) Computational systems biology. *Nature* 420:206–210
68. Koonin EV, Wolf YI et al (2002) The structure of the protein universe and genome evolution. *Nature* 420:218–223
69. Krogan NJ, Cagney G et al (2006) Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* 440(7084):637–43
70. Kumar A, Agarwal S et al (2002) Subcellular localization of the yeast proteome. *Genes Dev* 16(6):707–19
71. Launer RL, Wilkinson GN (1979) Robustness in statistics. Academic Press, New York
72. Lee I, Date SV et al (2004) A probabilistic functional network of yeast genes. *Science* 306(5701):1555–8
73. Lee TI, Rinaldi NJ et al (2002) Transcriptional regulatory net-

- works in *Saccharomyces cerevisiae*. *Science* 298(5594):799–804
74. Li S (2004) A map of the interactome network of the metazoan, *C. elegans*. *Science* 303:590–593
  75. Li W, Liu Y et al (2007) Dynamical systems for discovering protein complexes and functional modules from biological networks. *IEEE/ACM Trans Comput Biol Bioinform* 4(2):233–50
  76. Lockhart DJ, Dong H et al (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol* 14(13):1675–80
  77. Ma HW, Buer J et al (2004) Hierarchical structure and modules in the *Escherichia coli* transcriptional regulatory network revealed by a new top-down approach. *BMC Bioinformatics* 5:199
  78. Maclsaac KD, Wang T et al (2006) An improved map of conserved regulatory sites for *Saccharomyces cerevisiae*. *BMC Bioinformatics* 7:113
  79. Mangan S, Itzkovitz S et al (2006) The incoherent feed-forward loop accelerates the response-time of the gal system of *Escherichia coli*. *J Mol Biol* 356(5):1073–81
  80. Maslov S, Sneppen K (2002) Specificity and stability in topology of protein networks. *Science* 296:910–913
  81. Ma'ayan A, Jenkins SL et al (2005) Formation of regulatory patterns during signal propagation in a Mammalian cellular network. *Science* 309(5737):1078–83
  82. Milo R, Itzkovitz S et al (2004) Superfamilies of evolved and designed networks. *Science* 303(5663):1538–42
  83. Milo R, S Shen-Orr et al (2002) Network motifs: simple building blocks of complex networks. *Science* 298(5594):824–7
  84. Monod J, Cohen-Bazire G et al (1951) The biosynthesis of beta-galactosidase (lactase) in *Escherichia coli*; the specificity of induction. *Biochim Biophys Acta* 7(4):585–99
  85. Monod J, Jacob F (1961) Teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harb Symp Quant Biol* 26:389–401
  86. Nadvornik P, Drozen V (1964) Models of Neurons and Neuron Networks. *Act Nerv Super (Praha)* 6:293–302
  87. Newman MEJ (2002) Assortative mixing in networks. *Phys Rev Lett* 89:208701–208701
  88. Newman ME, Strogatz SH et al (2001) Random graphs with arbitrary degree distributions and their applications. *Phys Rev E Stat Nonlin Soft Matter Phys* 64(2 Pt 2):026118
  89. Novick A, Weiner M (1957) Enzyme Induction as an All-or-None Phenomenon. *Proc Natl Acad Sci USA* 43(7):553–66
  90. Oltvai ZN, Barabasi AL (2002) Life's complexity pyramid. *Science* 298:763–764
  91. Pastor-Satorras R, Vazquez A et al (2001) Dynamical and correlation properties of the Internet. *Phys. Rev. Lett.* 87:258701–258701
  92. Ptacek J, Devgan G et al (2005) Global analysis of protein phosphorylation in yeast. *Nature* 438(7068):679–84
  93. Qi Y, Klein-Seetharaman J et al (2005) Random forest similarity for protein-protein interaction prediction from multiple sources. *Pac Symp Biocomput*:531–42
  94. Rajewsky N (2006) microRNA target predictions in animals. *Nat Genet* 38 Suppl:S8–13
  95. Ravasz E, Barabasi AL (2003) Hierarchical organization in complex networks. *Phys Rev E Stat Nonlin Soft Matter Phys* 67:026112–026112
  96. Ravasz E, Somera AL et al (2002) Hierarchical organization of modularity in metabolic networks. *Science* 297(5586):1551–5
  97. Rives AW, Galitski T (2003) Modular organization of cellular networks. *Proc Natl Acad Sci USA* 100(3):1128–33
  98. Rouvray H (1990) The Origins of Chemical Graph Theory. In: Bonchev D, Rouvray DH (eds) *Chemical Graph Theory: Introduction and Fundamentals*, vol 41. Gordon and Breach Science Publishers, New York
  99. Rual JF, Venkatesan K et al (2005) Towards a proteome-scale map of the human protein-protein interaction network. *Nature* 437(7062):1173–8
  100. Schena M, Shalon D et al (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270(5235):467–470
  101. Schleif R (2000) Regulation of the L-arabinose operon of *Escherichia coli*. *Trends Genet* 16(12):559–65
  102. Schuster S, Pfeiffer T et al (2002) Exploring the pathway structure of metabolism: decomposition into subnetworks and application to *Mycoplasma pneumoniae*. *Bioinformatics* 18:351–361
  103. Shen-Orr SS, Milo R et al (2002) Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat Genet* 31(1):64–8
  104. Simon HA (1955) On a class of skew distribution functions. *Biometrika* 42:425–440
  105. Simonis N, Gonze D et al (2006) Modularity of the transcriptional response of protein complexes in yeast. *J Mol Biol* 363(2):589–610
  106. Simonis N, van Helden J et al (2004) Transcriptional regulation of protein complexes in yeast. *Genome Biol* 5(5):R33
  107. Smith LM, Fung S et al (1985) The synthesis of oligonucleotides containing an aliphatic amino group at the 5' terminus: synthesis of fluorescent DNA primers for use in DNA sequence analysis. *Nucleic Acids Res* 13(7):2399–412
  108. Smith LM, Sanders JZ et al (1986) Fluorescence detection in automated DNA sequence analysis. *Nature* 321(6071):674–9
  109. Snel B, Bork P et al (2002) The identification of functional modules from the genomic association of genes. *Proc Natl Acad Sci USA* 99:5890–5895
  110. Sole RV, R Pastor-Satorras et al (2002) A Model of Large-Scale Proteome Evolution. *Adv Complex Syst* 5:43–54
  111. Sood P, Krek A et al (2006) Cell-type-specific signatures of microRNAs on target mRNA expression. *Proc Natl Acad Sci USA* 103(8):2746–51
  112. Spellman PT, Sherlock G et al (1998) Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol Cell Biol* 19(12):3273–3297
  113. Spirin V, Mirny LA (2003) Protein complexes and functional modules in molecular networks. *Proc Natl Acad Sci USA* 100(21):12123–8
  114. St Onge RP, Mani R et al (2007) Systematic pathway analysis using high-resolution fitness profiling of combinatorial gene deletions. *Nat Genet* 39(2):199–206
  115. Stark A, Brennecke J et al (2005) Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell* 123(6):1133–46
  116. Stelzl U, Worm U et al (2005) A human protein-protein interaction network: a resource for annotating the proteome. *Cell* 122(6):957–68
  117. Strogatz SH (2001) Exploring complex networks. *Nature* 410(6825):268–76
  118. Stuart JM, Segal E et al (2003) A gene-coexpression network

- for global discovery of conserved genetic modules. *Science* 302:249–255
119. Tanaka R (2005) Scale-rich metabolic networks. *Phys Rev Lett* 94(16):168101
  120. Taylor RJ, Siegel AF et al (2007) Network motif analysis of a multi-mode genetic-interaction network. *Genome Biol* 8(8):R160
  121. Thieffry D, Huerta AM et al (1998) From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in *Escherichia coli*. *Bioessays* 20(5):433–40
  122. Tong AH, Lesage G et al (2004) Global mapping of the yeast genetic interaction network. *Science* 303(5659):808–13
  123. Tornow S, Mewes HW (2003) Functional modules by relating protein interaction networks and gene expression. *Nucleic Acids Res* 31:6283–6289
  124. Tsang J, Zhu J et al (2007) MicroRNA-mediated feedback and feedforward loops are recurrent network motifs in mammals. *Mol Cell* 26(5):753–67
  125. Uetz P, Giot L et al (2000) A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* 403(6770):623–7
  126. Wagner A (2001) The yeast protein interaction network evolves rapidly and contains few redundant duplicate genes. *Mol Biol Evol* 18(7):1283–92
  127. Wagner A, Fell DA (2001) The small world inside large metabolic networks. *Proc Biol Sci* 268(1478):1803–10
  128. Wall ME, Hlavacek WS et al (2004) Design of gene circuits: lessons from bacteria. *Nature Rev Genet* 5:34–42
  129. Watts DJ, Strogatz SH (1998) Collective dynamics of ‘small-world’ networks. *Nature* 393(6684):440–2
  130. Wen X, Fuhrman S et al (1998) Large-scale temporal gene expression mapping of central nervous system development. *Proc Natl Acad Sci USA* 95(1):334–339
  131. Winzeler EA (1999) Functional characterization of the *S cerevisiae* genome by gene deletion and parallel analysis. *Science* 285:901–906
  132. Wong SL, Zhang LV et al (2004) Combining biological networks to predict genetic interactions. *Proc Natl Acad Sci USA* 101(44):15682–7
  133. Wunderlich Z, Mirny LA (2006) Using the Topology of Metabolic Networks to Predict Viability of Mutant Strains. *Biophys J* 91(6):2304–2311
  134. Yeger-Lotem E, Sattath S et al (2004) Network motifs in integrated cellular networks of transcription-regulation and protein-protein interaction. *Proc Natl Acad Sci USA* 101(16):5934–9
  135. Yook SH, Oltvai ZN et al (2004) Functional and topological characterization of protein interaction networks. *Proteomics* 4(4):928–42
  136. Yu H, Gerstein M (2006) Genomic analysis of the hierarchical structure of regulatory networks. *Proc Natl Acad Sci USA* 103(40):14724–31
  137. Yu H, Luscombe NM et al (2003) Genomic analysis of gene expression relationships in transcriptional regulatory networks. *Trends Genet* 19(8):422–7
  138. Zhang L, King O et al (2005) Motifs, themes and thematic maps of an integrated *Saccharomyces cerevisiae* interaction network. *J Biol* 4(2):6
  139. Zhang LV, Wong SL et al (2004) Predicting co-complexed protein pairs using genomic and proteomic data integration. *BMC Bioinformatics* 5(1):38

## Books and Reviews

- Barabasi AL, Oltvai ZN (2004) Network biology: understanding the cell’s functional organization. *Nat Rev Genet* 5(2):101–13
- Diestel R (2005) *Graph Theory*, 3rd edn. Springer, Heidelberg

## Bivariate (Two-dimensional) Wavelets

BIN HAN

Department of Mathematical and Statistical Sciences,  
University of Alberta, Edmonton, Canada

### Article Outline

Glossary  
Definitions  
Introduction  
Bivariate Refinable Functions and Their Properties  
The Projection Method  
Bivariate Orthonormal and Biorthogonal Wavelets  
Bivariate Riesz Wavelets  
Pairs of Dual Wavelet Frames  
Future Directions  
Bibliography

### Glossary

**Dilation matrix** A  $2 \times 2$  matrix  $M$  is called a *dilation matrix* if all the entries of  $M$  are integers and all the eigenvalues of  $M$  are greater than one in modulus.

**Isotropic dilation matrix** A dilation matrix  $M$  is said to be *isotropic* if  $M$  is similar to a diagonal matrix and all its eigenvalues have the same modulus.

**Wavelet system** A wavelet system is a collection of square integrable functions that are generated from a finite set of functions (which are called wavelets) by using integer shifts and dilations.

### Definitions

Throughout this article,  $\mathbb{R}, \mathbb{C}, \mathbb{Z}$  denote the real line, the complex plane, and the set of all integers, respectively. For  $1 \leq p \leq \infty$ ,  $L_p(\mathbb{R}^2)$  denotes the set of all Lebesgue measurable bivariate functions  $f$  such that  $\|f\|_{L_p(\mathbb{R}^2)}^p := \int_{\mathbb{R}^2} |f(x)|^p dx < \infty$ . In particular, the space  $L_2(\mathbb{R}^2)$  of square integrable functions is a Hilbert space under the inner product

$$\langle f, g \rangle := \int_{\mathbb{R}^2} f(x) \overline{g(x)} dx, \quad f, g \in L_2(\mathbb{R}^2),$$

where  $\overline{g(x)}$  denotes the complex conjugate of the complex number  $g(x)$ .

In applications such as image processing and computer graphics, the following are commonly used isotropic dilation matrices:

$$\begin{aligned} M_{\sqrt{2}} &= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, & Q_{\sqrt{2}} &= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \\ M_{\sqrt{3}} &= \begin{bmatrix} 2 & -1 \\ 1 & -2 \end{bmatrix}, & dI_2 &= \begin{bmatrix} d & 0 \\ 0 & d \end{bmatrix}, \end{aligned} \quad (1)$$

where  $d$  is an integer with  $|d| > 1$ .

Using a dilation matrix  $M$ , a bivariate  $M$ -wavelet system is generated by integer shifts and dilates from a finite set  $\{\psi^1, \dots, \psi^L\}$  of functions in  $L_2(\mathbb{R}^2)$ . More precisely, the set of all the basic wavelet building blocks of an  $M$ -wavelet system generated by  $\{\psi^1, \dots, \psi^L\}$  is given by

$$\begin{aligned} X_M(\{\psi^1, \dots, \psi^L\}) \\ := \{\psi_{j,k}^{\ell,M} : j \in \mathbb{Z}, k \in \mathbb{Z}^2, \ell = 1, \dots, L\}, \end{aligned} \quad (2)$$

where

$$\psi_{j,k}^{\ell,M}(x) := |\det M|^{j/2} \psi^\ell(M^j x - k), \quad x \in \mathbb{R}^2. \quad (3)$$

One of the main goals in wavelet analysis is to find wavelet systems  $X_M(\{\psi^1, \dots, \psi^L\})$  with some desirable properties such that any two-dimensional function or signal  $f \in L_2(\mathbb{R}^2)$  can be sparsely and efficiently represented under the  $M$ -wavelet system  $X_M(\{\psi^1, \dots, \psi^L\})$ :

$$f = \sum_{\ell=1}^L \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}^2} h_{j,k}^\ell(f) \psi_{j,k}^{\ell,M}, \quad (4)$$

where  $h_{j,k}^\ell : L_2(\mathbb{R}^2) \mapsto \mathbb{C}$  are linear functionals. There are many types of wavelet systems studied in the literature. In the following, let us outline some of the most important types of wavelet systems.

### Orthonormal Wavelets

We say that  $\{\psi^1, \dots, \psi^L\}$  generates an *orthonormal  $M$ -wavelet basis* in  $L_2(\mathbb{R}^2)$  if the system  $X_M(\{\psi^1, \dots, \psi^L\})$  is an orthonormal basis of the Hilbert space  $L_2(\mathbb{R}^2)$ . That is, the linear span of elements in  $X_M(\{\psi^1, \dots, \psi^L\})$  is dense in  $L_2(\mathbb{R}^2)$  and

$$\begin{aligned} \langle \psi_{j,k}^{\ell,M}, \psi_{j',k'}^{\ell',M} \rangle &= \delta_{\ell-\ell'} \delta_{j-j'} \delta_{k-k'}, \\ \forall j, j' \in \mathbb{Z}, k, k' \in \mathbb{Z}^2, \ell, \ell' &= 1, \dots, L, \end{aligned} \quad (5)$$

where  $\delta$  denotes the *Dirac sequence* such that  $\delta_0 = 1$  and  $\delta_k = 0$  for all  $k \neq 0$ . For an orthonormal wavelet basis  $X_M(\{\psi^1, \dots, \psi^L\})$ , the linear functional  $h_{j,k}^\ell$  in (4) is given by  $h_{j,k}^\ell(f) = \langle f, \psi_{j,k}^{\ell,M} \rangle$  and the representation in (4) becomes

$$f = \sum_{\ell=1}^L \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}^2} \langle f, \psi_{j,k}^{\ell,M} \rangle \psi_{j,k}^{\ell,M}, \quad f \in L_2(\mathbb{R}^2) \quad (6)$$

with the series converging in  $L_2(\mathbb{R}^2)$ .

### Riesz Wavelets

We say that  $\{\psi^1, \dots, \psi^L\}$  generates a *Riesz  $M$ -wavelet basis* in  $L_2(\mathbb{R}^2)$  if the system  $X_M(\{\psi^1, \dots, \psi^L\})$  is a Riesz basis of  $L_2(\mathbb{R}^2)$ . That is, the linear span of elements in  $X_M(\{\psi^1, \dots, \psi^L\})$  is dense in  $L_2(\mathbb{R}^2)$  and there exist two positive constants  $C_1$  and  $C_2$  such that

$$\begin{aligned} C_1 \sum_{\ell=1}^L \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}^2} |c_{j,k}^\ell|^2 &\leq \left\| \sum_{\ell=1}^L \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}^2} c_{j,k}^\ell \psi_{j,k}^{\ell,M} \right\|_{L_2(\mathbb{R}^2)}^2 \\ &\leq C_2 \sum_{\ell=1}^L \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}^2} |c_{j,k}^\ell|^2 \end{aligned}$$

for all finitely supported sequences  $\{c_{j,k}^\ell\}_{j \in \mathbb{Z}, k \in \mathbb{Z}^2, \ell=1, \dots, L}$ . Clearly, a Riesz  $M$ -wavelet generalizes an orthonormal  $M$ -wavelet by relaxing the orthogonality requirement in (5). For a Riesz basis  $X_M(\{\psi^1, \dots, \psi^L\})$ , it is well-known that there exists a dual Riesz basis  $\{\tilde{\psi}^{\ell,j,k} : j \in \mathbb{Z}, k \in \mathbb{Z}^2, \ell = 1, \dots, L\}$  of elements in  $L_2(\mathbb{R}^2)$  (this set is not necessarily generated by integer shifts and dilates from some finite set of functions) such that (5) still holds after replacing  $\psi_{j',k'}^{\ell',M}$  by  $\tilde{\psi}^{\ell',j',k'}$ . For a Riesz wavelet basis, the linear functional in (4) becomes  $h_{j,k}^\ell(f) = \langle f, \tilde{\psi}^{\ell,j,k} \rangle$ . In fact,  $\tilde{\psi}^{\ell,j,k} := \mathcal{F}^{-1}(\psi_{j,k}^{\ell,M})$ , where  $\mathcal{F} : L_2(\mathbb{R}^2) \mapsto L_2(\mathbb{R}^2)$  is defined to be

$$\mathcal{F}(f) := \sum_{\ell=1}^L \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}^2} \langle f, \psi_{j,k}^{\ell,M} \rangle \psi_{j,k}^{\ell,M}, \quad f \in L_2(\mathbb{R}^2). \quad (7)$$

### Wavelet Frames

A further generalization of a Riesz wavelet is a wavelet frame. We say that  $\{\psi^1, \dots, \psi^L\}$  generates an  *$M$ -wavelet frame* in  $L_2(\mathbb{R}^2)$  if the system  $X_M(\{\psi^1, \dots, \psi^L\})$  is



a frame of  $L_2(\mathbb{R}^2)$ . That is, there exist two positive constants  $C_1$  and  $C_2$  such that

$$\begin{aligned} C_1 \|f\|_{L_2(\mathbb{R}^2)}^2 &\leq \sum_{\ell=1}^L \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}^2} |\langle f, \psi_{j,k}^{\ell,M} \rangle|^2 \\ &\leq C_2 \|f\|_{L_2(\mathbb{R}^2)}^2, \quad \forall f \in L_2(\mathbb{R}^2). \end{aligned} \quad (8)$$

It is not difficult to check that a Riesz  $M$ -wavelet is an  $M$ -wavelet frame. Then (8) guarantees that the *frame operator*  $\mathcal{F}$  in (7) is a bounded and invertible linear operator. For an  $M$ -wavelet frame, the linear functional in (4) can be chosen to be  $h_{j,k}^\ell(f) = \langle f, \mathcal{F}^{-1}(\psi_{j,k}^{\ell,M}) \rangle$ ; however, such functionals may not be unique. The fundamental difference between a Riesz wavelet and a wavelet frame lies in that for any given function  $f \in L_2(\mathbb{R}^2)$ , the representation in (4) is unique under a Riesz wavelet while it may not be unique under a wavelet frame. The representation in (4) with the choice  $h_{j,k}^\ell(f) = \langle f, \mathcal{F}^{-1}(\psi_{j,k}^{\ell,M}) \rangle$  is called the *canonical representation* of a wavelet frame. In other words,  $\{\mathcal{F}^{-1}(\psi_{j,k}^{\ell,M}) : j \in \mathbb{Z}, k \in \mathbb{Z}^2, \ell = 1, \dots, L\}$  is called the *canonical dual frame* of the given wavelet frame  $X_M(\{\psi^1, \dots, \psi^L\})$ .

### Biorthogonal Wavelets

We say that  $(\{\psi^1, \dots, \psi^L\}, \{\tilde{\psi}^1, \dots, \tilde{\psi}^L\})$  generates a pair of biorthogonal  $M$ -wavelet bases in  $L_2(\mathbb{R}^2)$  if each of  $X_M(\{\psi^1, \dots, \psi^L\})$  and  $X_M(\{\tilde{\psi}^1, \dots, \tilde{\psi}^L\})$  is a Riesz basis of  $L_2(\mathbb{R}^2)$  and (5) still holds after replacing  $\psi_{j',k'}^{\ell',M}$  by  $\tilde{\psi}_{j',k'}^{\ell',M}$ . In other words, the dual Riesz basis of  $X_M(\{\psi^1, \dots, \psi^L\})$  has the wavelet structure and is given by  $X_M(\{\tilde{\psi}^1, \dots, \tilde{\psi}^L\})$ . For a biorthogonal wavelet, the wavelet representation in (4) becomes

$$f = \sum_{\ell=1}^L \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}^2} \langle f, \tilde{\psi}_{j,k}^{\ell,M} \rangle \psi_{j,k}^{\ell,M}, \quad f \in L_2(\mathbb{R}^2). \quad (9)$$

Obviously,  $\{\psi^1, \dots, \psi^L\}$  generates an orthonormal  $M$ -wavelet basis in  $L_2(\mathbb{R}^2)$  if and only if  $(\{\psi^1, \dots, \psi^L\}, \{\psi^1, \dots, \psi^L\})$  generates a pair of biorthogonal  $M$ -wavelet bases in  $L_2(\mathbb{R}^2)$ .

### Dual Wavelet Frames

Similarly, we have the notion of a pair of dual wavelet frames. We say that  $(\{\psi^1, \dots, \psi^L\}, \{\tilde{\psi}^1, \dots, \tilde{\psi}^L\})$  generates a pair of dual  $M$ -wavelet frames in  $L_2(\mathbb{R}^2)$  if each of  $X_M(\{\psi^1, \dots, \psi^L\})$  and  $X_M(\{\tilde{\psi}^1, \dots, \tilde{\psi}^L\})$  is an

$M$ -wavelet frame in  $L_2(\mathbb{R}^2)$  and

$$\begin{aligned} \langle f, g \rangle &= \sum_{\ell=1}^L \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}^2} \langle f, \tilde{\psi}_{j,k}^{\ell,M} \rangle \langle \psi_{j,k}^{\ell,M}, g \rangle, \\ f, g &\in L_2(\mathbb{R}^2). \end{aligned}$$

It follows from the above identity that (9) still holds for a pair of dual  $M$ -wavelet frames. We say that  $\{\psi^1, \dots, \psi^L\}$  generates a tight  $M$ -wavelet frame if  $(\{\psi^1, \dots, \psi^L\}, \{\psi^1, \dots, \psi^L\})$  generates a pair of dual  $M$ -wavelet frames in  $L_2(\mathbb{R}^2)$ . For a tight wavelet frame  $\{\psi^1, \dots, \psi^L\}$ , the wavelet representation in (6) still holds. So, a tight wavelet frame is a generalization of an orthonormal wavelet basis.

### Introduction

Bivariate (two-dimensional) wavelets are of interest in representing and processing two dimensional data such as images and surfaces. In this article we shall discuss some basic background and results on bivariate wavelets.

We denote  $\Pi_J$  the set of all bivariate polynomials of total degree at most  $J$ . For compactly supported functions  $\psi^1, \dots, \psi^L$ , we say that  $\{\psi^1, \dots, \psi^L\}$  has  $J$  vanishing moments if  $\langle P, \psi^\ell \rangle = 0$  for all  $\ell = 1, \dots, L$  and all polynomials  $P \in \Pi_{J-1}$ .

The advantages of wavelet representations largely lie in the following aspects:

1. There is a fast wavelet transform (FWT) for computing the wavelet coefficients  $h_{j,k}^\ell(f)$  in the wavelet representation (4).
2. The wavelet representation has good time and frequency localization. Roughly speaking, the basic building blocks  $\psi_{j,k}^{\ell,M}$  have good time localization and smoothness.
3. The wavelet representation is sparse. For a smooth function  $f$ , most wavelet coefficients are negligible. Generally, the wavelet coefficient  $h_{j,k}^\ell(f)$  only depends on the information of  $f$  in a small neighborhood of the support of  $\psi_{j,k}^{\ell,M}$  (more precisely, the support of  $\tilde{\psi}_{j,k}^{\ell,M}$  if  $h_{j,k}^\ell(f) = \langle f, \tilde{\psi}_{j,k}^{\ell,M} \rangle$ ). If  $f$  in a small neighborhood of the support of  $\psi_{j,k}^{\ell,M}$  is smooth or behaves like a polynomial to certain degree, then by the vanishing moments of the dual wavelet functions,  $h_{j,k}^\ell(f)$  is often negligible.
4. For a variety of function spaces  $B$  such as Sobolev and Besov spaces, its norm  $\|f\|_B$ ,  $f \in B$ , is equivalent to certain weighted sequence norm of its wavelet coefficients  $\{h_{j,k}^\ell(f)\}_{j \in \mathbb{Z}, k \in \mathbb{Z}^2, \ell=1, \dots, L}$ .

For more details on advantages and applications of wavelets, see [6,8,14,16,21,58].

For a dilation matrix  $M = dI_2$ , the easiest way for obtaining a bivariate wavelet is to use the tensor product method. In other words, all the functions in the generator set  $\{\psi^1, \dots, \psi^L\} \subset L_2(\mathbb{R}^2)$  take the form  $\psi^\ell(x, y) = f^\ell(x)g^\ell(y)$ ,  $x, y \in \mathbb{R}$ , where  $f^\ell$  and  $g^\ell$  are some univariate functions in  $L_2(\mathbb{R})$ . However, tensor product (also called *separable*) bivariate wavelets give preference to the horizontal and vertical directions which may not be desirable in applications such as image processing ([9,16,50,58]). Also, for a non-diagonal dilation matrix, it is difficult or impossible to use the tensor product method to obtain bivariate wavelets. Therefore, nonseparable bivariate wavelets themselves are of importance and interest in both theory and application. In this article, we shall address several aspects of bivariate wavelets with a general two-dimensional dilation matrix.

### Bivariate Refinable Functions and Their Properties

In order to have a fast wavelet transform to compute the wavelet coefficients in (4), the generators  $\psi^1, \dots, \psi^L$  in a wavelet system are generally obtained from a refinable function  $\phi$  via a multiresolution analysis [6,15,16,58].

Let  $M$  be a  $2 \times 2$  dilation matrix. For a function  $\phi$  in  $L_2(\mathbb{R}^2)$ , we say that  $\phi$  is *M-refinable* if it satisfies the following refinement equation

$$\phi(x) = |\det M| \sum_{k \in \mathbb{Z}^2} a_k \phi(Mx - k), \quad \text{a.e. } x \in \mathbb{R}^2, \quad (10)$$

where  $a: \mathbb{Z}^2 \mapsto \mathbb{C}$  is a finitely supported sequence on  $\mathbb{Z}^2$  satisfying  $\sum_{k \in \mathbb{Z}^2} a_k = 1$ . Such a sequence  $a$  is often called a *mask* in wavelet analysis and computer graphics, or a *low-pass filter* in signal processing. Using the Fourier transform  $\hat{\phi}$  of  $\phi \in L_2(\mathbb{R}^2) \cap L_1(\mathbb{R}^2)$  and the Fourier series  $\hat{a}$  of  $a$ , which are defined to be

$$\begin{aligned} \hat{\phi}(\xi) &:= \int_{\mathbb{R}^2} \phi(x) e^{-ix \cdot \xi} dx \quad \text{and} \\ \hat{a}(\xi) &:= \sum_{k \in \mathbb{Z}^2} a_k e^{-ik \cdot \xi}, \quad \xi \in \mathbb{R}^2, \end{aligned} \quad (11)$$

the refinement equation (10) can be equivalently rewritten as

$$\hat{\phi}(M^T \xi) = \hat{a}(\xi) \hat{\phi}(\xi), \quad \xi \in \mathbb{R}^2, \quad (12)$$

where  $M^T$  denotes the transpose of the dilation matrix  $M$ . Since  $\hat{a}(0) = 1$  and  $\hat{a}$  is a trigonometric polynomial, one can define a function  $\hat{\phi}$  by

$$\hat{\phi}(\xi) := \prod_{j=1}^{\infty} \hat{a}((M^T)^{-j} \xi), \quad \xi \in \mathbb{R}^2, \quad (13)$$

with the series converging uniformly on any compact set of  $\mathbb{R}^2$ . It is known ([4,16]) that  $\phi$  is a compactly supported tempered distribution and clearly  $\phi$  satisfies the refinement equation in (12) with mask  $a$ . We call  $\phi$  the *standard refinable function* associated with mask  $a$  and dilation  $M$ . Now the generators  $\psi^1, \dots, \psi^L$  of a wavelet system are often obtained from the refinable function  $\phi$  via

$$\widehat{\psi^\ell}(M^T \xi) := \widehat{a^\ell}(\xi) \hat{\phi}(\xi), \quad \xi \in \mathbb{R}^2, \quad \ell = 1, \dots, L,$$

for some  $2\pi$ -periodic trigonometric polynomials  $\widehat{a^\ell}$ . Such sequences  $a^\ell$  are called *wavelet masks* or *high-pass filters*.

Except for very few special masks  $a$  such as masks for box spline refinable functions (see [20] for box splines), most refinable functions  $\phi$ , obtained in (13) from mask  $a$  and dilation  $M$ , do not have explicit analytic expressions and a so-called *cascade algorithm* or *subdivision scheme* is used to approximate the refinable function  $\phi$ . Let  $B$  be a Banach space of bivariate functions. Starting with a suitable initial function  $f \in B$ , we iteratively compute a sequence  $\{Q_{a,M}^n f\}_{n=0}^{\infty}$  of functions, where

$$Q_{a,M} f := |\det M| \sum_{k \in \mathbb{Z}^2} a_k f(M \cdot -k). \quad (14)$$

If the sequence  $\{Q_{a,M}^n f\}_{n=0}^{\infty}$  of functions converges to some  $f_\infty \in B$  in the Banach space  $B$ , then we have  $Q_{a,M} f_\infty = f_\infty$ . That is, as a fixed point of the *cascade operator*  $Q_{a,M}$ ,  $f_\infty$  is a solution to the refinement equation (10). A cascade algorithm plays an important role in the study of refinable functions in wavelet analysis and of subdivision surfaces in computer graphics. For more details on cascade algorithms and subdivision schemes, see [4,19,21,22,23,31,36,54] and numerous references therein.

One of the most important properties of a refinable function  $\phi$  is its smoothness, which is measured by its  $L_p$  smoothness exponent  $\nu_p(\phi)$  and is defined to be

$$\nu_p(\phi) := \sup\{\nu: \phi \in W_p^\nu(\mathbb{R}^2)\}, \quad 1 \leq p \leq \infty, \quad (15)$$

where  $W_p^\nu(\mathbb{R}^2)$  denotes the fractional Sobolev space of order  $\nu > 0$  in  $L_p(\mathbb{R}^2)$ .

The notion of sum rules of a mask is closely related to the vanishing moments of a wavelet system ([16,44]). For a bivariate mask  $a$ , we say that  $a$  satisfies the *sum rules* of order  $J$  with the dilation matrix  $M$  if

$$\sum_{k \in \mathbb{Z}^2} a_{j+Mk} P(j+Mk) = \sum_{k \in \mathbb{Z}^2} a_{Mk} P(Mk), \quad \forall P \in \Pi_{J-1}. \quad (16)$$

Or equivalently,  $\partial_1^{\mu_1} \partial_2^{\mu_2} \hat{a}(2\pi\gamma) = 0$  for all  $\mu_1 + \mu_2 < J$  and  $\gamma \in [(M^T)^{-1}\mathbb{Z}^2] \setminus \mathbb{Z}^2$ , where  $\partial_1$  and  $\partial_2$  denote the partial derivatives in the first and second coordinate, respectively. Throughout the article, we denote  $sr(a, M)$  the highest order of sum rules satisfied by the mask  $a$  with the dilation matrix  $M$ .

To investigate various properties of refinable functions, next we introduce a quantity  $v_p(a, M)$  in wavelet analysis (see [31]). For a bivariate mask  $a$  and a dilation matrix  $M$ , we denote

$$v_p(a, M) := -\log_{\rho(M)}[|\det M|^{1-1/p} \rho(a, M, p)], \quad 1 \leq p \leq \infty, \quad (17)$$

where  $\rho(M)$  denotes the *spectral radius* of  $M$  and

$$\rho(a, M, p) := \max \left\{ \limsup_{n \rightarrow \infty} \|a^{n, (\mu_1, \mu_2)}\|_{\ell_p(\mathbb{Z}^2)}^{1/n} : \mu_1 + \mu_2 = sr(a, M), \mu_1, \mu_2 \in \mathbb{N} \cup \{0\} \right\},$$

where the sequence  $a^{n, (\mu_1, \mu_2)}$  is defined by its Fourier series as follows: for  $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$ ,

$$\widehat{a^{n, (\mu_1, \mu_2)}}(\xi) := (1 - e^{-i\xi_1})^{\mu_1} (1 - e^{-i\xi_2})^{\mu_2} \hat{a}((M^T)^{n-1}\xi) \hat{a}((M^T)^{n-2}\xi) \cdots \hat{a}(M^T\xi) \hat{a}(\xi).$$

It is known that ([31,36])  $v_p(a, M) \geq v_q(a, M) \geq v_p(a, M) + (1/q - 1/p) \log_{\rho(M)} |\det M|$  for all  $1 \leq p \leq q \leq \infty$ .

Parts of the following result essentially appeared in various forms in many papers in the literature. For the study of refinable functions and cascade algorithms, see [4,10,12,15,16,22,23,25,29,31,45,48,53] and references therein. The following is from [31].

**Theorem 1** *Let  $M$  be a  $2 \times 2$  isotropic dilation matrix and  $a$  be a finitely supported bivariate mask. Let  $\phi$  denote the standard refinable function associated with mask  $a$  and dilation  $M$ . Then  $v_p(\phi) \geq v_p(a, M)$  for all  $1 \leq p \leq \infty$ . If the shifts of  $\phi$  are stable, that is,  $\{\phi(\cdot - k) : k \in \mathbb{Z}^2\}$  is a Riesz system in  $L_p(\mathbb{R}^2)$ , then  $v_p(\phi) = v_p(a, M)$ . Moreover, for every nonnegative integer  $J$ , the following statements are equivalent.*

1. *For every compactly supported function  $f \in W_p^J(\mathbb{R}^2)$  (if  $p = \infty$ , we require  $f \in C^J(\mathbb{R}^2)$ ) such that  $\hat{f}(0) = 1$  and  $\partial_1^{\mu_1} \partial_2^{\mu_2} \hat{f}(2\pi k) = 0$  for all  $\mu_1 + \mu_2 < J$  and  $k \in \mathbb{Z}^2 \setminus \{0\}$ , the cascade sequence  $\{Q_{a,M}^n f\}_{n=0}^\infty$  converges in the Sobolev space  $W_p^J(\mathbb{R}^2)$  (in fact, the limit function is  $\phi$ ).*
2.  $v_p(a, M) > J$ .

Symmetry of a refinable function is also another important property of a wavelet system in applications. For example, symmetry is desirable in order to handle the boundary of an image or to improve the visual quality of a surface ([16,21,22,32,50,58]).

Let  $M$  be a  $2 \times 2$  dilation matrix and  $G$  be a finite set of  $2 \times 2$  integer matrices. We say that  $G$  is a *symmetry group with respect to  $M$*  ([27,29]) if  $G$  is a group under matrix multiplication and  $MEM^{-1} \in G$  for all  $E \in G$ . In dimension two, two commonly used symmetry groups are  $D_4$  and  $D_6$ :

$$D_4 := \left\{ \pm \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \pm \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \pm \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \pm \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\}$$

and

$$D_6 := \left\{ \pm \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \pm \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}, \pm \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}, \right. \\ \left. \pm \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \pm \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix}, \pm \begin{bmatrix} -1 & 0 \\ -1 & 1 \end{bmatrix} \right\}.$$

For symmetric refinable functions, we have the following result ([32, Proposition 1]):

**Proposition 2** *Let  $M$  be a  $2 \times 2$  dilation matrix and  $G$  be a symmetry group with respect to  $M$ . Let  $a$  be a bivariate mask and  $\phi$  be the standard refinable function associated with mask  $a$  and dilation  $M$ . Then  $a$  is  $G$ -symmetric with center  $c_a$ :  $a_{E(k-c_a)+c_a} = a_k$  for all  $k \in \mathbb{Z}^2$  and  $E \in G$ , if and only if,  $\phi$  is  $G$ -symmetric with center  $c$ :*

$$\phi(E(\cdot - c) + c) = \phi, \quad \forall E \in G \quad \text{with } c := (M - I_2)^{-1} c_a.$$

In the following, let us present some examples of bivariate refinable functions.

**Example 3** Let  $M = 2I_2$  and  $\hat{a}(\xi_1, \xi_2) := \cos^4(\xi_1/2) \cos^4(\xi_2/2)$ . Then  $a$  is a tensor product mask with  $sr(a, 2I_2) = 4$  and  $a$  is  $D_4$ -symmetric with center 0. The associated standard refinable function  $\phi$  is the tensor product spline of order 4. The Catmull-Clark subdivision scheme for quadrilateral meshes in computer graphics is based on this mask  $a$  ([21]).

**Example 4** Let  $M = 2I_2$  and  $\hat{a}(\xi_1, \xi_2) := \cos^2(\xi_1/2) \cos^2(\xi_2/2) \cos^2(\xi_1/2 + \xi_2/2)$ . Then  $sr(a, 2I_2) = 4$  and  $a$  is  $D_6$ -symmetric with center 0. The associated refinable function  $\phi$  is the convolution of the three direction box spline with itself ([20]). The Loop subdivision scheme for triangular meshes in computer graphics is based on this mask  $a$  ([21]).

Another important property of a refinable function is interpolation. We say that a bivariate function  $\phi$  is *interpolating* if  $\phi$  is continuous and  $\phi(k) = \delta_k$  for all  $k \in \mathbb{Z}^2$ .

**Theorem 5** *Let  $M$  be a  $2 \times 2$  dilation matrix and  $a$  be a finitely supported bivariate mask. Let  $\phi$  denote the standard refinable function associated with mask  $a$  and dilation  $M$ . Then  $\phi$  is an interpolating function if and only if  $v_\infty(a, M) > 0$  and  $a$  is an interpolatory mask with dilation  $M$ , that is,  $a_0 = |\det M|^{-1}$  and  $a_{Mk} = 0$  for all  $k \in \mathbb{Z}^2 \setminus \{0\}$ .*

**Example 6** Let  $M = 2I_2$  and

$$\begin{aligned} \hat{a}(\xi_1, \xi_2) &:= \cos(\xi_1/2) \cos(\xi_2/2) \cos(\xi_1/2 + \xi_2/2) \\ &\quad [1 + 2 \cos(\xi_1) + 2 \cos(\xi_2) + 2 \cos(\xi_1 + \xi_2) \\ &\quad - \cos(2\xi_1 + \xi_2) - \cos(\xi_1 + 2\xi_2) \\ &\quad - \cos(\xi_1 - \xi_2)]/4. \end{aligned} \quad (18)$$

Then  $sr(a, 2I_2) = 4$ ,  $a$  is  $D_6$ -symmetric with center 0 and  $a$  is an interpolatory mask with dilation  $2I_2$ . Since  $v_2(a, 2I_2) \approx 2.44077$ , so,  $v_\infty(a, 2I_2) \geq v_2(a, 2I_2) - 1 \approx 1.44077 > 0$ . By Theorem 5, the associated standard refinable function  $\phi$  is interpolating. The butterfly interpolatory subdivision scheme for triangular meshes is based on this mask  $a$  (see [23]).

For more details on bivariate interpolatory masks and interpolating refinable functions, see [4,19,22,23,25,26,27,31,37,38,39,59].

### The Projection Method

In applications, one is interested in analyzing some optimal properties of multivariate wavelets. The projection method is useful for this purpose.

Let  $r$  and  $s$  be two positive integers with  $r \leq s$ . Let  $P$  be an  $r \times s$  real-valued matrix. For a compactly supported function  $\phi$  and a finitely supported mask  $a$  in dimension  $s$ , we define the *projected function*  $P\phi$  and *projected mask*  $Pa$  in dimension  $r$  by

$$\begin{aligned} \widehat{P\phi}(\xi) &:= \hat{\phi}(P^T \xi) \quad \text{and} \\ \widehat{Pa}(\xi) &:= \sum_{j=1}^t \hat{a}(P^T \xi + 2\pi \varepsilon_j), \quad \xi \in \mathbb{R}^r, \end{aligned} \quad (19)$$

where  $\hat{\phi}$  and  $\hat{a}$  are understood to be continuous and  $\{\varepsilon_1, \dots, \varepsilon_t\}$  is a complete set of representatives of the distinct cosets of  $[P^T \mathbb{R}^r]/\mathbb{Z}^s$ . If  $P$  is an integer matrix, then  $\widehat{Pa}(\xi) = \hat{a}(P^T \xi)$ .

Now we have the following result on projected refinable functions ([27,28,34,35]).

**Theorem 7** *Let  $N$  be an  $s \times s$  dilation matrix and  $M$  be an  $r \times r$  dilation matrix. Let  $P$  be an  $r \times s$  integer matrix such that  $PN = MP$  and  $P\mathbb{Z}^s = \mathbb{Z}^r$ . Let  $\hat{a}$  be a  $2\pi$ -periodic trigonometric polynomials in  $s$ -variables with  $\hat{a}(0) = 1$  and  $\phi$  the standard  $N$ -refinable function associated with mask  $a$ . Then  $sr(a, N) \leq sr(Pa, M)$  and  $\widehat{P\phi}(M^T \xi) = \widehat{Pa}(\xi) \widehat{P\phi}(\xi)$ . That is,  $P\phi$  is  $M$ -refinable with mask  $Pa$ . Moreover, for all  $1 \leq p \leq \infty$ ,*

$$\begin{aligned} v_p(\phi) &\leq v_p(P\phi) \quad \text{and} \\ |\det M|^{1-1/p} \rho(Pa, M, p) &\leq |\det N|^{1-1/p} \rho(a, N, p). \end{aligned}$$

If we further assume that  $\rho(M) = \rho(N)$ , then  $v_p(a, N) \leq v_p(Pa, M)$ .

As pointed out in [34,35], the projection method is closely related to box splines. For a given  $r \times s$  (direction) integer matrix  $\mathcal{E}$  of rank  $r$  with  $r \leq s$ , the Fourier transform of its associated box spline  $M_{\mathcal{E}}$  and its mask  $a_{\mathcal{E}}$  are given by (see [20])

$$\begin{aligned} \widehat{M_{\mathcal{E}}}(\xi) &:= \prod_{k \in \mathcal{E}} \frac{1 - e^{-ik \cdot \xi}}{ik \cdot \xi} \quad \text{and} \\ \widehat{a_{\mathcal{E}}}(\xi) &:= \prod_{k \in \mathcal{E}} \frac{1 + e^{-ik \cdot \xi}}{2}, \quad \xi \in \mathbb{R}^r, \end{aligned} \quad (20)$$

where  $k \in \mathcal{E}$  means that  $k$  is a column vector of  $\mathcal{E}$  and  $k$  goes through all the columns of  $\mathcal{E}$  once and only once. Let  $\chi_{[0,1]^s}$  denote the characteristic function of the unit cube  $[0,1]^s$ . From (20), it is evident that the box spline  $M_{\mathcal{E}}$  is just the projected function  $\mathcal{E}\chi_{[0,1]^s}$ , since  $\mathcal{E}\chi_{[0,1]^s} = M_{\mathcal{E}}$  by  $\widehat{\mathcal{E}\chi_{[0,1]^s}}(\xi) = \widehat{\chi_{[0,1]^s}}(\mathcal{E}^T \xi)$ ,  $\xi \in \mathbb{R}^r$ . Note that  $M_{\mathcal{E}}$  is  $2I_r$ -refinable with the mask  $a_{\mathcal{E}}$ , since  $\widehat{M_{\mathcal{E}}}(2\xi) = \widehat{a_{\mathcal{E}}}(\xi) \widehat{M_{\mathcal{E}}}(\xi)$ .

As an application of the projection method in Theorem 7, we have ([25, Theorem 3.5.]):

**Corollary 8** *Let  $M = 2I_s$  and  $a$  be an interpolatory mask with dilation  $M$  such that  $a$  is supported inside  $[-3, 3]^s$ . Then  $v_\infty(a, 2I_s) \leq 2$  and therefore,  $\phi \notin C^2(\mathbb{R}^s)$ , where  $\phi$  is the standard refinable function associated with mask  $a$  and dilation  $2I_s$ .*

We use proof by contradiction. Suppose  $v_\infty(a, 2I_s) > 2$ . Let  $P = [1, 0, \dots, 0]$  be a  $1 \times s$  matrix. Then we must have  $sr(a, 2I_s) \geq 3$ , which, combining with other assumptions on  $a$ , will force  $\widehat{Pa}(\xi) = 1/2 + 9/16 \cos(\xi) - 1/16 \cos(3\xi)$ . Since  $v_\infty(Pa, 2) = 2$ , by Theorem 7, we must have  $v_\infty(a, 2I_s) \leq v_\infty(Pa, 2) = 2$ . A contradiction. So,  $v_\infty(a,$

$2I_s) \leq 2$ . In particular, the refinable function in the butterfly subdivision scheme in Example 6 is not  $C^2$ .

The projection method can be used to construct interpolatory masks painlessly [27].

**Theorem 9** Let  $M$  be an  $r \times r$  dilation matrix. Then there is an  $r \times r$  integer matrix  $H$  such that  $M\mathbb{Z}^r = H\mathbb{Z}^r$  and  $H^r = |\det M|I_r$ . Let  $P := |\det M|^{-1}H$ . Then for any (tensor product) interpolatory mask  $a$  with dilation  $|\det M|I_r$ , the projected mask  $Pa$  is an interpolatory mask with dilation  $M$  and  $sr(Pa, M) \geq sr(a, |\det M|I_r)$ .

**Example 10** For  $M = M_{\sqrt{2}}$  or  $M = Q_{\sqrt{2}}$ , we can take  $H := M_{\sqrt{2}}$  in Theorem 9.

For more details on the projection method, see [25,27,28,32,34,35].

## Bivariate Orthonormal and Biorthogonal Wavelets

In this section, we shall discuss the analysis and construction of bivariate orthonormal and biorthogonal wavelets.

For analysis of biorthogonal wavelets, the following result is well-known (see [9,11,16,24,26,31,46,51,53] and references therein):

**Theorem 11** Let  $M$  be a  $2 \times 2$  dilation matrix and  $a, \tilde{a}$  be two finitely supported bivariate masks. Let  $\phi$  and  $\tilde{\phi}$  be the standard  $M$ -refinable functions associated with masks  $a$  and  $\tilde{a}$ , respectively. Then  $\phi, \tilde{\phi} \in L_2(\mathbb{R}^2)$  and satisfy the biorthogonality relation

$$\langle \phi, \tilde{\phi}(\cdot - k) \rangle = \delta_k, \quad k \in \mathbb{Z}^2, \quad (21)$$

if and only if,  $v_2(a, M) > 0$ ,  $v_2(\tilde{a}, M) > 0$ , and  $(a, \tilde{a})$  is a pair of dual masks:

$$\sum_{\gamma \in \Gamma_{M^T}} \widehat{a}(\xi + 2\pi\gamma) \widehat{\tilde{a}}(\xi + 2\pi\gamma) = 1, \quad \xi \in \mathbb{R}^2, \quad (22)$$

where  $\Gamma_{M^T}$  is a complete set of representatives of distinct cosets of  $[(M^T)^{-1}\mathbb{Z}^2]/\mathbb{Z}^2$  with  $0 \in \Gamma_{M^T}$ . Moreover, if (21) holds and there exist  $2\pi$ -periodic trigonometric polynomials  $\widehat{a}^1, \dots, \widehat{a}^{m-1}$ ,  $\widehat{\tilde{a}}^1, \dots, \widehat{\tilde{a}}^{m-1}$  with  $m := |\det M|$  such that

$$\mathcal{M}_{[\widehat{a}, \widehat{a}^1, \dots, \widehat{a}^{m-1}]}(\xi) \overline{\mathcal{M}_{[\widehat{\tilde{a}}, \widehat{\tilde{a}}^1, \dots, \widehat{\tilde{a}}^{m-1}]}(\xi)}^T = I_m, \quad \xi \in \mathbb{R}^2, \quad (23)$$

where for  $\xi \in \mathbb{R}^2$ ,

$$\mathcal{M}_{[\widehat{a}, \widehat{a}^1, \dots, \widehat{a}^{m-1}]}(\xi) := \begin{bmatrix} \widehat{a}(\xi + 2\pi\gamma_0) & \widehat{a}^1(\xi + 2\pi\gamma_0) & \dots & \widehat{a}^{m-1}(\xi + 2\pi\gamma_0) \\ \widehat{a}(\xi + 2\pi\gamma_1) & \widehat{a}^1(\xi + 2\pi\gamma_1) & \dots & \widehat{a}^{m-1}(\xi + 2\pi\gamma_1) \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{a}(\xi + 2\pi\gamma_{m-1}) & \widehat{a}^1(\xi + 2\pi\gamma_{m-1}) & \dots & \widehat{a}^{m-1}(\xi + 2\pi\gamma_{m-1}) \end{bmatrix} \quad (24)$$

with  $\{\gamma_0, \dots, \gamma_{m-1}\} := \Gamma_{M^T}$  and  $\gamma_0 := 0$ . Define  $\psi^1, \dots, \psi^{m-1}, \tilde{\psi}^1, \dots, \tilde{\psi}^{m-1}$  by

$$\begin{aligned} \widehat{\psi}^\ell(M^T\xi) &:= \widehat{a}^\ell(\xi) \widehat{\phi}(\xi) \quad \text{and} \\ \widehat{\tilde{\psi}}^\ell(M^T\xi) &:= \widehat{\tilde{a}}^\ell(\xi) \widehat{\tilde{\phi}}(\xi), \quad \ell = 1, \dots, m-1. \end{aligned} \quad (25)$$

Then  $(\{\psi^1, \dots, \psi^{m-1}\}, \{\tilde{\psi}^1, \dots, \tilde{\psi}^{m-1}\})$  generates a pair of biorthogonal  $M$ -wavelet bases in  $L_2(\mathbb{R}^2)$ .

As a direct consequence of Theorem 11, for orthonormal wavelets, we have

**Corollary 12** Let  $M$  be a  $2 \times 2$  dilation matrix and  $a$  be a finitely supported bivariate mask. Let  $\phi$  be the standard refinable function associated with mask  $a$  and dilation  $M$ . Then  $\phi$  has orthonormal shifts, that is,  $\langle \phi, \phi(\cdot - k) \rangle = \delta_k$  for all  $k \in \mathbb{Z}^2$ , if and only if,  $v_2(a, M) > 0$  and  $a$  is an orthogonal mask (that is, (22) is satisfied with  $\widehat{a}$  being replaced by  $\widehat{a}$ ). If in addition there exist  $2\pi$ -periodic trigonometric polynomials  $\widehat{a}^1, \dots, \widehat{a}^{m-1}$  with  $m := |\det M|$  such that  $\mathcal{M}_{[\widehat{a}, \widehat{a}^1, \dots, \widehat{a}^{m-1}]}(\xi) \overline{\mathcal{M}_{[\widehat{a}, \widehat{a}^1, \dots, \widehat{a}^{m-1}]}(\xi)}^T = I_m$ . Then  $\{\psi^1, \dots, \psi^{m-1}\}$ , defined in (25), generates an orthonormal  $M$ -wavelet basis in  $L_2(\mathbb{R}^2)$ .

The masks  $a$  and  $\tilde{a}$  are called *low-pass filters* and the masks  $a^1, \dots, a^{m-1}$ ,  $\tilde{a}^1, \dots, \tilde{a}^{m-1}$  are called *high-pass filters* in the engineering literature. The equation in (23) is called the *matrix extension problem* in the wavelet literature (see [6,16,46,58]). Given a pair of dual masks  $a$  and  $\tilde{a}$ , though the existence of  $\widehat{a}^1, \dots, \widehat{a}^{m-1}$ ,  $\widehat{\tilde{a}}^1, \dots, \widehat{\tilde{a}}^{m-1}$  (without symmetry) in the matrix extension problem in (23) is guaranteed by the Quillen–Suslin theorem, it is far from a trivial task to construct them (in particular, with symmetry) algorithmically. For the orthogonal case, the general theory of the matrix extension problem in (23) with  $\tilde{a} = a$  and  $\tilde{a}^\ell = a^\ell$ ,  $\ell = 1, \dots, m-1$ , remains unanswered. However, when  $|\det M| = 2$ , the matrix extension problem is trivial. In fact, letting  $\gamma \in \Gamma_{M^T} \setminus \{0\}$  (that is,  $\Gamma_{M^T} = \{0, \gamma\}$ , one defines  $\widehat{a}^1(\xi) := e^{-i\eta\xi} \widehat{a}(\xi + 2\pi\gamma)$  and  $\widehat{\tilde{a}}^1(\xi) := e^{-i\eta\xi} \widehat{\tilde{a}}(\xi + 2\pi\gamma)$ , where  $\eta \in \mathbb{Z}^2$  satisfies



$\eta \cdot \gamma = 1/2$ . It is not an easy task to construct nonseparable orthogonal masks with desirable properties for a general dilation matrix. For example, for the dilation matrix  $Q_{\sqrt{2}}$  in (1), it is not known so far in the literature [9] whether there is a compactly supported  $C^1$   $Q_{\sqrt{2}}$ -refinable function with orthonormal shifts. See [1,2,9,25,27,30,32,50,51,53,57] for more details on construction of orthogonal masks and orthonormal refinable functions.

However, for any dilation matrix, “separable” orthogonal masks with arbitrarily high orders of sum rules can be easily obtained via the projection method. Let  $M$  be a  $2 \times 2$  dilation matrix. Then  $M = E \operatorname{diag}(d_1, d_2) F$  for some integer matrices  $E$  and  $F$  such that  $|\det E| = |\det F| = 1$  and  $d_1, d_2 \in \mathbb{N}$ . Let  $a$  be a tensor product orthogonal mask with the diagonal dilation matrix  $\operatorname{diag}(d_1, d_2)$  satisfying the sum rules of any preassigned order  $J$ . Then the projected mask  $Ea$  (that is,  $(Ea)_k := a_{E^{-1}k}$ ,  $k \in \mathbb{Z}^2$ ) is an orthogonal mask with dilation  $M$  and  $sr(Ea, M) \geq J$ . See [27, Corollary 3.4.] for more detail.

Using the high-pass filters in the matrix extension problem, for a given mask  $a$ , dual masks  $\tilde{a}$  (without any guaranteed order of sum rules of the dual masks) of  $a$  can be obtained by the lifting scheme in [63] and the stable completion method in [3]. Special families of dual masks with sum rules can be obtained by the convolution method in [24, Proposition 3.7.] and [43], but the constructed dual masks generally have longer supports with respect to their orders of sum rules. Another method, which we shall cite here, for constructing all finitely supported dual masks with any preassigned orders of sum rules is the CBC (coset by coset) algorithm proposed in [5,25,26].

For  $\mu = (\mu_1, \mu_2)$  and  $\nu = (\nu_1, \nu_2)$ , we say  $\nu \leq \mu$  if  $\nu_1 \leq \mu_1$  and  $\nu_2 \leq \mu_2$ . Also we denote  $|\mu| = \mu_1 + \mu_2$ ,  $\mu! := \mu_1! \mu_2!$  and  $x^\mu := x_1^{\mu_1} x_2^{\mu_2}$  for  $x = (x_1, x_2)$ . We denote  $\Omega_M$  a complete set of representatives of distinct cosets of  $\mathbb{Z}^2 / [M\mathbb{Z}^2]$  with  $0 \in \Omega_M$ . The following result is from [25].

**Theorem 13 (CBC Algorithm)** *Let  $M$  be a  $2 \times 2$  dilation matrix and  $a$  be an interpolatory mask with dilation  $M$ . Let  $J$  be any preassigned positive integer.*

1. Compute the quantities  $h_\mu^a$ ,  $|\mu| < J$ , from the mask  $a$  by the recursive formula:

$$h_\mu^a := \delta_\mu - \sum_{0 \leq \nu < \mu} (-1)^{|\mu-\nu|} \frac{\mu!}{\nu!(\mu-\nu)!} h_\nu^a \sum_{k \in \mathbb{Z}^2} \bar{a}_k k^{\mu-\nu}.$$

2. For every nonzero coset  $\gamma \in \Omega_M \setminus \{0\}$  and pre-selected subsets  $E_\gamma \subset \mathbb{Z}^2$ , construct  $\tilde{a}$  on the coset  $\gamma + M\mathbb{Z}^2$ , with  $\tilde{a}_{\gamma+Mk} = 0$  for  $k \in \mathbb{Z}^2 \setminus E_\gamma$ , satisfying the equation:

tion:

$$\sum_{k \in E_\gamma} \tilde{a}_{\gamma+Mk} (\gamma + Mk)^\mu = |\det M|^{-1} h_\mu^a, \quad |\mu| < J.$$

3. Construct  $\tilde{a}$  at the zero coset  $M\mathbb{Z}^2$  by

$$\begin{aligned} \tilde{a}_{Mj} &:= |\det M|^{-1} \delta_k \\ &- \sum_{\gamma \in \Omega_M \setminus \{0\}} \sum_{k \in E_\gamma} \overline{a_{Mk-Mj+\gamma}} \tilde{a}_{\gamma+Mk}, \quad j \in \mathbb{Z}^2. \end{aligned}$$

Then mask  $\tilde{a}$  is a dual mask of  $a$  and satisfies the sum rules of order  $J$  with dilation  $M$ .

For a given interpolatory mask  $a$  and an arbitrary integer  $J$ , in fact any finitely supported dual mask  $\tilde{a}$  of  $a$ , with  $\tilde{a}$  satisfying the sum rules of order  $J$ , must come from the above CBC algorithm. The CBC algorithm also works well with symmetry. When  $a$  is symmetric, a dual mask  $\tilde{a}$  with symmetry and sum rules can be easily obtained via the CBC algorithm by appropriately selecting the sets  $E_\gamma$ . A general CBC algorithm is given in [26, Page 33 and Theorem 3.4.] for constructing finitely supported dual masks  $\tilde{a}$  with arbitrarily high orders of sum rules as long as the primal mask  $a$  possesses at least one finitely supported dual mask which may not have any order of sum rules. See [5,25,26] for more details on the CBC algorithm.

The projection method can be used to obtain some optimal properties of bivariate orthonormal and biorthogonal wavelets ([25,28,34,35]).

**Theorem 14** *Let  $M = dI_2$  and  $(a, \tilde{a})$  be a pair of dual masks with dilation  $M$ . Let  $P = [1, 0]$ . Suppose that the mask  $a$  is projectable, that is,  $\hat{a}(\xi, \gamma\pi/d) = 0$  for all  $\xi \in \mathbb{R}$  and  $\gamma = 1, \dots, d-1$ . Then  $(Pa, P\tilde{a})$  is also a pair of dual masks with dilation  $d$ . If (21) holds for  $\phi$  and  $\tilde{\phi}$ , where  $\phi$  and  $\tilde{\phi}$  are standard  $M$ -refinable functions with masks  $a$  and  $\tilde{a}$ , then (21) still holds with  $\phi$  and  $\tilde{\phi}$  being replaced by  $P\phi$  and  $P\tilde{\phi}$ . Moreover,  $v_p(\phi) \leq v_p(P\phi)$ ,  $v_p(\tilde{\phi}) \leq v_p(P\tilde{\phi})$ ,  $v_p(a, dI_2) \leq v_p(Pa, d)$  and  $v_p(\tilde{a}, dI_2) \leq v_p(P\tilde{a}, d)$  for all  $1 \leq p \leq \infty$ .*

Let us present a simple example to illustrate the CBC algorithm and the projection method on construction and analysis of bivariate biorthogonal wavelets.

**Example 15** Let  $M = 2I_2$  and  $a$  be the interpolatory mask given in (18). Note that the mask  $a$  is projectable. Let  $P = [1, 0]$ . Then  $\widehat{Pa}(\xi) = (2 - \cos(\xi)) \cos^4(\xi/2)$ . By the CBC algorithm, we see that the shortest symmetric dual mask  $b$  of  $Pa$  such that  $b$  satisfies the sum rules of order at least one is supported on  $[-4, 4]$  and is uniquely given by  $\hat{b}(\xi) := \cos^2(\xi/2)[14 - 5\cos(\xi) - 2\cos(2\xi) + \cos(3\xi)]/8$ .

By calculation, we have  $v_2(b, 2) \approx 0.59322$ . Take  $J = 2$  since  $sr(b, 2) = 2$ . By the CBC algorithm, any  $D_6$ -symmetric dual mask  $\tilde{a}$  of  $a$ , such that  $\tilde{a}$  is supported inside  $[-4, 4]^2$  and  $sr(\tilde{a}, 2I_2) = 2$ , must take the form

$$\frac{1}{64} \begin{bmatrix} 0 & 0 & 0 & 0 & t_2 \\ 0 & 0 & 0 & 0 & -t_3 \\ 0 & 0 & 1-2t_2 & t_3 & t_1 \\ 0 & 0 & t_3 & t_4 & 28+4t_1 \\ t_2 & -t_3 & t_1 & 28+4t_1 & 10-6t_1+6t_2 \\ 0 & t_3 & t_4 & 28+4t_1 & 28+4t_1 \\ 1-2t_2 & t_3 & t_1 & t_4 & t_1 \\ 0 & -t_3 & t_3 & t_3 & -t_3 \\ t_2 & 0 & 1-2t_2 & 0 & t_2 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1-2t_2 & 0 & t_2 \\ t_3 & t_3 & -t_3 & 0 \\ t_4 & t_1 & t_3 & 1-2t_2 \\ 28+4t_1 & t_4 & t_3 & 0 \\ 28+4t_1 & t_1 & -t_3 & t_2 \\ t_4 & t_3 & 0 & 0 \\ t_3 & 1-2t_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

with  $t_3 := 10 + 2t_1 + 4t_2$  and  $t_4 := -30 - 6t_1 - 4t_2$ , where  $\tilde{a}_{(0,0)} = 10 - 6t_1 + 6t_2$ . By Theorem 14 and  $P\tilde{a} = b$ , we have  $v_2(\tilde{a}, 2I_2) \leq v_2(b, 2) \approx 0.59322$ . Take  $t_1 = -17/4$  and  $t_2 = -1/2$ , by calculation, we have  $v_2(\tilde{a}, 2I_2) \approx 0.56715$ . So, the conditions in Theorem 11 are satisfied and  $\tilde{a}$  is near best in terms of the  $L_2$  smoothness with respect to the support of the dual mask.

### Bivariate Riesz Wavelets

Wavelet-based numerical algorithms have been successfully used in numerical solutions to partial differential equations and integral equations ([8,14]). The wavelets used in such wavelet-based methods are Riesz wavelets. For analysis on Riesz wavelets, see [10,33,40,41,42,47,56]. For applications based on Riesz wavelets, see [8,14,41,49].

**Theorem 16** Let  $M$  be a  $2 \times 2$  dilation matrix with  $m := |\det M|$  and  $a$  be a finitely supported bivariate mask. Let  $\hat{a}^1, \dots, \hat{a}^{m-1}$  be some  $2\pi$ -periodic trigonometric polynomials such that  $\det \mathcal{M}_{[\hat{a}, \hat{a}^1, \dots, \hat{a}^{m-1}]}(\xi) \neq 0$  for all  $\xi \in \mathbb{R}^2$  and  $\hat{a}^1(0) = \dots = \hat{a}^{m-1}(0) = 0$ . Let  $\phi$  be the standard refinable function associated with mask  $a$  and dilation  $M$ . If  $v_2(a, M) > 0$  and  $v_2(\hat{a}, M) > 0$ , where  $\hat{a}$  is the  $(1, 1)$ -entry of the matrix  $[\mathcal{M}_{[\hat{a}, \hat{a}^1, \dots, \hat{a}^{m-1}]}(\xi)]^{-1}$ , then  $\{\psi^1, \dots, \psi^{m-1}\}$ , which are defined in (25), generates a Riesz  $M$ -wavelet basis in  $L_2(\mathbb{R}^2)$ .

**Example 17** Let  $M = 2I_2$  and  $\hat{a}(\xi_1, \xi_2) := \cos^2(\xi_1/2) \cos^2(\xi_2/2) \cos^2(\xi_1/2 + \xi_2/2)$  in Example 4. Then

$sr(a, 2I_2) = 4$  and  $a$  is  $D_6$ -symmetric with center 0. Define ([41,60])

$$\begin{aligned} \hat{a}^1(\xi_1, \xi_2) &:= e^{-i(\xi_1 + \xi_2)} \overline{\hat{a}(\xi_1 + \pi, \xi_2)}, \\ \hat{a}^2(\xi_1, \xi_2) &:= e^{-i\xi_2} \overline{\hat{a}(\xi_1, \xi_2 + \pi)}, \\ \hat{a}^3(\xi_1, \xi_2) &:= e^{-i\xi_1} \overline{\hat{a}(\xi_1 + \pi, \xi_2 + \pi)}. \end{aligned}$$

Then all the conditions in Theorem 16 are satisfied and  $\{\psi^1, \psi^2, \psi^3\}$  generates a Riesz  $2I_2$ -wavelet basis in  $L_2(\mathbb{R}^2)$  ([41]). This Riesz wavelet derived from the Loop scheme has been used in [49] for mesh compression in computer graphics with impressive performance.

### Pairs of Dual Wavelet Frames

In this section, we mention a method for constructing bivariate dual wavelet frames. The following Oblique Extension Principle (OEP) has been proposed in [18] (and independently in [7]). Also see [17]) for constructing pairs of dual wavelet frames.

**Theorem 18** Let  $M$  be a  $2 \times 2$  dilation matrix and  $a, \tilde{a}$  be two finitely supported masks. Let  $\phi, \tilde{\phi}$  be the standard  $M$ -refinable functions with masks  $a$  and  $\tilde{a}$ , respectively, such that  $\phi, \tilde{\phi} \in L_2(\mathbb{R}^2)$ . If there exist  $2\pi$ -periodic trigonometric polynomials  $\Theta, \hat{a}^1, \dots, \hat{a}^L, \tilde{a}^1, \dots, \tilde{a}^L$  such that  $\Theta(0) = 1$ ,  $\hat{a}^1(0) = \dots = \hat{a}^L(0) = \tilde{a}^1(0) = \dots = \tilde{a}^L(0) = 0$ , and

$$\begin{aligned} \mathcal{M}_{[\Theta(M^T) \hat{a}, \hat{a}^1, \dots, \hat{a}^L]}(\xi) \overline{\mathcal{M}_{[\tilde{a}, \tilde{a}^1, \dots, \tilde{a}^L]}(\xi)}^T \\ = \text{diag}(\Theta(\xi + \gamma_0), \dots, \Theta(\xi + \gamma_{m-1})), \end{aligned} \quad (26)$$

where  $\{\gamma_0, \dots, \gamma_{m-1}\} = \Gamma_{M^T}$  with  $\gamma_0 = 0$  in Theorem 11. Then  $(\{\psi^1, \dots, \psi^L\}, \{\tilde{\psi}^1, \dots, \tilde{\psi}^L\})$ , defined in (25), generates a pair of dual  $M$ -wavelet frames in  $L_2(\mathbb{R}^2)$ .

For dimension one, many interesting tight wavelet frames and pairs of dual wavelet frames in  $L_2(\mathbb{R})$  have been constructed via the OEP method in the literature, for more details, see [7,17,18,24,30,52,61,62] and references therein. The application of the OEP in high dimensions is much more difficult, mainly due to the matrix extension problem in (26).

The projection method can also be used to obtain pairs of dual wavelet frames ([34,35]).

**Theorem 19** Let  $M$  be an  $r \times r$  dilation matrix and  $N$  be an  $s \times s$  dilation matrix with  $r \leq s$ . Let  $P$  be an  $r \times s$  integer matrix of rank  $r$  such that  $MP = PN$  and  $P^T(\mathbb{Z}^r \setminus [M^T \mathbb{Z}^r]) \subseteq \mathbb{Z}^s \setminus [N^T \mathbb{Z}^s]$ . Let  $\psi^1, \dots, \psi^r, \tilde{\psi}^1, \dots, \tilde{\psi}^L$  be compactly supported functions in  $L_2(\mathbb{R}^s)$  such that

$$v_2(\psi^\ell) > 0 \quad \text{and} \quad v_2(\tilde{\psi}^\ell) > 0 \quad \forall \ell = 1, \dots, L.$$

If  $(\{\psi^1, \dots, \psi^L\}, \{\tilde{\psi}^1, \dots, \tilde{\psi}^L\})$  generates a pair of dual  $N$ -wavelet frames in  $L_2(\mathbb{R}^s)$ , then  $(\{P\psi^1, \dots, P\psi^L\}, \{P\tilde{\psi}^1, \dots, P\tilde{\psi}^L\})$  generates a pair of dual  $M$ -wavelet frames in  $L_2(\mathbb{R}^r)$ .

**Example 20** Let  $\mathcal{E}$  be an  $r \times s$  (direction) integer matrix such that  $\mathcal{E}^T(\mathbb{Z}^r \setminus [2\mathbb{Z}^r]) \subseteq \mathbb{Z}^s \setminus [2\mathbb{Z}^s]$ . Let  $M = 2I_r$  and  $N = 2I_s$ . Let  $\{\psi^1, \dots, \psi^{2^s-1}\}$  be the generators of the tensor product Haar orthonormal wavelet in dimension  $s$ , derived from the Haar orthonormal refinable function  $\phi := \chi_{[0,1]^s}$ . Then by Theorem 19,  $\{\mathcal{E}\psi^1, \dots, \mathcal{E}\psi^{2^s-1}\}$  generates a tight  $2I_r$ -wavelet frame in  $L_2(\mathbb{R}^r)$  and all the projected wavelet functions are derived from the refinable box spline function  $\mathcal{E}\phi = M_{\mathcal{E}}$ .

### Future Directions

There are still many challenging problems on bivariate wavelets. In the following, we only mention a few here.

1. For any  $2 \times 2$  dilation matrix  $M$ , e.g.,  $M = Q_{\sqrt{2}}$  in (1), can one always construct a family of MRA compactly supported orthonormal  $M$ -wavelet bases with arbitrarily high smoothness (and with symmetry if  $|\det M| > 2$ )?
2. The matrix extension problem in (23) for orthogonal masks. That is, for a given orthogonal mask  $a$  with dilation  $M$ , find finitely supported high-pass filters  $a^1, \dots, a^{m-1}$  (with symmetry if possible) such that (23) holds with  $\tilde{a} = a$  and  $\tilde{a}^\ell = a^\ell$ .
3. The matrix extension problem in (23) for a given pair of dual masks with symmetry. That is, for a given pair of symmetric dual masks  $a$  and  $\tilde{a}$  with dilation  $M$ , find finitely supported symmetric high-pass filters  $a^1, \dots, a^{m-1}, \tilde{a}^1, \dots, \tilde{a}^{m-1}$  such that (23) is satisfied.
4. Directional bivariate wavelets. In order to handle edges of different orientations in images, directional wavelets are of interest in applications. See [13,55] and many references on this topic.

### Bibliography

#### Primary Literature

1. Ayache A (2001) Some methods for constructing nonseparable, orthonormal, compactly supported wavelet bases. *Appl Comput Harmon Anal* 10:99–111
2. Belogay E, Wang Y (1999) Arbitrarily smooth orthogonal nonseparable wavelets in  $\mathbb{R}^2$ . *SIAM J Math Anal* 30:678–697
3. Carnicer JM, Dahmen W, Peña JM (1996) Local decomposition of refinable spaces and wavelets. *Appl Comput Harmon Anal* 3:127–153
4. Cavaretta AS, Dahmen W, Micchelli CA (1991) Stationary subdivision. *Mem Amer Math Soc* 93(453)1–186
5. Chen DR, Han B, Riemenschneider SD (2000) Construction of multivariate biorthogonal wavelets with arbitrary vanishing moments. *Adv Comput Math* 13:131–165
6. Chui CK (1992) An introduction to wavelets. Academic Press, Boston
7. Chui CK, He W, Stöckler J (2002) Compactly supported tight and sibling frames with maximum vanishing moments. *Appl Comput Harmon Anal* 13:224–262
8. Cohen A (2003) Numerical analysis of wavelet methods. North-Holland, Amsterdam
9. Cohen A, Daubechies I (1993) Nonseparable bidimensional wavelet bases. *Rev Mat Iberoamericana* 9:51–137
10. Cohen A, Daubechies I (1996) A new technique to estimate the regularity of refinable functions. *Rev Mat Iberoamericana* 12:527–591
11. Cohen A, Daubechies I, Feauveau JC (1992) Biorthogonal bases of compactly supported wavelets. *Comm Pure Appl Math* 45:485–560
12. Cohen A, Gröchenig K, Vilelmoes LF (1999) Regularity of multivariate refinable functions. *Constr Approx* 15:241–255
13. Cohen A, Schlenker JM (1993) Compactly supported bidimensional wavelet bases with hexagonal symmetry. *Constr Approx* 9:209–236
14. Dahmen W (1997) Wavelet and multiscale methods for operator equations. *Acta Numer* 6:55–228
15. Daubechies I (1988) Orthonormal bases of compactly supported wavelets. *Comm Pure Appl Math* 41:909–996
16. Daubechies I (1992) Ten Lectures on Wavelets, CBMS-NSF Series. SIAM, Philadelphia
17. Daubechies I, Han B (2004) Pairs of dual wavelet frames from any two refinable functions. *Constr Approx* 20:325–352
18. Daubechies I, Han B, Ron A, Shen Z (2003) Framelets: MRA-based constructions of wavelet frames. *Appl Comput Harmon Anal* 14:1–46
19. Dahlke S, Gröchenig K, Maass P (1999) A new approach to interpolating scaling functions. *Appl Anal* 72:485–500
20. de Boor C, Hollig K, Riemenschneider SD (1993) Box splines. Springer, New York
21. DeRose T, Forsey DR, Kobbelt L, Lounsbery M, Peters J, Schröder P, Zorin D (1998) Subdivision for Modeling and Animation, (course notes)
22. Dyn N, Levin D (2002) Subdivision schemes in geometric modeling. *Acta Numer* 11:73–144
23. Dyn N, Gregory JA, Levin D (1990) A butterfly subdivision scheme for surface interpolation with tension control. *ACM Trans Graph* 9:160–169
24. Han B (1997) On dual wavelet tight frames. *Appl Comput Harmon Anal* 4:380–413
25. Han B (2000) Analysis and construction of optimal multivariate biorthogonal wavelets with compact support. *SIAM Math Anal* 31:274–304
26. Han B (2001) Approximation properties and construction of Hermite interpolants and biorthogonal multiwavelets. *J Approx Theory* 110:18–53
27. Han B (2002) Symmetry property and construction of wavelets with a general dilation matrix. *Lin Algeb Appl* 353:207–225
28. Han B (2002) Projectable multivariate refinable functions and biorthogonal wavelets. *Appl Comput Harmon Anal* 13:89–102
29. Han B (2003) Computing the smoothness exponent of a symmetric multivariate refinable function. *SIAM J Matrix Anal Appl* 24:693–714

30. Han B (2003) Compactly supported tight wavelet frames and orthonormal wavelets of exponential decay with a general dilation matrix. *J Comput Appl Math* 155:43–67
31. Han B (2003) Vector cascade algorithms and refinable function vectors in Sobolev spaces. *J Approx Theory* 124:44–88
32. Han B (2004) Symmetric multivariate orthogonal refinable functions. *Appl Comput Harmon Anal* 17:277–292
33. Han B (2006) On a conjecture about MRA Riesz wavelet bases. *Proc Amer Math Soc* 134:1973–1983
34. Han B (2006) The projection method in wavelet analysis. In: Chen G, Lai MJ (eds) *Modern Methods in Mathematics*. Nashboro Press, Brentwood, pp. 202–225
35. Han B (2008) Construction of wavelets and framelets by the projection method. *Int J Appl Math Appl* 1:1–40
36. Han B, Jia RQ (1998) Multivariate refinement equations and convergence of subdivision schemes. *SIAM J Math Anal* 29:1177–1199
37. Han B, Jia RQ (1999) Optimal interpolatory subdivision schemes in multidimensional spaces. *SIAM J Numer Anal* 36:105–124
38. Han B, Jia RQ (2002) Quincunx fundamental refinable functions and quincunx biorthogonal wavelets. *Math Comp* 71:165–196
39. Han B, Jia RQ (2006) Optimal  $C^2$  two-dimensional interpolatory ternary subdivision schemes with two-ring stencils. *Math Comp* 75:1287–1308
40. Han B, Jia RQ (2007) Characterization of Riesz bases of wavelets generated from multiresolution analysis. *Appl Comput Harmon Anal* 23:321–345
41. Han B, Shen Z (2005) Wavelets from the Loop scheme. *J Fourier Anal Appl* 11:615–637
42. Han B, Shen Z (2006) Wavelets with short support. *SIAM J Math Anal* 38:530–556
43. Ji H, Riemenschneider SD, Shen Z (1999) Multivariate compactly supported fundamental refinable functions, duals, and biorthogonal wavelets. *Stud Appl Math* 102:173–204
44. Jia RQ (1998) Approximation properties of multivariate wavelets. *Comp Math* 67:647–665
45. Jia RQ (1999) Characterization of smoothness of multivariate refinable functions in Sobolev spaces. *Trans Amer Math Soc* 351:4089–4112
46. Jia RQ, Micchelli CA (1991) Using the refinement equation for the construction of pre-wavelets II: Power of two. In: Laurent PJ, Le Méhauté A, Schumaker LL (eds) *Curves and Surfaces*. Academic Press, New York, pp. 209–246
47. Jia RQ, Wang JZ, Zhou DX (2003) Compactly supported wavelet bases for Sobolev spaces. *Appl Comput Harmon Anal* 15:224–241
48. Jiang QT (1998) On the regularity of matrix refinable functions. *SIAM J Math Anal* 29:1157–1176
49. Khodakovsky A, Schröder P, Sweldens W (2000) Progressive geometry compression. *Proc SIGGRAPH*
50. Kovačević J, Vetterli M (1992) Nonseparable multidimensional perfect reconstruction filter banks and wavelet bases for  $\mathbb{R}^n$ . *IEEE Trans Inform Theory* 38:533–555
51. Lai MJ (2006) Construction of multivariate compactly supported orthonormal wavelets. *Adv Comput Math* 25:41–56
52. Lai MJ, Stöckler J (2006) Construction of multivariate compactly supported tight wavelet frames. *Appl Comput Harmon Anal* 21:324–348
53. Lawton W, Lee SL, Shen Z (1997) Stability and orthonormality of multivariate refinable functions. *SIAM J Math Anal* 28:999–1014
54. Lawton W, Lee SL, Shen Z (1998) Convergence of multidimensional cascade algorithm. *Numer Math* 78:427–438
55. Le Pennec E, Mallat S (2005) Sparse geometric image representations with bandelets. *IEEE Trans Image Process* 14:423–438
56. Lorentz R, Oswald P (2000) Criteria for hierarchical bases in Sobolev spaces. *Appl Comput Harmon Anal* 8:32–85
57. Maass P (1996) Families of orthogonal two-dimensional wavelets. *SIAM J Math Anal* 27:1454–1481
58. Mallat S (1998) *A wavelet tour of signal processing*. Academic Press, San Diego
59. Riemenschneider SD, Shen Z (1997) Multidimensional interpolatory subdivision schemes. *SIAM J Numer Anal* 34:2357–2381
60. Riemenschneider SD, Shen ZW (1992) Wavelets and pre-wavelets in low dimensions. *J Approx Theory* 71:18–38
61. Ron A, Shen Z (1997) Affine systems in  $L_2(\mathbb{R}^d)$ : the analysis of the analysis operator. *J Funct Anal* 148:408–447
62. Ron A, Shen Z (1997) Affine systems in  $L_2(\mathbb{R}^d)$  II: dual systems. *J Fourier Anal Appl* 3:617–637
63. Sweldens W (1996) The lifting scheme: a custom-design construction of biorthogonal wavelets. *Appl Comput Harmon Anal* 3:186–200

## Books and Reviews

Cabrelli C, Heil C, Molter U (2004) Self-similarity and multiwavelets in higher dimensions. *Mem Amer Math Soc* 170(807):1–82

## Boolean Modeling of Biological Networks

PETER KRAWITZ<sup>1</sup>, ILYA SHMULEVICH<sup>2</sup>

<sup>1</sup> Fakultät für Physik, Ludwig Maximilians Universität, München, Germany

<sup>2</sup> Institute for Systems Biology, Department of Bioengineering University of Washington, Seattle, USA

## Article Outline

[Glossary](#)

[Definition of the Subject](#)

[Introduction](#)

[Theoretical Aspects of Boolean Networks](#)

[Boolean Models of Biological Networks](#)

[Network Inference](#)

[Future Directions](#)

[Bibliography](#)

## Glossary

**Connectivity matrix** In a Boolean network model, the information about *who* is interacting with *whom* is stored in the connectivity matrix. This static wiring diagram can be described by parameters from graph the-



ory, such as the degree distributions of its elements or nodes.

**Activity, sensitivity** Boolean functions define *how* elements of a network interact. Activity and sensitivity are parameters describing the dynamic properties of single network elements or nodes. The activity of a node measures its influence on another node and conversely, the sensitivity of a node measures how sensitive it is to random changes in one of its inputs.

**Phase transition and criticality** Regarding the propagation of perturbations, one can distinguish between an ordered phase, where perturbations will not influence the network dynamics and an chaotic phase, where small perturbations will increase over time. Networks that operate at the transition of these two phases, such that perturbations are neither attenuated nor amplified on average, are called critical.

**Attractors** In Boolean networks, dynamically stable patterns are called attractors. In their biological context, attractors may be interpreted as different cell types in models of cell differentiation or as cell states in models of cell signaling. Analyzing the influence of perturbations on attractor dynamics is thus of pivotal biological interest.

**Relevant and irrelevant nodes** Only the perturbation of certain elements will move the system from one attractor to another. These dynamically relevant nodes are potential lever points for system intervention.

## Definition of the Subject

Boolean Networks (BN) are a class of discrete dynamical systems that can be characterized by the interaction of a set of Boolean variables. Random Boolean Networks (RBN), which are ensembles of random network structures, were first introduced by Stuart Kauffman in 1969 as a simple model class for studying dynamical properties of gene regulatory networks [1,2]. Since then, Boolean Networks have been used as generic models for dynamics of complex systems of interacting entities, such as social and economic networks, neural networks, as well as gene and protein interaction networks. Despite their conceptual simplicity, Boolean Networks exhibit complex nonlinear behaviors that are, to this day, a challenging object of investigation for theoretical physicists and mathematicians. Further, a discretization of gene expression is often regarded as an experimentally justifiable simplification [3,4], making Boolean network models attractive tools for the application oriented biologist. In this chapter, the theoretical findings for RBNs of the last decade are reviewed and

the value of Boolean network models for the understanding of biological networks is illustrated.

## Introduction

Since the early 1960s, the hypothesis that different dynamically stable patterns of gene expression are the basic key to understanding the diversity of functionally and morphologically different cells of a single organism began to prevail and today, a gene expression profile of a cell is often regarded as a sufficient definition of its state [5]. In 1969, long before accurate gene numbers or the very order of the magnitude of genome sizes in higher developed organisms were known, Stuart A. Kauffman introduced random Boolean networks as a generic model to study and explain how an increasing number of dynamical stable states can arise from a growing network size [1]. The now classical model introduced by Kauffman consists of  $n$  binary variables  $x_i$  (also referred to as nodes or genes), that are each connected to  $k$  randomly chosen other nodes. A Boolean function  $f_i$  that specifies a value for every possible input combination,  $f_i : \{0, 1\}^k \rightarrow \{0, 1\}$ , is assigned randomly to each node  $x_i$ . For  $k$  input variables, there are  $2^k$  different functions. The state of a node  $x_i$  is then updated by calculating the Boolean function value of its  $k$  input nodes:  $x_i(t+1) = f_i(x_{i_1}(t), \dots, x_{i_k}(t))$ .

Under synchronous updating, a wiring diagram (i. e. which nodes are connected to which other nodes) and the Boolean functions assigned to every node define a deterministic system that assigns to any network state  $\mathbf{x}(t) \in \{0, 1\}^n$  a successor state  $\mathbf{x}(t+1) = F(\mathbf{x}(t))$ , where  $F : \{0, 1\}^n \rightarrow \{0, 1\}^n$  is a vector valued (multi-output) Boolean function that is comprised of the Boolean functions assigned to each of the nodes and that represents the state transition rules of the system. For a given initial state, a network thus defines a trajectory through the state space of size  $2^n$ . As this system is deterministic and the state space finite, a trajectory will eventually fall on a previously visited state and periodically circle through a certain number  $l$  of states from there on. These reoccurring and therefore dynamically stable states form an *attractor* of size  $l$  and may, in the context of gene regulatory networks, be interpreted as functionally salient states. For example, in the context of cellular differentiation, such attractors may correspond to cell types or cell fates [2]. All states from which the system eventually flows to a given attractor constitute that attractor's *basin of attraction*. The entire state space is naturally divided into disjoint basins of attraction. Kauffman observed in his numerical simulations that for  $k = 2$ , the mean number of attractors scaled approximately with the square root of the system size – a scaling behavior that



seemed to agree with the increasing number of different cell types in higher organized organisms of larger genome sizes.

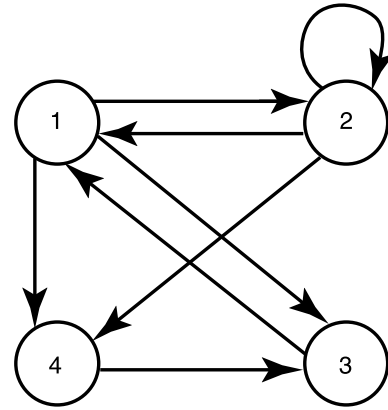
Although the scaling behavior of the mean number of attractors in such network ensembles as well as the estimated genome sizes are now outdated, Kauffman's ensemble approach to gene regulation was pioneering and is still valid [6]: define an ensemble of Boolean networks such that it fulfills certain known features of biological networks and then study random instances of these networks to learn more about general properties of these ensembles. This iterative and highly interdisciplinary process depends, on the one hand, on the input of new biological findings and might guide, on the other hand, the research of the experimentalist. The Kauffman ensembles of the parameter  $k = 2$ , for example, turned out to be "critical", lying at a phase transition between an ordered and a chaotic phase, in regard to the propagation of perturbations, and contributed to the hypothesis that living systems may be poised at the "edge of chaos" – a hypothesis for which experimental evidence has been mounting [7,8,9].

### Theoretical Aspects of Boolean Networks

In this section some key findings learned from the study of ensembles of Boolean networks will be illustrated. The understanding of dynamic properties of a network and how these are caused may help to identify the pivotal features in a biological Boolean network model. For a profound and detailed review of the theory of Random Boolean Networks, see also [10].

#### The Concept of Relevant Nodes

Given a particular starting state, the attractor to which the system will eventually flow starting from this state will be determined entirely by so-called *relevant nodes*. Some nodes of a network will always take on the same value regardless of the attractor on which the network finally settles down. These nodes establish the *frozen core* of a network. They are either updated by a constant Boolean function and are *frozen* to a fixed value right from the beginning or they freeze as the system transitions through its state space, in which case they are called *clamped*. Other nodes may have different values on different attractors, but their values are of no importance for determining the attractor to which the system will flow. These nodes are therefore called *non-frozen*, but *irrelevant*. Finally, there are the non-frozen *relevant* nodes that can be recursively defined as influencing other relevant nodes. Relevant nodes ultimately have an effect on their own state. The concept of classifying the nodes of a network



Boolean Modeling of Biological Networks, Figure 1

The connectivity graph of a Boolean network of four nodes

Boolean Modeling of Biological Networks, Table 1

Connectivity matrix of the example network

$x_i \backslash j$	1	2	3	4
$x_{i1}$	2	1	1	1
$x_{i2}$	3	2	4	2

Boolean Modeling of Biological Networks, Table 2

Boolean functions of the example network

$x_{i1}$	$x_{i2}$	$f_1$	$f_2$	$f_3$	$f_4$
0	0	0	1	0	0
0	1	0	0	1	1
1	0	1	1	1	1
1	1	1	1	0	1

into clamped, irrelevant, and relevant, regarding their importance for the attractor dynamics, is simple yet powerful for the understanding of Boolean network dynamics [11,12,13].

The example network, which is illustrated in Fig. (1) and whose connectivity matrix, Boolean functions and transition matrix are given in Tables 1–3, has two attractors, a fixed point,  $(x_1 x_2 x_3 x_4) = (1101)$  and an attractor of length two,  $(0101) \rightleftharpoons (1011)$ . Only the nodes 1 and 2 are relevant, whereas node 3 is non-frozen but irrelevant and node 4 is a clamped node, that freezes to the value  $x_4 = 1$  as soon as the system reaches an attractor.

Algorithmically, relevant nodes can often be identified by an iterative procedure:

First, links of frozen nodes to other nodes are cut, as this kind of input is not *variable* but *constant*.

Second, so-called 'degenerate' links have to be identified and removed from the connectivity matrix. In the example network a closer look at the Boolean function of node 1 reveals that it actually depends only on its first

**Boolean Modeling of Biological Networks, Table 3**  
Transition matrix of the example network

<b>x</b>				<b>F(x)</b>			
$x_1$	$x_2$	$x_3$	$x_4$	$f_1$	$f_2$	$f_3$	$f_4$
0	0	0	0	0	1	0	0
0	0	0	1	0	1	1	0
0	0	1	0	0	1	0	0
0	0	1	1	0	1	1	0
0	1	0	0	1	0	0	1
0	1	0	1	1	0	1	1
0	1	1	0	1	0	0	1
0	1	1	1	1	0	1	1
1	0	0	0	0	1	1	1
1	0	0	1	0	1	0	1
1	0	1	0	0	1	1	1
1	0	1	1	0	1	0	1
1	1	0	0	1	1	1	1
1	1	0	1	1	1	0	1
1	1	1	0	1	1	1	1
1	1	1	1	1	1	0	1

variable, which is node 2. Therefore link  $3 \rightarrow 1$  can be removed from the connectivity matrix.

Third, nodes that do not influence other nodes are irrelevant and can be removed from the network.

This procedure is repeated until no further nodes can be removed from the network. In our example network node 3 can be removed after cutting link  $3 \rightarrow 1$ . In a second round node 4 can be removed, which leaves us with the two relevant nodes 1 and 2. Some nodes identified by this method, however, are actually not relevant, but *freeze* to a constant value, as the following example illustrates:  $f_1 = x_1 \vee x_2$ ,  $f_2 = x_1 \vee \bar{x}_2$ . The properties of these self-freezing nodes were first described in [14].

### The Critical Notion of Criticality

Generally, a dynamic system is said to behave in an *ordered* or *robust* manner when similar initial conditions will lead to similar trajectories in the phase (state) space. A system is called *chaotic* when even infinitesimal differences will cause an increasing divergence in the ensuing trajectories. In Boolean networks, the difference between two states  $\mathbf{x}$  and  $\mathbf{x}'$  is usually quantified in terms of the normalized Hamming distance  $\eta$ :

$$\eta(\mathbf{x}, \mathbf{x}') = \frac{1}{n} \sum_{i=1}^n |x_i - x'_i|. \quad (1)$$

With this metric, the evolution of the overlap between two random states,  $a(t) = 1 - \eta(\mathbf{x}(t), \mathbf{x}'(t))$ , can be measured over time.

In the second half of the 1980s, Derrida, Flyvbjerg and others began to study the propagation of perturbations in Boolean networks with a mean field approach that had already successfully been applied in the theory of spin glasses [15,16,17]. Instead of considering the precise effects of all nodes on their linked nodes in a certain network instance, the mean field approach only deals with the average effect on a random node. If the overlap between two states,  $\mathbf{x}$  and  $\mathbf{x}'$ , is  $a(t)$  at a given time  $t$ , the probability that a random node  $i$  will receive the same input signal in the next step, is  $a(t)^k$ . If the Boolean function  $f_i$  assigned to the node is generated with *bias*  $p$ , meaning that the function value of a random input vector is 1 with probability  $p$ , then the probability that two random input vectors yield the same function value is  $P = p^2 + (1-p)^2$ . The expected overlap of the two states after one time step is given by:

$$a(t+1) = a(t)^k + P(1 - a(t))^k. \quad (2)$$

In the limit of  $t$  and  $n \rightarrow \infty$ ,  $a(t)$  asymptotically approaches the fixed point  $a^*$ , which obeys:

$$a^* = 1 - (1 - (a^*)^k)2p(1-p). \quad (3)$$

The critical transition curve between the ordered phase, where the overlap approaches  $a^* = 1$ , and a chaotic phase, where the overlap approaches a value  $a^* < 1$  is given by the relation

$$2kp(1-p) = 1. \quad (4)$$

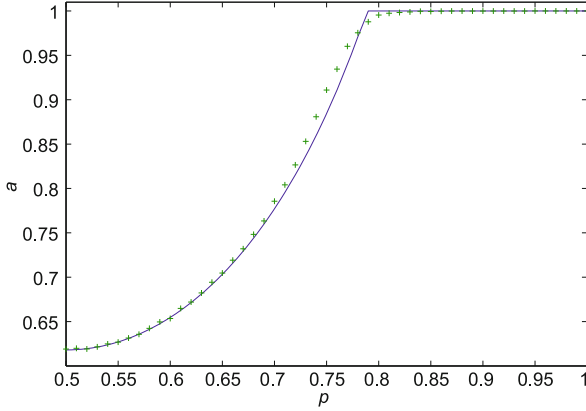
Figure 2 shows the averaged overlap  $a$  (+) of two random states of Hamming distance  $\eta(\mathbf{x}, \mathbf{x}') = 0.01$  after 200 time steps in network ensembles of  $n = 1000$  nodes and connectivity  $k = 3$  as a function of the bias  $p$ , as well as the results of the annealed approximation (blue line, Eq. 3). In ensembles with  $p \geq 0.7887$  the initial perturbation dies out and the overlap approaches  $a = 1$ .

Another order parameter of the phase transition, which is based on the concept of Boolean derivatives, is the network sensitivity  $s$  or the expected sensitivity of a random node [18,19]. The partial derivative of a Boolean function  $f$  with respect to  $x_j$  is defined as

$$\frac{\partial f}{\partial x_j} = f(\mathbf{x}^{(j,0)}) \oplus f(\mathbf{x}^{(j,1)}), \quad (5)$$

where  $\oplus$  is addition modulo 2 and  $\mathbf{x}^{(j,\sigma)} = (x_1, \dots, x_{j-1}, \sigma, x_{j+1}, \dots, x_k)$ ,  $\sigma = 0, 1$ . The  $n \times n$  Jacobian matrix  $F'_{ij}$  of a network at time  $t$  is then defined as

$$F'_{ij}(\mathbf{x}(t)) = \frac{\partial f_i(\mathbf{x}(t))}{\partial x_j}. \quad (6)$$



**Boolean Modeling of Biological Networks, Figure 2**

The averaged overlap  $a$ , of two random states  $\mathbf{x}, \mathbf{x}'$  with  $\eta(\mathbf{x}, \mathbf{x}') = 0.01$  in network ensembles  $B(n, k, p)$ ,  $n = 1000$ ,  $k = 3$  and  $p = [0.5, 1]$ , after 200 time steps is indicated by (+). The blue line shows the annealed approximation (Eq. 3)

The behavior of a perturbation vector  $\mathbf{x}$  is then determined by the successive products of the Jacobian matrices

$$Q(t) = F'(\mathbf{x}(0)) \cdot F'(\mathbf{x}(1)) \cdot \dots \cdot F'(\mathbf{x}(t)). \quad (7)$$

An initial perturbation disappears if  $Q(t)$  converges to a matrix  $Q^*$  consisting of zeros only.

In a mean field approach, the deterministic matrix  $F'(\mathbf{x}(t))$  is substituted by a random matrix  $R$  of the same form. The probability for a random Boolean function to have a partial derivative equal to 1 is equal to the probability that a change in its input  $x_j$  causes a change in the function value. In an ensemble with a given bias  $p$  this happens with probability  $2p(1-p)$ . The mean number of 'ones' in each row is accordingly  $2kp(1-p)$  and  $Q(t+1)$  becomes  $Q(t+1) = RQ(t)$ . The percentage  $q(t)$  of zeros in  $Q(t)$  evolves in the limit of large  $n$  as

$$q(t+1) = \lim_{n \rightarrow \infty} \left[ 1 - \frac{2p(1-p)k}{n} (1 - q(t)) \right]^n \quad (8)$$

$$= e^{-(1-q(t))2p(1-p)k}. \quad (9)$$

By analyzing this equation for the fixed point  $q^* = 1$  one finds again the critical transition between the ordered phase, where perturbations die out over time, and the chaotic phase, where even small perturbations spread, given by:

$$\left. \frac{q(t+1)}{q(t)} \right|_{q^*=1} = 2kp(1-p) > 1.$$

The activity  $\alpha_j^{f_i}$  of a variable  $x_j$  in a function  $f_i$  of  $k$  variables is defined as the mean value of the partial derivatives

over all possible inputs of this function

$$\alpha_j^{f_i} = \frac{1}{2^k} \sum_{\mathbf{x} \in \{0,1\}^k} \frac{\partial f_i(\mathbf{x})}{\partial x_j}. \quad (10)$$

It gives the probability that a change in the value of this variable will cause a change in the function's value. The sensitivity  $s^i$  of a node is the sum of the activities of its variables  $s^i = \sum_{j=1}^k \alpha_j^{f_i}$  and indicates how likely the function value is to change under a random change in its arguments. Critical random Boolean networks have an expected network sensitivity of  $s = \frac{1}{n} \sum_{i=1}^n s^i = 1$ , a perturbation in one node will on average change the value of one other node in the next time step. It is an appealing feature of the sensitivity parameter that it is independent of from the network architecture and entirely determined by the Boolean functions in a given network. The sensitivity of a Boolean model of a biological network and the activities of its nodes are easy to compute and may help conceptualize the dynamic characteristics of its elements. Using the mixing ratio of different classes of Boolean function as basis for the network dynamics the definition of criticality was further generalized in [20].

### Scaling Behaviors in RBNs

With increasing computational power, the mean number and length of attractors in larger network ensembles could be studied and the original hypothesis of a power law-like scaling of the mean number of attractors in critical networks was shown to be incorrect. In 2003, Samuelson and Troein showed by analytical means that the number of attractors grows superpolynomially in critical networks [21]. This astonishing result attracted new interest in the scaling behavior of other network parameters. Kaufman, Mihaljev and Drossel showed that the number of relevant nodes,  $n_r$ , in a general class of critical networks increases in a power law manner with the network size,  $n_r \sim n^{1/3}$  [20,22]. A further result of this work is that relevant nodes of random critical ensembles are only rarely regulated by more than one other relevant node – a finding whose implications on the ensemble approach will have to be carefully considered.

Another entity that increases only in critical network ensembles is the recently introduced *basin entropy*. A Boolean network classifies its state space into several basins of attraction by assigning a unique attractor to every state. The size  $w_\rho$  of basin  $\rho$  is the number of states belonging to that basin (including the attractor states, which are by definition part of the basin). Thus, the basin sizes describe the complexity of the state space partition.

In the example network, for instance, the basin of the fixed point only drains one forth of all states, whereas the rest flow into the second attractor. The basin entropy,  $h = -\sum_{\rho=1}^m w_{\rho} \ln w_{\rho}$ , where  $m$  is the total number of basins/attractors, serves as a measure for the complexity of this classification process. Equivalently, the basin entropy measures the uncertainty of a system that is started in a random state. Surprisingly, basin entropy only increases in critical network ensembles, whereas it approaches a finite limit in the ordered and highly chaotic phase [23]. This suggests that the informationally optimal partition of the state space is achieved when the system is operating at the critical boundary between the ordered and disordered phases. For biological networks this means that their informational complexity will only grow with size when their elements are organized in a critical manner.

Under asynchronous updating schemes many attractors of the synchronous scheme become unstable [24,25]. For the scaling behaviour of attractors under asynchronous updating, the mean number of attractors increases only polynomially with system size [26].

### Modifications to the Classical Model

The fixed in-degree distribution and the synchronous updating scheme are clearly among the most evident shortcomings of the classical Kauffman model from a biological point of view. The discovery of topological similarities in a variety of complex networks [27] spurred interest in the dynamical implications of ‘scale-free’ architectures of Boolean network models in which the degree distribution follows a power law. Aldana et al. showed that Boolean networks with a scale-free topology give rise to a much broader class of networks that exhibit robust network dynamics [28]. Only recently, realistic data for in- and out degree distributions in gene regulatory networks became available through genome-wide localization analysis of transcription factors [29]. These topological findings can now be incorporated into more realistic random network ensembles.

Harris studied in an extensive literature search the frequencies of logical rules in eukaryotic gene transcription [30] and showed that canalizing functions occur far more often than randomly expected. A variable  $x_i$  is called *canalizing* for a certain value  $x_i = \sigma_i \in \{0, 1\}$ , if it determines the value  $\sigma_j \in \{0, 1\}$  of the function  $f_j(\dots, x_i = \sigma_i, \dots) = \sigma_j$ . In the example network for instance in function  $f_4$  both variables  $x_1, x_2$  are canalizing for  $\sigma = 1$ , whereas in function  $f_3$  none of the variables are canalizing, meaning that both variables are always necessary to determine the value of the function. Canalizing and

related classes of functions were studied in the context of the sources of order and remarkable robustness observed in genetic regulatory networks [31] as well as in terms of stability [32]. The knowledge about which logical rules actually do occur in biological networks is not only important for more realistic ensembles but might also help to make an educated guess in the modeling of a concrete biological circuit, when the updating rules for a gene are uncertain.

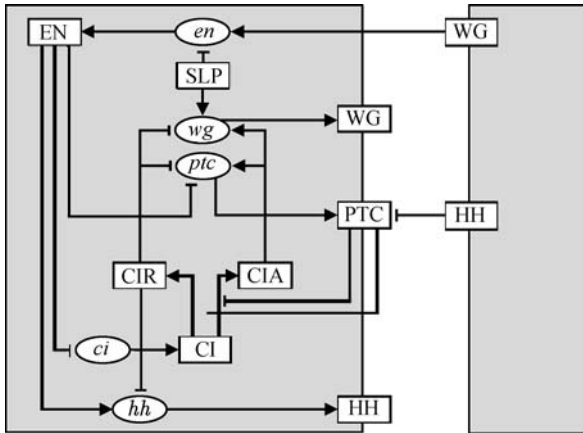
Although the synchronous updating scheme has been known to be biologically unrealistic, it has been applied in most analytical studies so far, as it is the only one that yields deterministic state transitions. Under asynchronous updating, many attractors of the synchronous scheme become unstable and decay into other dynamically stable cycles. However, the very definition of an attractor, except for fixed points remains difficult. Further, asynchronous updating is not per se GE biologically more realistic and has to be motivated carefully in every case not to fall victim to artefacts [33].

### Boolean Models of Biological Networks

Biological networks can often be modelled as logical circuits from well-known local interaction data in a straightforward way. This is clearly one of the advantages of the Boolean network approach. Though logical models may sometimes appear obvious and simplistic they help to understand the dynamic key properties of a regulatory process. They may also inspire the experimentalist to ask new question and to test them first in silico. The following three network models exemplify how Boolean networks may serve as useful in silico tools from cell differentiation processes to cell signaling.

#### The Segment Polarity Network of the Fruit Fly

The segment polarity genes represent the last step in the hierarchical cascade of gene families initiating the segmented body of the fruit fly *Drosophila melanogaster*. The stable maintenance of the segment polarity gene expression pattern is a crucial requirement in the further development of the embryo. The dynamics of the best characterized segment polarity genes have been studied in different modelling approaches in order to understand the stability and robustness of this pattern formation [33,34,35,36]. Albert and Othmer suggested a Boolean model of gene and protein interactions which is depicted in Fig. 3. This model correctly reproduces the characteristic expression patterns observed in the wild-type development. Further, distinct ‘knock down’ or ‘overexpression’ phenotypes can be simulated through the model by fixing



**Boolean Modeling of Biological Networks, Figure 3**

The network of interactions between the segment polarity genes. The *gray background layers* illustrate two neighboring cells indicating that some interactions in this network are inter-cellular [36]

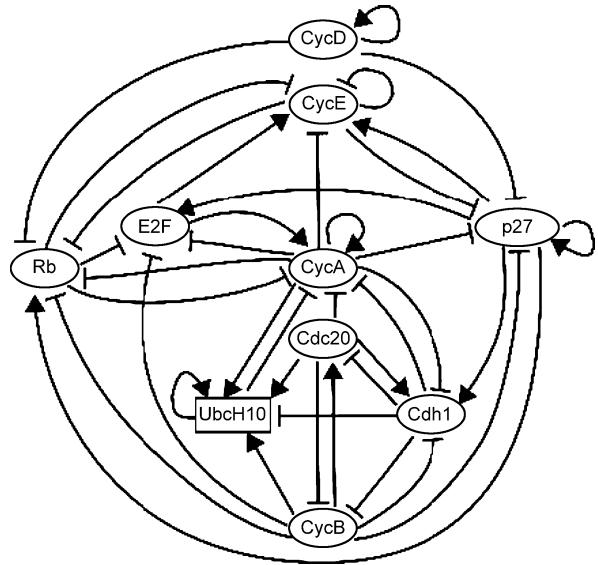
the state of the particular gene. An analysis of the robustness of the wild-type steady states under asynchronous updating schemes reveals that the model is very sensitive to temporal variations, whereas the wild-type development is stable as soon as certain differences in the temporal order of transcriptional and translational processes are incorporated into the updating procedure.

### Control of the Cell Cycle

Proliferating eukaryotic cells perform a series of highly regulated and coordinated actions collectively referred to as the cell cycle. As cancer development is associated with a loss of control over the cell cycle, an understanding of its regulatory dynamics is of paramount importance. The integration of the biological and mostly qualitative knowledge about interacting genes and proteins into a Boolean network model allows the analysis of the dynamic regulatory properties. Fauré et al. studied the effects of different updating schemes on the dynamics of a highly abstracted cell cycle control model [37]. Under the simplistic synchronous updating scheme the network gives rise to two different attractors, a stable state, matching the quiescent cell when growth factors are lacking and a complex dynamical cycle representing the cell cycle when the cyclin D complex is present (Fig. 4).

### T-Cell Receptor Signaling

T lymphocytes play a key role in the regulation of the adaptive immune system, as they are able to differentiate between self- and foreign-, potentially dangerous, antigens. T cells recognize foreign peptides that are presented



**Boolean Modeling of Biological Networks, Figure 4**

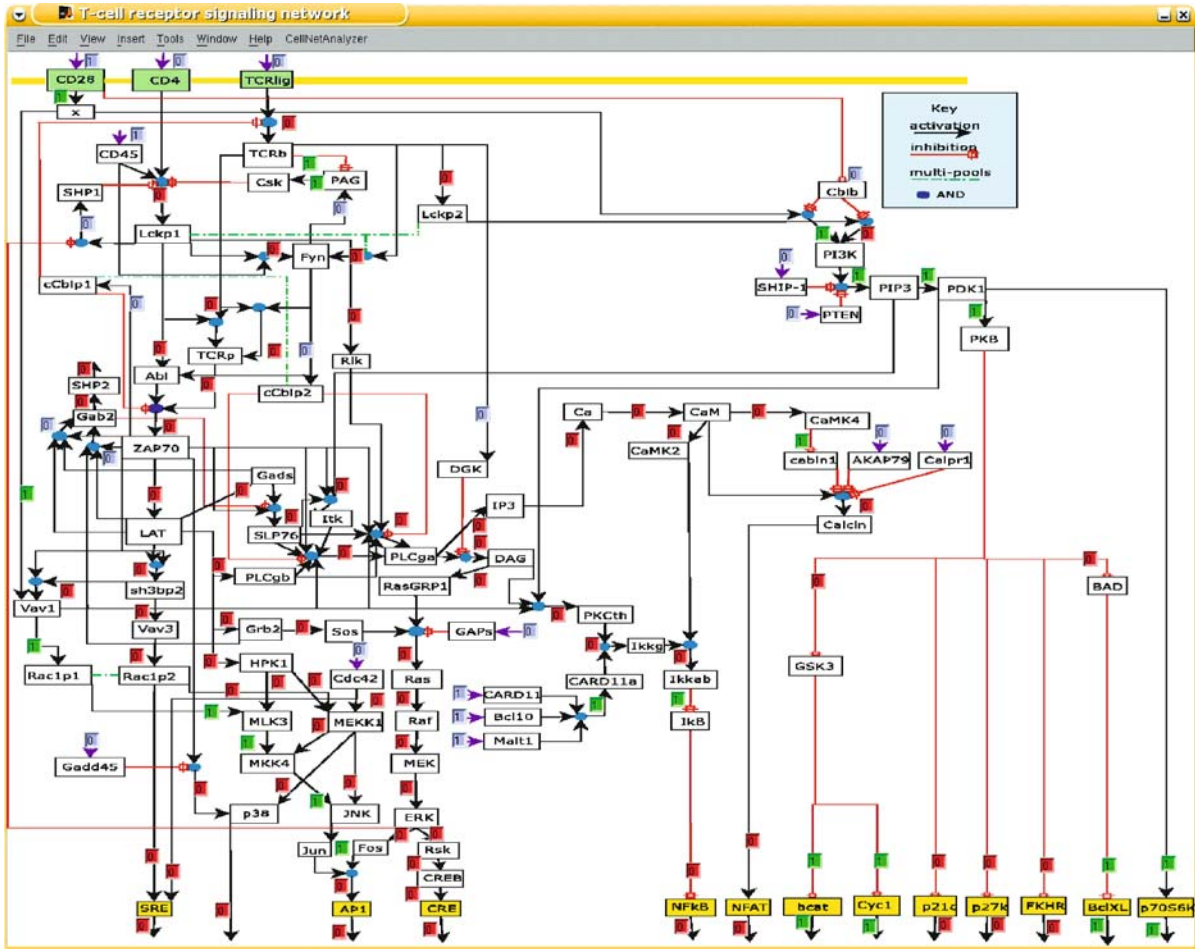
Logical regulatory graph for the mammalian cell cycle network. Each *node* represents the activity of a key regulatory element, whereas the *edges* represent cross-regulations. *Blunt arrows* stand for inhibitory effects, *normal arrows* for activations [37]

by antigen presenting cells by means of the T cell receptor (TCR) and costimulatory molecules (CD4/8/28). This stimulus initiates a signaling cascade within the T cell that eventually influences its activation status. The T cell response must be tightly regulated, as a response to a pathogen that is too weak endangers the whole organism, whereas an overshooting reaction may lead to autoimmune disorders. Saez-Rodriguez et al. constructed a complex Boolean model of this signaling network from local interaction knowledge (Fig. 5). In addition to reproducing known global behaviors, the model also predicted unexpected signaling events after antibody-mediated perturbation of CD28, which could experimentally be validated [38,39].

### Network Inference

While the inference of the network models described in the last section was clearly knowledge-driven and required the input of many interpreted biological experiments, network inference algorithms focus on automating or at least assisting this process. Different types of data can help reveal different aspects of the underlying network. For example, genome-wide transcription factor localization studies may reveal which elements are interacting, although they won't tell much about the logical rules representing their relationships. Time-series gene expression data can fill that gap and yield the required state transitions, in order to





Boolean Modeling of Biological Networks, Figure 5

Logical model of T cell signaling (screenshot from *CellNetAnalyzer*). The network consisting of 96 nodes describes the signaling from the T cell receptor and its coreceptors to the key transcription factors AP-1, NFAT and NF $\kappa$ B, which determine T cell activation and function [38]

derive the Boolean functions. However, it is often experimentally unrealistic to obtain a sufficiently rich time-series data set necessary to infer the Boolean rules non-ambiguously.

Suppose the connectivity graph of the example network is given and both attractor profiles are known, then Boolean rules for node 1, 2 and 4 can be derived, whereas this data is consistent with several logical rules for node 3:  $f_3 = \bar{x}_1 x_2$ , or  $x_1 \bar{x}_2 \vee \bar{x}_1 x_2$  or even  $x_1 \bar{x}_2 \vee \bar{x}_1 x_2 \vee \bar{x}_1 \bar{x}_2$ . Further, whenever one deals with real data, one has to take into account experimental noise. Thus, inference algorithms should not necessarily aim to identify the hypothetical single correct network, but instead, either identify a set of networks that optimally fit the data or construct a network model that incorporates stochasticity. The former approach, chosen by Martin et al., was based on time

series data from microarray expression profiling; the authors identified a set of networks that best fit the data by an algorithm proposed by Akutsu et al. [40] and analyzed the dynamics of this set [41]. For the latter approach, the model of Probabilistic Boolean Networks can be used – if the data does not allow a non-ambiguous choice of an updating rule, a set of probable rules can be implemented in the model [42,43,44].

### Future Directions

The theory of Boolean networks has always been enriched by data from biology. The theoretical description of criticality and its implications might be given as a paradigm. Today, new findings about the scale free and modular organization of networks are waiting for a theoretical expla-

nation. The increasingly precise data derived from high throughput genome-wide measurement technologies will yield more realistic in- and out-degree distributions of biological systems. The implications of these architectural features on the dynamics will also be numerically investigated by means of the ensemble approach. Also the effects of various update schemes on the network dynamics and how e.g. delays may help to filter out perturbations have to be further studied.

On the other hand the increasing quantity and quality of high-throughput gene and protein expression data and sophisticated algorithms for network inference provide the basis for a plethora of new biological network models, that can be analyzed by the Boolean formalism and its probabilistic extensions. Such models can then be used to make predictions about how the system would respond to various perturbations and environmental signals and, conversely, what kinds of intervention strategies would be most successful for controlling the system that is in a diseased state [45].

## Bibliography

- Kauffman SA (1969) Metabolic stability and epigenesis in randomly constructed genetic sets. *J Theor Biol* 22:437–467
- Kauffman SA (1993) *The Origins of Order*. Oxford University Press, New York
- Thomas R (1973) Boolean formalization of genetic control circuits. *J Theor Biol* 42:563–585
- Yuh CH, Bolouri H, Bower JM, Davidson EH (2001) A logical model of cis-regulatory control in a eukaryotic system. In: Bower JM, Bolouri H (eds) *Computational Modeling of Genetic and Biochemical Networks*. MIT Press, Cambridge, pp 73–100
- Jacob F, Monod J (1961) On the regulation of gene activity. *Cold Spring Harbor Symposia on Quantitative Biology*
- Kauffman SA (2004) The ensemble approach to understand genetic regulatory networks. *Physica A* 340:733–740
- Ramo P, Kesseli J, Yli-Harja O (2006) Perturbation avalanches and criticality in gene regulatory networks. *J Theor Biol* 242:164–170
- Shmulevich I, Kauffman SA, Aldana M (2005) Eukaryotic cells are dynamically ordered or critical but not chaotic. *Proc Natl Acad Sci USA* 102:13439–13444
- Nykter M, Price ND, Aldana M, Ramsey SA, Kauffman SA, Hood L, Yli-Harja O, Shmulevich I (2008) Gene Expression Dynamics in the Macrophage Exhibit Criticality. *Proc Natl Acad Sci USA* 105(6):1897–1900
- Drossel B (2007) Random Boolean networks. In: Schuster HG (ed) *Annual Review of Nonlinear Dynamics and Complexity*, vol 1. Wiley, Weinheim
- Bastolla U, Parisi G (1998) Relevant elements, magnetization and dynamical properties in Kauffman networks: A numerical study. *Physica D* 115:203–218
- Bastolla U, Parisi G (1998) The modular structure of Kauffman networks. *Physica D* 115:219–233
- Socolar JES, Kauffman SA (2003) Scaling in Ordered and Critical Random Boolean Networks. *Phys Rev Lett* 90:068702
- Paul U, Kaufman V, Drossel B (2006) Properties of attractors of canalizing random Boolean networks. *Phys Rev E* 73:026118
- Derrida B, Pomeau Y (1986) Random networks of automata: A simple annealed approximation. *Europhys Lett* 1:45–49
- Derrida B, Stauffer D (1986) Phase transition in two dimensional Kauffman cellular automata. *Europhys Lett* 2:739–745
- Flyvbjerg H (1988) An order parameter for networks of automata. *J Phys A: Math Theor* 21:955–960
- Luque B, Sole R (2000) Lyapunov exponents in random Boolean networks. *Physica D* 284:33–45
- Shmulevich I, Kauffman SA (2004) Activities and Sensitivities in Boolean Network Models. *Phys Rev Lett* 93:048701
- Mihaljev T, Drossel B (2006) Scaling in a general class of critical random Boolean networks. *Phys Rev E* 74:046101
- Samuelson B, Troein C (2003) Superpolynomial Growth in the Number of Attractors in Kauffman Networks. *Phys Rev Lett* 90:098701
- Kaufman V, Mihaljev T, Drossel B (2005) Scaling in critical random Boolean networks. *Phys Rev E* 72:046124
- Krawitz P, Shmulevich I (2007) Basin Entropy in Boolean Network Ensembles. *Phys Rev Lett* 98:158701
- Harvey I, Bossomaier T (1997) Time out of joint: Attractors in asynchronous random Boolean networks. In: Husbands P, Harvey I (eds) *Fourth European Conference on Artificial Life*. MIT Press, Cambridge, pp 67–75
- Klemm K, Bornholdt S (2005) Stable and unstable attractors in Boolean networks. *Phys Rev E* 72:055101
- Greil F, Drossel B (2005) Dynamics of Critical Kauffman Networks under Asynchronous Stochastic Update. *Phys Rev Lett* 95:048701
- Albert R, Barabasi AL (2002) Statistical mechanics of complex networks. *Rev Mod Phys* 74:47–97
- Aldana M, Cluzel P (2003) A natural class of robust networks. *Proc Natl Acad Sci USA* 100(15):8710–8714
- Barrera LO, Ren B (2006) The transcriptional regulatory code of eukaryotic cells – insights from genome-wide analysis of chromatin organization and transcription factor binding. *Curr Opin Cell Biol* 18(3):291–298
- Harris SE, Sawhill BK, Wuensche A, Kauffman SA (2002) A Model of Transcriptional Regulatory Networks Based on Biases in the Observed Regulation Rules. *Complexity* 7(4):23–40
- Shmulevich I, Lähdesmäki H, Dougherty ER, Astola J, Zhang W (2003) The role of certain Post classes in Boolean network models of genetic networks. *Proc Natl Acad Sci USA* 100(19):10734–10739
- Kauffman S, Peterson C, Samuelsson B, Troein C (2004) Genetic networks with canalizing Boolean rules are always stable. *Proc Natl Acad Sci USA* 101(49):17102–17107
- Chaves M, Sontag ED, Albert R (2006) Methods of robustness analysis for Boolean models of gene control networks. *IEE Proc Syst Biol* 153:154–167
- Dassow G, Meir E, Munro EM, Odell GM (2000) The segment polarity network is a robust developmental module. *Nature* 406:188–192
- Chaves M, Albert R, Sontag D (2005) Robustness and fragility of Boolean models for genetic regulatory networks. *J Theor Biol* 235:431–449
- Albert R, Othmer HG (2003) The topology of the regulatory interactions predicts the expression pattern of the *Drosophila* segment polarity genes. *J Theor Biol* 223:1–18
- Faure A, Naldi A, Chaouiya C, Thieffry D (2006) Dynamical anal-

ysis of a generic Boolean model for the control of the mammalian cell cycle. *Bioinformatics* 22(14):124–131

38. Saez-Rodriguez J, Simeoni L, Lindquist JA, Hemenway R, Bommhardt U, Arndt B, Haus UU, Weismantel R, Gilles ED, Klamt S, Schraven B (2007) A logical model provides insights into T cell receptor signaling. *PLoS Comput Biol* 3(8):1580–1590
39. Klamt S, Saez-Rodriguez J, Lindquist JA, Simeoni L, Gilles ED (2006) A methodology for the structural and functional analysis of signaling and regulatory networks. *BMC Bioinf* 7:56
40. Akutsu T, Miyano S, Kuhara S (2000) Inferring qualitative relations in generic networks and in metabolic pathways. *Bioinformatics* 16(8):727–734
41. Martin S, Zhang Z, Martino A, Faulon JL (2007) Boolean Dynamics of the Genetic Regulatory Networks Inferred from Microarray Time Series Data. *Bioinformatics* 23(7):866–874
42. Shmulevich I, Dougherty E, Zhang W (2002) From Boolean to Probabilistic Boolean Networks as Models of Genetic Regulatory Networks. *Proc IEEE* 90(11):1778–1792
43. Lähdesmäki H, Shmulevich I, Yli-Harja O (2003) On Learning Gene Regulatory Networks Under the Boolean Network Model. *Mach Learn* 52:147–167
44. Shmulevich I (2007) *Genomic Signal Processing*. Princeton University Press, Princeton
45. Datta A, Pal R, Choudhary A, Dougherty ER (2007) Control Approaches for Probabilistic Gene Regulatory Networks. *IEEE Signal Process Mag* 24:1:54–63

## Bootstrap Percolation

PAOLO DE GREGORIO<sup>1</sup>, AONGHUS LAWLOR<sup>2</sup>,  
KENNETH A. DAWSON<sup>2</sup>

<sup>1</sup> Department of Chemistry and Chemical Biology,  
Cornell University, Ithaca, USA

<sup>2</sup> School of Chemistry and Chemical Biology,  
University College Dublin, Dublin, Ireland

### Article Outline

[Glossary](#)

[Definition of the Subject](#)

[Introduction](#)

[Bootstrap Percolation with  \$m < 2\$](#)

[Sharp Thresholds:  \$m \geq d\$  on Regular Lattices](#)

[Void Instability and Rigorous Results](#)

[Corrections to the Asymptotic Behavior  
in the Presence of Unstable Voids](#)

[Bethe Lattice Calculations: The Cases  \$m \geq 2\$](#)

[Probabilistic Bootstrap Percolation](#)

[High Dimensional Lattices](#)

[k-core Percolation on Random Networks](#)

[Future Directions](#)

[Bibliography](#)

### Glossary

**Bethe lattice** A graph of nodes (sites) each being linked to exactly  $z$  other nodes in the graph ( $z$  is the coordination number). The Bethe lattice is an infinite object such that there is no boundary and each node has an identical topology. Given any two nodes, there is only one unique set of links connecting them, so there are no closed loops. This is a regular graph with connectivity  $z$ . The Cayley tree, also a loopless graph with connectivity  $z$ , differs from the Bethe lattice in that a finite fraction of nodes lead to dead ends at a boundary (where they have connectivity 1).

**Bootstrap percolation** Also known as  $k$ -core percolation, is defined as follows. Given a collection of occupied (or active) and vacant (inactive) sites on a lattice, subsets might exist with the following property: Elements in each subset must be occupied and have at least  $k$  occupied neighbors, each of those occupied neighbor in turn belonging to the same subset. Bootstrap percolation occurs if any subset exists that spans the system from one end to the other.

**Cluster** Any subset of occupied sites, each having  $k$  or more occupied sites belonging to the same subset. The emergence of a spanning cluster underlies the bootstrap percolation transition.

**Culling process** Starting from an initial configuration, culling means that occupied sites with less than  $k$  neighbors are rendered vacant, i. e. ‘culled’, and the same operation is iterated a sufficient number of times as to achieve a final configuration with no more sites that are candidates for culling. The culling process is a practical protocol that leads to identify all the relevant subsets (i. e. clusters) in bootstrap percolation.

**Lattice** A collection of points (‘sites’) generated from linear combinations with integral coefficients of some basis vectors. A  $d$ -dimensional lattice is generated from  $d$  basis vectors. Therefore, a  $d$ -dimensional lattice is a  $d$ -dimensional array or regularly spaced points. A lattice of finite size is such that all the integral coefficients have lower and upper bounds. The coordination number  $z$  is the number of sites (called the nearest neighbors) of equal least distance to any given site.

**Random networks** A network (or graph) of nodes (sites) linked to other randomly chosen nodes. The network is characterized by the distribution of the degree (the number of neighbors of a node). We will discuss bootstrap percolation (or  $k$ -core percolation) on simple random, uncorrelated networks with arbitrary degree distribution (we consider always undirected networks).

**Regular lattice** A regular lattice is a lattice generated from basis vectors that are Cartesian basis vectors, i. e. it is the set of points or a bounded proper subset of points in  $\mathbb{Z}^d$ . A square lattice is the 2-dimensional regular lattice whose basis vectors are  $a(1,0)$  and  $a(0,1)$ , a simple cubic lattice has basis vectors  $a(1,0,0)$ ,  $a(0,1,0)$  and  $a(0,0,1)$ . A hypercubic lattice is a regular lattice of dimensionality  $d > 3$ . The coordination number  $z$  of a regular lattice equals  $2d$ . The length  $a$  is called the lattice spacing (it is set to unity here for practical purposes).

### Definition of the Subject

In bootstrap percolation, we start with a lattice configuration of occupied and vacant sites. Occupied sites that have less than a certain prescribed number of occupied neighbors are rendered vacant, and as new occupied sites are found to satisfy the same condition, these are also rendered vacant. The process is iterated until eventually no more sites can be removed (if any exist). Bootstrap percolation endeavors to determine whether any occupied sites will survive the culling process, and what the macroscopic geometrical properties of the occupied ensembles are. In essence, it is a generalization of conventional percolation which has led to many fruitful insights. A complementary view which emphasizes the dynamical aspect of the bootstrap process, treats the vacant sites as invasive units and occupied sites as inert ones. Inert sites that are surrounded by too many invasive sites will become irreversibly infected and will start themselves to invade others. At the end, the question arises as to whether the ‘infection’ has been contained or it has been pervasive. Bootstrap percolation was first introduced in the late 1970s as a simple model to study the properties of some magnetic materials [20,43,56]. One application that it was proposed to model was crystal field interactions, which tend to suppress magnetic moments by forcing the atoms into their singlet state- they are then in competition with exchange interactions that favor the alignment of the magnetic moments of neighboring atoms. For the material to display macroscopic magnetic properties, exchange interactions must be everywhere locally dominant. Later it became apparent that the bootstrap percolation model is quite accurate in describing certain (low temperature) properties related to quadrupolar orientational ordering of molecular solids [5], for example when both quadrupolar and isotropic molecules are present. A paradigmatic example is solid molecular (ortho- $\text{H}_2$ ) $_x$ (para- $\text{H}_2$ ) $_{1-x}$ , or (para- $\text{D}_2$ ) $_x$ (ortho- $\text{D}_2$ ) $_{1-x}$  [1,5]. More recently bootstrap percolation, also called  $k$ -core percolation, has seen

a resurgence of interest in connection to simple kinetic models for the glass transition [29,39,42], and as a model system for jamming of both hard-sphere and soft-spheres molecules at high packing densities [63]. Other applications have been suggested in the study of porous fluid flow through cracked rocks [2], disease spreading [10], hysteresis [58], computer storage arrays resilience to failure [41], sand-pile formation [47], spreading of alerts on distributed networks [69] and many other problems for which standard nearest-neighbor percolation has already been proposed as a model system [66]. More generally, bootstrap percolation is a useful model to address certain physical properties of a system whose local interactions are strongly contingent on the relative states of neighboring units, whereby a prevailing driving ‘force’ only arises when many units act cooperatively ‘in phase’.

### Introduction

In the decades that preceded the introduction of bootstrap (or  $k$ -core) percolation models, a new paradigm in Statistical Physics had revolutionized our understanding of complex physical systems, especially in the presence of critical behavior and in its vicinity. From Ising onward [38], simple (often lattice) models became an increasingly popular way to treat highly complex problems, analytically and later by means of computer simulations. In some rare cases highly non-trivial exact solutions [53] provided the first hint that (already in two dimensions) one would discover important connections with real critical behavior of continuous systems. Quite unexpectedly, given the simplicity of lattice models, it was later demonstrated that in the proximity of critical points a surprising variety of important and correct general results can be derived all within the framework of renormalization group calculations for lattice models [75]. It emerged that in most cases dimensionality is the only relevant variable in play for systems belonging to the same universality classes. Thus, to rely on the use of lattice models was found to be much less problematic than one would have anticipated. The universality of behavior emerging from the study of criticality for simple lattice models was understood to apply under some very general conditions about the microscopic forces, as with short ranged interactions and if contributions from multi-body interactions were unimportant. Conventional (site or bond) percolation fits perfectly in this picture [65,66], providing a way to investigate further the geometrical properties of systems with critical behavior that are characterized by the same universal exponents as those of critical phenomena. On the other hand, it rep-



resented a confirmation of the rather surprising ubiquitousness of scaling [74] universality classes.

Bootstrap percolation provides an extremely useful platform to investigate other classes of phenomena. Can we construct a simple percolation model that produces other types of universality classes? In percolation theory, models that deal with exceptions to equilibrium continuous phase transition critical behavior are known to exist. Oriented and rigidity percolation are just some examples [59]. Bootstrap percolation is a model which has the flexibility to explore an ample spectrum of conditions over the local ‘interactions’ while still facilitating in many cases, through its simplicity, the development of rigorous mathematical results [7,18,19,35,36,62,70,72,73]. The underlying ‘interactions’ can be thought of as strictly local. However, over larger scales and unlike with critical phenomena, the results are strongly dependent on the specifics of the interactions, scaling can be sometimes problematic, and the outcomes can be quite diverse. Concurrently, bootstrap percolation can be viewed as a cellular automaton where basic dynamical properties of out-of-equilibrium or slowly evolving systems can be studied [2,27,39].

Let us make one concrete example. To set the discussion into context, we can think of some of the several lattice models that have recently been proposed to describe some dynamical properties of glass-forming systems [29,39,42], or jamming scenarios [63]. These models can be of two kinds [57], either lattice gas models in which the particle number is conserved [27,39,42], or spin models of the Ising type [29] in which some simple external field drives the equilibrium configurations to have low mobility at low temperatures. The distinction is not extremely important, similarly to the situation of Ising-like models for ferromagnetic transitions versus lattice models for liquid-vapor transitions. In any case, it is postulated that some particles can only move (or spins can flip) if the number of occupied neighbors (or of up-spins) does not exceed a given threshold number. Conversely, one may say that the dynamics is facilitated if a sufficient number of vacant sites (down-spins) exist in the neighborhood of a given particle to support the required density transformation. The outcome is clear. The dynamics are constrained locally in all-or-nothing possibilities, which depend on concentration of vacant (or mobile) space around. If this vacant space is not sufficient, the dynamics does simply not progress. As a consequence collective motion can become heterogeneous and dramatically slowed [29,39,45,57].

Having such a picture in mind, we can immediately recognize the connections with bootstrap percolation (BP). In BP, the sites of the lattice can be either occupied

with probability  $p$ , or vacant with probability  $q = 1 - p$ . Usually, multiple occupancies are not admitted in such models, and the initial probability for occupied sites is entirely random, but such conditions are of a practical nature, and variants are admitted [47]. We can refer to  $p$  as either the concentration, or particle density. The linear system size of the lattice is taken to be  $L$ . We then define a number  $1 \leq k \leq z$ ,  $z$  being the coordination number. For convenience, let us also define  $m = k - 1$  to make contact with several alternative definitions of the same parameter that are present in the literature. Then a culling process is carried out. All vacant sites as well as all sites that have  $k$  or more occupied sites among the nearest neighbors (the ‘constrained’ sites) are left unaltered, while all sites that have  $m$  or less occupied neighbors among their nearest neighbors are removed. Step by step, many particles are removed and the concentration of particles gradually decreases. As new space is emptied, other particles previously stable can be removed. We eventually reach a stationary situation.

There can be different scenarios for the stationary state at the end of the culling process. There can be either no particles at all, or there can be clusters of particles, one of which will eventually span the entire system (we shall see that in some cases there can be no finite-sized clusters, and the only two possibilities are that there be either no particles at all or a giant spanning cluster). The outcome depends on  $p$ ,  $k$ , the lattice geometry and on the system size  $L$ . For example, for not too small  $L$ ,  $k > 2$  and for small values of  $p$  we can expect that all particles will be eventually removed. A systematic study can be carried out using a similar procedure to that of conventional percolation. Given the system size  $L$ , the culling process is carried out at some low value of  $p$ . It is expected that there will not be spanning clusters in the final configurations.  $p$  is gradually incremented, until at a certain value  $p_c(L)$  that depends on  $L$ , 50% of the final configurations will have a spanning cluster of irremovable particles. By changing the system size  $L$  we can construct the relation  $p_c(L)$  between the bootstrap percolation finite-size critical concentration and the system size. For an infinitely extended system we define  $p_c = p_c(\infty)$  as the (‘thermodynamic’) critical fraction. Alternatively, one can fix  $p$  and let the system size vary from small to large  $L$  and identify the critical size  $L_c$ , known as the bootstrap percolation length- this is the size of system for a given  $p$  for which 50% of the configurations lead to a large spanning cluster in the final state. Then  $p$  is varied so that we can construct the relation  $L_c(p)$  between  $p$  and the corresponding bootstrap percolation critical size  $L_c$ , which is of course just the inverse of  $p_c(L)$ . Other quantities of interest are the same of conven-



tional percolation, such as the concentration of occupied sites in the clusters surviving the culling process, or the mass factor associated to the spanning cluster.

$p_c(L)$  and  $L_c(p)$  can be clearly evaluated in a systematic fashion from computer simulations. Analytically there are radically different techniques that can be used to find them, depending on the value of  $k$  and the lattice geometry. For low  $k$  the bootstrap percolation problem is essentially equivalent to conventional percolation [66], and therefore real-space renormalization group calculations are the obvious choice. As we will see, there typically is some value of  $k$ , call it  $k_{uv}$ , at which ‘unstable voids’ [1,2,3,7,23,43,70,72,73] emerge for any choice of  $k \geq k_{uv}$ . More precisely, while for  $k < k_{uv}$  clusters of either finite size or spanning the system can survive the culling procedure, for any  $k \geq k_{uv}$  self-sustaining (finite) clusters cannot survive the decimation process. This is typically due to the presence of voids that can become unstable when  $k \geq k_{uv}$ , i.e. are cores of empty sites of certain critical sizes that with high probability can nucleate and lead to a complete decimation of the particles in the system. For such models sophisticated analytical techniques have been devised and developed [7,11,18,19,60,62], in some cases leading to explicit rigorous mathematical results [35,36] or exact interpolations and predictions [23]. A third route is that of ‘mean-field’ calculations for Cayley trees (Bethe lattices) [20,34,49,63,64]. These have been investigated thoroughly in recent years especially because there can be values of the parameter  $k$  in the window between those values of  $k$  that can be treated with usual renormalization group techniques as in conventional percolation, and the value  $k_{uv}$  at which unstable voids are expected to arise. Some very interesting results have emerged already, in combination with simulations [63] and some generalizations of bootstrap percolation models [40,67,68]. However, there still are some controversies that will have to be clarified [54]. We expect that many of these issues will become clearer in the near future, possibly with new approaches [34] that will not rely entirely on mean-field calculations [20]. Finally, one obvious extension of bootstrap percolation is to random networks [25,30], which we discuss in Sect. “*k*-core Percolation on Random Networks” – this approach is already attracting considerable interest.

### Bootstrap Percolation with $m < 2$

The case  $k = 1$  ( $m = 0$ ) is trivially the same problem as conventional percolation [66], in any dimension and for any lattice geometry. In fact, only isolated occupied sites are removed, which in any case do not participate to

the construction of percolation clusters in conventional nearest-neighbor percolation. It is then straightforward to recognize that, on approach to criticality, the critical fraction  $p_c(L)$  is related to the system size and  $p_c = p_c(\infty)$  via the relation,

$$|p_c(L) - p_c| = AL^{-1/\nu}[1 + \epsilon(L)] \quad (1)$$

$$\lim_{L \rightarrow \infty} \epsilon(L) = 0$$

while the total fraction of sites belonging to the infinite cluster vanishes with  $p$  approaching  $p_c$  from above as,

$$P_\infty \propto (p - p_c)^\beta \quad (2)$$

The exponents  $\nu$  and  $\beta$  can be derived within real-space renormalization group calculations, or their relations derived from scaling theory, and estimated from computer simulations. They are therefore only dependent on the dimensionality  $d$  and not on the geometry of the lattice. All other exponents that are related to the mass fraction or to the percolation correlation length are dependent on the previous exponents and can be derived from those. Their values are given in this volume in the sections devoted to standard percolation.

In the case  $k = 2$  ( $m = 1$ ), only dangling ends are removed from the clusters of any size. It follows that this has no effect on the relation given in Eq. (1), representing correctly the relation between the concentration of particles and the size of the spanning cluster near the percolation transition. In regard to the other exponents relating the system size to the mass fraction or the probability that a site belongs to the spanning cluster, although not apparent from usual system sizes [15], it is found that the exponent  $\beta$  is also universal [16,21,48]. That is,  $\beta$  coincides with  $\beta$  from standard percolation. This becomes only evident if the system sizes are very large [21,48], and it is supported by general arguments [16]. It is noticeable that scaling arguments apply appropriately to both  $k = 1$  and  $k = 2$ , and one has the usual picture of continuous phase transitions.

### Sharp Thresholds: $m \geq d$ on Regular Lattices

The case that perhaps historically has attracted most interest in bootstrap percolation is when we pass from a picture of the type well described by the theory of critical points to one where ‘first-order’ type transitions emerge. Typically, the precursor of such transitions is a very sharp threshold in the concentration-dependent probability to percolate, which is driven by the indirect effects of the so called ‘unstable voids’ [1,3,7,22,23,27,39,43,70,72,73]. These are special objects whose interior is void of particles. They have

the potential capacity to eat away all the particles in the system, provided that they manage to grow beyond a certain threshold size. We can say that if their spreading is not contained at an early time, they will almost certainly end up ‘invading’ or ‘infecting’ the entire system. Whether there is any such object in the system depends of course on the initial concentration of occupied and vacant sites, and on the system size (if the system is finite in size). Their properties will be discussed extensively in this section. One interesting aspect of these models is that their studies have been characterized by an intriguing series of twists and turns, before a clearer and more thorough picture would finally emerge.

The simplest case in which dominance of unstable voids can emerge is when  $m \geq d$  ( $k \geq d + 1$ ) on regular lattices. Therefore,  $m \geq 2$  for the square lattice,  $m \geq 3$  for the simple cubic. However, other geometries can also lead to the same situation, and  $d$  might not always be the relevant threshold value for separating the different classes. For example, the same picture holds for the triangular lattice (the 2-dimensional lattice with coordination number  $z = 6$ ), whereby unstable voids will emerge if  $m \geq 3$ . Notice that for the triangular lattice, the threshold number for this class is 3 and it also coincides with  $z/2$ , as it does for the square lattice, and in general for the hypercubic lattice.

First, we shall be discussing informally some general preliminary properties in the case  $m = d$  for regular lattices. Next, some of the known results will be outlined. Finally some of them will be discussed more extensively with some derivations.

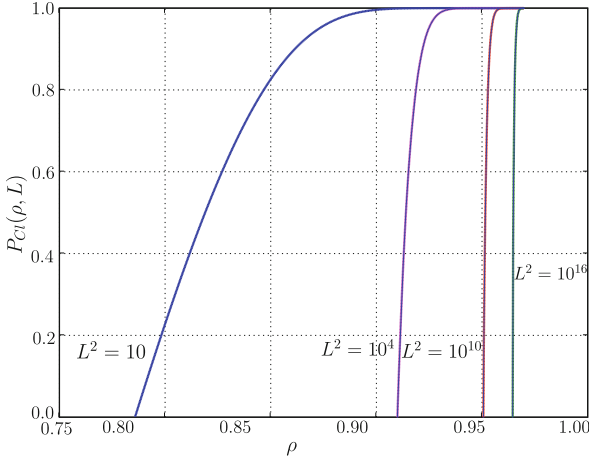
### Spanning Clusters: $m = d$

Let us take the case  $m = 2$  on the square lattice (i.e.  $k = m + 1 = 3$ ). This implies that only particles with less than 3 occupied neighbors can be removed. To start with, it is useful in general to know whether clusters of finite size (i.e. non-spanning) might survive the decimation process in the box of area  $L^2$ . Let us suppose that one such cluster exists. By definition, if it is finite, the cluster must have a boundary. But if it has a boundary, this implies that there exists a non-empty set of particles in such cluster that have at least two vacant sites as their nearest neighbors. This set cannot be empty because any finite cluster is a compact object (you may think that it has at least four ‘corners’, or vertices). Particles in such set are therefore removed at the next iteration of the culling process. But now we have again a finite cluster with a non-empty set of ‘corners’. It follows that the removal procedure cannot be stopped and the cluster will eventually be washed away. We conclude

that the only possibility for a cluster to survive the culling process is that it is not self-sustaining, i.e. only spanning clusters that are connected to themselves through the lattice boundary lines may exist at the end. Similar conclusions can be derived for any dimension for the simple cubic and hypercubic lattices.

It should start to transpire already why these models are special. Given a certain system size  $L$  and some value of  $p$ , after culling terminates only two situations can happen. Either the lattice is empty, or it contains one large (massive) spanning cluster, and we collect the statistics for many equivalent realizations at  $L$  and  $p$ . We cannot, as in the case of critical points in conventional percolation, identify a broad distribution of finite-sized clusters of all lengths, whereby each sampled configuration (whether percolating or not) would be similarly distributed in any element from the wide samples. Furthermore, from the above argument it is also clear that all the ‘roughness’ over the boundaries of any set of connected particles in the spanning cluster is washed away during the culling process. In other words, we should expect its fractal dimensionality to be the same as the volume it is embedded in. All these qualitative considerations suggest a scenario that is very typical of first order phase transitions rather than that of continuous ones. More careful examinations are consistent with this picture. These models are without doubt a departure from standard percolation problems, and also from bootstrap percolation models with small values of  $m$ .

In Fig. 1, we see an example of the probability of a spanning cluster in the final configuration, plotted as a function of particle concentration  $p$  and for different system sizes (from theory and for a modified model that is qualitatively equivalent to the two dimensional  $m = 2$  square lattice model). As we can see, the ‘transition’ in the probability, from essentially zero to some finite value, occurs over a very narrow range, and it displays an important size dependence. At large system sizes the sharp change is ‘almost’ discontinuous. Furthermore as said, whenever a spanning cluster survives the culling process, the fraction of occupied sites in it is typically finite (i.e. non-fractal), leading to a similar sharpness in the probability that any given site anywhere in the lattice belongs to the spanning cluster. The analysis of similar results in two and three dimension from computer simulations (which are limited as to the maximum system size available) led initially to the conjecture of a power-law behavior [43] for  $d = 3$ , in the form of Eq. (1) but with a critical concentration lower than that of the discontinuous jump, and it was not excluded that the critical concentration for an infinitely extended system would have been  $p_c < 1$ . This



**Bootstrap Percolation, Figure 1**

The typical probability  $P_{CI}$  that a spanning cluster survives the bootstrap process as a function of particle density and for different system sizes. Two dimensional modified model in the example. From left to right, system sizes correspond to  $L^2 = 10$ ,  $L^2 = 10^4$ ,  $L^2 = 10^{10}$  and  $L^2 = 10^{16}$

premature conjecture was only partially and very qualitatively consistent with the mean-field Bethe lattice behavior [20] studied theoretically (this topic will be discussed in another section). However, it was later rigorously proven that  $p_c = p_c(\infty) = 1$  [7,62,72,73]. More precisely, as we will see, the probability that an unstable void develops is always small but strictly positive, for any initial value of  $p < 1$  and in any dimension for regular lattices, whenever  $m \geq d$ . Therefore, if we suppose that the lattice is infinitely extended, we will encounter at least one unstable void with certainty.

### Void Instability and Rigorous Results

In recent years more and more rigorous, theoretical and simulation results have led to an impressively detailed understanding of these models [1,3,4,6,7,11,18,19,22,23,35,36,37,44,62,72]. Before stating some of these results we first define, for convenience, a variant model that has been described in great detail. It is called the modified bootstrap percolation model [22,23,35,36,62]. For any regular lattice, any particle can be removed if and only if in any of the coordinate directions, it is neighbored by *at least* one vacant site. This condition is similar to the case  $m = d = z/2$ , but with an additional requirement on how the vacant sites must be distributed around the particle (so to avoid too uneven distributions of voids). The set of particles that can be removed in modified bootstrap is a subset of those that can be removed for the standard (conventional)  $m = d$  case.

We can now look into some of the known rigorous results. They are listed here as relations between the threshold concentration for the first appearance of a spanning cluster, as functions of the system size  $L$ . This is the commonest way to visualize quickly the types of finite size corrections. However, later we will discuss why for other purposes not necessarily this is the most convenient form. The inverted form of these relations will be discussed later. In all the following,  $\epsilon(L)$  must be intended as an unspecified, *non-universal* (positive or negative) function that is vanishing in the limit  $\lim_{L \rightarrow \infty} \epsilon(L)$  is some positive term which is certainly bounded by two constants for sufficiently large  $L$ . We have for all regular lattices from square to hypercube,

$$p_c(L) = 1 - \frac{\pi^2}{18 \ln L} [1 + \epsilon(L)] \quad d = 2, \quad m = 2 \quad [39] \quad (3)$$

$$= 1 - \frac{\pi^2}{6 \ln L} [1 + \epsilon(L)] \quad d = 2, \quad \text{modified} \quad [39] \quad (4)$$

$$= 1 - \frac{B(L)}{\ln \ln L} \quad d = 3, \quad m = 3 \quad [7, 30] \quad (5)$$

$$= 1 - \frac{\pi^2}{6 \ln \ln L} [1 + \epsilon(L)] \quad d = 3 \quad \text{modified} \quad [40] \quad (6)$$

$$= 1 - \frac{B(L)}{\ln^{o(d-1)} L} \quad \text{any } d \geq 2, \quad m = d \quad [19] \quad (7)$$

$$= 1 - \frac{\pi^2}{6 \ln^{o(d-1)} L} [1 + \epsilon(L)] \quad \text{any } d \geq 2, \quad \text{modified} \quad [40] \quad (8)$$

$$= 1 - \frac{B(L)}{(\ln L)^{d-1}} \quad \text{any } d \geq 2, \quad m = z - 2 \quad [7] \quad (9)$$

$$= 1 - \frac{B(L)}{(\ln^{o(2d-m-1)} L)^{m-d+1}} \quad \text{any } d \geq 2, \quad d \leq m \leq 2d - 2 \quad [19, 29] \quad (10)$$

As usual,  $z$  is the coordination number,  $z = 2d$  in this case. The expression  $\ln^{o(n)}$  stands for the natural logarithm iterated  $n$  times. The explicit form of  $\epsilon(L)$  is expected to

vary considerably case by case, nevertheless it vanishes for large  $L$ , but it is unknown rigorously. From expressions from Eq. (3) to Eq. (10) we readily recognize the importance of finite-size effects. To see this, take for example the general expression for  $m = d$  outlined in Eq. (7), which is in fact the case of major interest. The leading term in the expression is the inverse of the logarithm iterated  $d - 1$  times, which converges to zero extremely slowly, the higher the dimension the slower it converges. This is very different from standard power law behavior, and explains why it was not immediately recognized that these systems have a critical concentration of 1. As it transpires, we have a whole new ‘universality class’, peculiar to bootstrap percolation. It is especially noticeable that in many cases, the asymptotic expressions are such that the multiplicative constant is known exactly. Another issue that may not appear so obvious at first is that the forms of  $\epsilon(L)$  are not rigorously known. Due to the slow convergence of the leading terms,  $\epsilon(L)$  may add substantial finite-size corrections that are in addition to the iterated logarithmic part. In fact, in principle  $\epsilon(L)$  in itself may converge to zero even slower than inverse logarithms, and it may be comparable to unity up until the largest sizes that are physically interesting. As we are going to see, already in two dimensions this effect is not at all negligible.

As we shall see, a more natural choice for these models is to express the critical bootstrap length in terms of the initial particle concentration, or rather of the initial concentration of empty sites  $q = 1 - p$ . There is a unique transformation from  $q = q(L)$  to the inverse relation  $L = L(q)$  that relates critical concentrations to critical sizes and viceversa. For example, we can preliminarily rewrite Eq. (7) as expressing the asymptotic result  $L_c(q) \sim \exp^{o(d-1)}(A/q)$ , i. e., an iterated exponential whose argument is of order  $O(1/q)$ . Therefore in  $d = 2$  we have

$$L_c(q) \sim \exp\left(\frac{A}{q}\right) \quad (11)$$

expressing the exact dramatic divergence of the critical size when the fully filled lattice limit is approached. It is useful however to formally invert Eqs. (3) and (4) by separating out the asymptotic contribution from the corrective term, and by defining an unknown function  $f(q)$  related to  $\epsilon(L)$ . We can do this by writing

$$\ln L = \frac{A}{q} + f(q)$$

and observing that

$$\epsilon(L) = \frac{qf(q)}{A} \quad (12)$$

must hold. Notice that  $A > 0$  and therefore  $f(q)$  and  $\epsilon(L)$  are of equal sign. From Eqs. (3) and (4) and the properties of  $\epsilon(L)$ , in the limit  $L \rightarrow \infty$  both  $q \rightarrow 0$  and  $\epsilon(L) \rightarrow 0$ , and therefore,

$$L_c(q) = \exp\left[\frac{A}{q} + f(q)\right] \quad (13)$$

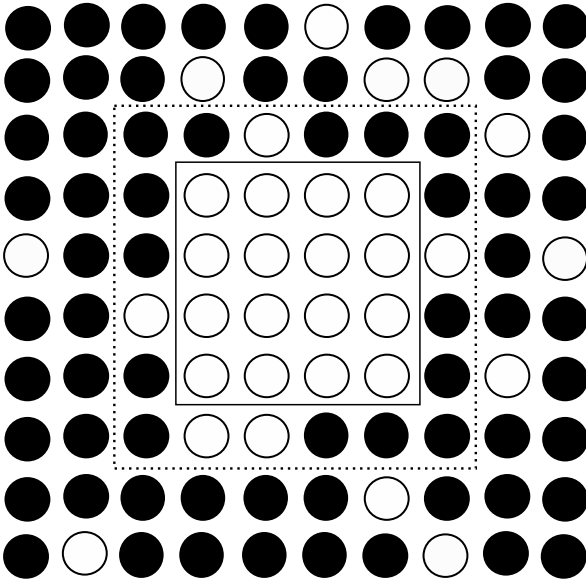
$$\lim_{q \rightarrow 0} q f(q) = 0 \quad (14)$$

As we can see,  $f(q)$  may in principle be a divergent function with vanishing  $q$ , although it must always diverge slower than  $1/q$ . Indeed, the observed critical bootstrap percolation size is often orders of magnitude off from the expected leading asymptotic behavior [1,3,22,35]. This subtlety will be briefly discussed later.

How can we derive Eq. (11) in  $d = 2$  and the corresponding iterated exponentials for  $d > 2$ ? First, we expand more explicitly on the concept of unstable voids (or critical droplets). To illustrate an elementary example of what an unstable void is, take the trivial case  $m = z - 1$ , that is, all occupied sites with at least one vacant neighbor are removed. Take a large lattice with a certain distribution of particles  $p = 1 - q < 1$ , and  $p$  very close to 1 (i. e.  $q$  very close to zero). If  $L^d \gg 1/q$ , we can be certain that there would be some empty sites in the system (despite  $q$  being small). According to the rule assigned, such voids will gradually expand and eat away all the particles in the system, without any barrier that may stop them from growing. Isolated voids are therefore always ‘unstable’ when  $m = z - 1$ , they start to grow immediately as soon as the culling process is started, and they cannot be stopped. Furthermore, if starting from the same initial configurations we subdivide the system into many subsystems of volume  $L'$  such that  $L'^d \ll 1/q \ll L^d$ , we end up with many separated subsystems entirely filled, and a very few containing one (or more) vacant sites. After the culling process is implemented separately on each one of the subsystems, a tiny fraction of the total number of them (those with vacant sites, i. e. unstable voids) are completely empty. All the others are entirely filled by particles. We thus see how in the presence of unstable voids, for a given value of  $p = 1 - q$  we can have different scenarios, depending on the system size, while at the same time for an ideal system that is infinitely large, any  $p < 1$  leads to all space to be empty after bootstrap.

The situation is of course much more complex in the more interesting cases, but some basic principles are unaltered (in a crude sense). Let us consider the case of the square lattice when  $m = 2$ , or consider the modified

model (they behave similarly). Imagine that we have created somewhere a rectangle whose inside is completely empty (note that it is often said that such rectangle is internally spanned, but we must not confuse this notion with that of a spanning cluster of percolating particles). In Fig. 2 we have a square as an example. On the boundary lines of its perimeter there can be any number of particles. But let us see what can happen if some sites on those lines are vacant. Due the rules for the removal of particles, any vacant site on any one of those lines is a one dimensional unstable void limited to that line, because each site on the line has already one vacant neighboring site which lies in the interior of the empty square. This means that if on any of the boundary segments there is one vacant site or there are more, that segment will be vacated. If the condition is satisfied in all four directions, the empty square can engulf its boundary and become 2 linear units larger



**Bootstrap Percolation, Figure 2**

Unstable voids in the square lattice. The internal  $4 \times 4$  square delimited by the solid line has been vacated previously. Both in the modified and in the conventional  $m = 2$  model, the square can engulf its four next sides leading to the vacation of the  $6 \times 6$  square delimited by the dotted line. From there, the growth by squares in the modified model would be stopped because of the impossibility to proceed in the West direction, while the growth by squares is still allowed in the conventional model up to the vacation of the largest  $10 \times 10$  square in the figure. Finally, allowing for the growth by elongated rectangles in the modified model, the void can momentarily expand in the North, East and South direction before all particles in the East direction can be vacated too. Differently from the squares-only process, this time the largest  $10 \times 10$  square in the figure will be vacant at the end, also in the modified model

per side. Again, let us consider  $p < 1$  but very close to 1, i.e.  $q$  positive but very small. For small squares with side length much smaller than  $1/q$  it is very unlikely that any such empty square will grow further. However, for side lengths much larger than  $1/q$  the opposite is true, it is almost certain that the square will grow to the next step, and to the steps after (as the side length becomes larger, it is increasingly likely that the boundary lines will be unstable with respect to the internal empty square). Voids of the first type are therefore stable, voids of the second type are expected to be unstable. Therefore when we turn our attention back to the entire lattice of size  $L$ , we can be certain that all particles will be removed provided that  $L$  is large enough so that the lattice contains one or more unstable voids (and so the lattice is sometimes said to be internally spanned by the voids. Again such terminology must not be confused here with spanning clusters of particles). In order to determine whether voids can ever become unstable even for very small  $q$ , we must determine rigorously some bounds for their probability to arise.

To quantify this argument we ask what is the probability that a vacant site can nucleate and grow indefinitely? By letting this process proceed via the formation of larger and larger squares we certainly obtain a lower bound to the nucleation probability (it is a lower bound because, more generally, nucleation can also proceed via rectangular voids of various elongations [23]). Given an empty square of side length  $l$ , the probability that any of the four external boundary lines can be vacated coincides the probability of there being at least one vacant site, which is formally given by  $1 - p^l$  (it is the complement of there being  $l$  particles, which is  $p^l$  from the random measure distribution on  $p$ ). The joint probability for all four sides *at the same time* is  $(1 - p^l)^4$ . Let us call the probability of indefinite nucleation  $\nu$ , and  $\nu_b < \nu$  is some lower bound. To evaluate this bound, we may start with a square of size  $2 \times 2$  as the nucleus. We may write,

$$\nu_b = q^4 \prod_{k=1}^{\infty} (1 - p^{2k})^4 \quad (15)$$

Here,  $k$  is the side length of the growing square (which is an even number at each step). Therefore,

$$\ln \nu_b = 4 \ln q + 4 \sum_{k=1}^{\infty} \ln (1 - p^{2k}) \quad (16)$$

To bound the sum we first expand the logarithm in infinite series and we exchange the order of summation. We then



use the inequality  $(1 - p^{2k})/p^{2k} > 2k(1 - p) = 2kq$ .

$$\begin{aligned} \sum_{k=1}^{\infty} \ln(1 - p^{2k}) &= - \sum_{k=1}^{\infty} \frac{1}{k} \frac{p^{2k}}{1 - p^{2k}} > - \frac{1}{2q} \sum_{k=1}^{\infty} \frac{1}{k^2} \\ &= - \frac{\pi^2}{12q} \end{aligned} \quad (17)$$

The last equality comes from Euler's sum which equals  $\pi^2/6$  (one can also show that when  $q \ll 1$  the inequality above is essentially an equality but for unimportant corrective terms). From (16) and (17) we therefore have a new bound for the overall nucleation probability, meaning,

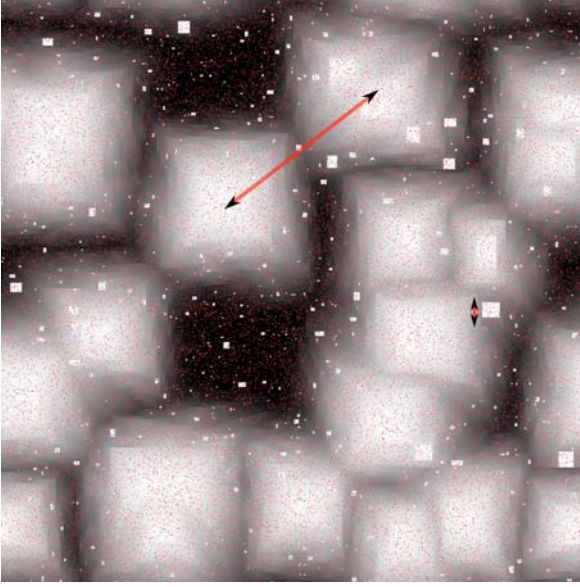
$$\nu > q^4 \exp\left(-\frac{\pi^2}{3q}\right) \quad (18)$$

Equation (18) shows that the nucleation probability, i. e. the probability that unstable voids exist, is a strictly positive quantity for any  $q > 0$ . This implies that for any system of size  $L^2 \nu_b \gg 1$ , i. e.  $L \gg q^{-2} \exp(\pi^2/6q)$ , we can be almost certain that no particle can survive decimation. As  $q \rightarrow 0$  ( $p \rightarrow 1$ ), the required  $L$  increases dramatically. Perhaps even more surprising is if we look at the latter expression, and compare it with Eq. (4), valid for the modified model. The two are fully consistent, with the correct constant. This is surprising because we started with a bounding argument and, apparently, we have ended up with the correct estimation of the asymptotic expression (at vanishing  $q$ ) for  $\nu$  and  $L_c$ . The proof that in fact the bound described above does furnish the correct asymptotics for both  $\nu$  and  $L_c$  is highly non-trivial [35], and despite it being a relatively recent result it is already a cornerstone of bootstrap percolation. One needs to calculate an upper bound which is, to the leading asymptotic order, coincident with the lower bound. In regard to the conventional (non modified)  $m = 2$  case, the argument employed for bounding  $\nu$  starting from the probability of growing square-like unstable voids can be refined. In that case, even if one boundary line is entirely filled with particles, one or more vacancies present in the adjacent line can also lead to all particles in both lines to be vacated (this is seen with the help of Fig. 2). This necessitates some technical adjustments in the calculation above in Eq. (15), but the rest of the argument remains unchanged [35]. One gets a new bound for the square-like process that is precisely in the form (18), but with a new constant given by  $\pi^2/9$ . Again, by means of the identification  $L_c^2 \sim \nu^{-1}$  for the critical size (part of the demonstration [35] shows that to be licit at  $q \ll 1$ , and simulations are consistent already at not too low  $q$  [22]), we have an expression consistent

with Eq. (3). The complete proof for the upper bound is essentially equivalent to that for the modified model. In general, the modified model renders some of the calculations more straightforward, but the two models are in essence very similar.

We thus learn that Eq. (11) (or more appropriately (13)–(14)) is a rigorous result, with the calculated constants  $A$  equal to  $\pi^2/6$  and  $\pi^2/18$  for the modified and conventional  $m = 2$  case on the square lattice, respectively. It would appear that we have at last gained a correct and complete picture, with exact rigorous mathematical results [35] that we can finally compare with computer simulations [44] with a sufficient degree of confidence. However, this is not quite as true yet [3,22]. In reality, fits that use Eq. (11) are typically consistent with estimated constants that are sensibly different [3,6] from the ones calculated within the theory (by factors as large as 2 or 3). This puzzle will be discussed in the next section. It has to do, as earlier pointed out briefly, with the more general expression in (13)–(14) and its subtle implications.

As we have seen, there is a non-zero probability to encounter unstable voids at any  $q > 0$  ( $p < 1$ ). When  $q \ll 1$ , voids can become unstable once they reach a side length much larger than of order  $\bar{k} \sim O(1/q)$  (a more rigorous estimate lends  $\bar{k} \sim O(\ln q / \ln p)$ ). To appreciate this statement qualitatively, we observe that the probability  $1 - p^k$  that a boundary line can be wiped out by the growing void is of order  $1 - \exp(-b) \sim 1$  if  $k = b/q$ ,  $b \gg 1$  and  $q \ll 1$  (notice that the latter also implies  $q \approx -\ln p$ ). Consequently, if on the one hand for  $k \ll O(1/q)$  the voids are stable with high probability, on the other hand for  $k \gg O(1/q)$  they are unstable with high probability. The length  $\bar{k} = 1/q$  is a novel length not present in standard percolation models with critical points (see Fig. 3). We see that contrary to those models, we might be in serious trouble if we were to assume that we can freely re-scale the lattice spacing by any finite factor (like in a self-similar spin-block picture of the type of Kadanoff), just as long as the scaling length is much smaller than the longest length. Here instead we have the larger length-scale  $L_c(q)$  which is always sensibly much greater than the length-scale  $\bar{k}(q)$ , for small  $q$ . On the other hand, on a practical scale, this feature can be of tremendous advantage. Indeed, it is sometimes sufficient to determine the properties of some physical quantities (e. g. correlations) up to within the distance  $\bar{k}$  and no further, when  $L_c \gg \bar{k}$ . A good example is when we want to evaluate the nucleation probability  $\nu$ . Rather than sample systematically systems of very large sizes, it is sufficient to determine whether some nuclei can grow further than  $\bar{k}$  [22], and what the associated probability is for a given concentration  $p$ . This saves an enormous amount



**Bootstrap Percolation, Figure 3**

Evolution of the bootstrap process for  $k = 3$  in 2D ( $p = 0.952$ ,  $L = 900$ ). The time at which particles are removed is depicted by the color- the first ones removed are *white* and the last ones removed are *black* (initial vacancies are *red*). It is seen that the bootstrap removal process starts at nucleation centres, also called connected holes. The *larger scale* indicated in the figure is the bootstrap length  $L = 1/\sqrt{p}$ , or the typical distance between connected holes. The *shorter scale*,  $\bar{k} \sim 1/(1-p)$ , is the critical length above which voids will become unstable

of computer running time for values of  $p$  that correspond to very large critical bootstrap lengths.

More articulated mathematical analysis lead to bounds and exact results in higher dimensions [18,19,36]. We can consider a qualitative argument to determine the bootstrap critical system size in more than 2 dimensions. As an example, let us consider  $d = 3 = m$  (also modified model). We need to establish whether voids can be unstable in this case, and this is indeed the case. Take for instance [72,73] some large cube all empty in the inside, and whose side length is of some value  $k$ . We look for a condition such that the empty cube can grow in a given direction by one step, thus absorbing any particle that might eventually lie within the boundary surfaces. In the simple case, we consider one square-like boundary surface per side at a time. The sites on one single surface form a  $k \times k$  square lattice, but additionally each site is also connected to one site that lies on the interior of the cube (which is certainly empty), and one on its exterior. Thus any site within such  $k \times k$  square lattice that forms the boundary surface of the cube has at least one neighboring vacant site (because the cube is empty), and only needs two more vacant neigh-

bors to be removed. A sufficient condition is that the two required vacant sites are among the 4 neighbors that lie on the  $k \times k$  square lattice itself, irrespective of the occupation state of any site from the exterior. We thus realize that if locally the boundary  $k \times k$  square lattice has at least one unstable void in the  $m = 2$  sense, then the entire square can be vacated. This can be repeated for any of the six sides that form the entire boundary of the cube. It follows that we have reduced the treatment of the one-step growth for the empty cube for  $m = 3$  to six equivalent  $m = 2$  sub-problems on its surface. A problem that we already know how to solve nevertheless. For example at  $q \ll 1$ , we know that if the side length  $k$  is of order  $\exp [O(1/q)]$ , then with high probability all particles within the square-shaped surfaces are likely to be removed. But this means that the empty cube will almost certainly grow by one step, and the same will happen at the next step. We have therefore established that, very approximately, we must expect the new  $\bar{k}$  to equal  $\exp [O(1/q)]$  in the present case (much larger than  $O(1/q)$  as for the square lattice).

All that is left is to evaluate the probability that such unstable void of the required length arises somewhere. In principle we would need an iterative formula of the type of Eq. (15), but it will suffice to determine a qualitative bound for that probability. A minimal condition for a cubical region at  $q \ll 1$  to be unstable is that it must contain  $\exp [O(1/q)]$  empty sites at least, in some appropriate positions. For example, if the 6 faces that constitute the frame of a cube are all empty, one can prove that the interior of the cube will be eaten away gradually, leading necessarily to an entirely empty cube of side length  $\exp [O(1/q)]$ , which is unstable for the argument above. This can happen with probability

$$q^{\exp [O(1/q)]} = \exp \{-\exp [O(1/q)] \ln (1/q)\}.$$

We then apply the usual condition that  $L_c^d$  multiplied by such probability be  $O(1)$ , and we get

$$L_c \sim \exp \{\exp [O(1/q)] \ln (1/q)\}.$$

By taking twice the logarithm we first neglect  $\ln \ln (1/q)$  compared to  $O(1/q)$  and we then obtain

$$q \sim \frac{1}{\ln \ln L_c}$$

which is the desired result as in Eqs. (5)–(6). Rigorous results for upper bounds require other elaborate treatments, not to be mentioned here. However, as above, in most cases a convenient scheme for the proofs requires one to employ some dimensional induction from  $d - 1$  (supposedly known) to  $d$  [19]. It is worth noticing that an (elaborate) induction argument on dimensionality also allows

one to derive Eq. (6) from Eq. (4) [36]. Similar arguments (alongside with their sophisticated rigorous counterparts) apply to dimensions higher than 3, giving rise to the iterated exponential (for  $L_c(q)$ ) and iterated logarithmic forms (for  $q_c(L)$ ) present in Eq. (7).

A word of caution however is in order here. The results so far have been illustrated assuming that  $d$  is kept fixed. From Eq. (7), which is valid for  $m = d$  on the hypercubic lattice, we see that every time we raise  $d$ , the required  $L$  for attaining an arbitrarily small  $q_c = 1 - p_c$  must increase, because each time we must add another iteration of the logarithmic function in the denominator. What happens then if we let  $d$  vary at the same time as varying  $L$ ? One notices that in reality Eq. (7) holds not only when  $d$  is fixed, but also when it is some extremely slowly growing function of  $L$  [13,19]. And so we ask if it is always guaranteed that  $q_c$  becomes arbitrarily small when  $d$  grows too (provided that  $L$  is large enough)? This must be answered in the negative [13]. In fact, there exists some monotonic  $d = d(L)$  such that in the limit of  $L \rightarrow \infty$  we have  $q_c = 1/2 = p_c$ . It is important to notice that  $d(L)$  is still a very slowly growing function of  $L$ , more precisely  $d(L) = (\ln \ln L)^2 \ln \ln L$ , but nevertheless it is sufficient for the result. In other words, there certainly exists a severe bound on how fast  $L$  must grow when  $d$  is systematically raised, in order that we still have  $p_c = 1$  for the system of infinite size. If the condition is not met, then  $p_c < 1$  ( $q_c > 0$ ) for a diverging  $L$ .

To complete this survey, we mention briefly the case  $m = z - 2$  [7] where we have relations that are described in Eq. (9). The simplest case is when  $m = 4$  for the simple cubic lattice. One can derive a bound by a recursion formula similar to Eq. (15), with partial probabilities of the type  $p^k$  in Eq. (15) now substituted by  $p^{k^2}$  (or more generally  $p^{k^{d-1}}$  in  $d$  dimensions). To see this, consider the simple cubic model with  $m = 4$  and the instability condition for cubic voids. We can reiterate an argument similar to (and yet simpler of) the argument above. If in any of the six boundary surfaces there lies at least one vacant site, then the empty cube can grow in that direction. If all six surfaces satisfy the condition, the whole cube will grow in all six directions concurrently. If the side length is given by  $k$ , the probability that one boundary surface is entirely filled with particles is given by  $p^{k^2}$ , and the complement is therefore  $1 - p^{k^2}$ , i. e. the probability that there lies one vacant site at least, or more. Rather than employing a full calculation (lengthy but trivial) for the bound such as the one in Eq. (15), we recognize that this situation is similar to the case  $m = 2$  on the square lattice. There the critical size for the instability of the voids had to satisfy the condition  $kq \gg 1$ , i. e. with high probability there

are vacant sites on the line. Here we have that a similar condition must hold, more precisely that  $k^2 q \gg 1$  must hold. We therefore have, roughly, that  $\bar{k} \sim 1/\sqrt{q}$  is the threshold size for voids' instability. By analogy with that case, we conclude that  $L_c \simeq O(\exp(1/\sqrt{q}))$ , in accordance with (9).

Additionally, some authors [9] have studied the behavior of isolated hypercubes for  $m = d - 2$ , for large values of  $d$ , showing that the critical concentration of empty sites sharply decays to 0 as  $q_c \sim 2^{-2\sqrt{d}}/d^2$ .

We conclude this survey on models with sharp thresholds by noticing that variants of bootstrap percolation may entail lengths that are diverging comparatively more severely than those outlined in this section, on approach to the critical point  $q = 0$ . One example is the square lattice with the asymmetric rule introduced by Gravner and Griffeath [31]. The neighborhood is extended to two additional sites, the second neighbors that lie, for example, on the horizontal directions. The total number of relevant neighbors is therefore  $4 + 2 = 6$ , the four nearest and two second-nearest neighbors. The value of  $m$  is taken to be  $m = 3$ . One finds [71] that at  $q \ll 1$ ,  $L_c \sim \exp[O(1/q) \ln^2(1/q)]$ . This is the same asymptotic behavior [50,72,73] as in another asymmetric bootstrap percolation model on the square lattice introduced by Duarte [26]. Only three out of four nearest neighbors are considered each time, and particles are removed if at least 2 of those neighbors are empty. An even sharper transition was found earlier for a special variant of the bootstrap model on the square lattice [61]. In this model, occupied sites can be removed as usual, but at the same time empty sites can also become occupied. The rule is therefore not strictly irreversible. At each iteration, any given site adjusts its state so that it conforms to the occupancy state of the majority of its neighbors. The key ingredient for the model is the assignment of a bias in all cases in which there is a parity situation, i. e. two neighbors are occupied and two are vacant. There is no majority, and to determine which state a site has to take a bias is introduced towards one of the two possibilities. For example we can codify the rule so that, for all parity situations, the site is removed with probability  $y > 1/2$ , and it becomes occupied with probability  $1 - y < 1/2$ . Therefore when 2 neighboring sites are empty and 2 are occupied, at the next step the site will become empty with a higher probability than to become occupied. In this case, for every choice of  $y > 1/2$ ,  $p < 1$  and for large enough system sizes, there is a high probability that the final state will be an empty lattice. The critical system size scales [61] as  $c_1(y)/q^2 < \ln L_c(q; y) < c_2(y)/q^2$ , where  $c_1$  and  $c_2$  are two unknown constants that depend only on  $y$ . In other words, in this case  $L_c \sim \exp[O(1/q^2)]$ .

### Corrections to the Asymptotic Behavior in the Presence of Unstable Voids

Typically, for the models that have a critical bootstrap length that grows exponentially or super-exponentially in  $1/q$  a detailed comparison of computer simulations inferences with rigorous mathematical results is unsuccessful [3,44]. For the two dimensional models, conventional  $m = 2$  or modified, the constant  $A$  of best fit from the expression in Eq. (11) differs sensibly from the expected values derived from Eqs. (3), (4), which are mathematically rigorous [35]. As an even more concerning issue, it has been suggested that the best fits can sometimes be obtained from different functional expressions altogether. For example in three dimensions, on the one hand we have the relation (5) that implies that we should expect the best fits to  $L_c$  to be in the form  $L_c \sim \exp \exp A/q$ , the closer we approach  $q = 0^+$  the better the fit. Contrarily to common intuition, however, it was proven from computer simulations [44] that for the largest sizes and the smallest  $q$  available a fit in the form  $L_c \sim \exp \exp A/q^2$  is far more reliable, the closer we are to  $q = 0^+$  the better the fit is compared to the expected  $L_c \sim \exp \exp A/q$ . At first sight this could erroneously be interpreted as a mere failure of either mathematicians or simulators. However this does not need be the case. Take for example the two-dimensional case. If we look carefully at what the asymptotic results really mean, we have two equivalent ways of expressing them in general terms. One is as in Eqs. (3) and (4), and the other one is the equivalent form Eq. (13), provided that Eq. (12) holds. From the rigorous analytical arguments usually employed we have no way (as yet) to infer the exact form of either  $\epsilon(L)$  or  $f(q)$ . We know however that  $\lim_{L \rightarrow \infty} \epsilon(L) = 0$  and  $\lim_{q \rightarrow 0} q f(q) = 0$ . However,  $f(q)$  might still be a diverging function of  $q$  in the limit  $q \rightarrow 0^+$ , albeit not as fast as  $1/q$ . Plus,  $f(q)$  enters Eq. (13) exponentially. This is worrying because it implies that the resulting discrepancy between Eq. (13) and Eq. (11) may be in the region of orders of magnitude, and indeed it is [22].

A natural question arises. Where did we lose track of all such non-asymptotic corrections? We may start looking backward at all the derivations and see where we might have made too gross assumptions. One of the main suspects is a relation that we have used all throughout in the discussion, although without rigorous proofs of its legitimacy, but rather tacitly assuming its general validity. Given the probability  $\nu$  (also called the ‘connected hole density’ [22,23,45]) that a vacant site can nucleate indefinitely (i. e. that it becomes an unstable void), we have repeatedly inferred that the relation  $L_c^d \nu \simeq 1$  holds reliably. Indeed, this is not strictly true in general. It tacitly assumes

that the successful nuclei are not correlated spatially in any important way. On the other hand, at high concentrations of particles such nuclei become very rare and therefore highly contingent on the specific distribution of particles and vacancies in their surrounding, meaning that the probability to find one nucleating center just a few lattice spacings away from another one is extremely low. Even if we were to assume that there is a typical average number  $n$  of nucleating centers that occasionally tend to stick together, the consequent correction to the usual assumption  $L_c^d \nu \simeq 1$  would be just a number, a factor of order  $n$  (eventually depending slowly on  $p$ ). Hardly something that would account for order of magnitudes anyway. The good news is that  $\nu$  is a quantity that is directly accessible from simulations, more efficiently than  $L_c$  indeed [22,45]. It is therefore possible to compare  $\nu$  to direct estimates [3,44] of  $1/L_c^d$  and operatively test the relation  $L_c^d \nu \simeq 1$ . Whenever comparisons have been made the two are in excellent agreement, provided  $p$  is not too small. We must look elsewhere to solve the mystery.

When we estimated  $\nu_b$  as a bound for  $\nu$  starting from Eq. (15), we explicitly stated that we were using a sufficient condition for the candidate unstable void, i. e. that this had to be a square all throughout the sequence. However this is not a necessary condition. In principle, even if the void cannot grow conserving its aspect ratio at some particular stage, it may still grow as a rectangle, and eventually become unstable later in the process, although this event cannot be accounted for when we estimate  $\nu_b$  from Eq. (15). In Fig. 2 we have an example of a configuration that for the modified model would lead the square process to stop before the largest is vacated entirely, while the same configuration would lead to an empty final state if we allow one small elongation. One can also prove that growth by rectangular voids of different sizes and elongations encompasses (i. e. exhausts) all self-efficient transformations through which a nucleating center can engulf one by one, neighbor by neighbor, the particles in the entire lattice. In other words, if a growing void cannot wipe out all particles in the system via rectangular transformations, it cannot do so on its own in any other way. The condition of growth by rectangles is not only a sufficient one (as it was in the case of squares), but also necessary. Consequently a calculation (if available) of the type derived in Eq. (15), but dealing with rectangles instead of squares as the intermediate transition states, would gain us access to  $\nu$  exactly.

For the modified model in  $d = 2$  this is exactly what can be done [23]. Instead of using Eq. (15) to estimate  $\nu_b$ , one can generalize the method and determine  $\nu$  accurately for any density, without the need to simulate the model.



The procedure entails mapping the growth by rectangles into an explicit form that takes the form,

$$P_{m,n}^i = \sum_j \left( c_{m,n}^{ij} P_{m-1,n}^j + d_{m,n}^{ij} P_{m,n-1}^j \right) \quad (19)$$

In Eq. (19),  $P_{m,n}^i = P_{m,n}^i(p)$  represents the probability that the nucleation process via rectangular growth will be found after  $m+n$  steps to occupy the state  $i$ . The state  $i$  at given  $m, n$  is typically a rectangle of sides  $m \times n$  that is entirely empty, but eventually for some of the lines on its contour that may be entirely occupied (which ones these are is in fact characteristic of the label  $i$  itself). The range of  $i$ 's and of  $j$ 's coincide. The map is constructed purposefully in such a way that the transitions from state to state can be expressed in the linear form Eq. (19). Concurrently, it has to be constructed so that the transitions coefficients  $c_{m,n}^{ij} = c_{m,n}^{ij}(p)$  and  $d_{m,n}^{ij} = d_{m,n}^{ij}(p)$  have explicit expressions (which are similar but not equal to the partial terms employed in the product Eq. (15) for squares). The mapping is not unique, as in the end we are going to take the limit  $m, n \rightarrow \infty$ , but any map in the form (19) has to satisfy the two conditions above. From the set of states with probabilities  $P_{m,n}^i$  there must be one, say  $i = 1$  for example, that represents an  $m \times n$  rectangle completely empty.  $P_{m,m}^1$  is therefore the probability for an empty square to arise after  $2m$  steps, and  $\nu$  will equal  $\nu = \lim_{m \rightarrow \infty} P_{m,m}^1$ . The map in (19) can be easily solved numerically.  $P_{m,m}^1$  converges sufficiently rapidly to  $\nu$  so that we can genuinely calculate numerically  $\nu$  for almost any value of  $p$ , well beyond anything of physical interest.

By solving Eq. (19) we therefore have the exact  $\nu$  for any particle density  $p$  for the  $d = 2$  modified model. Agreement with its value from computer simulations is beyond the most optimistic expectations [23]. From  $\nu \simeq 1/L_c^2$  we now have a reliable interpolation to the bootstrap critical length for any system size that is available from computer simulations, plus a prediction for  $\nu$  and therefore  $L_c$  well beyond scales of even physical interest, let alone one-day-to-be computational limits. This procedure lacks the insight and elegance of compact formulas such as the one in Eq. (4), but it has the merit (at least for two dimensional models) to wash away all uncertainties related to previously irreconcilable mathematical and computational differences.

Once we have  $\nu$  for any  $p$ , we can access many quantities that are normally available only via computer simulation, with the advantage that now we do not have any important limitation in regard to system sizes. Thus for example we may infer, approximately, the probability that a spanning cluster survives the culling process, given the

system size  $L$  and the particle density  $p$ . We have

$$P_{Cl}(p, L) \simeq \Theta(1 - L^2 \nu(p)) [1 - L^2 \nu(p)] \quad (20)$$

meaning that the average probability  $1 - P_{Cl}$  that a cluster does not survive the culling process is proportional to the average number of connected holes for that size. In essence, a re-elaboration of our (in)famous  $L^2 \nu \simeq 1$ . The pre-factor with the  $\Theta$  step-function (equal to 0 if its argument is negative and to 1 if positive) guarantees that  $P_{Cl}$  is positive definite. In Fig. 1 we have  $P_{Cl}$  as a function of  $p$  for four different system sizes. We have already appreciated the same figure in a previous explanatory section, and to whoever may have wondered how one could have those largest system sizes from computer simulations the trick is now revealed.

We may also try to fit  $L_c \simeq 1/\sqrt{\nu}$  using the expression given in Eq. (13). In essence we thus try to infer the form of  $f(q)$ . Although this is only a best fit, and has no rigorous analytical basis, it is suggestive, because we can access virtually any large value of  $L_c$  from the knowledge of  $\nu$  (even up to 10 to the order hundred). It is also insightful because it gives a flavor of the magnitude of the correction. An excellent fit at  $q \ll 1$  bears [24], for example,  $f(q)$  in the form  $f(q) \approx -Bq^{-\beta}$ , where  $B$  is a fitting constant of a few units, and the best estimate for  $\beta$  is  $\beta \approx 2/3$ . The modulating pre-factor in Eq. (13) given by  $\exp[f(q)]$  is therefore moderating the divergence of the asymptotic contribution. From Eq. (12) one sees that by successive iterations one can get  $\epsilon(L)$  (which is negative as  $f(q)$  is negative), but in general the form in Eq. (4) is less convenient (and less natural) than the one in Eq. (13). More explicitly, we have that if  $qf(q) \ll 1$  and  $f(q) \approx -Bq^{-\beta}$  holds, we estimate  $\epsilon(L) \approx -C/(\ln L)^{1-\beta}$  for diverging  $L$ , where  $C$  is a constant. Recently, Gravner and Holroyd [32] have proposed a rigorous proof that  $\epsilon(L)$  indeed must approach zero no faster than  $-\text{Const}/\sqrt{\ln L}$ . This implies  $\beta > 1/2$ , consistently with the extrapolated value of  $2/3$  above.

### Bethe Lattice Calculations: The Cases $m \geq 2$

As we have seen in the previous sections, we are always able to classify the behavior of bootstrap percolation model if either  $m < 2$  (i. e.  $k \leq 2$ ) or, for regular lattices, if  $m \geq d$  (i. e.  $k > d$ ). The former cases fall under the same universality classes of standard percolation and are valid for any lattice geometry and any dimension. The transition is continuous and characterized by universal exponents [16,21,48]. The latter class is that with a critical concentration  $p_c$  that is unity for an infinitely extended system, and is more reminiscent of first-order transition behavior. It is characterized by the absence of any genuine



power law behavior in the relevant quantities. Notice also that, for regular lattices,  $d = z/2$ . If we then generalize the condition  $m \geq d$  to the condition  $m \geq z/2$ , we are allowed to include the triangular lattice [43] in the same class discussed in the two previous sections, as well as the face centered cubic lattice (fcc), as simulation results suggest [45] (it would be interesting to determine rigorously whether  $m \geq z/2$  is a general requirement for belonging to the second class of models for any lattice geometry). A question arises as to whether, depending on the lattice geometry and the dimensionality, there can be other classes with behavior different from the two classes above. The rationale for this expectation has three motivations. One is the obvious observation that there are dimensions and lattice geometries for which some values of the parameter  $m$  are not covered by either class, and we cannot apply the general theoretical frameworks of above. Examples are the fcc lattice with  $m = 2, 3, 4, 5$  ( $k = 3, 4, 5, 6$ ) or the  $d = 4$  hypercube lattice with  $m = 2, 3$  ( $k = 3, 4$ ). A second motivation is that as we admit conditions for culling that include an extended neighborhood larger than the set of nearest neighbors, different classes of behavior do appear to arise [63,67]. The third motivation stems from the earliest results that were obtained within bootstrap percolation calculations on Bethe lattices [20], which were unexpected in nature as will be discussed briefly. The Bethe lattices are graphs with no closed loops and therefore many geometric specific correlations associated to Euclidean lattices are washed away. On the other hand, they are often regarded as mean-field models for the associated transitions (in this case the bootstrap transitions). In this sense, they might be seen as characteristic of genuine bootstrap behavior in infinite dimensions. Depending on the lattice geometry and on the dimension, remnants of the ideal Bethe-like behavior might be observed.

In a Bethe lattice (or Cayley tree) with coordination number  $z$ , every site is linked to  $z$  other sites. In turn each one of those sites is linked to  $z - 1$  other sites all disconnected with each other, and so on in a hierarchical fashion. As usual, each site is initially populated by particles with random probability  $p$ . The bootstrap percolation rule is unchanged, i.e. we assign a rule so that at each iteration occupied sites that have  $m$  or less occupied sites among the  $z$  neighboring are removed, while those with  $k = m + 1$  or more that are occupied remain occupied. After many iterations we check for the presence of a spanning cluster formed by the remaining particles. For  $m = 0$ , only isolated sites are removed, and we have as usual standard nearest neighbor percolation on the Bethe lattice, with a critical concentration  $p_c = 1/(z - 1)$  and a probability  $P_\infty(p)$  that a site is occupied by a particle

belonging to the spanning cluster that vanishes linearly in  $p - p_c$ , when  $p$  approaches  $p_c$  from above. Also for the case  $m = 1$   $P_\infty(p)$  vanishes at  $p_c$ , but this time with a non universal exponent of 2 [20]. Naturally, the case  $m = z - 1$  ( $k = z$ ) is trivial, because for any  $p < 1$  all particles can be removed one after the other. The most interesting case is when  $2 \leq m < z - 1$  ( $2 < k \leq z - 1$ ). One finds that for any given  $p$  the probability  $P_\infty(p)$  that a site hosts a particle belonging to the infinite  $k$ -cluster is given by,

$$P_\infty(p) = 0 \quad p < p_c \\ \simeq a + b(p - p_c)^{\frac{1}{2}}, \quad p \geq p_c, p \simeq p_c \quad (21)$$

We have therefore a discontinuity at  $p_c$ , and concurrently the order parameter has critical behavior when  $p \rightarrow p_c^+$ . Noticeably,  $\partial P_\infty / \partial p \simeq (p - p_c)^{-1/2}$  when  $p$  approaches  $p_c$  from above. Furthermore, a correlation length diverging at  $p_c$  (identified as the susceptibility) has been found to exist [63], thus showing that the transition on the Bethe lattice for  $m > 2$  is of a hybrid nature.

To proceed for the general case, we may define the probability  $R$  that a given site is *not* connected jointly with one specified neighbor to a  $k$ -cluster that may eventually span the system. This probability  $R$  must satisfy a condition of self-consistency of the type,

$$R = 1 - p + p \sum_{n=0}^{m-1} \binom{z-1}{n} R^{z-1-n} (1-R)^n \quad (22)$$

We also define  $T = 1 - R$  and we can illustrate the result [20,49,64] in Eq. (21) briefly for the case  $z = 4$  and  $m = 2$  as a concrete explanatory example. In this case all the particles in the surviving spanning cluster must have 3 or more occupied neighbors. We can imagine a hierarchical organization of connected links in which 3 links from upward merge into one site that has one other link going downward, and that will itself merge with 2 other links from sideways, and so on.  $T$  ( $= 1 - R$ ) is the probability that a site is connected to at least 2 among the 3 sites merging from above it, all of which are themselves connected to at least 2 of the 3 sites above, and so on indefinitely. If  $T = 0$ , we immediately conclude that there cannot be any spanning cluster with connectivity 3. On the other hand, if  $T > 0$ , the opposite is true. This is not straightforward, but it suffices to realize that if a site is occupied, and if at least 3 of the 4 nearest neighbors are each part of the substructure defined by the probability  $T$ , it follows that that particle must necessarily belong to a spanning cluster of connectivity 3. The probability  $T$  must satisfy the following relation of self-consistency,

$$T = p[T^3 + 3T^2(1 - T)] \quad (23)$$

which is equivalent to Eq. (22) above for  $R = 1 - T$ , while the probability that a site belongs to the spanning cluster is given by,

$$P_\infty = p[T^4 + 4T^3(1 - T)] \quad (24)$$

Equation (23) always admits the solution  $T = 0$  for any  $p$ . We can eliminate  $T$  on both sides and seek for the solutions with  $T \neq 0$ . We must solve,

$$2T^2 - 3T + \frac{1}{p} = 0 \quad (25)$$

which has as a physical solution the only real root such that  $T = 1$  when  $p = 1$ , which is

$$T = \frac{3}{4} + \sqrt{\frac{9}{16} - \frac{1}{2p}} \quad (26)$$

As we see, for any  $p < 8/9$ ,  $T$  has no non-zero real solutions. It has the solution  $T = 3/4$  when  $p = p_c = 8/9$ , which shows the discontinuity explicitly. We can write Eq. (26) in the convenient form,

$$T = \frac{3}{4} + \sqrt{\frac{p - p_c}{2pp_c}} \quad (27)$$

which takes the desired form, as  $p \rightarrow p_c^+$ ,

$$T = \frac{3}{4} + \frac{1}{p_c \sqrt{2}} (p - p_c)^{\frac{1}{2}} \quad (28)$$

From the above expression for  $p \rightarrow p_c^+$  combined with Eq. (24) we can easily verify that the relation (21) holds for  $z = 4$  and  $m = 2$ .

### Probabilistic Bootstrap Percolation

Regular lattices have the disadvantage that the coordination number  $z$  is relatively small, and we do not have a very wide spectrum of choices for the parameter  $m$ . One way to circumvent this problem in planar lattices is to extend the neighborhood of influence of a given site, so as to artificially enlarge the formal coordination number at each site. In this way, the bootstrap rule for removal of one particle is set to depend on the state (occupied or vacant) of many sites, beyond the mere nearest neighbors. Mixed transitions arise in these cases [63,67], although not exactly expressible in a form quantitatively identical to that given in Eq. (21).

Another approach that was taken by [16] was to introduce a probabilistic bootstrap model. This involves a simple extension to the model by considering mixtures of particles with different bootstrap rules. The particles are again

occupied with density  $p$ , but each particle is assigned its own threshold  $m_i$  as a criterion for removal during the bootstrap procedure. The author of [16] studied the case of a simple mixture of two types:  $m$  or  $m'$  with probability  $r$  and  $1 - r$  respectively. The motivation is to study non-integer values of  $m$  by controlling the mixing parameter  $r$ . The model is exactly solved and the results show that we can recover the full range of behavior from a continuous transition ( $\beta = 1$ ) to the standard discontinuous or hybrid transition ( $\beta = 1/2$ ).

The interesting cases are the ones which interpolate between continuous  $m < 3$  and discontinuous transitions  $m' \geq 3$  for the simple cubic lattice.

Using mixtures of particles with different rules is a promising approach for investigating the nature of bootstrap transition on real lattices [46] in 2 and 3 dimensions where simple nearest-neighbor models do not conclusively show the hybrid transition of Eq. (21).

Another case that has been studied is the  $m = 2$  case on the square lattice in a polluted environment [33], i. e. in the presence of a random fraction  $s$  of sites that are neither occupied nor empty. The occupied sites can be removed only if at least 2 neighboring sites are empty, and thus polluted sites may sometimes obstruct the culling process. It is found that even at low pollution  $s$ , if additionally  $q < \sqrt{s}$  then with finite probability many particles survive the bootstrap process, no matter how large the system.

The general result in Eq. (21) is suggestive also for other reasons. An identical relation (with the same exponent) to that of Eq. (21) holds if we replace  $P_\infty$  with the average number of touching hard spheres in the proximity of the jamming transition, at a given packing fraction [52]. Similar correspondences can be made with other quantities, still bearing the same exponents. It has been natural to investigate whether similar hybrid transitions as shown in the Bethe lattice survive if we pass from the loop-less case back to bootstrap percolation on Euclidean lattices. Undoubtedly this cannot be true in general, for Eq. (21) is valid for any  $z$  and for any  $m \geq 2$ , while we know already from the previous sections that Eq. (21) does not hold for the square lattice when  $m = 2$  and for the simple cubic lattice in the case  $m = 3$ . Furthermore, when very large system sizes are tested, it appears that also for the case  $m = 2$  for the simple cubic lattice all exponents lie in the same universality classes of standard percolation [17] (i. e. of continuous phase transitions). The former case ( $m = 2$  simple cubic) is particularly interesting. The critical concentration is estimated to lie in the vicinity of  $p_c \simeq 0.573$ , higher than that of the standard  $m = 0$  and  $m = 1$  value of  $p_c \simeq 0.312$ . All exponents but the exponent  $\beta$  (related to the vanishing of  $P_\infty$ ) were recognized

to be universal [4,43]. Later, results for sizes larger than tested before suggested that also  $\beta$  might be universal [17]. Therefore we conclude that we must rule out all nearest-neighbor bootstrap percolation models on regular lattices in  $d = 2, 3$  (square and simple cubic), if we want to recover the behavior as shown in Eq. (21).

### High Dimensional Lattices

Another scenario that we can envisage involves high-dimensional Euclidean lattices. It has been shown that high dimensionality expansions of the infinite dimensional Bethe lattice calculations, by means of perturbative methods [34], produce results that are qualitatively similar to Eq. (21) at sufficiently large  $d$ . In other words, for  $m \geq 2$  and for high  $d$  we might encounter hybrid mixed transitions also for regular Euclidean lattices. This result has been somewhat complemented by the work of [13] for  $m = d$  on hypercubic lattices as discussed towards the end of Sect. “Void Instability and Rigorous Results”, whereby if  $d$  grows with  $L$  at a certain (slow) rate  $d(L)$ , we find  $p_c = 1/2$ , in contrast to  $p_c = 1$  which is found for fixed  $d$  (a matter of future investigations will certainly be to determine whether  $p_c$  drops from 1 to 1/2 discontinuously for different  $d(L)$ ). Whether such hybrid transitions can be observed in high  $d$  geometries is a matter of present investigations. Simulations in the  $d = 4$  hypercubic lattice with  $m = 3$  seem to suggest [54] that a discontinuity does arise at  $p_c$ , but that this is not associated with a diverging susceptibility (which is always finite at the transition). This is a reflection of the fact that the putative critical point appears to be at a concentration lower than  $p_c$ . Thus the discontinuous jump at  $p_c$  would be purely first-order in nature.

### $k$ -core Percolation on Random Networks

The Bootstrap Percolation problem has been studied on random networks, where it is commonly known as  $k$ -core percolation. The  $k$ -core is the largest connected cluster of the network (occupation density  $p$ ) in which all the nodes have at least  $k$  neighbors, each of which is also in the sub-graph. The  $k$ -core is found by removing all nodes with less than  $k$  neighbors- this culling process is repeated until all remaining nodes (if there are any) have at least  $k$  neighbors. The  $k$ -core extends the simpler notion of the giant connected component of a network and provides a useful description of a network in terms of successively nested cores. The first exact solutions of this problem were obtained by Chalupa et al [20]. on the Bethe lattice (an infinite regular graph) as detailed above Eq. (21). After some progress was made in determining thresholds for the

emergence of a  $k$ -core in random networks [12,28,55], an exact solution was recently described by Dorogovtsev et al.[25] for uncorrelated, damaged or undamaged random networks with an arbitrary degree distribution.

We give a brief outline of the solution here. The starting point is the configuration model [14,51], an uncorrelated random network with a given degree distribution. An important point about this model is that the probability of there being more than one path between any pair of nodes is  $O(n^{-1})$  where  $n$  is the number of nodes, and so the local structure of the network is tree-like. Making use of the tree-like structure of the network it is observed that the  $k$ -core coincides with the infinite  $(k - 1)$ -ary sub-tree (in a  $k$ -ary sub-tree each node has at least  $k$  neighbors). We can therefore be certain that a randomly chosen node belongs to the  $k$ -core if at least  $k$  of its neighbors are the roots of  $(k - 1)$ -ary sub-trees. This quantity, which we call  $R$ , is the probability that a randomly chosen node is *not* the root of a  $(k - 1)$ -ary sub-tree. As  $R$  is defined it only describes one end of an edge, and is independent of whether the other end is in the  $k$ -core or not. If we choose an edge at random from the network, then it belongs to the  $k$ -core if both ends are the roots of  $(k - 1)$ -ary sub-trees- and so, with a factor of  $(1 - R)$  for each end of the edge we can express  $R$  in terms of the number of edges in the  $k$ -core,  $L_k$

$$\frac{L_k}{L} = (1 - R)^2 \quad (29)$$

The parameter  $R$  plays the role of the order parameter in  $k$ -core percolation and this expression relates it to simple measurable quantities of the networks (statistics of edges in the  $k$ -core).

Another basic quantity of interest is the relative size of the  $k$ -core  $M_k$ :

$$M_k = p \sum_{n \geq k} \sum_{q \geq n} P(q) C_n^q R^{q-n} (1 - R)^n \quad (30)$$

In this expression,  $C_n^q R^{q-n} (1 - R)^n$  is the probability that exactly  $n$  neighbors are roots and  $q - n$  are not roots of  $(k - 1)$ -ary trees (where  $C_n^q$  is the binomial coefficient) and  $P(q)$  is the probability that a node has  $q$  neighbors.

The nodes that comprise the  $k$ -core itself constitute an uncorrelated network and so their structure is encoded in the degree distribution of the  $k$ -core  $P_k(n)$  (the probability that a node in the  $k$ -core has  $n$  neighbors):

$$P_k(n) = \frac{M_k(n)}{M_k} \quad (31)$$

### Order Parameter for $k$ -core Percolation

In general, to find the physical quantities of interest, we must first solve for  $R$  for a given state of the network ( $R$  de-

depends on  $k$  and the damage  $p$ ). A given end of an edge is the root of a  $(k - 1)$ -ary sub-tree (with probability  $(1 - R)$ ), if at least  $(k - 1)$  of its children are also the roots of  $(k - 1)$ -ary sub-trees. The exact expression is [30]:

$$1 - R = p \sum_{n=k-1}^{\infty} \left[ \sum_{i=n}^{\infty} \frac{(i+1)P(i+1)}{z_1} C_n^i R^{i-n} (1-R)^n \right] \quad (32)$$

where  $z_1$  is the first moment of the degree distribution, and  $(i+1)P(i+1)/z_1$  is the number of edges of the node arrived at from a given edge. This equation bears some similarity to that for ordinary ( $k = 2$ ) percolation (see for example Eq. 1 of Chalupa et al [20]). As in  $k = 2$  percolation, it is found that there are two possible solutions:  $R = 1$ : leading to a continuous transition, and  $0 \leq R < 1$ : leading to a discontinuous/hybrid transition. The critical behavior for the  $R < 1$  case is found to be [25,30]:

$$R^* - R \propto (p - p_c)^{1/2} \quad (33)$$

where  $R^*$  is the value of order parameter at the point where the  $k$ -core vanishes. This hybrid critical behavior, characterized by a jump in the order parameter at  $p_c < 1$  and a square root singularity in the percolating phase, is shared by other quantities, such as the relative size of the  $k$ -core,  $M_k$ :

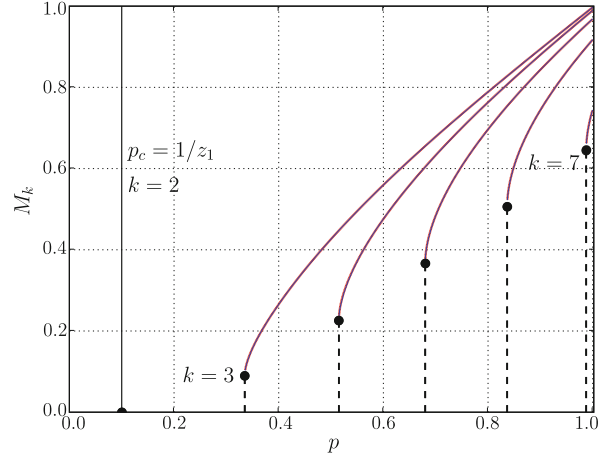
$$M_k - M_k^* \propto (p - p_c)^{1/2} \quad (34)$$

As the bootstrap process removes nodes at random, the cores are destroyed starting with the highest  $k$ -core. Each  $k$ -core collapses with the critical behavior of Eq. (33).

In Fig. 4 we show the relative sizes of all the possible  $k$ -cores in networks with Poisson degree distributions ( $z_1 = 10$ ) as a function of occupation density.

These results hold for networks with arbitrary degree distribution but  $z_2$  must also be finite. In scale free networks (degree distribution  $P(q) \propto 1/(q + c)^{-\gamma}$ ) with  $\gamma > 3$ , where  $z_2$  is finite, the critical behavior for  $k \geq 3$ -core percolation is found to be the hybrid phase transition (33). For  $2 < \gamma \leq 3$  the transition is infinite order, similar to the case for ordinary percolation in scale-free networks.

It is worth noting that the hybrid transition indicates there should be a diverging length-scale which has not been conclusively identified – it is the mean size of finite clusters which drives the conventional continuous transition, but for tree-like network models it is not obvious what the nature of divergence is since finite clusters are not supported for  $p < p_c$ .



**Bootstrap Percolation, Figure 4**

The relative size of the  $k$ -core for a classical Erdős-Rényi random network. The mean degree  $z_1 = 10$ . The case  $k = 2$  corresponds to ordinary percolation where there is a continuous transition at  $p_c = 1/z_1$ . For  $3 \leq k \leq 7$  we see that a core emerges with a discontinuous jump in  $M_k$  (see Eq. (34))

The first clues to understanding the nature of the hybrid transition were found by studying the critical behavior of the subset of the  $k$ -core known as the corona [63]. The corona is comprised of nodes in the  $k$ -core which have exactly  $k$  neighbors in the  $k$ -core (this is the minimum number they can have). The nodes in the corona are randomly distributed throughout the  $k$ -core and for  $p > p_c$  the corona is comprised of small finite clusters. It is an important observation that if a random node in the  $k$ -core is removed, then all the nodes in the corona clusters attached to that original node will be removed- leading to increasingly massive cascading removals, contingent on the removal of a single node, on approach to the transition.

The precise behavior of the mean size of clusters in the corona attached to a randomly chosen vertex in the  $k$ -core,  $N_{crn}$ , is known for random networks [25,30]. It diverges at the transition,

$$N_{crn} \propto (p - p_c)^{-1/2} \quad (35)$$

as was also found by [63] for the Bethe lattice. The picture that emerges is one of the collapse of the  $k$ -core at  $p_c$ , accompanied by a percolation transition of the corona cluster. For  $p > p_c$  the corona clusters are disconnected, while for  $p < p_c$  there is neither a  $k$ -core nor any corona clusters. As in ordinary percolation, we can identify a susceptibility belonging to the mean size of the clusters in the corona [30]. find that this susceptibility diverges also at the transition- indicating that the corona clusters undergo a percolation transition, growing in size as  $p \rightarrow p_c^{(+)}$ ,



percolating at  $p = p_c$ , and disappearing altogether for  $p < p_c$ .

There have been some attempts to apply the above picture of  $k$ -core percolation to real networks, such as the Internet (at the Autonomous System level) [8]. The identification of the  $k$ -cores provides a useful insight into the structural properties and regions of increasing centrality in these networks. Although the results are promising they are not in precise agreement with the predictions, a consequence of the fact that the theory deals only with uncorrelated networks.

### Future Directions

Future work will continue to try to understand the unusual scaling behaviour, and further develop rigorous results. The existence and the consequent classification of the transitions described above, including but not limited to hybrid mixed transitions, is likely to represent the next step for many working on bootstrap percolation transitions. The development of solutions for complex networks,  $k$ -core percolation, will lead to many new directions of interest including an understanding of the effect of correlations and loops. The broad range of applications that the bootstrap percolation problem has already found will continue to inspire fruitful research in diverse areas of physics and mathematics.

### Bibliography

- Adler J (1991) Bootstrap percolation. *Physica A* 171(3):453–470
- Adler J, Aharony A (1988) Diffusion percolation. I. Infinite time limit and bootstrap percolation. *J Phys A: Math Gen* 21(6):1387–1404
- Adler J, Lev U (2003) Bootstrap percolation: Visualisations and applications. *Braz J Phys* 33:641
- Adler J, Stauffer D (1990) Evidence for non-universal exponents in bootstrap percolation. *J Phys A Math Gen* 23:L1119
- Adler J, Palmer RG, Meyer H (1987) Transmission of order in some unusual dilute systems. *Phys Rev Lett* 58(9):882–885
- Adler J, Stauffer D, Aharony A (1989) Comparison of bootstrap percolation models. *J Phys A Math Gen* 22:L297
- Aizenman M, Lebowitz JL (1988) Metastability effects in bootstrap percolation. *J Phys A: Math Gen* 21:3801–3813
- Alvarez-Hamelin JI, Dall'Asta L, Barrat A, Vespignani A (2005)  $k$ -Core decomposition: A tool for the analysis of large scale internet graphs. eprint arXiv:cs/0504107
- Balogh J, Bollobás B (2006) Bootstrap percolation on the hypercube. *Probab Theory Relat Field* 134(4):624–648
- Balogh J, Pete G (1998) Random disease on the square grid. *Random Struct Algorithm* 13(3–4):409–422
- Balogh J, Bollobás B (2003) Sharp thresholds in bootstrap percolation. *Phys A: Stat Mech Appl* 326(3–4):305–312
- Balogh J, Pittel BG (2007) Bootstrap percolation on the random regular graph. *Random Struct Algorithm* 30(1–2):257–286
- Balogh J, Bollobás B, Morris R (2007) Majority bootstrap percolation on the hypercube. Arxiv preprint mathCO/0702373
- Bollobás B (2001) Random graphs. Cambridge University Press, Cambridge
- Branco N, Dos Santos R, de Queiroz S (1984) Bootstrap percolation: a renormalization group approach. *J Phys C* 17:1909–1921
- Branco NS (1993) Probabilistic bootstrap percolation. *J Stat Phys* 70:1035–1044
- Branco NS, Silva CJ (1999) Universality class for bootstrap percolation with  $m=3$  on the cubic lattice. *Int J Mod Phys C* 10:921–930
- Cerf R, Cirillo E (1999) Finite size scaling in three-dimensional bootstrap percolation. *Ann Probab* 27(4):1837–1850
- Cerf R, Manzo F (2002) The threshold regime of finite volume bootstrap percolation. *Stoch Process Appl* 101:69–82
- Chalupa J, Leath PL, Reich GR (1979) Bootstrap percolation on a bethe lattice. *J Phys C: Solid State Phys* 12(1):L31–L35
- Chaves C, Koiller B (1995) Universality, thresholds and critical exponents in correlated percolation. *Phys A* 218(3):271–278
- De Gregorio P, Lawlor A, Bradley P, Dawson KA (2004) Clarification of the bootstrap percolation paradox. *Phys Rev Lett* 93:025501
- De Gregorio P, Lawlor A, Bradley P, Dawson KA (2005) Exact solution of a jamming transition: Closed equations for a bootstrap percolation problem. *Proc Nat Acad Sci* 102:5669–5673
- De Gregorio P, Lawlor A, Dawson KA (2006) New approach to study mobility in the vicinity of dynamical arrest; exact application to a kinetically constrained model. *Europhys Lett* 74:287–293
- Dorogovtsev SN, Goltsev AV, Mendes JFF (2006)  $k$ -Core organization of complex networks. *Phys Rev Lett* 96:040601
- Duarte J (1989) Simulation of a cellular automaton with an oriented bootstrap rule. *Phys A* 157(3):1075–1079
- Ertel W, Frobröse K, Jäckle J (1988) Constrained diffusion dynamics in the hard-square lattice gas at high density. *J Chem Phys* 88:5027–5034
- Fernholz D, Ramachandran V (2003) The giant  $k$ -core of a random graph with a specified degree sequence. manuscript, UT-Austin, November
- Fredrickson GH, Andersen HC (1984) Kinetic ising model of the glass transition. *Phys Rev Lett* 53(13):1244–1247
- Goltsev AV, Dorogovtsev SN, Mendes JFF (2006)  $k$ -Core (bootstrap) percolation on complex networks: Critical phenomena and nonlocal effects. *Phys Rev E* 73:056101
- Gravner J, Griffeath D (1996) First passage times for threshold growth dynamics on  $\mathbb{Z}^2$ . *Ann Probab* 24(4):1752–1778
- Gravner J, Holroyd A (2008) Slow convergence in bootstrap percolation. *Ann Probab* 18(3):909–928
- Gravner J, McDonald E (1997) Bootstrap percolation in a polluted environment. *J Stat Phys* 87(3):915–927
- Harris A, Schwarz J (2005)  $1/d$  expansion for  $k$ -core percolation. *Phys Rev E* 72(4):46123
- Holroyd A (2003) Sharp metastability threshold for two-dimensional bootstrap percolation. *Probab Theory Relat Field* 125(2):195–224
- Holroyd A (2006) The metastability threshold for modified bootstrap percolation in  $d$  dimensions. *Electron J Probab* 11:418–433
- Holroyd A, Liggett T, Romik D (2004) Integrals, partitions and cellular automata. *Trans Amer Math Soc* 356:3349–3368



38. Ising E (1925) Beitrag zur Theorie des Ferromagnetismus. *Z Phys* 31:253–258
39. Jäckle J, Krönig A (1994) A kinetic lattice-gas model for the triangular lattice with strong dynamic correlations: I. cself-diffusion. *J Phys: Condens Matter* 6:7633–7653
40. Jeng M, Schwarz JM (2007) Comment on “jamming percolation and glass transitions in lattice models”. *Phys Rev Lett* 98(12):129601
41. Kirkpatrick S, Wilcke W, Garner R, Huels H (2002) Percolation in dense storage arrays. *Physica A: Stat Mech Appl* 314(1–4):220–229
42. Kob W, Andersen HC (1993) Kinetic lattice-gas model of cage effects in high-density liquids and a test of mode-coupling theory of the ideal-glass transition. *Phys Rev E* 48:4364–4377
43. Kogut PM, Leath PL (1981) Bootstrap percolation transitions on real lattices. *J Phys C* 14(22):3187–3194
44. Kurtsiefer D (2003) Threshold value of three dimensional bootstrap percolation. *Int J Mod Phys C* 14:529
45. Lawlor A, De Gregorio P, Bradley P, Sellitto M, Dawson KA (2005) Geometry of dynamically available empty space is the key to near-arrest dynamics. *Phys Rev E* 72:021401
46. Lawlor A, De Gregorio P, Cellai D, Dawson KA (2008) (to be published)
47. Manna SS (1998) Abelian cascade dynamics in bootstrap percolation. *Physica A: Stat Theor Phys* 261:351–358
48. Medeiros M, Chaves C (1997) Universality in bootstrap and diffusion percolation. *Physica A* 234(3):604–610
49. Moukarzel C, Duxbury PM, Leath PL (1997) Infinite-cluster geometry in central-force networks. *Phys Rev Lett* 78(8):1480–1483
50. Mountford TS (1995) Critical length for semi-oriented bootstrap percolation. *Stoch Process Appl* 56:185–205
51. Newman MEJ (2003) The structure and function of complex networks. *SIAM Rev* 45:167
52. O’Hern C, Silbert L, Liu A, Nagel S (2003) Jamming at zero temperature and zero applied stress: The epitome of disorder. *Phys Rev E* 68(1):11306
53. Onsager L (1944) Crystal statistics. I. A two-dimensional model with an order-disorder transition. *Phys Rev* 65(3–4):117–149
54. Parisi G, Rizzo T (2006) On  $k$ -core percolation in four dimensions. *Arxiv preprint cond-mat/0609777*
55. Pittel B, Spencer J, Wormald N (1996) Sudden emergence of a giant  $k$ -core in a random graph. *J Comb Theory B* 67:111–151
56. Pollak M, Riess I (1975) Application of percolation theory to 2D–3D Heisenberg ferromagnets. *Phys Stat Solidi (B)* 69(1):K15–K18
57. Ritort F, Sollich P (2003) Glassy dynamics of kinetically constrained models. *Adv Phys* 52:219–342
58. Sabhapandit S, Dhar D, Shukla P (2002) Hysteresis in the random-field Ising model and bootstrap percolation. *Phys Rev Lett* 88(19):197202
59. Sahimi M (1994) Applications of percolation theory. Taylor/Francis, London
60. Schonmann R (1990) Critical points of two-dimensional bootstrap percolation-like cellular automata. *J Stat Phys* 58(5):1239–1244
61. Schonmann R (1990) Finite size scaling behavior of a biased majority rule cellular automaton. *Phys A* 167(3):619–627
62. Schonmann R (1992) On the behavior of some cellular automata related to bootstrap percolation. *Ann Probab* 20(1):174–193
63. Schwarz JM, Liu AJ, Chayes LQ (2006) The onset of jamming as the sudden emergence of an infinite  $k$ -core cluster. *Europhys Lett (EPL)* 73(4):560–566
64. Sellitto M, Biroli G, Toninelli C (2005) Facilitated spin models on Bethe lattice: Bootstrap percolation, mode-coupling transition and glassy dynamics. *Europhys Lett* 69(4):496–502
65. Smirnov S, Werner W (2001) Critical exponents for 2D percolation. *Math Res Lett* 8:729–744
66. Stauffer D, Aharony A (1992) Introduction to percolation theory. Taylor/Francis, London
67. Toninelli C, Biroli G, Fisher DS (2006) Jamming percolation and glass transitions in lattice models. *Phys Rev Lett* 96(3):035702
68. Toninelli C, Biroli G, Fisher D (2007) Toninelli, Biroli, and Fisher reply. *Phys Rev Lett* 98(12):129602
69. Treaster M, Conner W, Gupta I, Nahrstedt K (2006) Contagalert: Using contagion theory for adaptive, distributed alert propagation. *nca* 0:126–136
70. van Enter A (1987) Proof of Straley’s argument for bootstrap percolation. *J Stat Phys* 48(3):943–945
71. van Enter A, Hulshof T (2007) Finite-size effects for anisotropic bootstrap percolation: logarithmic corrections. *J Stat Phys* 128(6):1383–1389
72. van Enter ACD, Adler J, Duarte JAMS (1990) Finite size effects for some bootstrap percolation models. *J Stat Phys* 60:323–332
73. van Enter A, Adler J, Duarte J (1991) Finite-size effects for some bootstrap percolation models, addendum. *J Stat Phys* 62:505–506
74. Widom B (1974) The critical point and scaling theory. *Phys* 73(1):107–118
75. Wilson K (1983) The renormalization group and critical phenomena. *Rev Mod Phys* 55:583–600

---

## Brain Pacemaker

PETER A. TASS<sup>1,2,3</sup>, OLEKSANDR V. POPOVYCH<sup>1</sup>,  
CHRISTIAN HAUPTMANN<sup>1</sup>

<sup>1</sup> Institute of Neuroscience and Biophysics 3 –  
Medicine and Virtual Institute of Neuromodulation,  
Research Center Jülich, Jülich, Germany

<sup>2</sup> Department of Stereotaxic and Functional  
Neurosurgery, University of Cologne,  
Cologne, Germany

<sup>3</sup> Brain Imaging Center West, Research Center Jülich,  
Jülich, Germany

## Article Outline

[Glossary](#)

[Definition of the Subject](#)

[Introduction](#)

[Standard High-Frequency Stimulation](#)

[Multisite Coordinated Reset Stimulation](#)

[Linear Multisite Delayed Feedback](#)

Nonlinear Delayed Feedback  
 Proportional-Integro-Differential Feedback  
 Plasticity  
 Summary  
 Bibliography

## Glossary

**Electrical deep brain stimulation (DBS)** is the standard therapy for medically refractory movements disorders, e. g., Parkinson's disease and essential tremor. It requires a surgical treatment, where depth electrodes are chronically implanted in target areas like the thalamic ventralis intermedius nucleus or the subthalamic nucleus and, as yet, an electrical high-frequency ( $>100$  Hz) stimulation is permanently delivered via depth electrodes. More sophisticated deep brain stimulation techniques are in the process of being established for clinical use.

**Delayed feedback** Delayed feedback is a method for the creation of a closed-loop forcing, where some portion of the measured output signal of a system is time delayed, linearly or non-linearly processed, and fed back into the system. This approach is often used to control the dynamic behavior of complex systems. In this article delayed feedback is used to control synchronization in ensembles of coupled oscillators, e. g., neurons.

**Order parameter** The order parameter is a quantity characterizing a phase transition or phase change in the transformation of a complex system from one phase (state) to another. The order parameter is convenient for characterizing the onset and extent of synchronization in larger ensembles: Perfect phase synchronization corresponds to a large value of the order parameter, whereas an incoherent (desynchronized) state is associated with a small value of the order parameter. In synergetics it has been shown that the dynamics of complex systems may be governed by only a few order parameters.

**Synchronization** Synchronization (from Greek *syn* = the same, common and *chronos* = time) means the adjustment of rhythms of self-sustained oscillators due to their weak interaction. The interacting oscillators can be regular (periodic) or chaotic. There are several different forms of synchronization including phase, complete, generalized, and lag synchronization, etc. In this article we focus on phase synchronization. In the simplest form, the oscillators, rotating with the same frequency, become phase synchronized (phase locked) to each other, if they tend to oscillate with the same repeating sequence of relative phase angles. Put other-

wise, the oscillators adjust their rhythms, while their amplitude dynamics need not be correlated.

## Definition of the Subject

A brain pacemaker is a medical device that is implanted into the brain with the purpose to stimulate nervous tissue with electrical signals. Brain pacemakers are used for the therapy of patients suffering for example from Parkinson's disease, epilepsy or mental disorders. Brain stimulation is either called deep brain stimulation (DBS) if structures deeply inside the brain are targeted or Stimulation, if the electrical contacts of the stimulator are positioned on the surface of the brain. Apart from direct brain stimulation, brain pacemakers may also be used to stimulate the spinal cord (e. g. for the treatment of pain) or the vagus nerve (for the treatment of epilepsy). The electrical stimulation of the nervous system has a long history which goes back to the 19th century where first tests with cortical stimulation were documented [23]. The first intraoperative deep brain stimulation was performed by Spiegel et al. in 1947 in a patient suffering from Huntington's chorea and in the eighties DBS was introduced as a treatment for motor disorders [4,11]. DBS was approved by the FDA as a treatment for essential tremor in 1997, for Parkinson's disease in 2002, and dystonia in 2003. The treatment of severe neurological and psychiatric diseases with brain pacemakers is a rapidly growing and promising field. Novel, model-based approaches, which use methods from synergetics, nonlinear dynamics, and statistical physics, to specifically restore brain function and connectivity, demonstrate how insights into the dynamics of complex systems contribute to the development of novel therapies.

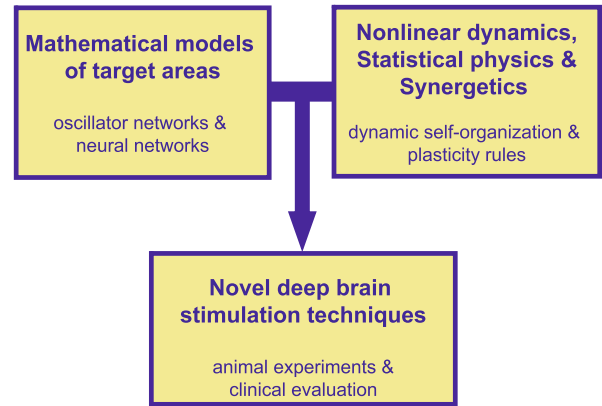
## Introduction

Self-organization processes are abundant in numerous fields of the natural sciences [28,29]. For instance, the nervous system elegantly utilizes self-organization principles for motor control purposes [30,32,43,72]. A classical example of a self-organization process is synchronization of populations of interacting oscillators, which is widely observed in physics [27,29,60,81], chemistry [46], biology [102], neuroscience [31,80], and medicine [16,54,86]. In the nervous system synchronization processes are important, e. g., in the context of information processing [76] and motor control [3]. However, pathological, excessive synchronization strongly impairs brain function [16,54]. In fact, pathological synchronization processes are the hallmark of several neurological diseases like Parkinson's disease (PD) or essential tremor [2,58]. For example, Parkinsonian resting tremor is caused by a pacemaker-like

population of neurons which fires in a synchronized and periodical manner [2,77]. In contrast, in healthy subjects these neurons fire in an uncorrelated, i. e. desynchronized manner [58].

Permanent deep brain stimulation (DBS) at high frequencies ( $> 100$  Hz) is the standard therapy for medically refractory patients suffering from Parkinson's disease and essential tremor [5,6,10], see Fig. 1. High-frequency DBS has been developed empirically, mainly based on experimental results and clinical observations. The mechanism of high-frequency DBS is still a matter of debate [7]. Clinical studies showed that high-frequency DBS essentially has similar effects as observed after tissue lesioning. Compared to lesioning, DBS is reversible and has a lower rate of side effects [71,82]. However, in spite of many beneficial effects, in some patients DBS may not help, or may cause side effects, or the therapeutic effects may disappear over time [14,20,66,82,101]. With the objective of finding better tolerated and more effective DBS techniques, a model-based development of novel stimulation methods has been initiated [37,38,39,41,62,63,64,86,91,93,94,98], see Fig. 2. In these studies, relevant neuronal target populations were modeled mathematically and stimulation techniques have been developed utilizing principles from nonlinear dynamics and statistical physics [86].

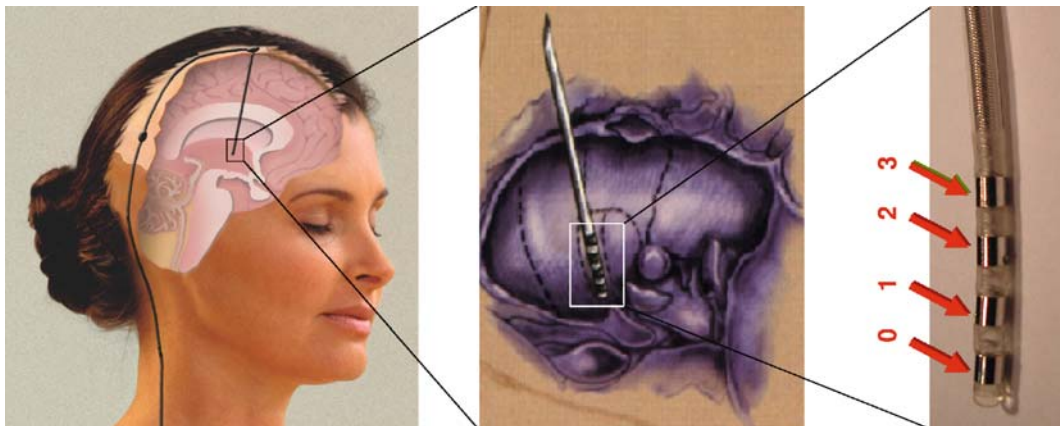
One goal of this approach is to control the pathological neural dynamics appropriately in order to achieve a mild and efficient relief of symptoms [86]. The second, more ambitious goal is to stimulate in a way that the formerly affected neuronal populations unlearn their pathological connectivity and, hence, their tendency to produce pathological synchronization [97]. Put otherwise, the very goal



**Brain Pacemaker, Figure 2**

*Model-based development of novel deep brain stimulation techniques:* Along the lines of a top-down approach target areas for deep brain stimulation are modeled by means of oscillator networks and physiology and anatomy based neural networks. Methods from nonlinear dynamics, statistical physics, and synergetics are employed to develop stimulation techniques which specifically utilize dynamic self-organization principles and plasticity rules. Experimental feedback from both animal experiments and clinical evaluation serves to validate, falsify or modify theoretical assumptions and predictions. This iterative approach aims at steadily improving the mathematically designed stimulation techniques and, hence, at establishing superior therapies

of this approach is to induce long-lasting therapeutic effects which outlast the cessation of stimulation [36,83,97]. To this end, stimulation algorithms have been developed and optimized to exploit dynamic self-organization principles and plasticity rules [36,41,83,96,97].



**Brain Pacemaker, Figure 1**

Standard DBS setup. A depth electrode is implanted into the target structure (e. g. the *subthalamic nucleus*). The electrode is subcutaneously connected with the generator of the high-frequency stimulation signal (not shown in this image). The stimulation signal is delivered through one or more of the four stimulation contacts labeled from 0 to 3

Three novel, particularly promising stimulation techniques have been developed in the past: The multisite coordinated reset stimulation (MCRS) [86,91,93,94], the multisite coordinated delayed feedback stimulation (MDFS) [37,38,39], and the nonlinear delayed feedback stimulation (NDFS) [62,63,64]. These techniques have the common objective of reducing the synchronized activity of the target population by reestablishing a normal physiological activity in a highly synchronized population of neurons.

MCRS uses short pulse trains to reset sub-populations of the neuronal network subsequently which induces a desynchronized state [86,91,93,94]. Here, the stimulation was applied through a small number of stimulation sites which were equally spaced within the neuronal population. MCRS induced desynchronization is achieved by utilizing self-organization principles, in particular, the slaving principle induced by the pathological neuronal interactions (i.e. interactions which have the potential to induce a pathological synchronization) [91,93,94]. MDFS [37,38,39] and NDFS [62,63,64] use delayed feedback for stabilizing a desynchronized state which was intended to be as close to the physiological mode of action as possible. Here, the local field potential (LFP) of the target population is measured, amplified, delayed, and fed back into the ensemble.

However, a novel, clinically extremely relevant aspect comes into play by considering the effect of such stimulation techniques on synaptic plasticity [97]. It has been shown experimentally, that synaptic plasticity enhances the pathological synchronization [59]. From the kindling phenomenon in the context of epilepsy it is well-known that neural networks may learn pathological strong interactions [56,79]. The novel desynchronizing stimulation protocols are designed to invert this pathological process, so that the affected neuronal populations unlearn their pathological connectivity, and physiological neuronal activity is re-established on a long-term basis. In a nutshell, the novel stimulation techniques aim at a well directed employment of fundamental principles of dynamic brain action to induce long-lasting therapeutic effects.

### Standard High-Frequency Stimulation

High-frequency deep brain stimulation (HFDBS) is the standard therapy for patients suffering from medically refractory PD or essential tremor [5,6]. To this end, depth electrodes are chronically implanted in the thalamic ventralis intermedius nucleus or the subthalamic nucleus [5,6] and a permanent high-frequency (> 100 Hz) periodic pulse train stimulation is applied (Fig. 1). HFDBS

has been developed empirically, mainly based on intra-operative observations. HFDBS strongly alters the neuronal firing and mimics the effect of tissue lesioning, e.g., by suppressing neuronal firing, which, in turn, suppresses the peripheral tremor [6,19,53,101]. However, as yet, the mechanism of HFDBS is not sufficiently understood [53].

During stimulation, high-frequency DBS seems to induce a regular bursting mode [9]. After a reduction of stimulation artifacts, robust bursting activity in subthalamic nucleus (STN) neurons was observed in slices from naive or reserpine-treated rats [8]. After offset of stimulation, a blockade of activity, i.e. a non-specific suppression of the neuronal activity in the target structures through a depolarization blockade which did not require synaptic transmission was observed [8].

Other hypotheses are that high-frequency DBS applied to a PD patient mimics the effect of tissue lesioning and appears to block neuronal activity in relevant target areas during stimulation [6]. The most probable hypothesis was offered recently by Benabid et al. [7], in which a mixture of different mechanisms was discussed. The contributing mechanisms resulting in the observed effects of high-frequency DBS might be membrane inhibition, jamming, excitation of excitatory and inhibitory afferents, excitation of efferents and plasticity [7].

To precisely evaluate the contribution of these different mechanisms, spatially extended multi-compartment neuron models were used to demonstrate the effects of extracellular stimulation on the different structures of the stimulated neuronal population [26]. Depending on the stimulation amplitude and the shape of the stimulation pulses, either the cells were activated directly or fibers mediating excitatory or strong inhibitory action were activated [26]. Modeling studies indicate that already at the level of single neurons, the activation of a larger number of structures can take place with different and possibly conflicting impacts on the single neuron dynamics [26]. The collective dynamics of neuronal populations further adds aspects which are important for the creation of synchronized activity: Cells responding differently to external inputs like somatosensory stimulation or stimulation due to active movements are present in the target tissue together with so called no-response cells [47]. High-frequency stimulation has a complex impact on these structures [6,74].

Recent experimental and modeling studies indicate that the *globus pallidum interior* GPi – one structure of the basal ganglia – might be strongly involved in the mechanisms of DBS [21,35,52,55,69]. The results of modeling studies indicate that under Parkinsonian conditions the rhythmic inhibition from GPi to the thalamus com-



promises the ability of thalamocortical relay cells to respond to depolarizing inputs, such as sensorimotor signals. High frequency stimulation of STN regularizes GPi firing, and this restores the responsiveness of the thalamus [69].

HF DBS is reversible and has a much lower rate of side effects than lesioning with thermocoagulation [71]. Although HF DBS is the golden standard for the therapy of medically refractory movement disorders, there are still limitations of HF DBS: On the one hand HF DBS may cause adverse effects like dysarthria, dysesthesia, cerebellar ataxia, and memory decline [20,66,101]. On the other hand HF DBS may be ineffective or its therapeutic effect may wear off over time [45,66]. For instance, 11–15% of PD patients have unsatisfactory outcomes although their depth electrodes are properly placed [48].

### Multisite Coordinated Reset Stimulation

To study the impact of pulsatile stimuli on single oscillators and, in particular, populations of oscillators in a biologically more realistic setting, it was necessary to take into account random forces [84,85,86]. To this end, a stochastic concept of phase resetting has been developed for populations of non-interacting [84,85] as well as interacting [86] oscillators in the presence of noise. In this approach limit-cycle oscillators were approximated by phase oscillators [34], so that the pulsatile stimulation only affects the oscillators' phases. If a single-pulse of the right intensity and duration is delivered to the population in the stable synchronized state, it causes an at least temporary desynchronization provided it hits the population at a vulnerable phase. Theoretically, single-pulse stimulation has also been studied in more complex networks, for instance, networks of coupled phase oscillators with inertia, modeling dendritic dynamics [15,50]. Based on the stochastic phase resetting theory and utilizing a phase oscillator as a model for a single neuron [34], demand-controlled single-pulse deep brain stimulation has been suggested for the therapy of movement disorders like Parkinson's disease or essential tremor [86,87].

However, there are drawbacks to single-pulse stimulation which decisively limit its applicability [89,104]: First, if the mutual coupling is not weak, the vulnerable phase range we have to hit in order to cause an effective desynchronization is only a small fraction (e.g., 5%) of a period of the collective oscillation. Second, the critical stimulation parameters required to achieve a good desynchronization depend on the dynamical state of the population. Thus, different stimulation parameters have to be used if the cluster is not in its stable synchronized state.

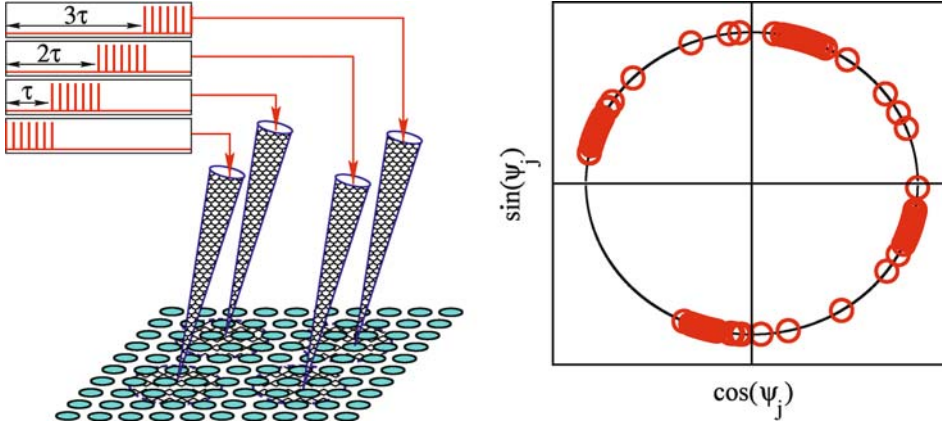
To overcome the limitations of single-pulse stimulation, double-pulse [88,89] stimulation has been proposed: Two qualitatively different stimuli are successively delivered. The first, stronger pulse resets (restarts) the collective oscillation irrespective of the initial state of the population. The second, weaker pulse is applied after a fixed time delay, where it hits the cluster in its vulnerable state and, hence, causes a desynchronization. There are different variants of double-pulse stimulation, depending on the type of stimuli used to achieve a reset or a desynchronization [90,91,92,93]. For instance, the first resetting pulse can be replaced by a brief high-frequency pulse train [90] or by a softly resetting low-frequency pulse train [91,92].

Although double-pulse stimulation causes a desynchronization irrespective of the initial dynamical state at which this stimulus is delivered, there are still limitations which may hinder an application to a biological system [94,95]: On the one hand, double-pulse stimulation requires a calibration. On the other hand, double-pulse stimulation is only effective if the system parameters are reasonably stable. The required quasi-stationarity of the system parameters combined with the possibly time consuming calibration may cause problems when applied to a real biological system, where fluctuations of model parameters are inevitable.

To provide a stimulation technique which is robust with respect to system parameters and which does not require calibration, multisite coordinated reset stimulation has been developed [94,95]. The idea behind this approach is to abstain from achieving a perfect desynchronization by a well-calibrated stimulus. Rather, by means of a robust and comparably mild stimulus the stimulated population is shifted into a dynamical state which is not the desired desynchronized state, but sufficiently close to it. Close in the sense that due to the pathologically strong coupling the population automatically relaxes into the desired desynchronized state. This approach essentially exploits the pathological tendency of the neuronal population to establish a synchronized state. Accordingly, multisite coordinated reset stimulation is in a way comparable to Asian martial arts, where ideally a minimal amount of energy (i.e. the multisite coordinated reset) is invested to control the adversary by utilizing the adversary's own energy (i.e. the neurons' pathological strong coupling).

The scheme of the stimulation setup is presented in Fig. 3 (left plot). Several stimulation sites are placed within the target network and weak resetting stimulation signals are administered via these stimulation sites. In this way the oscillatory population is split into several sub-populations, where each of them is assigned to the corresponding stimulation site and receiving the stimulation signal





Brain Pacemaker, Figure 3

Stimulation setup of the multisite coordinated reset stimulation method. Brief and mild resetting stimuli are administered at different sites at subsequent times (*left plot*) and cause an effective splitting of the stimulated population into several sub-populations such that their phases  $\psi_j$  form phase clusters distributed uniformly (or close to that) over the period of oscillations (*right plot*)

mostly from that stimulation site. Multi-site coordinated reset stimulation means that a synchronized population of neurons is stimulated with a sequence of brief resetting stimuli (typically brief high-frequency (HF) stimulus trains) via the different sites. The delay between the subsequent resetting stimuli can be chosen as  $\tau = T/n$  with respect to that at the preceding site, where  $n$  is the number of the stimulation sites,  $T$  approximates the mean period of the collective dynamics of the synchronized oscillators, and  $n$  is the number of stimulation sites [94,95].

The subsequent reset of the different sub-populations induces a so-called cluster state, i. e. the whole population splits into  $n$  sub-populations which differ with respect to their mean phase. This effect is illustrated in Fig. 3 (right plot), where a snapshot of the distribution of the phases  $\psi_j$  of stimulated oscillators is shown in the  $(\cos(\psi_j), \sin(\psi_j))$  plane after coordinated reset stimulation. The phases of the population of oscillators stimulated via, e. g., four sites form four phase clusters distributed uniformly (or close to that) over the unit circle. To estimate the extent and type of synchronization of the whole population of  $N$  oscillators, the cluster variables

$$Z_m(t) = R_m(t)e^{i\Psi_m(t)} = \frac{1}{N} \sum_{j=1}^N e^{im\psi_j(t)}, \quad (1)$$

can be used, where  $R_m(t)$  and  $\Psi_m(t)$  are the corresponding real amplitude and real phase, where  $0 \leq R_m(t) \leq 1$  for all time  $t$  [12,86]. Cluster variables are convenient for characterizing synchronized states of different types: Perfect in-phase synchronization corresponds to  $R_1 = 1$ , whereas an incoherent state, with uniformly distributed phases, is associated with  $R_m = 0$  ( $m = 1, 2, 3, \dots$ ).  $R_1 = 0$  com-

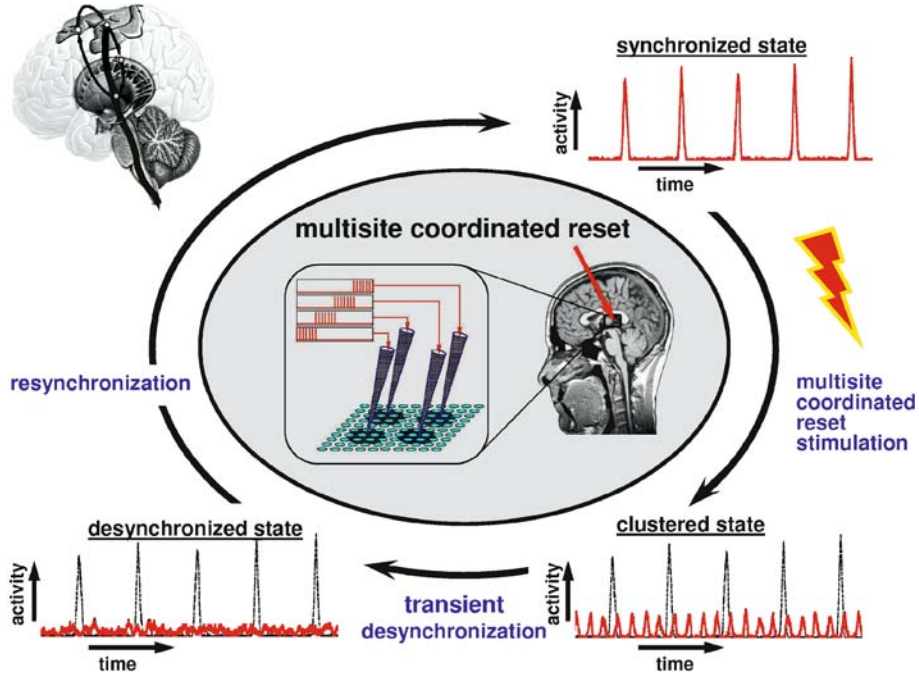
bined with large  $R_m$  is indicative of an  $m$ -cluster state consisting of  $m$  distinct and equally spaced clusters, where within each cluster all oscillators have similar phase. In Fig. 3 (right plot), for instance,  $R_1 \approx 0.02$ , whereas  $R_4 \approx 0.87$  indicating a four-cluster state induced by the multisite coordinated reset stimulation.

From the cluster state the neurons typically relax to a uniformly desynchronized state before they revert back to the in-phase synchronized state, if left unperturbed. To understand how a stimulus-induced clustering leads to an effective desynchronization, the dynamics of the leading modes  $Z_1, Z_2, \dots$  can be considered. When the coupling among oscillators becomes sufficiently large, e. g., it exceeds a certain critical value,  $Z_1$  from (1) becomes an *order parameter* [46], which according to the slaving principle [29] governs the dynamics of the other, stable modes  $Z_m$  ( $m = 2, 3, \dots$ ) on the center manifold [61]: The order parameter  $Z_1$  acts on a slow time scale, whereas the stable modes  $Z_m$  act on a fast time scale and relax to values given by the order parameter  $Z_1$  [29,103]. In a system with large number of oscillators this relationship reads [86]:

$$R_m \propto R_1^\nu \quad \text{with} \quad \nu \geq 2, \quad m = 2, 3, 4, \dots \quad (2)$$

Hence, to maintain a desynchronized neuronal firing, multisite coordinated reset stimuli have to be administered repetitively.

Multisite coordinated reset stimulation exploits transient responses which are due to the oscillators' (pathologically strong) interactions. The general stimulation protocol of the coordinated reset stimulation is illustrated in Fig. 4. Here, the collective dynamics is visualized by considering the collective firing of the neurons. A single firing/bursting model neuron fires/bursts whenever its phase



Brain Pacemaker, Figure 4

A general scheme of the stimulation with multisite coordinated reset. Desynchronized firing of neurons is maintained by repetitive administration of multisite coordinated reset stimuli intermingled with epochs of no stimulation

is close to zero (modulo  $2\pi$ ) [17,25,33,46,86]. The collective firing can be illustrated with the *relative number of neurons producing an action potential or burst at time  $t$*  given by

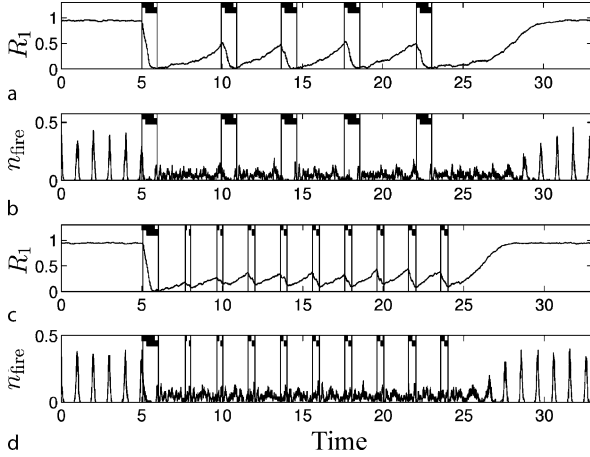
$$n_{\text{fire}}(t) = \frac{\text{number of neurons with } \cos \psi_j > 0.99}{N}. \quad (3)$$

$0 \leq n_{\text{fire}}(t) \leq 1$  for all  $t$ .  $n_{\text{fire}}(t) = 0$  means that no neuron fires/bursts, while all neurons fire/burst at time  $t$  if  $n_{\text{fire}}(t) = 1$ . Varying the threshold parameter 0.99 in a reasonable range does not change the results. As shown in Fig. 4, stimulation starts when the neurons are synchronized and the collective firing demonstrates high-amplitude rhythmic oscillations (upper-right insert in Fig. 4). After a few periods of stimulation the oscillatory population is set in a cluster state (bottom-right insert in Fig. 4). Then the stimulation is switched off and the ensemble returns to a synchronized state passing on this way through a uniform desynchronized state (bottom-left insert in Fig. 4). And the procedure is repeated such that the ensemble is kept in a transient desynchronized state. The relaxation to a cluster state is due to the system being attracted by the center manifold as characterized by Eq. (2). By imposing a cluster state, the stimulation does only half of the desynchronizing work. The rest, namely approach-

ing a uniformly desynchronized state, is done by the system itself. In this way the coupling, which causes the synchronization, is used for improving the desynchronizing effect. In the course of the post-stimulus transient  $R_1$  and according to (2) also  $R_2, R_3, \dots$  recover again. The system finally reaches its stable in-phase synchronized state again. In summary, by shifting the system into an unstable cluster state, the system reacts by automatically running through a desynchronized state. Finally, the system reverts back to the synchronized state, if left unperturbed.

The effectively desynchronizing multisite coordinated reset can be used to block the resynchronization. For this, the repetitive stimulus administration can be organized either regardless of the state of the stimulated ensemble (open-loop control) or in a demand-controlled way (closed-loop control), where the following three different control strategies can be utilized:

- (i) *Periodic administration of coordinated reset stimuli:* The most simple, open-loop type of stimulation is a periodic administration of coordinated reset stimuli.
- (ii) *Demand-controlled timing of the administration of identical stimuli:* Whenever the population tends to resynchronize, the same stimulus is administered



**Brain Pacemaker, Figure 5**

Time course of  $R_1$  from (1) (a, c) and of  $n_{\text{fire}}$  from (3) (b, d) during different types of stimulation. *Demand-controlled timing of stimulus administration (a, b)*: As soon as the amplitude  $R_1$  of the recovering order parameter reaches the value of 0.5, the stimulus is administered again. *Periodical stimulation with demand-controlled length of HF pulse train (c, d)*: The stimulus is administered periodically, where the length of the HF pulse trains is adapted to  $R_1$  according to (4) with  $M_{\text{max}} = 15$  and  $M_{\text{min}} = 0$

(Fig. 5). The stronger the synchronization among the neurons is, the more often a stimulus has to be administered to maintain an uncorrelated firing. In an experimental application one has to observe the synchronized oscillation during a sufficiently long period of time in order to perform a frequency analysis which yields the period  $T$  of the population in the absence of stimulation and, thus, the critical stimulation parameter  $\tau$  (the time delay between the two pairs of HF pulse trains, see Fig. 3).

- (iii) *Periodically administered HF pulse trains of demand-controlled length*: The stimuli are periodically administered with offset times  $t_k = k\nu T$ , where  $k = 0, 1, 2, 3, \dots$  is the index labeling the different stimuli,  $T = \tilde{T} + \varepsilon$  is a time interval in the range of the period  $\tilde{T}$  of the population without stimulation, and  $\nu$  is a small integer such as 2 or 3. This means that a  $1:\nu$  entrainment of the four sub-populations is performed, where the spontaneous frequency of the neurons is approximately  $\nu$  times larger compared to the frequency of stimulus administration. The smaller  $|\varepsilon|$ , the smaller is the stimulation strength necessary to achieve an entrainment.

The closed-loop variants (ii) and (iii) require that the ensembles activity can be measured appropriately. Here, either the start times of identical coordinated reset stimuli

or the length of periodically administered stimuli are calculated from the values of  $R_1$ . For example, in the case (ii) the stimulation is started if  $R_1$  becomes larger than a certain threshold (Fig. 5, upper two plots), whereas in the case (iii) the stimulation period is longer for larger values of  $R_1$  measured at the onset of the stimulation (Fig. 5, bottom two plots). In the latter case the length of the HF pulse trains increases linearly between a minimal value  $M_{\text{min}}$  and a maximal value  $M_{\text{max}}$  of single-pulses (except for rounding), where the latter is initially used for desynchronizing the fully synchronized population.  $R_1$  is measured at times  $t'_k = t_k - t_{\text{max}}$ , where  $t_{\text{max}}$  is the maximal duration of a HF pulse train (containing  $M_{\text{max}}$  single-pulses).  $R_1(t'_k)$  determines the number of pulses of the HF pulse trains administered via each of the stimulation sites of the  $k$ th stimulus according to

$$M_k = \min \left\{ \left\lceil \frac{R_1(t'_k)(M_{\text{max}} - M_{\text{min}})}{R_1(t_0)} \right\rceil + M_{\text{min}}, M_{\text{max}} \right\}, \quad (4)$$

where  $k = 0, 1, 2, 3, \dots$ ,  $[x]_{\mathbb{Z}}$  stands for rounding  $x$  to the nearest integer, and  $\min\{x_1, x_2\}$  stands for the minimum of  $\{x_1, x_2\}$ . The  $k$ th stimulus ends precisely at time  $t_k = k\nu T$ , whereas it starts somewhere between  $t'_k$  (for  $M_k = M_{\text{max}}$ ) and  $t_k$  (for  $M_k = M_{\text{min}} = 0$ ), depending on its duration. If the suppression of  $R_1$  is not sufficient one may (i) choose a larger intensity of stimulation, (ii) increase  $M_{\text{min}}$ , (iii) administer the stimuli at a higher rate, i.e. decrease  $\nu$ , so that the inter-stimulus interval  $t_{k+1} - t_k = \nu T$  gets smaller, and/or (iv) increase the duration of each single-pulse of the pulse trains. The feedback value of  $R_1$  can also be evaluated before time  $t'_k$ , especially in case of a slow order parameter dynamics (i.e. when the synchronization is weak with respect to the noise). One could also use the mean of  $R_1$  in a period of evaluation.

Applying the standard, permanent HF pulse train stimulation [5,10] (in a first approximation) corresponds to stimulating each neuron with the same HF pulse train. During a permanent HF stimulation a high-frequency entrainment of the order parameter  $Z_1$  captures  $Z_1$  in a small portion of the complex plane [90], so that the individual neurons' firing is stopped, but no desynchronization occurs. In contrast, during stimulation  $R_1$  can be even larger compared to its pre-stimulus level, and after stimulation the synchronous firing continues immediately. To suppress the firing with such a simple pulse train persistently, it has to be administered permanently. The number of single-pulses used to suppress the firing in the case of the standard permanent HF pulse train stimulation is

about 5 and 8 times larger than that used for blocking the resynchronization in Figs. 5a,b and 5c,d, respectively. This illustrates the effectiveness of the demand-controlled multisite coordinated reset stimulation. The latter can effectively desynchronize stimulated oscillators with a significantly smaller amount of stimulation current compared to the standard permanent HF pulse-train stimulation. Coordinated reset stimulation has been experimentally verified in a hybrid neuroelectronic system of paddlefish electroreceptors [57].

### Linear Multisite Delayed Feedback

Similarly as in the case of the coordinated reset stimulation, the linear multisite delayed feedback [37,38,39,40] is administered via several stimulation sites, e.g. via four sites as illustrated in Fig. 6 (left plot). Individual stimulation signals  $S_m(t)$  administered via each of the stimulation sites are derived from the delayed mean field of the stimulated ensemble using different time delays for different stimulation signals. The mean field characterizes the collective macroscopic dynamics of the oscillators and can be viewed as the ensemble average of the signals of individual oscillators  $Z(t) = N^{-1} \sum_{j=1}^N z_j(t)$ , where  $z_j(t)$ ,  $j = 1, \dots, N$  are the signals of the individual elements of the ensemble.

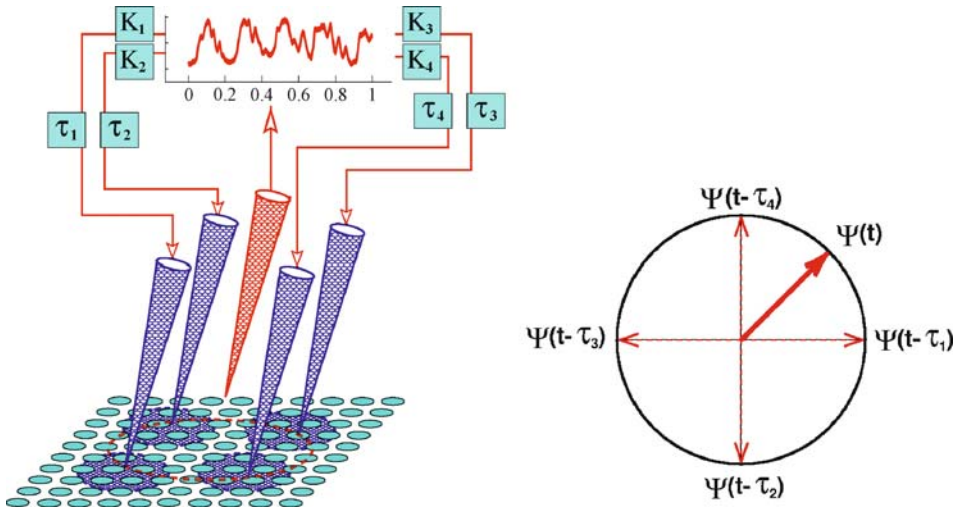
For  $n$  stimulation sites, the stimulation signals are calculated as  $S_m(t) = KZ(t - \tau_m)$ ,  $m = 1, \dots, n$ , where the

values of delay  $\tau_m$ , for example, for  $n = 4$  are calculated from the following relation

$$\tau_m = \frac{11 - 2(m-1)}{8} \tau, \quad m = 1, 2, 3, 4. \quad (5)$$

The delays  $\tau_m$  are symmetrically distributed with respect to the main delay  $\tau$ , where the smallest time delay between neighboring stimulation sites is chosen as  $\tau/4$ . In the case  $\tau = T$  (mean period of the ensemble), the delays  $\tau_m$  are uniformly distributed over the mean period  $T$ . In another realization, instead of four delays  $\tau_m$ ,  $m = 1, \dots, 4$  one can use only two of them, e.g.,  $\tau_1$  and  $\tau_2$ . One can put  $\tau_3 = \tau_1$  and  $\tau_4 = \tau_2$ , where the polarity of the stimulation signals  $S_3(t)$  and  $S_4(t)$  is reversed:  $S_3(t) = -S_1(t)$  and  $S_4(t) = -S_2(t)$ . Assuming that the mean field of the ensemble uniformly oscillates with period  $T = \tau$ , the alternating polarity of the signal corresponds to a shift in time by half a period. Therefore, under this condition the stimulation signal  $S_3(t) = -S_1(t) = -KZ(t - \tau_1)$  approximates the stimulation signal  $S_1(t + \tau/2)$  which is shifted in time by half of the period, which, in turn, is equal to  $KZ(t - \tau_3)$ , where  $\tau_3$  is calculated according to Eq. (5). Analogous arguments are applicable to the stimulation signal  $S_4(t) = -S_2(t) = -KZ(t - \tau_2)$ .

If the phase  $\Psi(t)$  of the mean field  $Z(t)$  uniformly rotates with a constant frequency  $\Omega = 2\pi/\tau$ , the phases  $\Psi_m(t) = \Psi(t - \tau_m)$  of the stimulation signals  $S_m(t)$  are distributed uniformly over the unit circle as illustrated in



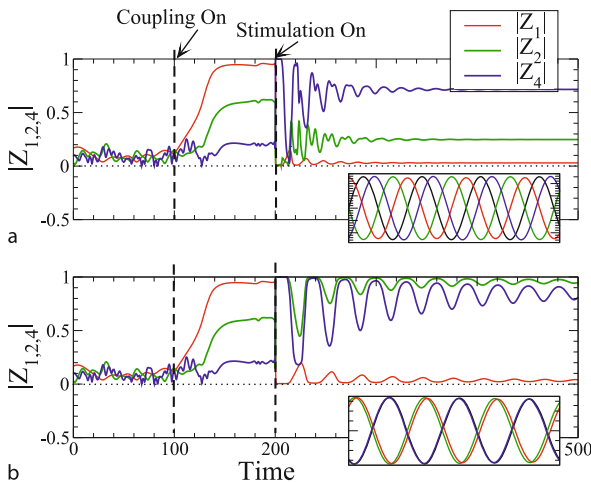
Brain Pacemaker, Figure 6

(Left plot) The macroscopic activity (mean field) of the controlled neuronal population is measured, delayed, amplified and fed back in a spatially coordinated way via several stimulation sites using different delays for different stimulation sites. (Right plot) Distribution of the delayed phases  $\Psi(t - \tau_m)$  of the stimulation signals  $S_m(t)$  of the multisite delayed feedback for four stimulation sites and for  $\tau = T$  in Eq. (5)



Fig. 6 (right plot). Then the phases  $\psi_j(t)$  of the stimulated neuronal sub-population assigned to the stimulation site  $m$  are attracted to the phase  $\Psi(t - \tau_m)$  of the corresponding stimulation signal. Hence, the phases of all oscillators stimulated with the multisite delayed feedback become redistributed symmetrically on the circle  $(0, 2\pi)$  in a cluster state. The order parameter  $R_1(t)$  gets thus minimized. Depending on the value of delay  $\tau$ , the stimulation can induce different cluster states in the stimulated ensemble, where the corresponding order parameter  $R_m$  attains large values.

As shown in Fig. 7, the in-phase synchronization in the stimulated ensemble is effectively suppressed (for time  $t > 200$ , where both coupling and stimulation are switched on), where the order parameter  $R_1(t) = |Z_1(t)|$  attains small values (Fig. 7, red curve). This indicates a symmetrical redistribution of the phases  $\psi_j(t)$  over the unit circle. For the parameter  $\tau$  close to the mean period  $T$  of the stimulation-free ensemble a four cluster state is induced by the stimulation, where the amplitudes of the cluster variables  $|Z_1|$  and  $|Z_2|$  are small, whereas  $|Z_4|$  is relatively large (Fig. 7a). In the subplot, where four trajectories from each of the stimulated sub-populations are depicted, the emerging four-cluster state induced by linear multi-site delayed feedback is illustrated. For  $\tau$  closer to, for example,  $2T$  the stimulation induces a two-cluster state, where  $|Z_1|$  is small, whereas  $|Z_2|$  and  $|Z_4|$  are large



**Brain Pacemaker, Figure 7**

**Control of synchronization by multisite delayed feedback stimulation.** Time courses of the the amplitudes of the cluster variables  $|Z_1|$ ,  $|Z_2|$ , and  $|Z_4|$  are depicted by red, green, and blue curves, respectively. In the subplots four trajectories from each of four stimulated sub-populations assigned to each of four different stimulation sites are shown for  $t \in (320, 340)$ . Parameter  $\tau = T$  in a and  $\tau = 2T$  in b

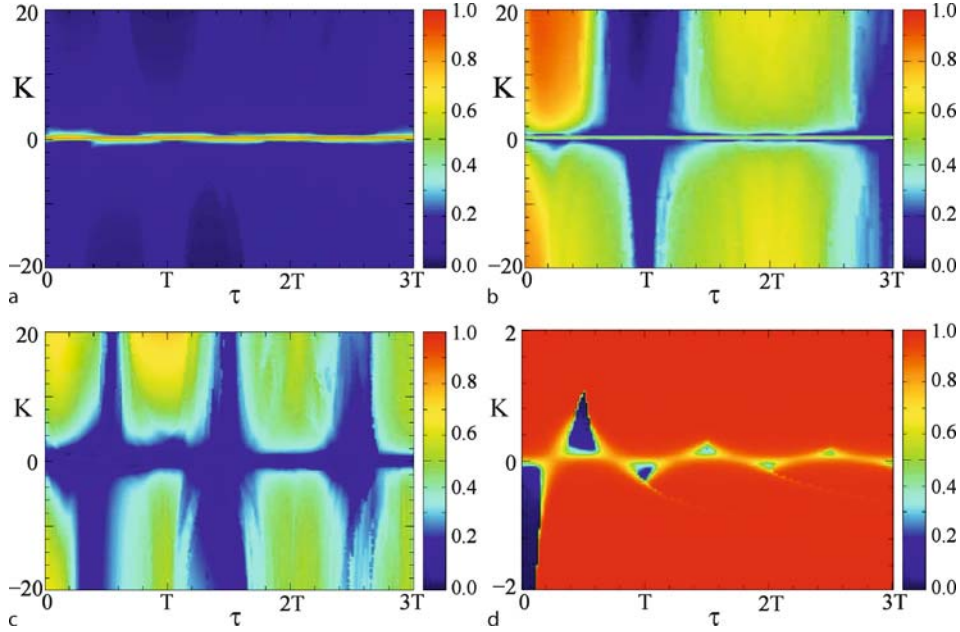
(Fig. 7b). The oscillators thus split into two clusters, which is also illustrated in the subplot in (Fig. 7b).

Linear multi-site delayed feedback robustly suppresses the in-phase synchronization as shown in Fig. 8a, where the order parameter  $R_1$  attains small values for a broad range of parameters  $\tau$  and  $K$ . On the other hand, depending on the values of parameters, either the second order parameter  $R_2$  (Fig. 8b) attains relatively large values (e. g., for  $\tau \approx 2T$ ) indicating a two-cluster state, or, if the values of  $R_2$  become small (e. g., for  $\tau \approx T$ ), the relatively large values of the fourth order parameter  $R_4$  (Fig. 8c) indicate a four-cluster state. Therefore, the whole stimulated population splits into two or four distinct sub-populations. Within the phase clusters the individual oscillators have phases close to each other, while the different phase clusters are equidistantly distributed within the cycle. Hence, depending on the values of the parameters  $\tau$  and  $K$ , linear multi-site delayed feedback with four stimulation sites may cause either a two-cluster state, where  $R_1$  is close to zero and  $R_2$  is maximal, or a four-cluster state, where both  $R_1$  and  $R_2$  are close to zero, but  $R_4$  is maximal. The cluster states become less pronounced and the phases redistribute on the circle even more uniformly if a local coupling as well as spatially decaying profile of the current spread is taken into account [37,40].

In Fig. 8d a similar two-parameter diagram for the averaged order parameter  $\langle R_1(t) \rangle$  is presented for *linear single-site delayed feedback*, suggested for synchronization control in Refs. [67,68]. The stimulation is performed via one stimulation electrode in such a way that all oscillators of the ensemble (in a first approximation) receive the same stimulation signal  $S(t)$ . In this case the stimulation signal  $S(t)$  attains the form  $S(t) = KZ(t - \tau)$ . For the stimulation with linear single-site delayed feedback in a corresponding two-parameter diagram (Fig. 8d) islands of perfect desynchronization are complemented by areas of stimulation-enhanced synchronization. In the limit  $N \rightarrow \infty$  the order parameter  $R_1 = 0$  in the desynchronization regions, where the phases are uniformly distributed on the circle  $(0, 2\pi)$  [67,68]. This is the state of complete desynchronization, where the stimulated oscillators rotate with different frequencies indicating an absence of any cluster state whatsoever.

The important property of the stimulation with the linear multi- and single-site delayed feedback is the inherited demand-controlled character of the methods. As soon as the desired desynchronized state is achieved, the value of the order parameter  $R_1(t)$ , i. e., the amplitude of the mean field becomes small. Along with the order parameter, in the desynchronized state the amplitude of the stimulation signal  $S(t)$  vanishes as well. The stimulation with





**Brain Pacemaker, Figure 8**

Impact of the multisite delayed feedback stimulation versus parameters  $\tau$  and stimulus amplification  $K$ . The time-averaged order parameters  $\langle R_1 \rangle$ ,  $\langle R_2 \rangle$ , and  $\langle R_4 \rangle$  are depicted in a, b, and c, respectively, and are encoded in color ranging from 0 (blue) to 1 (red). In plot d the impact of the linear single-site delayed feedback on the oscillatory population is illustrated, where the values of the order parameter  $\langle R_1 \rangle$  are depicted in color versus parameters  $\tau$  and  $K$

linear multi- and single-site delayed feedback thus represents a noninvasive control method for the desynchronization of coupled oscillators. The stimulated ensemble is then subjected to a highly effective control at a minimal amount of stimulation force.

### Nonlinear Delayed Feedback

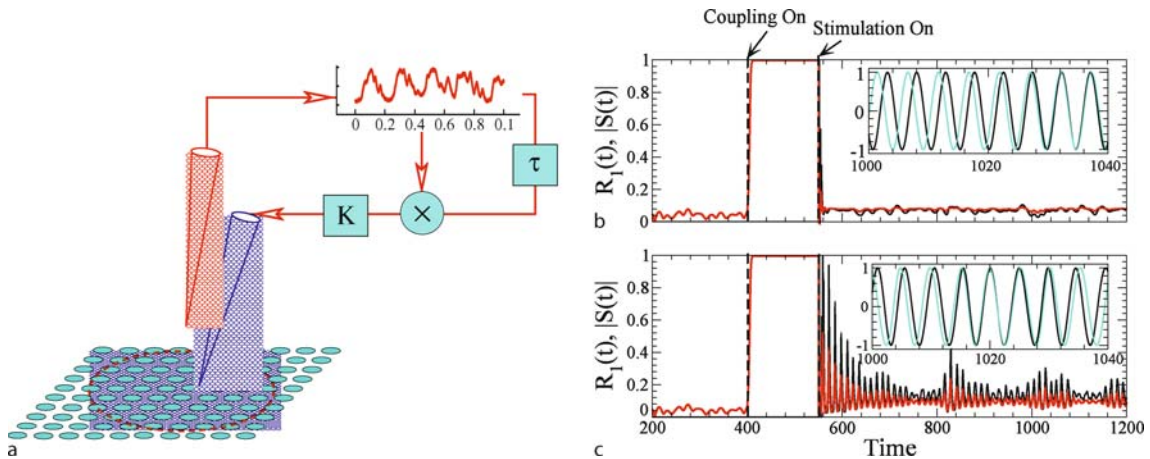
As for the case of the linear single-site delayed feedback, for the stimulation with nonlinear delayed feedback only one registering and one stimulating site is required, see Fig. 9a. All stimulated oscillators receive the same stimulation signal  $S(t)$  which is constructed from the measured mean field of the ensemble. It is assumed that the measured mean field  $Z(t)$  of the ensemble has the form of a complex analytic signal  $Z(t) = X(t) + iY(t)$ , where  $X(t)$  and  $Y(t)$  are the real and imaginary parts of  $Z(t)$ , respectively. In the case if only the real part  $X(t)$  of the mean field is measured, the imaginary part can be calculated, e.g., with the help of Hilbert transform [60]. The stimulation signal is then constructed by a nonlinear combination of a delayed complex conjugate mean field with the instantaneous mean field [62,63,64,98]

$$S(t) = KZ^2(t)Z(t - \tau), \quad (6)$$

where  $K$  is a stimulus amplification parameter,  $\tau$  is a time delay, and the asterisk denotes complex conjugacy.

The desynchronizing effect of the stimulation with nonlinear delayed feedback is illustrated in Fig. 9b,c, where the time courses of the order parameter  $R_1(t)$  (red curves) and the amplitude of the stimulation signal  $|S(t)|$  (black curves) are plotted for two values of delay  $\tau = T/2$  (Fig. 9b) and  $\tau = T$  (Fig. 9c), where  $T$  is the mean period of the stimulation-free population. The onset of stimulation at  $t = 550$  results in desynchronization of the stimulated oscillators and the order parameter reaches the values of approximately the same order of magnitude as in the uncoupled regime ( $t < 400$ ). This indicates a high level of desynchronization. The stimulation does not destroy the normal oscillatory activity of the individual phase oscillators. In the insets in Figs. 9b and 9c individual trajectories of two selected oscillators of the stimulated ensemble are plotted. The stimulated oscillators rotate with different individual frequencies just as in the coupling- and stimulation-free regime.

As soon as a desynchronized state is achieved, the stimulation force declines and the stimulated system is subjected to a highly effective control with a minimal amount of stimulation force. Also, as soon as a resynchronization occurs, the mean field starts to exhibit large-am-



Brain Pacemaker, Figure 9

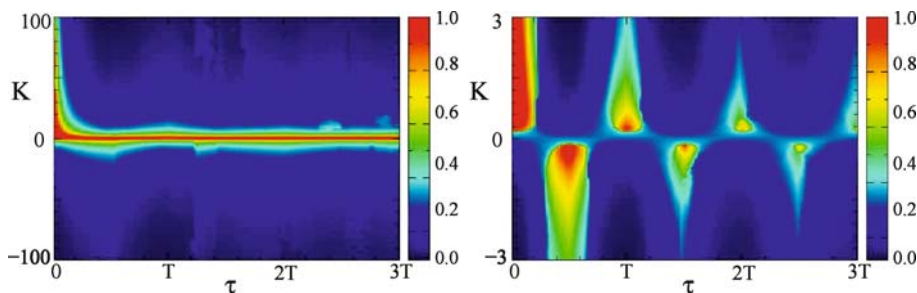
**a** The macroscopic activity (mean field) of the controlled population is measured, delayed, nonlinearly combined with the instantaneous mean field, amplified, and fed back via a single stimulation site. **b, c** Desynchronization of strongly synchronized oscillators by nonlinear delayed feedback. Time courses of the order parameter  $R_1(t)$  (red curves) and the amplitude of the stimulation signal  $|S(t)|$  (black curves) are plotted for delay **b**  $\tau = T/2$  and **c**  $\tau = T$ . In the subplots two selected trajectories of oscillators are depicted in the stimulated regime.  $T$  is the mean period of the stimulation-free ensemble

plitude oscillations and the stimulation signal increases its amplitude and brings the ensemble back to a desynchronized state. This demand-controlled character of the nonlinear delayed feedback is particularly well illustrated in Fig. 9c, where the onsets of resynchronization (increase of  $R_1(t)$ , red curve) at times around  $t \approx 850, 1050$ , and  $1200$  lead to an increase of the amplitude of the stimulation signal  $|S(t)|$  (black curve), which in turn results in a suppression of the resynchronization.

The impact of the nonlinear delayed feedback on the stimulated oscillators is twofold. On one hand, the stimulation can effectively desynchronize even strongly inter-

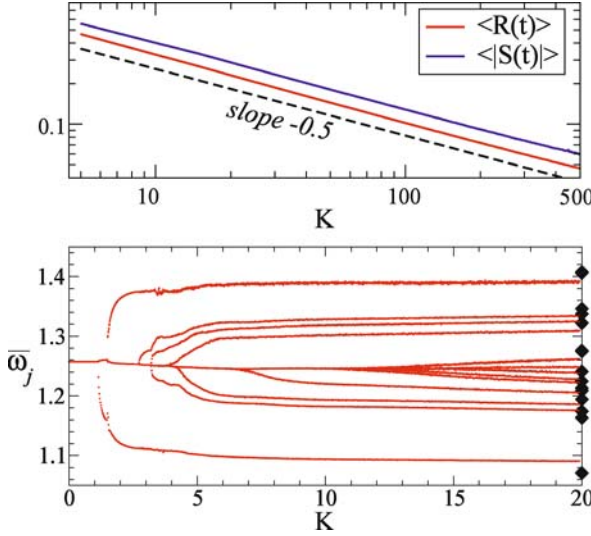
acting oscillators for a range of large values of the stimulus amplification  $K$ , see Fig. 10, left plot. This effect is very robust with respect to the variation of the delay  $\tau$  and, as a result, with respect to the variation of the mean frequency  $\Omega$  of the stimulated ensemble. On the other hand, in a weakly coupled ensemble the stimulation can induce synchronization in island-like regions of small values of the stimulus amplification  $K$  complemented by domain of desynchronization, see Fig. 10, right plot.

An increase of the stimulus amplification parameter  $K$  results in a gradual decay of the order parameter  $R_1$ , which indicates an onset of desynchronization in the stimulated



Brain Pacemaker, Figure 10

Nonlinear delayed feedback: Stimulation-induced desynchronization (left plot) and synchronization (right plot) in ensembles of strongly (left plot) or weakly (right plot) coupled oscillators. Note, nonlinear delayed feedback is a robust desynchronization technique, which enables to effectively counteract intermediate and strong coupling (left plot). However, in ensembles with weak coupling nonlinear delayed feedback with particularly chosen parameters may also enhance synchronization (right plot). The time-averaged values of the order parameter  $\langle R_1 \rangle$  are encoded in color ranging from red (synchronization) to blue (desynchronization) versus delay  $\tau$  and stimulus amplification  $K$



**Brain Pacemaker, Figure 11**

Time-averaged order parameter  $\langle R_1 \rangle$  and amplitude of the stimulation signal  $\langle |S(t)| \rangle$  (upper plot, log-log scale) and the observed individual frequencies  $\bar{\omega}_j$  of the stimulated oscillators (bottom plot) versus the stimulus amplification parameter  $K$ . Black diamonds depict values of the natural frequencies of the stimulated oscillators

ensemble. Simultaneously, the amplitude of the stimulation signal  $|S(t)|$  decays as well, indicating a demand-controlled character of the nonlinear delayed feedback stimulation. For a fixed delay  $\tau > 0$  the order parameter and the amplitude of the stimulation signal decay according to the following power law as  $|K|$  increases:

$$R_1 \sim |K|^{-1/2}, \quad |S| \sim |K|^{-1/2}, \quad (7)$$

see Fig. 11, upper plot. The desynchronization transition for increasing  $K$  also manifests itself in a sequence of frequency-splitting bifurcations, where the observed individual frequencies  $\bar{\omega}_j = \langle \dot{\psi}_j \rangle$  of the stimulated oscillators split, one after another from the mean frequency  $\Omega$  as  $K$  increases and approach the natural frequencies of the unperturbed oscillators (black diamonds in Fig. 11, bottom plot).

For large values of  $K$  all stimulated oscillators rotate with different frequencies which approach the natural frequencies  $\omega_j$  of the oscillators. The oscillators thus exhibit a uniform desynchronized dynamics without any kind of cluster states. Simultaneously, depending on the values of the delay  $\tau$ , the nonlinear delayed feedback can significantly change the mean frequency  $\Omega$ , i.e., the frequency of the mean field  $Z(t)$  of the stimulated ensemble [62,63]. The macroscopic dynamics can thus be either accelerated or slowed down, whereas the individual dynamics remains

close to the original. This opens an approach to frequency control of the oscillatory population stimulated with nonlinear delayed feedback.

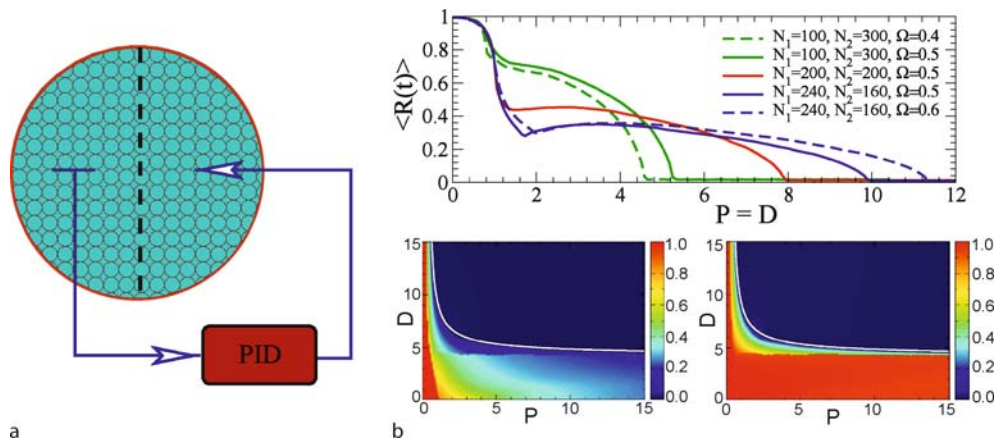
### Proportional-Integro-Differential Feedback

For a particularly difficult situation, where the measurement and stimulation are not possible at the same time and at the same place, there is another control method which is based on a proportional-integro-differential (PID) feedback. The scheme of this stimulation protocol is sketched in Fig. 12a. The controlled ensemble of  $N$  coupled oscillators is split into two separate sub-populations of  $N_1$  and  $N_2 = N - N_1$  oscillators, one being exclusively measured and the other being exclusively stimulated. In this way a separate stimulation-registration setup is realized, where the recording and stimulating sites are spatially separated and the measured signal is not corrupted by stimulation artifacts. The observed signal is considered to be the mean field  $Z^{(1)}$  of the measured sub-population. Below the main attention will be paid to the proportional-differential (PD) feedback only (for more details, see [65]). Then, the stimulation signal  $S(t)$  administered to the second, stimulated sub-population is constructed as

$$S(t) = PZ^{(1)}(t) + D\dot{Z}^{(1)}(t), \quad (8)$$

where the parameters  $P$  and  $D$  define the strengths of the proportional and differential feedback, respectively.

The effect of the stimulation with PD feedback is illustrated in Fig. 12b. As the strength of the feedback (parameters  $P$  and  $D$ ) increases the stimulation results in a complete desynchronization of both, measured and stimulated sub-populations. The threshold of the onset of desynchronization depends on the relative splitting  $N_1 : N_2$  of the oscillators between sub-populations and on the mean frequency  $\Omega$ : The threshold is larger for smaller number of oscillators  $N_2$  in the stimulated populations or for larger frequency  $\Omega$ . The later dependence can be eliminated if an integral component is included in the stimulation signal, see [65]. Moreover, if the coupling in the ensemble is rather weak, the desynchronization can be achieved by applying the proportional feedback only. In contrast, in the case of strong coupling the stimulation signal must also contain the differential feedback for robust desynchronization. As illustrated in the 2-parameter diagrams in Fig. 12b (bottom plots), there exists a certain threshold in parameters  $P$  and  $D$  values, where the stimulation with PID feedback desynchronizes both sub-populations in the target ensemble: Stimulated sub-population (Fig. 12b, bottom right plot) and also measured, non-stimulated sub-population (Fig. 12b, bottom left plot).



Brain Pacemaker, Figure 12

**PID Control:** a The mean field is measured in one part of the controlled ensemble and, after processing according to proportional-integro-differential (PID) algorithm, is administered to the other part of the ensemble. **b Upper plot:** The order parameter of the whole population versus the strength of the PD feedback (with  $P = D$ ) for different splitting  $N_1 : N_2$  and different mean frequencies  $\Omega$ . **Bottom plots:** The values of the order parameter (encoded in color) of the measured sub-population (left) and stimulated sub-population (right) versus stimulation parameters  $P$  and  $D$ . The white curve is the parameter threshold of the onset of desynchronization in sub-populations obtained in [65]

In this sense the PID feedback stimulation method appears to be very effective even in such a complicated stimulation protocol of separate stimulation-registration setup.

### Plasticity

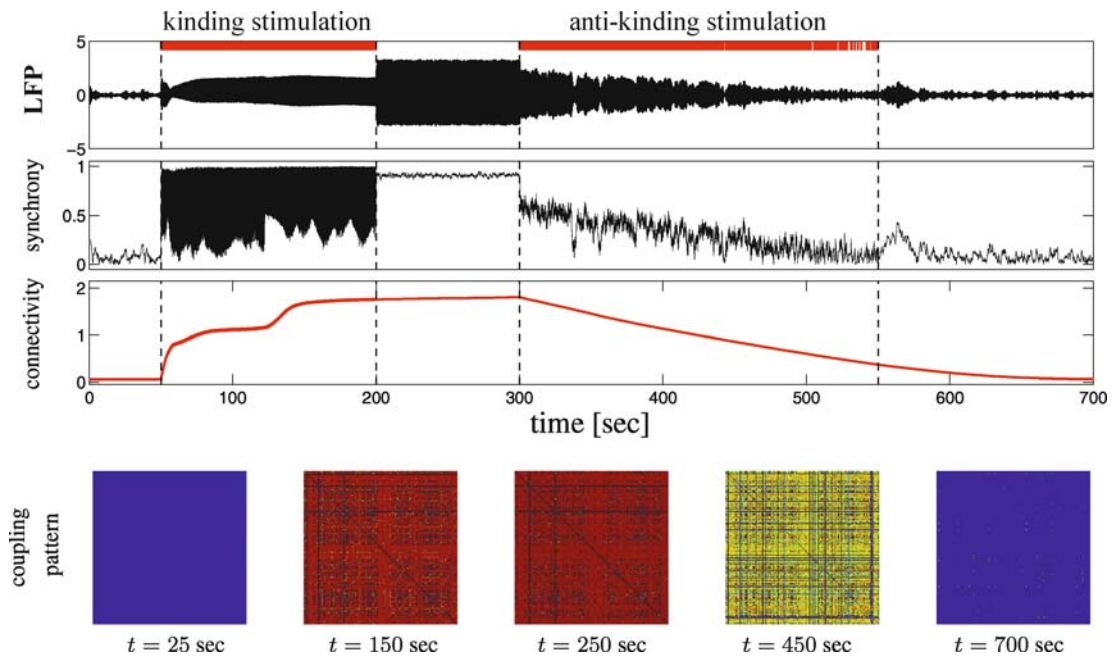
Plasticity is a fundamental property of the nervous system: In order to learn and to adapt to sensory inputs, neurons continuously adapt the strength of their synaptic connections in relation to the mutual timing properties of their firing or bursting [1,13,18,22,42,44,51,78,100,105]. However, plasticity may not only lead to desired learning and optimization processes. Rather neuronal populations can learn pathologically strong interactions which may lead, e.g., to the emergence of epilepsies [56,79]. This is well-known from the so-called kindling phenomenon [24], where preparatory stimulation induces the spontaneous production of epileptic seizures without gross morphological changes [56].

Recently, theoretical studies have initiated an approach which is targeted on unlearning pathologically strong synaptic interactions by desynchronizing brain stimulation [36,49,83,97]. This approach exploits plasticity in two different ways: On the one hand, due to plasticity desynchronizing stimulation may decrease the strength of the neurons' synapses by decreasing the rate of coincidences. On the other hand, neuronal networks with synaptic plasticity may exhibit bi- or multistability [36,49,73,83,97]. Accordingly, by decreasing the mean

synaptic weight, desynchronizing stimulation may shift a neuronal population from a stable synchronized (pathological) state to a stable desynchronized (healthy) state, where the neuronal population remains thereafter, if left unperturbed. In Fig. 13 an exemplary simulation of a network of 200 model neurons is displayed, for further details concerning the mathematical model we refer to [36,83]. Induced by appropriate stimulation protocols a switching between the different stable states is realizable. Starting from a desynchronized state, associated with physiological system dynamics, low-frequency stimulation can induce a kindling of the interneuronal connectivities and causes a stabilization of the synchronized state. After stimulation offset, the system remains in the pathological state, see Fig. 13. In contrast, desynchronizing (MCR) stimulation results in an anti-kindling of the pathological connectivities and finally the physiological state is reestablished, see Fig. 13.

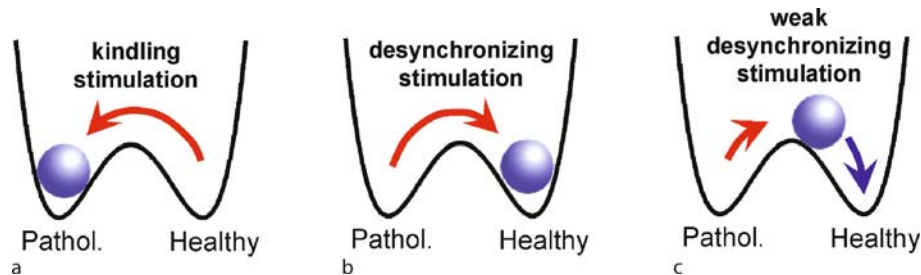
From a mathematical point of view, in a first approximation this situation may be illustrated by considering a double well potential, where each minimum corresponds to a stable attractor, surrounded by a basin of attraction (Fig. 14). The strongly synchronized state (*Pathol.* in Fig. 14) serves as a model for a disease state, whereas the uncorrelated or weakly synchronized state (*Healthy* in Fig. 14) is used as a model for a healthy state. As soon as the system, i.e. the neuronal population (illustrated by the ball in Fig. 14), enters a particular basin of attraction, it gets attracted by the corresponding attractor, so that it relaxes towards the corresponding minimum of the potential.





Brain Pacemaker, Figure 13

Effects of kindling and anti-kindling stimulation on a population of model neurons. Low-frequency stimulation is applied between 50 to 200 seconds and multi-site coordinated reset stimulation is applied between 300 to 550 seconds. The local field potential, the level of synchronization within the network and the mean connectivity is plotted (from top to bottom). Below five patterns representing the coupling topology of the network at different times are plotted. Blue (red) colors represent low (high) values of the interneuronal connectivity



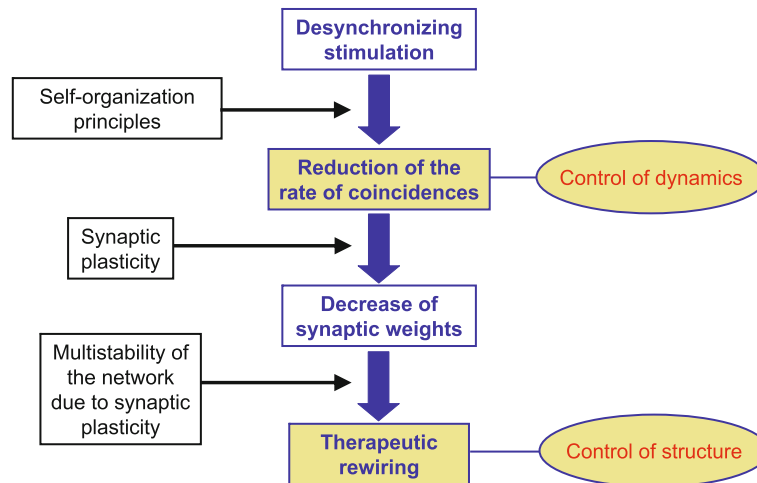
Brain Pacemaker, Figure 14

Anti-kindling and kindling stimulation can move the neuronal population from one attractor to another: Schematic plot of the attractors symbolizing pathological (*Pathol.*) or healthy (*Healthy*) dynamical model states. a Periodic, kindling stimulation (red trajectory) shifts the population from a healthy, desynchronized state (*Healthy*) to a pathological, synchronized state (*Pathol.*). b Conversely, desynchronizing stimulation shifts the population from a pathological state (*Pathol.*) to the healthy uncorrelated state (*Healthy*). c This anti-kindling is achieved by a desynchronizing stimulation of sufficient duration, so that after stimulus offset the population is close to the healthy state (red trajectory). d Alternatively, the same long-term anti-kindling effect can be achieved with a brief desynchronizing stimulation, which shifts the population to an intermediate state, which may still be connected with pronounced synchrony. However, since the intermediate state (blue ball) lies within the basin of attraction of a healthy state (*Healthy*) the population spontaneously relaxes to the healthy state without any further intervention (blue trajectory)

Appropriate stimulation protocols may shift the neuronal population from one state to another. Kindling stimulation of appropriate duration shifts the neuronal population from a desynchronized state close to a strongly synchronized state or at least into the basin of attraction of such a state (Fig. 14a, red trajectory from *Healthy*

to *Pathol.*). Conversely, anti-kindling can be achieved by means of a desynchronizing stimulation which shifts the neuronal population close to the desynchronized state [36,83,97] (Fig. 14c, red trajectory from *Pathol.* to *Healthy*). However, with respect to the long-term anti-kindling outcome, it is even sufficient to move the neuronal





**Brain Pacemaker, Figure 15**

**Schematic illustration of how desynchronizing stimulation may cause an unlearning of pathological connectivity:** Desynchronizing stimulation reduces the overall rate of coincidences in the neuronal population. This is effectively achieved by using dynamic self-organization principles. The reduction of the rate of coincidences, in turn, reduces the synaptic weights. Because of the multistability not only strongly synchronized states, but also uncorrelated and weakly synchronized states are stable. Accordingly, the neuronal population stably remains in a desynchronized or weakly synchronized state, after having been shifted into the basin of attraction of that state by means of desynchronizing stimulation. Hence, a suitable control of the dynamics of the network leads to a modulation of its connectivity

population from the synchronized state just into the basin of attraction of the desynchronized state (Fig. 14b, red trajectory from *Pathol.* to the intermediate state). After stimulus offset there may still be pronounced synchrony, but being captured within the basin of attraction of the desynchronized state, the neuronal population spontaneously relaxes to the desynchronized state, without further intervention (Fig. 14c, blue trajectory from the intermediate state to *Healthy*). It has to be considered that the short as well as the long desynchronizing stimulation in this schematic illustration have the same long-term anti-kindling outcome.

In PD neuronal populations of the basal ganglia are strongly synchronized [9,70,99] and synaptic plasticity results in a further amplification of the synchronized activity by a strengthening of the synaptic connections [59]. Properly designed electrical stimulation may be used to break this vicious circle and to induce an anti-kindling [36,83,97], which finally might reestablish the normal level of connectivity, associated with a mostly uncorrelated neuronal activity (Fig. 15).

## Summary

High-frequency deep brain stimulation is the golden standard for the treatment of medically refractory movement

disorders [5,101]. To overcome limitations of this type of therapy, a model-based development of novel deep brain stimulation techniques has been initiated [86]. The ingredients of this approach are dynamic neuronal self-organization principles combined with fundamental plasticity rules of the nervous system (Fig. 15) [36,83,96,97]. In forthcoming studies the mathematical modeling will be refined to incorporate further anatomical and physiological details, for instance, contributions of glial cells [75]. By the same token, control techniques will be optimized and further developed. As currently done for multisite coordinated reset stimulation, clinical studies will be performed to evaluate the therapeutic effects of the novel stimulation techniques. This interdisciplinary endeavor is performed to finally provide superior therapies for patients with neurological or psychiatric diseases.

## Bibliography

1. Abbott L, Nelson S (2000) Synaptic plasticity: Taming the beast. *Nat Neurosci* 3:1178–1183
2. Alberts WW, Wright EJ, Feinstein B (1969) Cortical potentials and parkinsonian tremor. *Nature* 221:670–672
3. Andres F, Gerloff C (1999) Coherence of sequential movements and motor learning. *J Clin Neurophysiol* 16(6):520–527
4. Benabid A, Pollak P, Louveau A, Henry S, de Rougemont JJ (1987) Combined (thalamotomy and stimulation) stereotac-

- tic surgery of the VIM thalamic nucleus for bilateral Parkinson disease. *Appl Neurophysiol* 50(1-6):344–346
5. Benabid AL, Pollak P, Gervason C, Hoffmann D, Gao DM, Hommel M, Perret JE, de Rougemont J (1991) Longterm suppression of tremor by chronic stimulation of ventral intermediate thalamic nucleus. *Lancet* 337:403–406
  6. Benabid AL, Benazzous A, Pollak P (2002) Mechanisms of deep brain stimulation. *Mov Disord* 17:73–74
  7. Benabid A-L, Wallace B, Mitrofanis J, Xia R, Piallat B, Chabardes S, Berger F (2005) A putative generalized model of the effects and mechanism of action of high frequency electrical stimulation of the central nervous system. *Acta neurol Belg* 105:149–157
  8. Beurrier C, Bioulac B, Audin J, Hammond C (2001) High-frequency stimulation produces a transient blockade of voltage-gated currents in subthalamic neurons. *J Neurophysiol* 85(4):1351–1356
  9. Beurrier C, Garcia L, Bioulac B, Hammond C (2002) Subthalamic nucleus: A clock inside basal ganglia? *Thalamus Relat Syst* 2:1–8
  10. Blond S, Caparros-Lefebvre D, Parker F, Assaker R, Petit H, Guieu J-D, Christiaens J-L (1992) Control of tremor and involuntary movement disorders by chronic stereotactic stimulation of the ventral intermediate thalamic nucleus. *J Neurosurg* 77:62–68
  11. Brice J, McLellan L (1980) Suppression of intention tremor by contingent deep-brain stimulation. *Lancet* 1(8180):1221–1222
  12. Daido H (1992) Order function and macroscopic mutual entrainment in uniformly coupled limit-cycle oscillators. *Prog Theor Phys* 88:1213–1218
  13. Debanne D, Gahweiler B, Thompson S (1998) Long-term synaptic plasticity between pairs of individual CA3 pyramidal cells in rat hippocampus slice cultures. *J Physiol* 507:237–247
  14. Deuschl G, Schade-Brittinger C, Krack P, Volkmann J, Schäfer H, Bötzel K, Daniels C, Deuschländer A, Dillmann U, Eisner W, Gruber D, Hamel W, Herzog J, Hilker R, Klebe S, Klotz M, Koy J, Krause M, Kupsch A, Lorenz D, Lorenz S, Mehdorn H, Morjane J, Oertel W, Pinsker M, Reichmann H, Reuß A, Schneider G-H, Schnitzler A, Steude U, Sturm V, Timmermann L, Tronnier V, Trottenberg T, Wojtecki L, Wolf E, Poewe W, Voges J (2006) A randomized trial of deep-brain stimulation for parkinsons disease. *N Engl J Med* 355:896–908
  15. Dolan K, Majtanik M, Tass P (2005) Phase resetting and transient desynchronization in networks of globally coupled phase oscillators with inertia. *Physica D* 211:128–138
  16. Elble RJ, Koller WC (1990) Tremor. John Hopkins University Press, Baltimore
  17. Ermentrout B, Kopell N (1991) Multiple pulse interactions and averaging in systems of coupled neural assemblies. *J Math Biol* 29:195–217
  18. Feldman D (2000) Timing-based LTP and LTD at vertical inputs to layer II/III pyramidal cells in rat barrel cortex. *Neuron* 27:45–56
  19. Filali M, Hutchison W, Palter V, Lozano A, Dostrovsky JO (2004) Stimulation-induced inhibition of neuronal firing in human subthalamic nucleus. *Exp Brain Res* 156:274–281
  20. Freund H-J (2005) Long-term effects of deep brain stimulation in parkinsons disease. *Brain* 128:2222–2223
  21. Garcia L, D'Alessandro G, Fernagut P-O, Bioulac B, Hammond C (2005) Impact of high-frequency stimulation parameters on the pattern of discharge of subthalamic neurons. *J Neurophysiol* 94:3662–3669
  22. Gerstner W, Kempter R, van Hemmen J, Wagner H (1996) A neuronal learning rule for sub-millisecond temporal coding. *Nature* 383:76–78
  23. Gildenberg P (2005) Evolution of neuromodulation. *Stereotact Funct Neurosurg* 83:71–79
  24. Goddar G (1967) Development of epileptic seizures through brain stimulation at low intensity. *Nature* 214:1020–1021
  25. Grannan ER, Kleinfeld D, Sompolinsky H (1993) Stimulus-dependent synchronization of neuronal assemblies. *Neural Comp* 5:550–569
  26. Grill WM, McIntyre CC (2001) Extracellular excitation of central neurons: Implications for the mechanisms of deep brain stimulation. *Thalamus Relat Syst* 1:269–277
  27. Haken H (1970) Laser theory vol XXV/2C. In: Flüge S (ed) *Encyclopedia of physics*. Springer, Berlin
  28. Haken H (1977) Synergetics. An introduction. Springer, Berlin
  29. Haken H (1983) Advanced synergetics. Springer, Berlin
  30. Haken H (1996) Principles of brain functioning. A synergetic approach to brain activity, behavior, and cognition. Springer, Berlin
  31. Haken H (2002) Brain dynamics. Synchronization and activity patterns in pulse-coupled neural nets with delays and noise. Springer, Berlin
  32. Haken H, Kelso J, Bunz H (1985) A theoretical model of phase transitions in human hand movements. *Biol Cybern* 51:347–356
  33. Hansel D, Mato G, Meunier C (1993) Phase dynamics of weakly coupled Hodgkin–Huxley neurons. *Europhys Lett* 23:367–372
  34. Hansel D, Mato G, Meunier C (1993) Phase reduction and neuronal modeling. *Concepts Neurosci* 4(2):193–210
  35. Hashimoto T, Elder C, Okun M, Patrick S, Vitek J (2003) Stimulation of the subthalamic nucleus changes the firing pattern of pallidal neurons. *J Neurosci* 23(5):1916–1923
  36. Hauptmann C, Tass PA (2007) Therapeutic rewiring by means of desynchronizing brain stimulation. *Biosystems* 89:173–181
  37. Hauptmann C, Popovych O, Tass PA (2005) Delayed feedback control of synchronization in locally coupled neuronal networks. *Neurocomputing* 65–66:759–767
  38. Hauptmann C, Popovych O, Tass PA (2005) Effectively desynchronizing deep brain stimulation based on a coordinated delayed feedback stimulation via several sites: A computational study. *Biol Cybern* 93:463–470
  39. Hauptmann C, Popovych O, Tass PA (2005) Multisite coordinated delayed feedback for an effective desynchronization of neuronal networks. *Stoch Dyn* 5(2):307–319
  40. Hauptmann C, Omelchenko O, Popovych OV, Maistrenko Y, Tass PA (2007) Control of spatially patterned synchrony with multisite delayed feedback. *Phys Rev E* 76:066209
  41. Hauptmann C, Popovych O, Tass P (2007) Desynchronizing the abnormally synchronized neural activity in the subthalamic nucleus: A modeling study. *Expert Rev Med Devices* 4(5):633–650
  42. Hebb D (1949) The organization of behavior. Wiley, New York
  43. Kelso J (1995) Dynamic patterns: The self-organization of brain and behavior. MIT Press, Cumberland
  44. Kilgard M, Merzenich M (1998) Cortical map reorganization enabled by nucleus basalis activity. *Science* 279:1714–1718
  45. Kumar R, Lozano A, Sime E, Lang A (2003) Long-term follow-

- up of thalamic deep brain stimulation for essential and parkinsonian tremor. *Neurology* 61:1601–1604
46. Kuramoto Y (1984) *Chemical oscillations, waves, and turbulence*. Springer, Berlin
  47. Lenz F, Kwan H, Martin R, Tasker R, Dostrovsky J, Lenz Y (1994) Single unit analysis of the human ventral thalamic nuclear group. Tremor-related activity in functionally identified cells. *Brain* 117:531–543
  48. Limousin P, Speelman J, Gielen F, Janssens M (1999) Multicentre European study of thalamic stimulation in parkinsonian and essential tremor. *J Neurol Neurosurg Psychiatry* 66(3):289–296
  49. Maistrenko Y, Lysyansky B, Hauptmann C, Burylko O, Tass P (2007) Multistability in the Kuramoto model with synaptic plasticity. *Phys Rev E* 75:066207
  50. Majtanik M, Dolan K, Tass P (2006) Desynchronization in networks of globally coupled neurons with dendritic dynamics. *J Biol Phys* 32:307–333
  51. Markram H, Lübke J, Frotscher M, Sakmann B (1997) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* 275:213–215
  52. McIntyre C, Grill W, Sherman D, Thakor N (2004) Cellular effects of deep brain stimulation: Model-based analysis of activation and inhibition. *J Neurophysiol* 91:1457–1469
  53. McIntyre CC, Savasta M, Goff KKL, Vitek J (2004) Uncovering the mechanism(s) of action of deep brain stimulation: Activation, inhibition, or both. *Clin Neurophysiol* 115:1239–1248
  54. Milton J, Jung P (eds) (2003) *Epilepsy as a dynamics disease*. Springer, Berlin
  55. Miocinovic S, Parent M, Butson C, Hahn P, Russo G, Vitek J, McIntyre C (2006) Computational analysis of subthalamic nucleus and lenticular fasciculus activation during therapeutic deep brain stimulation. *J Neurophysiol* 96:1569–1580
  56. Morimoto K, Fahnstock M, Racine R (2004) Kindling and status epilepticus models of epilepsy: Rewiring the brain. *Prog Neurobiol* 73:1–60
  57. Neiman A, Russell D, Yakusheva T, DiLullo A, Tass PA (2007) Response clustering in transient stochastic synchronization and desynchronization of coupled neuronal bursters. *Phys Rev E* 76:021908
  58. Nini A, Feingold A, Sloviter H, Bergmann H (1995) Neurons in the globus pallidus do not show correlated activity in the normal monkey, but phase-locked oscillations appear in the MPTP model of parkinsonism. *J Neurophysiol* 74:1800–1805
  59. Nowotny T, Zhigulin V, Selverston A, Abarbanel H, Rabinovich M (2003) Enhancement of synchronization in a hybrid neural circuit by spike-timing dependent plasticity. *J Neurosci* 23:9776–9785
  60. Pikovsky A, Rosenblum M, Kurths J (2001) *Synchronization, a universal concept in nonlinear sciences*. Cambridge University Press, Cambridge
  61. Pliss V (1964) Principal reduction in the theory of stability of motion. *Izv Akad Nauk SSSR Math Ser* 28:1297–1324
  62. Popovych OV, Hauptmann C, Tass PA (2005) Effective desynchronization by nonlinear delayed feedback. *Phys Rev Lett* 94:164102
  63. Popovych OV, Hauptmann C, Tass PA (2006) Control of neuronal synchrony by nonlinear delayed feedback. *Biol Cybern* 95:69–85
  64. Popovych OV, Hauptmann C, Tass PA (2006) Desynchronization and decoupling of interacting oscillators by nonlinear delayed feedback. *Int J Bif Chaos* 16(7):1977–1987
  65. Pyragas K, Popovych OV, Tass PA (2007) Controlling synchrony in oscillatory networks with a separate stimulation-registration setup. *Europhys Lett* 80:40002
  66. Rodriguez-Oroz M, Obeso J, Lang A, Houeto J, Pollak P, Rehn-crona S, Kulisevsky J, Albanese A, Volkmann J, Hariz M, Quinn N, Speelman J, Guridi J, Zamarbide I, Gironell A, Molet J, Pascual-Sedano B, Pidoux B, Bonnet A, Agid Y, Xie J, Benabid A, Lozano A, Saint-Cyr J, Romito L, Contarino M, Scerrati M, Fraix V, Blercom NV (2005) Bilateral deep brain stimulation in Parkinsons disease: A multicentre study with 4 years follow-up. *Brain* 128:2240–2249
  67. Rosenblum MG, Pikovsky AS (2004) Controlling synchronization in an ensemble of globally coupled oscillators. *Phys Rev Lett* 92:114102
  68. Rosenblum MG, Pikovsky AS (2004) Delayed feedback control of collective synchrony: An approach to suppression of pathological brain rhythms. *Phys Rev E* 70:041904
  69. Rubin J, Terman D (2004) High frequency stimulation of the subthalamic nucleus eliminates pathological thalamic rhythmicity in a computational model. *J Comput Neurosci* 16:211–235
  70. Schnitzler A, Timmermann L, Gross J (2006) Physiological and pathological oscillatory networks in the human motor system. *J Physiol Paris* 99(1):3–7
  71. Schuurman PR, Bosch DA, Bossuyt PM, Bonsel GJ, van Someren EJ, de Bie RM, Merkus MP, Speelman JD (2000) A comparison of continuous thalamic stimulation and thalamotomy for suppression of severe tremor. *N Engl J Med* 342:461–468
  72. Schöner G, Haken H, Kelso J (1986) A stochastic theory of phase transitions in human hand movement. *Biol Cybern* 53:247–257
  73. Seliger P, Young S, Tsimring L (2002) Plasticity and learning in a network of coupled phase oscillators. *Phys Rev E* 65:041906
  74. Shen K, Zhu Z, Munhall A, Johnson SW (2003) Synaptic plasticity in rat subthalamic nucleus induced by high-frequency stimulation. *Synapse* 50:314–319
  75. Silchenko A, Tass P (2008) Computational modeling of paroxysmal depolarization shifts in neurons induced by the glutamate release from astrocytes. *Biol Cybern* 98:61–74
  76. Singer W (1989) Search for coherence: A basic principle of cortical self-organization. *Concepts Neurosci* 1:1–26
  77. Smirnov D, Barnikol U, Barnikol T, Bezruchko B, Hauptmann C, Bührle C, Maarouf M, Sturm V, Freund H-J, Tass P (2008) The generation of parkinsonian tremor as revealed by directional coupling analysis. *Europhys Lett* 83:20003
  78. Song S, Miller K, Abbott L (2000) Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nat Neurosci* 3(9):919–926
  79. Speckmann E, Elger C (1991) The neurophysiological basis of epileptic activity: A condensed overview. *Epilepsy Res Suppl* 2:1–7
  80. Steriade M, Jones EG, Llinas RR (1990) *Thalamic oscillations and signaling*. Wiley, New York
  81. Strogatz SH (2003) *Sync: The emerging science of spontaneous order*. Hyperion, New York
  82. Tasker RR (1998) Deep brain stimulation is preferable to thalamotomy for tremor suppression. *Surg Neurol* 49:145–154
  83. Tass P, Hauptmann C (2007) Therapeutic modulation of

synaptic connectivity with desynchronizing brain stimulation. *Int J Psychophysiol* 64:53–61

84. Tass PA (1996) Phase resetting associated with changes of burst shape. *J Biol Phys* 22:125–155
85. Tass PA (1996) Resetting biological oscillators – a stochastic approach. *J Biol Phys* 22:27–64
86. Tass PA (1999) Phase resetting in medicine and biology: Stochastic modelling and data analysis. Springer, Berlin
87. Tass PA (2000) Stochastic phase resetting: A theory for deep brain stimulation. *Prog Theor Phys Suppl* 139:301–313
88. Tass PA (2001) Desynchronizing double-pulse phase resetting and application to deep brain stimulation. *Biol Cybern* 85:343–354
89. Tass PA (2001) Effective desynchronization by means of double-pulse phase resetting. *Europhys Lett* 53:15–21
90. Tass PA (2001) Effective desynchronization with a resetting pulse train followed by a single-pulse. *Europhys Lett* 55:171–177
91. Tass PA (2002) Desynchronization of brain rhythms with soft phase-resetting techniques. *Biol Cybern* 87:102–115
92. Tass PA (2002) Effective desynchronization with a stimulation technique based on soft phase resetting. *Europhys Lett* 57:164–170
93. Tass PA (2002) Effective desynchronization with bipolar double-pulse stimulation. *Phys Rev E* 66:036226
94. Tass PA (2003) A model of desynchronizing deep brain stimulation with a demand-controlled coordinated reset of neural subpopulations. *Biol Cybern* 89:81–88
95. Tass PA (2003) Desynchronization by means of a coordinated reset of neural sub-populations – a novel technique for demand-controlled deep brain stimulation. *Prog Theor Phys Suppl* 150:281–296
96. Tass PA, Hauptmann C (2006) Therapeutic rewiring by means of desynchronizing brain stimulation. *Nonlinear Phenom Complex Syst* 9(3):298–312
97. Tass PA, Majtanik M (2006) Long-term anti-kindling effects of desynchronizing brain stimulation: A theoretical study. *Biol Cybern* 94:58–66
98. Tass PA, Hauptmann C, Popovych OV (2006) Development of therapeutic brain stimulation techniques with methods from nonlinear dynamics and statistical physics. *Int J Bif Chaos* 16(7):1889–1911
99. Timmermann L, Florin E, Reck C (2007) Pathological cerebral oscillatory activity in parkinson's disease: A critical review on methods, data and hypotheses. *Expert Rev Med Devices* 4(5):651–661
100. van Hemmen J (2001) Theory of synaptic plasticity. In: Moss F, Gielen S (eds) *Handbook of biologicla physics*, vol 4. Elsevier, Amsterdam, pp 771–823
101. Volkmann J (2004) Deep brain stimulation for the treatment of Parkinson's disease. *J Clin Neurophysiol* 21:6–17
102. Winfree A (1980) *The geometry of biological time*. Springer, Berlin
103. Wunderlin A, Haken H (1975) Scaling theory for nonequilibrium systems. *Z Phys B* 21:393–401
104. Zhai Y, Kiss IZ, Tass PA, Hudson JL (2005) Desynchronization of coupled electrochemical oscillators with pulse stimulations. *Phys Rev E* 71:065202R
105. Zhou Q, Tao H, Poo M (2003) Reversal and stabilization of synaptic modifications in a developing visual system. *Science* 300:1953–1957

## Branching Processes

MIKKO J. ALAVA<sup>1</sup>, KENT BÆKGAARD LAURITSEN<sup>2</sup>

<sup>1</sup> Department of Engineering Physics, Espoo University of Technology, Espoo, Finland

<sup>2</sup> Research Department, Danish Meteorological Institute, Copenhagen, Denmark

### Article Outline

[Glossary](#)

[Definition of the Subject](#)

[Introduction](#)

[Branching Processes](#)

[Self-Organized Branching Processes](#)

[Scaling and Dissipation](#)

[Network Science and Branching Processes](#)

[Conclusions](#)

[Future Directions](#)

[Acknowledgments](#)

[Bibliography](#)

### Glossary

**Markov process** A process characterized by a set of probabilities to go from a certain state at time  $t$  to another state at time  $t + 1$ . These transition probabilities are independent of the history of the process and only depend on a fixed probability assigned to the transition.

**Critical properties and scaling** The behavior of equilibrium and many non-equilibrium systems in steady states contain critical points where the systems display scale invariance and the correlation functions exhibit an algebraic behavior characterized by so-called critical exponents. A characteristics of this type of behavior is the lack of finite length and time scales (also reminiscent of fractals). The behavior near the critical points can be described by scaling functions that are universal and that do not depend on the detailed microscopic dynamics.

**Avalanches** When a system is perturbed in such a way that a disturbance propagates throughout the system one speak of an avalanche. The local avalanche dynamics may either conserve energy (particles) or dissipate energy. The avalanche may also loose energy when it reaches the system boundary. In the neighborhood of a critical point the avalanche distribution is described by a power-law distribution.

**Self-organized criticality (SOC)** SOC is the surprising “critical” state in which many systems from physics to biology to social ones find themselves. In physics jar-

gon, they exhibit scale-invariance, which means that the dynamics – consisting of avalanches – has no typical scale in time or space. The really necessary ingredient is that there is a hidden, fine-tuned balance between how such systems are driven to create the dynamic response, and how they dissipate the input (“energy”) to still remain in balance.

**Networks** These are descriptions of interacting systems, where in graph theoretical language nodes or vertices are connected by links or edges. The interesting thing can be the structure of the network, and its dynamics, or that of a process on the top of it like the spreading of computer viruses on the Internet.

### Definition of the Subject

Consider the fate of a human population on a small, isolated island. It consists of a certain number of individuals, and the most obvious question, of importance in particular for the inhabitants of the island, is whether this number will go to zero. Humans die and reproduce in steps of one, and therefore one can try to analyze this fate mathematically by writing down what is called master equations, to describe the dynamics as a “branching process” (BP). The branching here means that if at time  $t = 0$  there are  $N$  humans, at the next step  $t = 1$  there can be  $N - 1$  (or  $N + 1$  or  $N + 2$  if the only change from  $t = 0$  was that a pair of twins was born). The outcome will depend in the simplest case on a “branching number”, or the number of offspring  $\lambda$  that a human being will have [1,2,3,4].

If the offspring created are too few, then the population will decay, or reach an “absorbing state” out of which it will never escape. Likewise if they are many (the Malthusian case in reality, perhaps), exponential growth in time will ensue in the simplest case. In between, there is an interesting twist: a phase transition that separates these two outcomes at a critical value of  $\lambda_c$ . As is typical of such transitions in statistical physics one runs into scale-free behavior. The lifetime of the population suddenly has no typical scale, and its total size will be a stochastic quantity, described by a probability distribution that again has no typical scale exactly at  $\lambda_c$ .

The example of a small island also illustrates the many different twists that one can find in branching processes. The population can be “spatially dispersed” such that the individuals are separated by distance. There are in fact two interacting populations, called “male” and “female”, and if the size of one of the populations becomes zero the other one will die out soon as well. The people on the island eat, and there is thus a hidden variable in the dynamics, the availability of food. This causes a history

effect which makes the dynamics of human population what is called “non-Markovian”. Imagine as above, that we look at the number of persons on the island at discrete times. A Markovian process is such that the probabilities to go from a state (say of)  $N$  to state  $N + \delta$  depends only on the fixed probability assigned to the “transition”  $N \rightarrow N + \delta$ . Clearly, any relatively faithful description of the death and birth rates of human beings has to consider the average state of nourishment, or whether there is enough food for reproduction.

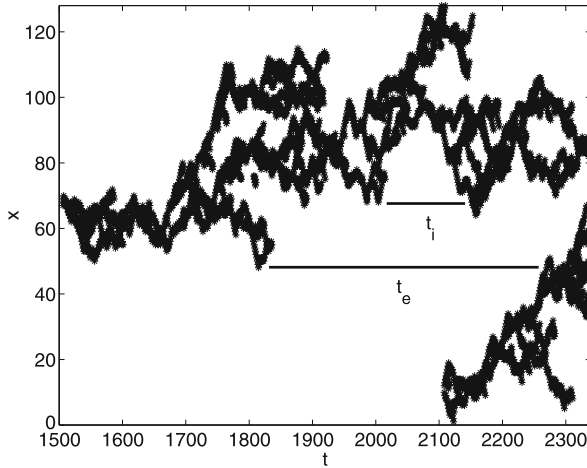
### Introduction

Branching processes are often perfect models of complex systems, or in other words exhibit deep complexity themselves. Consider the following example of a one-dimensional model of activated random walkers [5]. Take a line of sites  $x_i$ ,  $i = 1 \dots L$ . Fill the sites randomly to a certain density  $n = N/L$ , where  $N$  is the pre-set number of individuals performing the activated random walk. Now, let us apply the simple rule, that if there are two or more walkers at the same  $x_j$ , two of them get “activated” and hop to  $j - 1$  or  $j + 1$ , again at random. In other words, this version of the drunken bar-hoppers problem has the twist that they do not like each other.

If the system is “periodic” or  $i = 1$  is connected to  $i = L$ , then the dynamics is controlled by the density  $n$ . For a critical value  $n_c$  (estimated by numerical simulations to be about 0.9488... [6]) a phase transition takes place, such that for  $n < n_c$  the asymptotic state is the “absorbing one”, where all the walkers are immobilized since  $N_i = 1$  or 0. In the opposite case for  $n > n_c$  there is an active phase such that (in the infinite  $L$ -limit) the activity persists forever. This particular model is unquestionably non-Markovian if one only considers the number of active walkers or their density  $\rho$ . One needs to know the full state of  $N_i$  to be able to write down exact probabilities for how  $\rho$  changes in a discrete step of time.

The most interesting things happen if one changes the one-dimensional lattice by adapting two new rules. If a walker walks out (to  $i = 0$  or  $i = L + 1$ ), it disappears. Second, if there are no active ones ( $\rho = 0$ ), one adds one new walker randomly to the system. Now the activity  $\rho(t)$  is always at a marginal value, and the long-term average of  $n$  becomes a (possibly  $L$ -dependent) constant such that the first statement is true, i. e. the system becomes critical. With these rules, the model of activated random walkers is also known as the Manna model after the Indian physicist [5], and it exhibits the phenomenon dubbed “Self-Organized Criticality” (SOC) [7]. Figure 1 shows an example of the dynamics by using what is called an “activity plot”,





Branching Processes, Figure 1

We follow the “activity” in a one-dimensional system of random activated walkers. The walkers stroll around the  $x$ -axis, and the pattern becomes in fact scale-invariant. This system is such that some of the walkers disappear (by escaping through open boundaries) and to maintain a constant density new ones are added. One question one may ask is what is the waiting time, that another (or the same) walker gets activated at the same location after a time of inactivity. As one can see from the figure, this can be a result of old activity getting back, or a new “avalanche” starting from the addition of an outsider (courtesy of Lasse Laurson)

where those locations  $x_i$  are marked both in space and time which happen to contain just-activated walkers. One can now apply several kinds of measures to the system, but the figure already hints about the reasons why these simple models have found much interest. The structure of the activity is a self-affine fractal (for discussions about fractals see other reviews in this volume).

The main enthusiasm about SOC comes from the avalanches. These are in other words the bursts of activity that separate quiescent times (when  $\rho = 0$ ). The silence is broken by the addition of a particle or a walker, and it creates an integrated quantity (volume) of activity,  $s = \int_0^T \rho(t)dt$ , where for  $0 < t < T$ ,  $\rho > 0$  and  $\rho = 0$  at the endpoints  $t$  and  $T$ . The original boost to SOC took place after Per Bak, Chao Tang, and Kay Wiesenfeld published in 1987 a highly influential paper in the premium physics journal *Physical Review Letters*, and introduced what is called the Bak–Tang–Wiesenfeld (BTW) sandpile model – of which the Manna one is a relative [7]. The BTW and Manna models and many others exhibit the important property that the avalanche sizes  $s$  have a *scale-free* probability distribution but note that this simple criticality is not always true, not even for the BTW model which shows so-called multiscaling [8,9]. Scale-free probability distributions

are usually written as

$$P(s) \sim s^{-\tau_s} f_s(s/L^{D_s}), \quad (1)$$

here all the subscripts refer to the fact that we look at avalanches.  $\tau_s$  and  $D_s$  define the avalanche exponent and the cut-off exponent, respectively.  $f_s$  is a cut-off function that together with  $D_s$  includes the fact that the avalanches are restricted somehow by the system size (in the one-dimensional Manna model, if  $s$  becomes too large many walkers are lost, and first  $n$  drops and then  $\rho$  goes to zero). Similar statements can be made about the avalanche durations ( $T$ ), area or support  $A$ , and so forth [10,11,12,13,14].

The discovery of simple-to-define models, yet of very complex behavior has been a welcome gift since there are many, many phenomena that exhibit apparently scale-free statistics and/or bursty or intermittent dynamics similar to SOC models. These come from natural sciences – but not only, since economy and sociology are also giving rise to cases that need explanations. One particular field where these ideas have found much interest is the physics of materials ranging from understanding earthquakes to looking at the behavior of vortices in superconductors. The Gutenberg–Richter law of earthquake magnitudes is a power-law, and one can measure similar statistics from fracture experiments by measuring acoustic emission event energies from tensile tests on ordinary paper [15]. The modern theory of networks is concerned with graph or network structures which exhibit scale-free characteristics, in particular the number of neighbors a vertex has is often a stochastic quantity exhibiting a power-law-like probability distribution [16,17]. Here, branching processes are also finding applications. They can describe the development of network structures, e.g., as measured by the degree distribution, the number of neighbors a node (vertex) is connected to, and as an extension of more usual models give rise to interesting results when the dynamics of populations on networks are being considered.

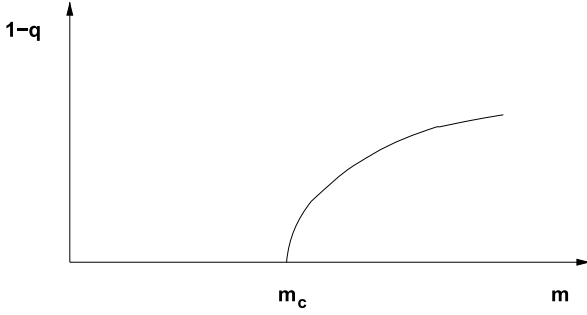
Whenever comparing with real, empirical phenomena models based on branching processes give a paradigm on two levels. In the case of SOC this is given by a combination of “internal dynamics” and “an ensemble”. Of these, the first means e.g. that activated walkers or particles of type  $A$  are moved around with certain rules, and of course there is an enormous variety of possible models. e.g. in the Manna-model this is obvious if one splits the walkers into categories  $A$  (active) and  $B$  (passive). Then, there is the question of how the balance (assuming this balance is true) is maintained. The SOC models do this by a combination of dissipation (by e.g. particles dropping off the edge of a system) and drive (by addition of  $B$ ’s), where the rates are chosen actually pretty carefully. In the case of growing

The theory of branching processes provide many answers to these questions, and in particular help to illustrate the influence of boundary conditions and modes of driving on the expected behavior. Thus, one gets a clear idea of the kind of complexity one can expect to see in the many different kinds of systems where avalanches and intermittency and scaling is observed.

The mathematical branching process is defined for a set of objects that do not interact. At each iteration, each object can give rise to new objects with some probability  $p$  (or in general it can be a set of probabilities). By continuing this iterative process the objects will form what is referred to as a cluster or an avalanche. We can now ask questions of the following type: Will the process continue for ever? Will the process maybe die out and stop after a finite number of iterations? What will the average lifetime be? What is the average size of the clusters? And what is (the asymptotic form of) the probability that the process is active after a certain number of iterations, etc.

Figure 1 shows a tree diagram representing the construction of a  $\sigma_3(p, t)$ -tree. The root node is black. It branches into two nodes: a black node labeled  $p$  and a white node labeled  $1-p$ . The black node  $p$  further branches into two black nodes, labeled  $1-p$  and  $p$ . The white node  $1-p$  branches into two white nodes, labeled  $1-p$  and  $p$ . Each of these four nodes then branches into two more nodes, resulting in a total of 16 leaf nodes. The diagram is labeled  $\sigma_3(p, t)$  on the right.

Schematic drawing of an avalanche in a system with a maximum of  $n = 3$  avalanche generations corresponding to  $N = 2^{n+1} - 1 = 15$  sites. Each black site relaxes with probability  $p$  to two new black sites and with probability  $1 - p$  to two white sites (i. e.,  $p_0 = 1 - p, p_1 = 0, p_2 = p$ ). The black sites are part of an avalanche of size  $s = 7$ , whereas the active sites at the boundary yield a boundary avalanche of size  $\sigma_3(p, t) = 2$



**Branching Processes, Figure 3**

Schematic drawing of the behavior of the probability of extinction  $q$ , of the branching process. The quantity  $1 - q$  is similar to the order parameter for systems in equilibrium at their critical points and the quantity  $m_c$  is referred to as the critical point for the BP

By taking derivatives of the generating function at  $s = 1$  it follows that:

$$Ez_n = m^n, \quad \text{Var } z_n = n\sigma^2. \quad (5)$$

For the BP defined in Fig. 2, it follows that  $m = 2p$  and  $\sigma^2 = 4p(1 - p)$ .

An important quantity is the probability for extinction  $q$ . It is obtained as follows:

$$q = P(z_n \rightarrow 0) = P(z_n = 0 \text{ for some } n) = \lim_n P(z_n = 0). \quad (6)$$

It can be shown that for  $m \leq 1$ :  $q = 1$ ; and for  $m > 1$ : there exists a solution that fulfills  $q = f(q)$ , where  $0 < q < 1$  [1]. It is possible to show that  $\lim_n P(z_n = k) = 0$ , for  $k = 1, 2, 3, \dots$ , and that  $z_n \rightarrow 0$  with probability  $q$ , and  $z_n \rightarrow \infty$ , with probability  $1 - q$ . Thus, the sequence  $\{z_n\}$  does not remain positive and bounded [1].

The quantity  $1 - q$  is similar to an order parameter for systems in equilibrium and its behavior is schematically shown in Fig. 3. The behavior around the value  $m_c = 1$ , the so-called critical value for the BP (see below), can in analogy to second order phase transitions in equilibrium systems be described by a critical exponent  $\beta$  defined as follows ( $\beta = 1$ , cf. [1]):

$$\text{Prob}(\text{survival}) = \begin{cases} 0, & m \leq m_c \\ (m - m_c)^\beta, & m > m_c. \end{cases} \quad (7)$$

### Avalanches and Critical Point

We will next consider the clusters, or avalanches, in more detail and obtain the asymptotic form of the probability distributions. The size of the cluster is given by the sum

$z = z_1 + z_2 + z_3 + \dots$ . One can also consider other types of clusters, e. g., the activity  $\sigma$  of the boundary (of a finite tree) and define the boundary avalanche (cf. Fig. 2).

For concreteness, we consider the BP defined in Fig. 2. The quantities  $P_n(s, p)$  and  $Q_n(\sigma, p)$  denote the probabilities of having an avalanche of size  $s$  and boundary size  $\sigma$  in a system with  $n$  generations. The corresponding generating functions are defined by [1]

$$f_n(x, p) \equiv \sum_s P_n(s, p)x^s \quad (8)$$

$$g_n(x, p) \equiv \sum_\sigma Q_n(\sigma, p)x^\sigma. \quad (9)$$

Due to the hierarchical structure of the branching process, it is possible to write down recursion relations for  $P_n(s, p)$  and  $Q_n(\sigma, p)$ :

$$f_{n+1}(x, p) = x[(1 - p) + pf_n^2(x, p)], \quad (10)$$

$$g_{n+1}(x, p) = (1 - p) + pg_n^2(x, p), \quad (11)$$

where  $f_0(x, p) = g_0(x, p) = x$ . The avalanche distribution  $D(s)$  is determined by  $P_n(s, p)$  by using the recursion relation (10). The stationary solution of Eq. (10) in the limit  $n \gg 1$  is given by

$$f(x, p) = \frac{1 - \sqrt{1 - 4x^2p(1 - p)}}{2xp}. \quad (12)$$

By expanding Eq. (12) as a series in  $x$ , comparing with the definition (8), and using Stirling's formula for the high-order terms, we obtain for sizes such that  $1 \ll s \lesssim n$  the following behavior:

$$P_n(s, p) = \frac{\sqrt{2(1 - p)/\pi p}}{s^{3/2}} \exp(-s/s_c(p)). \quad (13)$$

The cutoff  $s_c(p)$  is given by  $s_c(p) = -2/\ln[4p(1 - p)]$ . As  $p \rightarrow 1/2$ ,  $s_c(p) \rightarrow \infty$ , thus showing explicitly that the critical value for the branching process is  $p_c = 1/2$  (i. e.,  $m_c = 1$ ), and that the mean-field avalanche exponent, cf. Eq. (1), for the critical branching process is  $\tau = 3/2$ .

The expression (13) is only valid for avalanches which are not effected by the finite size of the system. For avalanches with  $n \lesssim s \lesssim N$ , it is possible to solve the recursion relation (10), and then obtain  $P_n(s, p)$  for  $p \geq p_c$  by the use of a Tauberian theorem [19,20,21]. By carrying out such an analysis one obtains after some algebra  $P_n(s, p) \approx A(p) \exp(-s/s_0(p))$ , with functions  $A(p)$  and  $s_0(p)$  which can not be determined analytically. Nevertheless, we see that for any  $p$  the probabilities  $P_n(s, p)$  will decay exponentially. One can also calculate the asymptotic form of  $Q_n(\sigma, p)$  for  $1 \ll \sigma \lesssim n$  and  $p \geq p_c$  by the use of

a Tauberian theorem [19,20,21]. We will return to study this in the next section where we will also define and investigate the distribution of the time to extinction.

### Self-Organized Branching Processes

We now return to discussing the link between self-organized criticality as mentioned in the introduction and branching processes. The simplest theoretical approach to SOC is mean-field theory [22], which allows for a qualitative description of the behavior of the SOC state. Mean-field exponents for SOC models have been obtained by various approaches [22,23,24,25] and it turns out that their values (e. g.,  $\tau = 3/2$ ) are the same for all the models considered thus far. This fact can easily be understood since the spreading of an avalanche in mean-field theory can be described by a front consisting of non-interacting particles that can either trigger subsequent activity or die out. This kind of process is reminiscent of a branching process. The connection between branching processes and SOC has been investigated, and it has been argued that the mean-field behavior of sandpile models can be described by a critical branching process [26,27,28].

For a branching process to be critical one must fine tune a control parameter to a critical value. This, by definition, cannot be the case for a SOC system, where the critical state is approached dynamically without the need to fine tune any parameter. In the so-called self-organized branching-process (SOBP), the coupling of the local dynamical rules to a global condition drives the system into a state that is indeed described by a critical branching process [29]. It turns out that the mean-field theory of SOC models can be exactly mapped to the SOBP model.

In the mean-field description of the sandpile model ( $d \rightarrow \infty$ ) one neglects correlations, which implies that avalanches do not form loops and hence spread as a branching process. In the SOBP model, an avalanche starts with a single active site, which then relaxes with probability  $p$ , leading to two new active sites. With probability  $1 - p$  the initial site does not relax and the avalanche stops. If the avalanche does not stop, one repeats the procedure for the new active sites until no active site remains. The parameter  $p$  is the probability that a site relaxes when it is triggered by an external input. For the SOBP branching process, there is a critical value,  $p_c = 1/2$ , such that for  $p > p_c$  the probability to have an infinite avalanche is non-zero, while for  $p < p_c$  all avalanches are finite. Thus,  $p = p_c$  corresponds to the critical case, where avalanches are power law distributed.

In this description, however, the boundary conditions are not taken into account – even though they are cru-

cial for the self-organization process. The boundary conditions can be introduced in the problem in a natural way by allowing for no more than  $n$  generations for each avalanche. Schematically, we can view the evolution of a single avalanche of size  $s$  as taking place on a tree of size  $N = 2^{n+1} - 1$  (see Fig. 2). If the avalanche reaches the boundary of the tree, one counts the number of active sites  $\sigma_n$  (which in the sandpile language corresponds to the energy leaving the system), and we expect that  $p$  decreases for the next avalanche. If, on the other hand, the avalanche stops before reaching the boundary, then  $p$  will slightly increase. The number of generations  $n$  can be thought of as some measure of the linear size of the system.

The above avalanche scenario is described by the following dynamical equation for  $p(t)$ :

$$p(t+1) = p(t) + \frac{1 - \sigma_n(p, t)}{N}, \quad (14)$$

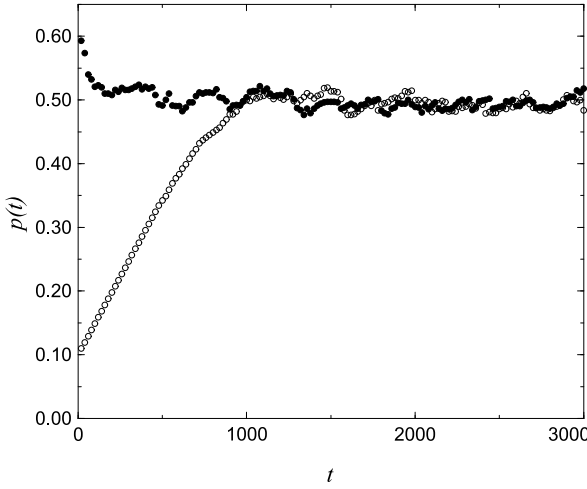
where  $\sigma_n$ , the size of an avalanche reaching the boundary, fluctuates in time and hence acts as a stochastic driving force. If  $\sigma_n = 0$ , then  $p$  increases (because some energy has been put into the system without any output), whereas if  $\sigma_n > 0$  then  $p$  decreases (due to energy leaving the system). Equation (14) describes the global dynamics of the SOBP, as opposed to the local dynamics which is given by the branching process. One can study the model for a fixed value of  $n$ , and then take the limit  $n \rightarrow \infty$ . In this way, we perform the long-time limit before the “thermodynamic” limit, which corresponds exactly to what happens in sandpile models.

We will now show that the SOBP model provides a mean-field theory of self-organized critical systems. Consider for simplicity the sandpile model of activated random walkers from the Introduction [5]: When a particle is added to a site  $z_i$ , the site will relax if  $z_i = 1$ . In the limit  $d \rightarrow \infty$ , the avalanche will never visit the same site more than once. Accordingly, each site in the avalanche will relax with the same probability  $p = P(z = 1)$ . Eventually, the avalanche will stop, and  $\sigma \geq 0$  particles will leave the system. Thus, the total number of particles  $M(t)$  evolves according to

$$M(t+1) = M(t) + 1 - \sigma. \quad (15)$$

The dynamical Eq. (14) for the SOBP model is recovered by noting that  $M(t) = NP(z = 1) = Np$ . By taking the continuum time limit of Eq. (14), it is possible to obtain the following expression:

$$\frac{dp}{dt} = \frac{1 - (2p)^n}{N} + \frac{\eta(p, t)}{N}, \quad (16)$$



**Branching Processes, Figure 4**

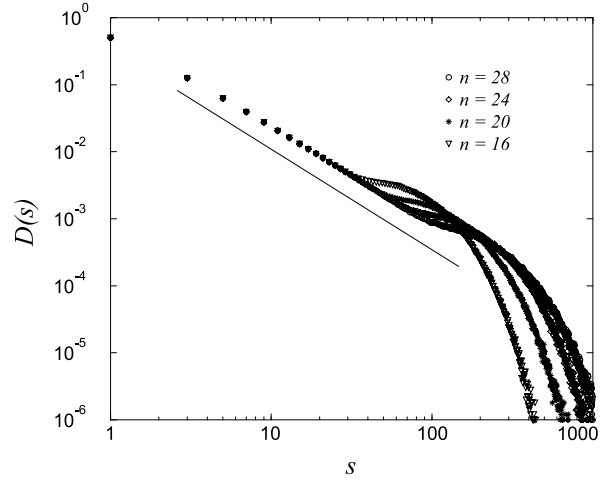
The value of  $p$  as a function of time for a system with  $n = 10$  generations. The two curves refer to two different initial conditions, above  $p_c$  (●) and below  $p_c$  (○). After a transient, the control parameter  $p(t)$  reaches its critical value  $p_c$  and fluctuates around it with short-range correlations

where  $\eta \equiv \langle \sigma_n \rangle - \sigma_n = (2p)^n - \sigma_n(p, t)$  describes the fluctuations in the steady state. Thus,  $\eta$  is obtained by measuring  $\sigma_n$  for each realization of the process. Without the last term, Eq. (16) has a fixed point ( $dp/dt = 0$ ) for  $p = p_c = 1/2$ . On linearizing Eq. (16), one sees that the fixed point is attractive, which demonstrates the self-organization of the SOBP model since the noise  $\eta/N$  will have vanishingly small effect in the thermodynamic limit [29].

Figure 4 shows the value of  $p$  as a function of time. Independent of the initial conditions, one finds that after a transient  $p(t)$  reaches the self-organized state described by the critical value  $p_c = 1/2$  and fluctuates around it with short-range correlations (of the order of one time unit). By computing the variance of  $p(t)$ , one finds that the fluctuations can be very well described by a Gaussian distribution,  $\phi(p)$  [29]. In the limit  $N \rightarrow \infty$ , the distribution  $\phi(p)$  approaches a delta function,  $\delta(p - p_c)$ .

### Avalanche Distributions

Figure 5 shows the avalanche size distribution  $D(s)$  for different values of the number of generations  $n$ . One notices that there is a scaling region ( $D(s) \sim s^{-\tau}$  with  $\tau = 3/2$ ), whose size increases with  $n$ , and characterized by an exponential cutoff. This power-law scaling is a signature of the mean-field criticality of the SOBP model. The distribution of active sites at the boundary,  $D(\sigma)$ , for different values of the number of generations falls off exponentially [29].



**Branching Processes, Figure 5**

Log-log plot of the avalanche distribution  $D(s)$  for different system sizes. The number of generations  $n$  increases from left to right. A line with slope  $\tau = 3/2$  is plotted for reference, and it describes the behavior of the data for intermediate  $s$  values, cf. Eq. (18). For large  $s$ , the distributions fall off exponentially

In the limit where  $n \gg 1$  one can obtain various analytical results and e. g. calculate the avalanche distribution  $D(s)$  for the SOBP model. In addition, one can obtain results for finite, but large, values of  $n$ . The distribution  $D(s)$  can be calculated as the average value of  $P_n(s, p)$  with respect to the probability density  $\phi(p)$ , i. e., according to the formula

$$D(s) = \int_0^1 dp \phi(p) P_n(s, p). \quad (17)$$

The simulation results in Fig. 4 yield that  $\phi(p)$  for  $N \gg 1$  approaches the delta function  $\delta(p - p_c)$ . Thus, from Eqs. (13) and (17) we obtain the power-law behavior

$$D(s) = \sqrt{\frac{2}{\pi}} s^{-\tau}, \quad (18)$$

where  $\tau = 3/2$ , and for  $s \gtrsim n$  we obtain an exponential cutoff  $\exp(-s/s_0(p_c))$ . These results are in complete agreement with the numerical results shown in Fig. 5. The deviations from the power-law behavior (18) are due to the fact that Eq. (13) is only valid for  $1 \ll s \lesssim n$ . One can also calculate the asymptotic form of  $Q_n(\sigma, p)$  for  $1 \ll \sigma \lesssim n$  and  $p \geq p_c$  by the use of a Tauberian theorem [19,20,21]; the result shows that the boundary avalanche distribution is

$$D(\sigma) = \int_0^1 dp \phi(p) Q_n(\sigma, p) = \frac{8}{n^2} \exp(-2\sigma/n), \quad (19)$$

which agrees with simulation results for  $n \gg 1$ , cf. [29].



The avalanche lifetime distribution  $L(t) \sim t^{-\gamma}$  yields the probability to have an avalanche that lasts for a time  $t$ . For a system with  $m$  generations one obtains  $L(m) \sim m^{-2}$  [1]. Identifying the number of generations  $m$  of an avalanche with the time  $t$ , we thus obtain the mean-field value  $\gamma = 2$ , in agreement with simulations of the SOBP model [29]. In summary, the self-organized branching process captures the physical features of the self-organization mechanism in sandpile models. By explicitly incorporating the boundary conditions it follows that the dynamics drives the system into a stationary state, which in the thermodynamic limit corresponds to the critical branching process.

### Scaling and Dissipation

Sometimes it can be difficult to determine whether the cutoff in the scaling is due to finite-size effects or due to the fact that the system is not *at* but rather only *close to* the critical point. In this respect, it is important to test the robustness of critical behavior by understanding which perturbations destroy the critical properties. It has been shown numerically [30,31,32] that the breaking of the conservation of particle numbers leads to a characteristic size in the avalanche distributions. We will now allow for dissipation in branching processes and show how the system self-organizes into a sub-critical state. In other words, the degree of nonconservation is a relevant parameter in the renormalization group sense [33].

Consider again the two-state model introduced by Manna [5]. Some degree of nonconservation can be introduced in the model by allowing for energy dissipation in a relaxation event. In a continuous energy model this can be done by transferring to the neighboring sites only a fraction  $(1 - \epsilon)$  of the energy lost by the relaxing site [30]. In a discrete energy model, such as the Manna two-state model, one can introduce dissipation as the probability  $\epsilon$  that the two particles transferred by the relaxing site are annihilated [31]. For  $\epsilon = 0$  one recovers the original two-state model.

Numerical simulations [30,31] show that different ways of considering dissipation lead to the same effect: a characteristic length is introduced into the system and the criticality is lost. As a result, the avalanche size distribution decays not as a pure power law but rather as

$$D(s) \sim s^{-\tau} h_s(s/s_c) . \quad (20)$$

Here  $h_s(x)$  is a cutoff function and the cutoff size scales as

$$s_c \sim \epsilon^{-\varphi} . \quad (21)$$

The size  $s$  is defined as the number of sites that relax in an avalanche. We define the avalanche lifetime  $T$  as the number of steps comprising an avalanche. The corresponding distribution decays as

$$D(T) \sim T^{-\gamma} h_T(T/T_c) , \quad (22)$$

where  $h_T(x)$  is another cutoff function and  $T_c$  is a cutoff that scales as

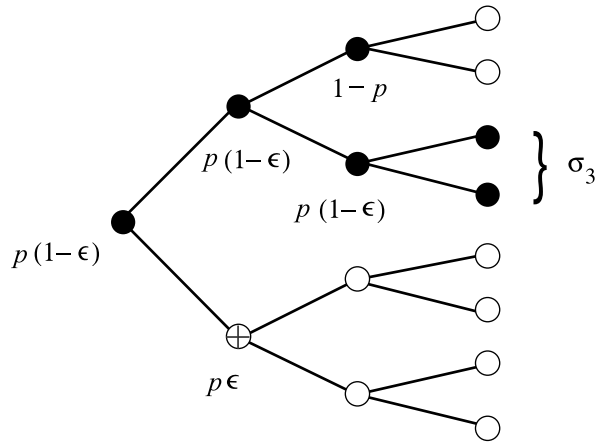
$$T_c \sim \epsilon^{-\psi} . \quad (23)$$

The cutoff or “scaling” functions  $h_s(x)$  and  $h_T(x)$  fall off exponentially for  $x \gg 1$ .

To construct the mean-field theory one proceeds as follows [34]: When a particle is added to an arbitrary site, the site will relax if a particle was already present, which occurs with probability  $p = P(z = 1)$ , the probability that the site is occupied. If a relaxation occurs, the two particles are transferred with probability  $1 - \epsilon$  to two of the infinitely many nearest neighbors, or they are dissipated with probability  $\epsilon$  (see Fig. 6).

The avalanche process in the mean-field limit is a branching process. Moreover, the branching process can be described by the *effective* branching probability

$$\tilde{p} \equiv p(1 - \epsilon) , \quad (24)$$



**Branching Processes, Figure 6**

Schematic drawing of an avalanche in a system with a maximum of  $n = 3$  avalanche generations corresponding to  $N = 2^{n+1} - 1 = 15$  sites. Each black site (●) can relax in three different ways: (i) with probability  $p(1 - \epsilon)$  to two new black sites, (ii) with probability  $1 - p$  the avalanche stops, and (iii) with probability  $p\epsilon$  two particles are dissipated at a black site, which then becomes a marked site (⊕), and the avalanche stops. The black sites are part of an avalanche of size  $s = 6$ , whereas the active sites at the boundary yield  $\sigma_3(p, t) = 2$ . There was one dissipation event such that  $\kappa = 2$ .

where  $\tilde{p}$  is the probability to create two new active sites. We know that there is a critical value for  $\tilde{p} = 1/2$ , or

$$p = p_c \equiv \frac{1}{2(1 - \epsilon)}. \quad (25)$$

Thus, for  $p > p_c$  the probability to have an infinite avalanche is non-zero, while for  $p < p_c$  all avalanches are finite. The value  $p = p_c$  corresponds to the critical case where avalanches are power law distributed.

### The Properties of the Steady State

To address the self-organization, consider the evolution of the total number of particles  $M(t)$  in the system after each avalanche:

$$M(t + 1) = M(t) + 1 - \sigma(p, t) - \kappa(p, t). \quad (26)$$

Here  $\sigma$  is the number of particles that leave the system from the boundaries and  $\kappa$  is the number of particles lost by dissipation. Since (cf. Sect. “Self-Organized Branching Processes”)  $M(t) = NP(z = 1) = Np$ , we obtain an evolution equation for the parameter  $p$ :

$$p(t + 1) = p(t) + \frac{1 - \sigma(p, t) - \kappa(p, t)}{N}. \quad (27)$$

This equation reduces to the SOBP model for the case of no dissipation ( $\kappa = 0$ ). In the continuum limit one obtains [34]

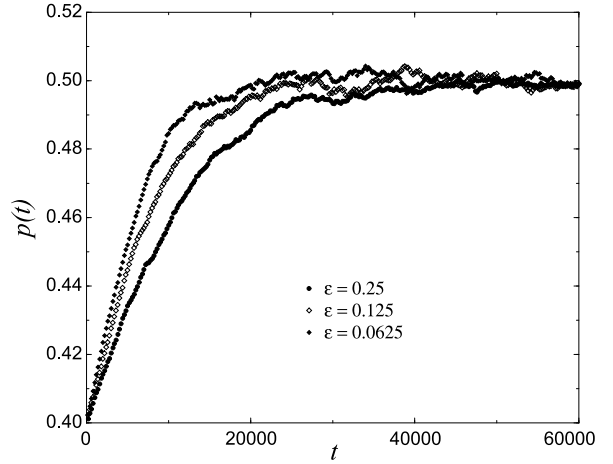
$$\frac{dp}{dt} = \frac{1}{N} \left[ 1 - (2p(1 - \epsilon))^n - p\epsilon H(p(1 - \epsilon)) \right] + \frac{\eta(p, t)}{N}. \quad (28)$$

Here, we defined the function  $H(p(1 - \epsilon))$ , that can be obtained analytically, and introduced the function  $\eta(p, t)$  to describe the fluctuations around the average values of  $\sigma$  and  $\kappa$ . It can be shown numerically that the effect of this “noise” term is vanishingly small in the limit  $N \rightarrow \infty$  [34].

Without the noise term one can study the fixed points of Eq. (28) and one finds that there is only one fixed point,

$$p^* = 1/2, \quad (29)$$

independent of the value of  $\epsilon$ ; the corrections to this value are of the order  $O(1/N)$ . By linearizing Eq. (28), it follows that the fixed point is attractive. This result implies that the SOBP model self-organizes into a state with  $p = p^*$ . In Fig. 7 is shown the value of  $p$  as a function of time for different values of the dissipation  $\epsilon$ . We find that independent of the initial conditions after a transient  $p(t)$  reaches the self-organized steady-state described by the



**Branching Processes, Figure 7**

The value of the control parameter  $p(t)$  as a function of time for a system with different levels of dissipation. After a transient,  $p(t)$  reaches its fixed-point value  $p^* = 1/2$  and fluctuates around it with short-range time correlations

fixed point value  $p^* = 1/2$  and fluctuates around it with short-range correlations (of the order of one time unit). The fluctuations around the critical value decrease with the system size as  $1/N$ . It follows that in the limit  $N \rightarrow \infty$  the distribution  $\phi(p)$  of  $p$  approaches a delta function  $\phi(p) \sim \delta(p - p^*)$ .

By comparing the fixed point value (29) with the critical value (25), we obtain that in the presence of dissipation ( $\epsilon > 0$ ) the self-organized steady-state of the system is *subcritical*. Figure 8 is a schematic picture of the phase diagram of the model, including the line  $p = p_c$  of critical behavior (25) and the line  $p = p^*$  of fixed points (29). These two lines intersect only for  $\epsilon = 0$ .

### Avalanche and Lifetime Distributions

In analogy to the results in Sect. “Self-Organized Branching Processes”, we obtain similar formulas for the avalanche size distributions but with  $\tilde{p}$  replacing  $p$ . As a result we obtain the distribution

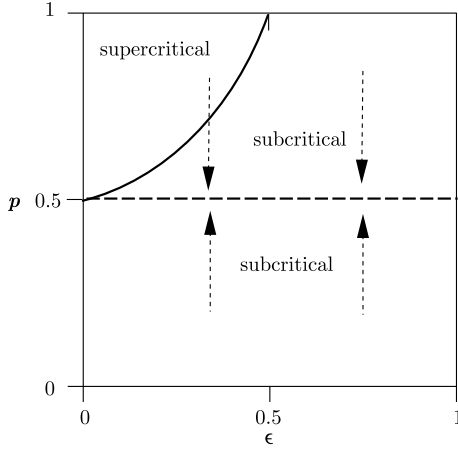
$$D(s) = \sqrt{\frac{2}{\pi}} \frac{1 + \epsilon + \dots}{s^\tau} \exp(-s/s_c(\epsilon)). \quad (30)$$

We can expand  $s_c(\tilde{p}(\epsilon)) = -2/\ln[4\tilde{p}(1 - \tilde{p})]$  in  $\epsilon$  with the result

$$s_c(\epsilon) \sim \frac{2}{\epsilon^\varphi}, \quad \varphi = 2. \quad (31)$$

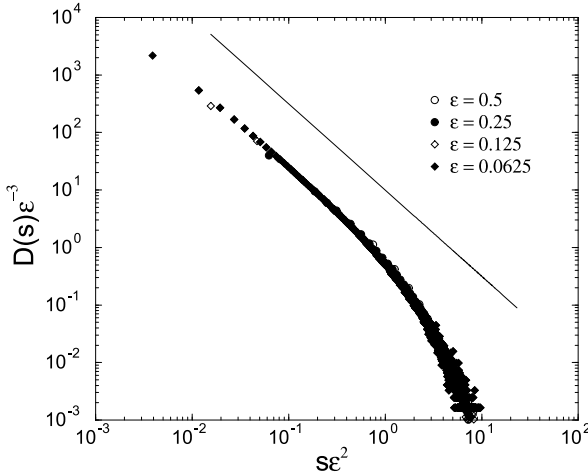
Furthermore, the mean-field exponent for the critical branching process is obtained setting  $\epsilon = 0$ , i. e.,

$$\tau = 3/2. \quad (32)$$



Branching Processes, Figure 8

Phase diagram for the SOBP model with dissipation. The dashed line shows the fixed points  $p^* = 1/2$  of the dynamics, with the flow being indicated by the arrows. The solid line shows the critical points, cf. Eq. (25)

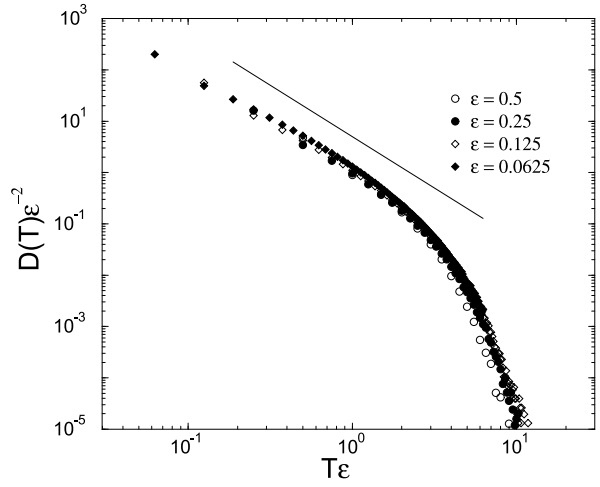


Branching Processes, Figure 9

Log-log plot of the avalanche distribution  $D(s)$  for different levels of dissipation. A line with slope  $\tau = 3/2$  is plotted for reference, and it describes the behavior of the data for intermediate  $s$  values, cf. Eq. (30). For large  $s$ , the distributions fall off exponentially. The data collapse is produced according to Eq. (30)

These results are in complete agreement with the SOBP model and the simulation of  $D(s)$  for the SOBP model with dissipation (cf. Fig. 9). The deviations from the power-law behavior (30) are due to the fact that Eq. (13) is only valid for  $1 \ll s \lesssim n$ .

Next, consider the avalanche lifetime distribution  $D(T)$  characterizing the probability to obtain an avalanche which spans  $m$  generations. It can be shown that the result



Branching Processes, Figure 10

Log-log plot of the lifetime distribution  $D(T)$  for different levels of dissipation. A line with slope  $\gamma = 2$  is plotted for reference. Note the initial deviations from the power law for  $\epsilon = 0$  due to the strong corrections to scaling. The data collapse is produced according to Eq. (33)

can be expressed in the scaling form [1,34]

$$D(T) \sim T^{-\gamma} \exp(-T/T_c), \quad (33)$$

where

$$T_c \sim \epsilon^{-\psi}, \quad \psi = 1. \quad (34)$$

The lifetime exponent  $\gamma$  was defined in Eq. (22), wherefrom we confirm the mean-field result

$$\gamma = 2. \quad (35)$$

In Fig. 10, we show the data collapse produced by Eq. (33) for the lifetime distributions for different values of  $\epsilon$ .

In summary, the effect of dissipation on the dynamics of the sandpile model in the mean-field limit ( $d \rightarrow \infty$ ) is described by a branching process. The evolution equation for the branching probability has a single attractive fixed point which in the presence of dissipation is not a critical point. The level of dissipation  $\epsilon$  therefore acts as a relevant parameter for the SOBP model. These results show, in the mean-field limit, that criticality in the sandpile model is lost when dissipation is present.

## Network Science and Branching Processes

An interesting mixture of two kinds of “complex systems” is achieved by considering what kind of applications one can find of branching processes applied to or on networks.

Here, network is a loose term in the way it is often used – as in the famous “six degrees” concept of how the social interactions of human beings can be measured to have a “small world” character [16,17,35]. These structures are usually thought of in terms of graphs, with vertices forming a set  $G$ , and edges a set  $E$  such that if  $e_{gg'} \in E$  it means that it connects two sites  $g$  and  $g' \in G$ . An edge or link can either be symmetric or asymmetric. Another article in this volume discusses in depth the modern view on networks, providing more information. The Structural properties of networks heterogeneous networks are most usually measured by the degree  $k_i$  of site  $i \in G$ , i. e. the number of nearest neighbors the vertex  $i$  is connected to. One can go further by defining weights  $w_{ij}$  for edges  $e_{ij}$  [36,37,38].

The interest in heterogeneous networks is largely due to two typical properties one can find. In a scale-free network the structure is such that the probability distribution has a power-law form  $P(k) \sim k^{-\alpha}$  up to a size-dependent cut-off  $k_{\max}$ . One can now show using mean-field-like reference models (such as the “configuration model”, which has no structural correlations and a prefixed  $P(k)$ ), that it follows that these networks tend to have a “small world” character, that the average vertex-vertex distance is only logarithmic in the number of vertices  $N$ :  $d(G) \sim \ln N$ . This property is correlated with the fact that the various moments of  $k$  depend on  $\alpha$ . For  $\langle k \rangle$  to remain finite,  $\alpha > 2$  is required. In many empirically measured networks it appears that  $\alpha < 3$ , which implies that  $\langle k^2 \rangle$  diverges (we are assuming, that  $P(k)$  is cut off at a maximum degree  $k_{\max}$  for which  $\int_{k_{\max}}^{\infty} \sim 1/N$ , from which the divergences follow) [16,17,35].

The first central question of the two we pose here is, what happens to branching processes on the top of such networks? Again, this can be translated into an inquiry about the “phase diagram” – given a model – and its particulars. We would like to understand what is the critical threshold  $\lambda_c$ , and what happens to sub-critical and super-critical systems (with appropriate values of  $\lambda$ ). The answers may of course depend on the model at hand, i. e. the universality classes are still an important issue.

One easy analytical result for un-correlated (in particular tree-like) networks is that activity will persist in the branching process if the number of offspring is at least one for an active “individual” or vertex. Consider now one such vertex,  $i$ . Look at the neighbors of neighbors. There are of the order  $O(k_i \times \langle k_{\text{neighbor of } i} \rangle - 1)$  of these, so  $\lambda_c \sim 1/\langle k^2 \rangle$ . This demonstrates the highly important result, that the *critical threshold may vanish* in heterogeneous networks, for  $\alpha \leq 3$  in the example case.

Pastor-Satorras and Vespignani were the first to point out this fact for the Susceptible-Infected-Susceptible (SIS)

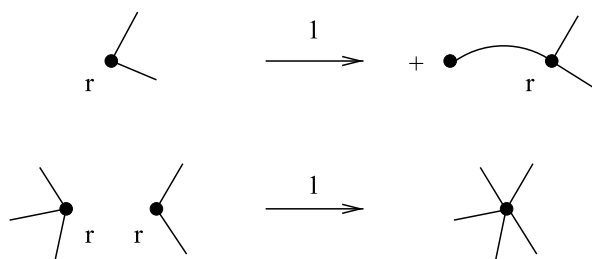
model [39]. The analysis is relatively easy to do using mean-field theory for  $\rho^k$ , the average activity for the SIS branching process on nodes of degree  $k$ . The central equation reads

$$\partial_t \rho^k(t) = -\rho^k(t) + \lambda k \left[ 1 - \rho^k(t) \right] \Theta_k(\lambda). \quad (36)$$

Here  $\Theta$  measures the probability that at least one of the neighbors of a vertex of degree  $k$  is infected. A MF-treatment of  $\Theta$ , excluding degree correlations, makes it possible to show that the SIS-model may have a zero threshold, and also establishes other interesting properties as  $\langle \rho_k \rangle$ . As is natural, the branching process concentrates on vertices with higher degrees, that is  $\rho_k$  increases with  $k$ .

The consequences of a zero epidemic threshold are plentiful and important [40]. The original application of the theory was to the spreading of computer viruses: the Internet is (on the Autonomous System level) a scale-free network [41]. So is the network of websites. The outcome even for subcritical epidemics is a long life-time, and further complications ensue if the internal dynamics of the spreading do not follow Poissonian statistics of time (e.g. the waiting time before vertex  $i$  sends a virus to a neighbor), but a power-law distribution [42]. Another highly topical issue is the behavior and control of human-based disease outbreaks. How to vaccinate against and isolate a dangerous virus epidemic depends on the structure of the network on which the spreading takes place, and on the detailed rules and dynamics. An intuitively easy idea is to concentrate on the “hubs”, on the most connected vertices of the network [43]. This can be an airport through which travellers acting as carriers move, or it can – as in the case of sexually transmitted viruses such as HIV – be the most active individuals. There are indeed claims that the network of sexual contacts is scale-free, with a small exponent  $\alpha$ . Of interest here is the recent result that for  $\lambda > \lambda_c$  and  $\alpha < 3$  the spreading of branching processes deviates from usual expectations. The outgrowth covers a finite fraction of the network in a short time (vanishing in the large- $N$  limit), and the growth in time of the infected nodes becomes polynomial, after an initial exponential phase [44].

Branching processes can also be applied to network structure. The “standard model” of growing networks is the Barabasi-Albert model. In it, from a small seed graph one grows an example network by so-called preferential attachment. That is, there are two mechanisms that operate: i) new vertices are added, and ii) old ones grow more links (degree increases) by getting linked to new ones. There is a whole zoo of various network growth models, but we next overview some ideas that directly connect to



Branching Processes, Figure 11

Two processes creating a growing or steady-state network: addition and merging of vertices

branching processes. These apply to cases in which vertices disappear by combining with other vertices, are generated from thin air, or split into “sub-vertices” such that old links from the original one are retained. One important mechanism, operative in cellular protein interaction networks, is the duplication of vertices (a protein) with slight changes to the copy’s connectivity from the parent vertex.

Figure 11 illustrates an example where the structure in the steady-state can be fruitfully described by similar mathematics as in other similar cases in networks. The basic rules are two-fold: i) a new vertex is added to the network, ii) two randomly chosen vertices merge. This kind of mechanisms are reminiscent of aggregation processes, of e.g. sticky particles that form colloidal aggregates in fluids. The simple model leads to a degree distribution of the asymptotic form  $P(k) \sim k^{-3/2}$ , reminiscent of the mean-field branching process result for the avalanche size distribution (where  $\tau = 3/2$ , cf. Sect. “Branching Processes”) [45,46]. The mathematical tool here is given by rate equations that describe the number of vertices with a certain degree,  $k$ . These are similar to those that one can write for the size of an avalanche in other contexts. It is an interesting question of what kind of generalizations one can find by considering cases where the edges have weights, and the elementary processes are made dependent in some way on those [47].

It is worth noting that the aggregation/branching process description of network structure develops easily deep complexity. This can be achieved by choosing the nodes to be joined with an eye to the graph geometry, and/or splitting vertices with some particular rules. In the latter case, a natural starting point is to maintain all the links of the original vertex and to distribute them among the descendant vertices. Then, one has to choose whether to link those to each other or not – either choice influences the local correlations. The same is naturally true, if one considers reactions of the type  $k + x' \rightarrow k''$ , which originate from joining two neighboring vertices with degrees  $k$  and

$x'$ . The likelihood of this rate is dependent on the conditional probability of a vertex with degree  $k$  having a neighbor with  $x'$ . Such processes are dependent on structural properties of the networks, which themselves change in time, thus the theoretical understanding is difficult due to the correlations. As a final note, in some cases the merging of vertices produces to the network a *giant component*, a vertex which has a finite fraction of the total mass (or, links in the whole). This is analogous to avalanching systems which are “weakly first-order”, in other words exhibit statistics which has a power-law, scale-free part and then a separate, singular peak. Often, this would be a final giant avalanche.

## Conclusions

In this short overview we have given some ideas of how to understand complexity via the tool of branching processes. The main issue has been that they are an excellent means of understanding “criticality” and “complexity” in many systems. We have concentrated on two particular applications, SOC and networks, to illustrate this. Many other important fields where BP-based ideas find use have been left out, from biology and the dynamics of species and molecules to geophysics and the spatial and temporal properties of say earthquakes. An example is the so-called ETAS model used for their modeling (see [48,49,50]).

## Future Directions

The applications of branching processes to complex systems continue on various fronts. One can predict interesting developments in a number of cases. First, as indicated the dynamics of BP’s on networks are not understood very well at all when it comes to the possible scenarios. A simple question is, whether the usual language of absorbing state phase transitions applies, and if not why – and what is the effect of the heterogeneous geometry of complex networks. In many cases the structure of a network is a dynamic entity, which can be described by generalized branching processes and there has been so far relatively little work into this direction.

The inclusion of spatial effects and temporal memory dynamics is another interesting and important future avenue. Essentially, one searches for complicated variants of usual BP’s to be able to model avalanching systems or cases where one wants to compute the typical time to reach an absorbing state, and the related distribution. Or, the question concerns the supercritical state ( $\lambda > \lambda_c$ ) and the spreading from a seed. As noted in the networks section this can be complicated by the presence of non-Poissonian temporal statistics. Another exciting future task is the de-



scription of non-Markovian phenomena, as when for instance the avalanche shape is non-symmetrical [51]. This indicates that there is an underlying mechanism which needs to be incorporated into the BP model.

## Acknowledgments

We are grateful to our colleague Stefano Zapperi with whom we have collaborated on topics related to networks, avalanches, and branching processes. This work was supported by the Academy of Finland through the Center of Excellence program (M.J.A.) and EUMETSAT's GRAS Satellite Application Facility (K.B.L.).

## Bibliography

### Primary Literature

- Harris TE (1989) *The Theory of Branching Processes*. Dover, New York
- Kimmel M, Axelrod DE (2002) *Branching Processes in Biology*. Springer, New York
- Athreya KB, Ney PE (2004) *Branching Processes*. Dover Publications, Inc., Mineola
- Haccou P, Jagers P, Vatutin VA (2005) *Branching Processes: Variation, Growth, and Extinction of Populations*. Cambridge University Press, Cambridge
- Manna SS (1991) J Phys A 24:L363. In this two-state model, the energy takes the two stable values,  $z_i = 0$  (empty) and  $z_i = 1$  (particle). When  $z_i \geq z_c$ , with  $z_c = 2$ , the site relaxes by distributing two particles to two randomly chosen neighbors
- Dickman R, Alava MJ, Munoz MA, Peltola J, Vespignani A, Zapperi S (2001) Phys Rev E 64:056104
- Bak P, Tang C, Wiesenfeld K (1987) Phys Rev Lett 59:381; (1988) Phys Rev A 38:364
- Tebaldi C, De Menech M, Stella AL (1999) Phys Rev Lett 83:3952
- Stella AL, De Menech M (2001) Physica A 295:1001
- Vespignani A, Dickman R, Munoz MA, Zapperi S (2000) Phys Rev E 62:4564
- Dickman R, Munoz MA, Vespignani A, Zapperi S (2000) Braz J Phys 30:27
- Alava M (2003) Self-Organized Criticality as a Phase Transition. In: Korutcheva E, Cuerno R (eds) *Advances in Condensed Matter and Statistical Physics*. arXiv:cond-mat/0307688,(2004) Nova Publishers, p 45
- Lubeck S (2004) Int J Mod Phys B18:3977
- Jensen HJ (1998) *Self-Organized Criticality*. Cambridge University Press, Cambridge
- Alava MJ, Nukala PKNN, Zapperi S (2006) Statistical models of fracture. Adv Phys 55:349–476
- Albert R, Barabási AL (2002) Statistical mechanics of complex networks. Rev Mod Phys 74:47
- Dorogovtsev SN, Mendes JFF (2003) *Evolution of Networks: From Biological Nets to the Internet and WWW*. Oxford University Press, Oxford; (2002) Adv Phys 51:1079; (2004) arXiv:cond-mat/0404593
- Bonachela JA, Chate H, Dornic I, Munoz MA (2007) Phys Rev Lett 98:115702
- Feller W (1971) *An Introduction to Probability Theory and its Applications*. vol 2, 2nd edn. Wiley, New York
- Asmussen S, Hering H (1983) *Branching Processes*. Birkhäuser, Boston
- Weiss GH (1994) *Aspects and Applications of the Random Walk*. North-Holland, Amsterdam
- Tang C, Bak P (1988) J Stat Phys 51:797
- Dhar D, Majumdar SN (1990) J Phys A 23:4333
- Janowsky SA, Laberge CA (1993) J Phys A 26:L973
- Flyvbjerg H, Sneppen K, Bak P (1993) Phys Rev Lett 71:4087; de Boer J, Derrida B, Flyvbjerg H, Jackson AD, Wettig T (1994) Phys Rev Lett 73:906
- Alstrøm P (1988) Phys Rev A 38:4905
- Christensen K, Olami Z (1993) Phys Rev E 48:3361
- García-Pelayo R (1994) Phys Rev E 49:4903
- Zapperi S, Lauritsen KB, Stanley HE (1995) Phys Rev Lett 75:4071
- Manna SS, Kiss LB, Kertész J (1990) J Stat Phys 61:923
- Tadić B, Nowak U, Usadel KD, Ramaswamy R, Padlewski S (1992) Phys Rev A 45:8536
- Tadic B, Ramaswamy R (1996) Phys Rev E 54:3157
- Vespignani A, Zapperi S, Pietronero L (1995) Phys Rev E 51:1711
- Lauritsen KB, Zapperi S, Stanley HE (1996) Phys Rev E 54:2483
- Newman MEJ (2003) The structure and function of complex networks. SIAM Rev 45:167
- Yook SH, Jeong H, Barabasi AL, Tu Y (2001) Phys Rev Lett 86:5835
- Barrat A, Barthelemy M, Pastor-Satorras R, Vespignani A (2004) Proc Natl Acad Sci USA 101:3747
- Barrat A, Barthelemy M, Vespignani A (2004) Phys Rev Lett 92:228701
- Pastor-Satorras R, Vespignani A (2001) Phys Rev Lett 86:3200
- Dorogovtsev SN, Goltsev AV, Mendes JFF (2007) arXiv:cond-mat/0750.0110
- Pastor-Satorras R, Vespignani A (2004) *Evolution and Structure of the Internet: A Statistical Physics Approach*. Cambridge University Press, Cambridge
- Vazquez A, Balazs R, Andras L, Barabasi AL (2007) Phys Rev Lett 98:158702
- Colizza V, Barrat A, Barthelemy M, Vespignani A (2006) PNAS 103:2015
- Vazquez A (2006) Phys Rev Lett 96:038702
- Kim BJ, Trusina A, Minnhagen P, Sneppen K (2005) Eur Phys J B43:369
- Alava MJ, Dorogovtsev SN (2005) Phys Rev E 71:036107
- Hui Z, Zi-You G, Gang Y, Wen-Xu W (2006) Chin Phys Lett 23:275
- Ogata Y (1988) J Am Stat Assoc 83:9
- Saichev A, Helmstetter A, Sornette D (2005) Pure Appl Geophys 162:1113
- Lippidello E, Godano C, de Arcangelis L (2007) Phys Rev Lett 98:098501
- Zapperi S, Castellano C, Colaioni F, Durin G (2005) Nat Phys 1:46

### Books and Reviews

- Albert R, Barabási AL (2002) Statistical mechanics of complex networks. Rev Mod Phys 74:47
- Asmussen S, Hering H (1983) *Branching Processes*. Birkhäuser, Boston

- Athreya KB, Ney PE (2004) *Branching Processes*. Dover Publications, Inc., Mineola
- Feller W (1971) *An Introduction to Probability Theory and its Applications*. vol 2, 2nd edn. Wiley, New York
- Haccou P, Jagers P, Vatutin VA (2005) *Branching Processes: Variation, Growth, and Extinction of Populations*. Cambridge University Press, Cambridge
- Harris TE (1989) *The Theory of Branching Processes*. Dover, New York
- Jensen HJ (1998) *Self-Organized Criticality*. Cambridge University Press, Cambridge
- Kimmel M, Axelrod DE (2002) *Branching Processes in Biology*. Springer, New York
- Newman MEJ (2003) The structure and function of complex networks. *SIAM Rev* 45:167
- Weiss GH (1994) *Aspects and Applications of the Random Walk*. North-Holland, Amsterdam

## Brittle Tectonics: A Non-linear Dynamical System

CHRISTOPHER H. SCHOLZ  
Lamont-Doherty Earth Observatory,  
Columbia University, New York, USA

### Article Outline

Glossary  
Definition of the Subject  
Introduction  
Scaling Relations: Self-similarity  
Earthquake and Fault Populations: Self-organization  
Models of the System  
Future Directions  
Bibliography

### Glossary

- Ductile shear zone** A quasi-planar tabular zone of localized shear deformation in the semi-brittle to fully plastic regimes.
- Earthquake** Dynamically running shear instability on a fault.
- Fault** A shear crack with friction between its interfaces.
- Mylonite** A metamorphic rock with a fabric produced by shear deformation.
- Suprafault** The shear relaxation structure that includes a fault and its associated ductile shear zone.

### Definition of the Subject

Brittle deformation is the primary mode of deformation of Earth's crust. At the long timescale it is manifested by

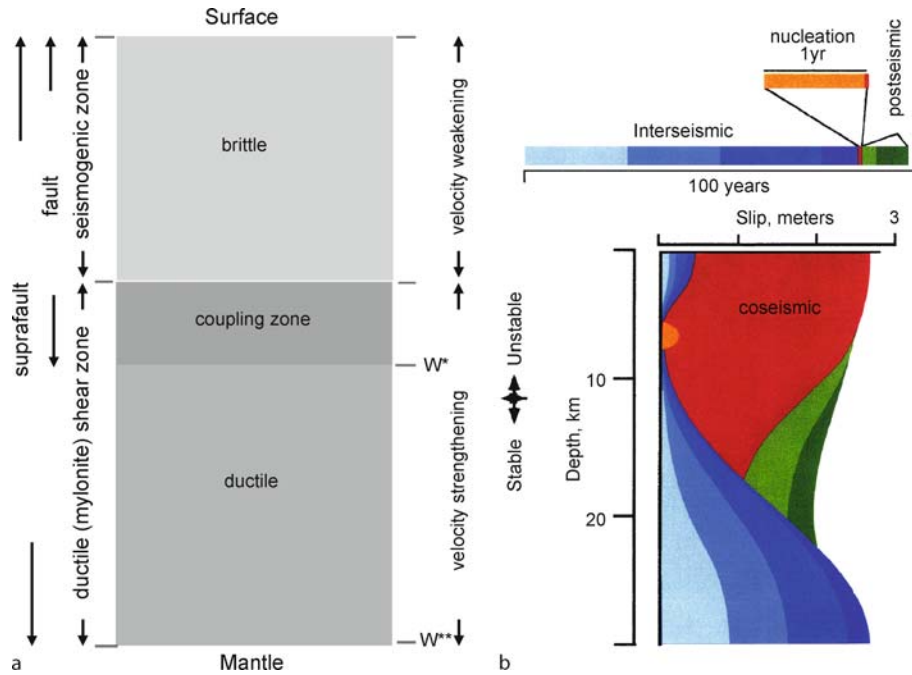
faulting, and on the short timescale by earthquakes. It is one of the best-known examples of a system exhibiting self-organized criticality. A full understanding of this system is essential to the evaluation of earthquake hazard.

### Introduction

The upper part of Earth's crust is brittle and under a state of all-round compression. It responds to deformation by faulting: the formation and propagation of shear cracks. The crack walls support normal stresses and hence fault propagation must overcome not only the rupture resistance of the fault tips but friction between its interior interfaces. This friction is usually velocity weakening, such that any slippage results in stick-slip instability. The resulting dynamically running crack-like shear instability radiates elastic waves, producing the shaking known as an earthquake. Thus brittle tectonics separates into two greatly different timescales: the long timescale of the faulting and the short timescale of the earthquakes. Faults propagate quasi-statically at rates set by tectonic driving rates, whereas earthquakes propagate at a rate set by the shear velocity of the rock. Almost all displacement on continental faults occurs by earthquakes, so that there is a coupling between the two timescales.

This system is illustrated in Fig. 1a. The term *fault* is usually restricted to the rupture of the brittle part, but large faults continue downwards as ductile shear zones, which, in the case of major faults, may extend as deep as the crust-mantle boundary. We will refer to the entire shear release structure, including both the fault and its corresponding ductile shear zone as the *suprafault*.

The division between brittle and ductile behavior occurs at what is usually called the brittle-ductile transition. This transition is gradual, with a change in deformation mechanism from brittle crack propagation and friction to fully plastic deformation through a region of mixed behavior known as semi-brittle. For our purposes, however, it is more useful to divide the suprafault into two regimes, of velocity weakening and velocity strengthening rheologies, at a depth  $S$ . These regions are predicted by a rate/state variable friction law that has been established in the laboratory (for a review, see [25]). Here, we also refer to the ductile regime as velocity strengthening because the creep law for velocity strengthening friction is the same form as for ductile creep.  $S$  corresponds to the upper onset of the brittle-ductile transition and it varies regionally because it depends on temperature (heat flow) and strain rate. For faults in continental crust it varies between about 10–20 kms. Earthquakes can only nucleate above  $S$ , so this constitutes the *seismogenic zone*. When earthquakes be-



**Brittle Tectonics: A Non-linear Dynamical System, Figure 1**

**a** A schematic longitudinal section of the rheological components of a crustal suprafault. **b** A simulation of a model with the properties of (a), shown throughout a seismic cycle (after Tse and Rice [37])

come large, they can, owing to the very high strain rates at their rupture tips, propagate below  $S$  to  $W^*$ , which defines the maximum width of earthquakes. Similarly, the maximum width of a suprafault is  $W^{**}$ .

Although the ductile region is velocity (strain-rate) strengthening, it is also, seemingly paradoxically, strain softening, which leads to strain localization into narrow shear zones. As the rock shears, it gradually reorganizes internally to form a mylonitic fabric, which facilitates further shearing. With finite strain, the mylonitic fabric develops further and the mylonitic foliation rotates closer to the plane of shear, progressively reducing the rock's resistance to shear in that particular direction and plane. It is this mechanism of *fabric weakening* that produces the shear localization.

The system has two characteristic length scales,  $W^*$  and  $W^{**}$ . An earthquake nucleates within the seismogenic zone and initially propagates in all directions along its perimeter, acting as a 3D crack. When its dimension exceeds  $W^*$ , it has breached the free surface and penetrated to  $W^*$  and then is prohibited from growing deeper. It is then restricted to propagating in the horizontal direction, acting as a 2D crack. Thus there is a symmetry breakage at the dimension  $W^*$ . *Small* earthquakes, with dimensions smaller than  $W^*$ , are not self-similar with *large* earth-

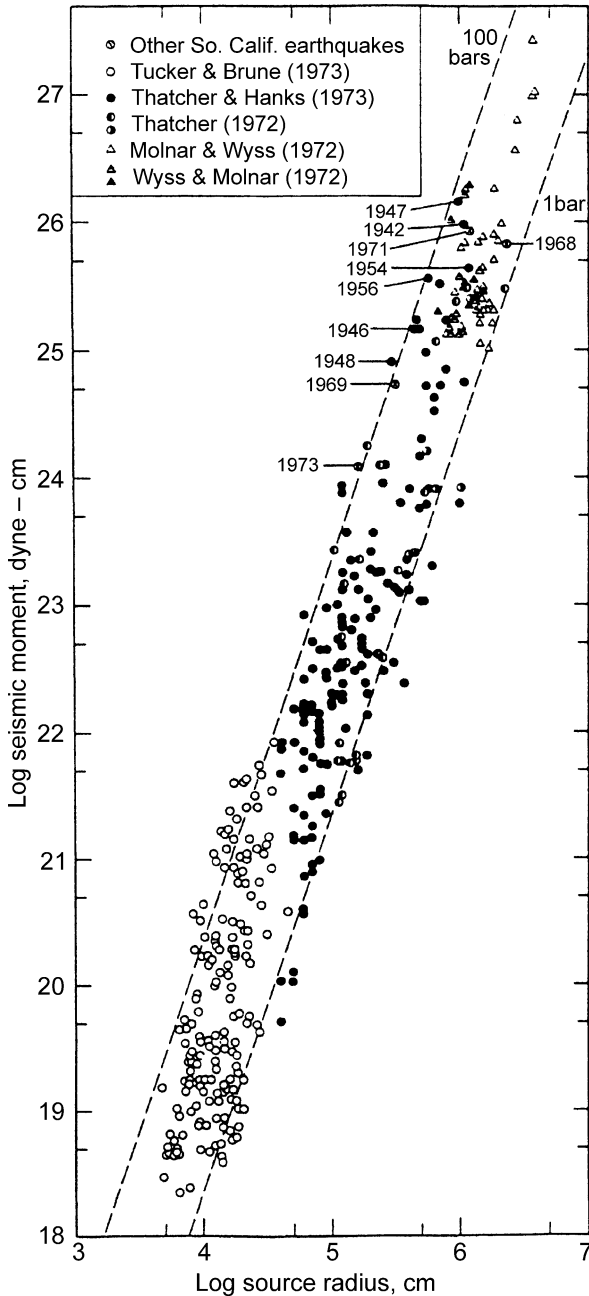
quakes, those with lengths larger than  $W^*$ . The same occurs for suprafaults at the dimension  $W^{**}$ .

The coupling of the system is illustrated in Fig. 1b. This simulation shows the deformation of a strike slip suprafault (one in which the slip is horizontal), over the seismic cycle, i.e. the recurrence time of a large earthquake. The recurrence time depends on the long-term (geologic) slip rate of the suprafault. During the *interseismic* period, the ductile region shears at a constant rate and gradually propagates upwards. This loads the upper region until nucleation occurs above  $S$  as a quasi-static crack that eventually goes dynamic, resulting in an earthquake that propagates over  $S$  and down to  $W^*$ . The ductile region is instantaneously reloaded by the downward propagation of the earthquake. This results in *postseismic* relaxation that exponentially decays over a period of years to decades.

## Scaling Relations: Self-similarity

### Earthquakes

The primary scaling relation for cracks is that between the shear displacement (slip),  $D$ , for faults or  $\Delta D$  for earthquakes, and the dimension  $L$  of the rupture. The ratio of these parameters is a stress-drop  $\Delta\sigma$  normalized by an



**Brittle Tectonics: A Non-linear Dynamical System, Figure 2**  
 A collection of data for small earthquakes showing the relationship between seismic moment  $M_0 = \mu \Delta D A$  and source radius. Dashed lines are of constant stress drop. From Hanks [15]

elastic constant. Figure 2 shows data for small earthquakes, in which seismic moment  $M_0 = \mu \Delta D A$  is plotted vs.  $L$  ( $\mu$  is the shear modulus and  $A$  is rupture area). Because these are 3D cracks the  $L^3$  lines are loci of equal stress-drop. Although there is considerable scatter, the data follow this

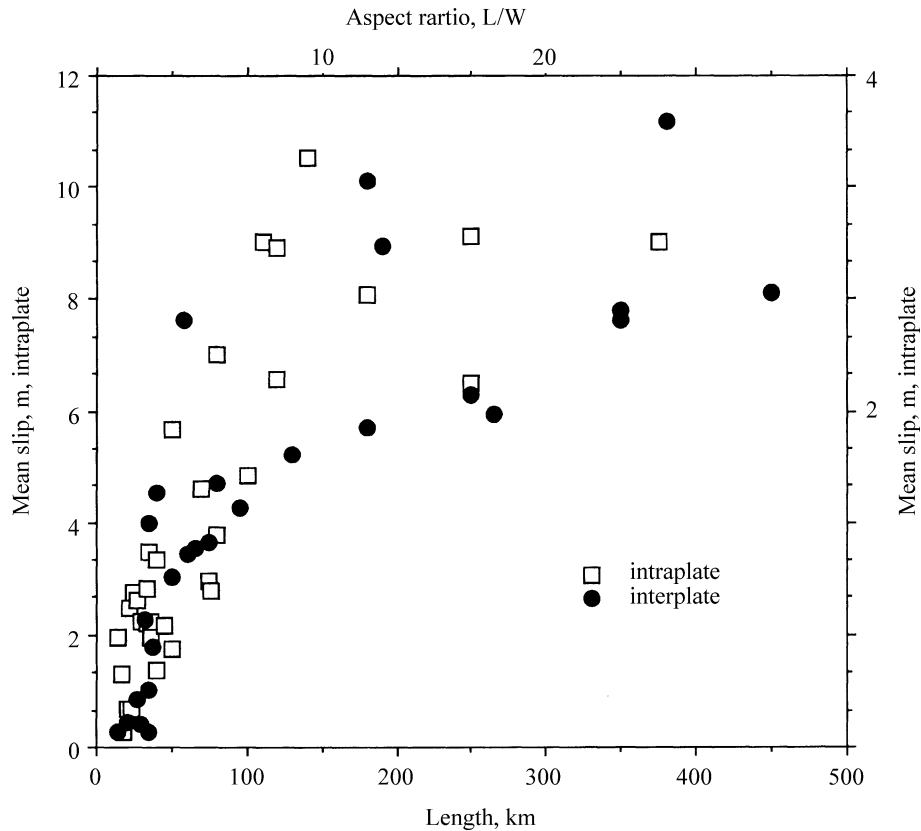
trend. Thus, while  $\Delta \sigma$  varies by about 2 orders of magnitude among earthquakes, it is scale-invariant. These earthquakes are thus self-similar.

A plot of slip vs. length for large earthquakes is shown in Fig. 3. What is seen is that initially  $\Delta D$  increases linearly with  $L$  and, following a long crossover, becomes constant. From the static analysis of a 2D crack, we would expect, for constant stress-drop, that  $\Delta D$  be proportional to  $W^*$ , and hence constant on a plot like this, but this is seen only for the rare earthquakes with aspect ratios greater than 10. A static scaling analysis suggests that the crossover be complete by an aspect ratio of 4. A numerical model of dynamic shear cracking produced the same delayed crossover [30]. Further study of the same model suggests that the long rollover results from significant slip below the seismogenic depth [31]. The scaling of large earthquakes thus exhibit two self-similar regimes, neither of which is self-similar with small earthquakes.

The two  $\Delta D$  scales in Fig. 3, for interplate and intraplate earthquakes, illustrate the rate effect on stress drop. Earthquakes on intraplate faults, for which geological sliding rates are several orders of magnitude less than for interplate faults, have, on average, stress drops about three times larger than interplate earthquakes. Such fault healing is expected from the rate/state friction constitutive law [4].

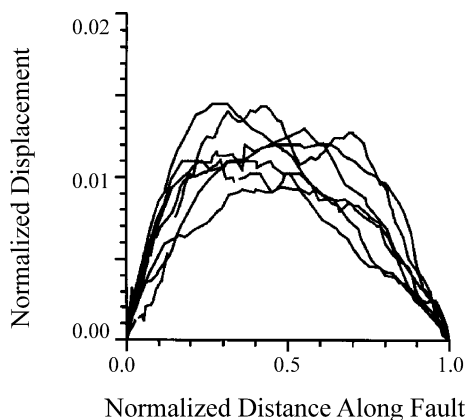
### Faults

The displacement profiles for some faults in the same rock type, ranging in length from 690 to 2200 m, are shown, normalized by length, in Fig. 4. The data collapse shows the self-similarity of these profiles. Displacement amplitude scales linearly with length, whereas the linear displacement tapers as the tips are approached are scale invariant. A global plot of  $D$  vs.  $L$  is shown in Fig. 5. There is, similar to earthquakes, considerable scatter, but the overall trend shows a linear relationship between  $D$  and  $L$ . Here more can be said about the origin of the scatter than in the case of earthquakes. Those faults with  $D/L$  ratios less than  $10^{-2}$  are in soft sedimentary rock at shallow depth; those with greater  $D/L$  are in strong metamorphic or igneous rock. Faults of length greater than 5 km extend increasingly deeper into the crust and exhibit an increase in  $D/L$  (stress-drop) owing to the increase of strength with pressure. Scatter is also produced by the interaction of faults through their stress fields, which distorts their displacement profiles (see pp. 126–129 in [27]). These effects of lithology and fault interaction also occur for the displacement tapers near the fault tips, which are another measure of strength [28]. In a study in which the faults are in a sin-



Brittle Tectonics: A Non-linear Dynamical System, Figure 3

Mean slip vs. rupture length for large earthquakes. Two scaling regimes are observed. Slip increases approximately linearly with  $L$  up to an aspect ratio of about 10, and is independent of length thereafter. After Scholz [22]



Brittle Tectonics: A Non-linear Dynamical System, Figure 4

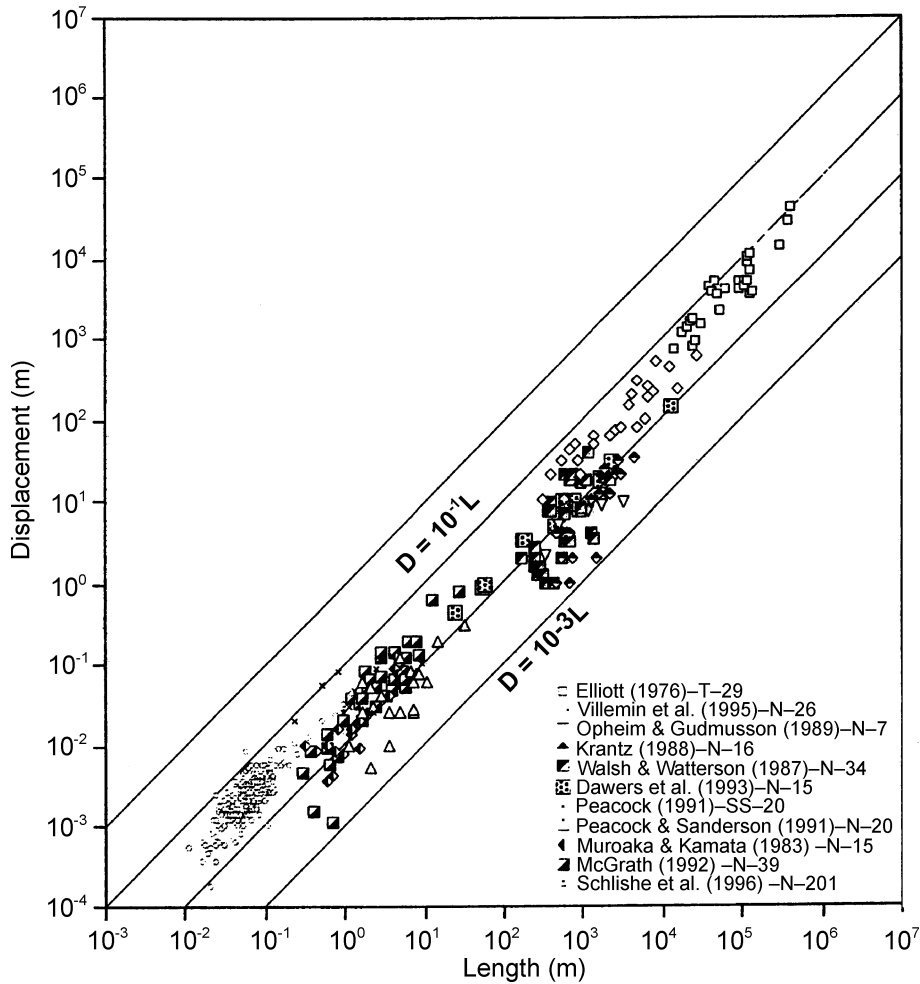
Displacement profiles along faults of different length in the Volcanic Tablelands of eastern California, normalized to length. Data from Dawers et al. [13]

gle lithology and were selected to be isolated [13], the scatter in both  $D/L$  and fault tip taper was correspondingly reduced (as in Fig. 4).

We do not have data for long enough faults to observe a crossover from 3D to 2D scaling, as we did for earthquakes. Because  $W^{**} \sim 40$  km, we would need data from faults of length exceeding 400 km to clearly see this effect. Although faults of that length and longer exist, they are predominately plate boundaries, which have no tips because their ends connect to other plate boundaries and are stress relieved by them. Therefore the scaling of Fig. 5 does not apply to them. The San Andreas fault in California, for example, terminates at its southern end at a spreading ridge. Its net slip is more than an order of magnitude greater than would be expected from simple application of the scaling shown in Fig. 5.

Faults, like earthquakes, are self-similar. The scaling parameters relating  $D$  and  $L$  differ greatly, however. For faults it is  $\sim 10^{-2}$  and for earthquakes,  $10^{-4}$ – $10^{-5}$ . Both represent stress-drops. For faults, it is the stress-drop from the fracture strength of intact rock to the residual frictional





Brittle Tectonics: A Non-linear Dynamical System, Figure 5

Maximum slip vs. length for fault data sets from various tectonic settings. Modified from Schlische et al. [20]

strength. For earthquakes, it is the stress-drop from static to dynamic friction.

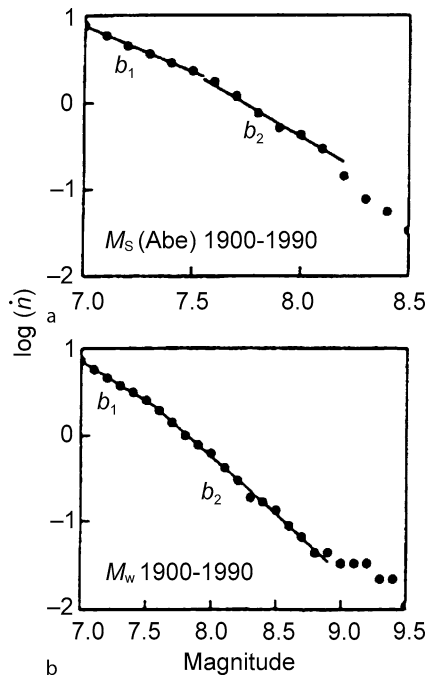
### Earthquake and Fault Populations: Self-organization

Earthquakes and faults both self-organize into populations with well defined statistical properties, thus exhibiting complexity in the physics of this system.

#### Earthquakes

The frequency  $N(M)$  of earthquakes greater than magnitude  $M$  has long been known to obey the Gutenberg–Richter law,  $\log N(M) = a - bM$  in which  $b \approx 1$  is found universally for small earthquakes. In laboratory experiments  $b$  is found to decrease with stress [21]. The fact that  $b$  is narrowly limited in the Earth's crust may be be-

cause stress in the crust is limited to a narrow range by the frictional strength of faults [36].  $M$  is a logarithmic measure of energy  $E$  or seismic moment,  $M_0$ . In terms of  $M_0$  this relation is  $N(M_0) = aM_0^{-B}$  where  $B = 2/3b$ . However, as shown in Fig. 6, in the global catalog of earthquakes there is a crossover from  $b = 1$  to  $b = 1.5$  at about  $M7.5$ . This corresponds to a rupture dimension of about 50 km, which is an average value of  $W^*$  for subduction zones, which dominate this data set. For a subset of continental earthquakes, the same crossover is observed at  $M6.5$ , corresponding to a rupture length of 10–20 km, similar to  $W^*$  for crustal faults. Thus in terms of a linear measure of earthquake size like  $E$  or  $M_0$ , earthquakes have a power law size distribution. Where there is a breakdown of self-similarity between small and large earthquakes there is a corresponding change in the exponent  $B$  in this distribution.



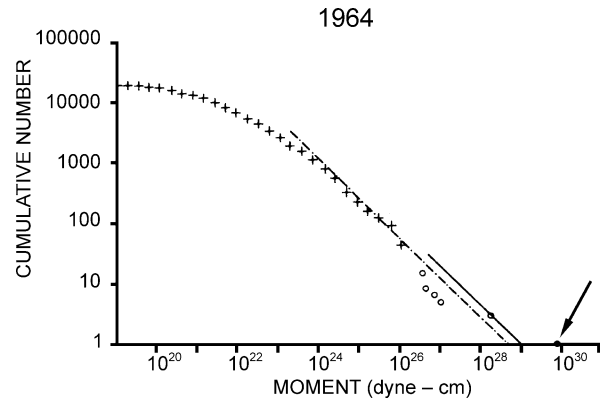
**Brittle Tectonics: A Non-linear Dynamical System, Figure 6**

The frequency size distribution from a global cataloger of Pacheco and Sykes [18] for two magnitude scales. From Pacheco et al. [19]

The size distribution discussed above applies to regions large enough to include many active faults or plate boundary segments. If we study instead a single fault or plate boundary segment that ruptures in a single large earthquake and consider earthquakes that occur within a single seismic cycle of that earthquake, a different picture emerges. An example is shown in Fig. 7. There we see that the small earthquakes during that interval obey a power law with  $B = 2/3$ , as usual, but that the single large earthquake is more than an order of magnitude greater than expected from extrapolation of the small events. This is another manifestation of large and small earthquakes belonging to different fractal sets.

## Faults

Fault data are usually obtained by geological mapping. Because mapping has a dynamic range limited to about one order of magnitude, it is difficult to convincingly demonstrate a power law size distribution with this method. When the results of mapping at several scales are combined [29], the case for a power law is strengthened, but the exponent is poorly determined. This problem was overcome with the data in Fig. 8a, from a Magellan SAR image of normal faults on the plains of Venus. This image has



**Brittle Tectonics: A Non-linear Dynamical System, Figure 7**

Distribution of small earthquakes in the rupture zone of the 1964 Alaskan  $M9.4$  earthquake, normalized to the recurrence time of that earthquake. An arrow indicates the 1964 earthquake. It is about  $1\frac{1}{2}$  times larger than an extrapolation of the small earthquakes would indicate. The roll off at  $M_0 < 3 \times 10^{23}$  dyne cm is caused by a lack of perceptibility of smaller events. From Davison and Scholz [11]

a far greater dynamic range, and unequivocally demonstrates a power law distribution, with an exponent  $-2$ . This is a frequency plot, rather than a cumulative plot such as shown in Fig. 7.

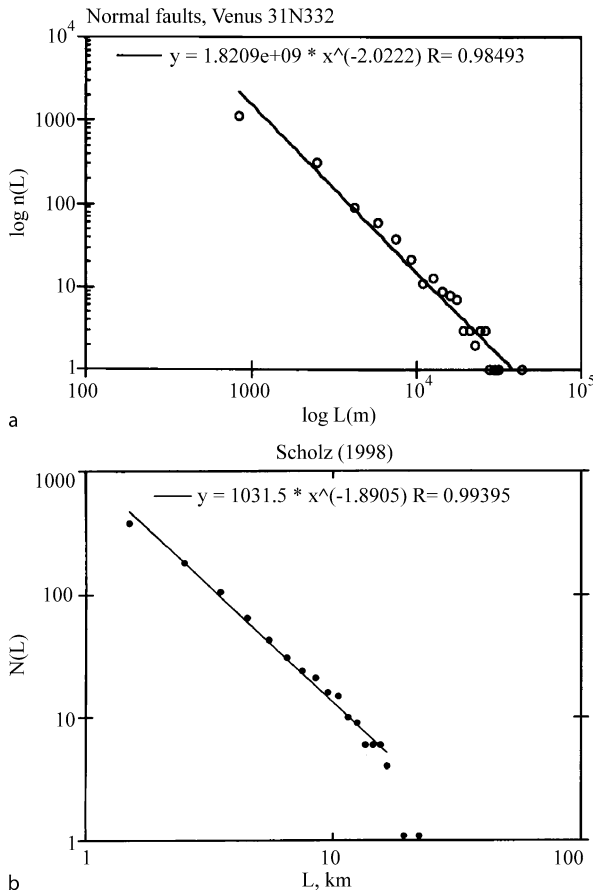
On close examination, faults are found to be not a continuous surface, but a complex of myriad strands or subfaults. A cumulative size distribution of the subfaults of the San Andreas fault is shown in Fig. 8b, where they are also shown to obey a power law. The upper fractal limit of the subfaults appears to be  $W^*$ .

The size distributions of earthquakes and faults are not unrelated. If one equates faults with large earthquakes and subfaults with small earthquakes, then by convolving over the fault distributions of Fig. 8 one can obtain the earthquake distributions of Figs. 6 and 7 [24,26].

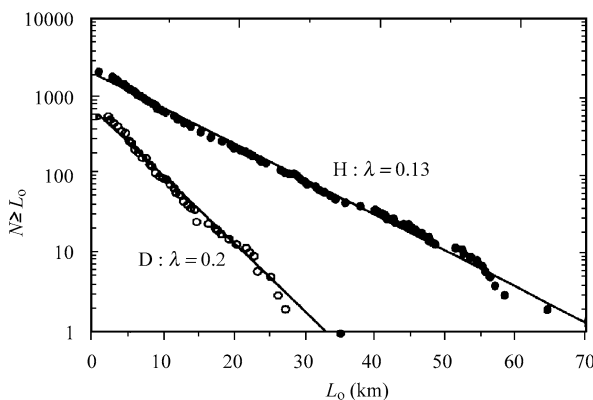
Not all fault distributions are power law, however. Figure 9 shows the size distribution of faults on the flanks of the East Pacific Rise, obtained by sonar swath mapping. These clearly obey an exponential rather than a power law distribution. One can also find examples of periodic distributions, such as the regularly spaced faults of the Basin and Range province of the Western US. We show later that three regimes, characterized by power law, exponential, and periodic distributions, evolve with increasing strain.

## Models of the System

The first model of the earthquake system was a bench-top apparatus of a string of weights connected by springs and pulled along a frictional surface [6]. Their main finding



**Brittle Tectonics: A Non-linear Dynamical System, Figure 8**  
Power law size distributions of faults. **a** Frequency-length distribution for faults on the plains of Venus from a Magellan SAR image [23]. **b** Cumulative length distribution of subfaults of the San Andreas fault [26]



**Brittle Tectonics: A Non-linear Dynamical System, Figure 9**  
The distribution of faults at two places on the flanks of mid-ocean ridges. These distributions are exponential, rather than power law. From Cowie et al. [9]

was that this model obeyed the Gutenberg–Richter law. Interest in this result was revived by the paper of Bak, Tang, and Wiesenfeld [3], which introduced the concept of *self-organized criticality*.

Bak et al. [3] illustrated this concept with a cellular automata model of a sandpile, in which the cells are governed by a nearest neighbor rule for avalanches and the system is driven by the slow random addition of grains to the pile. The result was a power law distribution of avalanches and a  $1/f$  energy flux. From this they hypothesized that systems containing many interacting elements may exhibit general statistical properties that are described by power laws. They called this behavior self-organized criticality (SOC). Although this concept is intuitively appealing, there is no clear-cut definition of what SOC is.

Because the Gutenberg–Richter law is the most famous of such power laws, adaptations of this cellular automata model to earthquakes soon followed (e.g. [2,5]). Carlson and Langer [7] explored a numerical spring-slider model that contained the full dynamics of the slider blocks. All these models reproduced the Gutenberg–Richter law and also explored other aspects of the system, for example, in the case of Brown et al. [5], the distribution of recurrence times of large earthquakes. Olami et al. [17] explored a non-conservative spring block model, and Christensen and Olami [8] showed how Omori's law of after-shocks might be explained without having to appeal to an added viscosity as did Burridge and Knopoff. Sornette and Virieux [32], using a mean field theory approach, showed that the scaling difference between small and large earthquakes results in the change in  $b$  value shown in Fig. 6. Extended discussion of these models and their results can be found in Turcotte [38] and Jensen [16].

The study of the self-organization of faults was done first in an analog model of continental deformation by Davy et al. [12], who found a geometrically fractal fault distribution. Cowie et al. [10] studied a model simulating a brittle sheet undergoing stretching. For large strains this model converges on a single system size crack, but at small strains power law size crack distributions were observed.

Spyropoulos et al. [34] expanded on this with a spring-block model of the stretching of a brittle sheet supported by a ductile substrate. They found an evolution of the crack population with strain, from power law to exponential distributions and finally to a periodic distribution of system size cracks. In this model, the brittle layer contains only nearest neighbor interactions but the ductile substrate provides longer-range force interactions. It is the longer-range interactions that result in the evolution of the system and is what differs this model from the sandpile family of models. The evolution from a power law

distribution to an exponential distribution of cracks has been demonstrated with physical models [1,33]. It has also been demonstrated in the field, for faults in the Asal rift of Ethiopia [14].

### Future Directions

Our understanding of brittle tectonics as a dynamical system is still very sketchy. Approaches to it tend to be piecewise, which leads to self-limiting results. For example, although it is now well known that earthquakes interact through their stress fields (e.g. [35]), this phenomenon is modeled by assuming that the background 'tectonic' stress field is uniform, thus ignoring the stress fields of faults, which also interact. The interaction of faults has so far been studied in only the simplest case, that of sub-parallel normal faults. For strike-slip faults, in which the fault tip stresses are asymmetric, more interesting possibilities exist, as yet unexplored. The San Andreas fault system in California is a complex of faults several hundred km wide that somehow act together to accommodate most of the Pacific-North America plate motion. How does this work? How do the faults interact in such a way to preserve continuity of plate motion along strike, as geodetic data indicate? How does the interaction between faults lead to the geometric evolution of fault patterns? Are fault slip rates stationary, as often assumed, or do they vary at intermediate time scales between that of earthquakes and that of the faults themselves? If so, can this explain the temporal clustering of earthquakes that seem to occur over spatial distances long compared to the stress fields of earthquakes? These are but a few of the problems one can imagine studying in the future, once the idea is accepted that the system must be considered in its entirety before being broken into digestible chunks.

### Bibliography

#### Primary Literature

- Ackermann RV, Schlische RW, Withjack MO (2001) The geometric and statistical evolution of normal fault systems: an experimental study of the effects of mechanical layer thickness on scaling laws. *J Struct Geol* 23:1803–1819
- Bak P, Tang C (1989) Earthquakes as a self-organized critical phenomenon. *J Geophys Res* 94:15635–15637
- Bak P, Tang C, Wiesenfeld K (1987) Self-organized criticality: An explanation of  $1/f$  noise. *Phys Rev Lett* 59:381–384
- Beeler NM, Hickman SH, Wong TF (2001) Earthquake stress drop and laboratory-inferred interseismic strength recovery. *J Geophys Res-Solid Earth* 106:30701–30713
- Brown SR, Scholz CH, Rundle JB (1991) A simplified spring-block model of earthquakes. *Geophys Res Lett* 18:215–218
- Burridge R, Knopoff L (1967) Model and theoretical seismicity. *Bull Seism Soc Am* 57:341–362
- Carlson JM, Langer JS (1989) Properties of earthquakes generated by fault dynamics. *Phys Rev Lett* 62:2632–2635
- Christensen K, Olami Z (1992) Variation of the Gutenberg-Richter B Values and Nontrivial Temporal Correlations in a Spring-Block Model for Earthquakes. *J Geophys Res-Solid Earth* 97:8729–8735
- Cowie PA, Scholz CH, Edwards M, Malinverno A (1993) Fault strain and seismic coupling on midocean ridges. *J Geophys Res-Solid Earth* 98:17911–17920
- Cowie PA, Sornette D, Vanneste C (1995) Multifractal scaling properties of a growing fault population. *Geophys J Int* 122:457–469
- Davison F, Scholz C (1985) Frequency-moment distribution of earthquakes in the Aleutian Arc: A test of the characteristic earthquake model. *Bull Seismol Soc Am* 75:1349–1362
- Davy P, Sornette A, Sornette D (1990) Some consequences of a proposed fractal nature of continental faulting. *Nature* 348:56–58
- Dawers NH, Anders MH, Scholz CH (1993) Growth of normal faults – displacement-length scaling. *Geology* 21:1107–1110
- Gupta A, Scholz CH (2000) Brittle strain regime transition in the Afar depression: implications for fault growth and seafloor spreading. *Geology* 28:1087–1090
- Hanks TC (1977) Earthquake Stress Drops, Ambient Tectonic Stresses and Stresses That Drive Plate Motions. *Pure Appl Geophys* 115:441–458
- Jensen HJ (1998) Self-organized criticality: Emergent complex behavior in physical and biological systems. Cambridge Univ. Press, Cambridge
- Olami Z, Feder HJS, Christensen K (1992) Self-organized criticality in a continuous, nonconservative cellular automaton modeling earthquakes. *Phys Rev Lett* 68:1244–1247
- Pacheco JF, Sykes LR (1992) Seismic moment catalog of large shallow earthquakes, 1900 to 1989. *Bull Seismol Soc Am* 82:1306–1349
- Pacheco JF, Scholz CH, Sykes LR (1992) Changes in frequency-size relationship from small to large earthquakes. *Nature* 355:71–73
- Schlische RW, Young SS, Ackermann RV, Gupta A (1996) Geometry and scaling relations of a population of very small rift-related normal faults. *Geology* 24:683–686
- Scholz CH (1968) The frequency-magnitude relation of microfracturing in rock and its relation to earthquakes. *Bull Seismol Soc Am* 58:399–415
- Scholz CH (1994) A reappraisal of large earthquake scaling. *Bull Seismol Soc Am* 84:215–218
- Scholz CH (1997) Earthquake and fault populations and the calculation of brittle strain. *Geowissenschaften* 3–4:124–130
- Scholz CH (1997) Size distributions for large and small earthquakes. *Bull Seismol Soc Am* 87:1074–1077
- Scholz CH (1998) Earthquakes and friction laws. *Nature* 391:37–42
- Scholz CH (1998) A further note on earthquake size distributions. *Bull Seismol Soc Am* 88:1325–1326
- Scholz CH (2002) The mechanics of earthquakes and faulting, 2nd edn. Cambridge University Press, Cambridge
- Scholz CH, Lawler TM (2004) Slip tapers at the tips of faults and earthquake ruptures. *Geophys Res Lett* 31:L21609, doi:10.1029/2004GL021030

29. Scholz CH, Dawers NH, Yu JZ, Anders MH (1993) Fault growth and fault scaling laws – preliminary-results. *J Geophys Res-Solid Earth* 98:21951–21961
30. Shaw BE, Scholz CH (2001) Slip-length scaling in large earthquakes: observations and theory and implications for earthquake physics. *Geophys Res Lett* 28:2995–2998
31. Shaw BE, Wesnouski SG (2008) Slip-length Scaling in large earthquakes: The role of deep penetrating slip below the seismogenic layer. *Bull Seismol Soc Am* 98:1633–1641
32. Sornette D, Virieux J (1992) Linking short-timescale deformation to long-timescale tectonics. *Nature* 357:401–403
33. Spyropoulos C, Griffith WJ, Scholz CH, Shaw BE (1999) Experimental evidence for different strain regimes of crack populations in a clay model. *Geophys Res Lett* 26:1081–1084
34. Spyropoulos C, Scholz CH, Shaw BE (2002) Transition regimes for growing crack populations. *Phys Rev E* 65:056105, doi:10.1103/PhysRevE.65.056105
35. Stein RS (1999) The role of stress transfer in earthquake occurrence. *Nature* 402:605–609
36. Townend J, Zoback MD (2000) How faulting keeps the crust strong. *Geology* 28:399–402
37. Tse S, Rice J (1986) Crustal earthquake instability in relation to the depth variation of frictional slip properties. *J Geophys Res* 91:9452–9472
38. Turcotte DL (1999) Seismicity and self-organized criticality. *Phys Earth Planet Inter* 111:275–293

## Books and Reviews

- Sornette D (2003) *Critical phenomena in natural systems: Chaos, fractals, self-organization, and disorder*. Springer, Berlin
- Turcotte DL (1997) *Fractals and chaos in geology and geophysics*. Cambridge, New York

## Business Policy and Strategy, System Dynamics Applications to

JAMES M. LYNEIS

Worcester Polytechnic Institute, Worcester, USA

## Article Outline

Glossary

Definition of the Subject

Introduction

Using System Dynamics Models in Policy  
and Strategy Formulation and Implementation

Theory Development – Understanding the Drivers  
of Business Dynamics

Applications and Case Examples

Future Directions

Bibliography

## Glossary

**Business policy and strategy** A firm's business strategy defines how and where it competes, and its approach to doing so. A business strategy typically specifies a firm's goals, the products and services offered and the markets served, and the basis for competing (price, service, quality, etc.). A strategy may also define the organization structure, systems and policies which implement the strategy. In addition, firm's will have systems and policies which focus on operations and functions, and are not truly "strategic" in nature. Nevertheless, these operational policies can be important in determining business performance.

**Business dynamics** Business dynamics is the study of how the structure of a business (or a part of the business), the policies it follows, and its interactions with the outside world (customers, competitors, suppliers) determine its performance over time. Business structure consists of feedback loops surrounding the stocks and flows of resources, customers, and competitive factors that cause change over time; business policies are important components of these feedback loops. Business dynamics is a means of determining the likely performance that will result from alternative business policies and strategies.

## Definition of the Subject

System dynamics has long been applied to problems of business performance. These applications range from operational/functional performance to overall strategic performance. Beginning with its founding at MIT's Sloan School of Management in 1957, an important focus of research, teaching, and application has been on understanding why companies and markets exhibit cycles, or underperform competitors in terms of growth or profitability. The original publication in the field was Forrester's *Industrial Dynamics* [26], which not only laid the theoretical foundations for the field, but also provided an understanding of the causes of instability in supply chains. Since that initial work, research and application has been widespread. It has addressed the dynamics underlying instability in manufacturing and service organizations, the processes which encourage or inhibit growth, the dynamics of research organizations, and the causes of cost and schedule overruns on individual projects. It has been applied in many industries, from manufacturing to high-tech to financial services and utilities, both by academics and consultants. Business theory and applications are taught at many universities, including but not limited to MIT, London Business School and others in England, Bergen (Nor-



way), Mannheim and Stuttgart (Germany) (see [62,72] for more details). Business policy and strategy has and will continue to be one of the major application areas for system dynamics.

## Introduction

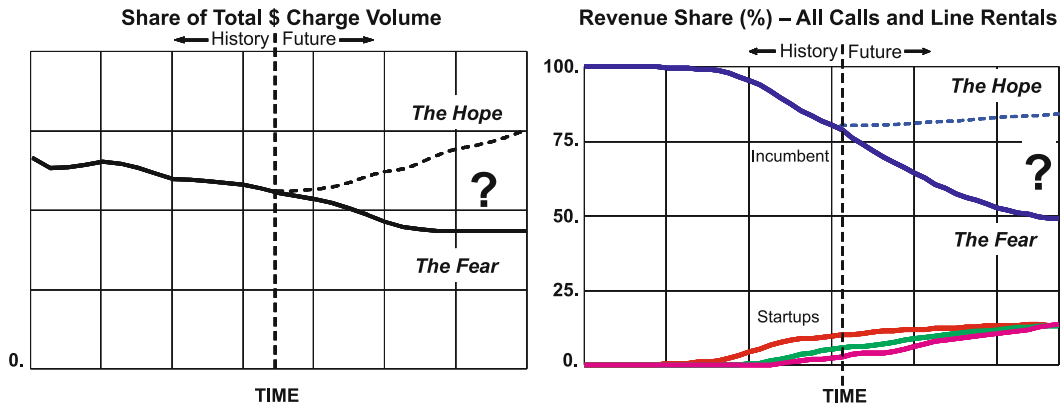
Business strategy, sometimes called simply ‘policy’ or ‘strategy’, is primarily concerned with how and where firm’s choose to compete. It includes such decisions as setting goals, selecting which products and services to offer in which markets, establishing the basis for competing (price, service, quality, etc.), determining the organization structure, systems and policies to accomplish the strategy, and designing policies for steering that strategy continually into the future. Academic and applied research on business strategy developed separately from system dynamics. That research, while widely disparate, has largely focused on static assessments and tools. For example, cross-sectional studies of many companies attempt to identify key differences that determine success or failure as a guide to management; “strategic frameworks” (for example, learning curves, growth share matrices, Porter’s five forces [79,80]) assist managers in framing strategy and intuitively assessing performance over time; scenario planning helps managers visualize alternative futures; the resource-based view of the firm and core competencies [81] help managers identify how resources and capabilities determine the best way to compete. While these tools provide valuable insights and frameworks, they leave the connection between a firm’s business strategy and the evolution of its performance over time to the intuition of managers – while traditional business strategy addresses the starting point and the desired end point, and the mechanisms that might allow the firm to transition between the two, the ability of those mechanisms to achieve that transition, and the path for getting between the two, is left unanswered.

Academic and applied research on operational and functional performance has similarly developed separately from system dynamics. Although it is difficult to generalize, this research is again typically static in nature and/or focused on the detailed management of a part of the organization over a relatively short period of time (for example, optimization of production scheduling during a month, quarter or year; optimal inventory management during a quarter or year). While this detailed management is necessary for running a business, it often overlooks the longer run implications of the policies established to manage the business in the short run, and of the impacts of one part of the business on other parts.

In contrast, system dynamics addresses how structure (feedback loops, stocks and flows) and policies determine performance over time – how does the firm, or a part of the firm, get from its current state to some future state. Evolving from this structural theory, system dynamicists have studied why firms and industries exhibit instability and cycles, and why firms grow or decline. Two real examples of problematic behavior over time are shown in Fig. 1. The example on the left shows the pattern of orders for commercial jet aircraft – a system dynamicist would try to understand why the orders are cyclical, and what can be done to make them less so (or to take advantage of the cycles by forecasting that cyclicity); the example on the right shows market shares of major players in a recently deregulated telecom market – a system dynamicist, working for the incumbent telecom, would try to understand the causes of market share loss and what can be done to reverse that loss.

From its beginnings in the late 1950s, system dynamics has been used to progressively develop structural theories to explain instability in supply chains [26], cycles of growth [28], boom and bust in product sales [100], and cost and schedule overrun on projects [54], to mention just a few. While system dynamics has at times borrowed, or in some cases reinvented, concepts from business policy and strategy, this structural theory development has until recently evolved largely independently of traditional business research and practice. There is, however, a great deal of potential synergy between traditional business research, particularly strategy research, and system dynamics that is increasingly being exploited.

This paper surveys the application of system dynamics to business policy and strategy. The next section discusses the role of system dynamics models in policy and strategy formulation and implementation. Topics include how system dynamics fits into the typical policy and strategy formulation process, how system dynamics offers synergies with more traditional approaches, and how to conduct a modeling effort in order to enhance the implementation of any changes in policy or strategy. In Sect. [Theory Development – Understanding the Drivers of Business Dynamics](#), system dynamics contribution to theory development is discussed – what are the structures underlying common business problems, and how can performance be improved? For example, what creates instability in supply chains, or “boom and bust” in product sales, and how can these behaviors be changed? Finally, in Sect. [Applications and Case Examples](#) applications to real world situations are presented – case studies that illustrate the value and impact of using system dynamics for policy and strategy development and implementation in specific firms and in-



Business Policy and Strategy, System Dynamics Applications to, Figure 1  
Examples of problematic behavior over time

dustries. As there has been a substantial amount of work done in this area, I must be selective, trying to touch on major themes and a representative sampling of work. Inevitably this will reflect my personal experiences, and my apologies to others that I have omitted either intentionally or unintentionally.

### Using System Dynamics Models in Policy and Strategy Formulation and Implementation

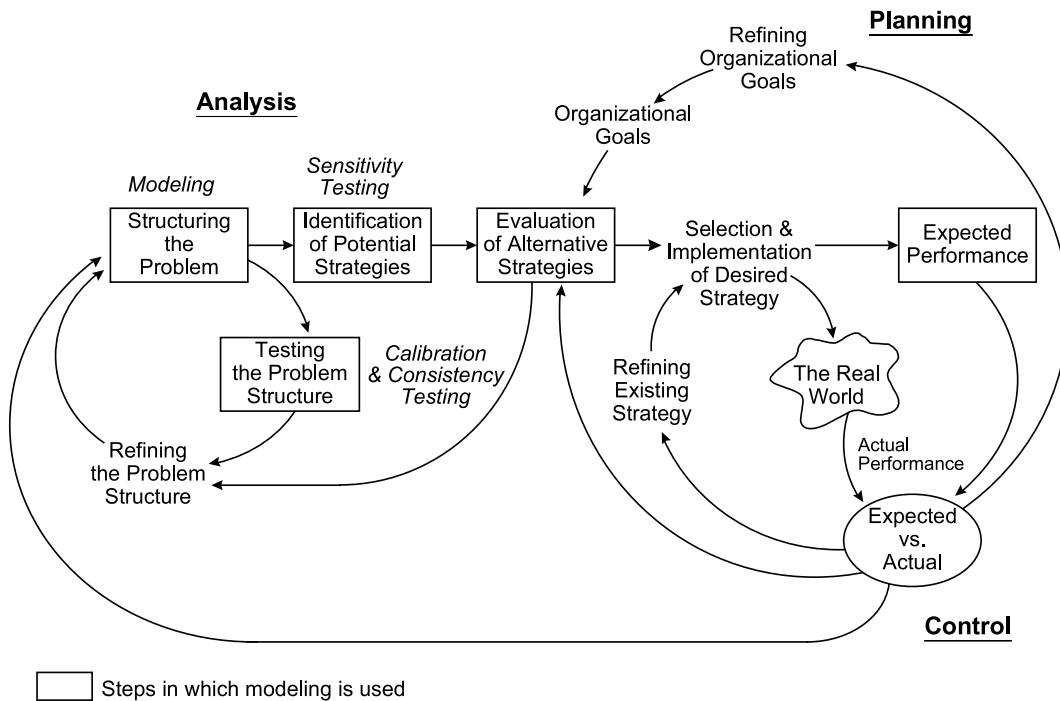
#### Role of System Dynamics Models in Policy and Strategy Formulation

There is general agreement among system dynamics modelers on the role that models play in the policy and strategy formulation process. This role has been depicted diagrammatically and described by Morecroft [65], Sterman [100], and Dyson et al. [23]. The role of the model in policy and strategy formulation is to act as a “virtual world” – a simpler and transparent version of the “real world”. The model serves as a vehicle for testing our understanding of the causes of behavior in the real world, and as a laboratory for experimentation with alternative policies and/or strategies.

One version of that role is shown in Fig. 2. In many cases, the process starts with the definition of a problem – an aspect of behavior that is problematic or threatening. This might be a decline in market share or profitability, or the threat posed by a new competitive product or service, as illustrated in the example of Fig. 1. Sometimes the problem can be expressed in terms of achieving business goals or objectives in the future. As illustrated in Fig. 2, the overall policy/strategy management process can be divided into three components: analysis, planning, and control. “Analysis” is usually triggered by a significant and/or persistent deviation between actual and expected perfor-

mance. It involves the iterative structuring, testing and refinement of an organization’s understanding of its operational or strategic problems and of the options open to it to deal with the performance gap. The model is the vehicle for this analysis – does our understanding of the system as reflected in model equations in fact produce behavior consistent with the observed problem, and if not, how can our understanding of structure be made more consistent with reality? The process of evaluating alternative policies/strategies often sheds new light on the problems faced by an organization or reveals the need for further analyzes. Note that the “modeling” cycle is iterative and compares simulated behavior to actual performance – the scientific method applied to strategy. Simulation of the model shows how business structure, policies, and external events together caused the past performance of the firm, and how future performance will evolve if structure, policies, and external events differ. The next phase, “planning”, is also an iterative process; it involves the evaluation, selection, and implementation of policies/strategies – some authors refer to this as “rehearsing” strategy. Evaluation of alternative policies/strategies depends not only on projected accomplishment of organizational goals, but also on the realities of current performance. The existing operational policies or existing strategy (vision, mission, strategic objectives) and goals are subject to refinement, as required, based on the successes and problems encountered, and in response to changing conditions.

A third phase of the policy/strategy formulation process is here called “control”. On-going policy/strategy management involves the continual, systematic monitoring of performance and the effective feeding back of successes, problems, threats, opportunities, experience, and lessons learned to the other components of the policy/strategy management process. The control phase is



**Business Policy and Strategy, System Dynamics Applications to, Figure 2**  
Policy/strategy management process and the role of system dynamics models [52]

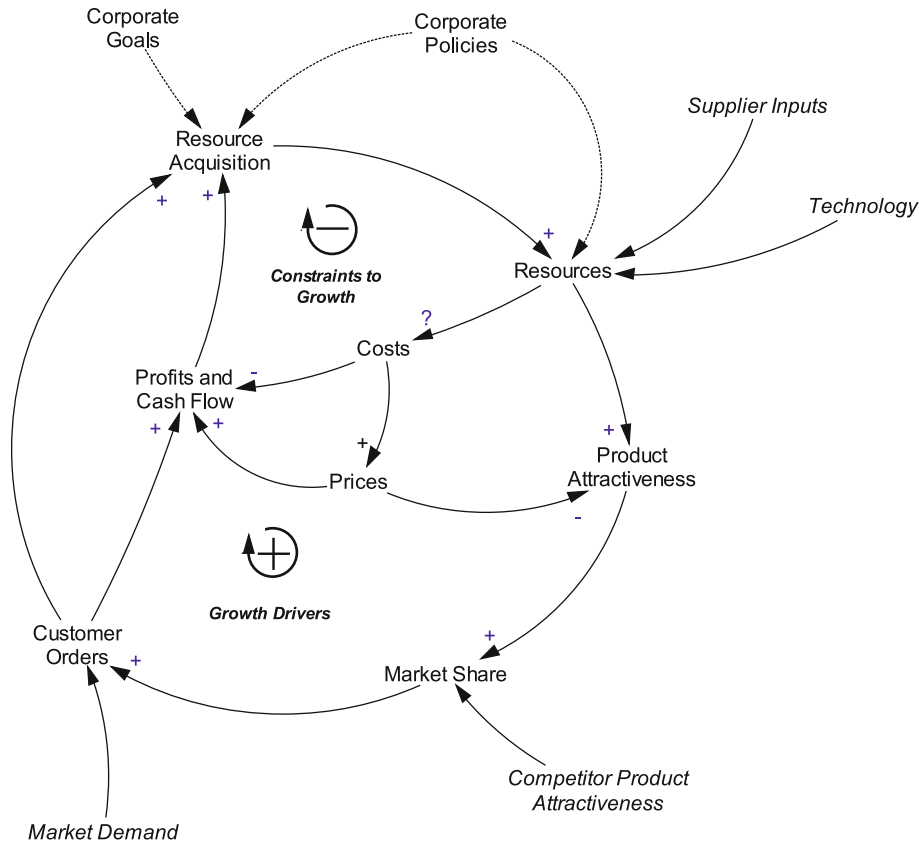
where organizations continue to learn. *The model provides an essential element to the control process – a forecast of expected performance against which actual performance can be monitored on a regular basis.* Deviations provide a signal for additional analysis: Has the policy/strategy been implemented effectively? Have conditions about the external environment changed? Are competitors acting differently than expected? Has the structure of the system changed? The model provides a means of assessing the likely causes of the deviation, and thereby provides an early warning of the need to act.

### Synergy Between Traditional Strategy and System Dynamics

While Fig. 2 illustrates how system dynamics models fit into the policy/strategy formulation process, there is also a synergy between system dynamics models and traditional strategy frameworks and concepts. Figure 3 illustrates the factors that drive business performance from a system dynamics perspective. Starting with resources, a firm's resources determine its product attractiveness; a firm's market share is based on that attractiveness compared to the attractiveness of competitor products; market share drives customer orders, which in turn generates profits and cash flow to finance the acquisition of ad-

ditional resources for further growth – thereby completing a growth-producing feedback around the outside of the figure (or as in the example of the telecom in Fig. 1, “growth” in the downward direction for the incumbent). However, the acquisition of additional resources can constrain future growth. To the extent increased resources increase costs, then profits and cash flow are reduced, and/or prices may need to increase. Both constrain growth (as might happen for the startups in the telecom example of Fig. 1).

There are a number of places in Fig. 3 where the system dynamics approach can be, and has been, connected to traditional strategy research and practice. For example, concepts such as learning curves, economics of scale, and economics of scope define possible connections between resources and costs – system dynamics models typically represent these connections. Figure 3 shows a number of factors external to the firm: market demand, competitor product attractiveness, technology, and supplier inputs. Strategy frameworks such as “five forces” and visioning approaches such as “scenario-based planning” [112], provide methods for thinking through these inputs – system dynamics models determine the consequences of alternative assumptions for the performance of the firm (note that system dynamics models also often internally represent structural dynamics of competi-

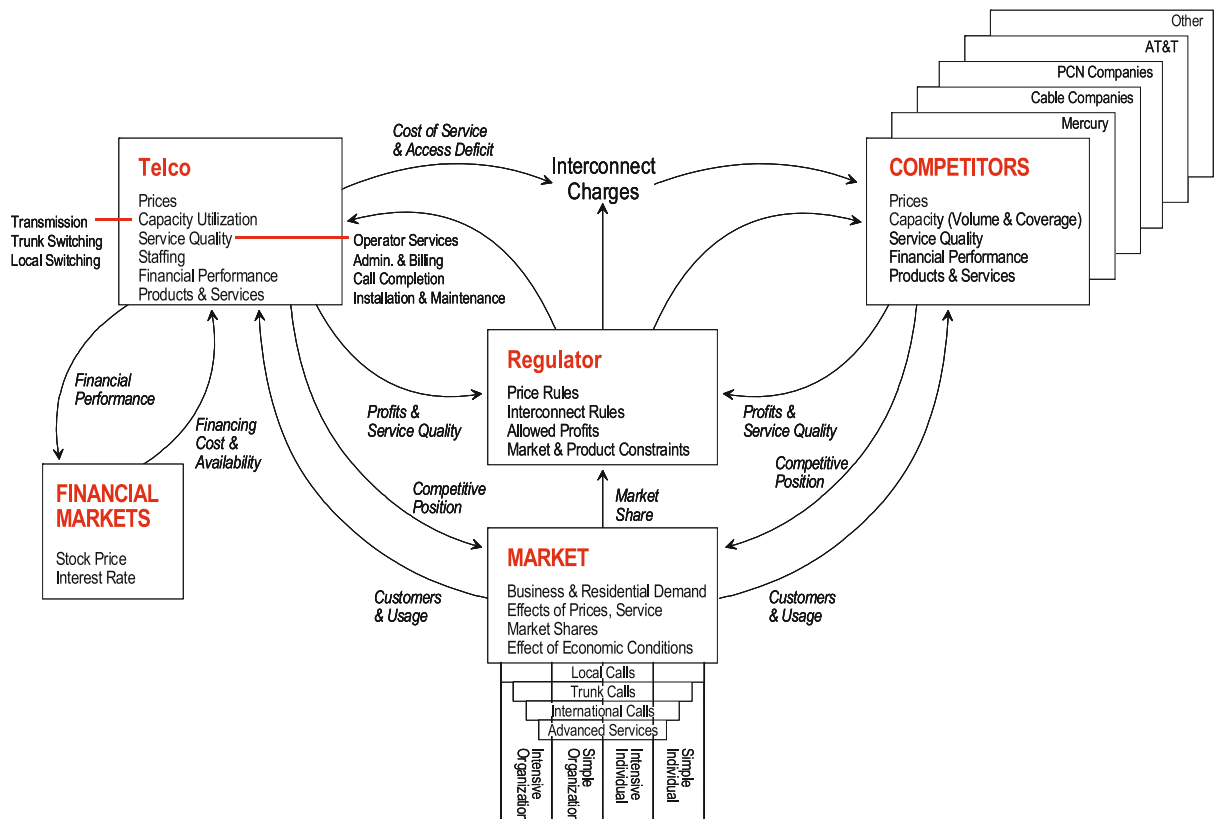


Business Policy and Strategy, System Dynamics Applications to, Figure 3  
Drivers of business performance (adapted from [49])

tors, suppliers, and the market as appropriate to explain the behaviors and issues of interest, rather than specifying them as exogenous inputs). For example, Fig. 4 shows a “sector” diagram of the major components of a strategy model developed by a consulting company for a telecom company dealing with loss of market share as in Fig. 1 above (from [35], originally developed by Lyneis). The model not only represents factors internal to the dynamics of the telecom firm, but also factors related to the internal dynamics of competitors, regulatory responses to telecom and competitor performance, financial market responses to telecom performance, and market reactions to telecom and competitor competitive position. This sector diagram connects to a number of the traditional strategy frameworks. In addition to these general connections, a number of system dynamics papers have detailed specific connections to strategy concepts: learning curves and product portfolios [59]; duopoly competition [97]; diversification [31,68]; and industry structure and evolution [93].

Many of the connections between system dynamics and traditional strategy practice and research are discussed in Warren’s *Competitive Strategy Dynamics* [105] and *Strategic Management Dynamics* [108]. More importantly, Warren’s book details and expands on the connections between the system dynamics concepts of structure, particularly the concepts of stocks and flows, and the well-established resource-based view (RBV) of strategy and performance. (See [63], and [30], for explanations of the development of RBV; a managerial explanation of how this theoretical perspective can be applied can be found in Chap. 5 in [36]. Again, system dynamics provides a means of simulating the consequences of alternative resource acquisition and allocation strategies on firm performance. As such, there would seem to be a strong synergy between system dynamics and this strategy approach.

In summary, while system dynamics and traditional strategy approaches developed largely independently, the potential synergies between the two are significant. Until recently, few researchers and practitioners have made the



Business Policy and Strategy, System Dynamics Applications to, Figure 4  
"Sector" diagram from a typical strategy model

effort to cross disciplines and exploit this synergy. More effort and publication are needed to demonstrate areas of synergy and get system dynamics into the mainstream of business strategy research and ultimately practice.

### Working with Management Teams to Achieve Implementation

While Fig. 2 depicts the overall role of the model in strategy management, the approach to developing and using the model itself involves an iterative, multi-phased process. That process has evolved over time as practitioners and researchers have learned from experience. Since its inception, system dynamics has been concerned with having an impact on business decisions. Jay Forrester, the founder of the field, stressed the importance of working on important problems – those that affect the success or failure of firms – and generating solutions that are relevant to those problems. Ed Roberts, one of the first researchers and practitioners in the field, was also involved in early research and experimentation on organizational change and how models can fit into that process [90,92].

The emphasis on having an impact has remained a central tenet of system dynamics, and over the years system dynamics practitioners have developed and refined methods of working with managers to not only solve problems, but also to enhance the likelihood of those solutions being implemented.

Starting with Roberts and his consulting firm Pugh-Roberts Associates (now a part of the PA Consulting Group), a number of researchers and practitioners have contributed to the evolving approach of working with managers to affect organizational change. These include, but are not limited to, Coyle and his group originally at Bradford University in the UK [14,15], Morecroft and his group at the London Business School [65,71], Richmond [88], Sterman [100] and Hines [37] at MIT, the "group model building" approach ([87,104]; ► [Group Model Building](#) and Peter Senge's organizational learning [95]. While there are some differences in emphasis and details, there is in general agreement on the high-level process of using system dynamics to affect corporate strategy development and change. In this section, I describe that process, discuss some specific areas where there is some



divergence in practice, and end with some examples to the approach in practice.

In the early years, the approach of most system dynamics to consulting was heavy on “product” and light on “process”. Like management science in general, many in system dynamics took the view that as experts we would solve the client’s problem for him, and present him with the solution. Practitioners gradually recognized that elegant solutions did not necessarily lead to implementation, and consulting styles changed to include increased client involvement [90,109]. At the same time, the “product” was evolving to meet the needs of clients (and to take advantage of the increased power of computers). Practitioners evolved from the smaller, policy-based models which characterized the original academic approach to more detailed models, along with the use of numerical time series data to calibrate these models and determine the expected numerical payoff to alternative strategies [52]. In addition, during the 1980s academic research began to focus more on “process:” the use of models in support of business strategy and on more effective ways to involve the client in the actual building of the model [65,66,87,88,104].

System dynamics practitioners now generally agree on a four-phased approach to accomplish these objectives:

1. Structuring the Problem
2. Developing an Initial Model and Generating Insight
3. Refining and Expanding the Model, and Developing a Strategy
4. On-going Strategy Management and Organizational Learning.

In practice, there is sometimes a fifth phase of work. Often modelers simplify a final project model in order to capture the core feedback loops that lie behind observed dynamics. Many of the generic structures discussed later in this paper arose in this way. In other cases, modelers will create management “games” and/or learning labs from the final project model.

As discussed below, the relative mix between “product” (detail complexity and calibration of model) and “process” (degree of involvement of client in model development and use) is perhaps the main difference in the style and approach of different practitioners of system dynamics. For those new to system dynamics, or for those seeking good examples, there is an excellent example of this process, including model development, by Kunc and Morecroft [44].

**Phase 1 – Structuring the Problem** The purpose of the first phase of analysis is to clearly define the problem of interest (either a past problem behavior or a desired fu-

ture trajectory), the likely causes of that problem (or desired trajectory), and any constraints that may arise in implementing a solution. It identifies the performance objectives of the organization and possible solutions – all to be rigorously tested in later phases. Similar to more traditional policy and strategy approaches, during this phase the consultant team reviews company documents, the business press, and available company data, and interviews company managers, and possibly customers and competitors. During this phase, the hypothesized drivers of business performance are identified: what compels customers to buy this product, what compels them to buy from one supplier versus another, what drives the internal acquisition and allocation of resources, what major externalities affect the business (e.g., the economy, regulations, etc.), perhaps drawing on frameworks such as “five forces” and SWOT analyzes. More importantly, these drivers are linked in a cause-effect model to form a working hypothesis of the reasons for company behavior. This hypothesis formation builds heavily on the tools and techniques of what is now commonly called “systems thinking”:

1. Behavior-over-time graphs (reference modes) – Graphs of problematic behavior over time, often with objectives for future performance highlighted (using actual data where readily available).
2. Causal-loop and mixed causal, stock-flow diagramming as a diagrammatic hypothesis of the causes of problematic behavior.
3. System archetypes, or common generic problem behaviors and structures observed over and over again in different businesses, as a means of identifying structure (see for example [95] and [43]); and
4. Mental simulation – does the hypothesis embodied in the conceptual model seem capable of explaining the observed problem(s)? Mental simulation is also used to identify the possible impact of alternative courses of action.

Note that the exercise to this point, as commonly practiced, is almost entirely qualitative. Warren [105,108] introduces quantitative dimensions even in this phase.

**Phase 2 – Developing an Initial Model and Generating Insight** The power of system dynamics comes from building and analyzing formal computer models. This is best done in two steps. In the first, a small, insight-based model is developed to understand the dynamics of the business so as to generate insights into the direction of actions needed to improve behavior. The small, insight-based model is also the next logical progression beyond “systems thinking” in the education of the client in the

methods and techniques of system dynamics modeling. In the second quantitative modeling step (Phase 3 below), a more detailed version of the first model is developed, and is often calibrated to historical data. Its purpose is to quantify the actions needed, to assure that the model accurately reflects all relevant knowledge, and to sell others.

Small models (anywhere from 20–100 equations) make it much easier to understand the relationship between structure and behavior: how is it that a particular set of positive feedback loops, negative feedback loops, and stocks and delays interact to create the behavior shown in the simulation output? This can only be determined by experimentation and analysis, which is very difficult with large models. The focus of the first model is on insight generation, communication, and learning, rather than determining a specific shift in strategic direction or investment. These models can help managers improve their intuition (mental models) about the nature of their business, and thereby to better understand the rationale behind more detailed strategies that evolve in later phases of model development.

In summary, Phase 2 delivers:

- A small model which recreates the observed pattern of behavior or hypothesized future behavior (and is roughly right quantitatively);
- Analysis and understanding of the principal causes of that pattern of behavior;
- Ideas of high leverage areas that could improve behavior into the future; and
- Recommendations as to where additional detail will improve the strategy advice, or will make the results of the model more usable and/or easier to accept by others.

**Phase 3 – Refining and Expanding the Model, and Developing a Strategy** The final phase of model development entails the iterative expansion of the model to include more detail, and often calibration to historical data, as deemed appropriate for the situation. One progressively adds detail and structure, initially to make the process manageable, and then as necessary to correct discrepancies between simulated output and data, or to add policy handles and implementation constraints. Further, model development is likely to continue in the “on-going learning” phase as additional structure and/or detail is required to address new issues that arise. The purpose of this more elaborate modeling phase is to:

1. *Assure that the model contains all of the structure necessary to create the problem behavior.* Conceptual models,

and even small, insight-based models, can miss dynamically important elements of structure, often because without data, the reference mode is incomplete or inaccurate (see [52] for examples of this).

2. *Accurately price out the cost-benefit of alternative choices.* Strategic moves often require big investments, and “worse-before-better” solutions. Knowing what is involved, and the magnitude of the risks and payoff, will make sticking with the strategy easier during implementation. Understanding the payoff and risks requires quantifying as accurately as possible the strengths of relationships.
3. *Facilitate strategy development and implementation.* Business operates at a detail level – information is often assembled at this level, and actions must be executed at that level. Therefore, the closer model information needs and results can be made to the normal business lines and planning systems of the company, the easier strategy development and implementation will be. And,
4. *Sell the results to those not on the client’s project team.* Few, if any, managers can dictate change – most often, change requires consensus, cooperation, and action by others. The “selling” of results may be required for a number of reasons. If, as in the optimal client situation, consensus among key decision-makers is achieved because they are all a part of the project team, then the only “selling” may be to bring on board those whose cooperation is needed to implement the change. Under less optimal client circumstances where the project is executed by advisors to key decision-makers, or by a support function such as strategy or planning, then selling to decision maker(s) and to other functions will be required.

There are two important elements in this phase of work: adding detail to the model; and possibly calibrating it to historical data. Adding detail to the small, insight-based model usually involves some combination of: (1) disaggregation of products, staff, customers, etc.; (2) adding cause and effect relationships, and feedback loops, often where the more detailed disaggregation requires representing allocations, etc., but also to represent additional feedback effects that may seem secondary to understanding key dynamics, but may come into play under alternative scenarios, or may later help to “prove” the feedbacks were not important; (3) including important external inputs, typically representing the economy, regulatory changes, etc.; and (4) adding detailed financial sectors, which entail numerous equations, with important feedback from profitability and cash flow to ability to invest, employment levels, pricing, and so on. Calibration is the iterative process

of adjusting model parameters, and revising structure, to achieve a better correspondence between simulated output and historical data. Whereas the Phase 2 model primarily relies on our store of knowledge and information about cause-effect structure, the Phase 3 model relies on our store of information about what actually happened over time.

In summary, Phase 3 delivers:

- An internally consistent data base of strategic information;
- A detailed, calibrated model of the business issue;
- A rigorous explanation and assessment of the causes of performance problems;
- Analyzes in support of strategic and/or tactical issues;
- Specific recommendations for actions; and
- Expectations regarding the performance of the business under the new strategy, and the most likely scenario.

**Phase 4 – On-going Strategy Management System and Organizational Learning** True strategy management (“control”) involves the on-going, systematic monitoring of performance and the effective feeding back of successes, problems, threats, opportunities, experience, and lessons learned to the other components of the strategy management process. The control phase is where organizations continue to learn. *The model provides an essential element to the control process – a forecast of expected performance against which actual performance can be monitored on a regular basis.* Deviations provide a signal for additional analysis: Has the strategy been implemented effectively? Have conditions about the external environment changed? Are competitors acting differently than expected? Has the structure of the system changed? The model provides a means of assessing the likely causes of the deviation, and thereby provides an early warning of the need to act. This feedback is only possible with a detailed, calibrated model.

### Differences in Emphasis and Style

While there is general agreement among system dynamics practitioners regarding the role of models in the strategy development process and of the basic steps in that process as described above, there are some differences in emphasis and style regarding: (1) the use of causal-loop diagrams (CLDs) vs. stock-flow (SF) diagrams; (2) whether you can stop after Phase 1 (i.e., after the “qualitative” phase of work); (3) is calibration necessary and/or cost effective; and (4) how much model detail is desirable.

**CLDs vs. SF** There are disagreements within the field about the value of causal loop diagramming (versus stock-flow diagrams). Causal-loop diagrams focus on the feedback loop structure that is believed to generate behavior; stock-flow diagrams also include key stocks and flows, and in the extreme correspond one-to-one with complete model equations. In my view, there is no “right” answer to this debate. The most important point is that in Phase 1 diagramming one is trying to develop a dynamic hypothesis that can explain the problem behavior, and that can form the basis of more detailed diagramming and modeling – whether that dynamic hypothesis is a CLD, a stock-flow diagram with links and loops labeled, or some combination depends in part on:

- Personal style and experience – some people, Jay Forrester perhaps being the best example, seem to always start with the key stocks and flows and work from there; Kim Warren [106] also argues for this approach as an effective means of connecting to the way managers view the problem and to the data;
- The structure of the system – some systems have “obvious” key chains of stocks and flows, and so starting there makes the most sense (for example, the aging chains in the urban dynamics model [29], the rework cycle on projects, and inventory control systems); other systems, without critical chains of stocks and flows, may be easier to address starting with CLDs;
- Whether or not you are doing the model for yourself or with a group – especially if the group is not conversant with the basics of system dynamics, starting with CLDs is easier, and it’s also easier to brainstorm with CLDs (which is different than developing a dynamic hypothesis); but again, it’s personal preference and nature of system as well. In practice, I have found that CLDs alone, or a mixed stock-flow/causal diagram, are extremely valuable for eliciting ideas in a group setting about the cause-effect structure of the business, and later for explaining the dynamics observed in simulation output. However, one cannot build a model literally from a causal diagram, and either explicit or implicit translation is required.

**Qualitative vs. Quantitative Modeling** Some practitioners of system dynamics believe that strategic insights can sometimes be obtained after the first phase of work, after the dynamic hypothesis and mental simulation (and note that much of traditional strategy practice relies on such qualitative insights). Coyle [17,18] argues for this; the popularity of “systems thinking” engendered by Senge’s work [95] has spawned a number of practitioners that use

only qualitative modeling [52]. Coyle's views generated a strong counter response from [41]. Wolstenholme [110] provides a history and discusses his view on the advantages and disadvantages of each approach. My own view is that while Phase 1 and the systems thinking that is a key part of it are a necessary start, it should not be the end point. Two problems limit its effectiveness in supporting business strategy. First, simple causal diagrams represented by system archetypes, while useful pedagogically, take a very narrow view of the situation (typically, one or two feedback loops). In reality, more factors are likely to affect performance, and it is therefore dangerous to draw policy conclusions from such a limited view of the system. A more complete representation of the problem considers more feedback effects and distinguishes stocks from flows, but introduces the second problem: research has shown that the human mind is incapable of drawing the correct dynamic insights from mental simulations on a system with more than two or three feedback loops [78,98]. In fact, without the rigor and check of a formal simulation model, a complex causal diagram might be used to argue any number of different conclusions. In addition to overcoming these limitations, as discussed below, formal modeling adds significant value to the development and implementation of effective business strategies. Warren (p. 347 in [107]) also stresses need to focus on quantitative behavior to achieve management consensus.

**Need for Data/Validation** The necessity of obtaining numerical data and calibrating model output to that data is also questioned by some practitioners. While I agree that curve fitting via exogenous variables is not a useful endeavor, proper calibration is an important part of the scientific method that involves systematically comparing simulation output to data, identifying causes of error, and correcting discrepancies by improving first the structure of the model and then its parametrization. In some cases, discrepancies are ignored because they are deemed to be caused by factors irrelevant to the problem of interest, or may be "fixed" by exogenous factors if these are deemed significant by the client and are consistent with the remaining model structure and calibration. As Homer [38,39] argues, the use of historical data and calibration is essential to scientific modeling.

In some cases, organizations lack the data on key factors felt to be essential to the dynamic performance of the business, and by implication essential to sound strategic management of the business. The modeling process can highlight these short-comings and, in the short-term, substitute educated assumptions for this data. In the longer-

term, companies can be encouraged to acquire this important data (and substitute it for much of the unimportant information companies generally pore over).

Accurate calibration can greatly enhance confidence in a model. This can be especially important when trying to convince others of the appropriateness of actions a management team is going to take, or to demonstrate to others the need to take action themselves based on the results of the model. Calibration can also be important for other reasons: (1) numerical accuracy is often necessary to evaluate the relative cost and benefits of changes in strategy, or to assess short-term costs before improvements occur; (2) calibration often uncovers errors in the data or other models, especially incomplete or incorrect mental models that form the basis for the dynamic hypothesis (see [52] for examples); and (3) the "control" feedback in the fourth phase of the strategy management process is only possible with a detailed, calibrated model.

**Level of Detail and Model Complexity** Some practitioners argue that large, complex models should be avoided, for a number of reasons: they can be even more like black boxes; they can be difficult to understand (not only for the non-modelers, but even the modelers); and they are costly to develop. Morecroft argues that a detailed model "loses its agility and becomes less effective as a basis for argument". (p. 227 in [65]) In practice, the first two issues can be avoided and/or minimized by executing the model development in three phases as discussed above. This allows the client to grow slowly with the concepts, and it allows the modeling team to develop a solid understanding of the model. The third problem is generally not an issue if you are working on significant problems – in my view the cost of the consulting engagement is trivial relative to the expected payoff. While I believe that the client obtains value, regardless of when you stop, strategy consulting is one case where the "80/20 rule" does not apply – the client does not get 80% of the value for 20% of the cost (which would be essentially at the end of Phase 1). In part this is a function of what I view as the objective of the project – providing tools, strategic analyzes, and advice in support of an important strategic and/or investment decision. In this situation, the "value" is back-end loaded. Finally, effective strategy management is only possible with a detailed, calibrated model.

In addition, detail and calibration are often necessary to sell the model to others. In many situations, everyone who may have an input to a strategic decision or be necessary for successful implementation cannot be a part of the client team. As surprising as it may seem, the selling of re-

sults (as opposed to understanding) is easier to accomplish with a detailed, calibrated model than with a small model. First, the numerical accuracy gives the model face validity. Second, a detailed model more often allows the modeler to counter the “have you considered (insert pet theory)?” criticism. I have often found that when you start explaining the model to others, they respond by asking “Have you considered this feedback? Or this effect?” And if you have not, that ends the discussion. Even though you may think that feedback or that effect may not have any impact, if it is not included in the model you cannot say “Yes, we looked at that and it did not have any impact”, and explain why. If it is not in the model the critic can argue that your results would be changed by the inclusion of their pet theory. One has a hard time countering that assertion without a convincing argument based on simulation results. Finally, a detailed, calibrated model helps tell a convincing story. The simulation output, which corresponds closely to the data, can be used to explain (again with output) why, for example, a loss of market share occurred. How price relative to the competitors’ price was the key factor, and/or how the factors affecting share changed over time. The simulation output can and should be tied to specific events. We have found that an explanation like this is compelling, and is important in enhancing the credibility of the model and the modeler.

The benefits of large, complex models in a consulting setting are also noted by Winch [111]. He specifically finds that “For the executive team to have confidence in the impartiality of the model, each person must feel it captures the detailed pressures and processes of his or her own sphere of responsibility yet produces a holistic view of the organization”. (pp. 295–6 in [111]), and that the model was essential to getting everyone to agree: “The process of building system dynamics models, in each case ostensibly as a forecasting and evaluation tool, enabled the managers eventually to develop a shared view, which formed the basis for formulating and agreeing upon a final strategy”. (p. 298).

### Process Examples

There are a number of published examples that support the four-phase process of applying system dynamics to business strategy:

- Lyneis [52] provides not only a more fully developed description of the detailed, calibrated-model Pugh–Roberts approach, but also illustrates its application to the credit card and airline manufacturing industries.
- Morecroft et al. [70] describe how a model was created and used to stimulate debate and discussion about growth management in a biotechnology startup firm. The paper highlights several novel features about the *process* used for capturing management team knowledge. A heavy emphasis was placed on mapping the operating structure of the factory and distribution channels. Qualitative modeling methods (structural diagrams, descriptive variable names, “friendly” algebra) were used to capture the management team’s descriptions of the business. Simulation scenarios were crafted to stimulate debate about strategic issues such as capacity allocation, capacity expansion, customer recruitment, customer retention, and market growth, and to engage the management team in using the computer to design strategic scenarios. The article concludes with comments on the impact of the project.
- Winch [111] examines the role that building and using a system dynamics model plays in developing consensus within management teams facing key strategic decisions: A shared view emerges within the team as individual views of the company, its industry, and the socioeconomic climate are articulated and compared. Examples are given based on two actual consulting assignments in which differing views concerning the competitive environment and the general business outlook initially pointed to quite different strategies. The emergence of consensus was considered a major benefit in addition to the forecasts and quantitative evaluations the model provided. In its analysis and examples, this article emphasizes both the “hard” benefits of forecasts and an objective framework for quantitative evaluations and the “soft” benefits of building consensus within management teams.
- Coyle [15,16] also has an approach that he discusses, with emphasis on CLDs (he terms these “influence diagrams”, and his group was instrumental in initial use of this technique).
- Snabe and Grossler [94] show how modeling can be supportive for strategy implementation in organizations and illustrate with a detailed case study from a high-tech company.
- A special issue of the Journal of the Operational Research Society on System Dynamics for Policy, Strategy, and Management, edited by Coyle and Morecroft [19], contains a number of papers which in part discuss consulting process issues [21,110,112] among others).
- The special issue Fall 2001 of *System Dynamics Review* on consulting practice contains papers by Thompson [102], Campbell [11], and Backus et al. [5] that focus on the consulting process.



### Theory Development – Understanding the Drivers of Business Dynamics

Another important contribution of system dynamics to business policy and strategy formulation is the development of structural theories to explain commonly observed patterns of behavior. Theory development provides us with: an understanding of the basic drivers of business dynamics; insights, enhanced mental models, and policy guidelines for improved performance; and building blocks of tested model equations for real applications (equations for the model must be provided for models to add to our base of theory). System dynamicists have developed structural theories to explain the basic patterns of business dynamics: (1) cycles and instability; (2) productivity and eroding performance; (3) life cycles; and (4) growth. Each is discussed in turn below.

#### Cycles and Instability: Stock Management, Supply Chains, and Manufacturing Systems

The very first applications of system dynamics were to understanding the tendencies of production-distribution systems, or “supply chains”, toward cycles and instability; these applications remain important to this day [25,26]. For example, the “Beer Game”, now distributed by the System Dynamics Society, was developed and refined at MIT beginning in the early 1960s and remains one of the most popular introductions to both system dynamics principles, and to supply chain issues.

Supply chains are an important component of all industrialized societies. They exist in any industry where goods are produced and distributed to consumers, for example, food and beverage production and distribution, or manufactured goods such as automobiles and appliances. Supply chains exhibit a classic behavior pattern which has impacts not only on the individual company, but also for the economy as a whole: as one moves up the supply chain from the end user, any variation in orders from the end user are progressively *amplified* and delayed at each additional stage in the chain – factory variations are greater than customer variations; raw materials production variations are greater than factory variations (see Chap. 17 and 20 in [100] for real world examples of this behavior). This behavior is also sometimes referred to as the “bull-whip” effect.

Figure 5 illustrates the structure of one stage of a typical supply chain and the causes of amplification: a stock of inventory is depleted by shipments (here assumed equal to demand) and replenished by production completions (or more generally, shipments from a supplier); the stock of goods in production (or goods being assembled and

shipped by a supplier) are increased by production and reduced, after the production (and/or shipping) delay, by production completions. This structure has a tendency to “amplify” any changes in demand – that is, “production” (or orders and reorders, depending on the system) increase or decrease more than any increase or decrease in demand, and tend to lag changes in demand. For example, in Fig. 5, when demand increases, even if production increases immediately inventory falls because production completions are delayed by the production (and/or shipping) delay. Therefore, production must increase higher than demand (amplification) in order to rebuild inventories. In addition to production and shipping delays, inventory might also fall because of delays caused by smoothing information about demand (such that production changes lag changes in demand). Production further increases above demand because of the need to increase inventories and production or supply lines to higher target levels. Intuitively, and verified by simulations, amplification is greater if: desired inventories are larger; production/transit delays are longer; and/or responses to inventory gaps are more aggressive (smaller adjustment time constant, as discussed below).

This basic structure in Fig. 5 also illustrates the “stock management” problem. In Fig. 5, the stock of finished goods inventory must be managed in order to serve customer demand in a timely fashion. Figure 6 details the structure typically used to control stocks, one of the most used and important structures in system dynamics:

$$\text{Production} = \text{Expected Demand} \\ + \text{Inventory Correction} + \text{Goods In Process Correction}$$

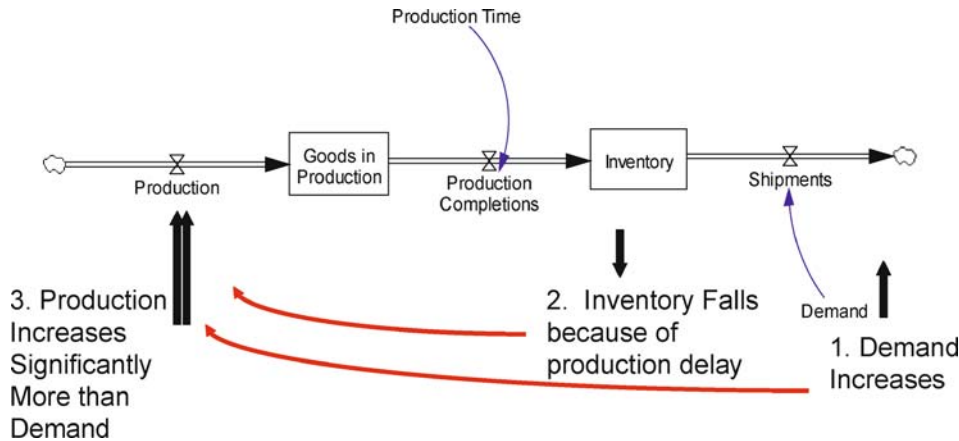
$$\text{Inventory Correction} \\ = (\text{Desired Inventory} - \text{Inventory}) / \text{Time to Correct}$$

$$\text{Inventory Desired Inventory} \\ = \text{Expected Demand} \times \text{Months Coverage Goal}$$

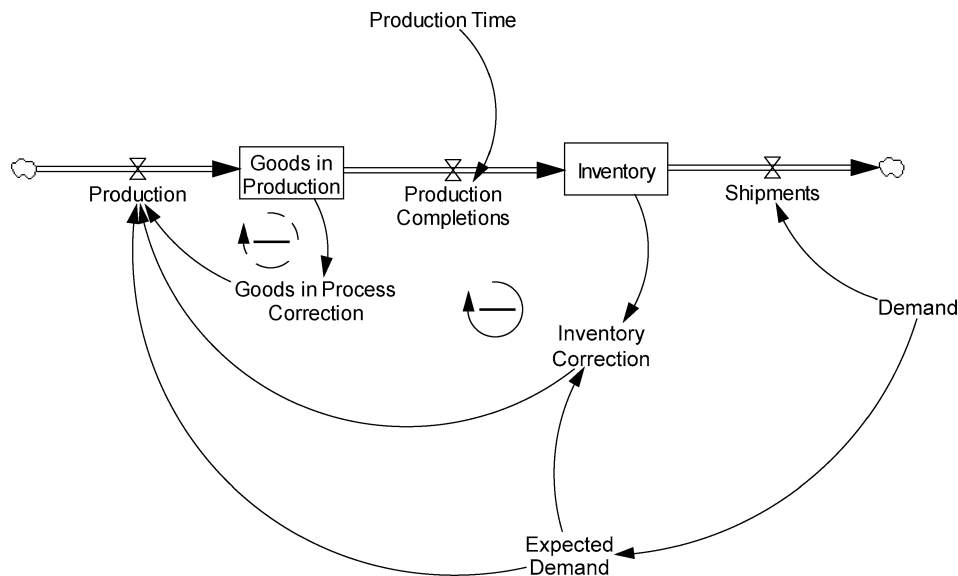
$$\text{Goods In Process Correction} = \\ (\text{Desired Goods In Production} - \text{Goods In Production}) \\ / \text{Time to Correct Inventory}$$

$$\text{Desired Goods In Production} \\ = \text{Expected Demand} \times \text{Production Time}$$

Stock management is complicated by delays in replenishing the stock, here a production delay. Depending on the pattern of demand, there is often a tradeoff between amplification and variations in inventory – less aggressive responses (longer time to correct inventory) generally reduce amplification but cause greater variations in inventory (and therefore may necessitate higher target levels to



Business Policy and Strategy, System Dynamics Applications to, Figure 5  
Structure and causes of amplification in one stage of a supply chain



Business Policy and Strategy, System Dynamics Applications to, Figure 6  
Stock management structure

reduce the likelihood of stockouts); more aggressive responses (shorter time to correct inventory) increase amplification and demands on manufacturing and suppliers, and potentially costs, but can result in more stable inventory levels. However, under some demand patterns and production conditions, aggressive responses can increase both amplification and inventory instability. While as noted below structural changes can significantly improve the overall performance of stock management and supply chain systems, nevertheless this fundamental trade-off between amplification and inventory levels will remain. The “optimal” solution will vary by firm, and over time as

the inherent pattern of demand changes. These dynamics and tradeoffs are discussed in depth in [49,100].

In a typical supply chain, there are multiple stages connected in series, for example, in the automotive industry: dealers, car manufacturers/assemblers, machine tool producers, parts manufacturers, raw material suppliers (with potentially several stock management stages in some of these main categories). The upstream stages suffer greater amplification than the downstream stages. In the main, this occurs because each stage uses as its demand signal the orders from the prior stage, which are amplified by that stage’s stock management policies as dis-

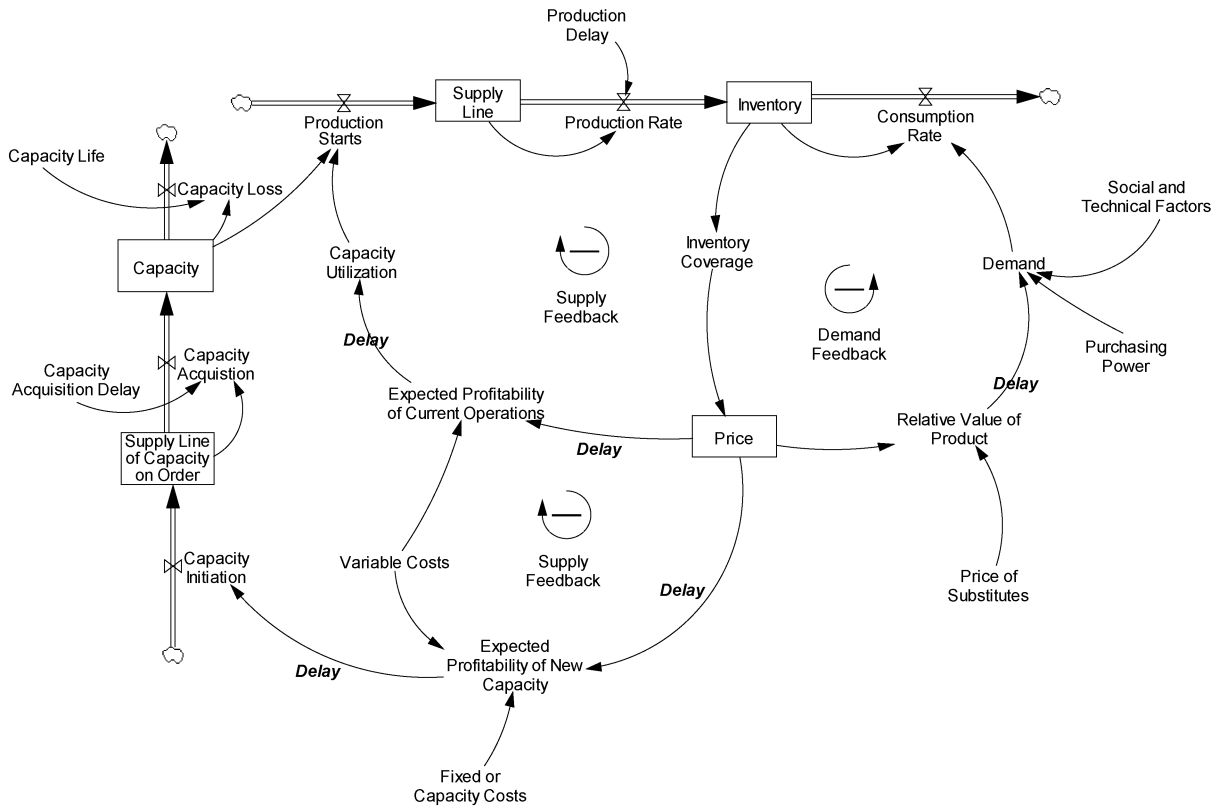
cussed above. Other reasons for increased upstream amplification include [2]: (1) in determining “expected demand”, each stage tends to extrapolate trends in orders from the prior stage; (2) order batching; (3) price fluctuations (in response to inventory levels); and (4) shortage gaming (ordering more than you really need to get a higher share of the rationed goods from the supplier; see [52] for an example in the aircraft industry). System dynamics analyses have identified a number of structural changes which can improve supply chain and stock management performance: (1) reduce delays; (2) reduce inventory; and (3) share information (for example, if upstream stages are aware of the downstream end user customer demand pattern, they can use that information rather than the amplified orders from the prior stage as the basis for their decisions, and at least partially avoid amplification [20]).

Applications of system dynamics to supply chain management, and production management, remain an important area of research and applications. Akkermans and Daellaert [1], in an article entitled “The Rediscovery of Industrial Dynamics: The Contribution of System Dynamics to Supply Chain Management in a Dynamic and Fragmented World”, provide an excellent survey of supply chain management and system dynamics potential role in moving that field forward. Additional work in this area includes Morecroft’s original analysis of the dynamics created by MRP systems ([64]); Gonçalves doctoral dissertation [32], some of which is summarized in [33,34]; Anderson and Fine [3] on capital equipment supply cycles, and Zahn et al. [114] on flexible assembly systems. Each of these discusses variations on the basic stock/production/supply chain management systems, and provides references for further research.

In addition to inventories, firms need to manage other stocks and resources, including raw materials, employees, capital equipment, and so on; the stock management structure described above for inventory applies to these other stocks as well. The management of stocks and resources is central to dynamic and strategy problems in many industries. First, the management of one stock often influences the ability to manage other stocks (for example, capital equipment and employees determine production). Not only does this interdependency create constraints, the additional negative feedback control in managing resources is another source of cyclical behavior (see Chap. 19 in [100], and Chap. 5 of [69]). Second, in addition to the stocks of resources, production is affected by the productivity of those resources. Dynamic drivers of productivity, such as experience and fatigue, are discussed in the next section.

While the negative control feedbacks described above are central to the observed cyclical behavior of supply chains and resource-based firms, an additional negative feedback through the market adds a further source of instability. This dynamic is perhaps clearest in commodity-based industries, which have also been extensively modeled by system dynamicists as first summarized by Meadows [58]. As illustrated in Fig. 7, these models integrate the supply chain with the dynamics created by supply and demand – in a typical commodity system, there are three major negative feedback loops: two supply feedbacks (one through production, often representing the resource labor, and one through the resource capacity), and one demand feedback (for example, an increase in inventory causes prices to fall, which increases demand and leads to a decrease in inventory from what it otherwise would be). Commodity industries typically exhibit behaviors that include cycles of two periodicities, one determined primarily by the shorter production feedback loop and another longer cycle driven by the capacity loop (see Chap. 20 in [100] for both detailed equations and for examples of the structure applied to the livestock and paper industries). The demand feedback loop, however, can play a role in the dynamics as well – if the demand feedback is strong and with a short delay, then demand corrections occur before the supply feedbacks operate and system stability is improved; however, if the magnitude of the delay in the demand loop is similar to the magnitude of the delays in either of the supply loops, the intensity of the corresponding cycle is increased as two negative feedback loops are both independently acting to “solve” the inventory problem. In addition, commodity industries, where they involve a depletable resource such as oil, can experience long-term resource depletion dynamics [101].

In conclusion, manufacturing and supply chain dynamics are central to many of the behaviors observed in businesses (see Chap. 20 in [100] for real world examples of these cycles). The supply chain, stock management, resource management, and commodity structures discussed above are therefore important components of many system dynamics models developed to support business policy and strategy. In some cases, the firm can change policies to reduce the severity of these cycles; in other cases, especially where the cycles are driven primarily by industry dynamics, the individual firm can use the enhanced understanding and forecasting of these cycles for more strategic decisions such as new product introduction and capacity planning (as in the commercial aircraft market case illustrated in Fig. 1 and discussed in Sect. [Applications and Case Examples](#)).



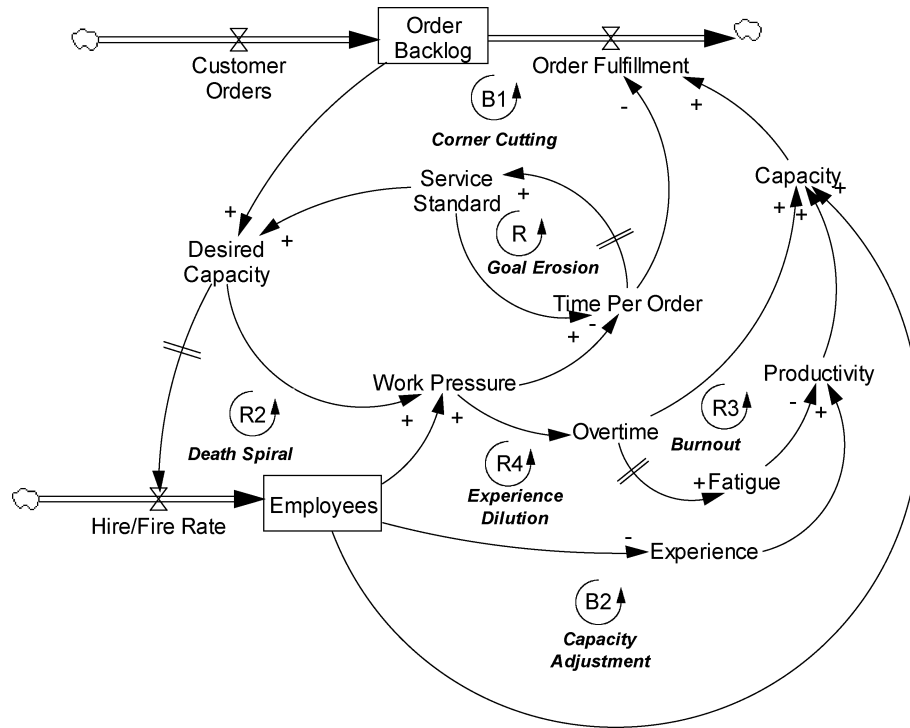
Business Policy and Strategy, System Dynamics Applications to, Figure 7

Commodity industry dynamics showing three controlling feedbacks (adopted from [100])

### Productivity and Eroding Performance: Service Industry Dynamics

Service-based firms (e. g. professional services, transportation, catalog and online shopping, etc.), and the service arms of manufacturing-based organizations, have a somewhat different set of structural dynamics. Firms in these industries have a number of characteristics that make them more difficult to manage than more manufacturing intensive industries: (1) their product is difficult if not impossible to inventory, and so something else must buffer changes in demand; (2) they are particularly dependent on the performance of people (although the productivity of resources is also important to manufacturing-based businesses as well); and (3) the performance of the system can be harder to detect, and so they are much more subject to a gradual erosion in performance and goals. "The major recurring problems observed in service industry – erosion of service quality, high turnover, and low profitability – can be explained by the organization's response to changes in work pressure." (see p. 28 of [75]).

One of the primary distinguishing features of service-based firms is that their end product is people-dependent and cannot be inventoried. While there may be inventories of products that support delivery of the service, that delivery must be performed based on current resources. As a result, work or order backlogs are the stock that buffers demand from "production". A simplified example is shown in Fig. 8. Customer orders (demand) fill an order backlog, which is depleted by order fulfillment. Order fulfillment is based on the firm's service "capacity" and the amount of time spent per order – for the same capacity, order fulfillment will be greater if less time is spent per order (although service quality may suffer). Capacity is dependent upon people, overtime, and productivity, as discussed further below. Employees are increased or decreased based on desired capacity, which in turn depends on order backlog relative to the firm's service standard (time per order). Work pressure depends on desired capacity relative to the current stock of employees – to the extent the number of employees does not increase as needed, work pressure builds which can result in increases in effective capacity



**Business Policy and Strategy, System Dynamics Applications to, Figure 8**  
**Drivers of service business dynamics**

via overtime, or reductions in time spent per order. Time spent per order depends on the firm's service standard for time spent per order, modified by work pressure – if work pressure is high, time spent can be reduced. The service standard often responds to actual performance.

Another important characteristic of service supply firms shown in Fig. 8 is that their capacity is particularly dependent on the performance of people, both in numbers and in productivity. While productivity is also a factor in manufacturing systems, the sensitivity of performance to people factors is generally less than in service-based firms. Therefore, models of such firms generally represent in some detail the factors that drive productivity, including: (1) skill and experience, often using an aging chain or “rookie-pro” structure [100,113]; (2) fatigue from sustained amounts of overtime; and (3) work intensity increasing productivity, but with “haste-makes-waste” impacts on errors and rework (not shown in Fig. 8). In these situations, when demand is growing there are considerable short-term forces which reduce productivity and cause a deterioration in service quality: adding people reduces productivity because of “experience dilution”; working overtime increases fatigue and reduces productivity; pressures to work more intensely increase errors and

cause additional work. As shown in Fig. 8, these productivity effects form reinforcing feedback loops which can drive down a service system's performance: an increase in work backlog and desired capacity causes the firm to hire more people; experience levels and productivity decline as a result, thereby reducing order fulfillment below what it otherwise would have been; order backlog does not fall as much as expected, necessitating additional capacity, further hires, and decreased experience; this completes the “experience dilution” R4 loop. The “burnout” loop through overtime and fatigue is similarly a reinforcing loop (R3).

Oliva [75] shows that how management responds to these work pressure problems can determine the long-term success or failure of a service-based organization, largely as a result of the third characteristic of such systems: performance of the system can be harder to detect, and so they are much more subject to a gradual erosion in performance and goals. Oliva demonstrates that if the firm reduces its service standards (goals) in response to deteriorating performance (loop R1 goal erosion), a death spiral can ensue in which the declining goals cause the firm to add fewer people, which locks in a situation of excessive work pressure and further declining performance (thereby



completing the “death spiral” loop R2). Unless there is a cyclical downturn in demand which alleviates the pressure, a firm’s service performance will gradually erode until competition captures the market. He further discusses solutions to these problems, including buffers and faster response. Oliva’s work, together with applications noted below, suggest that a service company should hire steadily rather than in spurts to avoid problems of inexperience, should hire enough workers to avoid overwork and a drift to low standards, and (in the case of equipment service) should give preventive maintenance high priority to avoid a spiral of equipment failures.

The resultant financial pressures engendered by the dynamics described above often drive service organizations to investments in process improvement and other cost containment initiatives to seek efficiency gains. Such investments, while offering perhaps the only long-term solution to remaining competitive, cause short-term workloads that further increase the demands on service personnel. This is demonstrated in the work of Repenning and Kaufman [83], and Repenning and Sterman [84,85].

Akkermans and Vos [2], and Anderson et al. [4] have studied the extent to which service industries have multi-stage supply chains similar to manufacturing industries, albeit with backlogs rather than inventories. Akkermans and Vos demonstrate that “inventory” cycles in service chains manifest themselves in terms of order backlog and workload cycles, and that while some of the causes of amplification existent in product supply chains apply to service supply chains (demand signaling and pricing), others, particularly those related to inventory management, do not (order batching, shortage gaming). They find that the real drivers of amplification in service supply chains come from the interactions of workloads, process quality, and rework. Because of delays in hiring and firing, capacity is slow to respond to changes, and is likely to exacerbate cycles. Anderson et al. [4] find that the bullwhip effect may or may not occur in service supply chains, depending on the policies used to manage each stage. However, when it does occur, they find that the systemic improvements that can often be achieved in physical supply chains by locally applied policies (e. g., reducing delay times and sharing information) do not have as many parallels in service chains. Instead service supply chains are characterized by numerous tradeoffs between improving local performance and improving system performance.

The modeling of service delivery has also had a long history in system dynamics, though the number of published works is more modest than in other areas. Much of the early work was in the area of health care and education [46]. Later works of note include models of Peo-

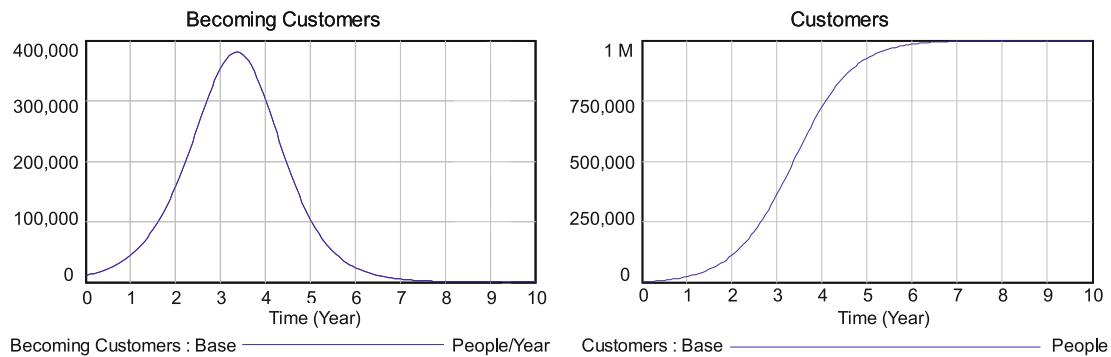
ple Express Airlines [98], Hanover Insurance claims processing [96]. NatWest Bank lending [74], and DuPont chemical plant equipment maintenance [12]. Homer [40] presents a case application for a major producer of equipment for semiconductor manufacturing that demonstrates many of the structures and policy issues enumerated above. These works incorporate the basic dynamic theory discussed above and illustrated in Fig. 8, and add another set of important structural theories to the building blocks for business strategy applications (note that the modeling of various effects on productivity is much more extensive in the area of project modeling, as discussed in [55]).

### Life Cycles of Products and Diffusion

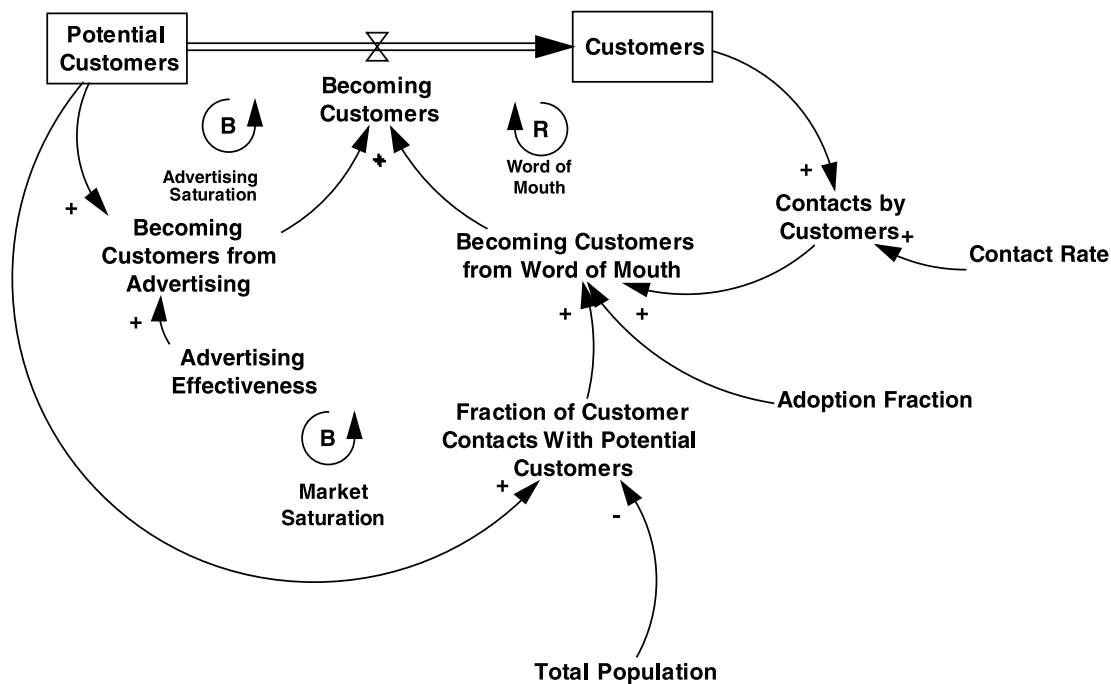
Another important pattern of behavior characteristic of many firms (or subsets of firms) is that of a life cycle (for the flow) and S-shaped pattern for the stock, as illustrated in Fig. 9: a gradual increase from a low level up to a peak, followed by a gradual decline either to zero or to some replacement level (sometimes referred to as “boom and bust” behavior). Sterman [100] and Oliva et al. [76] provide some real world examples of this behavior. The example shown is common for the sales of new products: the flow represents people becoming customers, and the stock, customers.

The structure which creates this “boom and bust” dynamics is shown in Fig. 10. In the marketing literature this structure is referred to as the “Bass Diffusion Model” after its original proponent [8]. The structure consists of three feedback loops: a reinforcing “word of mouth” loop that dominates behavior in the first half of customer sales growth; a balancing “market saturation” loop that constrains and eventually shuts down growth as the number of potential customers falls to zero; and another balancing loop “advertising saturation”, which represents other means of stimulating awareness, such as advertising, direct sales efforts, and media reports. These other channels are usually assumed to be proportional to the size of the pool of potential customers, and therefore initially stimulate the flow of “becoming customers” but then decline over time as the pool is depleted.

The dynamics of this structure, extensions to it (for example, loss of customers, competition, repeat sales), and policy implications are discussed in depth in Chap. 9 in [100] and Chap. 6 in [69]. This structure forms the basis of many system dynamics models that represent product sales, customer development, and the diffusion of innovations. Published examples include the work of Milling [60,61] and Maier [57] on the management of innovation diffusions and Oliva et al. [76] on boom and bust



Business Policy and Strategy, System Dynamics Applications to, Figure 9  
Life cycle behavior mode ("Boom and Bust")



Business Policy and Strategy, System Dynamics Applications to, Figure 10  
Basic structure generating boom and bust dynamics

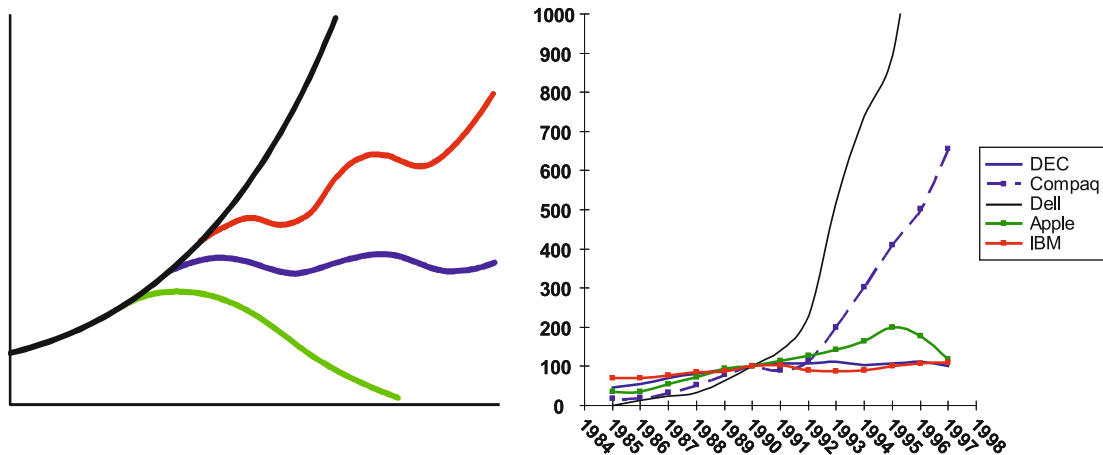
in e-commerce. Milling discusses the [Diffusion of Innovations, System Dynamics Analysis of the](#) in more depth.

Growth Dynamics

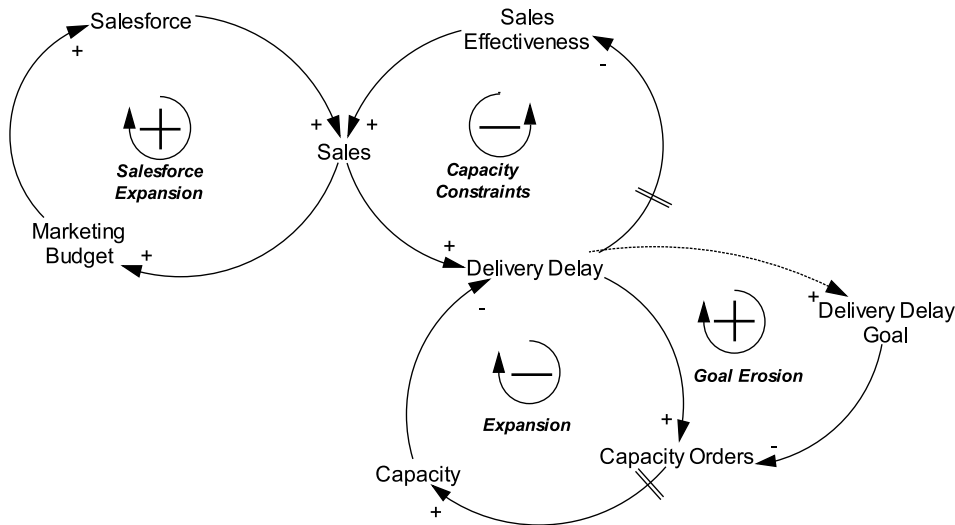
Growth is fundamentally a dynamic process, and therefore it is no surprise that since its early days system dynam- icists have shown an interest in the dynamics of corpo- rate growth. Forrester [27,28], Packer [77], Lyneis [48,49], Morecroft [67], Morecroft and Lane [70] and others stud- ied corporate growth in the field's early years. More re- cently, the People Express [98] and B&B ("Boom and

Bust") flight simulators [78] illustrate the field's interest in growth dynamics.

In his 1964 article, Forrester identified the range of possible growth patterns (see Fig. 11): smooth, steady growth; growth with repeated setbacks; stagnation; and decline. Examples of these patterns can be found in many real world industries, as illustrated for the computer in- dustry in Fig. 11 (see also [50] for examples). In his clas- sic article "Market Growth as Influenced by Capital In- vestment", Forrester detailed the three types of feedback loops which can create the range of possible growth pat- terns. These are illustrated in Fig. 11 (the equations for



Business Policy and Strategy, System Dynamics Applications to, Figure 11  
Stylized patterns of growth and examples from the computer hardware industry



Business Policy and Strategy, System Dynamics Applications to, Figure 12  
Feedback loops creating observed patterns of growth (adapted from Forrester 1968)

this model are provided in the original Forrester article; the model is also presented and discussed and analyzed in detail in Chap. 15 in [100], and Chap. 7 in [69]).

On the left in Fig. 12 is the reinforcing “salesforce expansion” loop: the salesforce generates sales, a portion of those sales are allocated to future marketing budgets, which allows an increase in the size of the salesforce and a further increase in sales. The salesforce expansion loop in isolation can create smooth growth forever (until the market is saturated). However, assuming a fixed capacity, the balancing “capacity constraints” loop activates: if sales exceed capacity, delivery delay increases such that, after a delay, sales effectiveness falls and sales decline. The

goal of the loop is to equate sales and capacity, and the two loops together can produce growth followed by stagnation (with fluctuations caused by delays in the balancing loop). In response to increasing delivery delay, however, firms often increase capacity (“capacity expansion” loop): when delivery delay exceeds the firm’s delivery delay goal, capacity orders increase, which after a delay increases capacity and thereby reduces delivery delay; the goal of this loop is to equate delivery delay to the firm’s delivery delay goal. However, once delivery delay is reduced, sales effectiveness and sales increase, thereby stimulating additional salesforce expansion, such that the growth with setbacks pattern of behavior can result. The final loop shown in

Fig. 12 is the reinforcing “goal erosion” loop: if the firm’s delivery delay goal responds to the actual delivery delay performance, a downward spiral can ensue – the goal increases, less expansion occurs than had before, capacity is less than needed, delivery delay increases, the goal is further increased, and so on (this loop is similar to the service standard goal erosion loop discussed above). The goal erosion loop can create the decline dynamics illustrated in Fig. 11 (although in actual practice the decline would likely occur over a much more extended period than shown).

In actual practice, there are numerous positive feedback loops through resources that might stimulate growth. These loops are listed below, and many are discussed and diagrammed in Chap. 10 in [100]. In each of these loops, an increase in sales causes management actions and/or investments that further increase the resource and sales:

- Sales channels – sales capability (which might include salesforce as discussed above, or retail stores), advertising, word-of-mouth contagion (as in the diffusion model), media hype (sales create media exposure which attracts potential customers and more sales)
- Price – Product attractiveness channels (operationalizing the link between resources and costs in Fig. 2 – spreading of fixed costs over more units, thereby lowering unit costs; learning curves; economies of scale; economies of scope; investments in process improvements)
- Market channels which increase the pool of potential customers – network effects (the more people using cell phones the greater their value), development of complementary goods (software applications for computers)
- Product investment channels – product improvement, new products
- Market power channels – over suppliers, over labor, over customers, cost of capital.

With all these positive feedback loops, how can anyone fail? In fact, there are also numerous constraints to growth, including: depletion of the pool of potential customers as discussed in the last section; growth of competition; delays in acquiring production capacity and/or service capacity; limits to financial capital (which can increase delays or limit acquiring productive assets); and increases in organizational size, complexity and administrative overheads (which might make the resources – costs loop in Fig. 2 revert to a positive connection, thereby constraining growth). Structures for representing these constraints are provided in the earlier references to this section, especially [49,69,100], and form the building blocks for many

of the practical applications of system dynamics to business growth strategy.

As system dynamicists have long recognized, managing growth is one of the more challenging management tasks. It entails fostering the positive, reinforcing feedback loops while simultaneously relaxing the constraining, negative feedback loops. While it is difficult to generalize without sounding platitudinous, a number of important lessons have emerged from studies of growth. Lyneis [50] discusses these in more detail, and references other work:

**Lesson 1** You won’t achieve what you don’t try for (if a firm is timid in its growth objectives, it will be timid in the acquisition of resources – balancing loops through for example delivery delay will then drive sales growth to the firm’s resource growth). Corollary 1: Don’t mistake forecasts for reality (a firm may be continually surprised by how accurate their sales forecasts are, because if resources are based on these forecasts, the balancing loops will drive sales to those resources). Corollary 2: Provide sufficient buffers and contingencies (these help minimize the risks that the balancing loops will become dominant).

**Lesson 2** Avoid the temptation to reduce objectives in the face of performance problems (the “goal erosion” loop in Figs. 8 and 12).

**Lesson 3** In a world of limited resources, something must limit growth. Proactively managing these limits is a key factor affecting performance. For example, if financial constraints are limiting expansion, with resultant delivery constraints on sales, why not increase prices to limit sales, and use the extra cash to finance expansion?

**Lesson 4** Make an effort to account for delays, especially in market response (for example, improvements in service will take a while to manifest themselves in improved sales, so avoid the temptation to cut back productive resources that will later be needed).

**Lesson 5** Account for short-term productivity losses such as fatigue and experience dilution in resource expansion decisions (in the short-term, you may be getting less capacity than you think).

**Lesson 6** Account for likely competitor responses in taking actions (it’s easy for competitors to follow price changes, and trigger a price war; improvements in other components of product attractiveness are harder to detect and replicate).

### Applications and Case Examples

The process and structural developments discussed in the last sections have formed the basis of numerous applications of system dynamics in support of business strategy.

In turn, these applications have provided the practice field through which the process has been refined, and the structural models, insights, and tools validated. While most of the published “real world” applications of system dynamics do not provide details of the models, they do nevertheless provide diagrams which show the nature of the model, representative results and policy conclusions, and the role the models played in business strategy formulation.

Space limitations preclude covering specific applications in depth. Therefore, I have chosen to reference applications in a number of different areas so readers can find references to the literature in their particular area of interest. Before getting to that, however, there are a couple of general references worthy of note: Roberts [91] covers many of the early published applications of system dynamics, with sections on manufacturing, marketing, research and development, and management and financial control. Coyle [14,15,16] touches on many of the applications initiated by his team at the University of Bradford, including work in the defense, natural resources, and utility industries. Richardson [86] provides an edited collection of academic journal articles containing some of the best work in system dynamics for business (and public) policy from its early years to the 1990s. Beyond these general compendiums, business strategy applications can perhaps best be classified by the industry of interest.

First, there have been a number of industry-wide models. The purpose of these models is typically to understand the drivers of change in the industry, and to forecast demand for use in other planning models. These include:

- Aircraft market as illustrated in Fig. 1 above [47,53]
- Health care market [42,103]
- Oil market [69,71,101]
- Shipping market [82].

Second, in addition to these industry-wide models, multiple applications to specific firms (which might also include some industry and/or competitive modeling), have been done in the following industry sectors:

- Utilities/Regulation – In the electric industry, work by Ford [24], Lyneis [51], Bunn and Larson [10,45] Ford covers some of this work elsewhere in ► [System Dynamics Models of Environment, Energy and Climate Change](#).
- Telecoms [35]
- Financial Services – work for MasterCard [52] and in the insurance industry [7,13,96,102].

Finally, applications to specific types of firms, particularly small and medium size enterprises [9].

For particular examples of where system dynamics has had an impact in changing or forming business strategy, the MasterCard application described by Lyneis [52] and the General Motors OnStar application described by Barabba et al. [6] are noteworthy. These applications provide some detail about the model structure and describe how the modeling process changed management intuition and thereby led to significant shifts in business strategy. The MasterCard model represents growth and competitive dynamics in some detail, and is used to illustrate the multi-stage development process detailed in Sect. [Using System Dynamics Models in Policy and Strategy Formulation and Implementation](#) above. The OnStar example describes the process of modeling an industry that does not yet exist. The model itself builds from the diffusion model discussed above, with significant elaboration of potential customers, provision of service, alliances, dealers, and financial performance. The paper details the significant role that the system dynamics model played in reshaping GM’s strategy.

### Future Directions

System dynamics has made significant theoretical and practical contributions to business strategy. These contributions fall into two general categories: first, the process through which models are developed, working with management teams to enhance model validity and implementation; and second, understanding of the business structures and policies which cause observed problem behavior within firms and industries. Nevertheless, the impact of system dynamics on business strategy has been modest – relatively few firms use system dynamics. In my view, several factors contribute to this slow uptake. These are listed below, with possible actions that could be taken to alleviate.

1. Knowledge of system dynamics and its potential in strategy formulation is limited. Senior executives of firm’s are unaware of system dynamics, or its potential [106]. In part this is because system dynamics is taught at relatively few business schools. Over time, this problem will be solved, but it will take a long time. But more importantly, researchers and practitioners of system dynamics rarely publish their work in publications which are widely read by senior management, such as the Harvard Business Review. This is a problem which can be addressed in the near future if researchers and practitioners made the effort to communicate with this market.
2. System dynamics is not well connected with traditional strategy research and practice. As noted earlier, system



dynamics and traditional strategy developed largely independently. While the potential synergies between the two are significant, until recently few researchers and practitioners have made the effort to cross disciplines. More effort and publication are needed to demonstrate areas of synergy and get system dynamics into the mainstream of business strategy research and ultimately practice. Warren [108] makes a start on this.

3. System dynamics is hard. Building system dynamics models, calibrating them to data, and analyzing their behavior to improve business strategy requires significant skill and experience. This has traditionally been developed via an apprenticeship program, either in university or consulting firms. Much more can be done to hasten this process. First, tried and true model structures that can be used as building blocks for models must be better documented, and gaps closed. While the theoretical basis for business dynamics described above is a good start, and reflects structures which all practicing system dynamicists should know, the underlying models and building blocks are widely dispersed and difficult to access. For example, the models in [100] are contained within a 900 page introductory textbook; the models in [49] are in old software and the book is out of print. In both cases, the models are only starting points and much unpublished work has occurred that expands these introductory models to make them more relevant and directly applicable to real business strategy problems. Efforts could and should be made to expand the library of business strategy building blocks. In addition, the process of building models and working with managers needs to be better documented and disseminated, both in formal courses and in published works, so as to facilitate the apprenticeship learning process.

## Bibliography

1. Akkermans HA, Daellaert N (2005) The Rediscovery of Industrial Dynamics: The Contribution of System Dynamics to Supply Chain Management in a Dynamic and Fragmented World. *Syst Dyn Rev* 21(3):173–186
2. Akkermans HA, Vos B (2003) Amplification in service supply chains: an exploratory case study from the telecom industry. *Prod Oper Manag* 12(2):204–223
3. Anderson E, Fine C (1999) Business Cycles and Productivity in Capital Equipment Supply Chains. In: Tayur et al(eds) *Quantitative Models for Supply Chain Management*. Kluwer Academic Publishers, Norwell
4. Anderson E, Morrice D, Lundeen G (2005) The ‘physics’ of capacity and backlog management in service and custom manufacturing supply chains. *Syst Dyn Rev* 21(3):187–216
5. Backus G, Schwein MT, Johnson ST, Walker RJ (2001) Comparing expectations to actual events: the post mortem of a Y2K analysis. *Syst Dyn Rev* 17(3):217–235
6. Barabba V, Huber C, Cooke F, Pudar N, Smith J, Paich M (2002) A Multimethod Approach for Creating New Business Models: The General Motors OnStar Project. *Interfaces* 32(1):20–34
7. Barlas Y, Cirak K, Duman E (2000) Dynamic simulation for strategic insurance management. *Syst Dyn Rev* 16(1):43–58
8. Bass FM (1969) New product growth model for consumer durables. *Manag Sci* 15:215–227
9. Bianchi C (2002) Editorial to Special Issue on Systems Thinking and System Dynamics in Small-Medium Enterprises. *Syst Dyn Rev* 18(3):311–314
10. Bunn DW, Larsen ER (eds) (1997) *Systems Modeling for Energy Policy*. Wiley, Chichester
11. Campbell D (2001) The long and winding (and frequently bumpy) road to successful client engagement: one team’s journey. *Syst Dyn Rev* 17(3):195–215
12. Carroll JS, Sterman JD, Marcus AA (1998) Playing the maintenance game: How mental models drive organizational decisions. In: Halpern JJ, Stern RN (eds) *Nonrational Elements of Organizational Decision Making*. Cornell University Press, Ithaca
13. Doman A, Glucksman M, Mass N, Sasportes M (1995) The dynamics of managing a life insurance Company. *Syst Dyn Rev* 11(3):219–232
14. Coyle RG (1996) *System Dynamics Modelling: A Practical Approach*. Chapman and Hall, London
15. Coyle RG (1997) System Dynamics at Bradford University: A Silver Jubilee Review. *Syst Dyn Rev* 13(4):311–321
16. Coyle RG (1998) The Practice of System Dynamics: Milestones, Lessons and Ideas From 30 Years Experience. *Syst Dyn Rev* 14(4):343–365
17. Coyle RG (2000) Qualitative and Quantitative Modeling in System Dynamics: Some Research Questions. *Syst Dyn Rev* 16(3):225–244
18. Coyle RG (2001) Rejoinder to Homer and Oliva. *Syst Dyn Rev* 17(4):357–363
19. Coyle RG, Morecroft JDW (1999) System Dynamics for Policy, Strategy and Management Education. *J Oper Res Soc* 50(4)
20. Croson R, Donohue K (2005) Upstream versus downstream information and its impact on the bullwhip effect. *Syst Dyn Rev* 21(3):187–216
21. Delaunzun F, Mollona E (1999) Introducing system dynamics to the BBC World Service: an insider perspective. *J Oper Res Soc* 50(4):364–371
22. Doman A, Glucksman M, Mass N, Sasportes M (1995) The dynamics of managing a life insurance company. *Syst Dyn Rev* 11(3):219–232
23. Dyson RG, Bryant J, Morecroft J, O’Brien F (2007) The Strategic Development Process. In: O’Brien FA, Dyson RG (eds) *Supporting Strategy*. Wiley, Chichester
24. Ford AJ (1997) System Dynamics and the Electric Power Industry. *Syst Dyn Rev* 13(1):57–85
25. Forrester JW (1958) Industrial dynamics: a major breakthrough for decision makers. *Harv Bus Rev* 36(4):37–66
26. Forrester JW (1961) *Industrial Dynamics*. MIT Press, Cambridge (now available from Pegasus Communications, Waltham)
27. Forrester JW (1964) Common Foundations Underlying Engineering and Management. *IEEE Spectr* 1(9):6–77
28. Forrester JW (1968) Market growth as influenced by capital investment. *Industrial Management Review* 9(2):83–105.

- Reprinted in Forrester JW (1975) *Collected Papers of Jay W Forrester*. Pegasus Communications, Waltham
29. Forrester JW (1969) *Urban Dynamics*. Pegasus Communications, Waltham
  30. Foss NJ (ed) (1997) *Resources, Firms and Strategies*. Oxford University Press, Oxford
  31. Gary MS (2005) Implementation Strategy and Performance Outcomes in Related Diversification. *Strateg Manag J* 26:643–664
  32. Gonçalves PM (2003) *Demand Bubbles and Phantom Orders in Supply Chains*. Unpublished Dissertation Sloan School of Management. MIT, Cambridge
  33. Gonçalves PM (2006) The Impact of Customer Response on Inventory and Utilization Policies. *J Bus Logist* 27(2):103–128
  34. Gonçalves P, Hines J, Sterman J (2005) The Impact of Endogenous Demand on Push-Pull Production Systems. *Syst Dyn Rev* 21(3):187–216
  35. Graham AK, Walker RJ (1998) *Strategy Modeling for Top Management: Going Beyond Modeling Orthodoxy at Bell Canada*. Proceedings of the 1998 International System Dynamics Conference, Quebec
  36. Grant RM (2005) *Contemporary Strategy Analysis*, 5th edn. Blackwell, Cambridge
  37. Hines JH, Johnson DW (1994) *Launching System Dynamics*. Proceedings of the 1994 International System Dynamics Conference, Business Decision-Making, Stirling
  38. Homer JB (1996) Why We Iterate: Scientific Modeling In Theory and Practice. *Syst Dyn Rev* 12(1):1–19
  39. Homer JB (1997) Structure, Data, and Compelling Conclusions: Notes from the Field. *Syst Dyn Rev* 13(4):293–309
  40. Homer JB (1999) Macro- and Micro-Modeling of Field Service Dynamics. *Syst Dyn Rev* 15(2):139–162
  41. Homer J, Oliva R (2001) Maps and Models in System Dynamics: A Response to Coyle. *Syst Dyn Rev* 17(4):347–355
  42. Homer J, Hirsch G, Minniti M, Pierson M (2004) Models for Collaboration: How System Dynamics Helped a Community Organize Cost-Effective Care For Chronic Illness. *Syst Dyn Rev* 20(3):199–222
  43. Kim DH, Lannon C (1997) *Applying Systems Archetypes*. Pegasus Communications Inc, Waltham
  44. Kunc M, Morecroft J (2007) System Dynamics Modelling for Strategic Development. In: O'Brien FA, Dyson RG (eds) *Supporting Strategy*. Wiley, Chichester
  45. Larsen ER, Bunn DW (1999) Deregulation in electricity: understanding strategic and regulatory risks. *J Oper Res Soc* 50(4):337–344
  46. Levin G, Roberts EB, Hirsch GB (1976) *The Dynamics of Human Service Delivery*. Ballinger, Cambridge
  47. Liehr M, Großler A, Klein M, Milling PM (2001) Cycles in the sky: understanding and managing business cycles in the airline market. *Syst Dyn Rev* 17(4):311–332
  48. Lyneis JM (1975) Designing Financial Policies to Deal With Limited Financial Resources. *Financ Manag* 4(1)
  49. Lyneis JM (1980) *Corporate Planning and Policy Design: A System Dynamics Approach*. M.I.T. Press, Cambridge
  50. Lyneis JM (1998) Learning to Manage Growth: Lessons From a Management Flight Simulator, Proceedings of the 1998 International System Dynamics Conference (Plenary Session), Quebec City, 1998
  51. Lyneis JM (1997) Preparing for a Competitive Environment: Developing Strategies for America's Electric Utilities. In: Bunn DW, Larsen ER (eds) *Systems Modeling for Energy Policy*. Wiley, Chichester
  52. Lyneis JM (1999) System dynamics for business strategy: a phased approach. *Syst Dyn Rev* 15(1):37–70
  53. Lyneis JM (2000) System Dynamics for Market Forecasting and Structural Analysis. *Syst Dyn Rev* 16(1):3–25
  54. Lyneis JM, Cooper KG, Els SA (2001) Strategic Management of Complex Projects: A Case Study Using System Dynamics. *Syst Dyn Rev* 17(3):237–260
  55. Lyneis JM, Ford DN (2007) System Dynamics Applied to Project Management: A Survey, Assessment, and Directions for Future Research. *Syst Dyn Rev* 23(2/3):157–189
  56. Magee J (1958) *Production Planning and Inventory Control*. McGraw-Hill, London
  57. Maier FH (1998) New product diffusion models in innovation management – a system dynamics perspective. *Syst Dyn Rev* 14(4):285–308
  58. Meadows DL (1970) *Dynamics of commodity production cycles*. Pegasus Communications, Waltham
  59. Merten PP, Löffler R, Wiedmann KP (1987) *Portfolio Simulation: A Tool to Support Strategic Management*. *Syst Dyn Rev* 3(2):81–101
  60. Milling PM (1996) Modeling innovation processes for decision support and management Simulation. *Syst Dyn Rev* 12(3):211–23
  61. Milling PM (2002) Understanding and managing innovation processes. *Syst Dyn Rev* 18(1):73–86
  62. Milling PM (2007) A Brief History of System Dynamics in Continental Europe. *Syst Dyn Rev* 23(2/3):205–214
  63. Montgomery CA (ed) (1995) *Resource-Based and Evolutionary Theories of the Firm*. Kluwer, Boston
  64. Morecroft JDW (1983) A Systems Perspective on Material Requirements Planning. *Decis Sci* 14(1):1–18
  65. Morecroft JDW (1984) *Strategy Support Models*. *Strateg Manag J* 5:215–229
  66. Morecroft JDW (1985) The Feedback View of Business Policy and Strategy. *Syst Dyn Rev* 1(1):4–19
  67. Morecroft JDW (1986) The Dynamics of a Fledgling High-Technology Growth Market. *Syst Dyn Rev* 2(1):36–61
  68. Morecroft JDW (1999) Management attitudes, learning and scale in successful diversification: a dynamic and behavioural resource system view. *J Oper Res Soc* 50(4):315–336
  69. Morecroft JDW (2007) *Strategic Modelling and Business Dynamics: A Feedback Systems View*. Wiley, Chichester
  70. Morecroft JDW, Lane DC, Viita PS (1991) Modeling growth strategy in a biotechnology startup firm. *Syst Dyn Rev* 7(2):93–116
  71. Morecroft JDW, van der Heijden KAJM (1994) Modeling the Oil Producers: Capturing Oil Industry Knowledge in a Behavioral Simulation Model. In: Morecroft JDW, Sterman JD (eds) *Modeling for Learning Organizations*. Productivity Press, Portland
  72. Morecroft JDW, Wolstenholme E (2007) System Dynamics in the UK: A Journey from Stirling to Oxford and Beyond. *Syst Dyn Rev* 23(2/3):205–214
  73. O'Brien FA, Dyson RG (eds) (2007) *Supporting Strategy*. Wiley, Chichester
  74. Oliva R (1996) *A Dynamic Theory of Service Delivery: Implications for Managing Service Quality*. Ph.D. Thesis, Sloan School of Management, Massachusetts Institute of Technology

75. Oliva R (2001) Tradeoffs in responses to work pressure in the service industry. *Calif Manag Rev* 43(4):26–43
76. Oliva R, Sterman JD, Giese M (2003) Limits to growth in the new economy: exploring the ‘get big fast’ strategy in e-commerce. *Syst Dyn Rev* 19(2):83–117
77. Packer DW (1964) *Resource Acquisition in Corporate Growth*. M.I.T. Press, Cambridge
78. Paich M, Sterman JD (1993) Boom, Bust, and Failures to Learn in Experimental Markets. *Manag Sci* 39(12)
79. Porter ME (1980) *Competitive Strategy*. The Free Press, New York
80. Porter ME (1985) *Competitive Advantage*. The Free Press, New York
81. Prahalad CK, Hamel G (1990) The Core Competence of the Corporation. *Harv Bus Rev* (May–June):71–91
82. Randers J, Gölluke U (2007) forecasting turning points in shipping freight rates: lessons from 30 years of practical effort. *Syst Dyn Rev* 23(2/3):253–284
83. Repenning N, Kofmann F (1997) Unanticipated Side Effects of Successful Quality Programs: Exploring a Paradox of Organizational Improvement. *Manag Sci* 43(4)
84. Repenning NP, Sterman JD (2001) Nobody Ever Gets Credit for Fixing Problems that Never Happened: Creating and Sustaining Process Improvement. *Calif Manag Rev* 43(4):64–88
85. Repenning NP, Sterman JD (2002) Capability traps and self-confirming attribution errors in the dynamics of process improvement. *Adm Sci Q* 47:265–295
86. Richardson GP (1996) *Modelling for Management: Simulation in Support of Systems Thinking*. Dartmouth Publishing Company, Aldershot
87. Richardson GP, Andersen DF (1995) Teamwork in Group Model Building. *Syst Dyn Rev* 11(2):113–138
88. Richmond B (1997) The Strategic Forum: Aligning Objectives, Strategy and Process. *Syst Dyn Rev* 13(2):131–148
89. Risch J, Troyano-Bermúdez L, Sterman J (1995) Designing Corporate Strategy with System Dynamics: A Case Study in the Pulp and Paper Industry. *Syst Dyn Rev* 11(4):249–274
90. Roberts EB (1977) Strategies for the Effective Implementation of Complex Corporate Models. *Interfaces* 8(1):26–33
91. Roberts EB (ed) (1978) *Managerial Applications of System Dynamics*. The MIT Press, Cambridge
92. Roberts EB (2007) Making System Dynamics Useful: A Personal Memoir. *Syst Dyn Rev* 23(2/3):119–136
93. Rockart SF, Lenox MJ, Lewin AY (2007) Interdependency, Competition, and Industry Dynamics. *Manag Sci* 53(4):599–615
94. Snabe B, Grossler A (2006) System Dynamics Modelling for Strategy Implementation—Case Study and Issues. *Syst Res Behav Sci* 23:467–481
95. Senge PM (1990) *The Fifth Discipline: The Art and Practice of The Learning Organization*. Doubleday, New York
96. Senge PM, Sterman JD (1992) Systems thinking and organizational learning: Acting locally and thinking globally in the organization of the future. *Eur J Oper Res* 59(1):137–150
97. Sice P, Mosekilde E, Moscardini A, Lawler K, French I (2000) Using system dynamics to analyse interactions in duopoly competition. *Syst Dyn Rev* 16(2):113–133
98. Sterman JD (1989) Modeling management behavior: misperceptions of feedback in a dynamic decision making experiment. *Manag Sci* 35:321–339
99. Sterman JD (1989) *Strategy Dynamics: the Rise and Fall of People Express*. Memorandum D-3939-1 Sloan School of Management, M.I.T
100. Sterman JD (2000) *Business Dynamics: Systems Thinking and Modeling for a Complex World*. McGraw-Hill, New York
101. Sterman JD, Richardson G, Davidsen P (1988) Modeling the estimation of petroleum resources in the United States. *Technol Forecast Soc Chang* 33(3):219–249
102. Thompson JP (1999) Consulting approaches with system dynamics: three case studies. *Syst Dyn Rev* 15(1):71–95
103. Thompson JP (2006) Making sense of US health care system dynamics. *Proceedings of the 2006 International System Dynamics Conference*, Nijmegen, The Netherlands
104. Vennix JAM (1996) *Group Model Building: Facilitating Team Learning Using System Dynamics Field*. Wiley, Chichester
105. Warren KD (2002) *Competitive Strategy Dynamics*. Wiley, Chichester
106. Warren KD (2004) Why Has Feedback Systems Thinking Struggled to Influence Strategy and Policy Formulation? Suggestive Evidence, Explanations and Solutions. *Syst Res Behav Sci* 21:331–347
107. Warren KD (2005) Improving Strategic Management With the Fundamental Principles of System Dynamics. *Syst Dyn Rev* 21(4):329–350
108. Warren KD (2007) *Strategic Management Dynamics*. Wiley, Chichester
109. Weil HB (1980) The Evolution of an Approach for Achieving Implemented Results from System Dynamics Projects. In: Randers J (ed) *Elements of the System Dynamics Method*. MIT Press, Cambridge
110. Wolstenholme EF (1999) Qualitative vs. Quantitative Modelling: The Evolving Balance. *J Oper Res Soc* 50(4):422–428
111. Winch GW (1993) Consensus building in the planning process: benefits from a “hard” modeling approach. *Syst Dyn Rev* 9(3):287–300
112. Winch GW (1999) Dynamic Visioning for Dynamic Environments. *J Oper Res Soc* 50(4):354–361
113. Winch GW (2001) Management of the “skills inventory” in times of major change. *Syst Dyn Rev* 17(2):151–159
114. Zahn E, Dillerup R, Schmid U (1998) Strategic evaluation of flexible assembly systems on the basis of hard and soft decision criteria. *Syst Dyn Rev* 14(4):263–284