



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m12e/2025/09.19.12.14-TDI

**ESTUDO COMPARATIVO ENTRE MÉTODOS DE
PRODUÇÃO DE MAPAS DE INTENSIDADE NA
MITIGAÇÃO DO RUÍDO 1/F PRESENTE EM
OBSERVAÇÕES DO SINAL DE 21 CM DO HI**

Helissa Helen da Costa

Dissertação de Mestrado do Curso de Pós-Graduação em Astrofísica, orientada pelos Drs. Carlos Alexandre Wuensche de Souza, e Eduardo Jubini de Merícia, aprovada em 11 de setembro de 2025.

URL do documento original:

<<http://urlib.net/8JMKD2USNRW34T/4E9H9N5>>

INPE
São José dos Campos
2025

PUBLICADO POR:

Instituto Nacional de Pesquisas Espaciais - INPE
Coordenação de Ensino, Pesquisa e Extensão (COEPE)
Divisão de Biblioteca (DIBIB)
CEP 12.227-010
São José dos Campos - SP - Brasil
Tel.:(012) 3208-6923/7348
E-mail: pubtc@inpe.br

**CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO DA PRODUÇÃO
INTELECTUAL DO INPE - CEPPII (PORTARIA Nº 176/2018/SEI-
INPE):**

Presidente:

Dr. Thales Sehn Korting - Coordenação-Geral de Ciências da Terra (CGCT)

Membros:

Dr. Antonio Fernando Bertachini de Almeida Prado - Conselho de Pós-Graduação (CPG)

Dr. Evandro Marconi Rocco - Coordenação-Geral de Engenharia, Tecnologia e Ciência Espaciais (CGCE)

Dr. Heyder Hey - Coordenação-Geral de Infraestrutura e Pesquisas Aplicadas (CGIP)

Simone Angélica Del Ducca Barbedo - Divisão de Biblioteca (DIBIB)

BIBLIOTECA DIGITAL:

Dr. Gerald Jean Francis Banon

Clayton Martins Pereira - Divisão de Biblioteca (DIBIB)

REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:

Simone Angélica Del Ducca Barbedo - Divisão de Biblioteca (DIBIB)

André Luis Dias Fernandes - Divisão de Biblioteca (DIBIB)

EDITORAÇÃO ELETRÔNICA:

Ivone Martins - Divisão de Biblioteca (DIBIB)

André Luis Dias Fernandes - Divisão de Biblioteca (DIBIB)



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m12e/2025/09.19.12.14-TDI

**ESTUDO COMPARATIVO ENTRE MÉTODOS DE
PRODUÇÃO DE MAPAS DE INTENSIDADE NA
MITIGAÇÃO DO RUÍDO 1/F PRESENTE EM
OBSERVAÇÕES DO SINAL DE 21 CM DO HI**

Helissa Helen da Costa

Dissertação de Mestrado do Curso de Pós-Graduação em Astrofísica, orientada pelos Drs. Carlos Alexandre Wuensche de Souza, e Eduardo Jubini de Merícia, aprovada em 11 de setembro de 2025.

URL do documento original:

<<http://urlib.net/8JMKD2USNRW34T/4E9H9N5>>

INPE
São José dos Campos
2025

Dados Internacionais de Catalogação na Publicação (CIP)

Costa, Helissa Helen da.

C823e Estudo comparativo entre métodos de produção de mapas de intensidade na mitigação do ruído 1/F presente em observações do sinal de 21 cm do HI / Helissa Helen da Costa. – São José dos Campos : INPE, 2025.

xx + 138 p. ; (sid.inpe.br/mtc-m12e/2025/09.19.12.14-TDI)

Dissertação (Mestrado em Astrofísica) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2025.

Orientadores : Drs. Carlos Alexandre Wuensche de Souza, e Eduardo Jubini de Merícia.

1. Oscilações acústicas de bárions. 2. Mapeamento de intensidade. 3. Linha de 21 cm do HI. 4. Ruído 1/f. 5. Separação de componentes. I.Título.

CDU 523.03



Esta obra foi licenciada sob uma Licença [Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada](https://creativecommons.org/licenses/by-nc/3.0/).

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](https://creativecommons.org/licenses/by-nc/3.0/).



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÃO



INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

FOLHA DE APROVAÇÃO

Aluno(a): Helissa Helen da Costa

Título: "ESTUDO COMPARATIVO ENTRE MÉTODOS DE PRODUÇÃO DE MAPAS DE INTENSIDADE NA MITIGAÇÃO DO RUÍDO 1/F PRESENTE EM OBSERVAÇÕES DO SINAL DE 21 CM DO H I"

Aprovado(a) pela Banca Examinadora em cumprimento ao requisito exigido para obtenção do Título Mestra (a) em Astrofísica.

(assinado eletronicamente)

Dr. Odylio Denys de Aguiar - Membro Interno - Presidente (PGAST/INPE & DIAST/INPE)
() Aprovado () Reprovado

(assinado eletronicamente)

Dr. José Carlos Neves de Araújo - Membro Interno (PGAST/INPE & DIAST/INPE)
() Aprovado () Reprovado

(assinado eletronicamente)

Dr. Carlos Alexandre Wuensche De Souza - Orientador (PGAST/INPE & DIAST/INPE)
() Aprovado () Reprovado

(assinado eletronicamente)

Dr. Eduardo Jubini de Merícia - Orientador (DIAST/INPE)
() Aprovado () Reprovado

(assinado eletronicamente)

Dr. Ribamar Rondon Rezende dos Reis - Membro Externo (IF-UFRJ)
() Aprovado () Reprovado

(assinado eletronicamente)

Dra. Larissa Carlos de Oliveira Santos - Membro Externo (Yangzhou University, Chi-na)

Aprovado Reprovado

Este trabalho foi aprovado por:

maioria simples

unanimidade

São José dos Campos, 11 de setembro de 2025



Documento assinado eletronicamente por **Odylio Denys de Aguiar, Pesquisador**, em 09/01/2026, às 12:58 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **José Carlos Neves de Araújo, Pesquisador**, em 09/01/2026, às 13:16 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ribamar Rondon de Rezende dos Reis, Usuário Externo**, em 09/01/2026, às 13:27 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Carlos Alexandre Wuensche de Souza, Pesquisador**, em 09/01/2026, às 14:20 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Eduardo Jubini de Merícia, Usuário Externo**, em 10/01/2026, às 22:23 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Larissa Carlos de Oliveira Santos (E), Usuário Externo**, em 26/02/2026, às 15:26 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site <https://sei.mcti.gov.br/verifica.html>, informando o código verificador **13423837** e o código CRC **BF0363A8**.

“ É um negócio perigoso, Frodo, sair da sua porta. Você pisa na estrada, e, se não controlar seus pés, não há como saber até onde você pode ser levado.”

- J.R.R. Tolkien

AGRADECIMENTOS

Agradeço o programa de pós-graduação em astrofísica do INPE pela oportunidade de realização deste trabalho e à CAPES pelo apoio financeiro. Agradeço também a todos os professores e pesquisadores que me ensinaram e me permitiram chegar neste ponto da minha formação acadêmica.

Agradeço especialmente ao meu orientador Dr. Carlos Alexandre, por todo o conhecimento compartilhado e por confiar na minha capacidade e ao meu orientador Dr. Eduardo Merícia, por toda a paciência, dedicação, pelas longas horas de conversa, por todo o esforço empregado na realização deste trabalho e também por ter sido um grande amigo.

Agradeço também ao Dr. Giancarlo de Gasperis, por gentilmente disponibilizar as rotinas computacionais utilizadas neste trabalho, pelo auxílio oferecido e pelas dúvidas respondidas.

Agradeço à minha família que nunca deixou de me apoiar, especialmente ao meu noivo Bruno Bizarria, por não me deixar desacreditar do meu potencial e por estar sempre ao meu lado. Agradeço aos meus amigos por todas as conversas e desa-bafos, especialmente aos amigos do Grupo de Cosmologia Observacional e da pós-graduação.

Agradeço à Deus por me ensinar tanto diante das adversidades.

RESUMO

As oscilações acústicas de bárions (BAO) constituem uma das principais ferramentas para investigar os efeitos da energia escura na evolução cósmica. Elas produzem padrões característicos na distribuição da matéria, que funcionam como uma régua padrão em cosmologia. A detecção das BAO na faixa de rádio pode ser realizada por meio do mapeamento de intensidade (IM) da linha de 21 cm do hidrogênio neutro (HI). O radiotelescópio BINGO (*BAO from Integrated Neutral Gas Observations*), atualmente em construção em Aguiar (PB) sob liderança de instituições brasileiras, tem como objetivo detectar as BAO através da distribuição de HI ao longo do tempo cósmico. Um dos principais desafios do BINGO é lidar com as variações de ganho dos amplificadores, conhecidas como ruído $1/f$. Esse ruído apresenta correlação temporal e espectral, afetando a qualidade dos dados e dificultando a recuperação do sinal cosmológico. Neste trabalho investigamos dois métodos de produção de mapas (*naive* e *iterative Generalized Least Squares* – iGLS) e dois métodos de separação de componentes (*Principal Component Analysis* – PCA e *Generalized Needlet Internal Linear Combination* – GNILC) aplicados a simulações do BINGO. Os resultados indicam que a presença do ruído $1/f$ impossibilita a recuperação adequada do sinal de HI, independentemente do método de *map-making* empregado. O iGLS apresentou desempenho inferior ao esperado, enquanto o método *naive* produziu os melhores mapas. Entre as técnicas de separação, o GNILC mostrou resultados superiores. Entretanto, a ausência de varreduras com *cross-linking*, a complexidade da matriz de covariância e a suavidade espectral do ruído reduziram a eficácia das técnicas testadas. Concluimos que o ruído $1/f$ com correlação espectral compromete a recuperação do sinal de 21 cm nas simulações do BINGO. Ademais, o iGLS apresenta limitações em instrumentos operando em *drift-scan*, e o GNILC encontra dificuldades para remover ruídos espectralmente suaves. Para que o BINGO alcance seu objetivo científico, serão necessárias estratégias adicionais de calibração e controle de ganho, como o uso de correlacionadores.

Palavras-chave: Oscilações acústicas de bárions. Mapeamento de intensidade. Linha de 21 cm do HI. Ruído $1/f$. Separação de componentes.

COMPARATIVE STUDY OF MAP-MAKING METHODS FOR MITIGATING $1/f$ NOISE IN 21 CM HI SIGNAL OBSERVATIONS

ABSTRACT

Baryonic Acoustic Oscillations (BAO) are a key tool to probe the effects of dark energy on cosmic evolution. They leave characteristic patterns in the matter distribution, which act as a standard ruler in cosmology. The detection of BAO in the radio band can be achieved through intensity mapping (IM) of the 21 cm line of neutral hydrogen (HI). The BINGO telescope (BAO from Integrated Neutral Gas Observations), currently under construction in Aguiar (PB) and led by Brazilian institutions, aims to detect BAO from the distribution of HI across cosmic time. A major challenge for BINGO is dealing with gain variations in the amplifiers, known as $1/f$ noise. This noise shows temporal and spectral correlations, degrading data quality and hindering the recovery of the cosmological signal. In this work, we investigate two map-making methods (*naive* and *iterative Generalized Least Squares* – iGLS) and two component-separation techniques (*Principal Component Analysis* – PCA and *Generalized Needlet Internal Linear Combination* – GNILC) applied to BINGO simulations. Our results indicate that the presence of $1/f$ noise prevents an accurate recovery of the HI signal, regardless of the map-making method. The iGLS showed a lower performance than expected, while the *naive* method produced the best maps. Among separation techniques, GNILC yielded the best results. However, the lack of cross-linking scans, the complexity of the covariance matrix, and the spectral smoothness of the noise limited the efficiency of the tested methods. We conclude that spectrally correlated $1/f$ noise severely compromises the recovery of the 21 cm HI signal in BINGO simulations. Furthermore, iGLS shows limitations for drift-scan instruments, and GNILC struggles to remove spectrally smooth noise. For BINGO to achieve its scientific goals, additional calibration strategies and gain-control elements, such as correlators, will be required.

Keywords: Baryonic Acoustic Oscillations (BAO). Intensity Mapping (IM). 21 cm HI line. $1/f$ noise. Component separation.

LISTA DE FIGURAS

	<u>Pág.</u>
2.1 Ilustração dos anéis de concentração de matéria que determinam as regiões de formação de estruturas.	10
2.2 Gráfico da função de correlação de dois pontos para distancia comóvel entre galaxias.	11
2.3 Comparação entre a observação de um mesmo campo para medidas pontuais e para o mapeamento de intensidade.	13
3.1 Projeto de engenharia do arranjo do sistema óptico do radiotelescópio BINGO.	28
3.2 Representação artística do arranjo do sistema óptico do radiotelescópio BINGO.	29
3.3 Concepção artística de uma das cornetas do BINGO	30
3.4 Representação do plano focal do BINGO com 28 cornetas	31
3.5 Representação da cadeia receptora de correlação do radiotelescópio BINGO.	32
3.6 Séries temporais produzidas com a IM <i>pipeline</i> para os canais de frequência 0, 14 e 29 da corneta 1, com $f_k = 0,001$ Hz e $\beta = \{0, 0,25, 0,50, 0,75, 1,0\}$, em um trecho de duas horas da observação total.	38
4.1 A ferramenta de <i>map-making</i> como uma etapa intermediária na determinação de parâmetros cosmológicos.	40
5.1 Arranjo original (AO) com 28 cornetas no plano focal para a Fase 1 do BINGO.	65
5.2 Arranjo original (AO) com 14 cornetas no plano focal para a Fase 1 do BINGO.	66
5.3 Arranjo compacto inferior (ACI) com 14 cornetas no plano focal para a Fase 1 do BINGO.	67
6.1 Matrizes de correlação das simulações com presença do sinal de HI e do ruído $1/f$	84
6.2 Comparação da estrutura espectral do ruído $1/f$ para $\beta = 0,1$ e $\beta = 0,25$, com a estrutura do HI.	86

6.3	Mapas de temperatura em projeção cartesiana, simulados no canal 14, para os arranjos de cornetas original (AO; superior) e compacto inferior (ACI; inferior), com $N_{side} = 128$ e gerados a partir de 3 dias de observação contendo HI e ruído $1/f$ com $\beta = 0,25$	88
6.4	Mapas de temperatura em projeção gnomônica, simulados no canal 14, para os arranjos de cornetas original (AO; esquerda) e compacto inferior (ACI; direita), com $N_{side} = 128$ e gerados a partir de 3 dias de observação contendo HI e ruído $1/f$ com $\beta = 0,25$	89
6.5	Séries temporais produzidas com a IM <i>pipeline</i> para os canais de frequência 0, 14 e 29 da corneta 1, com $f_k = 0,001$ Hz e $\beta = 0,25$ em um intervalo de duas horas da observação total.	91
6.6	Mapas <i>naive</i> da Simulação 1 para o canal 0, após diferentes etapas de pré-processamento.	93
6.7	Mapas <i>naive</i> da Simulação 1 para o canal 14, após diferentes etapas de pré-processamento.	94
6.8	Mapas <i>naive</i> da Simulação 1 para o canal 29, após diferentes etapas de pré-processamento.	95
6.9	Histogramas normalizados pelo <i>z-score</i> , gerados a partir dos mapas de observação simulados do BINGO, ilustrando as diferentes etapas do pré-processamento dos dados da Simulação 1.	96
6.10	Mapa <i>prior</i> de HI usado na Simulação 1.	99
6.11	Mapa de entrada de HI (<i>target</i>) usado na Simulação 1.	99
6.12	Mapa de entrada de HI (superior; com feixe ajustado), mapa de HI reconstruído pelo GNILC (central, <i>naive</i>) e mapa de resíduo (inferior, HI entrada - HI reconstruído pelo GNILC) do canal 0 da Simulação 1 do BINGO.	100
6.13	Mapa de entrada de HI (superior), mapa de HI reconstruído pelo GNILC (central, <i>naive</i>) e mapa de resíduo (inferior, HI entrada - HI reconstruído pelo GNILC) do canal 14 da Simulação 1 do BINGO.	101
6.14	Mapa de entrada de HI (superior; com feixe ajustado), mapa de HI reconstruído pelo GNILC (central, <i>naive</i>) e mapa de resíduo (inferior, HI entrada - HI reconstruído pelo GNILC) do canal 29 da Simulação 1 do BINGO.	102
6.15	Histogramas dos mapas de entrada de HI (linha superior), dos mapas de HI reconstruídos pelo GNILC (<i>naive</i> ; linha central) e dos mapas de resíduo, que são a diferença entre o mapa de HI de entrada e o mapa de HI reconstruído pelo GNILC (<i>naive</i> ; linha inferior).	103

6.16	Coeficiente de Pearson para a Simulação 1, realizada com <i>map-making naive</i> e separação de componentes feita pelo GNILC, calculados para os três canais de frequência.	105
6.17	Coeficiente de Pearson calculado para cada um dos 30 canais e coeficiente de Pearson médio para a Simulação 1, realizada com <i>map-making naive</i> e separação de componentes feita pelo GNILC, calculado em todos os canais.	106
6.18	Comparação entre o espectro de potência angular calculado para o mapa real de HI e para o mapa recuperado com o GNILC.	107
6.19	Mapas recuperados de HI para a simulação 1, realizada com <i>map-making naive</i> para o canal 0 após aplicação do PCA e retirada de 7 à 11 modos.	108
6.20	Mapas recuperados de HI para a simulação 1, realizada com <i>map-making naive</i> para o canal 14 após aplicação do PCA e retirada de 7 à 11 modos.	109
6.21	Mapas recuperados de HI para a simulação 1, realizada com <i>map-making naive</i> para o canal 29 após aplicação do PCA e retirada de 7 à 11 modos.	110
6.22	Espectros de potência angular dos mapas de HI recuperados para a Simulação 1, obtidos com o método <i>naive</i> de <i>map-making</i> após a remoção de 7 a 11 modos com o PCA.	111
6.23	Erro médio dos espectros de potência angular (η_{abs}) para os mapas de HI recuperados na Simulação 1, obtidos com o métodos <i>naive</i> de <i>map-making</i> e PCA de separação de componentes.	112
6.24	Comparação entre o erro médio dos C_ℓ dos mapas recuperados de HI, para a Simulação 1, produzidos com <i>map-making naive</i> e recuperados com o GNILC (curva amarela) e com o PCA para a retirada de 8 modos (curva azul) e 9 modos (curva verde).	113
6.25	Mapa de HI de entrada (superior) e mapa reconstruído pelo GNILC (inferior) a partir do mapa observado (HI + 1/f; Simulação 1), obtido com o iGLS, para os canais 0, 14 e 29	115
6.26	Espectros de potência angular para o mapa real de HI e para o mapa iGLS recuperado, obtidos com <i>rebeaming</i> no espaço de pixels e separação de componentes realizada pelo GNILC.	116
6.27	Comparação entre os espectros de potência angular do sinal para o mapa real de HI, para o mapa iGLS recuperado e para o mapa <i>naive</i> recuperado, com <i>rebeaming</i> realizado no espaço de pixels e separação de componentes realizada com o GNILC.	117
6.28	Coeficiente de correlação de Pearson obtido para o mapa iGLS recuperado quando comparado ao mapa real de HI, apresentado nos canais 0, 14 e 29.	118

6.29	Coeficiente de Pearson por canal e coeficiente de Pearson médio para o mapa iGLS e para o mapa <i>naive</i> , com <i>rebeamng</i> no espaço de pixels e separação de componentes realizada pelo GNILC.	119
6.30	Erros de reconstrução dos C_ℓ 's obtidos para o mapa <i>naive</i> e para o mapa iGLS quando comparados ao mapa de HI real.	120
6.31	Erros de reconstrução do espectro de potência do HI para os mapas <i>naive</i> gerados com $N_{\text{side}} = 128$ e $N_{\text{side}} = 512$ com ajuste de feixe no espaço de <i>pixels</i> e no espaço harmônico.	124
6.32	Erros de recuperação dos espectros de potência angular do HI para os mapas iGLS gerados com $N_{\text{side}} = 128$ e $N_{\text{side}} = 512$ com ajuste de feixe no espaço de <i>pixels</i> e no espaço harmônico.	125

LISTA DE TABELAS

	<u>Pág.</u>
3.1 Parâmetros do projeto do BINGO - Fase 1.	27
3.2 Contribuições para T_{sys}	34
5.1 Parâmetros adotados para as simulações de missão do radiotelescópio BINGO.	68
5.2 Tamanho angular do feixe do BINGO para cada canal de frequências. . .	70
6.1 Sumário das simulações realizadas e seus parâmetros.	87
6.2 Desvio padrão (K) dos mapas simulados para três canais de frequência distintos em quatro etapas de pré-processamento.	97
6.3 Desvio padrão (σ) dos mapas normalizados por <i>z-score</i> nos canais 0, 14 e 29.	102

SUMÁRIO

	<u>Pág.</u>
1 INTRODUÇÃO	1
2 REFERENCIAL TEÓRICO	5
2.1 O modelo cosmológico padrão	7
2.2 Oscilações acústicas de bárions	9
2.3 Mapeamento de intensidade em linhas	12
2.4 Física da emissão de 21 cm do HI neutro	14
2.5 Contaminantes astrofísicos	18
3 INSTRUMENTAÇÃO	23
3.1 Radiotelescópios	24
3.2 Radiotelescópio BINGO	25
3.3 Ruídos instrumentais	30
3.3.1 Ruído térmico	32
3.3.2 Ruído 1/f	34
4 PRODUÇÃO DE MAPAS E SEPARAÇÃO DE COMPONENTES 39	
4.1 Métodos e fundamentos de <i>map-making</i>	41
4.1.1 Método <i>Naive</i>	46
4.1.2 Método <i>iterative Generalized Least Squares</i> (iGLS)	48
4.2 Ferramentas de simulação	54
4.2.1 IM <i>pipeline</i>	54
4.2.2 ROMA - Roma Optimal Map-making Algorithm	56
4.3 Métodos de separação de componentes	58
4.3.1 PCA	60
4.3.2 GNILC	61
5 PREPARAÇÃO DOS MAPAS E FERRAMENTAS DE COM- PARAÇÃO	63
5.1 Parâmetros das simulações	63
5.2 Arranjo do sistema óptico	64
5.3 Outros parâmetros usados nas simulações	64
5.4 Preparação dos mapas	68

5.5	Ajuste de feixe	72
5.5.1	Ajuste de feixe no espaço harmônico	72
5.5.2	Ajuste de feixe no espaço de pixels	75
5.6	Ferramentas de comparação	76
5.6.1	Mapas de intensidade	76
5.6.2	Ferramentas estatísticas gerais	77
5.6.3	Espectro de potência angular	78
5.6.4	Dispersão e coeficiente de Pearson	79
5.6.5	Erro de reconstrução do espectro	80
6	RESULTADOS	83
6.1	Arranjo das cornetas e resolução espacial	87
6.2	Séries temporais	90
6.3	Pré-processamento dos mapas de intensidade	92
6.4	Separação de componentes	97
6.4.1	Resultados GNILC	98
6.4.2	Resultados PCA	105
6.4.3	Comparação entre os métodos GNILC e PCA	112
6.5	Comparação entre os métodos <i>naive</i> e iGLS	113
6.6	Testes complementares	122
7	CONCLUSÕES	127
	REFERÊNCIAS BIBLIOGRÁFICAS	131

1 INTRODUÇÃO

Estimativas cada vez mais precisas dos parâmetros cosmológicos descritos pelo modelo cosmológico padrão (MCP) são essenciais para o pleno entendimento do universo e para uma boa descrição de seus fenômenos, podendo resultar em uma quebra de paradigmas para a cosmologia, tornando-a uma ciência cada vez mais baseada em observações e dados. Sob tal pretexto é que se constrói a cosmologia observacional, uma área das ciências que busca encontrar provas empíricas de propriedades e características descritas pelo MCP, além de formas de aprimorá-lo. Considerada um dos pilares do MCP, a expansão do universo foi observada pela primeira vez por Edwin Hubble e entendida como ocorrendo de forma acelerada por [Riess et al. \(1998\)](#) e [Perlmutter \(1999\)](#). Apesar de muito estudado, esse efeito ainda é pouco entendido, sendo a abordagem mais convencional associá-lo a uma componente exótica que permeia todo o cosmo chamada de energia escura, responsável por cerca de 70% da densidade de energia do universo e representada teoricamente pela chamada “Constante Cosmológica”.

No entanto, suas propriedades ainda são pouco conhecidas e estudá-las mostra-se também um processo complexo. Em tal contexto, instrumentos capazes de detectar sinais cosmológicos com grande precisão, como o Planck ([TAUBER et al., 2004](#)), o SKA (*Square Kilometre Array*) ([DEWDNEY et al., 2009](#)), o LSST (*Large Synoptic Survey Telescope*) ([TYSON, 2002](#)) e o BINGO (*BAO from Integrated Neutral Gas Observations*) ([WUENSCHÉ; COLLABORATION, 2019](#)), vem se tornando cada vez mais importantes para o suporte observacional do que hoje conhecemos como “cosmologia de precisão”. Uma das formas de se restringir os valores dos parâmetros cosmológicos que descrevem o setor escuro está em medir a distribuição de matéria em grande escala no universo e analisar como esta tem evoluído dinamicamente ([FURLANETTO et al., 2006](#)). A medição das oscilações acústicas de bárions (do inglês, baryonic acoustic oscillations - BAO), flutuações na distribuição de matéria ordinária produzidas por pequenas instabilidades no universo primordial que deixaram uma assinatura na distribuição de matéria em larga escala, a chamada escala de BAO, tem se estabelecido como uma excelente régua cósmica ([WEINBERG et al., 2013](#)). Um dos traçadores utilizados para mapear essa distribuição espacial de matéria é a emissão da transição hiperfina do hidrogênio neutro (doravante, HI), elemento mais abundante no universo, a linha de 21 cm ([PETERSON et al., 2009](#)).

A cosmologia de 21 cm tem se destacado como uma área promissora para a investigação da distribuição do HI no universo, oferecendo informações sobre o processo de

evolução das estruturas cósmicas. Nesse contexto, o radiotelescópio BINGO (*BAO from Integrated Neutral Gas Observations*) (WUENSCHÉ et al., 2022), projeto liderado pelo Brasil com forte contribuição de instituições chinesas, surge como uma ferramenta inovadora para o estudo das oscilações acústicas de bárions por meio da observação da emissão do HI na faixa de 21 cm. Utilizando-se das medidas do sinal de 21 cm, serão construídos mapas de intensidade do mesmo, produzindo uma “tomografia” da distribuição do HI ao longo de um intervalo de *redshifts*, o que permitirá o estudo da evolução cósmica e suas correlações.

O instrumento será localizado em um sítio com baixa contaminação de sinais eletromagnéticos produzidos pelo homem, no Parque Estadual da Serra do Urubu no estado da Paraíba. O BINGO é um radiotelescópio de trânsito, formado por 2 espelhos refletores fixos e um arranjo de 28 cornetas organizadas em uma estrutura de colmeia, em que as cornetas são alocadas de forma a minimizar o espaço entre si (ABDALLA et al., 2022b). O sinal de origem cosmológica incide no espelho principal, paraboloide, com semi-eixo maior de 25,5 m, sendo posteriormente refletido para um espelho secundário, hiperboloide, com semi-eixo maior de 17,8 m e por fim, captado pelo arranjo de cornetas localizado no plano focal. O BINGO operará em uma faixa de baixas frequências, na banda delimitada por 980 MHz e 1260 MHz (WUENSCHÉ et al., 2022).

Por se tratar de um telescópio de trânsito, o instrumento poderá revisitar uma mesma região do céu a cada dia, de forma a cobrir uma região de céu total de mais de 2000 deg². As medidas realizadas pelo instrumento retornarão na forma de um conjunto de séries temporais contendo o sinal de 21 cm medido em cada instante de tempo para cada região do céu observada. Tais séries temporais serão então incorporadas a algoritmos que cruzarão os dados obtidos com características instrumentais e observacionais, de forma a obter a distribuição e a intensidade do sinal em cada região do céu, construindo assim os mapas de intensidade do sinal de 21 cm.

Para o presente trabalho, serão realizadas simulações de missão do radiotelescópio BINGO com o objetivo de testar as vantagens dos algoritmos de produção de mapas e obter o método mais otimizado para tal tarefa. Este estudo propõe-se a analisar e comparar dois métodos de produção de mapas simulados no âmbito do projeto BINGO: o método denominado *naive*, que baseia-se na média das medidas obtidas por cada pixel do detector para determinar a intensidade no ponto de interesse e o método mínimos quadrados generalizados iterativos, em uma abordagem de máxima

verossimilhança, que incorpora em suas especificações uma pré-análise dos dados que auxilia na remoção de parte do ruído instrumental, o qual espera-se ser mais eficiente. Serão testadas e comparadas rotinas computacionais produzidas e aprimoradas pela colaboração e outras disponibilizadas para a realização do trabalho. A aplicação desses métodos na geração de mapas de intensidade visa explorar as vantagens da abordagem de máxima verossimilhança e conduzir uma análise comparativa de seus resultados com relação ao método convencional. Ao analisar e discutir a importância da produção de mapas de intensidade e da recuperação do sinal de HI no contexto do radiotelescópio BINGO, este estudo almeja contribuir para o aprimoramento das técnicas de análise de dados astrofísicos do projeto, e por conseguinte para a realização de medidas de BAO na banda de rádio.

O trabalho está organizado da seguinte forma: o Capítulo 2 apresenta uma introdução ao referencial teórico adotado, no qual discorre-se sobre o modelo cosmológico padrão, em especial sobre a expansão acelerada do Universo e sobre a utilização das medidas das oscilações acústicas de bárions como elemento para medir-se tal expansão. Em seguida apresenta-se o mapeamento de intensidade da linha de 21 cm como um técnica utilizada para medir a escala de BAO e os detalhes sobre a emissão de tal linha e por fim, uma breve descrição dos principais contaminantes astrofísicos na faixa de frequência de interesse. O Capítulo 3 apresenta uma breve descrição sobre a instrumentação em radioastronomia, contando com os diferentes tipos de radiotelescópios; uma descrição do radiotelescópio BINGO, instrumento adotado neste trabalho para as simulações de observação da linha de 21 cm e por fim, uma descrição de alguns ruídos instrumentais, com especial interesse para o ruído $1/f$. Já o Capítulo 4 apresenta os fundamentos matemáticos e os métodos de produção de mapas adotados neste trabalho, bem como as ferramentas de simulação e os métodos de separação de componentes adotados. O Capítulo 5 conta com uma descrição das etapas de pré-processamento as quais os mapas foram submetidos, como a remoção do nível mínimo de ruído, o ajuste de feixe e a aplicação de uma máscara apodizada. Também são apresentados nesse Capítulo as ferramentas e os parâmetros utilizados para comparar os resultados obtidos para as ferramentas de produção de mapas e separação de componentes. O Capítulo 6 discute os resultados obtidos durante a realização desse trabalho, como os mapas produzidos para diferentes arranjos de cornetas, as séries temporais para diferentes parâmetros espectrais do ruído, a eficácia dos métodos de separação de componentes e os mapas obtidos com o método *naive* e o método iGLS. Por fim, o Capítulo 7 apresenta as conclusões e as perspectivas relacionadas ao trabalho e aos resultados obtidos, discutindo acerca do impacto da estratégia de observação na redução de efeitos sistemáticos, como o introduzido pelo

ruído $1/f$.

2 REFERENCIAL TEÓRICO

As bases da cosmologia moderna foram estruturadas através das hipóteses determinadas pelo Modelo Cosmológico Padrão, que se fundamenta em quatro evidências observacionais principais: a expansão do Universo, a aceleração dessa expansão, a presença da radiação cósmica de fundo em micro-ondas e a abundância de elementos leves. A observação e estudo desses fenômenos levaram à construção de um modelo teórico-matemático que descreve a evolução do Universo a partir de interações gravitacionais, que regem a matéria em escalas cosmológicas.

No início do século 20, a cosmologia teórica tomou forma com as descrições matemáticas do Universo propostas por Einstein, Friedmann e de Sitter (RYDEN, 2017; DODELSON; SCHMIDT, 2020). O desenvolvimento da relatividade geral permitiu a descrição da ação de forças no Universo em grandes escalas, através da ação de campos gravitacionais gerados pela distorção do tecido espaço-tempo, possibilitando uma forma unificada de estudar o Universo em escalas cosmológicas. Einstein descreveu a ação da gravidade sobre a matéria através de suas equações de campo, partindo de um modelo que considerava um Universo espacialmente homogêneo, isotrópico e estático, sem expansão, representado pela presença da constante cosmológica, Λ .

No entanto, a cosmologia observacional também em ascensão à época, apresentou provas da expansão cosmológica. Uma das primeiras propostas de um Universo não estático foi feita por Edwin Hubble (1889 — 1953) em seu trabalho de observação de galáxias (HUBBLE, 1929). Hubble argumentou que as galáxias estão gradativamente se afastando umas das outras de forma acelerada, ao observar a mudança de posição de suas linhas espectrais. O deslocamento das linhas espectrais permitiu que Hubble determinasse a velocidade de afastamento entre galáxias, obtendo a primeira evidência observacional de um Universo não estático (HUBBLE, 1929; ASSIS et al., 2008). Este fato levou ao estabelecimento da Lei de Hubble, que quantifica a expansão cosmológica e desde então, diversos trabalhos (BAHCALL, 2015; WEINBERG et al., 2013; YOO; WATANABE, 2012; FRIEMAN et al., 2008) vêm sendo desenvolvidos para compreender tal comportamento.

O formato e o comportamento expansivo do Universo foram descritos através das equações de Friedmann, a partir da métrica de Friedmann-Lemaître-Robertson-Walker. Tais equações representam uma forma de descrever a dinâmica do Universo dentro do modelo cosmológico escolhido, assumindo parâmetros cosmológicos para a curvatura do Universo, a dinâmica deste e a densidade de seus constituintes de matéria e radiação, entre outros (PADMANABHAN, 2003). A constante Λ foi

inicialmente adicionada nos modelos para contrabalancear a ação da gravidade e permitir a existência de um Universo estático. No entanto, após as descobertas de Hubble, a ideia foi abandonada, e o Universo passou a ser entendido como dinâmico. A constante cosmológica hoje é vista como responsável pela expansão acelerada representando uma energia com pressão negativa (PADMANABHAN, 2003). A interpretação de Λ , assumindo o princípio cosmológico relativo à densidade de massa-energia do Universo, é de que esta está associada à energia escura.

Considerando-se as proposições anteriores, tem-se que o Modelo Cosmológico Padrão, também denominado Λ CDM, fundamenta-se, portanto, em um Universo isotrópico, homogêneo, em expansão acelerada, dominado por energia escura (Λ) e matéria escura fria (do inglês, Cold Dark Matter - CDM) (LIDDLE; LYTH, 1993; DODELSON; SCHMIDT, 2020). Este modelo está em concordância com as proposições do Big Bang, considerando um Universo que evoluiu a partir de um estado altamente denso e quente que resfriou ao expandir-se, respaldado pelas observações da Radiação Cósmica de Fundo em Microondas (do inglês, Cosmic Microwave Background - CMB) (AGHANIM et al., 2018), pela organização da estrutura em grande escala e pela abundância de elementos químicos leves no Universo.

Esse capítulo descreve a fundamentação teórica adotada neste trabalho, centrando-se em discutir as bases teóricas que levaram aos estudos da expansão acelerada do Universo e como quantificá-la, a fim de restringir valores aos parâmetros cosmológicos adotados nos modelos teóricos, buscando uma descrição mais precisa da dinâmica e da estruturação do Universo. Na Seção 2.1 descrevemos o Modelo Cosmológico Padrão, a descrição matemática da estrutura e dinâmica do Universo, bem como de sua expansão acelerada e os processos físicos envolvidos na sua evolução. Já na Seção 2.2 constam os fundamentos das oscilações acústicas de bárions, seu processo de formação e uso como régua padrão em cosmologia. A Seção 2.3 trata do uso da técnica de mapeamento de intensidade como traçador da distribuição de matéria em larga escala e especificamente do mapeamento da linha de 21 cm do HI, apresentando as vantagens da técnica. Na Seção 2.4 consta a física da emissão de 21 cm, os fenômenos e processos físicos que levam à transição de estado e a descrição matemática do contraste de temperatura medido. Por fim, a Seção 2.5 versa sobre os principais contaminantes astrofísicos na faixa de rádio e a modelagem da sua emissão.

2.1 O modelo cosmológico padrão

Considerando para a estruturação teórica deste trabalho um Universo descrito pelo Modelo Cosmológico Padrão, ou seja, um Universo isotrópico, homogêneo e em expansão acelerada, dominado por energia escura e matéria escura fria (DODELSON; SCHMIDT, 2020; LIDDLE; LYTH, 1993), assume-se, portanto, um Universo descrito pela métrica de Friedmann-Lemaître-Robertson-Walker e pelas equações de Friedmann-Lemaître.

As equações de Friedmann-Lemaître representam soluções exatas para as equações de campo formuladas por Einstein. Einstein expressou a ação da gravidade sobre a matéria por meio dessas equações de campo na seguinte forma:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + g_{\mu\nu}\Lambda = \frac{8\pi G}{c^4}T_{\mu\nu}, \quad (2.1)$$

onde $R_{\mu\nu}$ representa o tensor de curvatura de Ricci, que descreve a curvatura intrínseca do espaço-tempo em cada ponto; $g_{\mu\nu}$ é o tensor métrico, que descreve a distância infinitesimal entre dois pontos próximos em uma variedade diferenciável; R é o escalar de curvatura, que fornece uma medida global da curvatura do espaço-tempo; Λ é a constante cosmológica, inicialmente incluída para permitir uma solução estática para o Universo; G é a constante gravitacional de Newton; c é a velocidade da luz no vácuo; e $T_{\mu\nu}$ é o tensor de energia-momento, que descreve a distribuição de energia, momento e pressão em uma região específica do espaço-tempo.

Para possibilitar o estudo cosmológico, assume-se um tensor energia-momento de um fluido perfeito para descrever a distribuição de fluxo de energia e momento associados ao conteúdo material/energético que permeia o espaço-tempo (DODELSON; SCHMIDT, 2020), de forma que este é dado por:

$$T_{\mu\nu} = (\rho_T + p_T)\frac{u_\mu u_\nu}{c^2} - p_T g_{\mu\nu}. \quad (2.2)$$

No qual ρ_T representa a densidade de energia total do fluido (energia por volume); p_T é a pressão do fluido; o termo $u_\mu u_\nu$ representa o quadri-vetor velocidade que descreve o movimento do fluido no espaço-tempo e $g_{\mu\nu}$ representa a métrica adotada.

Identificamos as componentes de $g_{\mu\nu}$ a partir da métrica de Friedmann-Lemaître-Robertson-Walker, dada por

$$ds^2 = (cdt)^2 - a^2(t) \left[\frac{dr^2}{1 - Kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \right], \quad (2.3)$$

onde $a(t)$ é o fator de escala, que descreve como as distâncias entre objetos no Universo mudam ao longo do tempo, sendo uma grandeza dependente do tempo e do modelo cosmológico adotado. Com $a(t) = 1$, tem-se um Universo estático, sem expansão ou contração. Na Equação (2.3), dt representa o diferencial de tempo para dois eventos; dr é a diferença de distância radial; $d\theta$ e $d\phi$ são os diferenciais de ângulo para coordenadas esféricas; e K representa a curvatura espacial do Universo.

Ao aplicar as Equações 2.3 e 2.2 na 2.1, obtém-se as equações

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho_T - \frac{K}{a^2} + \frac{\Lambda}{3} \quad (2.4)$$

e

$$2\frac{\ddot{a}}{a} + \frac{\dot{a}^2 + Kc^2}{a^2} = \frac{-8\pi G}{c^2}p_T + \Lambda c^2, \quad (2.5)$$

onde \dot{a} é a derivada temporal de a , G é a constante gravitacional e Λ é a constante cosmológica. Essas equações, conhecidas como Equações de Friedmann, representam uma forma de descrever a dinâmica do Universo dentro do modelo cosmológico escolhido, assumindo parâmetros cosmológicos para a curvatura do Universo, a dinâmica deste e a densidade de seus constituintes de matéria e radiação, entre outros.

A constante Λ foi inicialmente adicionada nos modelos para contrabalançar a ação da gravidade e permitir a existência de um Universo estático. No entanto, após as descobertas de Hubble, a ideia foi abandonada, e o Universo passou a ser entendido como dinâmico. A constante cosmológica hoje é vista como responsável pela expansão acelerada do Universo, representando uma energia com pressão negativa (AGHANIM et al., 2018). A interpretação de Λ , assumindo o princípio cosmológico da densidade de massa-energia do Universo, está associada à energia escura. Neste contexto, o setor escuro do Universo, que compreende a matéria e a energia escuras, é essencial para descrever o modelo e a dinâmica cosmológica, representando 96% dos constituintes do Universo (AGHANIM et al., 2018).

Para estudar a expansão acelerada, é necessário definir elementos de medida em

escala cosmológica que permitam traçar sua evolução e esclarecer a sua ação sobre a matéria no Universo. Para isso, é comum utilizar as oscilações acústicas de bárions, que fornecem uma medida de distância bem definida nas estruturas em grande escala e podem ser usadas como réguas padrão em cosmologia (EISENSTEIN et al., 2007; KIM et al., 2014; WANG, 2006). A seção seguinte abordará esse elemento, explicando como se formam e sua aplicação no estudo da expansão cosmológica.

2.2 Oscilações acústicas de bárions

As pequenas perturbações na homogeneidade do Universo primordial favoreceram o surgimento de zonas de maior densidade de matéria, o que permitiu que esta se aglutinasse via ação gravitacional, formando poços de potencial gravitacional. A matéria aglutinada aqueceu e passou a emitir radiação, o que gerou uma tensão entre a pressão de radiação e a atração gravitacional (BASSETT; HLOZEK, 2010). Tais instabilidades geraram oscilações na matéria que se propagaram como ondas sonoras, denominadas oscilações acústicas de bárions (do inglês, baryon acoustic oscillations - BAO), carregando consigo fótons e bárions até o momento do desacoplamento. A presença da tensão entre pressão dos fótons e atração gravitacional pode ser descrita pela expressão (BASSETT; HLOZEK, 2010)

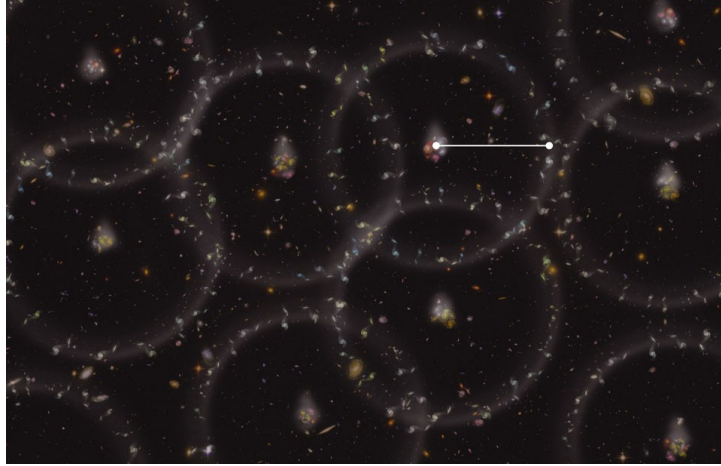
$$\ddot{\delta} + (\text{Pressão} - \text{Gravidade})\delta = 0, \quad (2.6)$$

na qual δ representa o contraste de densidade $\delta(x, t)$ relativo à uma densidade média ρ_0 e os dois pontos representam sua segunda derivada temporal, o que permite escrever a equação que representa o movimento do fluido primordial

$$\ddot{\delta}_\gamma + \frac{\mathcal{H}R}{1+R}\delta_\gamma + c_s^2 k^2 \delta_\gamma = -\frac{4}{3}k^2\Phi + \frac{4\dot{R}}{1+R}\dot{\Phi} + 4\ddot{\Phi}, \quad (2.7)$$

onde δ_γ representa a perturbação da densidade de fótons e $\ddot{\delta}_\gamma$ é a sua segunda derivada temporal, \mathcal{H} é o produto entre o fator de escala a e o parâmetro de Hubble H ; R é proporcional à razão entre densidades médias de bárions e fótons; c_s a velocidade de propagação do som no plasma de bárions e fótons; k é o número de onda e Φ é o potencial gravitacional em um dado ponto x e instante de tempo t . O lado esquerdo da Equação 2.7 nos apresenta o comportamento da pressão produzida pelos fótons e o lado direito a variação da gravidade por meio do potencial gravitacional produzido por bárions e matéria escura (WEINBERG et al., 2013).

Figura 2.1: Ilustração dos anéis de concentração de matéria que determinam as regiões de formação de estruturas.

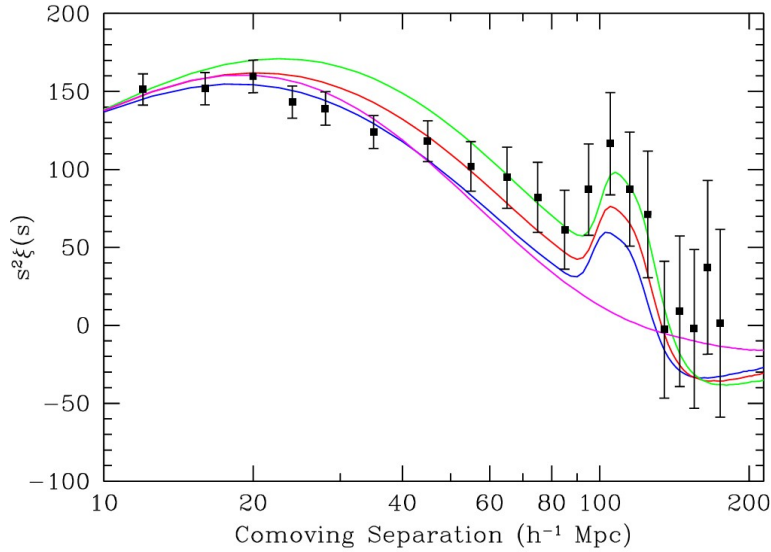


Os pontos coloridos representam galáxias, enquanto os anéis esbranquiçados ilustram a escala característica de BAO. A linha branca demarca o raio acústico, r_s .
 Fonte: Rostomian (2025).

Quando os fótons se dissociaram dos bárions devido a queda da temperatura média, no evento denominado desacoplamento, estes permaneceram em anéis concêntricos formados pela propagação das oscilações, sendo seu centro dado pela região de maior densidade, onde a matéria escura permaneceu (EISENSTEIN et al., 2005). Os fótons, por sua vez, livres da ação gravitacional, dissiparam-se no cosmos. Os anéis formados caracterizam regiões de acúmulo de densidade de matéria, as “sementes” da formação das primeiras estruturas no Universo, que estabelecem um padrão de distribuição de matéria em escala cosmológica, permitindo a determinação de réguas cósmicas (EISENSTEIN et al., 2007; COLE et al., 2005). O raio médio dos anéis formados pelas oscilações estabelece uma escala que determina a organização da matéria e a formação de estruturas em grande escala. Tal raio, também chamado de escala de comprimento acústico, r_s , é dado pela distância percorrida pelas ondas geradas no Big Bang até o momento do desacoplamento e pode ser expresso pela expressão (WEINBERG et al., 2013):

$$r_s = \int_0^{t_{\text{rec}}} \frac{c_s(t)}{a(t)} dt = \int_{z_{\text{rec}}}^{\infty} \frac{c_s(z)}{H(z)} dz, \quad (2.8)$$

Figura 2.2: Gráfico da função de correlação de dois pontos para distância comóvel entre galáxias.



O eixo y apresenta a função de correlação de dois pontos, $s^2\xi(s)$ e o eixo x a separação comóvel entre galáxias, s . O segundo pico representa uma separação típica dada por r_s . As diferentes curvas representam diferentes modelos cosmológicos, com diferentes valores para o parâmetro de densidade de matéria não-relativística. A curva magenta, especialmente, representa um modelo CDM puro, sem energia escura.

Fonte: Eisenstein et al. (2005).

na qual os limites t_{rec} e z_{rec} representam, respectivamente, o tempo e o *redshift* da ocorrência da recombinação e c_s representa a velocidade do som no plasma em função do tempo e do *redshift*. Teorias de perturbação cosmológica e medidas realizadas indicam uma escala de aproximadamente 150 Mpc para r_s .

A Figura 2.1 ilustra a distribuição de galáxias no Universo, destacando a organização preferencial das estruturas em escalas angulares correspondentes a um raio físico r_s , uma consequência das oscilações acústicas de bárions. Esse excesso de probabilidade de encontrar pares de galáxias a uma determinada distância é quantificado pela função de correlação de dois pontos. A Figura 2.2 exibe essa função, calculada por Eisenstein et al. (2005) a partir de dados do SDSS (*Sloan Digital Sky Survey*) (YORK et al., 2000). O pico proeminente no gráfico, conhecido como “pico de BAO”, indica uma separação preferencial entre galáxias em r_s . Em outras palavras, há um excesso estatisticamente significativo de pares de galáxias separados pela distância do horizonte acústico, o que constitui uma evidência observacional da presença das BAO.

Com a expansão acelerada do Universo, tais estruturas evoluíram e o raio característico das BAO alterou-se em conjunto, aumentando progressivamente. Portanto, obter as medidas das BAO em momentos distintos da evolução do Universo permite estudar a sua expansão e delimitar propriedades do setor escuro, ver, p.ex., Abdalla et al. (2022a). No entanto, o estudo da energia e da matéria escura permanece como um grande desafio para a cosmologia moderna (BUCKHOLTZ, 2024; ZHUANG et al., 2024), uma vez que o setor escuro não interage eletromagneticamente com a matéria ordinária e seus elementos estruturantes são ainda desconhecidos. Com tal plano de fundo, a cosmologia de 21 cm apresenta-se como um caminho para encontrar os parâmetros cosmológicos necessários à compreensão e delimitação das propriedades do setor escuro bem como a evolução do Universo.

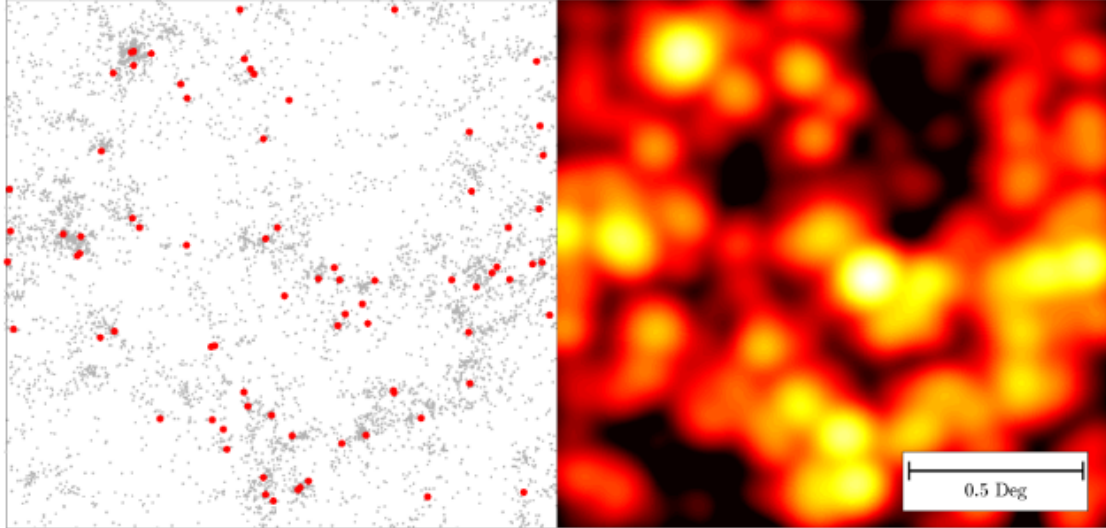
2.3 Mapeamento de intensidade em linhas

O mapeamento de intensidade em linhas é uma técnica observacional fundamental na astrofísica e especialmente na cosmologia, empregada na observação de grandes regiões do céu com baixa resolução (KOVETZ et al., 2017; PETERSON et al., 2009). Seu uso se dá principalmente no estudo de emissões integradas em um volume de *redshift*, permitindo a construção de um mapa tridimensional do céu (PRITCHARD; LOEB, 2012). Esta técnica fundamenta-se na coleta da emissão integrada de linhas espectrais de um determinado elemento químico e na análise de suas variações espaciais para estudar a formação e desenvolvimento da estrutura cósmica (PETERSON et al., 2009). As variações na emissão têm uma ligação direta com a estrutura em grande escala do Universo, pois permitem observar como a matéria está distribuída espacialmente ao longo do tempo, enquanto a dependência da frequência revela detalhes sobre a distribuição da emissão na linha de visada ao longo do *redshift* (KOVETZ et al., 2017). Isto permite o mapeamento da distribuição de matéria em um volume cósmico, definido pela profundidade em *redshift* e pela região de observação (PETERSON et al., 2009).

Uma das vantagens do mapeamento de intensidade é que este não exige altas resoluções angulares e permite que grandes áreas do céu sejam cobertas em um tempo relativamente pequeno, uma vez que, nesta técnica, a emissão de todas as fontes em uma determinada linha de visada é contabilizada (BULL et al., 2015). Têm-se, portanto, a medida de determinado sinal em uma grande área, com baixa resolução (KOVETZ et al., 2017). A resolução não é um problema pois não se têm interesse em resolver pontualmente as fontes, mas sim na emissão total do meio.

A Figura 2.3 reproduzida por Kovetz et al. (2017) apresenta uma comparação da

Figura 2.3: Um campo simulado de 2,5 graus quadrados mostrando as localizações das galáxias (esquerda) ao lado do mapa de intensidade de CO correspondente (direita).



As fontes que são suficientemente brilhantes para serem detectadas com 1 hora de observação no VLA estão indicadas em vermelho.

Fonte: Breysse (2025).

simulação de observação de um mesmo campo utilizando duas técnicas distintas, compreendendo um campo de galáxias pontuais observadas pelo *Very Large Array* (VLA) (NAPIER et al., 1983) à esquerda e o mapeamento de intensidade da mesma região, feito utilizando linhas de CO pelo *CO Mapping Array Project* (COMAP) (CLEARY et al., 2022) e exemplifica a atuação da técnica. Diversas linhas espectrais podem ser utilizadas para a realização do mapeamento de intensidade como a emissão do CO associada a formação estelar (BREYSSE et al., 2014), $H-\alpha$ e $Ly-\alpha$ em regiões de gás ionizado no interior de galáxias (SILVA et al., 2017; PULLEN et al., 2014) e linhas de $[OIII]$ utilizadas para estudos em altos *redshifts* (MORIWAKI et al., 2018). Para o estudo da distribuição da matéria em larga escala, há um interesse especial na emissão na linha de 21cm do hidrogênio neutro, pois representa um bom traçador de matéria, sendo o hidrogênio o elemento mais abundante no Universo.

O mapeamento de intensidade da linha de 21 cm é uma técnica fundamental na cosmologia moderna, permitindo investigar a evolução do Universo desde a chamada ‘Idade das Trevas’ até a formação das primeiras galáxias e a época da reionização. Durante as primeiras centenas de milhões de anos após o Big Bang, o Universo

passou por fases importantes, como a recombinação, quando os átomos neutros se formaram, e a subsequente “Idade das Trevas”, onde não havia fontes de luz visíveis (DODELSON; SCHMIDT, 2020). Entretanto, há uma lacuna significativa entre essa época e a formação das primeiras galáxias, um intervalo conhecido como “*cosmic dawn*” (aurora cósmica). O mapeamento de intensidade de 21 cm busca preencher essa lacuna (PRITCHARD; LOEB, 2012).

A técnica consiste em medir a radiação emitida pelo hidrogênio neutro (HI) ao longo de diferentes *redshifts*, através do mapeamento de intensidade em 21 cm, permitindo reconstruir a distribuição de matéria e das estruturas cósmicas em uma “tomografia” e, então, determinar a evolução do Universo ao longo do tempo (PRITCHARD; LOEB, 2012; BULL et al., 2015). A principal vantagem desse método é a sua capacidade de oferecer uma nova janela de observação do Universo em uma faixa de *redshift* onde o hidrogênio neutro ainda está presente, após a reionização e no período de formação das primeiras estruturas (KOVETZ et al., 2017). Após a reionização, o sinal de 21 cm revela como os halos de matéria escura colapsam e como a distribuição do hidrogênio neutro se agrupa ao longo do campo de densidade de matéria (VILLAESCUSA-NAVARRO et al., 2018). Além disso, ao contrário de observações ópticas que se interessam por galáxias individuais, o mapeamento de 21 cm mede a intensidade integrada em grandes volumes (KOVETZ et al., 2017), proporcionando uma visão estatística da estrutura cósmica e o estudo do Universo em grandes escalas, o que permite sua aplicabilidade nas investigações acerca das oscilações acústicas de bárions (ABDALLA et al., 2022a).

Afim de compreender melhor o átomo de hidrogênio neutro e a emissão do seu sinal característico, está apresentada a seguir uma descrição da linha de 21 cm e os processos físicos que levam à sua emissão.

2.4 Física da emissão de 21 cm do HI neutro

A linha de 21 cm corresponde à transição hiperfina do hidrogênio neutro (HI), elemento mais abundante no Universo. Essa transição ocorre no nível de energia mais baixo do átomo de hidrogênio, conhecido como estado fundamental (1S). O hidrogênio, composto por um próton e um elétron, apresenta uma interação entre os momentos magnéticos dessas duas partículas, resultando em uma pequena divisão de energia dentro desse estado, formando a chamada estrutura hiperfina (COHEN-TANNOUJJI et al., 1986). A transição hiperfina se dá quando o átomo de HI sofre uma transição para um desses dois estados hiperfinos, com emissão ou absorção de um fóton. Essa diferença de energia é extremamente pequena, aproximadamente

$5,9 \times 10^{-6} eV$, o que corresponde a um comprimento de onda de 21,1 cm e uma frequência de 1420 MHz (COHEN-TANNOUJDI et al., 1986). Esse valor é um dos mais precisos em astrofísica e foi fundamental para o desenvolvimento do campo da radioastronomia (PRITCHARD; LOEB, 2012).

A física por trás dessa transição envolve a interação do momento magnético do próton com o do elétron. Quando os spins estão alinhados (paralelos), o sistema possui uma energia ligeiramente maior do que quando estão desalinhados (antiparalelos). A transição para o estado de menor energia (antiparalelo) resulta na emissão de um fóton com comprimento de onda de 21 cm (MESINGER et al., 2013). Também é possível que ocorra o processo inverso, em que o átomo transita do estado antiparalelo para o estado paralelo, ocorrendo absorção de um fóton com comprimento de onda de 21 cm (COHEN-TANNOUJDI et al., 1986). Apesar de ser uma transição rara, com uma taxa de decaimento espontâneo da ordem de $A_{10} \approx 2,85 \times 10^{-15} s^{-1}$, o que implica que um átomo de hidrogênio isolado levaria, em média, cerca de 10 milhões de anos para emitir um fóton, a enorme quantidade de hidrogênio no Universo gera um sinal detectável (PRITCHARD; LOEB, 2012).

A transição entre os níveis de energia ocorre através de três processos: emissão espontânea, emissão estimulada e absorção (COHEN-TANNOUJDI et al., 1986). A emissão espontânea é um fenômeno probabilístico que ocorre sem necessidade de estímulo externo, quando um átomo em um estado de energia mais alto decai para um estado de menor energia, emitindo um fóton correspondente a diferença de energia entre os estados. Já a emissão estimulada ocorre quando um átomo em um estado excitado interage com um fóton de mesma frequência, resultando na perda de energia do átomo e sua transição para um estado energético inferior. Isto ocorre através da emissão de um fóton com as mesmas características físicas como frequência, fase, direção e polarização do fóton estimulante. Por fim, a absorção ocorre quando o átomo absorve um fóton, ganhando energia e transitando para um estado excitado. Esta interação pode ocorrer através de colisões ou interações com campos eletromagnéticos.

Usualmente no campo da cosmologia de 21 cm, refere-se ao sinal do hidrogênio neutro através de uma medida de temperatura, mas não no sentido termodinâmico. A intensidade da radiação associada ao sinal de 21 cm pode ser descrita em termos de temperaturas através da função de Planck. Neste caso, a intensidade da radiação, $I_\nu(T)$, estando o emissor em equilíbrio termodinâmico, é equivalente ao brilho, $B_\nu(T)$, de forma que a equação de Planck fica dada por

$$I_\nu(T) = \frac{2h\nu^3/c^2}{\exp(h\nu/kT) - 1}, \quad (2.9)$$

em que h é a constante de Planck, ν representa a frequência da radiação, k é a constante de Boltzmann e T é a temperatura absoluta do emissor. Como estamos trabalhando na região de rádio do espectro eletromagnético, vale a aproximação de Rayleigh-Jeans e a Equação 2.9 se reduz a

$$I_\nu = \frac{2\nu^2 k}{c^2} T_B, \quad (2.10)$$

em que T_B é a temperatura de brilho, a temperatura de corpo negro associada ao brilho do emissor.

No entanto, ao longo de sua trajetória desde sua emissão até a captação do sinal, o feixe de radiação sofre variações na sua intensidade devido a interações com o meio no trajeto entre a fonte e o observador. A variabilidade do feixe de radiação pode ser descrita pela Equação de Transferência Radiativa (ETR). A ETR pode ser descrita como a variabilidade da intensidade da radiação ao longo de uma direção (RYBICKI; LIGHTMAN, 2024),

$$\frac{dI_\nu}{d\tau_\nu} = -I_\nu + B_\nu(T), \quad (2.11)$$

em que τ_ν é a profundidade óptica do meio. É conveniente escrever a ETR em função das temperaturas dos elementos envolvidos. Para tal, aplica-se a Equação 2.10 em 2.11, de forma que a ETR pode ser expressa por

$$\frac{dT_B}{d\tau_\nu} = -T_B + T_R, \quad (2.12)$$

em que T_R é a temperatura da radiação de fundo, no caso desta análise, a temperatura da CMB, T_{CMB} . A solução da ETR da Equação 2.12 em função das temperaturas e da profundidade óptica (FURLANETTO et al., 2006), pode ser expressa como

$$T_B(\nu) = T_s(1 - e^{-\tau_\nu}) + T_{\text{CMB}}(\nu)e^{-\tau_\nu}, \quad (2.13)$$

onde T_s é chamada de “temperatura de spin” e está associada a energia de excitação do átomo de HI quando da geração do sinal de 21 cm. No entanto, para o estudo do sinal de 21 cm do HI através do mapeamento de intensidade, a quantidade de interesse é a flutuação da temperatura da emissão em relação a temperatura de fundo, $\delta T_B(\nu)$. Esta depende fortemente da definição da profundidade óptica e de quais interações, com quais tipos de campos de radiação, o feixe sofrerá ao longo de sua trajetória (FURLANETTO et al., 2006). Uma descrição mais completa da variação da temperatura de brilho, considerando as interações da emissão de 21 cm com diversos agentes de campo externo é, por fim, dada por

$$\begin{aligned} \delta T_b(\nu) \approx 9X_{HI}(1 + \delta_b) & \left[1 - \frac{T_{\text{CMB}}(z)}{T_s} \right] \\ & \times (1 + z)^{1/2} \left[\frac{H(z)/(1 + z)}{d\nu_{\parallel}/dr_{\parallel}} \right] \text{ [mK]}. \end{aligned} \quad (2.14)$$

Na Equação 2.14, também denominada como contraste da temperatura de brilho, X_{HI} é fração de hidrogênio neutro no Universo em função da evolução cosmológica e está relacionada com a quantidade de hidrogênio neutro presente no meio intergaláctico; δ_b é a variação da densidade de bárions; $T_{\text{CMB}}(z)$ é a temperatura da CMB em função do *redshift*; a expressão $(1 + z)^{1/2}$ representa o fator de escala cosmológico, a , considerado devido aos efeitos da expansão do Universo; $H(z)$ é o parâmetro de Hubble e $d\nu_{\parallel}/dr_{\parallel}$ o gradiente de velocidade radial na linha de visada, que introduz alterações no perfil de linha da emissão. Uma descrição mais completa dos efeitos de cada componente ao longo da linha de visada na intensidade do sinal medido consta em Loeb e Wyithe (2008) e Furlanetto et al. (2006).

Para a realização do mapeamento de intensidade em 21 cm, necessita-se portanto, de instrumentos refinados que permitam medir pequenas variações espaciais na temperatura de brilho da radiação de 21 cm em relação a temperatura da CMB. Com a capacidade de observar o sinal em $z \approx 1 - 3$ (LOEB; WYITHE, 2008), uma região chave na transição de dominância da matéria para a energia escura, tal técnica tem o potencial de melhorar as restrições sobre parâmetros cosmológicos, como a massa do neutrino e a evolução da densidade de matéria (PRITCHARD; LOEB, 2012). Além disso, a realização do mapeamento de intensidade é particularmente importante porque permite observar uma região muito maior do Universo do que os levantamentos de galáxias.

O MI em 21 cm foi proposto como uma ferramenta para estudar uma série de fenô-

menos cosmológicos, incluindo a medição das oscilações acústicas de bárions e a energia escura (BULL et al., 2015). Ao estudar essas oscilações, é possível restringir a equação de estado da energia escura, um componente fundamental que acelera a expansão do Universo. Têm-se, portanto, que o IM em 21 cm pode trazer diversos avanços quanto à realização de medidas da distribuição de matéria no Universo (VILLAESCUSA-NAVARRO et al., 2018). A seção seguinte trata sobre alguns dos contaminantes astrofísicos que dificultam a observação do sinal de interesse e comprometem a qualidade dos mapas de intensidade observados e devem ser tratados através de métodos de separação de componentes.

2.5 Contaminantes astrofísicos

A seguir estão descritos alguns dos principais sinais astrofísicos presentes na faixa de rádio que agem como contaminantes do sinal de HI. Felizmente, seus espectros de emissão são bem conhecidos e modelados, o que possibilita que seu sinal seja devidamente separado do sinal do hidrogênio (BIGOT-SAZY et al., 2015; BATTYE et al., 2013).

- Síncroton

A radiação síncroton é um tipo de emissão eletromagnética gerada por partículas carregadas que se movem em trajetórias espirais ao redor de linhas de campo magnético com velocidades relativísticas, o que faz com que os elétrons emitam energia na forma de radiação não térmica. A radiação síncroton apresenta um espectro característico em forma de lei de potência, com a intensidade decrescendo em frequências mais altas e grande variabilidade espacial (OLIVEIRA-COSTA et al., 2008). É amplamente observada em vários contextos astrofísicos, como remanescentes de supernovas, núcleos galácticos ativos (AGNs), pulsares e também na Via Láctea, onde é produzida pela interação de elétrons relativísticos com o campo magnético galáctico.

O espectro de emissão síncroton, em termos da temperatura de brilho, pode ser escrito como (OLIVEIRA-COSTA et al., 2008)

$$T(\nu) \propto \nu^{\beta+C \times \ln\left(\frac{\nu}{\nu_0}\right)}, \quad (2.15)$$

onde ν é a frequência de emissão; β é o índice espectral da radiação Síncroton; C é uma constante relacionada à variação na relação entre a temperatura de brilho e

a frequência e ν_0 é uma frequência de referência. No caso do BINGO, a radiação síncrotron é um dos principais contaminantes astrofísicos, pois pode dominar a faixa de frequência de interesse para a detecção da linha de 21 cm do hidrogênio neutro, uma vez que a intensidade do seu sinal chega a ser 4 ordens de grandeza maior do que o sinal do HI.

- Bremsstrahlung

A radiação Bremsstrahlung, ou radiação de frenagem, é um tipo de emissão contínua gerada quando partículas carregadas, como elétrons, sofrem deflexão ao interagir com campos elétricos de outras partículas, como íons. Durante essa interação, a mudança na trajetória e na velocidade do elétron faz com que ele perca energia, emitida na forma de radiação eletromagnética. Essa radiação é particularmente importante em meios ionizados, como o plasma interestelar, e sua intensidade depende da densidade do meio e da energia das partículas envolvidas.

A temperatura de brilho da radiação Bremsstrahlung fora do plano galáctico pode ser descrita como (OLIVEIRA-COSTA et al., 2008)

$$T_{\text{bs}} \approx 90 \text{ mK} \left(\frac{T_e}{\text{K}} \right)^{-3.5} \left(\frac{\nu}{\text{GHz}} \right)^{-2.1} \left(\frac{EM}{\text{cm}^{-6} \text{ pc}} \right), \quad (2.16)$$

onde EM é a medida de emissão, ligada à densidade do plasma ao longo da linha de visada; T_e é a temperatura de elétrons, ligada a energia dos elétrons no plasma e ν é a frequência de emissão de tal radiação. Tem-se, portanto, que a temperatura de brilho da emissão obedece uma lei de potências com índice espectral $\beta = -2, 1$.

- Fontes Puntiformes Extragalácticas

As fontes puntiformes extragalácticas são objetos localizados fora da Via Láctea que se destacam como pontos de alta intensidade em mapas de rádio devido à sua pequena extensão angular. Entre os principais exemplos estão os quasares, que são núcleos galácticos ativos extremamente luminosos, as galáxias de rádio, caracterizadas por emissões intensas em frequências de rádio, núcleos galácticos ativos (AGNs), regiões compactas no núcleo de galáxias que emitem intensamente devido a processos de alta energia, e as galáxias com formação estelar intensa, onde o elevado índice de formação de estrelas gera emissões significativas.

As fontes mais intensas, que podem ser resolvidas espacialmente, são facilmente determinadas e podem ser subtraídas em uma análise posterior das medidas feitas pelo instrumento. Já as fontes fracas não podem ser resolvidas e formam um fundo contínuo de rádio, que gera uma elevação do nível de ruído acima do ruído instrumental. Pode-se modelar tal emissão de fontes fracas pela seguinte expressão

$$T_{ps}(\nu, \hat{n}) = \left(\frac{dB}{dT} \right)^{-1} \Omega_{\text{pix}}^{-1} \sum_{i=1}^N S_i(\nu), \quad (2.17)$$

em que $T_{ps}(\nu, \hat{n})$ é a temperatura de brilho da radiação de uma fonte puntiforme em uma direção específica \hat{n} e para uma frequência ν ; $\left(\frac{dB}{dT} \right)^{-1}$ é um termo que representa o inverso da derivada do brilho específico B em função da temperatura T ; Ω_{pix} é o ângulo sólido correspondente ao tamanho de um pixel no mapa do céu e por fim, $\sum_{i=1}^N S_i(\nu)$ é o somatório da intensidade de radiação de N fontes puntiformes. Cada fonte i contribui com uma intensidade $S_i(\nu)$.

- Radiação Cósmica de Fundo em Micro-ondas

A Radiação Cósmica de Fundo em Microondas (RCFM) é uma das principais fontes de radiação de micro-ondas no Universo e sua emissão pode ser modelada como uma radiação de corpo negro com uma temperatura muito bem definida de cerca de 2,725 K, sendo remanescente do período de recombinação, aproximadamente 380.000 anos após o Big Bang. Sua emissão é homogênea e isotrópica em grandes escalas, mas contém anisotropias que fornecem informações importantes sobre a estrutura inicial do Universo (FIXSEN, 2009). Na faixa de interesse dos radiotelescópios, a emissão da RCFM pode ser aproximada para o limite de Rayleigh-Jeans, o que faz com que sua emissão apareça como um fundo contínuo com pequenas flutuações em torno de 2,7 K.

Ao apresentar o referencial teórico adotado, observamos que a técnica de mapeamento da intensidade da emissão de 21 cm, proveniente do hidrogênio neutro, se mostra eficaz para medir este importante traçador de matéria em grande escala. Essa técnica é um método eficiente para examinar a evolução das oscilações acústicas de bárions ao longo da história cósmica e, assim, entender o efeito da expansão acelerada do Universo nas suas estruturas. Este estudo possibilita uma descrição mais precisa dos parâmetros cosmológicos e a exploração das propriedades do setor escuro. Daí decorre a importância de elaborar ferramentas direcionadas a essa técnica de mapeamento e enfrentar os desafios de medir sinais de intensidade muito baixa,

que são afetados por contaminações de vários agentes astrofísicos e instrumentais. Com isto em mente, no Capítulo 3, discutiremos os princípios da instrumentação em radioastronomia, os tipos diferentes de radiotelescópios, e suas limitações instrumentais inerentes. Também apresentaremos o radiotelescópio BINGO, um projeto brasileiro com significativas colaborações internacionais, que irá medir o sinal de HI e será avaliado como estudo de caso neste trabalho. Analisaremos como variadas abordagens de produção de mapas, relacionadas ao mapeamento de intensidade, podem auxiliar na melhor detecção do sinal de 21 cm medido pelo BINGO.

3 INSTRUMENTAÇÃO

O desenvolvimento e a execução de experimentos astrofísicos são altamente dependentes da precisão e confiabilidade dos instrumentos empregados. A instrumentação, em particular nos projetos de cosmologia observacional e radioastronomia, é projetada para detectar sinais extremamente fracos, os quais são mascarados pelos ruídos intrínsecos tanto do equipamento utilizado quanto do ambiente de observação. A caracterização e mitigação desses ruídos constituem etapas fundamentais para o êxito das medições e análises subsequentes.

Entre os desafios predominantes na instrumentação científica incluem-se a incidência de ruídos eletrônicos, inevitáveis devido às limitações físicas inerentes aos componentes e ao ambiente operacional e a presença de contaminantes astrofísicos, que para os estudos da CMB e da emissão de 21 cm são diversas ordens de grandeza mais intensos do que o sinal de interesse. O ruído térmico, derivado da movimentação de cargas elétricas em resistores e amplificadores e o ruído $1/f$, emergente em baixas frequências como consequência das flutuações no ganho dos instrumentos, são alguns dos principais ruídos sistemáticos envolvidos em observações de radiotelescópios e serão discutidos nas seções seguintes. A adoção de técnicas sofisticadas, tais como o resfriamento criogênico e o emprego de amplificadores de baixo ruído, possibilita atenuar esses efeitos, embora não os elimine por completo. Para além dos ruídos intrínsecos ao equipamento adotado, interferências externas, notadamente as interferências de radiofrequência (RFI), representam um desafio significativo. Fontes artificiais, tais como satélites, radares e dispositivos eletrônicos terrestres, introduzem sinais espúrios capazes de obscurecer ou mesmo saturar as observações. Assim, estratégias de mitigação como filtragem, monitoramento contínuo e a seleção de janelas espectrais menos contaminadas são essenciais para assegurar a qualidade dos dados coletados.

Portanto, uma compreensão aprofundada dos aspectos instrumentais e dos ruídos associados é fundamental para o desenvolvimento de técnicas eficazes de calibração e de processamento de dados. A combinação de uma instrumentação de alta sensibilidade com métodos robustos de processamento de sinais é essencial para a detecção e análise de fenômenos astrofísicos de baixa intensidade, tais como o problema científico que motiva esta dissertação. A seção seguinte apresentará um panorama sobre a instrumentação em radioastronomia, com destaque para o radiotelescópio BINGO, além de uma discussão dos ruídos sistemáticos que devem ser abordados para viabilizar a recuperação do sinal de HI.

3.1 Radiotelescópios

Os radiotelescópios são instrumentos cuja finalidade principal é detectar e analisar a radiação nos comprimentos de onda de rádio proveniente de diversas fontes astrofísicas. Em contraposição aos telescópios ópticos que operam na faixa do espectro visível, os radiotelescópios são capazes de captar sinais de baixa frequência, entre dezenas de MHz e dezenas de GHz, emitidos por fenômenos astrofísicos em nuvens de gás frio, pulsares, buracos negros e galáxias distantes.

O princípio de funcionamento desses instrumentos fundamenta-se na utilização de diferentes tipos de antenas, que atuam captando e direcionando as ondas de rádio para um receptor de alta sensibilidade situado no ponto focal. Entre os principais formatos das antenas destacam-se as antenas parabólicas, como as do *Green Bank Telescope* (GBT) (PRESTAGE et al., 2009) e do *Five-hundred-meter Aperture Spherical Telescope* (FAST) (NAN et al., 2011); as antenas dipolo, organizadas em arranjos como o *Low-Frequency ARray* (LOFAR) (HAARLEM et al., 2013) e o *Square Kilometre Array – Low Frequency* (SKA-Low) (DEWDNEY et al., 2009); e as estruturas cilíndricas, utilizadas por radiotelescópios como o *Canadian Hydrogen Intensity Mapping Experiment* (CHIME) (BANDURA et al., 2014) e o *Hydrogen Intensity and Real-time Analysis eXperiment* (HIRAX) (NEWBURGH et al., 2016). As antenas parabólicas e cilíndricas são também chamadas de refletoras, pois refletem o sinal observado. Instrumentos com apenas uma antena, como o FAST, são chamados de radiotelescópios de refletor único.

Além disso, os radiotelescópios podem operar de forma autônoma ou em rede, como no *Very Large Array* (VLA) (KASSIM et al., 1993), onde múltiplas antenas são utilizadas, resultando em imagens com alta resolução espacial. Cada antena da rede coleta os sinais de rádio simultaneamente, e os dados são posteriormente combinados, considerando diferenças no tempo de chegada dos sinais. Radiotelescópios com esse tipo de configuração são chamados de interferômetros. Essa técnica permite atingir uma resolução angular equivalente à de uma antena do tamanho da distância entre os radiotelescópios, possibilitando imagens detalhadas de objetos celestes.

Além das antenas parabólicas convencionais, existem também radiotelescópios que utilizam arranjos de dipolos, como nos observatórios de baixa frequência, a exemplo do LOFAR (*Low-Frequency Array*) (HAARLEM et al., 2013; RÖTTGERING, 2003). Esses instrumentos são formados por antenas menores, distribuídas em vastas áreas, que operam em conjunto para captar sinais de frequências muito baixas.

Finalmente, os radiotelescópios podem operar no modo de varredura ou de ponto fixo. No modo de varredura, o instrumento acompanha uma região do céu ao longo do tempo, permitindo o monitoramento de mudanças ou fenômenos transientes. Neste modo o telescópio permanece fixo e o céu se move acima dele, devido a rotação da Terra. No modo de ponto fixo, o radiotelescópio permanece direcionado para uma fonte específica, acumulando dados durante longos períodos, essencial para a detecção de sinais fracos e estáveis. Os radiotelescópios ainda podem ser divididos em duas categorias: fixos, cujas antenas permanecem estáticas, e móveis, que permitem direcionar as antenas para diferentes partes do céu desejadas para observação. O BINGO, instrumento de interesse deste trabalho, é um radiotelescópio que opera no modo de varredura, também chamado de telescópio de trânsito e possui antenas e refletores fixos.

No campo da cosmologia, os radiotelescópios desempenham um papel crucial no estudo da estrutura em larga escala do universo. A linha de 21 cm, observada na faixa de rádio, é amplamente utilizada para investigar as oscilações acústicas de bárions, oferecendo uma régua cósmica para medir a expansão do universo. Para este propósito está sendo construído no Brasil o radiotelescópio BINGO, instrumento utilizado no contexto deste trabalho para analisar as potencialidades e ganhos de uma modelagem mais complexa para os ruídos sistemáticos. A seção seguinte aborda os objetivos de sua construção, o local de observação, a configuração óptica do instrumento e outros detalhes de sua estratégia de observação e funcionamento.

3.2 Radiotelescópio BINGO

No contexto apresentado anteriormente, o radiotelescópio BINGO destaca-se como um importante instrumento brasileiro capaz de realizar medidas da emissão de 21 cm. O projeto BINGO, acrônimo para *Baryon Acoustic Oscillations In Neutral Gas Observations*, consiste em uma colaboração internacional liderada pelo Brasil para o desenvolvimento de um radiotelescópio com o objetivo de medir as oscilações acústicas de bárions através do mapeamento de intensidade do hidrogênio neutro. O radiotelescópio está em construção e será alocado na Serra do Urubu, no município de Aguiar, na Paraíba.

O instrumento realizará mapeamentos de intensidade do sinal de 21 cm do hidrogênio neutro no intervalo de *redshift* $0,127 \leq z \leq 0,449$, na faixa de frequências de 980 MHz a 1260 MHz, de forma a permitir traçar a distribuição de matéria em grande escala e assim estimar os parâmetros cosmológicos e delimitar propriedades da energia escura (WUENSCHÉ et al., 2022; ABDALLA et al., 2022a). Pretende-se, a partir

dos mapas de intensidade, produzir espectros de potência do céu em função da escala angular e do *redshift*, bem como um espectro de potência angular em 2D para cada fatia de *redshift* (BATTYE et al., 2013; BIGOT-SAZY et al., 2015; ABDALLA et al., 2022a; WUENSCHKE et al., 2022; DE MERICIA, Eduardo J. et al., 2023).

A banda de frequências em que o instrumento operará corresponde a um intervalo de *redshift*, o que permite acessar um volume cosmológico. Como cada frequência equivale a um *redshift*, ao analisar os mapas em cada frequência pode-se realizar uma “tomografia” do volume observado. Pretende-se com os dados do instrumento montar mapas de intensidade do sinal do hidrogênio e utilizá-los como um traçador de matéria, usando as réguas cosmológicas definidas pelas BAO como uma referência comparativa para a evolução da distribuição de matéria no universo, representando assim, uma maneira de realizar a observação da evolução cosmológica.

O instrumento será composto de um sistema ótico compacto, do tipo *Crossed-Dragone*, com um espelho principal refletor parabólico e um espelho secundário hiperbólico, ambos sem partes móveis, que iluminará o plano focal, com 28 cornetas. O instrumento foi planejado para ser um telescópio de trânsito, aproveitando-se do movimento da Terra para cobrir uma faixa do céu, o que permite revisitar tal faixa a cada dia. O campo de visão será tal que fornecerá uma imagem instantânea de $14,75^\circ(\text{DEC}) \times 6^\circ(\text{RA})$ (WUENSCHKE et al., 2022). Uma lista das características instrumentais e dos parâmetros do radiotelescópio BINGO está apresentada na Tabela 3.1, retirada de Wuensche et al. (2022) e produzida pela colaboração BINGO. Outros detalhes sobre as configurações técnicas do radiotelescópio podem ser obtidos em Abdalla et al. (2022b) e Wuensche et al. (2022).

O BINGO será composto de 2 espelhos, um espelho primário parabólico de semi-eixo maior igual a 25,5 m e um espelho secundário hiperbólico de semi-eixo maior igual a 18,3 m, além de 28 cornetas, organizadas em uma estrutura de colmeia conforme representado na Figura 3.4 produzida pela colaboração BINGO. A Figura 3.1 mostra o projeto de engenharia do instrumento, enquanto a Figura 3.2 mostra uma representação artística do BINGO no seu sítio na Serra do Urubu, na Paraíba. O espelho primário está voltado para a direção sul, com um deslocamento de -8° em relação ao zênite, que somado a latitude local, de -7° , resulta em uma declinação central do espelho de -15° . O sinal astrofísico incide sobre o espelho primário e é então refletido para o espelho secundário, que por sua vez está aproximadamente vertical em relação ao solo. O sinal é novamente refletido pelo secundário, sendo redirecionado para as cornetas, que servirão como um guia de onda para o sinal

Tabela 3.1: Parâmetros do projeto do BINGO - Fase 1.

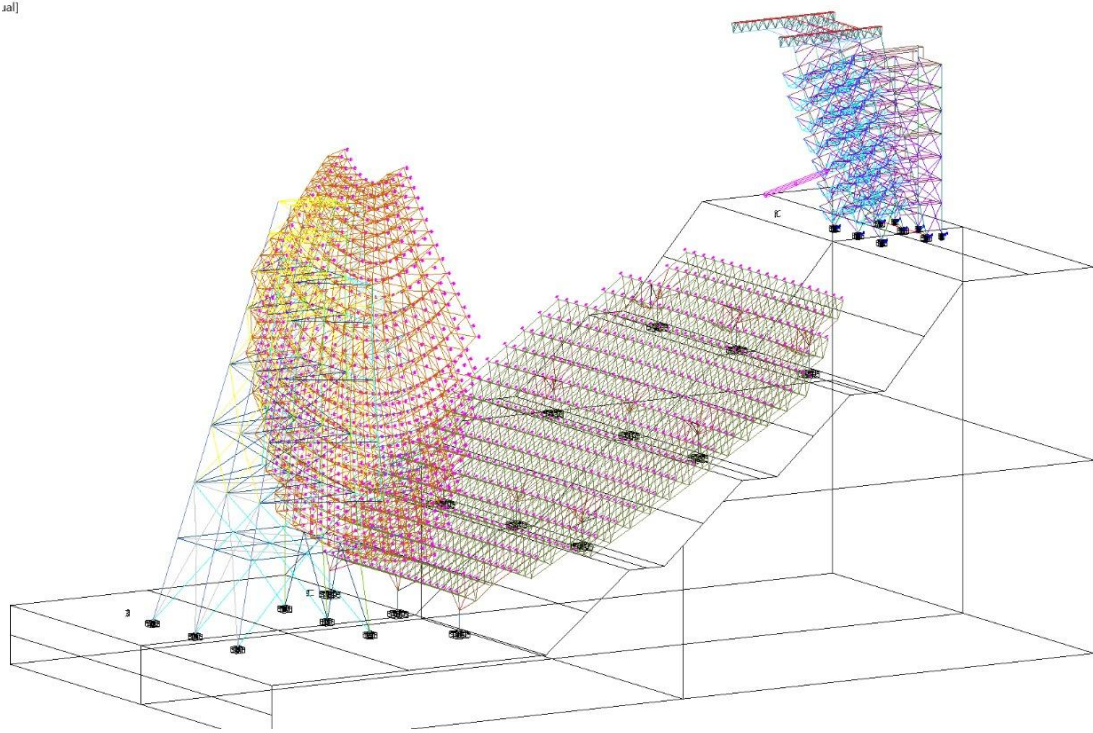
Parâmetro	Valor	Unidade
Receptor		
T_{sys} (Temperatura de sistema)	70	K
Menor frequência (f_{min})	980	MHz
Maior frequência (f_{max})	1260	MHz
Intervalo de frequência (Banda - B)	280	MHz
Número de intervalos de <i>redshift</i>	30	
Tempo de amostragem (τ_s)	10	Hz
Telescópio		
Espelho primário (Parabolóide)		
Semi-eixo maior	25.5	m
Semi-eixo menor	20.0	m
Concavidade	0.7	m
Espelho secundário (Hiperbolóide)		
Semi-eixo maior	18.3	m
Semi-eixo menor	17.8	m
Concavidade	0.5	m
Área instantânea observada	14.75×6.0	(Dec, RA) deg.
Ótica		
Distância focal	63.2	m
Angulo sólido do feixe (Ω_{beam})	0.35	sqr deg
Número de cornetas (Fase 1)	28	

(ABDALLA et al., 2022b).

As cornetas consistem em elementos cônicos corrugados (com ondulações), possuindo 1,7 m de diâmetro e 4,90 m de comprimento. Cada uma delas será sustentada por uma estrutura hexagonal, que irá compor o plano focal do instrumento. As corrugações buscam diminuir os lobos laterais do feixe e contribuir para a polarização circular do sinal. A Figura 3.3 mostra a representação de uma corneta individual, ligada ao polarizador e ao *magic-tee* e sustentada pela estrutura hexagonal. A Figura 3.4, por sua vez, apresenta o arranjo do plano focal para as 28 cornetas na Fase 1 do BINGO.

Associado às cornetas está o polarizador e o sistema eletrônico do receptor do instrumento, composto por amplificadores, filtros de banda, integradores do sinal, conversores analógicos-digitais e gravadores dos dados obtidos. Os elementos próximos a parte física do receptor, que equivale às cornetas, são usualmente chamados de *front end*, enquanto os mais próximos do gravador são chamados de *back end*. Uma representação da cadeia do receptor pode se encontrada na Figura 3.5.

Figura 3.1: Projeto de engenharia civil do arranjo do sistema óptico do radiotelescópio BINGO.



À esquerda, em um tom amarelo-rosado, está representado o espelho secundário. Ao centro está o espelho primário e à direita, o arranjo de cornetas e os receptores.

Os *front ends* do BINGO são compostos por um elemento de transição, polarizador e *magic tee*. O elemento de transição conecta a base da corneta ao polarizador. O polarizador, por sua vez, polariza linearmente o sinal, que possui uma polarização circular ao passar pelas cornetas. O sinal subsequentemente atravessa os *magic tee*, transformadores híbridos na forma de “T” que dividem as duas polarizações em duas cadeias distintas de receptores (WUENSCHÉ et al., 2022).

Na cadeia receptora, o sinal passa por amplificadores, que aumentam a relação sinal-ruído ao amplificar o sinal de interesse enquanto atenuam o ruído. Os filtros de banda filtram o sinal fora da banda de interesse, o que reduz a contaminação de outras frequências utilizadas para fins comerciais, como para uso de empresas de telefonia móvel e de tráfego aéreo. O sinal então passa por outro conjunto de híbridos, que desta vez recombina as duas polarizações e é então encaminhado para o *back end* do sistema eletrônico do radiotelescópio, que realiza a decomposição em frequências do sinal (WUENSCHÉ et al., 2022). Este passa por um conversor analógico-digital e é

Figura 3.2: Representação artística do arranjo do sistema óptico do radiotelescópio BINGO no seu sítio na Serra do Urubu.

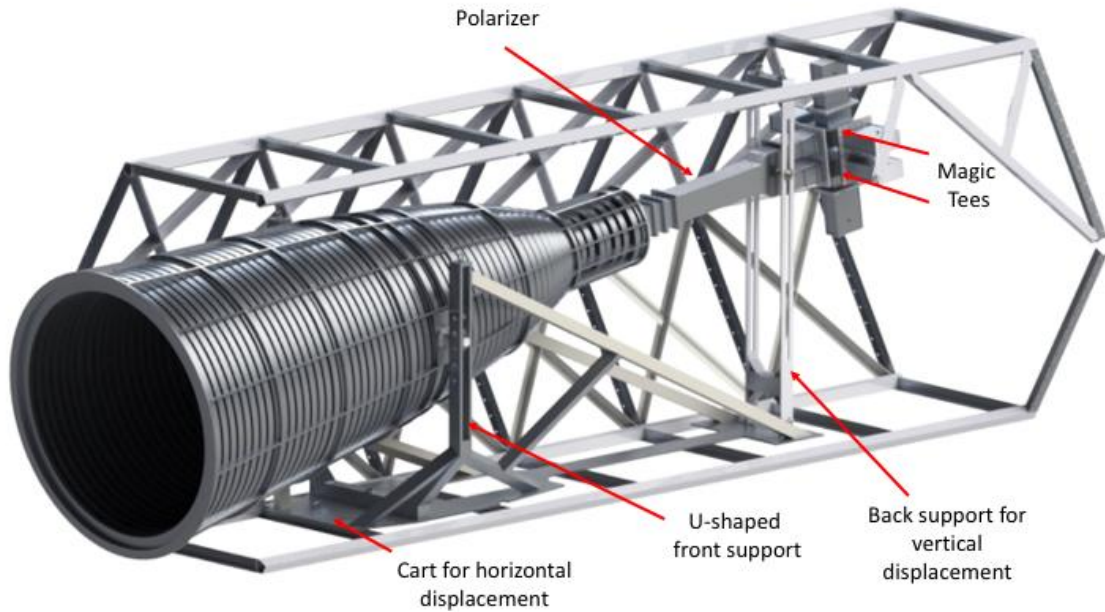


À esquerda são representadas as cornetas, enquanto ao centro e ao lado direito da imagem estão o espelho primário e secundário, respectivamente.

por fim gravado e integrado no tempo na forma de séries temporais.

A cadeia receptora do BINGO foi projetada para ser um receptor de correlação. Os amplificadores do instrumento introduzem um elemento de ruído correlacionado no tempo devido a flutuações no ganho dos amplificadores. Este ruído, chamado $1/f$, é melhor discutido nas seções seguintes e é tema de interesse da execução deste trabalho. A cadeia receptora de correlação busca minimizar o efeito do ruído ao diminuir a frequência de joelho (do inglês *knee frequency* (f_k)), que define o ponto em frequência em que o ruído $1/f$ deixa de dominar a distribuição espectral de potência e o ruído térmico passa a dominar (WUENSCHÉ et al., 2022). Esse processo diminui a escala, em frequência e em tempo, de domínio do ruído correlacionado. Foi adotada neste trabalho uma abordagem conservadora, na qual os correlacionadores não são capazes de lidar com o ruído correlacionado, tendo este de ser tratado em etapas seguintes da análise de dados.

Figura 3.3: Concepção artística de uma das cornetas do BINGO, ligada ao polarizador e ao *magic tee*.



Observa-se a estrutura hexagonal de suporte.

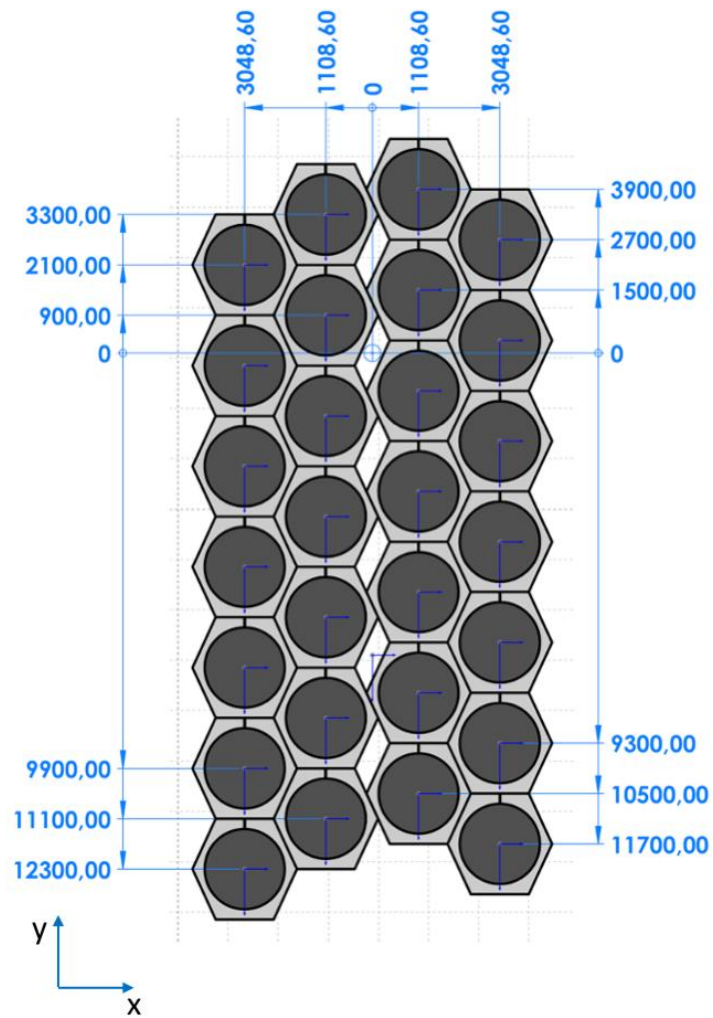
Fonte: Wuensche et al. (2022).

3.3 Ruídos instrumentais

Apesar de sua relevância para a pesquisa científica, a operação e a obtenção de dados medidos por radiotelescópios são processos que enfrentam desafios consideráveis, entre eles, a mitigação de ruídos instrumentais. Os ruídos instrumentais associados a um radiotelescópio são aqueles gerados dentro do próprio sistema receptor, independentemente do sinal coletado do céu ou do ambiente externo.

Entre os principais tipos de ruído instrumental estão o térmico e o $1/f$. Tais ruídos surgem de efeitos dos próprios componentes eletrônicos do instrumento e são inerentes ao seu funcionamento (BURKE et al., 2019). Em variações temporais de longo período, o ruído $1/f$ é especialmente problemático, pois domina a densidade espectral de potência (PSD, do inglês *Power Spectral Density*), tornando indispensável a utilização de ferramentas para minimizá-lo (SEIFFERT et al., 2002; MILOTTI, 2002). A seguir, são apresentados mais detalhes sobre os ruídos térmico e $1/f$.

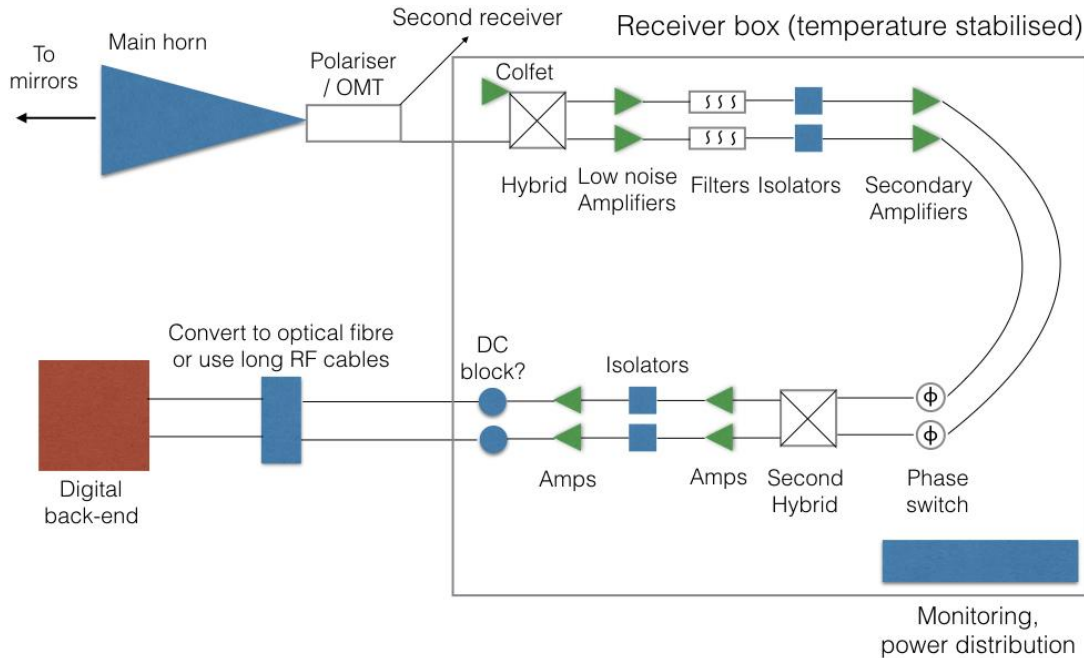
Figura 3.4: Representação do plano focal do BINGO com 28 cornetas.



As dimensões são dadas em milímetros.

Fonte: Wuensche et al. (2022).

Figura 3.5: Representação da cadeia receptora de correlação do radiotelescópio BINGO.



A região demarcada apresenta os amplificadores, os *magic tee* e os filtros de banda.
 Fonte: Wuensche et al. (2022).

3.3.1 Ruído térmico

O ruído térmico é uma fonte importante de interferência em experimentos e observações astrofísicas, especialmente na medição de sinais de baixa intensidade, como a CMB e o sinal de 21 cm do HI, que exigem alto nível de sensibilidade. O ruído térmico surge da agitação térmica de cargas presentes nos componentes eletrônicos do sistema devido às flutuações de Johnson-Nyquist, de forma que torna-se impossível que este não esteja presente em qualquer medida realizada utilizando tais componentes, já que é um fenômeno físico fundamentado nas leis da termodinâmica (NYQUIST, 1928). Essas flutuações térmicas podem afetar a precisão dos dados observacionais, tornando difícil a diferenciação entre o sinal de interesse e o ruído gerado pelo próprio equipamento. Sua PSD, que descreve a distribuição de energia do sinal ao longo do espectro de frequências, é aproximadamente constante (espectro plano).

A sensibilidade de um detector na faixa de radiofrequência, para o caso da presença apenas de ruído térmico aleatório e estatisticamente bem comportado, é dada pela equação do radiômetro. Ela define as variações de temperatura de brilho que o detector é capaz de medir acima do nível de ruído de fundo, que neste caso é dado pelo ruído térmico (BURKE et al., 2019). Na ausência de fonte, o detector ainda mediria um nível mínimo de sinal, advindo das agitações térmicas de cargas nos seus componentes. A intensidade deste sinal de fundo medido é dado pelo seu valor quadrático médio (do inglês, *root mean square - rms*). É possível distinguir apenas variações de temperatura que sejam maiores que a intensidade de tal ruído de fundo. Tem-se, portanto, que a sensibilidade de um detector é equivalente ao valor quadrático médio do ruído térmico deste mesmo sistema (BURKE et al., 2019).

A rms do ruído térmico em um dado canal de frequência ν e em uma dada direção do céu (ou em um determinado *pixel*) é dada por

$$\Delta T \approx \sigma_{\text{term}} = \frac{T_{\text{sys}}}{\sqrt{\delta\nu t_{\text{pix}}}}, \quad (3.1)$$

onde $\delta\nu$ é a largura de um canal de frequência, t_{pix} é o tempo de integração do respectivo *pixel* e T_{sys} é a temperatura de sistema, descrita por

$$T_{\text{sys}} = T_{\text{rec}} + T_{\text{sky}} + T_{\text{spill}} + T_{\text{ground}}, \quad (3.2)$$

onde T_{rec} é a contribuição do ruído do receptor, T_{sky} é a contribuição do céu (por exemplo, síncrotron galáctico e radiação cósmica de fundo), T_{spill} é a parcela associada à radiação captada fora da direção desejada (da atmosfera ou da estrutura do próprio instrumento) e T_{ground} é a componente relativa à radiação produzida pelo solo (BURKE et al., 2019).

O ruído térmico pode ser modelado por uma distribuição uniforme e segundo a Equação 3.1, o desvio padrão do mesmo é diretamente proporcional à temperatura de sistema (BURKE et al., 2019). Apesar de poder ser reduzido controlando-se a temperatura do receptor ou aumentando o tempo de observação, ele nunca pode ser completamente eliminado. Nos sistemas eletrônicos, esse ruído pode limitar o desempenho de dispositivos como amplificadores, sensores e equipamentos de comunicação, sendo um fator crítico na relação sinal-ruído (SNR, do inglês *signal-to-noise ratio*) de receptores de rádio, micro-ondas e outros sistemas sensíveis.

Tabela 3.2: Contribuições para T_{sys} .

Contribuições	T(K)
CMB	2.7
Sinais galácticos/extragalácticos	~ 5.0
Solo	2.0
Atmosfera	4.5
Refletor parabólico	1.5
Refletor hiperbólico	1.5
Cornetas (medido)	6.8
Polarizador (medido)	9.1
<i>Magic tee</i> (medido)	2.8
Elemento de transição (medido)	9.1
Amplificadores (medido)	24.9
Total	~ 69.9

Uma maneira de minimizar o impacto do ruído térmico é utilizar técnicas de refrigeração criogênica para manter os detectores a baixas temperaturas, reduzindo assim as flutuações térmicas. Outra abordagem é reduzir a largura de banda do sistema para limitar a quantidade de ruído captada, uma vez que a intensidade do ruído térmico é proporcional a largura de banda. Além disso, a utilização de técnicas de filtragem e processamento de sinais também é comum para separar os sinais desejados do ruído térmico, aumentando a relação sinal-ruído e melhorando a qualidade das observações.

No caso específico do radiotelescópio BINGO, foi realizada por parte da sua equipe uma medição no laboratório de cosmologia do Instituto Nacional de Pesquisas Espaciais (INPE), com o objetivo de estimar a temperatura de sistema do instrumento considerando todos os elementos do sistema receptor, como cornetas, polarizadores, elementos de transição, *magic tees* e receptores. Uma relação das temperaturas associadas a cada elemento instrumental e cada elemento de contaminação astrofísica é apresentada na Tabela 3.2, adaptada de Wuensche et al. (2022). As contribuições demarcadas foram medidas em laboratório enquanto as demais constituem estimativas teóricas. A soma de todas as contribuições resulta na temperatura de sistema do instrumento, T_{sys} , que foi estimada em 69,9 K.

3.3.2 Ruído 1/f

O ruído 1/f ocorre devido a flutuações de longa duração no ganho dos amplificadores do instrumento. Sua PSD é inversamente proporcional à frequência temporal das

flutuações, o que resulta em uma dominância desse ruído em baixas frequências temporais (SEIFFERT et al., 2002; HARPER et al., 2018). Em experimentos de CMB, o ruído $1/f$ aparece como um dos principais desafios para medir anisotropias de temperatura em escalas angulares maiores, podendo mascarar sinais cosmológicos sutis.

Nas séries temporais, o ruído $1/f$ aparece como pequenas variações multiplicativas em torno da temperatura de sistema que podem ser representadas por

$$\Delta T(t, \nu) = \delta G(t, \nu) T_{\text{sys}}(t, \nu), \quad (3.3)$$

onde $\Delta T(t, \nu)$ representa a flutuação de temperatura em função do tempo t e em um dado canal de frequência ν , associada ao ruído $1/f$, sendo uma combinação da flutuação instantânea no ganho, δG , e da temperatura de sistema, $T_{\text{sys}}(t, \nu)$. Nesse trabalho considerou-se a temperatura de sistema como constante (69,9 K) e igual para todos os canais de frequência.

A PSD do ruído térmico + $1/f$ é dada pela equação

$$\text{PSD}(f) = \frac{T_{\text{sys}}^2}{\delta \nu} \left[1 + \left(\frac{f_k}{f} \right)^\alpha \right], \quad (3.4)$$

onde T_{sys} representa a temperatura de sistema do instrumento, $\delta \nu$ é a largura de um canal de frequência, f é a frequência associada às flutuações temporais do ruído $1/f$ e o termo α é o índice espectral do ruído, assumido nesse trabalho como igual a 1, correspondendo ao chamado ruído rosa. Já f_k representa a frequência de joelho, definida como a frequência na qual as potências do ruído térmico (branco) e do ruído $1/f$ se igualam no domínio da frequência. Para $f \gg f_k$, o termo $\left(\frac{f_k}{f}\right)^\alpha$ se torna pequeno, e o comportamento do ruído tende a um valor constante, semelhante ao ruído branco. Para $f \ll f_k$, o termo $\left(\frac{f_k}{f}\right)^\alpha$ domina e o ruído segue o comportamento inverso típico do $1/f^\alpha$. O termo f_k pode ser interpretado como o limite acima do qual flutuações de longo prazo não têm mais influência significativa (o ruído branco passa a dominar) (MILOTTI, 2002).

Em experimentos de MI cada receptor mede o sinal em diversos canais de frequência, o que permite obter informação do sinal cosmológico ao longo de um intervalo de *redshift*. No entanto, isto também introduz uma correlação significativa no sinal

medido entre canais (HARPER et al., 2018) e no ruído $1/f$. Para incorporar essa correlação entre canais de frequência, a Equação 3.4 pode ser reformulada considerando uma nova lei de potências, expressa por

$$\text{PSD}(f, \omega) = \frac{T_{\text{sys}}^2}{\delta\nu} \left[1 + C(\beta, N_\nu) \left(\frac{f_k}{f} \right)^\alpha \left(\frac{1}{\omega\Delta\nu} \right)^{\frac{1-\beta}{\beta}} \right], \quad (3.5)$$

onde ω representa o número de onda associado à frequência espectroscópica inversa, $\Delta\nu$ é a largura de banda total do receptor, β corresponde ao índice espectral das correlações em frequência, e $C(\beta, N_\nu)$ é uma constante de normalização, cuja definição é apresentada adiante. O primeiro termo da Equação 3.5 representa a contribuição do ruído térmico para a PSD, enquanto o segundo termo corresponde à componente associada ao ruído $1/f$. Essa equação descreve uma PSD bidimensional, que leva em conta tanto as flutuações temporais quanto as correlações espectrais características do ruído $1/f$.

O índice espectral das correlações em frequência é estabelecido no intervalo $0 \leq \beta \leq 1$, onde $\beta = 0$ implica flutuações de ruído $1/f$ iguais em todos os canais de frequência, enquanto $\beta = 1$ caracteriza um ruído $1/f$ totalmente independente em cada canal (HARPER et al., 2018). A constante $C(\beta, N_\nu)$ é um fator de normalização que mantém a variância temporal do ganho, δG , constante, independentemente do valor de β adotado (HARPER et al., 2018) para modelar as correlações em frequência. Essa constante é expressa por

$$C(\beta, N_\nu) = \frac{N_\nu}{1 + \left(\frac{\sum_{i=1}^{N_\nu/2} H(\omega_i, \beta) \Delta\omega}{\int H(\omega, 0) d\omega} \right) (N_\nu - 1)}, \quad (3.6)$$

onde N_ν é o número de canais de frequência e $H(\omega, \beta)$ descreve a estrutura espectral do ruído $1/f$, conforme

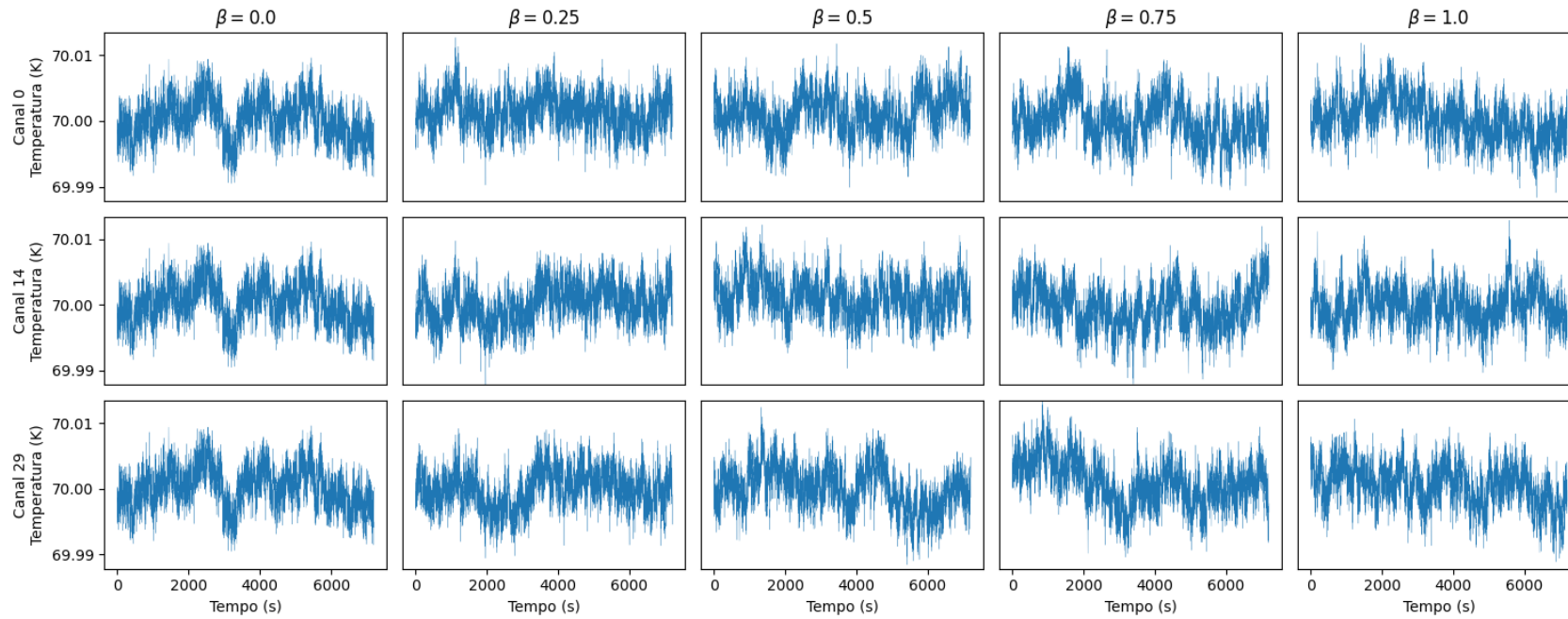
$$H(\omega, \beta) = \left(\frac{\omega_0}{\omega} \right)^{\frac{1-\beta}{\beta}}, \quad (3.7)$$

onde ω denota o modo de Fourier referente à frequência espectral (ou seja, o número de onda correspondente às variações ao longo dos canais de frequência), e ω_0 é o menor número de onda considerado, que corresponde ao modo fundamental da transformada discreta de Fourier, $1/\Delta\nu$, onde $\Delta\nu$ é a largura total da banda ob-

servada. Essa modelagem garante que, ao variar o parâmetro β , a função $H(\omega, \beta)$ controla o grau de correlação entre canais de frequência, conforme pode ser melhor visualizado na Figura 3.6: para $\beta = 1$, o ruído é branco em frequência (sem correlação), enquanto para $\beta = 0$, o ruído apresenta correlação máxima entre todos os canais (HARPER et al., 2018). Quando não há correlação, as séries temporais do sinal são totalmente independentes e a Equação 3.5 se reduz à Equação 3.4. Quando há correlação total, as flutuações são idênticas para todos os canais de frequência.

As variáveis apresentadas na Equação 3.5, bem como nas demais, representam parâmetros que contribuem para a estruturação do ruído $1/f$ e influenciam a intensidade do ruído na banda de frequências de interesse. Neste trabalho, analisaremos especialmente o efeito dos parâmetros β e f_k na recuperação do sinal cosmológico de 21 cm do HI. Com base no referencial teórico apresentado, a construção de mapas de intensidade confiáveis exige um tratamento adequado do ruído instrumental, de modo a evitar que suas flutuações em baixas frequências impeçam a recuperação do sinal de interesse. Para isso, algumas ferramentas de produção de mapas possuem rotinas computacionais específicas para atenuar ou remover o ruído $1/f$ ainda na etapa de processamento dos dados. No próximo capítulo, são discutidos os fundamentos dos métodos de *map-making* e de separação de componentes empregados neste trabalho.

Figura 3.6: Séries temporais produzidas com a IM *pipeline* para os canais de frequência 0, 14 e 29 da corneta 1, com $f_k = 0,001$ Hz e $\beta = \{0, 0,25, 0,50, 0,75, 1,0\}$, em um trecho de duas horas da observação total.



As TODs indicam que, à medida que o coeficiente β aumenta, a correlação entre canais diferentes diminui.

4 PRODUÇÃO DE MAPAS E SEPARAÇÃO DE COMPONENTES

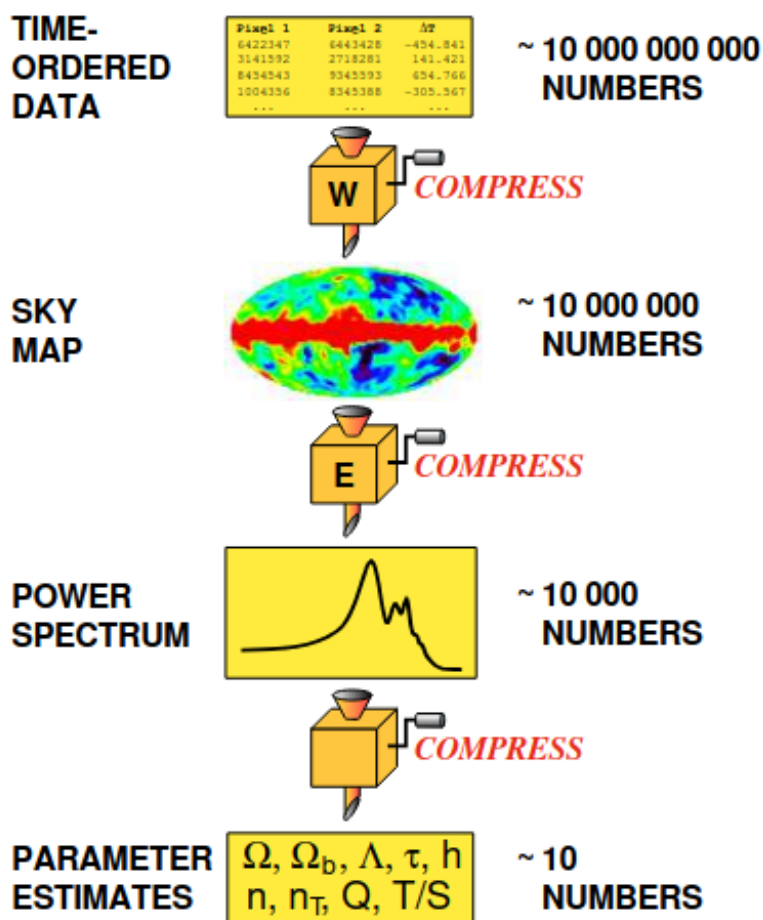
O processamento de dados em experimentos de mapeamento de intensidade é uma etapa fundamental para a recuperação do sinal cosmológico de interesse. Uma das etapas deste processamento, a produção de mapas de intensidade, envolve a conversão dos dados observacionais (ou simulados), organizados em função do tempo - conhecidos como séries temporais (*time ordered data* - TOD) - em mapas bidimensionais da emissão do céu, por meio de técnicas de produção de mapas (*map-making*). A produção dos mapas tem como objetivo organizar os dados de uma forma compacta, sem perda significativa de informação (TEGMARK, 1997a) e permitir a visualização da região observada. A partir desses mapas de intensidade, aplicam-se métodos de separação de componentes para isolar o sinal do HI dos *foregrounds* astrofísicos e do ruído instrumental.

A realização e aplicação de tais etapas de análise de dados, no entanto, possuem desafios a serem superados. Entre os desafios enfrentados nessa etapa, destaca-se a necessidade da aplicação de técnicas de mitigação dos ruídos, com especial interesse neste trabalho para a mitigação do ruído $1/f$, que apresenta forte correlação temporal e pode contaminar significativamente os dados em baixas frequências temporais, dificultando a recuperação do sinal cosmológico.

Como o *map-making* representa uma etapa intermediária crucial na análise de dados (entre a aquisição dos dados brutos e a estimativa dos parâmetros cosmológicos, conforme a Figura 4.1 extraída de Tegmark (1997b)), é imprescindível que os mapas do sinal cosmológico de 21 cm sejam pouco afetados por contaminações sistemáticas. Em experimentos como o BINGO, cujo objetivo principal é medir a escala de BAO com precisão e restringir as propriedades do setor escuro do universo, isso é especialmente relevante.

No contexto apresentado acima, é essencial a escolha de métodos eficientes de *map-making*, que sejam capazes de atenuar através de processamentos internos o ruído $1/f$ e preservar as características estatísticas do sinal cosmológico de 21 cm. Neste trabalho, são comparados os desempenhos de duas técnicas de *map-making*: o método *naive*, que consiste no cálculo da média direta das observações em cada *pixel*, e o método iGLS (*iterative Generalized Least Squares*), que utiliza o conhecimento das propriedades do ruído, como sua distribuição espectral de potência, para produzir um mapa da região observada com uma melhor relação sinal-ruído. Além disso, aliado a estes métodos de *map-making*, são testados dois métodos de separação de componentes: o *Principal Component Analysis* (PCA) e o *Generalized Needlet In-*

Figura 4.1: A imagem produzida por Tegmark (1997b) apresenta o processo de *map-making* como uma etapa intermediária na determinação de parâmetros cosmológicos.



No topo são representados os dados obtidos pelos instrumentos de observação, em grande volume e organizados na forma de séries temporais. As ferramentas de *map-making* atuam transformando as TOD's em mapas, compactando o volume de dados. Estes mapas são utilizados para o cálculo de quantidades estatísticas, como o espectro de potência do sinal, que por sua vez é utilizado na determinação dos parâmetros cosmológicos.

ternal Linear Combination (GNILC). Estes métodos foram escolhidos devido a sua simplicidade de operação, sua ampla utilização na análise de dados cosmológicos e à expertise do grupo de cosmologia do INPE para com as ferramentas.

Este capítulo inicia com a fundamentação teórica dos métodos de *map-making* e de separação de componentes avaliados nesse trabalho. Em seguida, descreve-se a *pipeline* de *intensity mapping* (IM *pipeline*) utilizada para gerar as séries temporais das observações simuladas do BINGO, bem como para produzir os mapas *naive*. Por fim, apresenta-se a ferramenta ROMA, baseada no método iGLS de geração de mapas de intensidade.

4.1 Métodos e fundamentos de *map-making*

Consideremos d_t a medida feita por um detector de um radiotelescópio no instante t . Supondo que em um determinado intervalo de tempo esse detector faça N_d medidas, pode-se organizar tais medidas realizadas por esse detector ao longo do tempo na forma de uma série temporal. A série temporal produzida por ele pode ser disposta em um vetor \mathbf{d} de dimensões $N_d \times 1$ dado por

$$\mathbf{d} = \begin{bmatrix} d_1 \\ \vdots \\ d_{N_d} \end{bmatrix}. \quad (4.1)$$

A construção de um mapa de intensidade, que nada mais é do que uma representação 2-D das medidas feitas sobre uma determinada região do céu, consiste em associar tais medidas organizadas na forma de série temporal às respectivas posições do céu de onde as mesmas foram obtidas. Essas regiões são discretizadas em *pixels*, que representam a menor unidade espacial de um mapa e servem como referência para atribuir as contribuições de cada região observada do céu. Cada pixel corresponde a uma direção no céu, definida por coordenadas angulares na esfera celeste. A associação é dada pelo esquema de pixelização do céu que foi adotado. Neste trabalho, adotou-se o modelo de pixelização do HEALPix (*Hierarchical Equal Area isoLatitude Pixelisation*), um modelo amplamente utilizado em cosmologia que discretiza a superfície da esfera na qual o céu é projetado na forma de pixels de mesma área (GORSKI et al., 2005).

No contexto do mapeamento de intensidade, o termo “intensidade” refere-se a uma grandeza astrofísica (como a temperatura de brilho) registrada ao longo do tempo

nas séries temporais produzidas pelos detectores do instrumento. Assim, um mapa com N_m *pixels*, em que m representa a grandeza medida, pode ser escrito como um vetor $N_m \times 1$ conforme

$$\mathbf{m} = \begin{bmatrix} m_1 \\ \vdots \\ m_{N_m} \end{bmatrix}. \quad (4.2)$$

Considerando que a relação entre as intensidades medidas e a série temporal é linear, conforme a abordagem adotada por [Wright et al. \(1996\)](#), pode-se representar o vetor série temporal na forma ([TEGMARK, 1997a](#); [POUTANEN et al., 2006](#))

$$\mathbf{d} = \mathbf{P}\mathbf{m} + \mathbf{n}, \quad (4.3)$$

onde \mathbf{P} é a matriz de apontamento do instrumento, de dimensões $N_d \times N_m$ e definida conforme as especificações do experimento e sua estratégia de observação. Cada elemento P_{tp} da matriz de apontamento relaciona a medida feita no instante t com a intensidade do pixel p do mapa através de

$$d_t = \sum_{p=1}^{N_m} P_{tp} m_p + n_t, \quad (4.4)$$

onde d_t é a medida feita pelo detector no instante t , P_{tp} simboliza um elemento da matriz de apontamento, m_p é a intensidade do pixel p do mapa do céu e n_t o ruído associado a cada medida. Tem-se, portanto, que os dados na forma de série temporal são definidos pela intensidade medida, a matriz de apontamento do instrumento (definida pelo arranjo dos detectores e pela estratégia de observação) e pelo ruído associado.

A matriz de apontamento pode incluir ou não a operação de convolução com o feixe do instrumento. No caso de um feixe simétrico, é possível considerar que o vetor \mathbf{m} já contempla os efeitos do mesmo no mapa de intensidade ([TRISTRAM et al., 2011](#); [NATOLI et al., 2001](#)). Assumindo ainda que o detector é do tipo potência total (como no caso do BINGO), que mede continuamente a potência total do sinal recebido pela antena e que ele mede apenas uma direção por vez, pode-se concluir que a matriz de apontamento nesse caso é esparsa, com um valor não nulo por linha, correspondente

ao pixel observado naquele instante (TEGMARK, 1997a). Como cada linha de \mathbf{P} está associada a uma medida d_t , a posição do elemento não nulo da mesma indica para qual *pixel* do mapa de intensidade o centro do feixe está apontando no instante t . Voltando à Equação 4.3, \mathbf{n} é o vetor de ruído, com dimensões $N_d \times 1$, dado por

$$\mathbf{n} = \begin{bmatrix} n_1 \\ \vdots \\ n_{N_d} \end{bmatrix}. \quad (4.5)$$

onde cada elemento n_t é a parcela de ruído instrumental presente em cada medida d_t , conforme Equação 4.4.

O equacionamento apresentado até aqui pode ser expandido para N_{det} detectores, considerando que nesse caso seriam produzidas N_{det} séries temporais (SUTTON et al., 2010). Assim, seja \mathbf{d}^i o vetor contendo a série temporal produzida pelo detector i e \mathbf{d} redefinido como o vetor composto pelos vetores contendo as séries temporais de todos os detectores, de forma que

$$\mathbf{d} = \begin{bmatrix} \mathbf{d}^1 \\ \vdots \\ \mathbf{d}^{N_{\text{det}}} \end{bmatrix}. \quad (4.6)$$

Seguindo o mesmo raciocínio para o vetor de ruído e para a matriz de apontamento, obtém-se

$$\mathbf{n} = \begin{bmatrix} \mathbf{n}^1 \\ \vdots \\ \mathbf{n}^{N_{\text{det}}} \end{bmatrix} \quad (4.7)$$

e

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}^1 \\ \vdots \\ \mathbf{P}^{N_{\text{det}}} \end{bmatrix}. \quad (4.8)$$

Conforme já apresentado no Capítulo 3, as medidas feitas pelos detectores são con-

taminadas principalmente pelo ruído de origem instrumental. Pode-se dizer que esse ruído possui uma parcela associada à agitação térmica das partículas portadoras de carga nos componentes eletrônicos (ruído térmico) e outra às flutuações de ganho dos amplificadores (ruído $1/f$).

O ruído térmico apresenta uma densidade espectral de potência (*power spectral density* - PSD) aproximadamente constante em uma ampla faixa de frequências, característica típica de um ruído branco. Além disso, suas amplitudes podem ser bem descritas por uma distribuição normal de média zero e variância finita, o que justifica sua modelagem como um ruído branco gaussiano.

Conforme discutido anteriormente, o vetor de ruído \mathbf{n} é formado por duas componentes: o vetor de ruído térmico e o vetor de ruído $1/f$, de forma que

$$\mathbf{n} = \mathbf{n}_{\text{term}} + \mathbf{n}_{1/f}. \quad (4.9)$$

Considerando o ruído térmico como estacionário (com propriedades estatísticas constantes no tempo), ele pode ser tratado como um ruído branco gaussiano, caracterizado por valor esperado nulo, ou seja, com flutuações em torno de zero, de forma que $\langle \mathbf{n}_{\text{term}} \rangle = \mathbf{0}$, e ausência de correlação temporal entre as suas realizações. Assim, a matriz de covariância temporal do ruído térmico, denotada por \mathbf{C}_{term} , possui dimensões $N_d \times N_d$ e é dada por

$$\mathbf{C}_{\text{term}} = \langle \mathbf{n}_{\text{term}} \mathbf{n}_{\text{term}}^T \rangle = \begin{bmatrix} \sigma_{\text{term}}^2 & & \\ & \ddots & \\ & & \sigma_{\text{term}}^2 \end{bmatrix} = \sigma_{\text{term}}^2 \mathbf{I}, \quad (4.10)$$

onde σ_{term} é o valor *rms* (*root mean square*) do ruído térmico, dado pela Equação 3.1, e \mathbf{I} é a matriz identidade.

O ruído $1/f$ surge devido a pequenas flutuações nos ganhos dos amplificadores do sistema receptor. Como o período dessas oscilações é muito maior que o tempo de amostragem do instrumento, o reflexo do ruído $1/f$ nos mapas de intensidade se dá na forma de estruturas em grandes escalas angulares. No caso do BINGO, devido à sua estratégia de varredura do céu (*drift scan*), o ruído $1/f$ aparece na forma de listras na direção de observação do instrumento (BIGOT-SAZY et al., 2015). Considerando o ruído $1/f$ como um processo estacionário, ou seja, cujas propriedades

estatísticas permanecem constantes ao longo do tempo, o mesmo apresenta correlações temporais significativas. Isso faz com que sua matriz de covariância, definida como $\mathbf{C}_{1/f} = \langle \mathbf{n}_{1/f} \mathbf{n}_{1/f}^T \rangle$, seja não esparsa (densa), apresentando elementos significativos fora da diagonal principal devido às correlações temporais. Os valores dos elementos dessa matriz diminuem com o aumento da separação temporal entre as medidas, refletindo a estrutura de correlação do ruído ao longo do tempo, conforme

$$C_{1/f,ij} = \langle n_{1/f,i} n_{1/f,j} \rangle = \sigma_{1/f}^2 \cdot r(\tau_{ij}), \quad (4.11)$$

onde $r(\tau_{ij})$ é a função de correlação temporal, que descreve o grau de correlação entre dois valores de um processo estocástico em dois instantes de tempo separados por $\tau_{ij} = |t_i - t_j|$ e é dado por

$$r(\tau_{ij}) = \langle n(t) \cdot n(t + \tau_{ij}) \rangle, \quad (4.12)$$

onde $n(t)$ representa o valor do ruído no instante t , τ_{ij} é a separação temporal e $\langle \cdot \rangle$ denota o valor esperado ou a média estatística calculada sobre várias realizações do ruído. Por definição, $r(0) = 1$, indicando correlação máxima quando avaliada no mesmo instante. Para o ruído térmico, $r(\tau_{ij}) = \delta(\tau_{ij})$, onde $\delta(\cdot)$ é a função delta de Dirac, o que indica que só a correlação no mesmo instante temporal é não nula. Para as demais separações temporais ($\tau \neq 0$), a correlação é nula. Já para o ruído $1/f$, $r(\tau)$ decai lentamente com τ , o que reflete a presença de correlações de longo alcance temporal.

Resumindo, o problema de *map-making* consiste em fazer a melhor estimativa possível do vetor \mathbf{m} , que aqui é chamada de $\hat{\mathbf{m}}$, a partir do vetor de medidas \mathbf{d} , da matriz de apontamento \mathbf{P} e do conhecimento *a priori* das características estatísticas do ruído (TEGMARK, 1997a).

Pode-se definir possíveis soluções para a Equação 4.3 adotando uma modelagem linear, dada por

$$\hat{\mathbf{m}} = \mathbf{Wd}, \quad (4.13)$$

em que \mathbf{W} representa a matriz $N_m \times N_d$ que descreve o método aplicado para a resolução e \mathbf{d} é o vetor que contém a série temporal. Substituindo a Equação 4.13

na Equação 4.3 e assumindo que $\mathbf{WP} = \mathbf{I}$, para que o erro de reconstrução seja independente do sinal (TEGMARK, 1997a), obtém-se

$$\hat{\mathbf{m}} = \mathbf{m} + \mathbf{W}\mathbf{n}. \quad (4.14)$$

Assim, ao minimizar o termo $\mathbf{W}\mathbf{n}$, faz-se a melhor estimativa $\hat{\mathbf{m}}$ do mapa \mathbf{m} . Em outras palavras, busca-se um método, representado por uma matriz \mathbf{W} que seja capaz de reproduzir com a maior fidelidade possível o céu real, ao mesmo tempo em que minimiza a contribuição do ruído \mathbf{n} . Existem diversos métodos para a construção de mapas produzidas durante observações com radiotelescópios, que foram descritos em mais detalhes em Tegmark (1997a). Neste trabalho, foram avaliados dois desses métodos: o método *naive* (TEGMARK, 1997a), que assume uma média simples das observações por *pixel*, e o método de máxima verossimilhança, com aplicação da técnica iGLS (OH et al., 1999; GASPERIS et al., 2005; CANTALUPO et al., 2010), que utiliza informações estatísticas sobre o ruído e simplifica a matriz de covariância do mapa para estimar mapas de intensidade com maior precisão. A seguir, são apresentados mais detalhes sobre ambas as abordagens.

4.1.1 Método *Naive*

Um dos principais desafios na construção de mapas de intensidade é a escolha do método adequado para processar as medidas obtidas ao longo do tempo para cada *pixel*. Como abordagem inicial, é comum utilizar o método *naive*, considerado um ponto de partida no estudo do processo de produção de mapas. Esse método consiste em calcular a média simples das observações registradas em cada *pixel* ao longo do tempo, assumindo esse valor como a estimativa final da emissão naquela respectiva direção do céu. Para isso, cada valor da série temporal correspondente ao *pixel* é somado e dividido pelo número total de observações por *pixel*, sem levar em conta correlações temporais ou espectrais do ruído.

O método *naive* pode ser descrito pela expressão

$$m_p = \frac{1}{N_{\text{obs},p}} \sum_{i \in \mathcal{J}_p} d_i, \quad (4.15)$$

onde m_p é o valor médio estimado no *pixel* p , $N_{\text{obs},p}$ é o número de medidas associadas ao *pixel* p , \mathcal{J}_p é o conjunto de índices das medidas associadas ao *pixel* p e d_i é o valor

da medida i dentro do conjunto de medidas associadas ao pixel p .

Para o método *naive* a matriz \mathbf{W} pode ser obtida ao se minimizar a diferença entre o mapa observado e o mapa esperado caso não houvesse nenhum ruído. Ela é dada por

$$\mathbf{W} = [\mathbf{P}^T \mathbf{P}]^{-1} \mathbf{P}^T. \quad (4.16)$$

Adotando o \mathbf{W} apresentado na Equação 4.16, a solução para a Equação 4.3 pode ser expressa pela equação

$$\hat{\mathbf{m}} = (\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T \mathbf{d}, \quad (4.17)$$

onde o produto $\mathbf{P}^T \mathbf{d}$ projeta os dados temporais no domínio de *pixels*, somando o valor de todas as amostras observadas em cada *pixel*. Em seguida, o termo $(\mathbf{P}^T \mathbf{P})^{-1}$ normaliza essa soma pelo número de observações de cada *pixel*, convertendo-a em uma média e produzindo assim o mapa *naive*. Substituindo a Equação 4.17 na Equação 4.3, obtém-se

$$\hat{\mathbf{m}} = \mathbf{m} + (\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T \mathbf{n}. \quad (4.18)$$

O termo $(\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T \mathbf{n}$ representa a contribuição do ruído para o mapa *naive* produzido. Essa contribuição depende diretamente das propriedades estatísticas do vetor de ruído \mathbf{n} . Como visto anteriormente, um ruído puramente térmico pode ser modelado como não correlacionado no tempo, com média igual a zero e variância constante, ou seja, $\langle \mathbf{n} \rangle = \mathbf{0}$ e $\langle \mathbf{n} \mathbf{n}^T \rangle = \sigma^2 \mathbf{I}$. Nesse caso, cada componente do ruído no mapa *naive* produzido pode ser interpretado como a média aritmética das amostras de ruído associadas a um determinado *pixel*, conforme

$$\left[(\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T \mathbf{n} \right]_p = \frac{1}{N_{\text{obs},p}} \sum_{i \in \mathcal{I}_p} n_i. \quad (4.19)$$

Como o ruído térmico é não correlacionado, a média dessas amostras tende a zero com o aumento de $N_{\text{obs},p}$. Assim, a contribuição do ruído térmico para o *pixel* p do mapa *naive* é, em média, nula, e decai proporcionalmente a $1/\sqrt{N_{\text{obs},p}}$. Dessa forma,

o método *naive* é eficiente para reduzir o impacto do ruído térmico, desde que o número de observações por *pixel* seja suficientemente grande.

Quando o ruído \mathbf{n} possui uma componente do tipo $1/f$, ele apresenta correlação temporal significativa. Nesse caso, as amostras de ruído associadas a um mesmo *pixel* deixam de ser estatisticamente independentes, de modo que a média dessas amostras não tende mais a zero, mesmo com grande número de observações. Isso faz com que o termo $(\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T \mathbf{n}$ contribua significativamente para o mapa final, introduzindo flutuações espúrias, especialmente em grandes escalas angulares (baixos multipolos). Como consequência, o mapa reconstruído pelo método *naive* fica contaminado com ruído de longa correlação temporal, comprometendo a recuperação do sinal cosmológico.

Apesar de sua simplicidade e baixo custo computacional, o método *naive* mostra-se ineficaz para mitigar ou remover a contribuição do ruído $1/f$. Durante a conversão dos dados temporais em mapas, esse método pode causar perda de informação, pois comprime os dados sem considerar as propriedades estatísticas do ruído (TEGMARK, 1997a). Para lidar com essas limitações, são empregados métodos mais sofisticados de *map-making*, que modelam explicitamente a correlação temporal do ruído $1/f$ através da sua matriz de covariância. Entre eles, destaca-se o método de Máxima Verossimilhança (*Maximum Likelihood* - ML), também conhecido como Mínimos Quadrados Generalizados (*Generalized Least Squares* - GLS), cuja versão iterativa (*iterative Generalized Least Squares* - iGLS) tem sido amplamente aplicada em experimentos de cosmologia. O método GLS foi utilizado, por exemplo, na análise dos dados do COBE (*COsmic Background Explorer*) (JANSSEN et al., 1993), enquanto o iGLS foi empregado em experimentos como o BOOMERanG (*Balloon Observations Of Millimetric Extragalactic Radiation And Geophysics*) Lange et al. (1995) e o Planck (TAUBER et al., 2004; NATOLI et al., 2001; GASPERIS et al., 2005). Neste trabalho, além do método *naive*, utilizou-se o iGLS, que é descrito com mais detalhes na próxima seção.

4.1.2 Método *iterative Generalized Least Squares* (iGLS)

O método GLS busca minimizar o erro quadrático ponderado pela inversa da matriz de covariância do ruído, dado por

$$\chi^2 = (\mathbf{d} - \mathbf{P}\mathbf{m})^T \mathbf{C}_n^{-1} (\mathbf{d} - \mathbf{P}\mathbf{m}), \quad (4.20)$$

onde $\mathbf{C}_n = \langle \mathbf{nn}^T \rangle$ é a matriz de covariância do ruído. Para encontrar o mínimo de χ^2 , basta derivar a Equação 4.20 com relação a \mathbf{m} e igualar a zero, conforme

$$\frac{\partial \chi^2}{\partial \mathbf{m}} = -2\mathbf{P}^T \mathbf{C}_n^{-1} (\mathbf{d} - \mathbf{Pm}) = \mathbf{0}. \quad (4.21)$$

Assim, a solução obtida com o método de GLS é

$$\hat{\mathbf{m}} = (\mathbf{P}^T \mathbf{C}_n^{-1} \mathbf{P})^{-1} \mathbf{P}^T \mathbf{C}_n^{-1} \mathbf{d}. \quad (4.22)$$

Para o caso particular em que o ruído possui apenas componente térmica, sua matriz de covariância é $\mathbf{C}_n = \sigma^2 \mathbf{I}$. Nessa situação, a solução se reduz exatamente àquela obtida com o método *naive*, dada por $\hat{\mathbf{m}} = (\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T \mathbf{d}$, conforme Equação 4.17.

Voltando à Equação 4.22, o produto $\mathbf{C}_n^{-1} \mathbf{d}$ atenua as contribuições dos modos temporais mais afetados pelo ruído, ou seja, aqueles de maior variância, como é o caso dos modos de baixa frequência (longa duração temporal), associados ao ruído $1/f$. Assim, o mapa de intensidade reconstruído com o método GLS preserva as componentes dos dados com maior relação sinal-ruído.

A função de verossimilhança no contexto de *map-making*, $\mathcal{L}(\mathbf{d} | \mathbf{m})$, é definida como a probabilidade de se obter o vetor de série temporal \mathbf{d} dado um mapa de intensidade \mathbf{m} . Considerando que o ruído siga uma distribuição gaussiana multivariada com média zero e matriz de covariância \mathbf{C}_n , pode-se escrever

$$\mathcal{L}(\mathbf{d} | \mathbf{m}) = \frac{1}{(2\pi)^{N_d/2} \det(\mathbf{C}_n)^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{d} - \mathbf{Pm})^T \mathbf{C}_n^{-1} (\mathbf{d} - \mathbf{Pm}) \right], \quad (4.23)$$

onde $\det(\mathbf{C}_n)$ é o determinante da matriz de covariância do ruído. Pelo teorema de Bayes (LUPTON, 1993), segue que

$$\mathcal{P}(\mathbf{m} | \mathbf{d}) \propto \mathcal{L}_m(\mathbf{d}) \mathcal{P}(\mathbf{m}), \quad (4.24)$$

onde $\mathcal{P}(\mathbf{m})$ e $\mathcal{P}(\mathbf{d} | \mathbf{m})$ são, respectivamente, as probabilidades *a priori* e *a posteriori* de \mathbf{m} . Assumindo $\mathcal{P}(\mathbf{m})$ uniforme (nenhum conhecimento *a priori* sobre o

mapa), conclui-se que o estimador com maior probabilidade *a posteriori* é o de máxima verossimilhança. Como o logaritmo é uma função crescente, pode-se maximizar o logaritmo da função de verossimilhança, dado por

$$\ln \mathcal{L}(\mathbf{d} \mid \mathbf{m}) = -\frac{1}{2} [(\mathbf{d} - \mathbf{P}\mathbf{m})^T \mathbf{C}_n^{-1} (\mathbf{d} - \mathbf{P}\mathbf{m})] - \frac{1}{2} \ln [\det(\mathbf{C}_n)] - \frac{N_d}{2} \ln(2\pi). \quad (4.25)$$

Assim, maximizar a Equação 4.25 é o mesmo que minimizar a Equação 4.20, ou seja, assumindo ruído gaussiano com média zero e com covariância conhecida, o método de máxima verossimilhança se reduz ao problema dos mínimos quadrados generalizados, cuja solução é dada pela Equação 4.22.

Voltando à solução obtida com o método GLS, descrita na Equação 4.22, observa-se que o termo $\mathbf{P}^T \mathbf{C}_n^{-1} \mathbf{P}$ resulta em uma matriz de dimensão $N_p \times N_p$, onde N_p é significativamente menor que o número total de amostras temporais N_d . Além disso, a matriz \mathbf{P} é esparsa, o que reduz consideravelmente o custo computacional associado às multiplicações matriciais. Por outro lado, a matriz \mathbf{C}_n , de dimensões $N_d \times N_d$, é densa no caso de ruído $1/f$, o que torna a sua inversão direta inviável, tanto pelo tempo de processamento como pelo uso de memória computacional.

Para superar essa limitação, ao invés de calcular \mathbf{C}_n^{-1} , transforma-se o produto por essa matriz como uma operação linear sobre vetores. Essa operação pode ser implementada através de técnicas como a Transformada Rápida de Fourier (*Fast Fourier Transform* - FFT), filtragem no domínio do tempo ou da frequência, ou ainda por meio da resolução iterativa de sistemas lineares utilizando métodos como o do Gradiente Conjugado (*Conjugate Gradient* - CG), combinado muitas vezes com a aplicação de um pré-condicionamento, ou seja, uma transformação do sistema de equações original em outro melhor condicionado, visando acelerar a convergência.

O método de Mínimos Quadrados Generalizados Iterativos (*iterative Generalized Least Squares* - iGLS) resolve o problema de *map-making* sem construir nem inverter as matrizes de grandes dimensões. Isso é feito resolvendo um sistema linear $\mathbf{A}\mathbf{x} = \mathbf{b}$ dado por

$$\mathbf{P}^T \mathbf{C}_n^{-1} \mathbf{P} \hat{\mathbf{m}} = \mathbf{P}^T \mathbf{C}_n^{-1} \mathbf{d}, \quad (4.26)$$

onde

$$\mathbf{A} = \mathbf{P}^T \mathbf{C}_n^{-1} \mathbf{P} \quad (4.27)$$

e

$$\mathbf{b} = \mathbf{P}^T \mathbf{C}_n^{-1} \mathbf{d}, \quad (4.28)$$

usando um método iterativo, como por exemplo o CG (OH et al., 1999). Isso é feito começando por uma etapa de preparação, onde são definidos os seguintes operadores:

- matriz de apontamento \mathbf{P} : relaciona o espaço de *pixels* com o espaço dos dados temporais (TOD), projetando um mapa de céu simulado no domínio temporal de acordo com a estratégia de varredura do instrumento;
- matriz de apontamento transposta \mathbf{P}^T : realiza a operação inversa, projetando os dados temporais (TOD) no espaço de *pixels*, acumulando todas as observações de cada *pixel* do mapa;
- matriz de covariância do ruído inversa \mathbf{C}_n^{-1} : fornece o peso ótimo a ser aplicado nas observações, suprimindo frequências dominadas pelo ruído; na prática, sua aplicação é realizada no espaço de Fourier, por meio de FFT's.

Feito isso, o próximo passo é a inicialização do método CG. Primeiro, adota-se uma estimativa inicial para o mapa \mathbf{m}_0 , que pode ser, por exemplo, um mapa nulo ou um mapa *naive*. Em seguida, calcula-se o resíduo inicial, conforme

$$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{m}_0. \quad (4.29)$$

Por fim, define-se a direção inicial \mathbf{p}_0 , ou seja, a primeira direção de busca no espaço das soluções, que coincide com a direção do gradiente descendente (direção em que o erro decresce mais rapidamente a partir da estimativa inicial \mathbf{m}_0) dada por

$$\mathbf{p}_0 = \mathbf{r}_0. \quad (4.30)$$

Em seguida, inicia-se a etapa das iterações segundo o método CG. Seja i o índice das

iterações que devem ser repetidas até a convergência. Primeiro, calcula-se o operador do sistema sobre a direção \mathbf{p}_i , conforme

$$\mathbf{q}_i = \mathbf{A}\mathbf{p}_i. \quad (4.31)$$

Isso é feito em três etapas: primeiro calcula-se o vetor $\mathbf{p}_{\text{TOD}} = \mathbf{P}\mathbf{p}_i$; depois, aplica-se a matriz de covariância do ruído por meio de FFT, ou seja, sem inverter a mesma, de forma que $\mathbf{p}'_{\text{TOD}} = \mathbf{C}_{\mathbf{n}}^{-1}\mathbf{p}_{\text{TOD}}$; por fim, calcula-se $\mathbf{q}_i = \mathbf{P}^T\mathbf{p}'_{\text{TOD}}$.

Obtido o operador \mathbf{q}_i , calcula-se o passo, dado por

$$\alpha_i = \frac{\mathbf{r}_i^T \mathbf{r}_i}{\mathbf{p}_i^T \mathbf{q}_i}; \quad (4.32)$$

depois, atualiza-se a solução

$$\mathbf{m}_{i+1} = \mathbf{m}_i + \alpha_i \mathbf{q}_i \quad (4.33)$$

e o resíduo

$$\mathbf{r}_{i+1} = \mathbf{r}_i - \alpha_i \mathbf{q}_i. \quad (4.34)$$

Feito isso, verifica-se a convergência conforme

$$\|\mathbf{r}_{i+1}\| / \|\mathbf{b}\| < \varepsilon. \quad (4.35)$$

O processo iterativo é interrompido quando a razão entre a norma do vetor de resíduo $\|\mathbf{r}_{i+1}\|$ e a norma do vetor $\|\mathbf{b}\|$, definido na Equação 4.28, se torna inferior à tolerância ε . Caso a convergência não seja atingida, calcula-se o fator β_i , dado por

$$\beta_i = \frac{\mathbf{r}_{i+1}^T \mathbf{r}_{i+1}}{\mathbf{r}_i^T \mathbf{r}_i}, \quad (4.36)$$

atualiza-se a direção, conforme

$$\mathbf{p}_{i+1} = \mathbf{r}_{i+1} + \beta_i \mathbf{p}_i. \quad (4.37)$$

e avança-se para a próxima iteração.

O método CG não necessita que grandes matrizes sejam formadas explicitamente. Todas as suas operações dependem de produtos matriciais rápidos, uma vez que a matriz de apontamento \mathbf{P} é esparsa e que o produto da inversa da matriz de covariância do ruído \mathbf{C}_n por um vetor \mathbf{v} pode ser calculado por meio de FFTs. Para um ruído estacionário e com covariância que dependa somente do intervalo temporal, conforme já discutido anteriormente, a matriz \mathbf{C}_n é do tipo Toeplitz, ou seja, constante ao longo das diagonais. Se, além disso, o ruído for periódico, \mathbf{C}_n pode ser aproximada como uma matriz circulante. Assim, pode-se usar uma FFT para aplicar \mathbf{C}_n^{-1} como um filtro no domínio da frequência, conforme

$$\mathbf{C}_n^{-1} \mathbf{v} \approx \mathfrak{F}^{-1} \left[\frac{1}{P(f)} \cdot \mathfrak{F}(\mathbf{v}) \right], \quad (4.38)$$

onde $\mathfrak{F}(\cdot)$ representa a Transformada de Fourier, $\mathfrak{F}^{-1}(\cdot)$ a sua inversa, e $P(f)$ é a densidade espectral de potência do ruído, que pode ser estimada a partir do conhecimento prévio sobre as características do instrumento. Assim, multiplicar a matriz \mathbf{C}_n^{-1} pelo vetor \mathbf{v} equivale a transformar este para o domínio da frequência através de FFT, dividir cada componente pela densidade espectral de potência e transformar o resultado disso de volta para o domínio do tempo por meio de FFT.

A matriz $\mathbf{P}^T \mathbf{C}_n^{-1} \mathbf{P}$ é tipicamente mal condicionada devido às correlações temporais do ruído $1/f$, o que torna a convergência do método CG lenta. Para evitar isso, utiliza-se uma matriz pré-condicionadora \mathbf{M} , mais simples e de fácil inversão, que melhora o condicionamento do sistema, acelerando a convergência do método, conforme

$$\mathbf{M}^{-1} \mathbf{A} \mathbf{x} = \mathbf{M}^{-1} \mathbf{b}. \quad (4.39)$$

A matriz \mathbf{M} deve ser uma boa aproximação de \mathbf{A} ou seja, $\mathbf{M}^{-1} \mathbf{A} \approx \mathbf{I}$. Além disso, a matriz pré-condicionadora deve ser simples e de rápida inversão ou aplicação. Ela atua como um filtro, reduzindo a variância dos autovalores de $\mathbf{M}^{-1} \mathbf{A}$, o que melhora a eficiência do método CG. O uso combinado de uma técnica de pré-condicionamento

com o método CG é conhecido como Conjugado Gradiente Pré-condicionado (*Pre-conditioned Conjugated Gradient* - PCG) (CANTALUPO et al., 2010).

4.2 Ferramentas de simulação

A produção dos mapas de intensidade a partir de dados simulados ou observacionais pode ser realizada por meio de diferentes rotinas computacionais, cada uma implementando métodos específicos para etapas como simulação de séries temporais, *map-making* e separação de componentes. Essas ferramentas requerem arquivos de entrada que contêm informações como parâmetros instrumentais, estratégia de observação, características do ruído e modelos astrofísicos, os chamados *inputs*. A partir desses *inputs*, as rotinas computacionais executam as operações necessárias, produzindo arquivos de saída (*outputs*), como, por exemplo, os mapas de intensidade reconstruídos. Nesta subseção são apresentadas as ferramentas computacionais utilizadas neste trabalho, incluindo seu funcionamento, método implementado e parâmetros de entrada de cada uma.

4.2.1 IM *pipeline*

A *pipeline* utilizada para simular as observações do BINGO para *intensity mapping*, IM *pipeline*, combina um conjunto de parâmetros de entrada para simular as observações realizadas pelo instrumento. Ela foi desenvolvida na Universidade de Manchester e está sendo aprimorada pelo INPE, permitindo simular a operação de outros experimentos de IM além do BINGO. Os parâmetros de entrada são definidos no arquivo de parâmetros da simulação, que contém informações do instrumento, da estratégia de observação, do espectro de potência do sinal de HI, dos contaminantes astrofísicos e dos ruídos sistemáticos. Este arquivo define a estratégia de observação e as características espectrais dos sinais considerados, garantindo que a simulação esteja alinhada com as condições reais do instrumento. A *pipeline* então combina tais parâmetros e reproduz o que o instrumento observaria, retornando as séries temporais das observações e os mapas de intensidade *naive* (BIGOT-SAZY et al., 2015; BATTYE et al., 2013).

Os *inputs* para a simulação são o modelo do sinal de HI, o modelo de emissão dos *foregrounds* astrofísicos e do ruído instrumental. O sinal cosmológico de HI é simulado a partir do seu espectro de potência angular, que é gerado a partir dos parâmetros cosmológicos adotados (ABDALLA et al., 2022a).

Os principais contaminantes astrofísicos na faixa de frequências do BINGO são as

emissões síncrotron, devido aos raios cósmicos produzidos por elétrons, a CMB, *bremstrahlung* do gás difuso ionizado, provenientes da emissão difusa da galáxia e as fontes extragalácticas de rádio pontuais, fracas e não resolvidas (galáxias e quasares), que dominam o céu fora do plano galáctico, contaminando os mapas de intensidade em torno de ~ 1 GHz e dificultando a recuperação do sinal de HI. Seus sinais de emissão são simulados através de modelos espectrais bem estabelecidos, que foram apresentados na Seção 2.5. A IM *pipeline* convolui cada componente do sinal astrofísico com o feixe do instrumento e soma suas contribuições pixel a pixel. Pode-se modelar o sinal recebido pelo instrumento pela expressão

$$T_{\text{sky}}(\nu, p) = T_{\text{CMB}}(\nu, p) + T_{\text{gal}}(\nu, p) + T_{\text{ps}}(\nu, p) + T_{\text{HI}}(\nu, p), \quad (4.40)$$

na qual $T_{\text{sky}}(\nu, p)$ é o sinal obtido, representado pela temperatura de brilho do céu. $T_{\text{CMB}}(\nu, p)$ é a temperatura da radiação cósmica de fundo em micro-ondas, $T_{\text{gal}}(\nu, p)$ é a temperatura da radiação galáctica difusa, $T_{\text{ps}}(\nu, p)$ é a temperatura de emissão de fontes pontuais extragalácticas e $T_{\text{HI}}(\nu, p)$ é a temperatura de emissão do HI. Assim, $T_{\text{sky}}(\nu, p)$ é a temperatura de brilho na frequência ν e na direção p do céu.

Após a construção do sinal simulado conforme será observado pelo instrumento, são somadas a este as contribuições do ruído sistemático, constituído pelo ruído térmico e pelo ruído $1/f$. O ruído também é simulado através de modelos espectrais, apresentados na Seção 3.3. O sinal total é composto pelo sinal de 21 cm do HI, por contaminantes astrofísicos e por ruído instrumental (ABDALLA et al., 2022a; WUENSCHKE; COLLABORATION, 2019). Após a simulação, este sinal é disposto ao longo do tempo na forma de séries temporais. As séries temporais podem ser definidas como um conjunto de medidas feitas em um dado intervalo de tempo. No caso do BINGO, esse conjunto de dados é composto pelas medidas de temperatura em um dado canal ($\Delta\nu_i$) no intervalo 980-1260 MHz, numa dada coordenada celeste (α, δ) e em um dado instante de observação (t_i) (DE MERICIA, Eduardo J. et al., 2023).

As TODs constituem um dos *outputs* da IM *pipeline*. Outro conjunto de *outputs* são os mapas de emissão, gerados no formato HEALPix a partir das TODs por meio da técnica *naive* de *map-making*, conforme descrito na Seção 4.1.1. A IM *pipeline* empregada neste trabalho produz separadamente os mapas de *foregrounds* astrofísicos, o mapa de HI e o mapa da observação simulada, que inclui, além dos dois mapas anteriores, a contribuição do ruído instrumental. Os mapas podem ser

definidos como uma projeção cartográfica da distribuição das temperaturas médias observadas em cada direção do céu (DE MERICIA, Eduardo J. et al., 2023), descrita por 4.15.

4.2.2 ROMA - Roma Optimal Map-making Algorithm

O algoritmo ROMA constitui uma ferramenta para a produção de mapas, desenvolvida com o propósito de processar e reduzir dados provenientes de múltiplos canais de frequência em experimentos relacionados à CMB. O algoritmo é capaz de tratar medidas de anisotropia e polarização, que são essenciais para a extração de informações cosmológicas das observações da CMB (NATOLI et al., 2001). Destaca-se no ROMA a sua implementação em processamento paralelo, a qual permite que os cálculos sejam distribuídos entre diferentes processadores, possibilitando assim o tratamento eficiente de grandes volumes de dados e a implementação de uma etapa de pré-condicionamento, que simplifica a matriz de covariância do ruído, conforme apresentado na Subseção 4.1.2. O algoritmo emprega o método GLS de forma iterativa, para a minimização do ruído correlacionado e a otimização da precisão dos mapas reconstruídos (GASPERIS et al., 2005; NATOLI et al., 2001). O ROMA foi aplicado com êxito em dados reais e simulados do experimento BOOMERanG (LANGE et al., 1995) e do Planck (NATOLI et al., 2001), atestando a sua capacidade de produzir mapas de alta fidelidade das anisotropias e padrões de polarização da CMB (GASPERIS et al., 2005). Para a aplicação neste trabalho, o algoritmo foi adaptado para gerar mapas de intensidade de HI, com rotinas computacionais que permitiam que o algoritmo produzisse os mapas para os 30 canais em uma única realização.

O funcionamento do ROMA se baseia em algumas suposições fundamentais. Primeiramente, assume-se que o feixe do instrumento é axisimétrico, apresentando simetria em torno do eixo óptico, de forma a manter a estrutura da matriz de apontamento simples (GASPERIS et al., 2005). Em relação ao ruído, considera-se que os dados temporais apresentam ruído estacionário, cujas propriedades estatísticas não mudam com o tempo, e não correlacionado entre diferentes detectores (GASPERIS et al., 2005). Também é importante notar que a resolução do ROMA para o problema de *map-making* baseia-se em um experimento de receptor único e, para o caso de múltiplos detectores, a sua cobertura de céu deve ser ao menos minimamente sobreposta (NATOLI et al., 2001).

Para resolver o problema inicial do *map-making* apresentado na Equação 4.3, o ROMA utiliza o método dos mínimos quadrados generalizados iterativos, minimi-

zando o erro de recuperação do mapa χ^2 , conforme apresentado na Seção 4.1.2. Para lidar com a complexidade da matriz de covariância do ruído na presença do ruído $1/f$, é necessário aplicar a técnica de minimização do gradiente conjugado, apresentada em detalhes em Axelsson (1987). No entanto, o conjunto de dados precisa ser pré-condicionado para garantir que a técnica convirja. Para tal, o algoritmo aplica um pré-condicionador nos dados ao simplificar a matriz de covariância do ruído, \mathbf{C}_n , definindo uma nova matriz, mais simples que a original, que é chamada de pré-condicionador, \mathbf{M} . O pré-condicionador é escolhido de forma que se aproxime da matriz original, mas que seja mais rápido de computar e inverter. Para o ROMA, a matriz de pré-condicionamento adotada é dada em geral pela parte diagonal da matriz de covariância do mapa, que é simétrica, positiva e quase totalmente diagonal para o ruído térmico. Mais detalhes sobre a implementação da técnica do gradiente conjugado e da definição do pré-condicionante no algoritmo ROMA podem ser encontradas na Subseção 4.1.2 e em Natoli et al. (2001). Aplicada a técnica, a ferramenta resolve o sistema da Equação 4.3 e produz os mapas de intensidade a partir das TODs da simulação e de uma estimativa do ruído.

A utilização do ROMA para produção do mapa iGLS, por sua vez, envolve uma série de etapas que asseguram a conversão e o processamento adequados dos dados de entrada. Os arquivos de entrada do ROMA consistem nas TODs das observações, na matriz de apontamento dos receptores do instrumento e de uma estimativa da distribuição espectral inversa de potência do ruído (*Inverse Power Spectral Density - IPSD*). As TODs são geradas utilizando a própria IM *pipeline*. A matriz de apontamento é gerada através das TODs e identifica quais pixels do mapa do céu estão sendo observados por quais detectores do instrumento. A IPSD, por sua vez, é gerada através de um código em Python disponibilizado pelos autores do algoritmo e apresentado em Gasperis et al. (2005) e Natoli et al. (2001), a partir da Equação 3.5 e adaptado neste trabalho para receber valores de β diferentes de 1.

Para obter-se os arquivos de entrada descritos acima, se faz necessário converter as TODs geradas pela *pipeline*, que originalmente estão no formato HDF5 (*Hierarchical Data Format version 5*), para o formato FITS (*Flexible Image Transport System*). Essa conversão é crucial, uma vez que o formato FITS assegura a compatibilidade com as ferramentas de análise de dados empregadas na reconstrução do mapa. A seguir, é preciso montar um arquivo FITS que contenha a matriz de apontamento do instrumento, utilizando as informações presentes na TOD. A matriz de apontamento é essencial para associar cada medida temporal registrada pelo instrumento a uma posição específica no céu. Para sua construção, os dados de apontamento

do telescópio são extraídos da TOD e organizados de forma estruturada no novo arquivo FITS.

O passo seguinte envolve a simulação do ruído instrumental. Para isso, utiliza-se a IM *pipeline* para gerar conjuntos de TODs sintéticos contendo apenas o ruído esperado nas observações reais. A partir dessas simulações, é possível estimar a densidade espectral de potência inversa do ruído (*Inverse Power Spectrum Density* - IPSD), parâmetro crucial para a mitigação de artefatos sistemáticos na reconstrução do mapa. A IPSD quantifica a distribuição do ruído em função da frequência, permitindo a construção da matriz de covariância e seu uso na resolução do método, permitindo que o algoritmo ROMA corrija parte dos efeitos do ruído e melhore a qualidade do mapa final (GASPERIS et al., 2005; NATOLI et al., 2001).

Com os arquivos necessários devidamente preparados, o algoritmo ROMA é então executado, recebendo como entrada as TODs da simulação, a matriz de apontamento do instrumento e a estimativa da IPSD do ruído. O ROMA emprega um método fundamentado no princípio de máxima verossimilhança para reconstruir a distribuição do sinal celeste, utilizando as informações da matriz de apontamento para mapear os dados observacionais. A implementação da IPSD permite que o algoritmo ajuste o impacto do ruído instrumental, aprimorando a recuperação do sinal real de HI, enquanto a aplicação do pré-condicionador simplifica a matriz de covariância do mapa e permite a resolução da equação de *map-making* (NATOLI et al., 2001). Por fim, a saída do ROMA consiste em um conjunto de mapas do tipo iGLS, com um mapa gerado para cada canal de frequência observado. Esses mapas retratam a distribuição do sinal reconstruído e podem ser posteriormente analisados para a extração de parâmetros cosmológicos e astrofísicos (NATOLI et al., 2001; GASPERIS et al., 2005).

4.3 Métodos de separação de componentes

O objetivo da separação de componentes é isolar o sinal de HI dos contaminantes, para que este possa ser recuperado e estudado. Existem diversos contaminantes do sinal na faixa de frequência de rádio, alguns exemplos são as emissões galácticas, como a emissão anômala de poeira (*Anomalous Microwave Emission* - AME) e extragalácticas, como a emissão de rádio-galáxias, além da interferência em rádio frequência (*Radio Frequency Interference* - RFI) (OLIVARI et al., 2016). Além disso, a intensidade do sinal de HI é consideravelmente menor do que a dos contaminantes, especialmente quando comparado às emissões síncrotron galáctica e às das fontes puntiformes extragalácticas (DE MERICIA, Eduardo J. et al., 2023). Tais emissões

misturam-se ao sinal de 21 cm, prejudicando seu estudo. A separação de componentes nada mais é do que identificar cada um dos termos da Equação 4.40, de forma a subtraí-los do sinal total coletado pelo radiotelescópio e isolando o sinal de interesse (HI).

O problema de separação de componentes, ou de recuperação de um sinal desejado, pode ser modelado escrevendo os dados de observação numa dada frequência ν e pixel p como (DE MERICIA, Eduardo J. et al., 2023)

$$x_\nu(p) = b_\nu(p) * y_\nu(p) + n_\nu(p), \quad (4.41)$$

em que $x_\nu(p)$ representa o mapa do sinal observado, $b_\nu(p)$ é o modelo do feixe do instrumento, $y_\nu(p)$ é o mapa de emissão do céu, $n_\nu(p)$ é o mapa de ruído, e $*$ representa a operação de convolução. O mapa de emissão do céu é composto pela soma da contribuição de diferentes componentes e pode ser escrito como

$$y_\nu(p) = \sum_c a_{\nu c}(p) s_c(p), \quad (4.42)$$

em que $s_c(p)$ representa cada componente do sinal observado e $a_{\nu c}(p)$ é um coeficiente que relaciona as componentes à frequência ν . O somatório é feito sobre o total de componentes, c . A Equação 4.42 representa as mesmas componentes do céu apresentadas em 4.40, mas desta vez atribuindo um peso para cada componente dependendo da frequência ν .

Ao substituir a Equação 4.42 na Equação 4.41, obtém-se a expressão

$$x_\nu(p) = b_\nu(p) * \left(\sum_c a_{\nu c}(p) s_c(p) \right) + n_\nu(p) = b_\nu(p) * [\mathbf{A}s]_\nu(p) + n_\nu(p) \quad (4.43)$$

onde $\mathbf{A}_\nu(p)$ é a matriz de combinação cujos elementos são $a_{\nu c}(p)$. O processo de separação de componentes consiste em encontrar uma solução para a Equação 4.43, recuperando o sinal cosmológico de interesse, em uma determinada frequência ν (DE MERICIA, Eduardo J. et al., 2023).

Os métodos de separação de componentes podem ser classificados como paramétricos ou não-paramétricos. Os métodos paramétricos utilizam modelos estatísticos e

espectrais dos diversos contaminantes, o que auxilia na identificação dos sinais. Já os métodos não-paramétricos realizam a recuperação do sinal cosmológico usando apenas os dados observacionais (DE MERICIA, Eduardo J. et al., 2023). Neste trabalho foram testados dois métodos não-paramétricos, que serão descritos nas próximas subseções. Como o objetivo é analisar os efeitos dos diferentes métodos de *map-making* na mitigação do ruído e seu posterior efeito na recuperação do sinal, não será realizada uma análise aprofundada da separação de componentes. Pode-se encontrar um estudo focado nesta etapa em Merícia (2019), Remazeilles et al. (2011) e Olivari et al. (2016).

4.3.1 PCA

A Análise de Componentes Principais (*Principal Component Analysis* - PCA), é um método não-paramétrico que transforma os mapas espectrais em modos ortogonais independentes (BIGOT-SAZY et al., 2015). O PCA recebe como *input* os mapas de intensidade em diferentes frequências e com isso transforma o espectro do sinal observado em cada pixel em uma nova base de modos ortogonais, chamada de componente principal. Essa transformação é realizada através da decomposição em autovalores e autovetores da matriz de covariância entre os canais, baseando-se na covariância entre frequências e buscando encontrar padrões de variação comuns entre os pixels. Cada uma dessas componentes representa uma combinação linear dos mapas de entrada, ordenadas de acordo com a grandeza dos respectivos autovalores, permitindo separar os modos dominantes ou componentes principais, geralmente associados aos *foregrounds*, daqueles de menor variação, que é o caso do sinal de HI.

Os mapas para os diferentes canais espectrais apresentam alta correlação entre si devido à dominância dos *foregrounds*, que se estendem suavemente ao longo de todos os canais de frequência. Cada componente é descrita por um vetor que associa pesos às frequências espectrais. A ferramenta então realiza, para cada pixel, um produto escalar de cada componente com o espectro do sinal, obtendo quanto daquela componente está nele presente. Esta operação permite reconstruir o espectro em cada pixel e separar as principais componentes do resíduo, que em geral é dado pelo sinal cosmológico. O PCA identifica as componentes com maior variância, que em geral apresentam alta relação sinal/ruído e estão associadas aos *foregrounds* dominantes. Essas componentes podem então ser removidas, isolando o sinal cosmológico mais fraco.

4.3.2 GNILC

O segundo método testado para a separação de componentes foi o *Generalized Needlet Internal Linear Combination* (GNILC), um método não paramétrico desenvolvido por [Remazeilles et al. \(2011\)](#) para análise de dados da CMB e adaptado por [Olivari et al. \(2016\)](#) para a realização do mapeamento de intensidade do HI. O GNILC é um método mais complexo quando comparado ao PCA, pois realiza a decomposição do sinal no espaço harmônico, não se restringindo ao espaço de *pixels*, o que permite a realização de uma análise local e em diferentes escalas angulares, adaptando-se às variações dos próprios *foregrounds* astrofísicos.

O método resolve o problema apresentado na Equação 4.43 realizando, para cada região do céu e escala angular, uma decomposição espectral da matriz de covariância dos dados no espaço de pixels e no espaço harmônico. O GNILC define uma base vetorial em ambos os espaços e define assim os modos espectrais, análogos às componentes principais do PCA. Para cada região do céu e em cada escala angular, é estimada a matriz de covariância entre os canais de frequência observados, analisando como os sinais variam em conjunto entre diferentes frequências ([REMAZEILLES et al., 2011](#)). Em paralelo, o GNILC utiliza uma matriz de covariância do sinal cosmológico estimado, que pode ser obtida a partir de simulações ou modelos teóricos. A informação sobre o sinal cosmológico, neste caso o sinal de 21 cm do HI, é introduzida através do modelo teórico de emissão do HI contido no mapa denominado *prior*. A comparação entre essas duas matrizes permite avaliar, em cada região e escala, quais modos espectrais estão dominados pelos *foregrounds* e quais são compatíveis com o sinal cosmológico ([REMAZEILLES et al., 2011](#)). Diferentemente do PCA, que remove cegamente os modos de maior variância, o GNILC utiliza essa comparação para determinar de forma adaptativa e local o número de modos a serem considerados contaminados.

Uma vez identificado o subespaço espectral que corresponde ao sinal, ou seja, quais modos estão atrelados ao sinal, o GNILC aplica uma técnica chamada *Internal Linear Combination* (ILC) dentro desse subespaço. O ILC combina os mapas em diferentes frequências por meio de uma soma ponderada que minimiza a variância total, mas com a restrição de preservar o sinal cosmológico ([REMAZEILLES et al., 2011](#); [OLIVARI et al., 2016](#)). No GNILC, essa combinação é feita somente nos modos que não foram classificados como contaminados. O resultado deste processo é dado pelo mapa de HI reconstruído, com contaminação por *foregrounds* minimizada.

É importante notar que o GNILC não trata o ruído instrumental, mas apenas re-

move a contaminação dos *foregrounds* astrofísicos. A ferramenta praticamente não remove o ruído térmico pois o mesmo possui as mesmas propriedades espectrais do sinal de HI. Já a remoção do ruído $1/f$ com o GNILC depende do grau de correlação entre frequências do mesmo. Um ruído com alta correlação espectral é mais fácil de ser removida pelo método. Tal condição demonstra a importância de se tratar o ruído instrumental ainda na etapa de *map-making*, permitindo uma melhor recuperação do sinal cosmológico. Uma descrição detalhada do funcionamento do método e a fundamentação matemática adotada pelo GNILC no cenário da atuação do radiotelescópio BINGO foi realizada no trabalho de Merícia (2019) e DE MERICIA, Eduardo J. et al. (2023). Como a separação de componentes foi utilizada neste trabalho apenas como uma das ferramentas comparativas dos resultados obtidos para o *map-making*, decidiu-se por não explorá-la profundamente.

5 PREPARAÇÃO DOS MAPAS E FERRAMENTAS DE COMPARAÇÃO

Afim de garantir uma análise acurada dos mapas utilizados durante a realização do trabalho, garantindo que estes retornassem a melhor estimativa possível do mapa de HI teórico e para uma boa realização da etapa de separação de componentes, é necessário realizar uma série de preparações nos mapas de intensidade. Estas preparações ou etapas de pré-processamento incluem o ajuste do feixe do instrumento entre os canais de frequência, a remoção da contribuição dos elementos do sistema do radiotelescópio nas medidas dos mapas de intensidade e a aplicação de uma máscara apodizada para restringir a região de análise dos mapas à região de observação do instrumento. Além disso, após a aplicação das ferramentas e a recuperação do mapa de HI, é preciso definir os métodos de comparação que serão adotados para distinguir e avaliar os resultados obtidos. O capítulo a seguir descreve os parâmetros adotados para as simulações, as etapas de preparação dos mapas de intensidade e as ferramentas de comparação utilizadas para a análise dos resultados obtidos com ambas as ferramentas de *map-making*, buscando encontrar aquela que obteve melhores resultados para os mapas recuperados através das ferramentas de separação de componentes.

5.1 Parâmetros das simulações

Como o objetivo deste trabalho é testar e comparar diferentes métodos de *map-making* aplicados às simulações de observação do radiotelescópio BINGO, é necessário determinar parâmetros para as simulações que condizem com o esperado para a atuação do instrumento. Estes parâmetros incluem características observacionais e do arranjo óptico como o tempo de observação simulado, a região do céu observada, o tipo de observação adotado, o número de cornetas do instrumento e sua posição no plano focal, o modelo de feixe adotado, entre outros. Devem também ser definidos quais *foregrounds* astrofísicos e ruídos sistemáticos serão adicionados às simulações e quais suas características espectrais. Quanto mais realista for a modelagem do instrumento e da missão de observação, com a inclusão de efeitos sistemáticos e detalhes do sistema óptico e eletrônico, mais complexa e fidedigna se torna a simulação, e por consequência, mais difícil se torna isolar o sinal de 21 cm do HI.

A seguir estão organizados alguns dos parâmetros selecionados no desenvolvimento deste trabalho para as simulações de missão do BINGO, levando em consideração sua primeira fase de funcionamento (WUENSCHÉ et al., 2022; ABDALLA et al., 2022a). Alguns parâmetros foram mantidos fixos, enquanto outros, especialmente os

ligados a características espectrais dos ruídos, foram alterados de forma a visualizar seu efeito nos mapas produzidos e na recuperação do sinal. Na Seção 5.2 são apresentados os diversos arranjos de cornetas para o sistema óptico que foram testados, considerando a Fase 1 do instrumento, com 14 cornetas. Já na Seção 5.3 são apresentados os parâmetros espectrais e características adotadas para as simulações e para os mapas. Na Seção 5.4 são apresentadas as etapas de processamento e preparação dos mapas de intensidade, necessárias para a aplicação das ferramentas de separação de componentes. A Seção 5.5 trata de uma discussão acerca dos dois espaços de parâmetros em que se pode realizar o ajuste do feixe do instrumento. Por fim, na Seção 5.6 são apresentados os critérios e os métodos comparativos utilizados para avaliar o desempenho das ferramentas de *map-making* adotadas e seus resultados constam no Capítulo 6.

5.2 Arranjo do sistema óptico

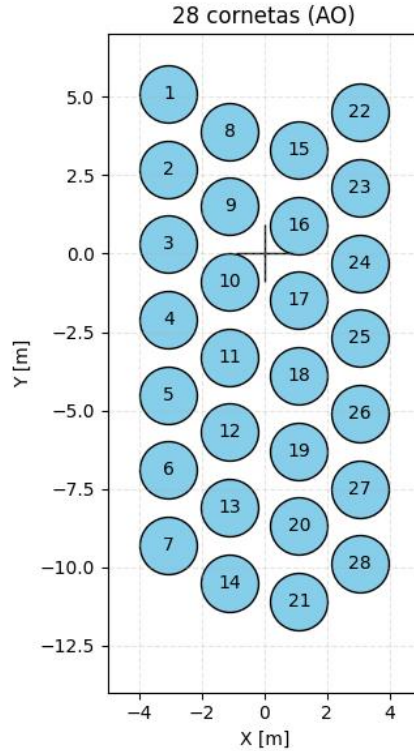
Foram mantidas fixas algumas características do sistema óptico do instrumento, como o tamanho dos espelhos que compõem sua estrutura, enquanto outras, como o número de cornetas e seu arranjo no plano focal, foram alteradas. O arranjo original (AO) de cornetas previsto para Fase 1 do BINGO possui 28 cornetas agrupadas de forma compacta em 4 colunas de 7 (WUENSCHÉ et al., 2022), conforme Figura 5.1.

Considerando o tempo necessário para a fabricação, transporte e montagem das cornetas no plano focal, avaliou-se a possibilidade de iniciar as observações com apenas 14 cornetas, dispostas segundo o arranjo original (AO) em duas colunas de sete elementos, conforme apresentado na Figura 5.2. Porém, devido ao espaçamento entre as cornetas e à resolução angular adotada, esse arranjo resultou em mapas com faixas sem cobertura, ou seja, com *pixels* não observados. Para contornar essa limitação, optou-se neste trabalho por um arranjo compacto, mantendo o mesmo número de cornetas, agora agrupadas na parte inferior (arranjo compacto inferior - ACI) do plano focal, conforme Figura 5.3.

5.3 Outros parâmetros usados nas simulações

Quanto aos demais parâmetros das simulações, foram variados a frequência de jorlo, f_k , e o índice espectral de correlação em frequência, β , de modo a observar o efeito de diferentes características espectrais do ruído $1/f$ na produção de mapas de intensidade do HI. Os valores de f_k foram variados entre 0, 1, 0,01 e 0,001 Hz. Os valores de β foram variados entre 0, 1, 0,25 e 1. No entanto, para a análise final dos resultados, foram consideradas apenas as simulações com $f_k = 0,001$ Hz e $\beta = 0,25$,

Figura 5.1: Arranjo original (AO) com 28 cornetas no plano focal para a Fase 1 do BINGO.

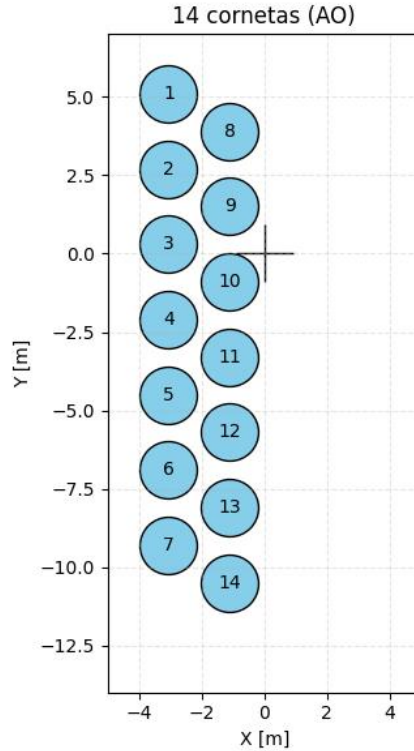


o que está melhor desenvolvido no Capítulo 6.

Conforme apresentado no Capítulo 3, o índice espectral das correlações em frequência, β , é definido dentro do intervalo $0 \leq \beta \leq 1$, sendo $\beta = 0$ para flutuações idênticas entre canais, ou seja, totalmente correlacionadas, e $\beta = 1$ para flutuações totalmente independentes. Como espera-se que o ruído $1/f$ seja altamente correlacionado, os valores de β estão mais próximos de zero (HARPER et al., 2018). Um β próximo de 1 representaria um ruído com características próximas do ruído branco, tornando sua remoção um processo mais penoso, dificultando assim a recuperação do sinal de interesse (HARPER et al., 2018). Os valores explorados para f_k , por sua vez, impactam nas escalas angulares onde há contaminação dos mapas. O impacto do ruído $1/f$ no espectro de potência angular recuperado é mais significativo quando a intensidade do ruído é comparável ou excede a potência angular do HI (HARPER et al., 2018). Além disso, quanto maior o valor de f_k , maiores serão as escalas em que o ruído dominará, contaminando ainda mais o mapa de intensidade.

Quanto à resolução do mapa do céu, foi adotado o valor de $N_{\text{side}} = 128$ e posteriormente, $N_{\text{side}} = 512$, de acordo com o modo de pixelização definido pelo HEALPix

Figura 5.2: Arranjo original (AO) com 14 cornetas no plano focal para a Fase 1 do BINGO.

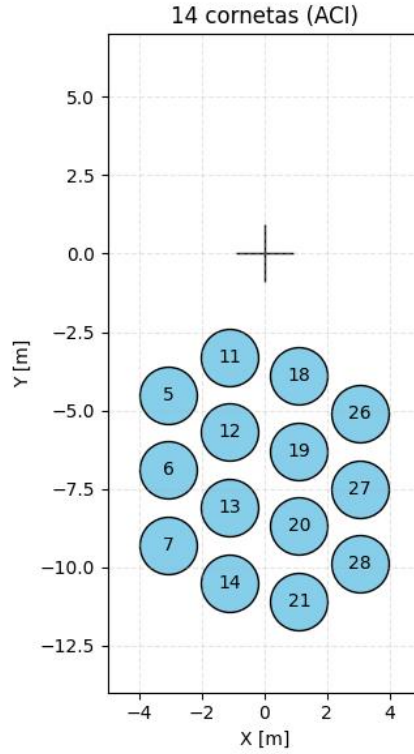


(*Hierarchical Equal Area isoLatitude Pixelisation*) (GORSKI et al., 2005), um método de pixelização globalmente adotado em cosmologia. Para sua escolha, foram consideradas a resolução angular que espera-se que o instrumento seja capaz de obter e os resultados preliminares dos mapas produzidos. Quanto maior o valor do N_{side} , maior é a resolução angular, e maior a capacidade de poder se distinguir entre elementos pequenos em um mapa do céu.

O tempo de observação foi definido para que cada *pixel* pudesse ser observado um número suficiente de vezes, sem que as simulações tornassem-se excessivamente grandes, do ponto de vista computacional. O valor inicial adotado foi de 3 dias de observação, considerando a discussão realizada por Doré et al. (2001). Demais parâmetros, como largura de banda, número de canais e a largura de cada canal, foram selecionados considerando a banda de operação do instrumento (980 - 1260 MHz) e testes preliminares realizados em trabalhos da colaboração (DE MERICIA, Eduardo J. et al., 2023; LICCARDO et al., 2022).

A Tabela 5.1 sintetiza os parâmetros espectrais e instrumentais adotados para as simulações de missão do BINGO utilizadas na realização deste trabalho. Os *fore-*

Figura 5.3: Arranjo compacto inferior (ACI) com 14 cornetas no plano focal para a Fase 1 do BINGO.



foregrounds astrofísicos não foram considerados nas simulações deste trabalho, pois o objetivo do mesmo é analisar os efeitos do ruído $1/f$ na produção de mapas de HI a partir de observações simuladas do BINGO. Apesar de estas não serem simulações realistas, elas são suficientes para testar os métodos aplicados de *map-making* na mitigação do $1/f$, já que o ruído correlacionado surge de componentes instrumentais, de forma que seu efeito é observado independentemente dos sinais astrofísicos considerados. Os efeitos dos *foregrounds* astrofísicos na recuperação do sinal de 21 cm do HI a partir de observações simuladas do BINGO foram explorados nos trabalhos Merícia (2019), DE MERICIA, Eduardo J. et al. (2023) e Liccardo et al. (2022).

Para realizar a separação de componentes, as ferramentas empregadas baseiam-se no comportamento espectral do sinal e de seus contaminantes. De forma simplificada, componentes com espectro suave e covariantes em frequência são atribuídas aos *foregrounds*, enquanto aquelas com alta variabilidade espectral e ausência de covariância são associadas ao sinal cosmológico. Como o ruído térmico resulta de flutuações aleatórias, seu espectro não é suave e se assemelha ao do sinal de HI. Por esse motivo, os métodos de separação de componentes não conseguem distinguir entre o ruído térmico e o sinal cosmológico, resultando em mapas recuperados que

Tabela 5.1: Parâmetros adotados para as simulações de missão do radiotelescópio BINGO.

Descrição	Parâmetro	Valor adotado
Temperatura de sistema	T_{sys}	70 K
Frequência de joelho	f_k	0.1; 0.01 e 0.001 Hz
Parâmetro de correlação em frequência	β	0.1; 0.25 e 1
Resolução do mapa	N_{side}	128 e 512
Número de cornetas	N_{horns}	14
Tempo de observação	t_{obs}	3 dias
Frequência de amostragem	f_{sr}	1 Hz
Largura de banda	$\Delta\nu$	280 MHz
Número de canais	N_ν	30
Largura de canal	$\delta\nu$	9.33 MHz

contêm ambos de forma indistinta. Isso comprometeria a avaliação da eficácia da reconstrução do sinal de HI. Por essa razão, o ruído térmico não foi incluído nas simulações apresentadas neste trabalho. Mais detalhes sobre essa escolha podem ser encontrados na Seção 4.3. Embora esteja fora do escopo deste estudo, a separação entre ruído térmico e sinal de HI pode ser realizada posteriormente, na etapa de construção dos espectros de potência angular, por meio da técnica de *debiasing*, conforme discutido em (DE MERICIA, Eduardo J. et al., 2023).

5.4 Preparação dos mapas

Antes de aplicar os métodos de separação de componentes, é necessário preparar os mapas de entrada, de forma a colocá-los em condições adequadas para análise. Essa preparação inclui a subtração do nível médio do sinal observado, a correção da resolução angular dos mapas com base na largura do feixe em cada canal de frequência (*rebeaming*) e a aplicação de uma máscara apodizada, que visa minimizar os efeitos de borda no espectro de potência angular, pois a transição descontínua entre as regiões observada e não-observada introduz alterações e “buracos” na forma do espectro de potência.

Inicialmente, para a análise dos mapas produzidos a partir das TODs, faz-se necessária a subtração do valor médio de intensidade associado aos componentes eletrônicos, conhecido como nível DC (*Direct Current*). Em sistemas eletrônicos, o nível DC refere-se a uma componente constante de tensão ou corrente, presente independentemente das variações introduzidas por sinais reais. No contexto de radiotelescópios, essa contribuição vem dos próprios elementos eletrônicos do instrumento e pode ser

estimada somando-se as contribuições individuais de cada componente, previamente caracterizadas em laboratório. No caso do BINGO, essa contribuição é de cerca de 70 K. Como esse valor não corresponde a um sinal astrofísico, mas sim a uma característica instrumental, ele deve ser removido dos dados. Essa remoção foi feita de duas maneiras distintas, conforme a ferramenta de *map-making* utilizada. No caso do ROMA, que utiliza o método iGLS, a subtração foi feita diretamente nas TODs, antes de serem processadas, o que é considerado o procedimento mais adequado, pois a presença do nível DC pode introduzir um viés na estimação estatística dos mapas. Isso ocorre pois métodos como o iGLS assumem ruído com média igual a zero ao longo do tempo. Assim, a presença de um nível constante pode ser confundida com o sinal real, comprometendo a reconstrução do mapa. Já na IM *pipeline*, que aplica o método *naive*, a subtração do nível DC foi feita *a posteriori*, diretamente nos mapas produzidos, por meio da remoção do valor médio de intensidade de cada *pixel*. Essa abordagem foi adotada devido à configuração atual da *pipeline*, que não permite executar separadamente as etapas de geração das TODs e de construção de mapas.

Após a subtração do nível DC, os mapas passaram pelo processo de *rebeamng*, ou ajuste de feixe, necessário para uniformizar a resolução angular entre os canais de frequência. No mapeamento de intensidade, o sinal de HI é observado em diversos canais, definidos pela largura de banda total ($\Delta\nu$) e pela largura individual de cada canal ($\delta\nu$). De acordo com a equação da difração de Rayleigh,

$$\theta_{\text{FWHM}}(\nu_{\text{obs}}) \approx \frac{c}{\nu_{\text{obs}}D}, \quad (5.1)$$

em que $\theta_{\text{FWHM}}(\nu_{\text{obs}})$ é a largura do feixe a meia altura (*Full Width at Half Maximum* - FWHM) na frequência observada ν_{obs} , c é a velocidade da luz e D é o diâmetro do refletor do instrumento, observa-se que a resolução angular é inversamente proporcional à frequência. Como o diâmetro D é fixo, frequências mais baixas (associadas a *redshifts* mais altos) correspondem a feixes mais largos e, portanto, a mapas com menor resolução.

Na Equação 5.1, omite-se o fator 1,22 presente na fórmula da difração para uma abertura circular (padrão de Airy), pois assume-se neste trabalho que o perfil do feixe é aproximadamente gaussiano. Nesse caso, a largura do feixe pode ser estimada por uma expressão proporcional a λ/D , com constante da ordem da unidade. Essa aproximação é geralmente empregada em experimentos de mapeamento de in-

Tabela 5.2: Tamanho angular do feixe do BINGO para cada canal de frequências.

Canal	f_c (MHz)	FWHM (arcmin)	Canal	f_c (MHz)	FWHM (arcmin)
0	984.6	45.5	15	1124.7	39.8
1	994.0	45.1	16	1134.0	39.5
2	1003.3	44.7	17	1143.3	39.2
3	1012.7	44.2	18	1152.7	38.9
4	1022.0	43.8	19	1162.0	38.6
5	1031.3	43.4	20	1171.3	38.2
6	1040.7	43.0	21	1180.7	37.9
7	1050.0	42.7	22	1190.0	37.6
8	1059.3	42.3	23	1199.3	37.4
9	1068.7	41.9	24	1208.7	37.1
10	1078.0	41.6	25	1218.0	36.8
11	1087.3	41.2	26	1227.3	36.5
12	1096.7	40.9	27	1236.7	36.2
13	1106.0	40.5	28	1246.0	36.0
14	1115.3	40.2	29	1255.3	35.7

tensidade, nos quais o feixe do instrumento é bem descrito por um perfil gaussiano simétrico.

Como a banda de frequências do BINGO, entre 980 MHz e 1260 MHz, esta dividida em 30 canais obtêm-se 30 valores diferentes para o tamanho angular do feixe do instrumento ao aplicar-se a Equação 5.1. Nas simulações realizadas, os canais estão ordenados por frequência crescente, de modo que canais com menor índice possuem feixes maiores. O tamanho angular do feixe associado a cada canal de frequência esta apresentado na Tabela 5.2, em que f_c representa a frequência central do canal. Essa variação de resolução dificulta a aplicação de métodos de separação de componentes, que assumem que todos os mapas possuem a mesma resolução angular.

Para contornar essa limitação, realiza-se o *rebeamng*, que consiste em remover o efeito do feixe original de cada mapa e aplicar um novo feixe comum a todos, igual (ou um pouco maior) ao maior entre os feixes originais — garantindo assim que todos os mapas passem a ter a mesma resolução, correspondente à resolução mínima entre os canais. Essa operação pode ser realizada tanto no espaço de *pixels* quanto no espaço dos harmônicos esféricos, conforme detalhado na Seção 5.5.

Por fim, aplica-se aos mapas uma máscara apodizada, com o objetivo de suavizar a transição entre a região observada e a não observada. Mapas que cobrem todo o céu são chamados de mapas *full sky*, ou mapas de céu inteiro, e são produzidos por

experimentos capazes de observar a esfera celeste em sua totalidade, como o Planck (TAUBER et al., 2004), WMAP (BENNET et al., 2003) e COBE (MATHER et al., 1991). Porém, instrumentos em solo observam apenas uma fração do céu, limitados por sua localização geográfica e por restrições instrumentais. O BINGO é um desses casos: suas observações cobrem aproximadamente $14,75^\circ$ em declinação e $6,0^\circ$ em ascensão reta, totalizando uma área de ~ 5400 graus quadrados ao longo de um dia de observação com 28 cornetas, durante a Fase 1 de operação (WUENSCHKE et al., 2022).

Como consequência, os mapas produzidos contêm regiões com dados observacionais e regiões sem informação. Para lidar com isso, é comum utilizar uma máscara binária, composta por valores 1 para os *pixels* observados, e 0 para os não observados. Essa máscara é aplicada por meio de uma multiplicação *pixel a pixel* com o mapa original, garantindo que análises seguintes considerem apenas os *pixels* com informação. Essa etapa é importante porque os *pixels* não observados, se não forem mascarados, podem ser erroneamente interpretados como regiões com valores nulos associados a dados reais, embora não contenham nenhuma informação verdadeira.

Como as operações envolvendo os mapas geralmente são realizadas em toda a sua extensão, a presença de transições abruptas entre regiões observadas e não observadas pode introduzir efeitos de borda e *mode mixing* (vazamento de potência de modos de ℓ baixos para modos de ℓ altos) no cálculo do espectro de potência angular. Para mitigar esses efeitos, é comum usar máscaras apodizadas, que suavizam a transição entre essas regiões, fazendo com que a intensidade do sinal decresça gradualmente em direção às bordas da área observada.

Neste trabalho foi usada uma máscara do tipo C2, gerada com o pacote NaMaster (ALONSO et al., 2019), com um grau de apodização de 2,5 graus. Esse tipo de apodização garante que tanto a máscara quanto a sua derivada sejam contínuas, reduzindo o vazamento espectral entre diferentes modos harmônicos.

A apodização é implementada como um filtro multiplicativo aplicado ao mapa e, no caso do tipo C2, é definida em função da distância angular θ até a borda da região observada por

$$M(\theta) = \begin{cases} 1, & \theta \geq \theta_{\text{apo}} \\ \left[1 - \left(1 - \cos\left(\pi \frac{\theta}{\theta_{\text{apo}}}\right)\right)^2\right], & 0 < \theta < \theta_{\text{apo}} \\ 0, & \text{fora da região observada,} \end{cases} \quad (5.2)$$

onde $\theta_{\text{apo}} = 2.5^\circ$ é o raio angular de apodização. Essa função garante uma transição suave e contínua entre os valores 1 (regiões totalmente observadas) e 0 (regiões não observadas), diminuindo gradualmente o peso atribuído aos *pixels* próximos da borda. Como consequência, a aplicação da máscara resulta em uma atenuação da potência nas bordas dos mapas, o que pode visualmente se manifestar como uma leve degradação da resolução angular nessas regiões. Ainda assim, esse procedimento é crucial para garantir melhores estimativas do espectro de potência angular, minimizando as contaminações associadas à borda da área observada.

O resultado final, após os processos de remoção do nível DC, ajuste do feixe entre canais e aplicação da máscara apodizada, é dado pelo mapa de HI mais ruído $1/f$ para a área observada e com feixe comum para todos os canais de frequência. Desta forma, os mapas estão preparados para serem adicionados como *inputs* nas ferramentas de separação de componentes. O Capítulo 6 é dedicado a apresentar os mapas produzidos com os parâmetros descritos na Tabela 5.1, que foram adequados segundo os processos descritos acima e submetidos a comparações estatísticas quantitativas, apropriadas aos métodos utilizados.

5.5 Ajuste de feixe

Conforme discutido anteriormente, o ajuste de feixe (*rebeamng*) é uma etapa crucial da análise, pois garante que todos os mapas possuam a mesma resolução angular em todos os canais de frequência. Esse procedimento pode ser realizado em diferentes domínios, cada um com suas vantagens e limitações. A seguir, é apresentada uma discussão sobre a realização do ajuste de feixe tanto no espaço harmônico (espaço dos multipolos ℓ) quanto no espaço de *pixels*.

5.5.1 Ajuste de feixe no espaço harmônico

Ao realizar uma observação do céu, cada instrumento apresenta um feixe característico, que descreve sua resposta angular às estruturas do céu. Estruturas com escalas angulares comparáveis ou maiores que o tamanho do feixe podem ser resolvidas espacialmente, enquanto aquelas com escalas menores tendem a ser suavizadas, o que resulta em perda de informação. Assim, para o estudo de estruturas finas (variações espaciais do sinal em escalas angulares pequenas), é desejável que o feixe do instrumento tenha sempre uma resolução menor ou igual ao tamanho da estrutura que se deseja observar. No entanto, no contexto do mapeamento de intensidade, em que o objetivo não é resolver individualmente as fontes, mas sim mapear a distribuição estatística do sinal ao longo do céu, não é necessário que o feixe seja estreito

(pequeno).

No caso de instrumentos que tem por objetivo medir a escala de BAO, como o BINGO, o tamanho do feixe não pode ser muito maior do que a escala angular característica das BAO. Essa escala varia inversamente com o *redshift*, sendo menor conforme o *redshift* aumenta. Assim, os feixes de instrumentos voltados para o mapeamento de intensidade variam entre 30 e 60 minutos de arco, dependendo do *redshift* observado.

Observa-se, portanto, que a resposta do feixe depende das escalas angulares envolvidas, sendo mais significativa em pequenas escalas. Para quantificar o efeito do feixe em cada escala, define-se o perfil do feixe no espaço harmônico, B_ℓ , o qual descreve como o feixe atenua as contribuições do mapa em função do multipolo ℓ . Assim, B_ℓ associa a cada escala angular um fator de supressão correspondente à resposta do instrumento.

Ao realizar o estudo das emissões do céu, é comum que se realize a decomposição do campo de temperatura na esfera celeste através de harmônicos esféricos. Nesta decomposição, cada harmônico é representado pelo termo $Y_{\ell m}$ e está relacionado a um coeficiente de expansão harmônica, $a_{\ell m}$, que representa a amplitude das oscilações de temperatura. O campo de temperatura, nesta abordagem, é dado por

$$T(\theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\theta, \phi) \quad (5.3)$$

em que ℓ é o número de multipolo e representa a escala angular das flutuações de temperatura, m é o número azimutal, que varia de $-\ell$ até ℓ , e está relacionado à simetria da flutuação ao redor do eixo polar, com $\ell \approx 1/\theta$. θ e ϕ representam o ângulo polar e azimutal, respectivamente, sendo utilizados para determinar a posição no céu. Tem-se assim, uma forma de representar as variações do campo de temperatura do céu ao longo das diversas escalas angulares.

A capacidade do instrumento de detectar variações angulares do sinal depende da sua resolução angular, caracterizada pelo perfil do feixe no espaço harmônico B_ℓ . Para um feixe com formato gaussiano simétrico, esse perfil é dado por

$$B_\ell = \exp\left(-\frac{\ell(\ell+1)\sigma^2}{2}\right), \quad (5.4)$$

onde σ é a largura do feixe em radianos e está relacionada à largura a meia altura (*full width at half maximum* - FWHM), θ_{FWHM} , do mesmo por

$$\sigma = \frac{\theta_{\text{FWHM}}}{\sqrt{8\ln 2}}. \quad (5.5)$$

A função B_ℓ descreve como o feixe do instrumento atua como um filtro de convolução sobre os coeficientes harmônicos do sinal observado, suavizando principalmente as estruturas em pequenas escalas angulares (altos valores de ℓ). No espaço harmônico, essa suavização é expressa pela relação

$$a_{\ell m}^{\text{obs}} = a_{\ell m}^{\text{true}} \cdot B_\ell, \quad (5.6)$$

onde $a_{\ell m}^{\text{obs}}$ são os coeficientes harmônicos do mapa observado em um dado canal de frequência, e $a_{\ell m}^{\text{true}}$ representam os coeficientes do sinal verdadeiro do céu.

Como discutido na Seção 5.4, cada canal de frequência possui uma resolução angular diferente, determinada pelo respectivo valor de θ_{FWHM}^i . Essa variação ocorre devido às propriedades do instrumento, uma vez que o feixe depende da frequência observada. Consequentemente, os mapas produzidos em diferentes canais são suavizados por diferentes perfis de feixe, representados por B_ℓ^i , onde o índice i indica o canal correspondente.

Uma das técnicas utilizadas para garantir que os mapas de todos os canais possuam a mesma resolução angular é conhecida como *rebeamng* no espaço harmônico. Essa técnica consiste em substituir o perfil de feixe original de cada canal B_ℓ^i , por um feixe comum, B_ℓ^{comm} , associado ao canal com o feixe mais largo (menor resolução angular). No caso deste trabalho, B_ℓ^{comm} corresponde ao perfil do canal de menor frequência (canal 0), conforme apresentado na Equação 5.1.

O procedimento de *rebeamng* é realizado aplicando-se a seguinte transformação aos coeficientes harmônicos observados:

$$a_{\ell m}^{\text{rebeam},i} = a_{\ell m}^{\text{obs},i} \cdot \frac{B_\ell^{\text{comm}}}{B_\ell^i}, \quad (5.7)$$

onde $a_{\ell m}^{\text{obs},i}$ são os coeficientes harmônicos do mapa observado no canal i ; $a_{\ell m}^{\text{rebeam},i}$ são os coeficientes após o *rebeamng*; B_ℓ^i é o perfil do feixe original do canal i ; B_ℓ^{comm}

é o perfil do feixe mais largo, comum a todos os canais após o *rebeaming*. Essa operação remove o efeito do feixe original B_ℓ^i e aplica o efeito do feixe mais largo, B_ℓ^{comm} , resultando em mapas com resolução angular homogênea em todos os canais. O *rebeaming* é realizado separadamente para cada um dos canais de frequência simulados.

5.5.2 Ajuste de feixe no espaço de pixels

O ajuste no espaço de *pixels*, por sua vez, é realizado diretamente sobre os valores de intensidade atribuídos a cada *pixel* do mapa. Para uniformizar a resolução angular entre diferentes canais, aplica-se um processo de suavização aos mapas com maior resolução (isto é, aqueles com feixes mais estreitos). Essa suavização é feita por meio da convolução com um *kernel* gaussiano, que é uma matriz de dimensão reduzida representando uma função gaussiana centrada no *pixel* de interesse. O *kernel* possui valor máximo no centro e decai suavemente para os vizinhos, de modo a redistribuir a intensidade de cada *pixel* para os arredores, simulando o efeito de um feixe mais largo. Assim, o mapa é reprocessado para possuir a resolução angular correspondente ao canal com o feixe mais largo.

Seja $T_i(p)$ o mapa do canal i , cuja largura de feixe original θ_{orig}^i e θ_{comm} a largura de feixe comum desejada. Define-se então a largura efetiva da suavização a ser aplicada como

$$\theta_{\text{FWHM}}^{\text{smooth},i} = \sqrt{(\theta_{\text{FWHM}}^{\text{comm}})^2 - (\theta_{\text{FWHM}}^{\text{orig},i})^2}, \quad (5.8)$$

desde que $\theta_{\text{FWHM}}^{\text{comm}} > \theta_{\text{FWHM}}^{\text{orig},i}$. A suavização é realizada através da convolução com um *kernel* gaussiano discreto em HEALPix dado por

$$K_i(p, p') = \frac{1}{2\pi\sigma_i^2} \exp\left(-\frac{\theta_{pp'}^2}{2\sigma_i^2}\right), \quad (5.9)$$

onde $K_i(p, p')$ é o valor do *kernel* que determina o quanto o *pixel* p' contribui para o *pixel* p ; $\theta_{pp'}$ é a distância angular entre os centros dos *pixels* p e p' , calculada pelo produto escalar entre os vetores unitários associados a esses *pixels*; σ_i é a largura do feixe (desvio padrão), relacionada à $\theta_{\text{FWHM}}^{\text{smooth},i}$ por (conforme Equação 5.5)

$$\sigma_i = \frac{\theta_{\text{FWHM}}^{\text{smooth},i}}{\sqrt{8 \ln 2}}. \quad (5.10)$$

Assim, a convolução no espaço de *pixels* é dada por

$$T_i^{\text{rebeamed}}(p) = \sum_{p'} K(p, p') T_i(p'), \quad (5.11)$$

onde $T_i(p')$ é a temperatura no *pixel* p' do mapa associado ao canal de frequência i e a soma é calculada considerando apenas os vizinhos próximos de p (dentro de um raio máximo pré-definido).

5.6 Ferramentas de comparação

A avaliação comparativa dos mapas gerados pela IM *pipeline* e pelo ROMA requer a aplicação de diferentes métodos de análise, que possibilitam investigar as propriedades estatísticas e estruturais dos mapas produzidos. Esses métodos fornecem informações complementares sobre a qualidade dos mapas, a presença de contaminantes e a eficiência das técnicas empregadas para mitigar ruídos. Afim de poder comparar os mapas gerados pelas ferramentas de *map-making*, foram analisados alguns aspectos e informações extraídas dos *outputs* destes programas, como a análise dos histogramas dos mapas, suas médias e desvio padrão; os espectros de potência angular dos mapas e como estes se comportam, comparando-os com o espectro de potência teórico do HI e, por fim, ambos os mapas foram processados pelo GNILC (REMAZEILLES et al., 2011) e pelo PCA para definir qual dos algoritmos retorna um melhor resultado para a recuperação do mapa de HI. Realizada a separação de componentes, para quantificar a recuperação do sinal foram analisados o coeficiente de correlação de Pearson, aplicado aos mapas e o erro de reconstrução do espectro de potência angular, C_ℓ .

Cada elemento comparativo é descrito nas seções seguintes. Seus resultados serão apresentados no Capítulo 6

5.6.1 Mapas de intensidade

Os mapas de intensidade representam os produtos primários da realização do *map-making* e contêm a informação contida nas TODs, agora de forma compacta e em 2D, na forma de mapa do céu. Eles mostram a região do céu observada e as intensidades

associadas ao sinal, o que permite identificar a presença de estruturas no mapa e observar a potência do sinal. No entanto, de forma isolada, os mapas não são capazes de fornecer informações suficientes para definir a melhor ferramenta de *map-making*. Assim, foram utilizados outros parâmetros comparativos e ferramentas estatísticas, que são explorados adiante.

No contexto deste trabalho, e considerando os parâmetros apresentados na Tabela 5.1, foram gerados, para cada simulação e para cada método de *map-making*, 30 mapas correspondentes aos diferentes canais de frequência, totalizando 60 mapas por simulação. Cada mapa contém a combinação do sinal de 21 cm do HI com ruído $1/f$. Além disso, foram salvos separadamente os mapas com apenas a contribuição do sinal de HI, de forma a servir como referência para a avaliação da recuperação do sinal. Como *priors* para o método GNILC, também foram geradas 30 realizações independentes dos mapas de HI, utilizadas como *input* para o processo de separação de componentes.

Para a realização das demais análises, e especialmente da separação de componentes, estes mapas tiveram que ser tratados e preparados, conforme descrito na Seção 5.4. Os mapas obtidos após tratamento, para três dos 30 canais de frequência, estão apresentados na Seção 6.3. Foram escolhidos os canais 0, 14 e 29 para evitar redundância nos resultados.

5.6.2 Ferramentas estatísticas gerais

O cálculo dos histogramas, da média e do desvio padrão são ferramentas de análise estatística que podem ser aplicadas aos resultados obtidos. Os histogramas dos mapas ilustram a distribuição estatística dos valores de intensidade em cada pixel, facilitando a identificação da dispersão da temperatura de brilho e de possíveis tendências sistemáticas. Caso os histogramas das duas metodologias apresentem diferenças significativas, isso pode indicar que um dos métodos introduz um viés ou que existem diferenças na remoção de contaminantes. Ademais, a média e o desvio padrão dos mapas fornecem uma perspectiva quantitativa sobre a variabilidade do sinal. A média permite uma estimativa do nível de brilho médio dos mapas, enquanto o desvio padrão expõe a amplitude das flutuações presentes. Uma discrepância substancial entre os desvios padrão dos mapas pode sugerir que um dos métodos demonstra maior sensibilidade ao ruído instrumental ou às variações na emissão do sinal.

5.6.3 Espectro de potência angular

Quanto à análise do espectro de potência e a aplicação de técnicas de separação de componentes, para este trabalho, estes representam ferramentas comparativas mais específicas para o estudo de dados cosmológicos. O espectro de potência angular proporciona uma análise quantitativa e estatística da distribuição da potência do sinal em escalas angulares distintas.

A partir da decomposição do campo de temperatura em harmônicos esféricos, conforme apresentado na Equação 5.3, é possível definir o espectro de potência angular, C_ℓ , que quantifica a distribuição da potência do sinal em diferentes escalas angulares (DODELSON; SCHMIDT, 2020). Esse espectro fornece uma medida estatística da variância associada a cada modo ℓ e pode ser calculado através da expressão

$$C_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2. \quad (5.12)$$

O espectro de potência angular, C_ℓ , é calculado a partir dos coeficientes $a_{\ell m}$ obtidos por meio da transformada de Fourier esférica do campo de temperatura do mapa $T(\theta, \phi)$. Esses coeficientes representam a contribuição de cada modo harmônico para o mapa. Uma vez determinados os $a_{\ell m}$, o espectro C_ℓ é obtido conforme a Equação 5.12. De forma análoga, é possível produzir um mapa de temperatura a partir do espectro C_ℓ teórico, por meio da geração de coeficientes $a_{\ell m}$ com estatísticas adequadas. O espectro de potência permite, portanto, avaliar como as estruturas do mapa variam com a escala angular, além de fornecer uma base quantitativa para a comparação entre mapas reconstruídos e a previsão teórica do sinal de 21 cm do HI.

Na realização deste trabalho foram adotados os pseudo- C_ℓ (\tilde{C}_ℓ) ao invés do C_ℓ real. Isto se deve ao fato de que o C_ℓ é o espectro de potência angular real, calculado a partir de um mapa de céu inteiro. Já o \tilde{C}_ℓ é o espectro estimado a partir de um mapa parcial, como os analisados na realização deste trabalho. O \tilde{C}_ℓ é calculado como uma combinação linear do C_ℓ e de uma matriz de pesos, dado por:

$$\langle \tilde{C}_\ell \rangle = \sum_{\ell'} M_{\ell\ell'} C_{\ell'}. \quad (5.13)$$

Em que ℓ' representam os modos reais e ℓ os modos ligados ao \tilde{C}_ℓ . Cada modo ℓ do pseudo-espectro não representa isoladamente o modo ℓ' do espectro real, mas está

contaminado por modos vizinhos, devido à perda de ortogonalidade imposta pela máscara. Ou seja, os modos ℓ deixam de ser independentes e passam a misturar-se.

Para obter-se uma estimativa do C_ℓ a partir do \tilde{C}_ℓ é necessário realizar uma operação de deconvolução da matriz de pesos com o pseudo- C_ℓ . Isto é feito invertendo a relação da Equação 5.13, obtendo-se

$$C_\ell = \sum_{\ell'} (M^{-1})_{\ell\ell'} \tilde{C}_{\ell'}, \quad (5.14)$$

de forma que pode-se obter o C_ℓ real dos mapas a partir do pseudo- C_ℓ .

Ao calcular os espectros de potência dos mapas, verifica-se a precisão dos métodos em reproduzir as características estatísticas previstas para o sinal de HI. Foram comparados neste trabalho os espectros de potência teóricos do HI com os espectros de potência dos mapas recuperados pelas ferramentas de separação de componentes. A proximidade entre os espectros teóricos e reconstruídos indica uma boa recuperação do sinal. Discrepâncias significativas em uma das escalas angulares podem sinalizar problemas relacionados à filtragem do ruído ou à remoção de contaminantes. A presença de excesso de potência em certas escalas pode indicar a presença de sistemáticos não corrigidos por um dos métodos.

5.6.4 Dispersão e coeficiente de Pearson

No contexto deste trabalho, o coeficiente de Pearson (ρ) é uma medida estatística que quantifica a correlação linear entre o mapa de HI reconstruído pelo GNILC (ou outro método de separação de componentes) e o mapa de HI de entrada. Seu valor varia entre -1 e 1 : valores próximos de 1 indicam correlação linear positiva forte, ou seja, o mapa reconstruído se aproxima muito bem do mapa real; enquanto valores próximos de -1 indicam correlação linear negativa forte, isto é, o mapa reconstruído tende a ser inversamente correlacionado com o mapa real; e valores próximos de 0 indicam ausência de correlação linear significativa entre os mapas. Em resumo, o coeficiente de Pearson fornece uma medida da qualidade da reconstrução do sinal de HI em termos de correspondência espacial. O coeficiente de correlação de Pearson em um dado canal de frequência ν é definido como

$$\rho_\nu = \frac{\sum_p (\hat{s}_\nu(p) - \mu_{\hat{s}_\nu})(s_\nu(p) - \mu_{s_\nu})}{\sqrt{\sum_p (\hat{s}_\nu(p) - \mu_{\hat{s}_\nu})^2 \sum_p (s_\nu(p) - \mu_{s_\nu})^2}}, \quad (5.15)$$

em que $\hat{s}_\nu(p)$ e $s_\nu(p)$ representam, respectivamente, os valores dos mapas reconstruído e real de HI, em um dado pixel p e canal de frequência ν . Os termos $\mu_{\hat{s}_\nu}$ e μ_{s_ν} correspondem às médias dos *pixels* dos mapas reconstruído e do mapa de HI real (DE MERICIA, Eduardo J. et al., 2023).

Além do valor da correlação de Pearson, é possível obter uma representação visual da correlação entre os mapas real e reconstruído por meio de um gráfico de dispersão. Nesse gráfico, os valores dos *pixels* do mapa recuperado são plotados contra os valores correspondentes do mapa real. Quando há forte correlação, os pontos tendem a se concentrar ao longo da diagonal, indicando semelhança entre os mapas. Já uma dispersão maior dos pontos em relação à diagonal mostra que há fraca correlação, ou seja, que há uma recuperação ruim do sinal.

5.6.5 Erro de reconstrução do espectro

Para medir o erro de reconstrução do espectro de potência angular dos mapas de HI, conforme apresentado na Seção 5.6.3, foi adotado o parâmetro $\eta(\nu)$, calculado individualmente para cada canal de frequência. Esse parâmetro expressa o erro percentual médio absoluto entre o espectro de potência do mapa reconstruído, C_ℓ^{rec} , e do mapa real de HI, C_ℓ^{real} , normalizado por C_ℓ^{real} (DE MERICIA, Eduardo J. et al., 2023). Sua definição é dada por

$$\eta(\nu) = \frac{1}{N_\ell} \sum_{\ell=\ell_{\min}}^{\ell_{\max}} \left| \frac{C_\ell^{\text{rec}}(\nu) - C_\ell^{\text{real}}(\nu)}{C_\ell^{\text{real}}(\nu)} \right| \times 100\%, \quad (5.16)$$

onde N_ℓ representa o número total de multipolos considerados. O intervalo de multipolos adotado, $30 \leq \ell \leq 270$, foi definido com base nas limitações associadas à cobertura do céu observado e pela resolução angular do instrumento, de forma a garantir que as análises fossem feitas em escalas angulares corretas.

Cada método de comparação empregado fornece uma perspectiva distinta sobre a qualidade dos mapas recuperados. A análise estatística inicial por meio dos histogramas, médias e desvios padrão permite uma inspeção geral dos resultados obtidos com os diferentes métodos. O espectro de potência angular oferece um diagnóstico da consistência dos mapas em diferentes escalas angulares. Finalmente, a aplicação do GNILC e do PCA testa a robustez dos mapas no cenário da separação de componentes. A combinação dessas análises possibilita uma avaliação abrangente da eficácia dos métodos empregados na produção de mapas. Dessa forma, torna-se

possível identificar qual método de *map-making* oferece maior fidelidade na reconstrução do sinal de HI, contribuindo para a melhoria das técnicas de análise de dados em experimentos de mapeamento de intensidade.

O Capítulo 4 apresentou as ferramentas e os métodos utilizados para simular as observações do instrumento BINGO e produzir os mapas de intensidade do sinal de 21 cm do HI. Enquanto no Capítulo 5 foram discutidas as etapas de preparação destes mapas e os parâmetros adotados para a comparação dos mapas recuperados com ambos os métodos de *map-making*. O Capítulo 6, a seguir, apresenta os resultados obtidos após as etapas descritas anteriormente, sendo estes as TODs simuladas, os mapas de intensidade obtidos, os mapas de HI recuperados após a aplicação da separação de componentes, seus espectros de potência angular e os coeficientes de correlação de Pearson, bem como o erro de reconstrução do C_ℓ recuperado.

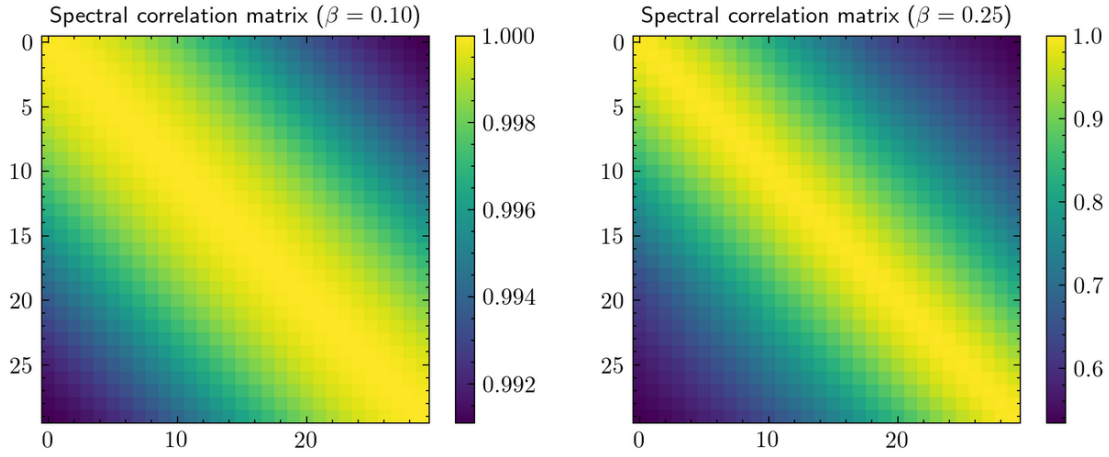
6 RESULTADOS

Diversos parâmetros (relativos ao ruído $1/f$, configuração do receptor, estratégia de observação, etc.) influenciam diretamente as características dos mapas de intensidade gerados a partir das observações simuladas, bem como a qualidade da reconstrução do sinal de HI feita a partir dos mesmos. Considerando os parâmetros descritos na Tabela 5.1, foram realizadas diversas simulações neste trabalho. Inicialmente, as diferenças entre as simulações foram definidas pela escolha da frequência de Joelho (f_k) e pelo parâmetro de correlação em frequência (β). Porém, devido a limitações impostas pela estratégia de observação do instrumento e pelas características espectrais do ruído $1/f$, as simulações com $f_k = 0,1$ Hz e $f_k = 0,01$ Hz foram desconsideradas, pois não foi possível recuperar o sinal de HI a partir dessas configurações. Assim, o conjunto de simulações foi reduzido a uma única configuração inicial, à qual foram aplicadas diferentes técnicas de processamento.

A justificativa para a redução do número de simulações se deu pelo fato de que, conforme apresentado por Harper et al. (2018), à medida que o valor de f_k aumenta, a relação sinal-ruído dos mapas reconstruídos de HI diminui. Ou seja, espera-se melhores resultados para pequenos valores de f_k . Além disso, conforme o valor de f_k aumenta, o ruído $1/f$ passa a dominar em escalas de tempo menores e também em escalas angulares menores. Este fato introduz mais ruído em pequenas escalas angulares e dificulta ainda mais a separação de componentes em altos valores de multipolos. No trabalho de Bigot-Sazy et al. (2015), realizaram-se simulações que foram submetidas à separação de componentes para valores de f_k que variaram de 1 mHz à 10 Hz, com um modelo de ruído assumindo uma correlação perfeita entre canais ($\beta = 0$). Neste caso, não houve diferenças significativas na recuperação do sinal com a variação de f_k . Observa-se, portanto, que o parâmetro mais significativo na efetividade da recuperação do sinal é a correlação espectral. Devido à configuração dos receptores do BINGO, é esperado que a frequência de Joelho em operação seja $f_k \sim 1$ mHz (WUENSCHÉ et al., 2022). Assim, optou-se por realizar as simulações com o valor fixo de $f_k = 1$ mHz, um cenário próximo da frequência-alvo.

Quanto ao valor do parâmetro de correlação, β , foram testados três valores distintos: 0,1; 0,25 e 1. Pode-se visualizar os efeitos do parâmetro de correlação no perfil do ruído, além dos seus efeitos no sinal, através da matriz de correlação. Ela mostra o nível de correlação entre os canais de frequência do sinal medido. Para o caso do HI puro, espera-se que a matriz de correlação apresente um comportamento totalmente diagonal, simbolizando uma forte autocorrelação entre canais e uma fraca correlação

Figura 6.1: Matrizes de correlação das simulações com presença do sinal de HI e do ruído $1/f$.



À esquerda consta a matriz para $\beta = 0,1$ e à direita para $\beta = 0,25$. Os eixos são dados pelos 30 canais espectrais, entre 980 e 1260 MHz. A região amarelada indica alta correlação e a azulada, baixo nível de correlação.

com canais distintos. Para o sinal de HI contaminado com ruído $1/f$, espera-se um comportamento diagonal com uma queda suave na forma de uma curva $1/x$ para os elementos fora da diagonal.

O perfil de correlação do ruído pode ser observado na Figura 6.1, na qual são apresentadas as matrizes de correlação para as simulações com $\beta = 0,1$ e $\beta = 0,25$. Os eixos representam os canais espectrais, enquanto a barra de cores indica o nível de correlação, sendo amarelo quando alto e azulado quando baixo.

Observa-se que para o gráfico da esquerda ($\beta = 0,1$), a faixa amarelada é mais larga, o que demonstra a alta correlação do sinal entre canais distantes. Os altos limites da barra de cores também destacam a forte correlação. A queda branda da diagonal principal, em direção às bordas da imagem, evidencia o perfil espectral suave do ruído quando o valor de beta é baixo o suficiente. Enquanto isso, o gráfico da direita ($\beta = 0,25$) possui uma faixa amarelada mais estreita, além do intervalo dos limites da barra de cores ser mais amplo. Portanto, há uma menor suavidade espectral entre os canais neste caso, uma vez que a correlação cai mais rapidamente.

Para $\beta = 1$, caso de total ausência de correlação espectral no ruído $1/f$, a expressão 3.5 é reduzida, o que torna a determinação da IPSD mais simples (GASPERIS et al., 2005). No entanto, como os canais são totalmente independentes entre si, o ruído

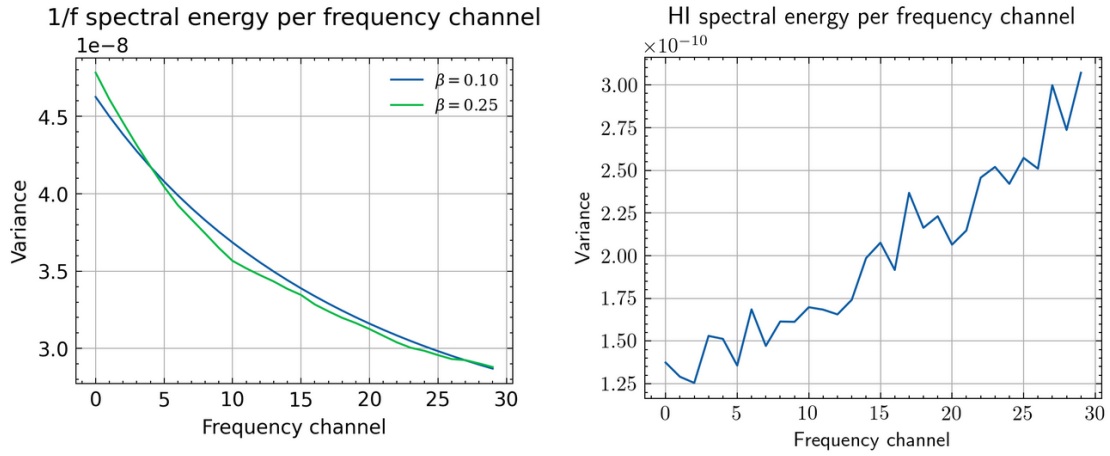
$1/f$ torna-se muito parecido com o ruído branco e são necessários muitos modos para descrevê-lo, o que dificulta a ação dos métodos de separação de componentes (HARPER et al., 2018). Assim, como esperado, as simulações com $\beta = 1$ resultaram em mapas de observação a partir dos quais não foi possível recuperar o sinal de HI. Nesse caso específico, o GNILC utiliza todos os modos, ou graus de liberdade, para caracterizar o ruído $1/f$. Isto faz com que o mapa recuperado seja nulo, não sobrando componente associada ao sinal de HI. Desta forma, os resultados para tais simulações foram desconsiderados por não retornarem mapas válidos para o estudo.

De forma semelhante, o GNILC e o PCA foram testados em um cenário onde os mapas simulados de observação contêm apenas as componentes de HI e ruído $1/f$ com $\beta = 0,1$. Teoricamente, esperava-se que ambos os métodos de separação de componentes apresentassem bom desempenho nessa configuração, uma vez que o ruído $1/f$ com β baixo possui alta correlação espectral, ou seja, varia suavemente ao longo dos canais de frequência, ao contrário do sinal de HI, que apresenta baixa correlação entre canais de frequência. Essa diferença entre as estruturas espectrais das duas componentes deveria, em princípio, permitir uma separação de componentes eficaz por meio da análise das matrizes de covariância locais, nas quais o ruído, por ser suave, poderia ser descrito por poucos modos principais, enquanto o sinal de HI seria descrito pelos demais modos residuais, geralmente a maioria, conforme o número de canais de frequência disponíveis.

No entanto, os resultados obtidos contrariaram as expectativas. Os mapas de HI reconstruídos com o GNILC foram nulos, indicando que o método não conseguiu separar o sinal de HI do ruído. Esse resultado sugere que, apesar de as duas componentes serem estatisticamente separáveis, o ruído $1/f$ com $\beta = 0,1$ pode ainda dominar as variações espectrais em regiões de baixa razão sinal-ruído, mascarando o sinal de HI mesmo quando sua estrutura espectral se destaca. Outra hipótese é que o GNILC, ao estimar o número de modos dominados pelo ruído com base na decomposição das matrizes de covariância, tenha identificado todos os modos relevantes como contaminados, suprimindo, assim, também o sinal de HI.

A Figura 6.2 reforça essa análise ao mostrar a comparação entre a estrutura espectral do ruído $1/f$, para diferentes valores de β , e a do sinal de HI. É possível notar uma diferença de cerca de 2 ordens de grandeza nas escalas de variação espectral entre o ruído e o sinal. Apesar dessa diferença evidenciar que as duas componentes são teoricamente separáveis, ela também mostra que, para $\beta = 0,1$, o ruído ainda apresenta variações de amplitude suficientemente significativas em relação ao sinal

Figura 6.2: Comparação da estrutura espectral do sinal de HI (direita) com a do ruído $1/f$ para $\beta = 0,1$ e $\beta = 0,25$ (esquerda).



Observa-se que o ruído apresenta uma variação espectral muito mais suave, principalmente para $\beta = 0,1$ (curva azul), caracterizando forte correlação entre canais. O sinal de HI, por outro lado, exibe flutuações mais proeminentes e menos correlacionadas entre canais de frequência. A diferença de escala entre os espectros é de cerca de duas ordens de grandeza.

de HI, principalmente nas escalas onde a potência do HI é naturalmente baixa. Esse comportamento pode levar o GNILC a interpretar toda a variância do mapa como ruído, resultando em mapas reconstruídos nulos.

Por fim, entre todas as configurações de parâmetros testadas, os únicos mapas de observação a partir dos quais foi possível recuperar o sinal de HI por meio dos métodos de separação de componentes empregados neste trabalho foram aqueles gerados com $\beta = 0,25$. Nesse caso, o espectro do ruído era suficientemente suave para que os métodos conseguissem distingui-lo do sinal de HI como uma componente independente. Ao mesmo tempo, a correlação espectral do ruído ainda era alta o bastante para que sua contribuição pudesse ser modelada com um número reduzido de modos espectrais. A combinação desses fatores favoreceu a eficácia da separação estatística usada pelos métodos, permitindo a recuperação, mesmo que parcial, do sinal de HI.

Assim, todos os resultados da separação de componentes apresentados neste trabalho referem-se a simulações com ruído gerado com frequência de joelho $f_k = 0,001$ Hz e índice espectral $\beta = 0,25$. A Tabela 6.1 organiza e sumariza as simulações realizadas e a nomenclatura adotada. Conforme detalhado no Capítulo 5, foram exploradas diferentes resoluções angulares dos mapas e variações nos domínios em que foi rea-

lizado o ajuste do feixe.

Tabela 6.1: Sumário das simulações realizadas e seus parâmetros.

Simulação	t_{obs} (dias)	N_{side}	Domínio
Simulação 1	3	128	Pixels
Simulação 2	3	512	Pixels
Simulação 3	3	128	Harmônico
Simulação 4	3	512	Harmônico

Definidos os parâmetros do ruído, é feita a análise dos resultados obtidos a partir da aplicação de duas abordagens distintas de *map-making*. Essas ferramentas são comparadas quanto à sua eficácia na remoção do ruído $1/f$ presente nas simulações de missão, com base nos métodos descritos no Capítulo 5. Inicialmente, são apresentadas as séries temporais geradas, seguidas pelos mapas construídos a partir das mesmas, de acordo com cada método de processamento e ajuste aplicado.

Na etapa de separação de componentes, foram avaliados os desempenhos do GNILC e do PCA, com o objetivo de identificar qual método se mostrou mais eficaz na recuperação do sinal de HI nas condições simuladas. A partir da seleção da técnica de separação que apresentou melhores resultados, foi feita a comparação entre os métodos de *map-making*, utilizando como critério os mapas reconstruídos e os respectivos erros de recuperação do sinal de HI.

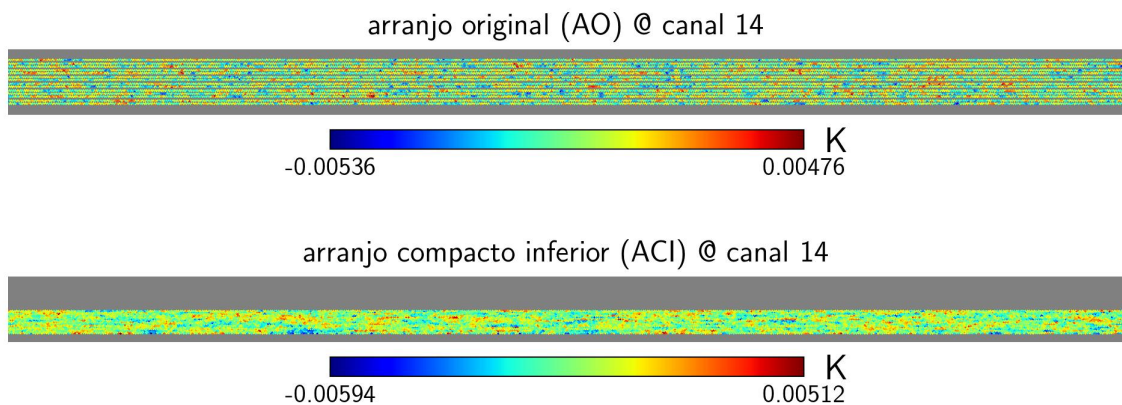
Os resultados são apresentados nas seções seguintes, organizadas de acordo com as ferramentas de processamento de dados adotadas e as etapas envolvidas na preparação dos mapas.

6.1 Arranjo das cornetas e resolução espacial

Os primeiros testes com as simulações foram realizados com o intuito de observar o efeito do arranjo de cornetas no plano focal do instrumento e da resolução espacial adotada nos mapas simulados. Inicialmente, observou-se que para o arranjo original (AO) de cornetas do BINGO e para um $N_{side} = 128$, o que equivale a uma resolução espacial de $\sim 27,5$ minutos de arco, os mapas apresentaram faixas não observadas em isolatitude, resultando em *gaps* na cobertura do céu reconstruído. Essa limitação pode ser vista tanto no mapa superior da Figura 6.3 (projeção cartesiana) quanto no mapa da esquerda da Figura 6.4 (projeção gnomônica), que representam a mesma região celeste.

Esses *gaps* nos mapas das observações simuladas podem comprometer as etapas subsequentes da análise de dados, como a separação de componentes, a estimativa do espectro de potência angular e a própria reconstrução do sinal de HI. Embora um arranjo compacto das cornetas cubra uma área menor do céu, ele apresenta uma cobertura uniforme, gerando mapas contínuos e sem *gaps*. Isso o torna mais adequado para análises que exigem homogeneidade na área celeste coberta, como as abordagens utilizadas neste trabalho.

Figura 6.3: Mapas de temperatura simulados no canal 14 para dois arranjos de cornetas: arranjo original (AO; mapa superior) e arranjo compacto inferior (ACI; mapa inferior).

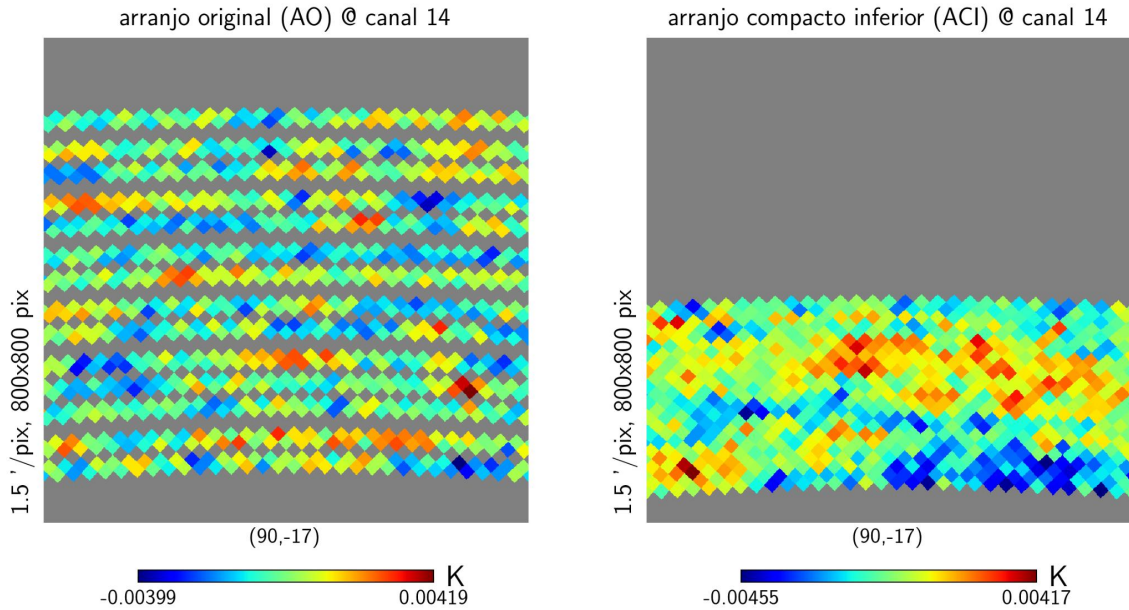


Os mapas são apresentados em projeção cartesiana, com $N_{\text{side}} = 128$, e foram gerados a partir de 3 dias de observação simulada contendo o sinal de HI e ruído $1/f$ com $\beta = 0,25$, utilizando o método *naive*. As faixas não observadas de declinação constante são consequência do AO de cornetas do BINGO.

Uma possível solução para mitigar a presença das falhas seria diminuir a resolução dos mapas, pois, com a diminuição do N_{side} , o tamanho de cada pixel aumenta e passa a cobrir uma região maior do céu. No entanto, essa abordagem implicaria perda da informação em pequenas escalas. Assim, optou-se por manter a resolução original e estudar novas configurações para o arranjo de cornetas.

Foi testada uma configuração alternativa ao arranjo original, denominado Arranjo Compacto Inferior (ACI), constituído pelas cornetas da região inferior do plano focal do BINGO, conforme ilustrado na Figura 5.3. A configuração compacta permite posicionar os eixos centrais das cornetas mais próximos entre si, possibilitando uma cobertura do céu contínua, sem áreas “descobertas”, como aparece no mapa superior

Figura 6.4: Mapas de temperatura simulados no canal 14 para dois arranjos de cornetas: arranjo original (AO; mapa da esquerda) e arranjo compacto inferior (ACI; mapa da direita).



Os mapas são apresentados em projeção gnomônica, com $N_{\text{side}} = 128$, e foram gerados a partir de 3 dias de observação simulada contendo o sinal de HI e ruído $1/f$ com $\beta = 0,25$, utilizando o método *naive*. As faixas não observadas de declinação constante (em isolatitude) são consequência do AO de cornetas.

da Figura 6.3.

Os mapas resultantes do ACI não apresentaram regiões não observadas em declinação constante. Para esse arranjo, utilizando a resolução de $N_{\text{side}} = 128$, foram obtidos mapas contínuos, embora com uma faixa de observação mais estreita. Esse resultado pode ser visto na Figura 6.3, em projeção cartesiana, e em visão ampliada na Figura 6.4, em projeção gnomônica. Assim, a partir deste ponto, o Arranjo Compacto Inferior será denominado simplesmente Arranjo Compacto (AC) e será utilizado na realização de todas as simulações seguintes.

Com base nos resultados anteriores sobre a escolha do arranjo ideal de cornetas, as simulações de missão descritas na Tabela 6.1 foram realizadas com o AC. As simulações foram geradas com a IM *pipeline*, cujos *outputs* consistem nas TODs das observações, que são posteriormente convertidas em mapas por meio das ferramentas de *map-making*. A seção seguinte discute as TODs geradas e analisa o efeito do coeficiente de correlação espectral no comportamento das mesmas ao longo dos

diferentes canais de frequência.

6.2 Séries temporais

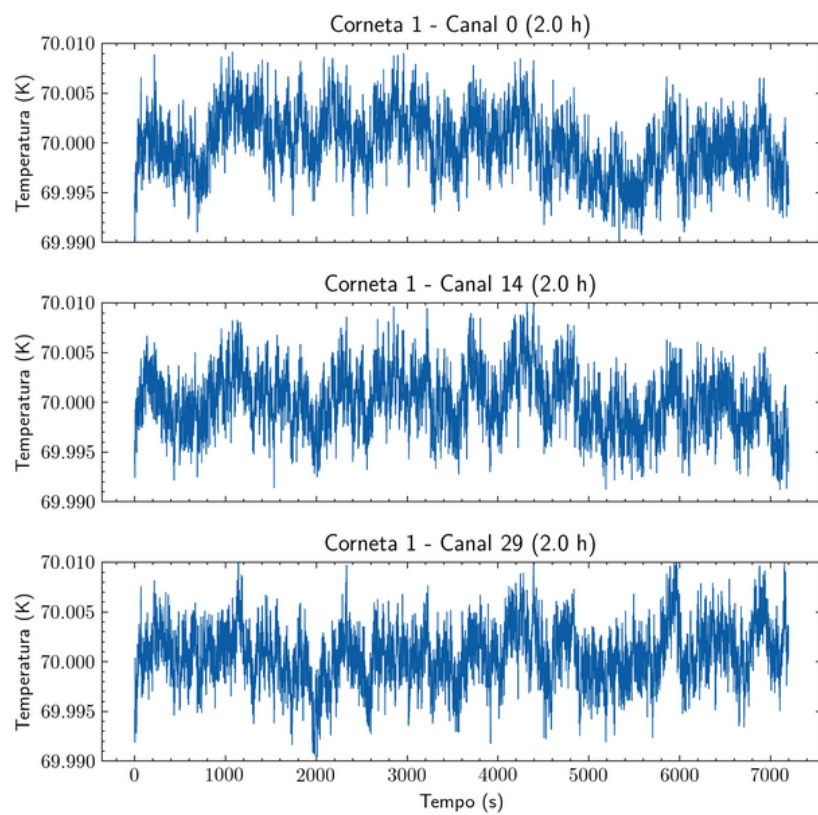
As simulações de missão do BINGO foram geradas com a IM *pipeline*, uma ferramenta que produz séries temporais de observação a partir de um conjunto de parâmetros de simulação definidos conforme as características do instrumento, a estratégia de observação, as propriedades dos ruídos instrumentais, os modelos de *foregrounds* astrofísicos considerados e a cosmologia adotada. As TODs constituem o principal *output* da *pipeline* e servem como *input* para as ferramentas de *map-making*, que geram os mapas de observação. As TODs contêm tanto o sinal de HI simulado, quanto as contribuições de ruído, permitindo visualizar diretamente as estruturas típicas do ruído $1/f$.

A Figura 6.5 apresenta as séries temporais para a Simulação 1, com três dias de observação, $\beta = 0,25$ e $N_{\text{side}} = 128$. As TODs são apresentadas para três canais de frequência e em um intervalo de duas horas, compreendendo apenas uma parte da simulação total. A presença do ruído $1/f$ pode ser observada pelas suas impressões características no formato das TODs, que aparecem na forma das oscilações periódicas de longa duração em torno da temperatura média de 70 K. O comportamento oscilatório depende da frequência de joelho do sinal, de forma que, quando o valor de f_k aumenta, a potência do ruído $1/f$ em escalas temporais menores aumenta.

Quanto ao parâmetro de correlação β , valores próximos de zero indicam que as TODs são muito semelhantes entre os canais (isto é, apresentam alta correlação em frequência), sendo idênticas quando $\beta = 0$. À medida que β aumenta e se aproxima de 1, as TODs tornam-se totalmente independentes (isto é, não correlacionadas). A Figura 6.5 apresenta as TODs referentes à Simulação 1, com $\beta = 0,25$, um valor considerado baixo e, portanto, indicativo de alta correlação. De acordo com Harper et al. (2018), valores de $\beta < 0,5$ são classificados como baixos, enquanto valores de $\beta > 0,5$ são considerados altos. Observa-se, neste caso, que as TODs apresentam grande semelhança entre os canais, tendo o mesmo padrão de oscilações em escalas temporais mais longas, com pequenas variações apenas nas escalas menores.

As TODs apresentadas na Figura 6.5 são utilizadas para gerar mapas de intensidade através do processo de *map-making*, conforme apresentado na Equação 4.3. Esse processo consiste em combinar os dados das séries com a matriz de apontamento do instrumento, utilizando as ferramentas e métodos estudados neste trabalho. Cada ferramenta apresenta um funcionamento distinto, descrito nas Seções 4.1.2 e 4.1.1,

Figura 6.5: Séries temporais produzidas com a IM *pipeline* para os canais de frequência 0, 14 e 29 da corneta 1, com $f_k = 0,001$ Hz e $\beta = 0,25$ em um intervalo de duas horas da observação total.



A forma semelhante entre as TODs de diferentes canais é um reflexo da alta correlação em frequência do sinal.

exigindo diferentes etapas de pré-processamento para gerar e analisar os mapas de saída, conforme detalhado no Capítulo 5. As seções seguintes apresentam os resultados de cada uma dessas etapas, às quais as séries temporais e os mapas foram submetidos, com o objetivo de viabilizar a aplicação dos métodos de separação de componentes e a análise *a posteriori* de seus resultados.

6.3 Pré-processamento dos mapas de intensidade

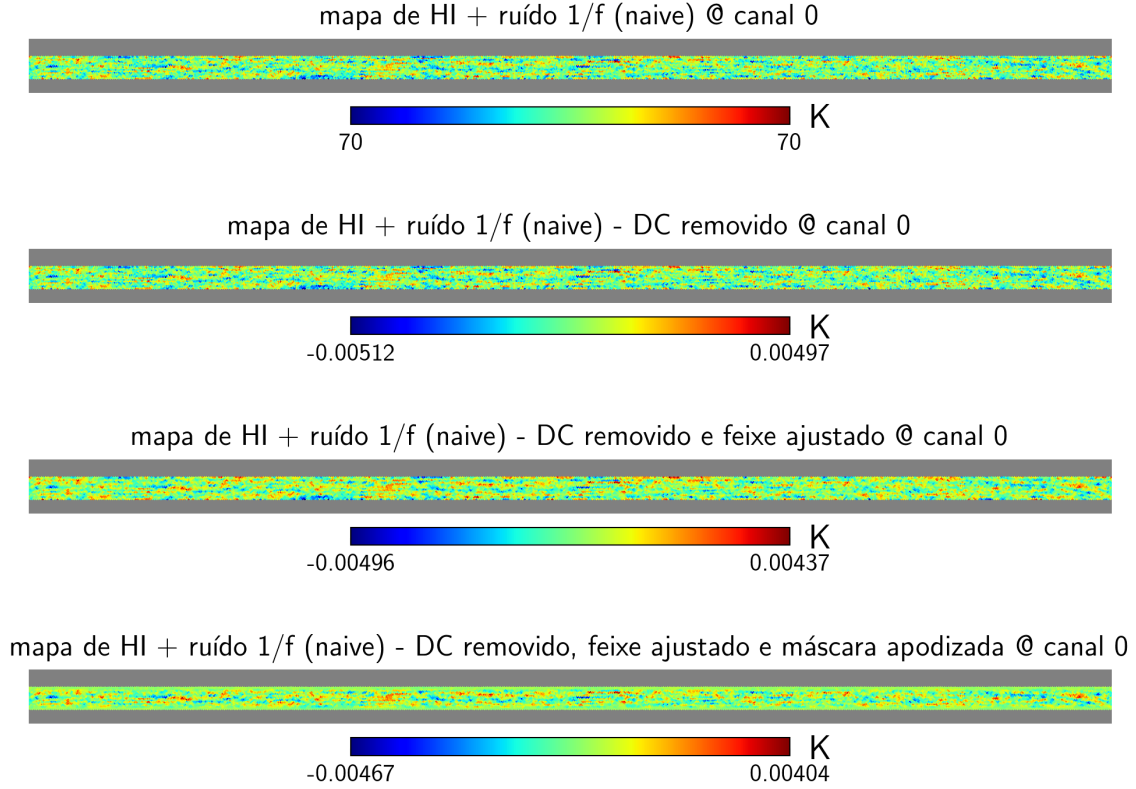
Esta seção apresenta o resultado de cada etapa do pré-processamento dos mapas produzidos com os métodos *naive* e iGLS. Por conveniência, são exibidos apenas os mapas da Simulação 1, obtidos pelo método *naive* e produzidos com a IM *pipeline*. Essa escolha deve-se ao fato de que as etapas são comuns a todas as simulações e a ambas as ferramentas de *map-making*, com exceção da subtração do nível DC, realizada internamente pelo ROMA. Os resultados após a separação de componentes, objeto de interesse final deste trabalho, são apresentados para as duas ferramentas na Seção 6.5. Na sequência, a análise do erro de recuperação do sinal permite comparar a eficácia de sua aplicação aos mapas produzidos por cada ferramenta.

Os mapas de intensidade devem ser gerados de forma a preservar o sinal de HI e permitir a extração posterior de informações cosmológicas úteis. Para alcançar esse objetivo, é fundamental tratar os mapas de modo que a separação de componentes e a recuperação do sinal de HI se tornem viáveis. As etapas desse tratamento, bem como seus objetivos e formas de aplicação, são detalhadas no Capítulo 5. A seguir, são apresentados os mapas resultantes dessas etapas de pré-processamento.

As Figuras 6.6, 6.7 e 6.8 apresentam os mapas *naive* obtidos para três canais (0, 14 e 29). Esses canais foram escolhidos por corresponderem, respectivamente, ao primeiro, ao centro aproximado e ao último canal de frequência da banda adotada pelo BINGO. Em cada figura, são mostrados quatro mapas que ilustram o resultado da aplicação sequencial das etapas de pré-processamento, realizadas sempre na ordem apresentada: inicialmente o mapa *naive* bruto, sem nenhum tratamento; em seguida, o mapa após remoção do nível DC; depois, o mapa resultante do ajuste de feixe; e, por fim, o mapa com aplicação da máscara apodizada.

O primeiro mapa não passou por nenhum tratamento, sendo um *output* direto da IM *pipeline* produzido pelo método *naive*, ou seja, o mapa bruto. A presença do nível DC é evidenciada pelos limites da barra de cores, fixos em 70 K. Após a remoção do nível DC, as variações de intensidade presentes no mapa podem ser melhor observadas nos limites da barra de cores, o que está apresentado no segundo mapa. A terceira

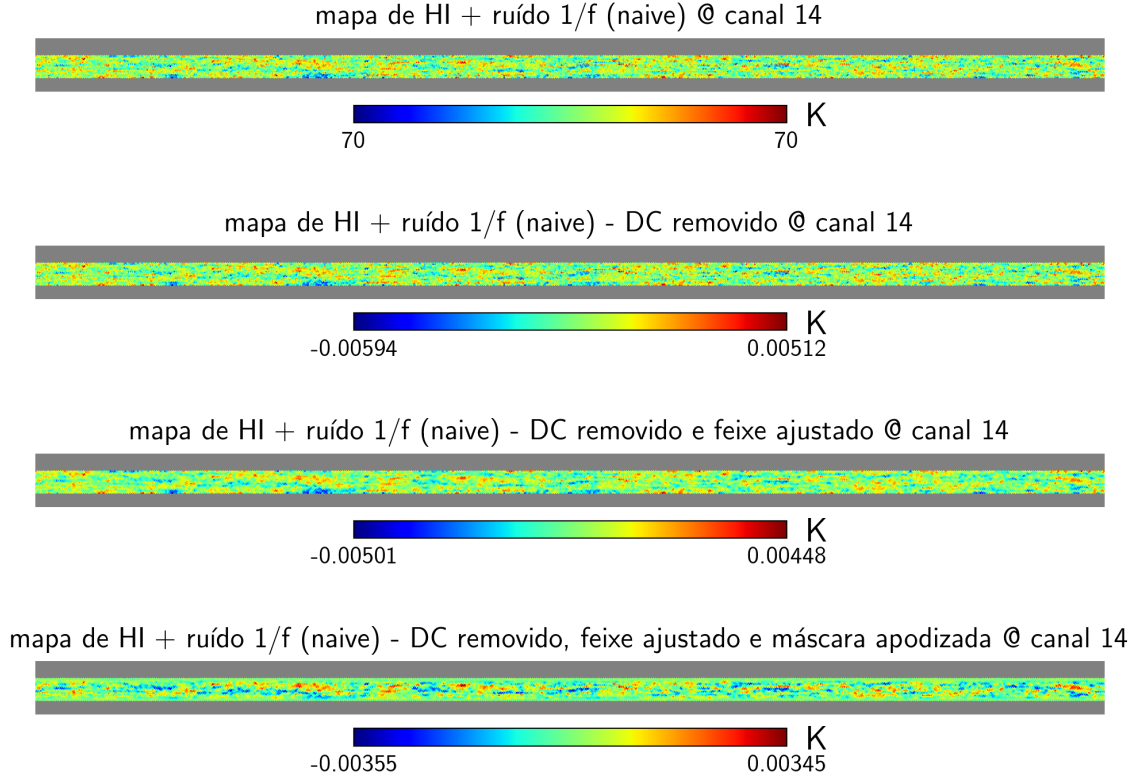
Figura 6.6: Mapas *naive* da Simulação 1 para o canal 0, após diferentes etapas de processamento (de cima para baixo): (i) mapa original bruto; (ii) com nível DC removido; (iii) com nível DC removido e feixe ajustado (no domínio de *pixels*); (iv) com nível DC removido, feixe ajustado (no domínio de *pixels*) e máscara apodizada.



imagem, por sua vez, apresenta os mapas *naive* após a realização do ajuste de feixe no espaço de pixels, no qual observa-se a perda de potência principalmente nos canais de frequência mais alta, conforme evidenciado nas Figuras 6.7 e 6.8. Estes, originalmente, possuem as menores larguras de feixe associadas, portanto, são os mais afetados pelo efeito do *rebeamng*. Por fim, é possível observar o efeito da aplicação da máscara no último mapa. Observa-se a perda de potência geral, evidenciada pelos limites da barra de cores e pela perda de informação das estruturas nas bordas superior e inferior, agora com menores flutuações em torno de zero. A máscara age suavizando tais regiões, para permitir uma transição suave até a região não observada do céu.

Os histogramas dos mapas de cada etapa do pré-processamento foram normalizados usando o *z-score*, definido por

Figura 6.7: Mapas *naive* da Simulação 1 para o canal 14, após diferentes etapas de processamento (de cima para baixo): (i) mapa original bruto; (ii) com nível DC removido; (iii) com nível DC removido e feixe ajustado (no domínio de *pixels*); (iv) com nível DC removido, feixe ajustado (no domínio de *pixels*) e máscara apodizada.

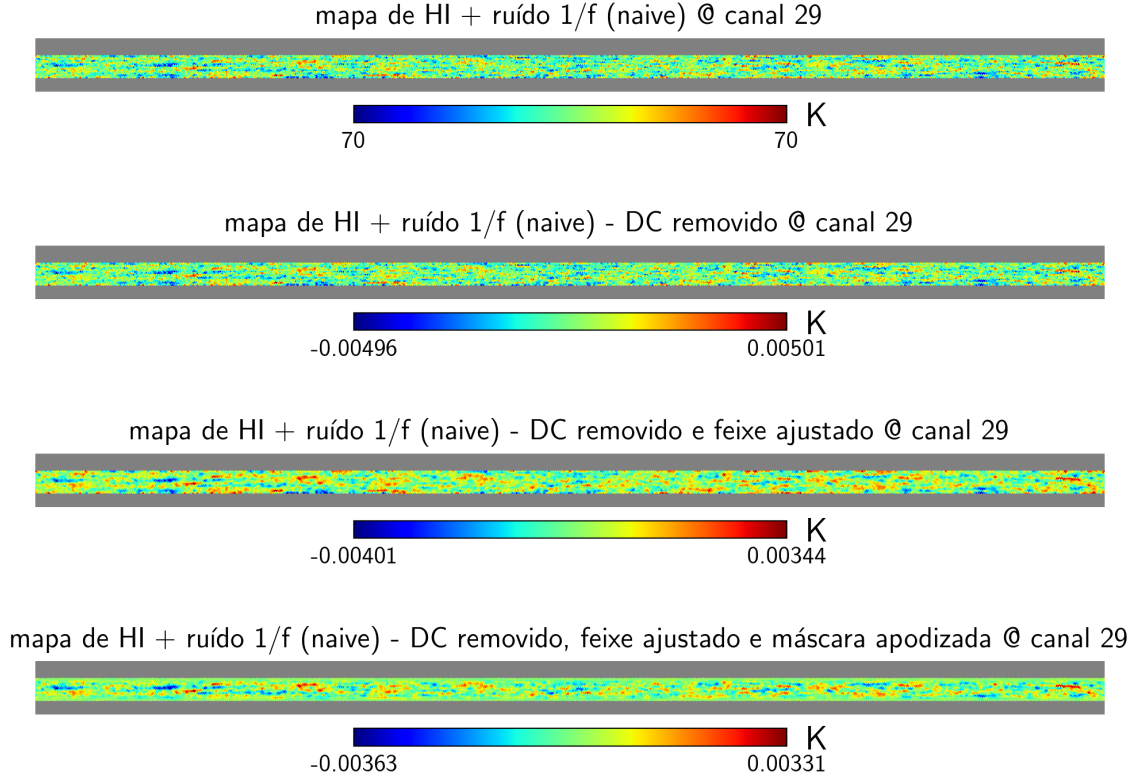


$$x_i^{\text{norm}} = \frac{x_i - \mu}{\sigma}, \quad (6.1)$$

onde x_i é o valor do *pixel*, μ a média e σ o desvio padrão de todos os *pixels* válidos do mapa. Valores de *z-score* positivos indicam *pixels* com valores acima da média, valores negativos indicam *pixels* abaixo, e zero corresponde à média. Essa normalização é adequada para dados com distribuição aproximadamente normal, como nas simulações deste trabalho, permitindo comparação direta entre mapas em diferentes escalas.

A Figura 6.9 apresenta os histogramas normalizados (*z-score*) dos mapas simulados do BINGO (sinal de HI + ruído 1/f, com 12.288 *pixels* válidos) em cada etapa do pré-processamento. Os mapas brutos já exibem formato próximo de uma distribuição gaussiana, com desvio padrão médio $\sigma_{\text{médio}} = 1,31 \times 10^{-3}\text{K}$. Após a remoção do nível DC, os histogramas permanecem centrados em zero, com distribuição aproxi-

Figura 6.8: Mapas *naive* da Simulação 1 para o canal 29, após diferentes etapas de processamento (de cima para baixo): (i) mapa original bruto; (ii) com nível DC removido; (iii) com nível DC removido e feixe ajustado (no domínio de *pixels*); (iv) com nível DC removido, feixe ajustado (no domínio de *pixels*) e máscara apodizada.

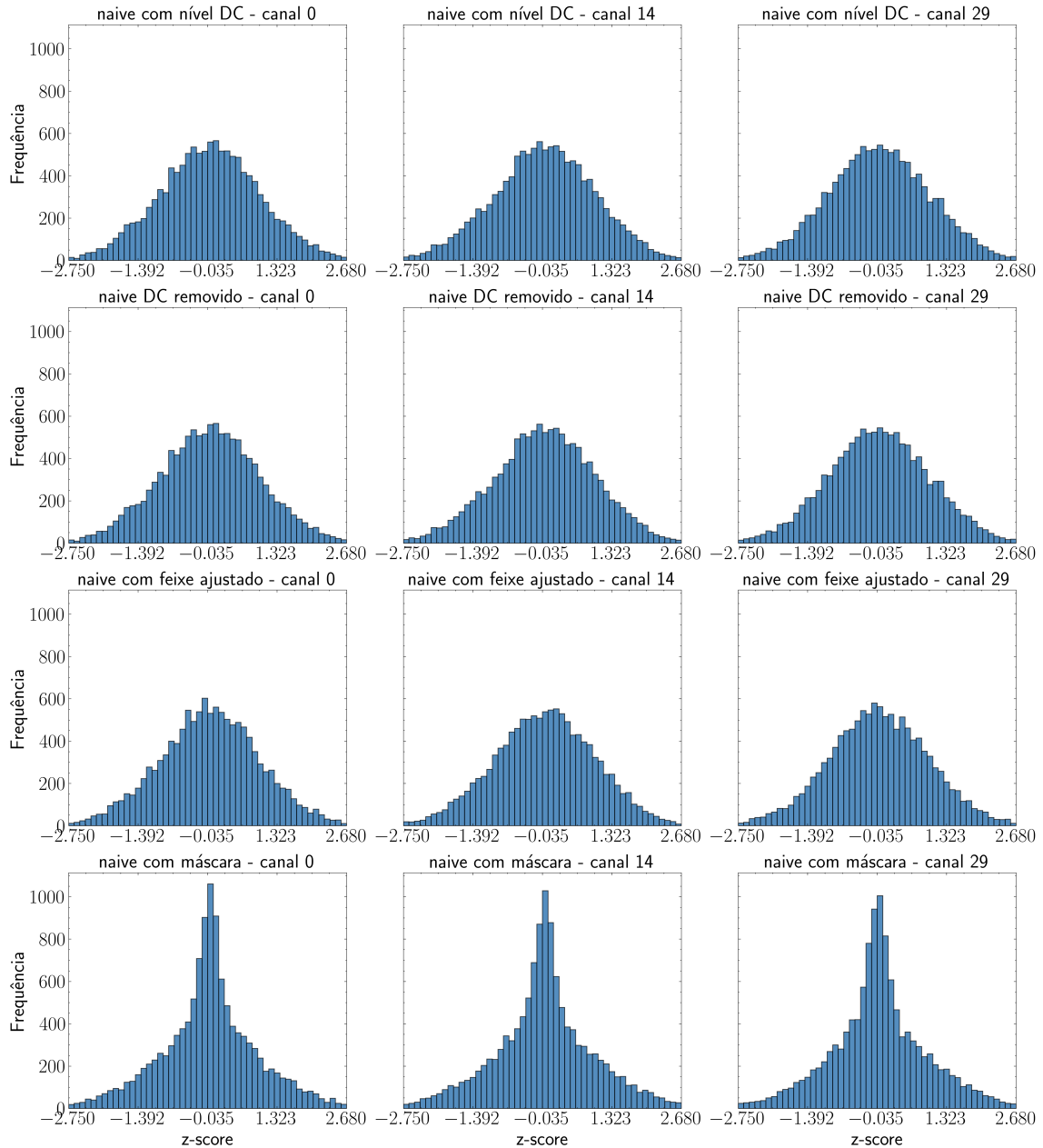


madamente normal e $\sigma_{\text{médio}} = 1,31 \times 10^{-3}\text{K}$. O ajuste do feixe mantém a simetria da distribuição, mas provoca um leve estreitamento, com $\sigma_{\text{médio}} = 1,06 \times 10^{-3}\text{K}$, refletindo a pequena redução na potência geral do mapa.

A aplicação da máscara apodizada aumenta esse efeito, concentrando ainda mais os valores em torno de zero e resultando em histogramas com picos mais acentuados e caudas mais longas, com $\sigma_{\text{médio}} = 8,14 \times 10^{-4}\text{K}$. Esse comportamento decorre da multiplicação do mapa por valores contínuos entre 0 e 1, atenuando suavemente os *pixels* próximos às bordas e preservando os centrais. Esses resultados demonstram a importância do pré-processamento aliado à normalização pelo *z-score* para preparar os dados para análises posteriores. A Tabela 6.2 apresenta os desvios padrão de cada canal em todas as etapas.

Após a análise dos mapas e histogramas, fica claro o efeito do pré-processamento, destacando a importância de sua aplicação. A remoção do nível DC centraliza os

Figura 6.9: Histogramas normalizados pelo z -score, gerados a partir dos mapas de observação simulados do BINGO, ilustrando as diferentes etapas do pré-processamento dos dados da Simulação 1.



A primeira linha apresenta os histogramas normalizados dos mapas brutos, com média próxima a 70 K, refletindo a amplitude do sinal sem tratamento. As linhas seguintes exibem, respectivamente, os histogramas dos mapas após remoção da componente DC, o ajuste do feixe e a aplicação da máscara apodizada. Nesses casos, as médias são próximas de zero, indicando a correta centralização dos dados. Os histogramas mantêm formato próximo ao de uma distribuição normal, com pequenas variações na dispersão e na forma, mais perceptíveis após a aplicação da máscara, em que assumem um formato com picos acentuados.

Tabela 6.2: Desvio padrão (K) dos mapas simulados para três canais de frequência distintos em quatro etapas de pré-processamento.

Etapa de pré-processamento	Canal 0	Canal 14	Canal 29
Mapas brutos	$1,29 \times 10^{-3}$	$1,26 \times 10^{-3}$	$1,28 \times 10^{-3}$
Remoção DC	$1,29 \times 10^{-3}$	$1,26 \times 10^{-3}$	$1,28 \times 10^{-3}$
Ajuste de feixe	$1,15 \times 10^{-3}$	$9,90 \times 10^{-4}$	$9,36 \times 10^{-4}$
Máscara apodizada	$8,78 \times 10^{-4}$	$8,64 \times 10^{-4}$	$7,17 \times 10^{-4}$

dados em torno de zero, enquanto o ajuste do feixe (realizado tanto no espaço harmônico como no de *pixels*) corrige a resolução dos mapas sem alterar significativamente a amplitude do sinal ou do ruído. A aplicação da máscara apodizada reduz a amplitude dos *pixels* próximos às bordas, pois o mapa é multiplicado por valores entre 0 e 1, atenuando essas regiões, mas preservando as estruturas centrais. Os mapas resultantes desse tratamento são então submetidos às ferramentas de separação de componentes, nas quais o sinal de HI deve ser separado do ruído $1/f$. Além do mapa de HI, os resultados incluem um conjunto de coeficientes estatísticos que permitem quantificar a eficiência da recuperação do mesmo. Esses resultados são apresentados na seção seguinte.

6.4 Separação de componentes

A fim de obter uma base de comparação consistente entre os mapas *naive* e iGLS, foram aplicadas técnicas de separação de componentes a ambos. Como o objetivo do *map-making* é gerar mapas que permitam uma melhor recuperação do sinal de interesse, submeter os resultados a ferramentas como GNILC e PCA, fornece um critério final e determinante para avaliar qual método produz o mapa mais fiel da distribuição de HI. Desta forma, foi necessário escolher qual técnica de separação de componentes seria adotada nas comparações. Para isso foram testadas as duas ferramentas apresentadas na Seção 4.3: GNILC e PCA. Os mapas pré-processados apresentados na Seção 6.3 foram submetidos a ambas as técnicas, e seus *outputs* comparados e avaliados segundo parâmetros estatísticos próprios para este tipo de análise: o coeficiente de correlação de Pearson (ρ) e o erro de reconstrução dos espectros de potência (η_{abs}), conforme já aplicado em DE MERICIA, Eduardo J. et al. (2023), Merícia (2023).

A seguir são apresentados os resultados obtidos com ambas as ferramentas: mapas de HI recuperados, seus histogramas e espectros de potência angular, coeficientes de Pearson e erros de recuperação dos espectros. Novamente, por conveniência, são

mostrados apenas os resultados obtidos com os mapas *naive* da Simulação 1, uma vez que todos os conjuntos de mapas foram submetidos aos mesmos processos e o objetivo destes testes foi identificar a ferramenta de separação de componentes com melhor desempenho nos cenários estudados.

6.4.1 Resultados GNILC

Para utilizar o GNILC é preciso adicionar como *inputs*, além dos mapas das simulações, o mapa de HI teórico, também denominado *prior* e o mapa de HI com efeito instrumental, denominado mapa de entrada ou mapa *target*. Ambos são gerados internamente utilizando a IM *pipeline*, a partir do modelo teórico selecionado determinado pelos parâmetros cosmológicos adotados. O *prior* consiste no mapa que contém apenas o sinal de HI e é utilizado para gerar o sinal de interesse nos mapas analisados, representando a melhor estimativa dos mapas de HI, baseada no conhecimento sobre o espectro de potência teórico do sinal.

Já o *target* é formado pelo mapa contendo o sinal de HI com efeito de feixe que idealmente deve ser recuperado após as etapas de tratamento de dados, gerado em uma realização independente. É assim chamado pois será o mapa que o GNILC utilizará como referência para separar as componentes do sinal. Ele é, essencialmente, o mapa que se espera obter após a separação de componentes caso não haja nenhuma contaminação por outros sinais ou ruídos. Nas simulações realizadas o *target* possuía apenas efeito do feixe. Assim, para permitir uma comparação correta entre o mapa reconstruído e o mapa de HI *target*, este foi também submetido ao processo de ajuste de feixe e aplicação da máscara apodizada. Em conjunto com o mapa da simulação, esse trio de mapas constitui os *inputs* do GNILC, que serão processados para gerar o mapa do sinal de HI recuperado.

A Figura 6.10 mostra os mapas *prior* de HI para cada um dos três canais de interesse, já com feixe ajustado e com a máscara apodizada aplicada. Pode-se observar como nos canais de menor número, ligados a menores frequências, a potência do sinal também é menor. Isto se deve ao fato de que os canais menores estão ligados a *redshifts* maiores, onde o HI estava disperso, ainda não organizado nas grandes estruturas. Ao comparar os mapas da Figura 6.10 com os das Figuras 6.6, 6.7 e 6.8 na seção anterior, pode-se notar o seu menor nível de potência, uma vez que são mapas apenas do sinal de HI, sem a presença do ruído. A Figura 6.11, por sua vez, apresenta os mapas de entrada (*target*) após ajuste do feixe e aplicação da máscara apodizada, para os três canais de frequência.

Figura 6.10: Mapa *prior* de HI usado na Simulação 1, após passar pelos processos de ajuste de feixe (no domínio de *pixels*) e aplicação da máscara apodizada, apresentado para os três canais de interesse.

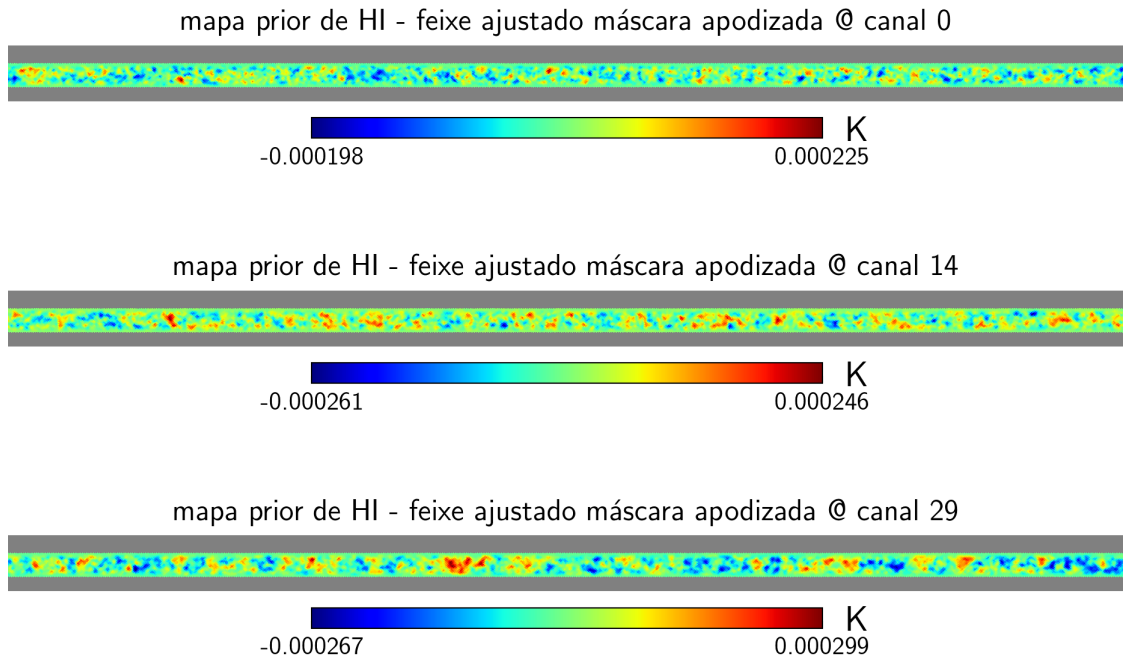
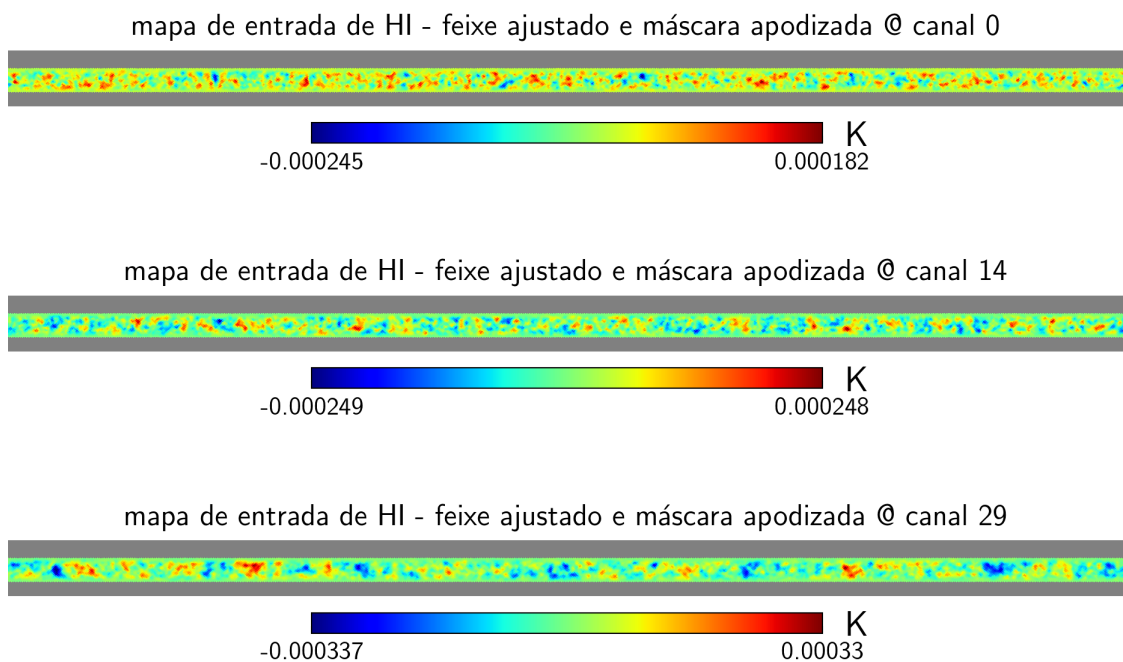
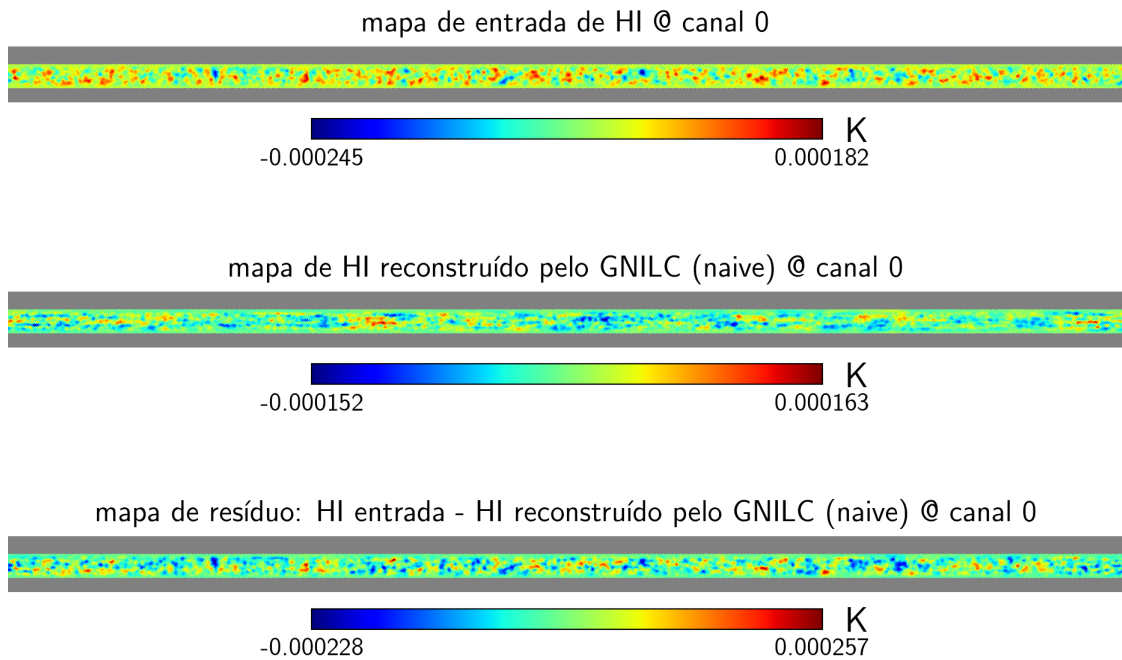


Figura 6.11: Mapa de entrada de HI usado na Simulação 1, após passar pelos processos de ajuste de feixe (no domínio de *pixels*) e aplicação da máscara apodizada.



Após a preparação dos inputs, o GNILC realiza a separação de componentes, produzindo mapas recuperados de HI. Para avaliar o desempenho do método, comparou-se os mapas reconstruídos pelo GNILC com os mapas de HI de entrada (*target*). A seguir, são apresentados os mapas obtidos a partir dos dados gerados pela IM *pipeline* para a Simulação 1, cujos parâmetros estão descritos na Tabela 6.1.

Figura 6.12: Mapas do canal 0 da Simulação 1 do BINGO.

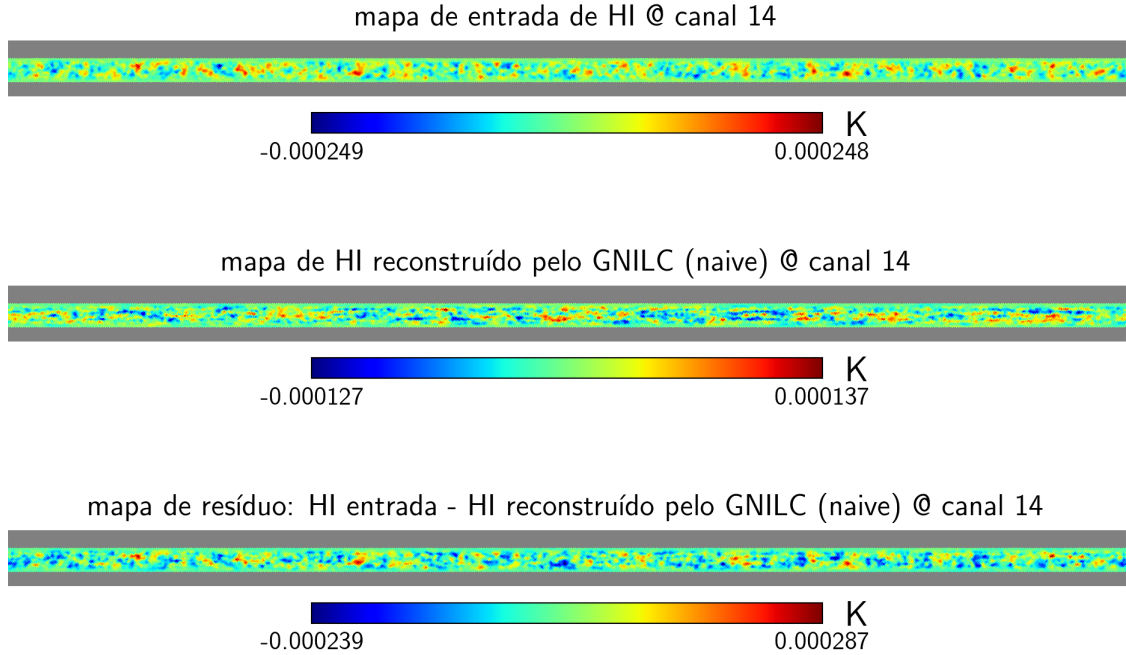


O painel superior mostra o mapa de entrada de HI (com feixe ajustado). O painel central apresenta o mapa de HI reconstruído pelo método GNILC (naive). O painel inferior corresponde ao mapa de resíduo, obtido subtraindo-se o mapa reconstruído do mapa de entrada (HI entrada - HI reconstruído). As barras de cores indicam a amplitude da temperatura em K.

As Figuras 6.12, 6.13 e 6.14 apresentam, para os canais 0, 14 e 29, respectivamente, os mapas de HI de entrada (sinal que se deseja recuperar), os mapas de HI reconstruídos pelo GNILC (*naive*) e os mapas de resíduos, obtido pela diferença entre ambos.

Nota-se de imediato que os mapas reconstruídos apresentam níveis de intensidade inferiores aos do mapa de entrada. Esse comportamento decorre do fato de que o GNILC remove parte do sinal de HI em regiões de baixa relação sinal-ruído, onde o método tem dificuldade em distinguir corretamente o subespaço associado ao HI daquele dominado pelo ruído $1/f$. Além disso, como o ruído $1/f$ com $\beta = 0,25$

Figura 6.13: Mapas do canal 14 da Simulação 1 do BINGO.

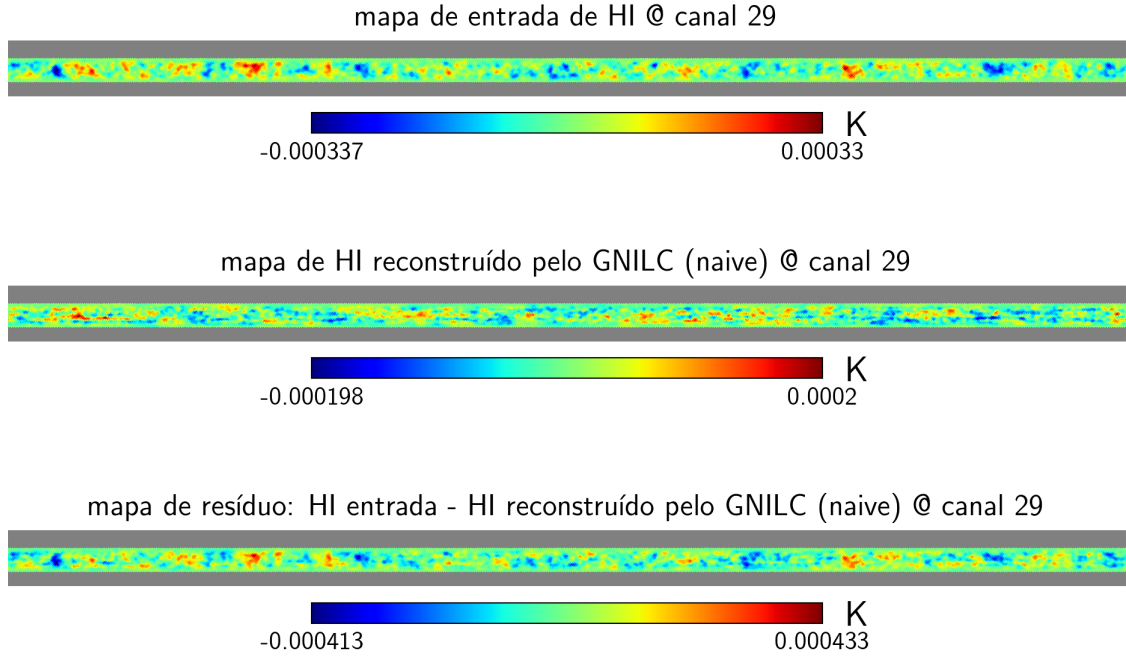


O painel superior mostra o mapa de entrada de HI. O painel central apresenta o mapa de HI reconstruído pelo método GNILC (naive). O painel inferior corresponde ao mapa de resíduo, obtido subtraindo-se o mapa reconstruído do mapa de entrada (HI entrada - HI reconstruído). As barras de cores indicam a amplitude da temperatura em K.

não é perfeitamente correlacionado em frequência, parte de sua contribuição não é identificada como contaminante e, portanto, não é removida pelo GNILC. Isso leva o método a confundir parcialmente sinal e ruído, resultando em mapas contaminados tanto por perdas do HI real quanto por resíduos de ruído instrumental.

Também se observam, nos mapas reconstruídos, estruturas alongadas na direção horizontal (semelhantes a faixas), ausentes no mapa de HI de entrada. Esse padrão é característico do ruído $1/f$, que tende a produzir listras alinhadas com a direção de varredura de instrumentos como o BINGO, reforçando que tal contribuição não foi completamente mitigada. Os histogramas correspondentes a esses mapas são mostrados na Figura 6.15, enquanto seus desvios padrão estão apresentados na Tabela 6.3. Nota-se ainda que o mapa de HI de entrada apresenta maior concentração de *pixels* com valores em torno de zero e uma dispersão mais ampla de valores, enquanto os mapas reconstruídos exibem distribuições mais suaves, reflexo da filtragem exercida pelo método, que atenua não apenas o ruído, mas também parte do sinal de interesse. Pode-se observar também que o desvio padrão dos resíduos são sempre

Figura 6.14: Mapas do canal 29 da Simulação 1 do BINGO.



O painel superior mostra o mapa de entrada de HI. O painel central apresenta o mapa de HI reconstruído pelo método GNILC (*naive*; com feixe ajustado). O painel inferior corresponde ao mapa de resíduo, obtido subtraindo-se o mapa de HI reconstruído pelo GNILC (*naive*) do mapa de entrada de HI. As barras de cores indicam a amplitude da temperatura em K.

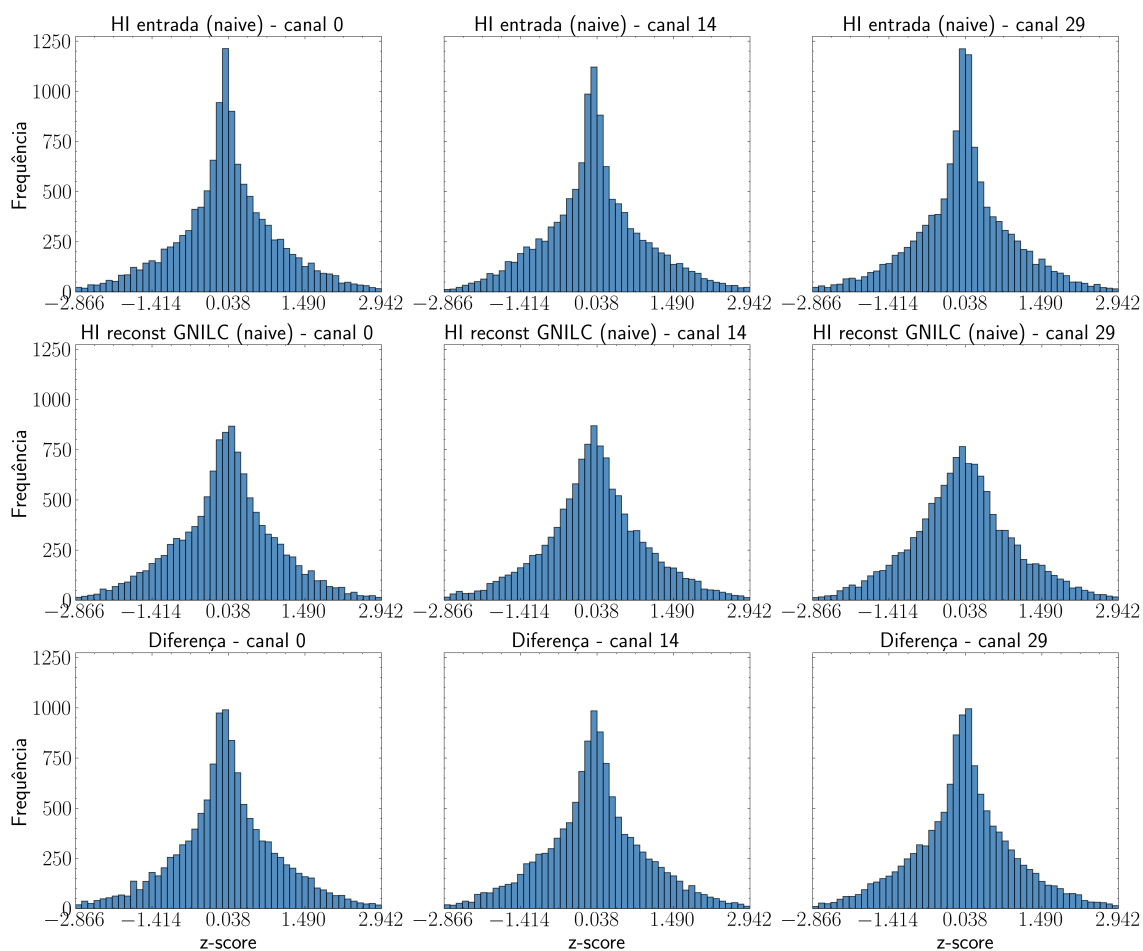
maiores do que os do mapa de entrada e o recuperado, o que também caracteriza uma recuperação insatisfatória do sinal.

Tabela 6.3: Desvio padrão (σ) dos mapas normalizados por *z-score* nos canais 0, 14 e 29.

Tipo de mapa	Canal 0	Canal 14	Canal 29
HI entrada	$4,44 \times 10^{-5}$	$5,25 \times 10^{-5}$	$7,27 \times 10^{-5}$
HI reconstruído GNILC (<i>naive</i>)	$3,17 \times 10^{-5}$	$3,42 \times 10^{-5}$	$4,93 \times 10^{-5}$
Resíduo	$5,59 \times 10^{-5}$	$6,33 \times 10^{-5}$	$8,98 \times 10^{-5}$

Para uma análise quantitativa dos resultados da separação de componentes, foram utilizados dois parâmetros: o parâmetro de correlação de Pearson (ρ) e o erro relativo médio dos espectros de potência angular (C_ℓ), denominado aqui η_{abs} . O gráfico de dispersão permite visualizar a correspondência ponto a ponto entre os mapas de HI

Figura 6.15: Histogramas dos mapas de entrada de HI (linha superior), dos mapas de HI reconstruídos pelo GNILC (naive; linha central) e dos mapas de resíduo, que são a diferença entre o mapa de HI de entrada e o mapa de HI reconstruído pelo GNILC (naive; linha inferior).



Ambos possuem comportamento parecido, embora os mapas recuperados apresentem maior dispersão e suavidade nos seus histogramas.

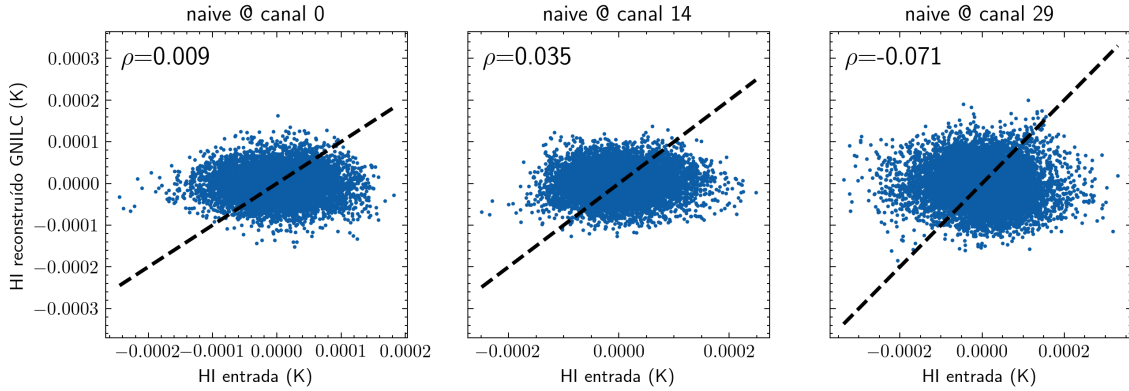
de entrada e reconstruído, enquanto o coeficiente de Pearson quantifica linearmente essa correlação, indicando o quanto do sinal original foi preservado na reconstrução. A Figura 6.16 apresenta o gráfico de dispersão entre os valores dos *pixels* do mapa de entrada de HI (eixo horizontal) e os respectivos valores do mapa de HI reconstruído (eixo vertical), ambos em K, para os três canais de interesse. A proximidade dos pontos em relação à diagonal indica o grau de correlação entre os dois mapas, destacando a eficiência do processo de reconstrução. Cada gráfico apresenta também o valor do coeficiente de correlação de Pearson calculado para o par de mapas correspondente, medindo a qualidade do processo de recuperação do sinal de HI.

Nos três gráficos, os pontos estão espalhados em torno de uma nuvem quase circular ou oval (na horizontal) centrada em zero. Isso indica que não há uma relação linear clara entre o mapa de entrada e o mapa reconstruído. Um indicativo de boa reconstrução é uma nuvem mais alongada ao longo da diagonal (linha preta tracejada).

Os valores de ρ são consistentemente pequenos entre os canais, o que é evidenciado visualmente pela dispersão dos pontos e indica baixo grau de correlação entre os mapas de entrada e reconstruído. A Figura 6.17 apresenta o coeficiente de Pearson (eixo vertical) calculado para cada um dos 30 canais de frequência adotados para o BINGO neste trabalho (eixo horizontal). O valor médio dos coeficientes de Pearson calculado sobre todas as frequências é $\bar{\rho} = 0,004$, variando muito pouco entre os canais. Os valores próximos de zero indicam baixa correlação entre os mapas de entrada e reconstruídos. Ou seja, o sinal de HI não foi recuperado de forma satisfatória nessa configuração adotada para a Simulação 1 (*naive map-making* e separação de componentes GNILC).

Para obter uma avaliação complementar da eficácia do método de separação de componentes na recuperação do sinal de HI, pode-se comparar também os espectros de potência angular dos mapas, como mostrado na Figura 6.18. Essa análise fornece informações distintas das apresentadas nas Figuras 6.16 e 6.17, pois avalia a potência do sinal de HI em diferentes escalas angulares. Observa-se que, apesar de que em escalas menores (altos valores de ℓ) os espectros tendem a se aproximar, ainda existem discrepâncias significativas de quase uma ordem de magnitude nas escalas maiores e intermediárias. Como o espectro de potência do mapa reconstruído é inferior ao do HI real, conclui-se que o GNILC removeu parte do sinal cosmológico, classificando parte dele como ruído. Em outras palavras, diversos modos foram interpretados como contaminantes e suprimidos, o que levou à perda de informação do sinal de HI.

Figura 6.16: Coeficiente de Pearson para a Simulação 1, realizada com *map-making naive* e separação de componentes feita pelo GNILC, calculados para três canais de frequência.



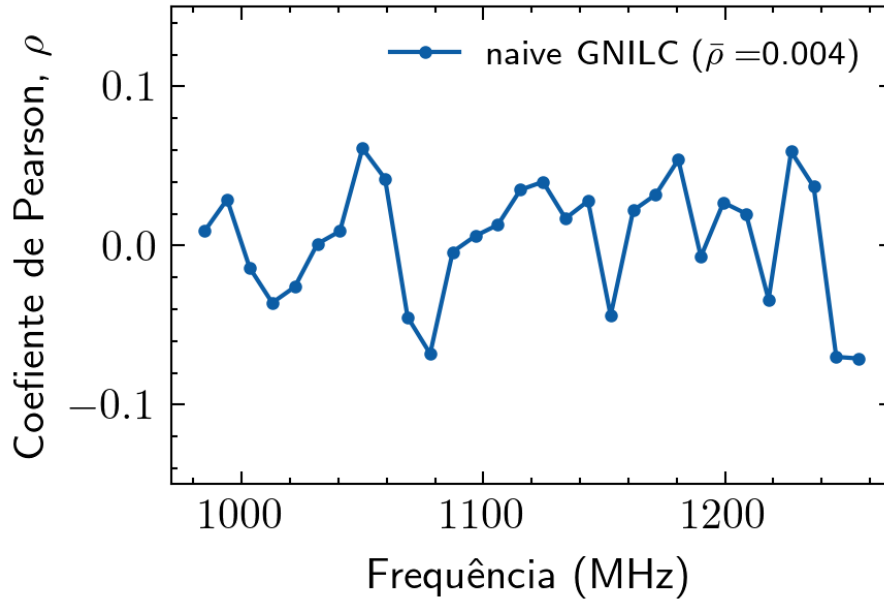
O eixo horizontal representa o mapa real de HI e o eixo vertical o mapa reconstruído. A alta dispersão dos pontos e sua forma aproximadamente circular indica baixa correlação entre ambos.

O conjunto de tais resultados mostra a dificuldade encontrada pelo GNILC em realizar a separação de componentes para os mapas *naive* contaminados com ruído correlacionado. As razões da remoção de parte do sinal de HI em mapas simulados com ruído $1/f$ e correlação espectral $\beta = 0,25$ pelo GNILC ainda não são bem compreendidas. Uma correlação espectral mais alta também não apresentou melhores resultados. Um estudo mais detalhado das razões do fraco desempenho do GNILC, bem como do entendimento do impacto do parâmetro de correlação na recuperação do sinal cosmológico foi deixada para um trabalho futuro. Em função dos resultados desta seção, decidiu-se testar o desempenho do PCA como ferramenta de separação de componentes e compara-los com os resultados obtidos com o GNILC. Os resultados obtidos estão descritos na próxima seção.

6.4.2 Resultados PCA

Para o PCA, foram usados como *inputs* os mapas simulados de HI com ruído $1/f$, já com nível DC removido, com ajuste do feixe e com a máscara aplicada, conforme apresentado nas Figuras 6.6, 6.7 e 6.8. Como o PCA é uma técnica mais simples, não são necessários outros elementos de entrada para seu funcionamento. Ele identifica os modos que correspondem às maiores dispersões estatísticas, geralmente associadas ao ruído e aos contaminantes e permite separar esses componentes das contribuições de menor variância, frequentemente relacionadas ao sinal. Foi testada durante a sua

Figura 6.17: Coeficiente de Pearson calculado para cada um dos 30 canais e coeficiente de Pearson médio para a Simulação 1, realizada com *map-making naive* e separação de componentes feita pelo GNILC, calculado em todos os canais de frequência.



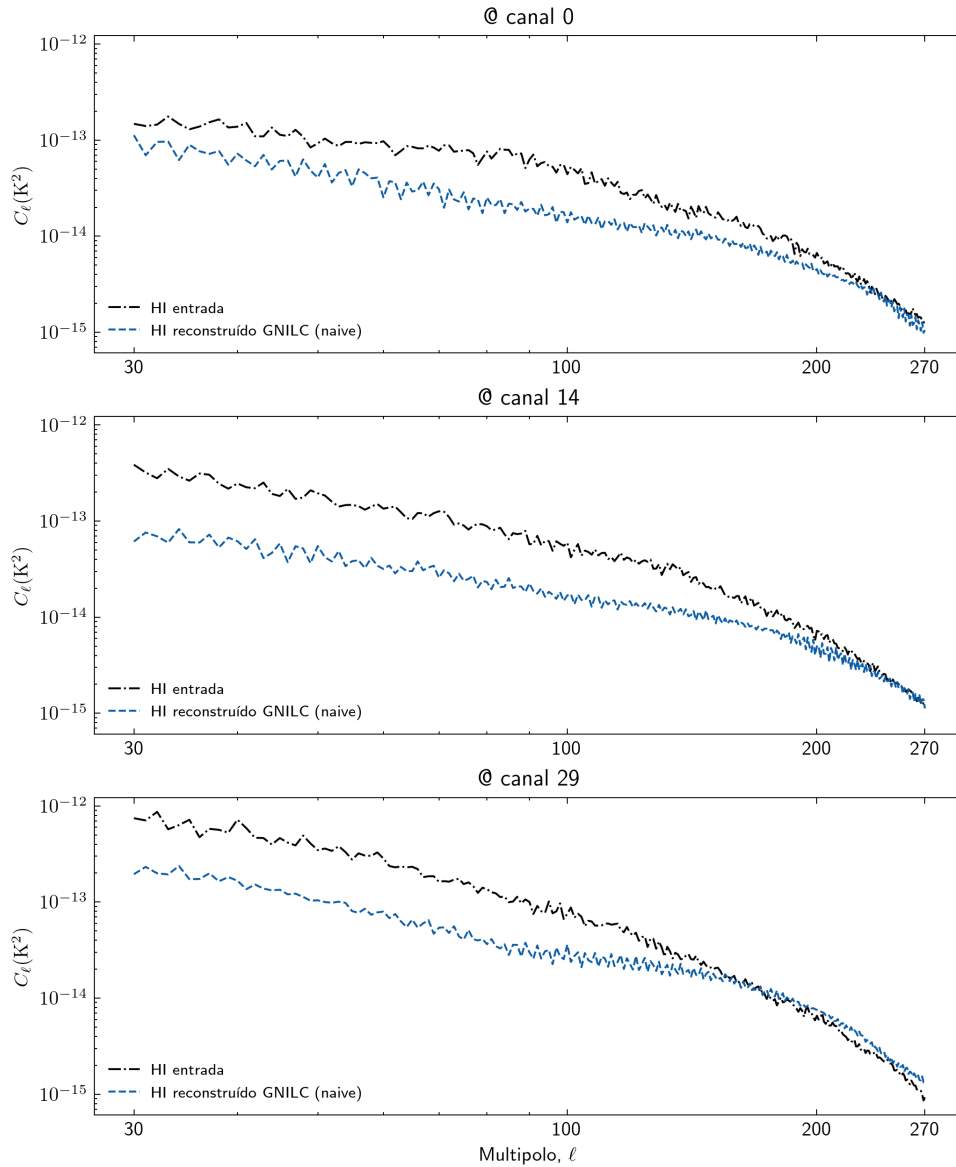
O valor de ρ é próximo de zero em todos os canais, resultando em um valor médio de $\bar{\rho} = 0.004$. Esse é mais um indicativo de que praticamente não há correlação entre os mapas.

aplicação a remoção de até 11 destes modos de maior intensidade, tendo obtido o melhor resultado com a remoção de 9 modos.

As Figuras 6.19, 6.20 e 6.21 apresentam os mapas obtidos após a remoção de 7 a 11 componentes principais (modos principais) identificados pelo PCA. O primeiro mapa presente nas três figuras é o mapa de HI de entrada, que também passou pelo processo de ajuste de feixe no espaço de *pixels*, detalhado na Seção 5. Este é o mesmo utilizado para comparar os resultados obtidos com a recuperação efetuada pelo GNILC (mapa *target*).

A perda progressiva de potência com o aumento do número de modos principais removidos é esperada, pois os primeiros componentes principais concentram a maior parte da variância do sinal observado, geralmente associada aos *foregrounds*. Porém, ao remover um número elevado de modos, parte da informação do HI também é eliminada, reduzindo a potência recuperada do sinal cosmológico. Ainda podem ser observadas as estruturas residuais alongadas no sentido horizontal, típicas do ruído $1/f$, mesmo após a remoção de diversos modos principais pelo PCA. As diferenças

Figura 6.18: Comparação entre o espectro de potência angular calculado para o mapa real de HI, em preto e para o mapa recuperado com o GNILC, em azul.

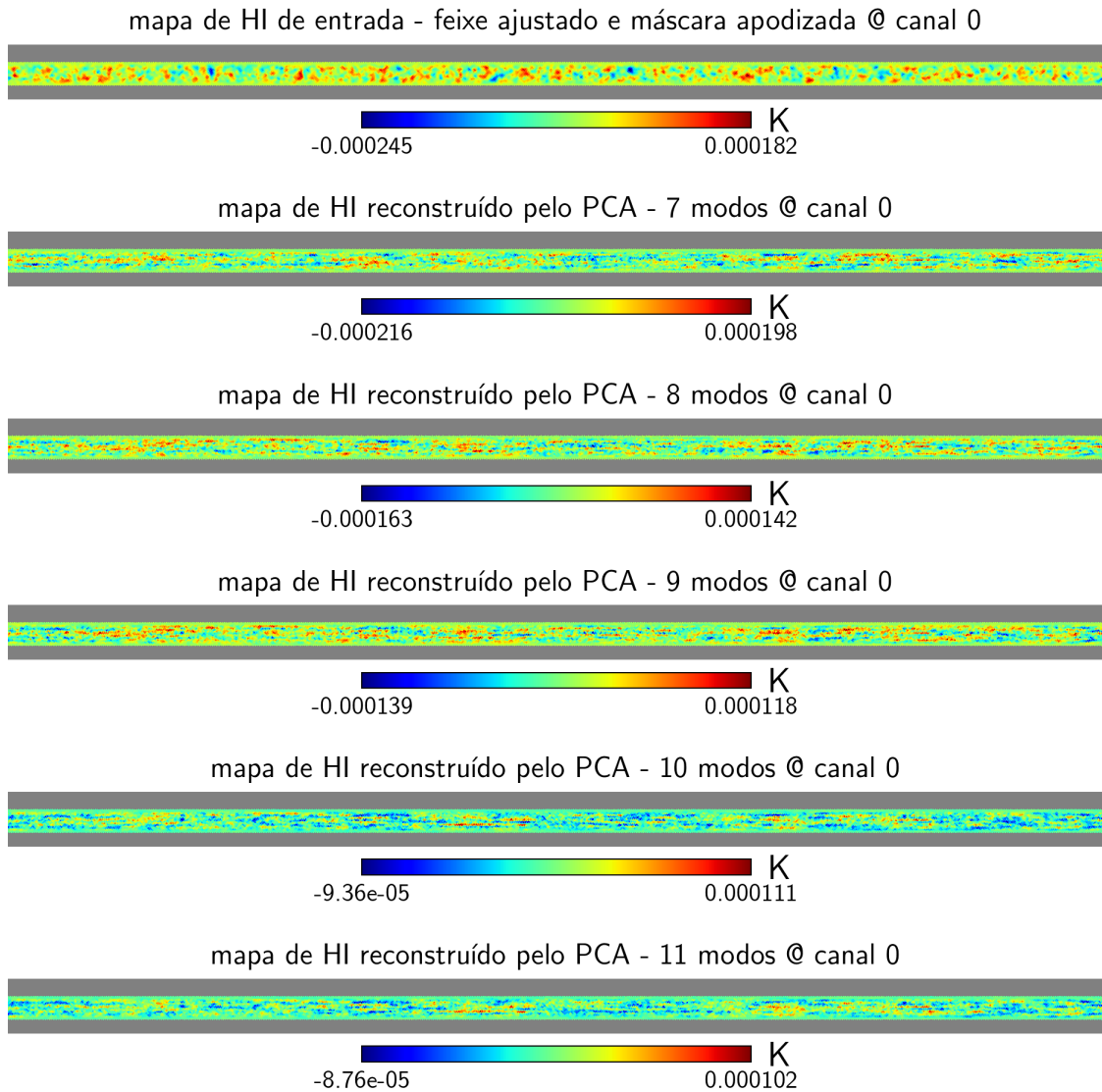


Nota-se a diferença entre ambos em pequenos e médios multipolos, embora haja uma recuperação coerente para menores escalas.

entre os mapas recuperados e o mapa de entrada também podem ser percebidas em uma simples inspeção visual.

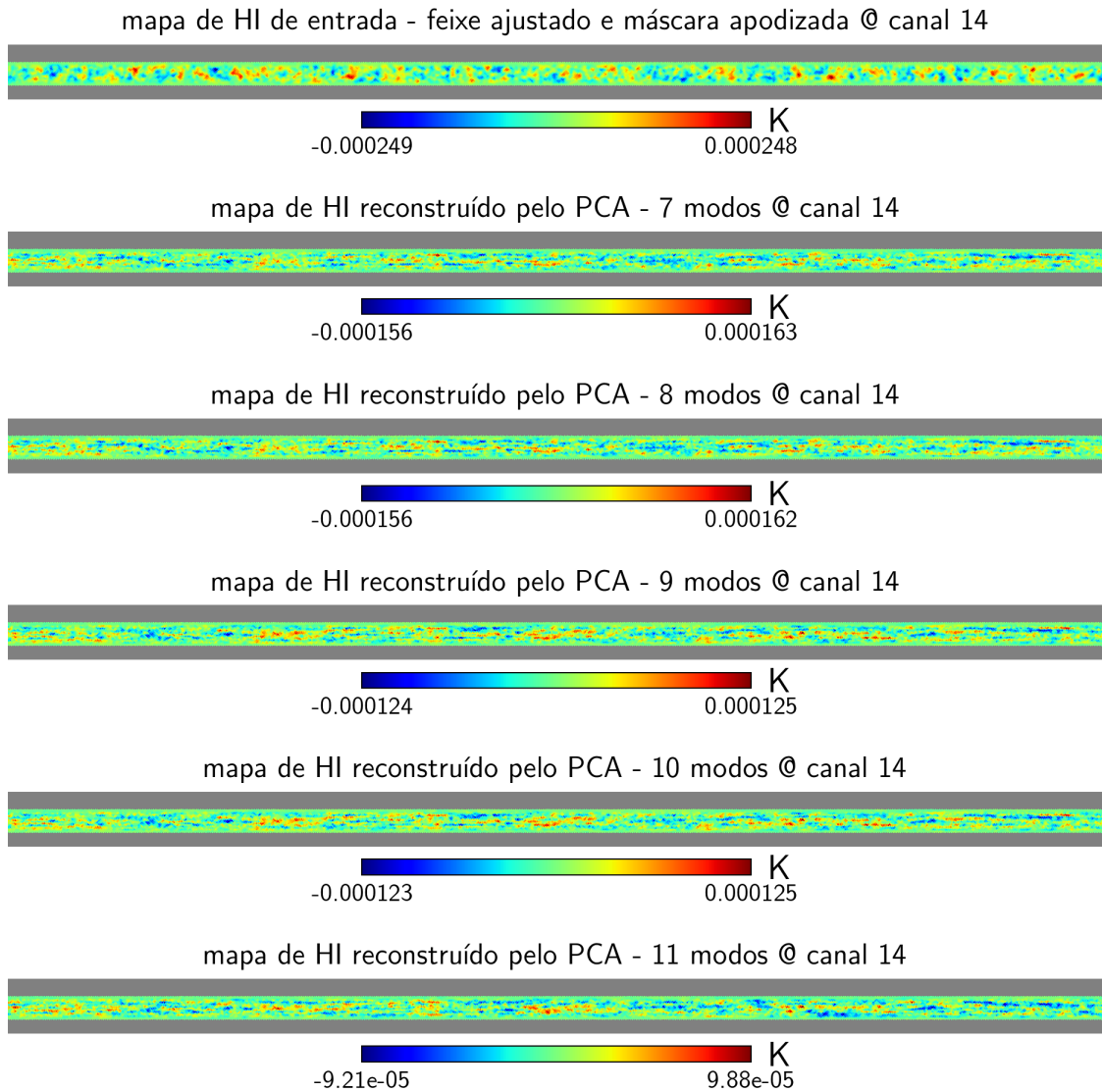
Também foram comparados os espectros de potência angular dos mapas de HI recuperados pelo PCA com o mapa observado (HI + ruído $1/f$) e com o mapa de entrada de HI. A Figura 6.22 apresenta os espectros de potência angular (C_ℓ) do

Figura 6.19: Mapas recuperados de HI para a simulação 1, realizada com *map-making naive* para o canal 0 após aplicação do PCA e retirada de 7 à 11 modos.



mapa de HI com ruído, do mapa apenas de HI e dos mapas de HI reconstruídos pelo PCA, onde se observa claramente a perda de potência associada à remoção dos modos de maior intensidade. Assim como o que ocorre com os mapas recuperados pelo GNILC, o espectro de potência dos mapas obtidos com PCA também se afasta do espectro obtido com o mapa de entrada de HI. Novamente, os espectros reconstruídos encontram-se abaixo do espectro de referência do HI, indicando que parte significativa do sinal de interesse foi removida por ter sido interpretada como ruído. Por outro lado, em altos multipolos, as curvas se invertem, mostrando que o PCA superestimou a contribuição do sinal.

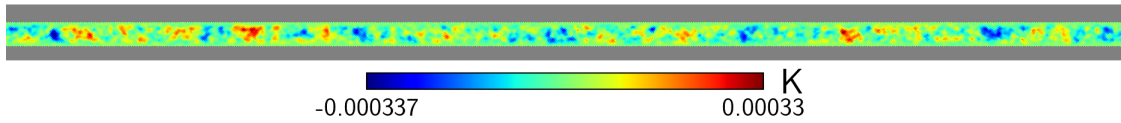
Figura 6.20: Mapas recuperados de HI para a simulação 1, realizada com *map-making naive* para o canal 14 após aplicação do PCA e retirada de 7 à 11 modos.



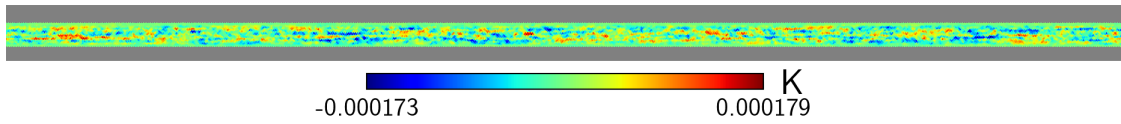
Com o objetivo de quantificar a recuperação do sinal, calculou-se o erro de recuperação dos espectros de potência (η_{abs}) ao longo dos canais para cada número de modos removidos, conforme apresentado na Figura 6.23. Observa-se que o erro diminui progressivamente até a retirada de oito ou nove modos, ponto em que atinge seu menor valor (59,4 %), voltando a aumentar em seguida. Assim, considera-se que a melhor recuperação do sinal de HI com o PCA ocorre quando são removidos 8 ou 9 modos de maior intensidade. No entanto, mesmo nesse caso, o erro permanece da ordem de 60 %, o que é indesejável e evidencia também a baixa eficiência do PCA na recuperação do sinal de HI na presença de ruído $1/f$.

Figura 6.21: Mapas recuperados de HI para a simulação 1, realizada com *map-making naive* para o canal 29 após aplicação do PCA e retirada de 7 à 11 modos.

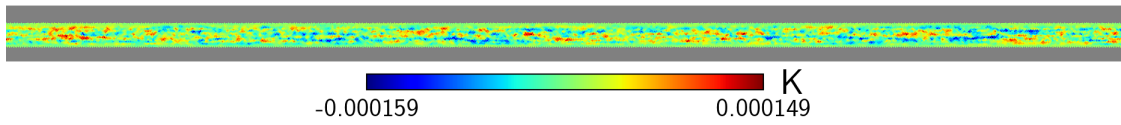
mapa de HI de entrada - feixe ajustado e máscara apodizada @ canal 29



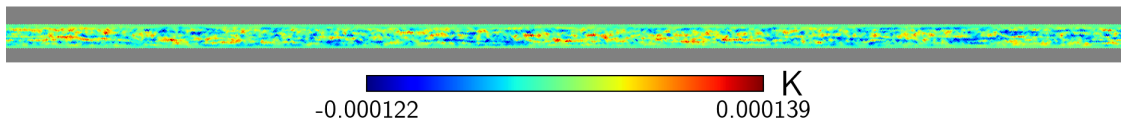
mapa de HI reconstruído pelo PCA - 7 modos @ canal 29



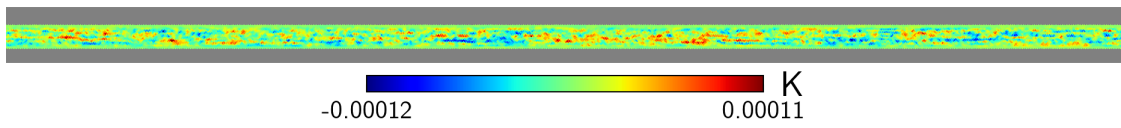
mapa de HI reconstruído pelo PCA - 8 modos @ canal 29



mapa de HI reconstruído pelo PCA - 9 modos @ canal 29



mapa de HI reconstruído pelo PCA - 10 modos @ canal 29



mapa de HI reconstruído pelo PCA - 11 modos @ canal 29

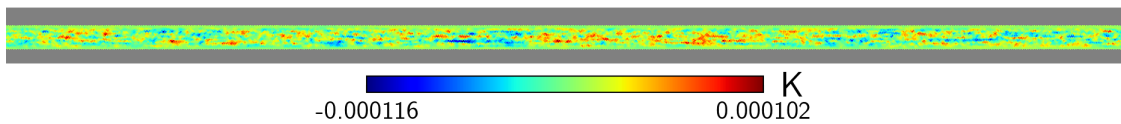
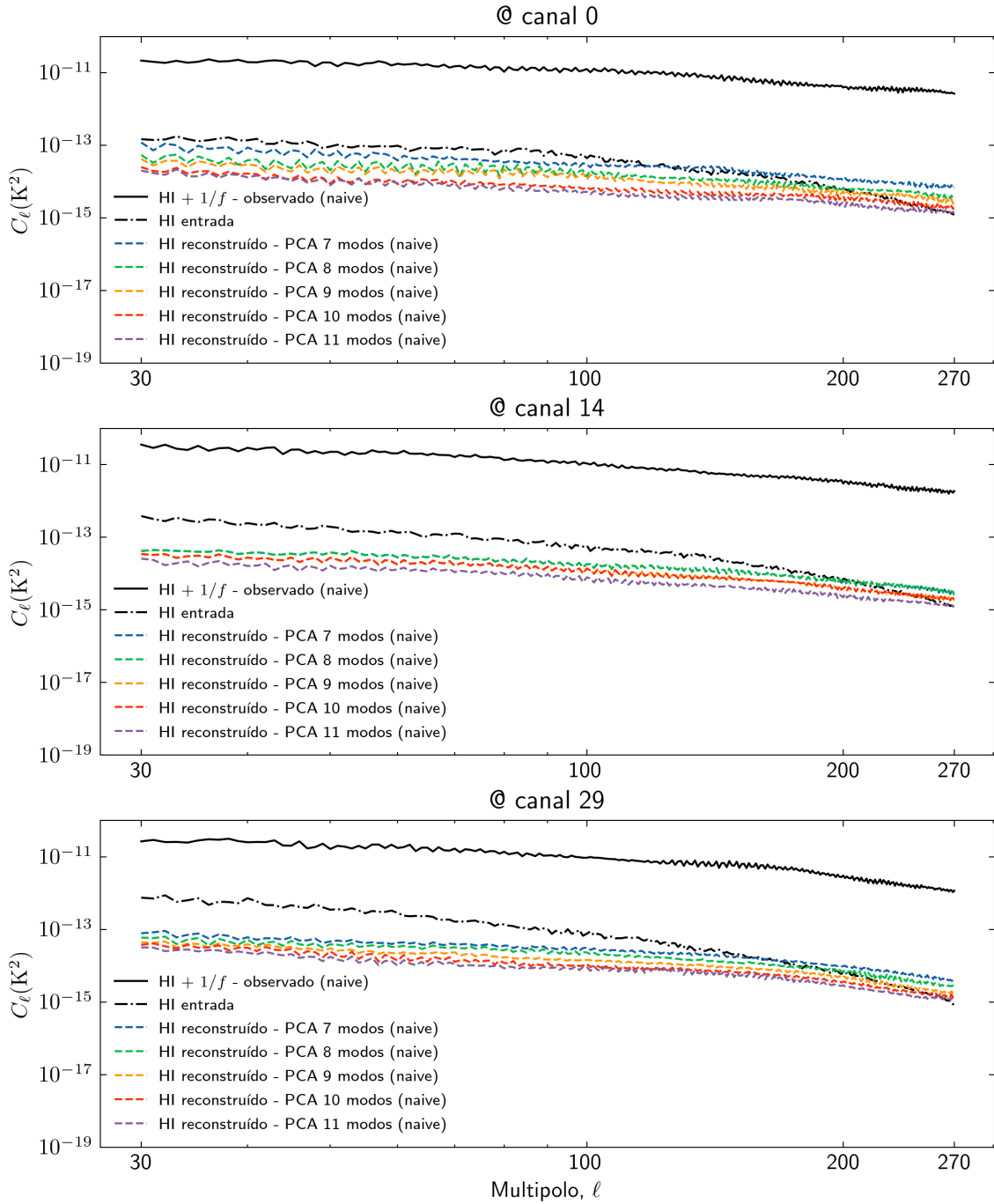
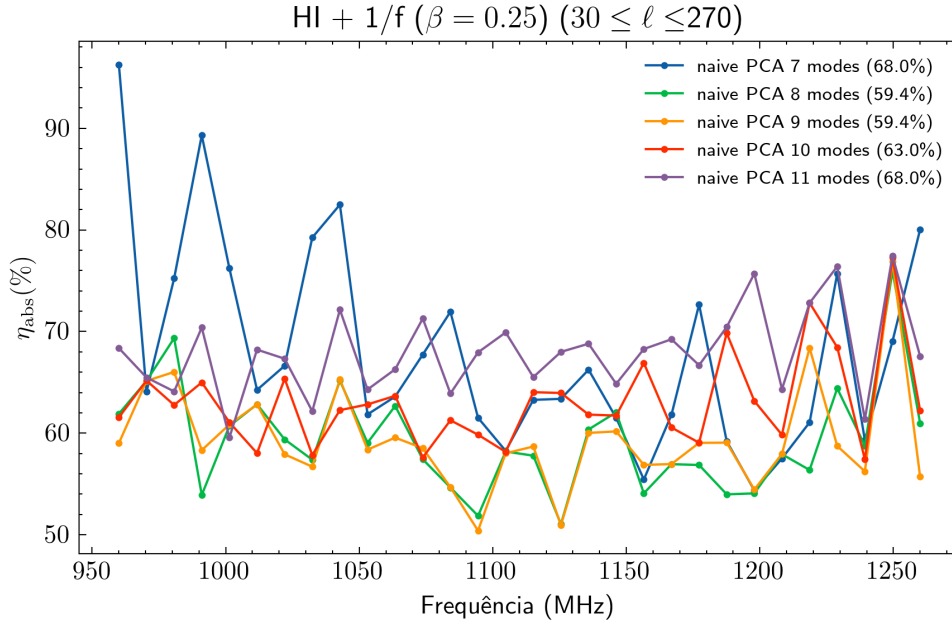


Figura 6.22: Espectros de potência angular dos mapas de HI recuperados para a Simulação 1, obtidos com o método *naive* de *map-making* após a remoção de 7 a 11 modos com o PCA.



São mostrados, para três canais de frequência (0, 14 e 29), os espectros do mapa observado (HI + ruído $1/f$), do HI de entrada e dos mapas reconstruídos.

Figura 6.23: Erro médio dos espectros de potência angular (η_{abs}) para os mapas de HI recuperados na Simulação 1, obtidos com o métodos *naive* de *map-making* e PCA de separação de componentes.



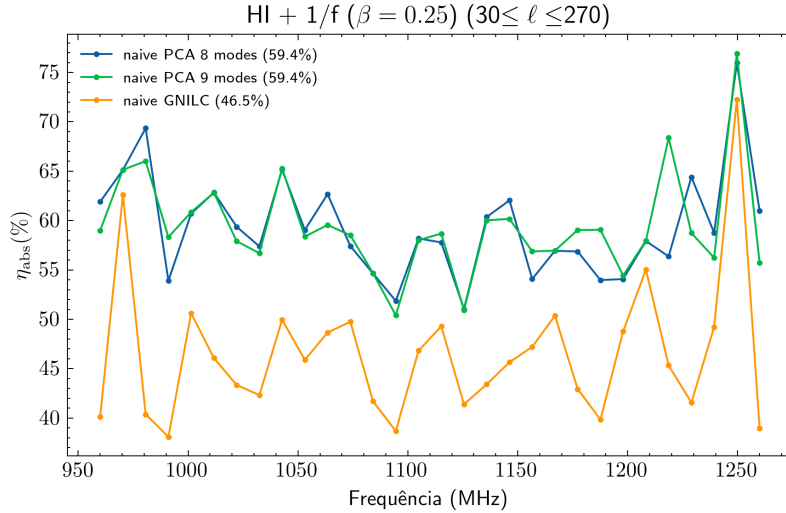
O erro é mostrado em função da frequência para diferentes números de modos removidos. O menor valor de erro médio (59,4 %) foi alcançado com a retirada de 8 ou 9 modos principais, indicado pelas linhas verde e amarela.

6.4.3 Comparação entre os métodos GNILC e PCA

Com base nesses resultados, realizou-se a comparação da eficácia das duas ferramentas de separação de componentes na recuperação do sinal de HI a partir dos mapas observados simulados. Para isso, analisaram-se os erros relativos de recuperação (η_{abs}) de ambas as técnicas, calculados a partir da diferença entre o C_ℓ do mapa real e o do mapa reconstruído após a separação de componentes. O desempenho do GNILC foi comparado ao do PCA na configuração de remoção de 8 e 9 modos. A Figura 6.24 apresenta os erros médios dos espectros de potência angular para os mapas de HI recuperados por cada método, considerando o intervalo de multipolos de interesse $30 \leq \ell \leq 270$.

No caso do PCA, com remoção de 8 modos (linha azul) e 9 modos (linha verde), o erro médio de recuperação foi praticamente o mesmo para ambos, $\eta_{\text{abs}} \simeq 59,4 \%$. Já para o GNILC (linha laranja), o erro médio de recuperação foi menor, $\eta_{\text{abs}} \simeq 46,5 \%$, indicando melhor desempenho na recuperação do sinal de HI.

Figura 6.24: Comparação entre o erro médio dos C_ℓ dos mapas recuperados de HI, para a Simulação 1, produzidos com *map-making naive* e recuperados com o GNILC (curva amarela) e com o PCA para a retirada de 8 modos (curva azul) e 9 modos (curva verde).



Este resultado se deve ao fato de que o GNILC é uma ferramenta de separação de componentes mais sofisticada, realizando a decomposição do sinal no espaço harmônico (OLIVARI et al., 2016). Isso permite determinar o número ideal de componentes para cada escala angular, evitando que o sinal de HI seja superestimado ou subestimado durante a recuperação. Uma discussão mais detalhada sobre a comparação entre o GNILC e o PCA pode ser encontrada em (OLIVARI et al., 2016). Como o GNILC apresentou melhor desempenho na recuperação do sinal de HI nas simulações realizadas neste trabalho, ele será adotado como a ferramenta padrão de separação de componentes em todas as simulações a seguir.

6.5 Comparação entre os métodos *naive* e iGLS

Esta seção apresenta os mapas de intensidade para a Simulação 1, produzidos com o *map-making naive* e com o iGLS, bem como a comparação entre ambos os resultados. As TODs geradas com a IM *pipeline* foram usadas para produzir mapas de intensidade com o *map-making naive*, feito pela IM *pipeline*, e com o *map-making iGLS*, feito pelo ROMA.

Estes mapas passaram pelas etapas de pré-processamento descritos na Seção 5 e pela separação de componentes realizada pelo GNILC, conforme explicado na Seção 4.3. Os resultados para ambas as técnicas de *map-making* foram comparados através dos parâmetros descritos na Seção 5.6.

Os resultados da Simulação 1, realizada para um tempo de observação de 3 dias utilizando $f_{\text{knee}} = 0,001$ Hz, $\beta = 0,25$, e $N_{\text{side}} = 128$, gerados pela IM *pipeline*, foram discutidos nas Seções 6.2, 6.3 e 6.4. A seguir, são apresentados os resultados da mesma Simulação 1, mas executada com a ferramenta de *map-making* ROMA (método iGLS), incluindo uma comparação com os resultados obtidos com o método *naive*. São apresentados os mapas recuperados após a separação de componentes, os espectros de potência angular correspondentes, os coeficientes de Pearson e os erros de recuperação dos C_ℓ , calculados para cada um dos mapas recuperados em relação ao mapa real de HI.

- **Mapas de HI recuperados pelo GNILC**

A seguir são apresentados os resultados obtidos com o GNILC para a recuperação do sinal de HI a partir dos mapas produzidos com o ROMA (mapas iGLS). A Figura 6.25 mostra o mapa de HI de entrada (real) e os mapas iGLS recuperados para os canais 0, 14 e 29, respectivamente. Nota-se uma perda significativa de potência em relação ao mapa real, especialmente nos canais de menor frequência, indicando a remoção de parte do sinal do HI. Visualmente, os mapas apresentam baixa correlação com o mapa de entrada de HI.

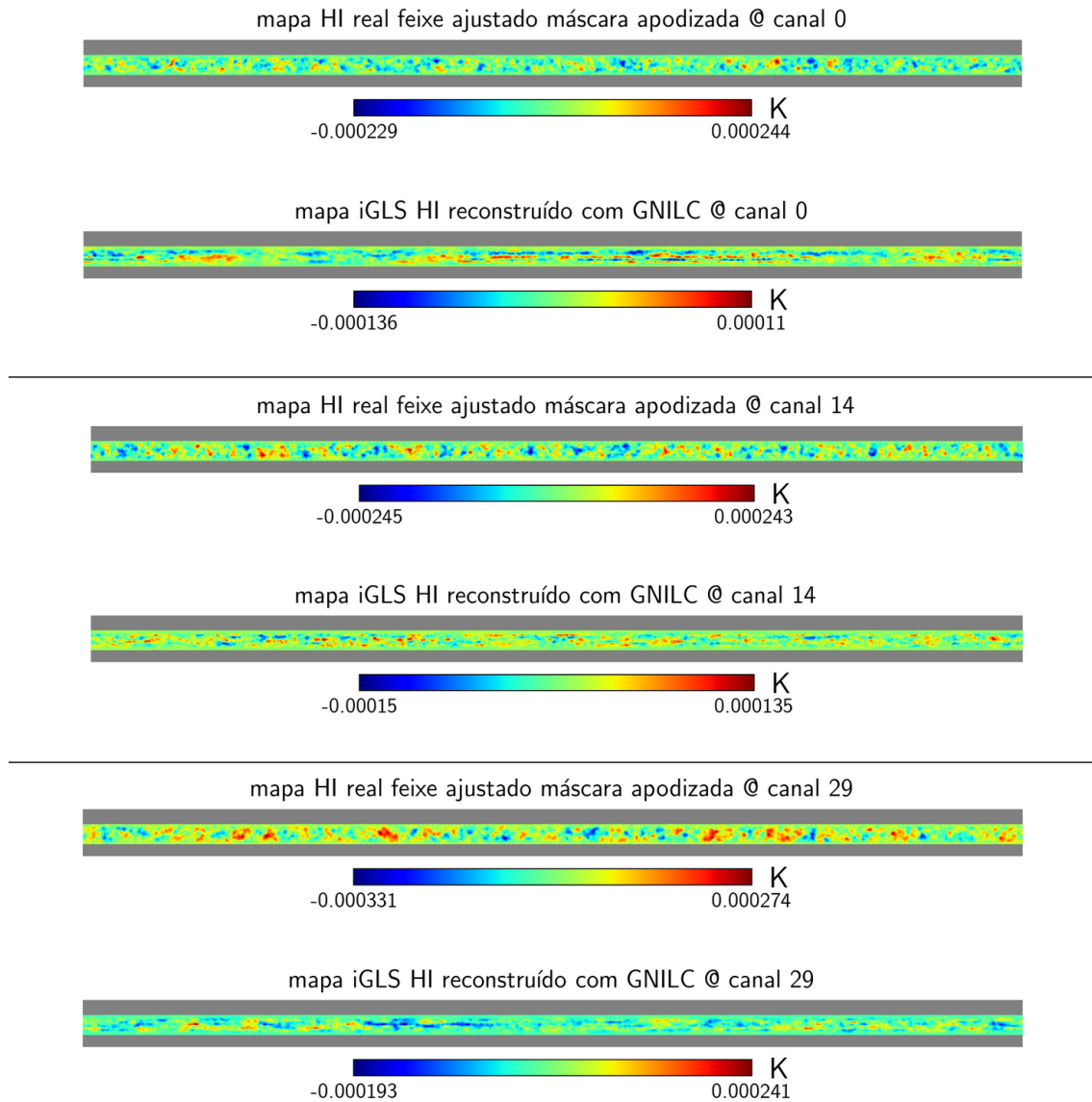
Assim como nos mapas produzidos com a IM *pipeline*, os mapas obtidos com o ROMA apresentam estruturas alongadas na direção horizontal de varredura, típicas do ruído $1/f$. A presença dessas estruturas indica que o ruído não foi completamente removido e as discrepâncias em relação ao mapa real mostram que a separação de componentes não recuperou adequadamente o sinal de HI.

- **Espectros de potência angular**

Foram também plotados os espectros de potência angular dos mapas da Figura 6.25, que foram novamente comparados com o espectro do HI de entrada, apresentados na Figura 6.26 para os canais de interesse. Nela, a curva preta representa o espectro de potência do mapa de HI real e a curva azul, o do mapa iGLS recuperado. Observa-se ainda, mesmo para os mapas iGLS, a discrepância entre os C_ℓ 's do mapa real e dos mapas recuperados, especialmente em grandes escalas angulares. O fato da curva para o sinal reconstruído estar abaixo da curva para o sinal real indica que parte do HI foi removido como ruído.

Mesmo para pequenas escalas ainda há diferenças, que tendem a diminuir para

Figura 6.25: Mapa de HI de entrada (superior) e mapa reconstruído pelo GNILC (inferior) a partir do mapa observado (HI + 1/f; Simulação 1), obtido com o iGLS para os canais 0 (superior), 14 (centro) e 29 (inferior).

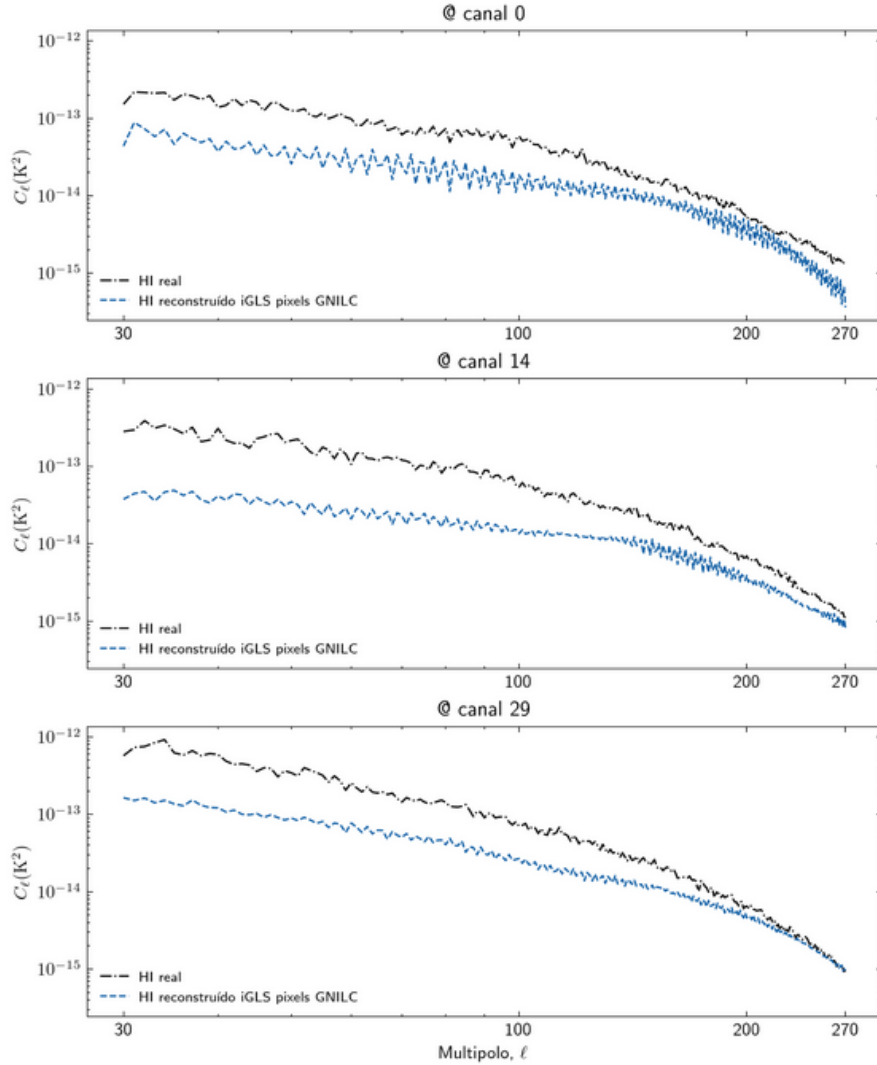


Os mapas de cada canal estão separados por uma linha horizontal.

canais de frequências maiores, conforme observado no painel inferior da Figura 6.26, referente ao canal 29. O espectro de potência recuperado está abaixo do espectro do HI de entrada para os três canais, indicando que uma parcela considerável do sinal foi removida pelo GNILC em todos os canais de frequência.

A Figura 6.27 apresenta os espectros de potência angular *naive* e iGLS: a curva

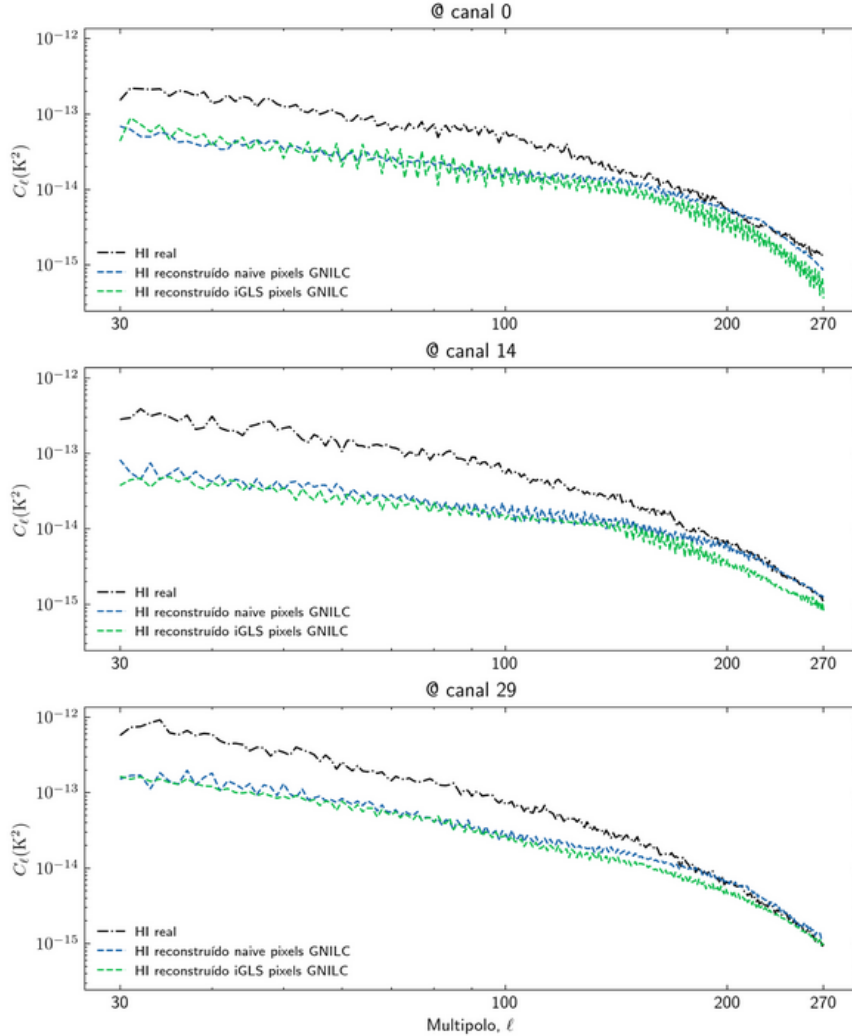
Figura 6.26: Espectros de potência angular para o mapa real de HI (preto) e para o mapa iGLS recuperado (azul), obtidos com *rebeaming* no espaço de pixels e separação de componentes realizada pelo GNILC.



Observa-se as discrepâncias entre ambas as curvas e a instabilidade do espectro do mapa recuperado, especialmente no canal 0 e em altos multipoles.

azul representa o espectro recuperado para o mapa *naive* e a curva verde o espectro recuperado para o mapa iGLS. O espectro angular do mapa de HI de entrada é representado pela linha preta. Em ambos o *rebeaming* foi realizado no espaço de pixels. Para ambos os casos recuperados, os espectros estão distantes do espectro de entrada, mostrando uma baixa recuperação do sinal de HI independente da ferramenta de *map-making* adotada.

Figura 6.27: Comparação entre os espectros de potência angular para o mapa real de HI (preto), para o mapa iGLS recuperado (verde) e para o mapa *naive* recuperado (azul), com *rebeaming* realizado no espaço de pixels e separação de componentes realizada com o GNILC.



Os espectros para ambos os mapas recuperados estão distantes do espectro do HI, embora para altos multipoles o espectro do mapa *naive* aproxime-se do HI real.

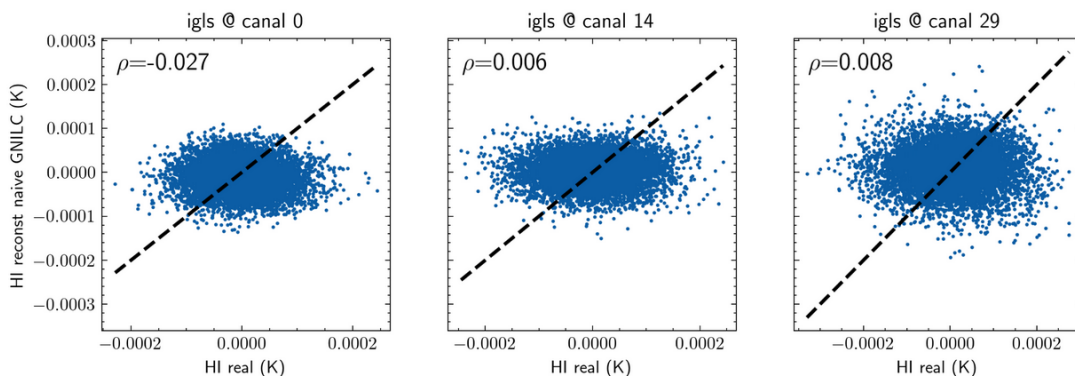
No entanto, as curvas referentes aos mapas *naive* recuperados apresentam menor instabilidade quando comparadas às curvas para os mapas iGLS, o que pode ser observado para a imagem do canal 0, no primeiro painel da Figura 6.27. Estas se aproximam do espectro real em altos multipoles, com bons resultados quando $\ell > 200$, especialmente para o canal 29, no último painel da mesma figura. Informações quantitativas acerca da recuperação do sinal para os diferentes métodos de *map-*

making são apresentadas a seguir a partir de parâmetros estatísticos apropriados.

- **Coefficiente de Pearson (ρ) e erro de reconstrução dos espectros de potência angular (η_{abs})**

Para quantificar e comparar o desempenho das duas ferramentas de *map-making*, calculou-se o coeficiente de Pearson (ρ) para cada uma delas. Como esse coeficiente mede o grau de correlação entre o mapa recuperado e o mapa real, ele fornece uma medida direta da qualidade da recuperação do sinal. A Figura 6.17, exibida anteriormente na Seção 6.4, apresenta o valor de ρ calculado para os mapas *naive* e a Figura 6.28 o valor de ρ para os mapas iGLS, nos canais 0, 14 e 29. Na Figura 6.28, os pontos do gráfico de dispersão se distribuem afastados da diagonal, indicando a baixa correlação entre os dois mapas analisados. Esse fato também é evidenciado pelo baixo valor do coeficiente ρ nos canais analisados.

Figura 6.28: Coeficiente de correlação de Pearson obtido para o mapa iGLS recuperado quando comparado ao mapa real de HI, apresentado nos canais 0, 14 e 29.

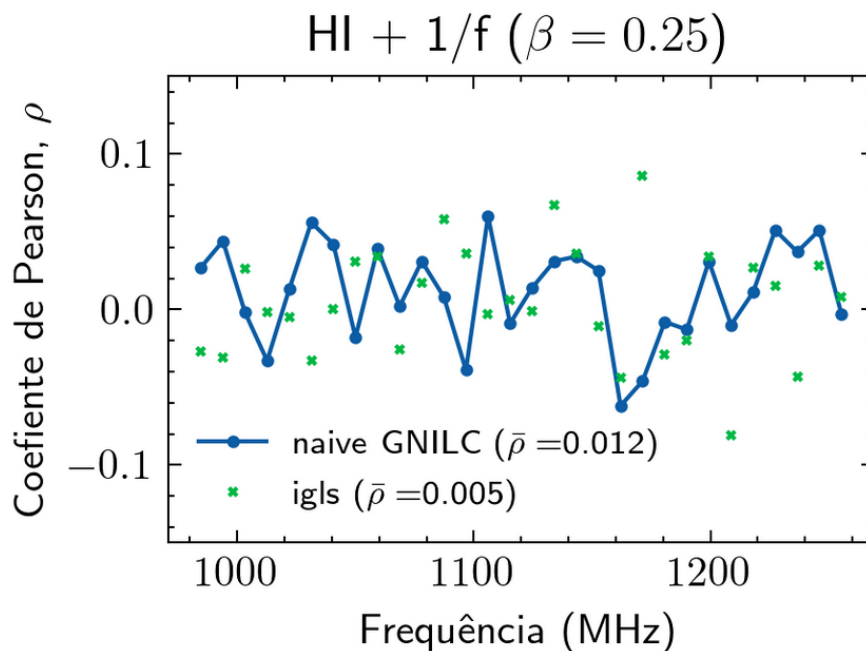


O coeficiente foi calculado ponto a ponto entre os mapas para cada um dos canais. Seu formato ovalado, com alta dispersão e distante da diagonal representada pela linha tracejada indica baixa correlação entre o mapa de entrada e o mapa recuperado.

A Figura 6.29 apresenta uma comparação entre os coeficientes de Pearson (ρ) calculados para os mapas de HI reconstruídos pelo GNILC a partir dos mapas observados produzidos com os métodos *naive* (curva azul) e iGLS (curva verde) de *map-making*, bem como o valor do coeficiente de Pearson médio ($\bar{\rho}$) calculado sobre os mapas de todos os canais de frequência considerados em ambos os casos.

Para os mapas *naive*, obteve-se um coeficiente de Pearson médio $\bar{\rho} = 0,012$, enquanto para os mapas iGLS o valor encontrado foi $\bar{\rho} = 0,005$, ambos muito próximos de zero. Isso indica uma correlação fraca entre os mapas de HI de entrada e os reconstruídos pelo GNILC, e um desempenho insatisfatório na recuperação do sinal de HI. Ainda assim, o valor ligeiramente maior de $\bar{\rho}$ calculado a partir dos mapas *naive* sugere que essa técnica produziu reconstruções um pouco mais próximas do mapa real em comparação ao método iGLS nas simulações do BINGO.

Figura 6.29: Coeficiente de Pearson por canal para o mapa iGLS (pontos verdes em “x”) e para o mapa *naive* (pontos e curva azul), com *rebeaming* no espaço de pixels e separação de componentes realizada pelo GNILC.

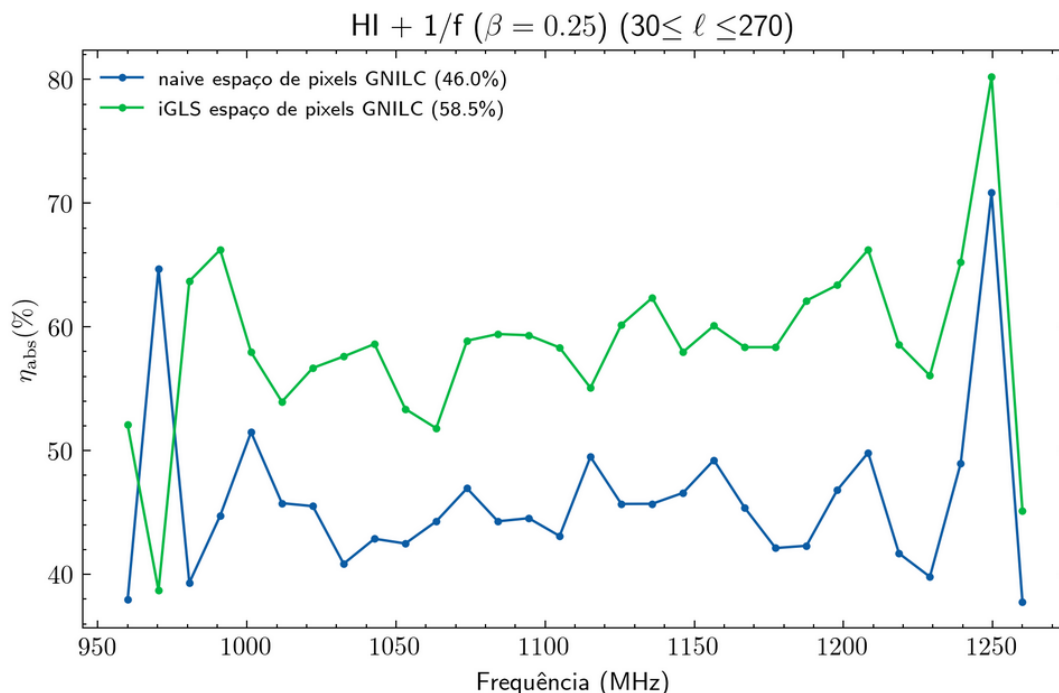


A Figura 6.30 apresenta o erro de reconstrução dos espectros de potência angular dos mapas de HI reconstruídos pelo GNILC a partir de mapas gerados pelos métodos *naive* (curva azul) e iGLS (curva verde), em função dos canais de frequência, no intervalo de multipolos $30 \leq \ell \leq 270$, que corresponde aproximadamente à região angular associada à escala característica de BAO no intervalo de *redshifts* do BINGO.

A curva dos mapas *naive* fica abaixo da curva dos mapas iGLS em todos os canais de frequência exceto no canal 1, indicando de forma consistente um menor erro de reconstrução dos C_ℓ nos mapas *naive*. Para os mapas *naive* processados pelo GNILC, o erro médio calculado sobre todos os canais foi de $N_{\text{abs}} = 46 \%$, enquanto

para os mapas iGLS o valor encontrado foi $N_{\text{abs}} = 58,5 \%$. O erro de reconstrução do espectro de potência angular ($> 40\%$) reforça as limitações enfrentadas por ambas as técnicas de *map-making* ao lidar com o ruído $1/f$ parcialmente correlacionado.

Figura 6.30: Erros de reconstrução dos C_ℓ 's obtidos para cada canal de frequências para o mapa *naive* (azul) e para o mapa iGLS (verde) quando comparados ao mapa de HI real.



O erro médio, calculado em todos os canais, associado ao mapa *naive* é dado por $\eta_{\text{abs}} = 46 \%$ e o erro médio para o mapa iGLS é dado por $\eta_{\text{abs}} = 58,5 \%$.

Os espectros de potência angular produzidos a partir dos mapas de intensidade de HI reconstruídos pelo GNILC, foram analisados neste capítulo, juntamente com os erros de reconstrução e os coeficientes de correlação de Pearson. Esses resultados indicam que a ferramenta de *map-making* que proporcionou a melhor recuperação do sinal cosmológico de HI nos cenários propostos foi o método *naive*.

Mesmo que seus resultados não tenham sido satisfatórios, pois seus mapas apresentaram baixa correlação com o mapa de HI real, além de altos erros de reconstrução do espectro de potência do sinal, seu erro de reconstrução do C_ℓ foi menor e seu coeficiente de correlação foi ligeiramente maior quando comparado ao resultado obtido com o ROMA. Portanto, a aplicação das técnicas de pré-processamento do ROMA,

como o pré-condicionamento e a análise estatística da distribuição espectral de potência do ruído, não foi adequada para lidar com a contaminação do ruído $1/f$ nas simulações de missão do BINGO produzidas neste trabalho.

O ROMA foi desenvolvido originalmente como um algoritmo de processamento de dados da CMB, sendo aplicado com sucesso em experimentos como BOOMERanG (BERNARDIS et al., 2000) e Planck (TAUBER et al., 2004). Portanto, sua estrutura foi pensada levando em conta as características instrumentais desses experimentos. O BOOMERanG, um instrumento lançado a bordo de balão estratosférico, mediu as anisotropias da CMB mapeando o céu por meio de varreduras com velocidade constante (BERNARDIS et al., 2000). O satélite Planck, por sua vez, realizava um movimento que permitia observar a mesma região do céu várias vezes sob diferentes ângulos ou direções de varredura, aumentando a redundância das medições e permitindo separar melhor o sinal cosmológico do ruído instrumental, numa técnica conhecida como *cross-linking* (TAUBER et al., 2004).

Com a aplicação desta técnica diminui-se o efeito das faixas na direção de observação (*striping*) ao mudar a maneira como o ruído se projeta no mapa. Diversos radiotelescópios mitigam o efeito do ruído $1/f$ já na fase de observação, movimentando os detectores ou algum elemento do sistema óptico (WANG et al., 2021). Ao varrer o céu em intervalos de tempo menores do que o tempo de estabilidade do sistema (período durante o qual os amplificadores mantêm ganho aproximadamente constante) é possível reduzir a contaminação causada pela correlação temporal do ruído $1/f$ (HARPER et al., 2018).

O BINGO, no entanto, é um telescópio de trânsito, no qual os detectores permanecem fixos e o céu passa sobre eles (*drift scan*), não havendo varredura do céu, no sentido descrito acima. Essa característica limita a redundância angular das observações e contribui para o desempenho insatisfatório do ROMA como ferramenta de *map-making* neste contexto. Dessa forma, o ruído $1/f$ não pode ser tratado na fase de aquisição de dados e seu efeito precisa ser mitigado diretamente nos mapas, através de ferramentas alternativas de *map-making* ou de separação de componentes (BIGOT-SAZY et al., 2015). Para ruídos com alta correlação espectral ($\beta \leq 0.25$), as ferramentas testadas não mostraram eficácia suficiente, comprometendo a recuperação do sinal de HI (HARPER et al., 2018).

As simulações realizadas neste estudo demonstraram uma recuperação ineficaz, que falhou em recompor o sinal cosmológico ao não conseguir mitigar a contaminação diretamente nos mapas de intensidade. Isso sublinha a necessidade de investigar

outras técnicas e ferramentas capazes de atenuar os efeitos do ruído $1/f$, possibilitando, assim, o uso dos dados obtidos com o radiotelescópio BINGO em análises cosmológicas mais sofisticadas.

6.6 Testes complementares

Foram realizados alguns testes adicionais buscando alternativas para a recuperação do sinal cósmico de HI, entre elas alterações na resolução espacial dos mapas e no ajuste do feixe entre os canais de frequência. Esses testes avaliaram como as etapas de processamento influenciam o mapa reconstruído.

O ajuste do feixe entre os canais espectrais pode ser realizado tanto no espaço dos *pixels* quanto no espaço dos harmônicos esféricos. A escolha do domínio em que essas operações são realizadas pode impactar na recuperação do sinal, já que ao realizar a transição dos dados entre tais espaços pode haver perda ou diluição de informação. Complementarmente, a resolução espacial na qual os mapas são gerados influencia a maneira como a informação de intensidade é observada entre as suas regiões. Em um mapa de maior resolução, pode-se obter informações mais detalhadas da intensidade do sinal em pequenas escalas angulares, enquanto em mapas com resoluções menores, o sinal aparece mais difuso devido ao efeito do feixe do instrumento. Portanto, a resolução do mapa impacta diretamente a forma como o sinal de interesse é obtido e, conseqüentemente, a análise posterior dos dados.

Buscando analisar o efeito desses elementos na recuperação do sinal de HI, foram realizados dois testes adicionais, considerando os processos acima citados. O primeiro consistiu na realização do *rebeamng* no espaço harmônico, em contraste aos mapas anteriores, em que este era realizado no espaço de *pixels*. O segundo envolveu a simulação dos mapas com uma resolução maior, $N_{\text{side}} = 512$, em comparação com os mapas anteriores, que possuíam $N_{\text{side}} = 128$. Observou-se uma melhora nos resultados dos mapas produzidos com o ROMA, embora o resultado global não tenha sido significativamente alterado.

- **Resultados do ajuste de feixe em diferentes domínios e resoluções espaciais dos mapas**

A seguir, são apresentados os resultados dos testes adicionais, realizados ao alterar o domínio para a realização do ajuste do feixe e ao aumentar a resolução espacial dos mapas simulados. Para tentar melhorar a recuperação do sinal de HI, optou-se por gerar mapas simulados com uma resolução espacial maior, equivalente a $N_{\text{side}} = 512$.

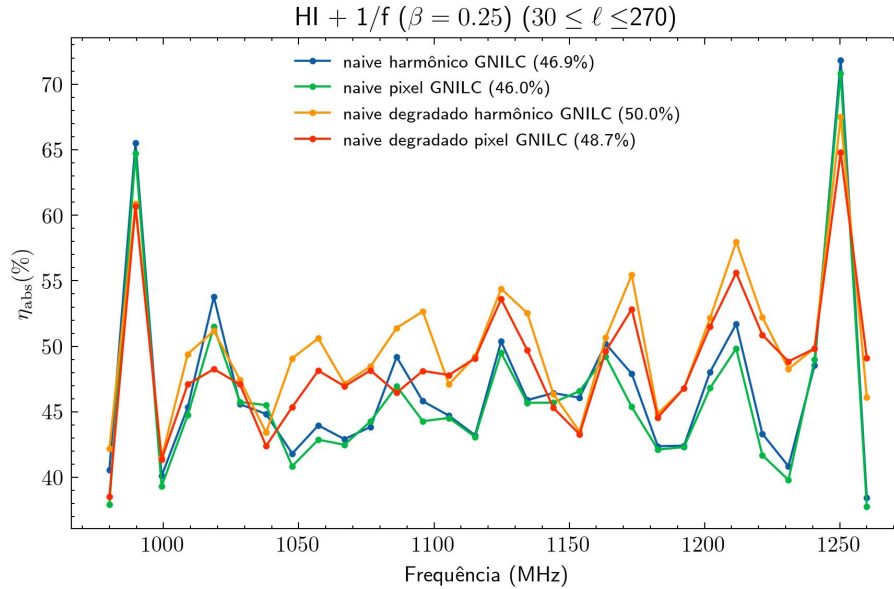
Mesmo utilizando o arranjo compacto de cornetas, esses mapas apresentaram falhas em isolatitude devido à disposição das cornetas no plano focal, conforme abordado nas Seções 5.2 e 6.1. Para evitar que essas falhas comprometessem a separação de componentes, os mapas de observação foram então degradados para uma resolução de $N_{\text{side}} = 128$ ainda no pré-processamento, logo após a remoção do nível DC. Desta forma, furtou-se da presença das falhas e foi possível minimizar o efeito das mesmas e reduzir os erros associados ao processo de pixelização (DORÉ et al., 2001).

Na Figura 6.31 são apresentados os erros absolutos de reconstrução dos espectros de potência angular do HI para os mapas *naive*, correspondentes às quatro simulações listadas na Tabela 6.1. Em todas as simulações foram utilizados os mesmos parâmetros espectrais do ruído, $f_k = 0,001$ Hz e $\beta = 0,25$, diferindo apenas pelo valor do N_{side} adotado e pelo domínio em que foi realizado o ajuste do feixe. As curvas denominadas com o termo “degradado” correspondem às simulações geradas com $N_{\text{side}} = 512$ e posteriormente degradadas para $N_{\text{side}} = 128$, logo após a remoção do nível DC, enquanto as demais foram geradas diretamente com $N_{\text{side}} = 128$. Já os termos “harmônico” e “pixel”, referem-se ao domínio no qual foi realizado o ajuste do feixe. Especificamente, a curva azul refere-se ao ajuste de feixe no espaço harmônico com $N_{\text{side}} = 128$ (erro absoluto médio $\bar{\eta}_{\text{abs}} = 46,9\%$), a curva verde ao ajuste no espaço de *pixels* com $N_{\text{side}} = 128$ (erro absoluto médio $\bar{\eta}_{\text{abs}} = 46,0\%$), a curva laranja ao ajuste no espaço harmônico com $N_{\text{side}} = 512$ degradado para $N_{\text{side}} = 128$ (erro absoluto médio $\bar{\eta}_{\text{abs}} = 50,0\%$), e a curva vermelha ao ajuste no espaço de *pixels* com $N_{\text{side}} = 512$ degradado para $N_{\text{side}} = 128$ (erro absoluto médio $\bar{\eta}_{\text{abs}} = 48,7\%$).

A curva verde representa a Simulação 1, realizada com método *naive* de *map-making*, com $f_k = 0,001$ Hz, $\beta = 0,25$, gerada com $N_{\text{side}} = 128$ e com tempo de observação igual a 3 dias, abordada anteriormente neste trabalho (Figura 6.30). Entre os mapas *naive*, o melhor resultado para a reconstrução do espectro de potência, isto é, o menor erro relativo associado ao C_ℓ , foi obtido quando o ajuste de feixe foi realizado no espaço de *pixels* com $N_{\text{side}} = 128$. Nesse caso, o erro médio absoluto calculado em todos os canais foi de $\eta_{\text{abs}} = 46,0\%$ para o mapa reconstruído de HI.

Em contrapartida, para os mapas iGLS, obteve-se uma melhora na recuperação ao alterar os métodos de processamento. A Figura 6.32 contém as curvas para o erro de recuperação dos C_ℓ 's dos mapas produzidos com o ROMA. De forma análoga à Figura 6.31, o termo “degradado” refere-se aos mapas produzidos com $N_{\text{side}} = 512$ e degradados para $N_{\text{side}} = 128$, enquanto os termos “harmônico” e “*pixel*” também referem-se ao domínio em que foi realizado o ajuste de feixe. Foram analisadas as

Figura 6.31: Erros de reconstrução do espectro de potência do HI para os mapas *naive* com tipos distintos de pré-processamento.

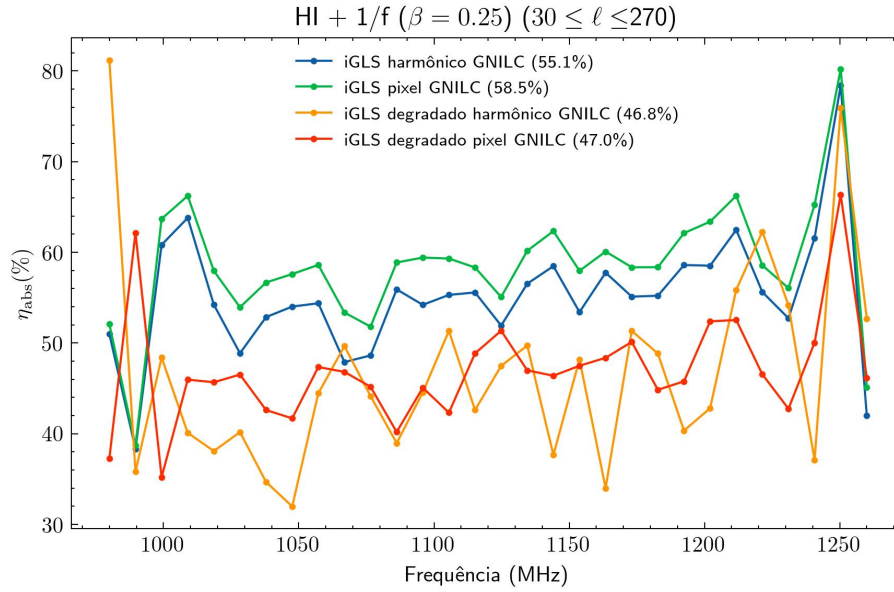


Mapa gerado com $N_{\text{side}} = 128$ e *rebeaming* no espaço harmônico (azul); mapa gerado com $N_{\text{side}} = 128$ e *rebeaming* no espaço de *pixels* (verde); mapa gerado com $N_{\text{side}} = 512$ e degradado para $N_{\text{side}} = 128$ com *rebeaming* no espaço harmônico (laranja) e mapa gerado com $N_{\text{side}} = 512$ e degradado para $N_{\text{side}} = 128$ com *rebeaming* no espaço de *pixels* (vermelho).

mesmas simulações consideradas para os mapas *naive*. A curva verde representa a simulação de referência (mapa de observação gerado com o método iGLS, ajuste de feixe no espaço de *pixels*, e mapa de HI reconstruído pelo GNILC) conforme representado na Figura 6.30. As demais curvas de erro referem-se aos testes adicionais realizados.

Observa-se que de fato houve uma melhora nos resultados após as adaptações na análise. O erro dos mapas iGLS caiu para todos os casos analisados, atingindo o valor mínimo de 46,8% quando o ajuste de feixe foi realizado no espaço harmônico e o mapa teve sua resolução degradada. Neste caso, o seu resultado tornou-se competitivo com o resultado para o caso *naive* com *rebeaming* no espaço de *pixels* e N_{side} original, com uma diferença de apenas 0,8 % no erro. A resolução do mapa parece ter sido a principal responsável pela melhora dos resultados, tendo sido obtida uma diminuição de 11,5 % e 8,3 % nos erros ao fixar o domínio do ajuste de feixe e alterar a resolução de referência dos mapas. Ou seja, quando compara-se a curva verde com a curva vermelha e a curva azul com a laranja, respectivamente. Embora

Figura 6.32: Erros de recuperação dos espectros de potência angular do HI para os mapas iGLS com tipos distintos de pré-processamento.



Mapa gerado com $N_{\text{side}} = 128$ e *rebeaming* no espaço harmônico (azul); mapa gerado com $N_{\text{side}} = 128$ e *rebeaming* no espaço de *pixels* (verde); mapa gerado com $N_{\text{side}} = 512$ e degradado para $N_{\text{side}} = 128$ com *rebeaming* no espaço harmônico (laranja) e mapa gerado com $N_{\text{side}} = 512$ e degradado para $N_{\text{side}} = 128$ com *rebeaming* no espaço de *pixels* (vermelho). Neste caso, a melhor recuperação ocorreu para a simulação com o ajuste de feixe no espaço harmônico e com resolução degradada, representa pela curva amarela cujo erro médio é igual a 46,8%.

globalmente, ao comparar todos os mapas, o melhor resultado tenha sido obtido com o método *naive* de *map-making*, as adaptações melhoraram os mapas iGLS, trazendo uma redução de 11,7% do erro de recuperação. Mesmo que os resultados continuem não sendo satisfatórios para a recuperação do sinal, a melhora indica que feitos os devidos ajustes durante o processamento, o ROMA pode competir com o *map-making* da IM *pipeline*.

7 CONCLUSÕES

Considerando os resultados apresentados ao longo do Capítulo 6, e levando em conta as ferramentas comparativas discutidas na Seção 5.6, conclui-se que, no âmbito das técnicas de *map-making*, a aplicação do método *naive* com a IM *pipeline* apresentou o melhor desempenho na geração dos mapas de intensidade nos cenários propostos neste trabalho. Já entre os métodos de separação de componentes testados, o GNILC mostrou-se mais eficiente na recuperação do sinal de HI.

Quando comparado à melhor configuração do PCA, obtida com a remoção de 8 a 9 modos principais, o GNILC apresentou um erro de recuperação dos espectros de potência consideravelmente menor ($\eta_{\text{abs}} = 46,0\%$) quando comparado ao PCA ($\eta_{\text{abs}} = 60,9\%$). Este resultado indica que o GNILC proporcionou uma separação de componentes mais eficaz, produzindo mapas do sinal de 21 cm de HI mais próximos do mapa teórico. Sua boa performance está associada à decomposição dos mapas no espaço harmônico, o que pode melhorar os resultados, especialmente em pequenas escalas. Devido a este bom desempenho, o GNILC foi adotado como a ferramenta de separação de componentes utilizada ao longo deste trabalho.

Quanto às ferramentas de *map-making*, após a separação de componentes, os mapas com menor erro de recuperação do espectro de potência foram aqueles produzidos pela IM *pipeline* (método *naive*). Além disso, esses mapas também apresentaram um coeficiente de Pearson ligeiramente maior e uma menor dispersão em relação ao mapa de HI teórico, bem como seus espectros de potência apresentaram um comportamento mais estável, com uma curva menos oscilante.

Este resultado não reflete uma superioridade natural do método *naive*, que em cenários adequados seria superado pelo iGLS. O iGLS depende da redundância fornecida pelo *cross-linking* para separar de forma eficaz o sinal de HI do ruído $1/f$. No BINGO, cada *pixel* é amostrado apenas em uma direção, determinada pela posição fixa dos detectores, o que limita a capacidade do iGLS de modelar corretamente o ruído $1/f$ no tempo e sua parcela não correlacionada em frequência.

A partir destes resultados, conclui-se que o ROMA não foi capaz de mitigar de forma eficaz os efeitos do ruído $1/f$ nas simulações do BINGO analisadas neste trabalho. Esse algoritmo foi originalmente desenvolvido para o tratamento de dados de CMB, obtidos por instrumentos com características técnicas e estratégias observacionais distintas das do BINGO. Ele foi aplicado inicialmente em dados obtidos pelo BOOMERanG (LANGE et al., 1995) e pelo Planck (TAUBER et al., 2004;

NATOLI et al., 2001), ambos os instrumentos que lidavam com a presença do ruído correlacionado através de suas estratégias de varredura, mudando a direção do mapeamento em escalas de tempo menores do que as variações do ruído. No caso do BINGO, por se tratar de um instrumento de trânsito, essa estratégia não é viável. Assim, as diferenças estruturais entre os instrumentos provavelmente contribuíram para o desempenho menos satisfatório do ROMA quando aplicado às simulações do BINGO.

Além disso, o método iGLS aplicado pelo ROMA depende da estrutura da matriz de covariância do ruído. Como o método utiliza um pré-condicionador para aproximar a inversa de \mathbf{C}_n , ele se beneficia do fato de que, para ruído quase branco, \mathbf{C}_n é aproximadamente diagonal e esparsa, permitindo usar uma matriz diagonal positiva como pré-condicionador. No entanto, na presença de ruído $1/f$, \mathbf{C}_n apresenta elementos fortemente correlacionados, tornando-se densa (maioria dos elementos diferentes de zero) e não diagonal dominante (valor absoluto do elemento da diagonal é menor que a soma dos valores absolutos dos outros elementos da linha). Essa complexidade dificulta a eficácia do pré-condicionador e a convergência do método iGLS, o que explicaria os resultados insatisfatórios obtidos para os mapas.

No entanto, após as adaptações realizadas no N_{side} e no domínio de realização do *rebeaming*, houve uma melhora de 11,7% no erro de recuperação do espectro de potência dos mapas gerados com o ROMA. Com essas adequações, o algoritmo se mostrou competitivo em relação ao método *naive*, o que possibilita seu uso futuro, embora o erro de recuperação ainda seja alto e revisões futuras sejam necessárias.

Os elevados erros de recuperação obtidos pelo GNILC evidenciam a dificuldade do método em lidar com ruído fortemente correlacionado, mesmo na ausência de *foregrounds* astrofísicos. No caso estudado, com apenas o sinal cosmológico de HI e ruído $1/f$, a matriz de covariância do ruído não é diagonal nem esparsa, apresentando correlações significativas entre diferentes canais e *pixels*. O GNILC estima localmente a matriz de covariância dos dados para separar o sinal de interesse do ruído dominante, assumindo que as componentes podem ser diferenciadas com base em suas variações espaciais e espectrais. Porém, a estrutura suave e correlacionada do ruído $1/f$ aumenta a variância estimada nos modos principais da análise, fazendo com que o algoritmo projete parte do sinal de HI sobre o subespaço do ruído e remova mais modos do que o necessário. Esse efeito resulta em uma supressão parcial do sinal cosmológico, explicando os elevados erros de recuperação observados e evidenciando a limitação do GNILC para lidar com ruído fortemente correlacionado, mesmo em

cenários sem *foregrounds* astrofísicos.

Conforme discutido por Harper et al. (2018), quando o ruído $1/f$ apresenta baixa correlação, ele se comporta de forma semelhante ao ruído branco, variando rapidamente ao longo dos canais de frequência. Nessas condições, o ruído não domina nenhum modo específico nem é identificado como *foreground*, e também não pode ser completamente isolado como ruído térmico. Como resultado, ele compete diretamente com o sinal cosmológico de HI, projetando-se nos mesmos modos de baixa potência e dificultando a sua recuperação.

Em outras palavras, independente do nível de correlação, o ruído $1/f$ representa um desafio complexo para a recuperação do sinal cosmológico no mapeamento de intensidade, impedindo que o sinal de HI seja isolado e tornando impossível recuperá-lo sem que o ruído seja tratado em etapas anteriores. Este é um resultado importante desta dissertação, raramente detalhado em trabalhos anteriores. A dificuldade é evidenciada pelos baixos valores do coeficiente de correlação de Pearson, pela acentuada dispersão dos dados e pelos elevados erros de recuperação do espectro de potência angular em todas as simulações. É imprescindível, portanto, desenvolver métodos eficazes de mitigação do ruído $1/f$ para possibilitar a recuperação confiável do sinal de HI.

Uma vez que não foi possível remover o efeito do ruído $1/f$ na etapa de *map-making* através das ferramentas utilizadas, é necessário que busquem-se outras estratégias de mitigação do ruído, idealmente antes da etapa de produção de mapas, através de otimização do hardware (eletrônica) para que o radiotelescópio BINGO alcance seu objetivo científico. Usualmente o ruído $1/f$ é tratado ainda na fase de aquisição dos dados, através da movimentação dos detectores do radiotelescópio. Como o BINGO foi projetado para ser um instrumento fixo, isto não é possível. Restam, portanto, a utilização de técnicas de calibração precisas e a implementação de sistemas que sejam capazes de controlar o ganho dos amplificadores. O sistema de amplificadores do instrumento foi previamente testado, obtendo valores equivalente a $\beta \approx 0,25$ (BIGOT-SAZY et al., 2015; WUENSCHÉ et al., 2022) e as estratégias de calibração mencionadas em Wuensche et al. (2022) ainda não foram testadas. A utilização de correlacionadores, que corrigem a variação de ganho dos amplificadores, pode auxiliar na mitigação do ruído $1/f$, mas sua efetividade depende de uma boa calibração do instrumento. Espera-se que com várias estratégias de calibração combinadas seja possível recuperar o sinal de 21 cm do HI (LICCARDO et al., 2022).

Como perspectivas para trabalhos futuros, outras ferramentas de *map-making*, que

utilizem métodos de resolução alternativos podem ser exploradas, especialmente aquelas capazes de modelar de forma mais precisa a matriz de covariância do ruído, visando melhorar a recuperação do sinal de HI. Ferramentas de separação de componentes mais sofisticadas, como métodos paramétricos (assumem conhecimento prévio sobre as componentes que se deseja remover), também podem ser testadas, com o objetivo de reduzir a supressão do sinal cosmológico causada por ruído $1/f$. Além disso, aprimorar a IM *pipeline*, utilizada para fazer simulações de experimentos de mapeamento de intensidade, como é o caso do BINGO, e implementar um módulo de calibração do instrumento permitiria caracterizar de maneira mais precisa os sinais medidos, aumentando a eficácia das etapas de *map-making* e separação de componentes.

Por fim, este trabalho destacou-se como um dos primeiros da colaboração BINGO em caracterizar e quantificar o efeito do ruído $1/f$ na recuperação do sinal de 21 cm do HI. Os resultados demonstram a importância crítica de encontrar estratégias de mitigação dos ruídos instrumentais, permitindo que o instrumento alcance seus objetivos científicos. O ruído correlacionado apresenta-se ainda como um problema desafiador a ser solucionado, seja através da estratégia de observação, da calibração, na etapa de produção de mapas ou na separação de componentes. A IM *pipeline* demonstrou melhores resultados na criação de mapas de intensidade com o método *naive*, enquanto o GNILC se destacou pela sua maior capacidade de recuperação do sinal de HI, especialmente em pequenas escalas angulares. A dificuldade do ROMA em lidar com esse tipo de ruído nas observações do BINGO reflete as limitações do algoritmo em tal cenário e da estratégia de observação do instrumento, o que exige ajustes futuros. Diante desses desafios, torna-se claro que a aplicação e o desenvolvimento de técnicas para a mitigação do ruído $1/f$ são essenciais para a recuperação com boa relação sinal-ruído do sinal cosmológico e para o estudo da estrutura em larga escala no Universo.

REFERÊNCIAS BIBLIOGRÁFICAS

ABDALLA, E.; FERREIRA, E. G. M.; LANDIM, R. G.; COSTA, A. A.; FORNAZIER, K. S. F.; ABDALLA, F. B.; BAROSI, L.; BRITO, F. A.; QUEIROZ, A. R.; VILLELA, T.; WANG, B.; WUENSCHÉ, C. A.; MARINS, A.; NOVAES, C. P.; LICCARDO, V.; SHAN, C.; ZHANG, J.; ZHANG, Z.; ZHU, Z.; BROWNE, I.; DELABROUILLE, J.; SANTOS, L.; SANTOS, M. V. dos; XU, H.; ANTON, S.; BATTYE, R.; CHEN, T.; DICKINSON, C.; MA, Y.-Z.; EI, B. M.; MERICIA, E. J. de; MOTTA, P.; OTOBONE, C. H. N.; PEEL, M. W.; ROYCHOWDHURY, S.; REMAZEILLES, M.; RIBEIRO, R. M.; SANG, Y.; SANTOS, J. R. L.; SANTOS, J. F. R. dos; SILVA, G. B.; VIEIRA, F.; VIEIRA, J.; XIAO, L.; ZHANG, X.; ZHU, Y. The bingo project-i. baryon acoustic oscillations from integrated neutral gas observations. **Astronomy & Astrophysics**, v. 664, p. A14, 2022. 12, 14, 25, 26, 54, 55, 63

ABDALLA, F. B.; MARINS, A.; MOTTA, P.; ABDALLA, E.; RIBEIRO, R. M.; WUENSCHÉ, C. A.; DELABROUILLE, J.; FORNAZIER, K. S. F.; LICCARDO, V.; MAFFEI, B.; MERICIA, E. J. de; OTOBONE, C. H. N.; SANTOS, J. F. R. dos; SILVA, G. B.; VIEIRA, J.; BARRETOS, J. A. M.; BAROSI, L.; BRITO, F. A.; QUEIROZ, A. R.; VILLELA, T.; WANG, B.; COSTA, A. A.; FERREIRA, E. G. M.; LANDIM, R. G.; NOVAES, C. P.; PEEL, M. W.; SANTOS, L.; SANTOS, M. V. dos; ZHANG, J. The bingo project-iii. optical design and optimization of the focal plane. **Astronomy & Astrophysics**, v. 664, p. A16, 2022. 2, 26, 27

AGHANIM, N.; AKRAMI, F. A.; ASHDOWN, M.; AUMONT, J.; BACCIGALUPI, C.; BALLARDINI, M.; BANDAY, A.; BARREIRO, R.; BARTOLO, N. Planck results. **Astronomy and Astrophys.-A &A**, v. 641, p. 2020, 2018. 6, 8

ALONSO, D.; SANCHEZ, J.; SLOSAR, A.; COLLABORATION, L. D. E. S. A unified pseudo- ℓ framework. **Monthly Notices of the Royal Astronomical Society**, v. 484, n. 3, p. 4127–4151, 2019. 71

ASSIS, A. K.; NEVES, M. C.; SOARES, D. S. A cosmologia de hubble: de um universo finito em expansão a um universo infinito no espaço e no tempo. In: NEVES, M. (Ed.). **Evoluções e revoluções: o mundo em transição**. [S.l.]: Massoni; LCV, 2008. p. 199–221. 5

AXELSSON, O. A generalized conjugate gradient, least square method. **Numerische Mathematik**, v. 51, n. 2, p. 209–227, 1987. 57

BAHCALL, N. A. Hubble's law and the expanding universe. **Proceedings of the National Academy of Sciences**, v. 112, n. 11, p. 3173–3175, 2015. 5

BANDURA, K. et al. Canadian hydrogen intensity mapping experiment (chime) pathfinder. In: SPIE. **Ground-based and Airborne Telescopes V**. [S.l.], 2014. v. 9145, p. 738–757. 24

BASSETT, B.; HLOZEK, R. Baryon acoustic oscillations. In: RUIZ-LAPUENTE, P. (Ed.). **Dark energy: observational and theoretical approaches**. Cambridge: Cambridge University Press, 2010. p. 246. 9

BATTYE, R.; BROWNE, I.; DICKINSON, C.; HERON, G.; MAFFEI, B.; POURTSIDOU, A. High intensity mapping: a single dish approach. **Monthly Notices of the Royal Astronomical Society**, v. 434, n. 2, p. 1239–1256, 2013. 18, 26, 54

BENNET, C. L.; HALPERN, M.; HINSHAW, G.; JAROSIK, N.; KOGUT, A.; LIMON, M.; MEYER, S. S.; PAGE, L.; SPERGEL, D. N.; TUCKER, G. S.; WOLLACK, E.; WRIGTH, E. L.; BARNES, C.; GREASON, M. R.; HILL, R. S.; KOMATSU, E.; NOLTA, M. R.; ODEGARD, N.; PEIRIS, H. V.; VERDE, L.; WEILAND, J. L. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps And Basic Results. **The Astrophysical Journal Supplement Series**, v. 148, n. 1, p. 1–27, set. 2003. 71

BERNARDIS, P. de et al. A flat universe from high-resolution maps of the cosmic microwave background radiation. **Nature**, v. 404, n. 6781, p. 955–959, 2000. 121

BIGOT-SAZY, M.-A.; DICKINSON, C.; BATTYE, R. A.; BROWNE, I.; MA, Y.-Z.; MAFFEI, B.; NOVIELLO, F.; REMAZEILLES, M.; WILKINSON, P. Simulations for single-dish intensity mapping experiments. **Monthly Notices of the Royal Astronomical Society**, v. 454, n. 3, p. 3240–3253, 2015. 18, 26, 44, 54, 60, 83, 121, 129

BREYSSE, P. C.; KOVETZ, E. D.; KAMIONKOWSKI, M. Carbon monoxide intensity mapping at moderate redshifts. **Monthly Notices of the Royal Astronomical Society**, v. 443, n. 4, p. 3506–3512, 2014. 13

BUCKHOLTZ, T. J. Characterizations that help explain particle and cosmic data. **Journal of High Energy Physics, Gravitation and Cosmology**, v. 10, n. 3, p. 1304–1357, 2024. 12

BULL, P.; FERREIRA, P. G.; PATEL, P.; SANTOS, M. G. Late-time cosmology with 21 cm intensity mapping experiments. **The Astrophysical Journal**, v. 803, n. 1, p. 21, 2015. 12, 14, 18

BURKE, B. F.; GRAHAM-SMITH, F.; WILKINSON, P. N. **An introduction to radio astronomy**. [S.l.]: Cambridge University Press, 2019. 30, 33

CANTALUPO, C.; BORRILL, J.; JAFFE, A.; KISNER, T.; STOMPOR, R. Madmap: a massively parallel maximum likelihood cosmic microwave background map-maker. **The Astrophysical Journal Supplement Series**, v. 187, n. 1, p. 212, 2010. 46, 54

CLEARY, K. A. et al. Comap early science. i. overview. **The Astrophysical Journal**, v. 933, n. 2, p. 182, 2022. 13

COHEN-TANNOUJDI, C.; DIU, B.; LALOE, F. Quantum mechanics, volume 2. **Quantum Mechanics**, v. 2, p. 626, 1986. 14, 15

COLE, S. et al. The 2df galaxy redshift survey: power-spectrum analysis of the final data set and cosmological implications. **Monthly Notices of the Royal Astronomical Society**, v. 362, n. 2, p. 505–534, 2005. 10

DE MERICIA, Eduardo J. et al. Testing synchrotron models and frequency resolution in bingo 21 cm simulated maps using gnild. **Astronomy & Astrophysics**, v. 671, p. A58, 2023. Disponível em: <<<https://doi.org/10.1051/0004-6361/202243804>>>. 26, 55, 56, 58, 59, 60, 62, 66, 67, 68, 80, 97

DEWDNEY, P. E.; HALL, P. J.; SCHILIZZI, R. T.; LAZIO, T. J. L. The square kilometre array. **Proceedings of the IEEE**, v. 97, n. 8, p. 1482–1496, 2009. 1, 24

DODELSON, S.; SCHMIDT, F. **Modern cosmology**. [S.l.]: Academic Press, 2020. 5, 6, 7, 14, 78

DORÉ, O.; TEYSSIER, R.; BOUCHET, F.; VIBERT, D.; PRUNET, S. Mapcumba: a fast iterative multi-grid map-making algorithm for cmb experiments. **Astronomy & Astrophysics**, v. 374, n. 1, p. 358–370, 2001. 66, 123

EISENSTEIN, D. J. et al. Detection of the baryon acoustic peak in the large-scale correlation function of sdss luminous red galaxies. **The Astrophysical Journal**, v. 633, n. 2, p. 560, 2005. 10, 11

EISENSTEIN, D. J.; SEO, H.; SIRKO, E.; SPERGEL, D. N. Improving cosmological distance measurements by reconstruction of the baryon acoustic peak. **The Astrophysical Journal**, v. 664, n. 2, p. 675–679, ago. 2007. ISSN 1538-4357. Disponível em: <<<http://dx.doi.org/10.1086/518712>>>. 9, 10

FIXSEN, D. The temperature of the cosmic microwave background. **The Astrophysical Journal**, v. 707, n. 2, p. 916, 2009. 20

FRIEMAN, J. A.; TURNER, M. S.; HUTERER, D. Dark energy and the accelerating universe. **The Annual Review of Astronomy and Astrophysics**, v. 46, n. 1, p. 385–432, 2008. 5

FURLANETTO, S. R.; OH, S. P.; BRIGGS, F. H. Cosmology at low frequencies: the 21 cm transition and the high-redshift universe. **Physics Reports**, v. 433, n. 4-6, p. 181–301, 2006. 1, 16, 17

GASPERIS, G. D.; BALBI, A.; CABELLA, P.; NATOLI, P.; VITTORIO, N. Roma: a map-making algorithm for polarised cmb data sets. **Astronomy & Astrophysics**, v. 436, n. 3, p. 1159–1165, 2005. 46, 48, 56, 57, 58, 84

GORSKI, K. M.; HIVON, E.; BANDAY, A. J.; WANDEL, B. D.; HANSEN, F. K.; REINECKE, M.; BARTELMANN, M. Healpix: a framework for high-resolution discretization and fast analysis of data distributed on the sphere. **The Astrophysical Journal**, v. 622, n. 2, p. 759, 2005. 41, 66

HAARLEM, M. P. van et al. Lofar: the low-frequency array. **Astronomy & astrophysics**, v. 556, p. A2, 2013. 24

HARPER, S.; DICKINSON, C.; BATTYE, R.; ROYCHOWDHURY, S.; BROWNE, I.; MA, Y.-Z.; OLIVARI, L.; CHEN, T. Impact of simulated 1/f noise for hi intensity mapping experiments. **Monthly Notices of the Royal Astronomical Society**, v. 478, n. 2, p. 2416–2437, 2018. 35, 36, 37, 65, 83, 85, 90, 121, 129

HUBBLE, E. A relation between distance and radial velocity among extra-galactic nebulae. **Proceedings of the National Academy of Sciences**, v. 15, n. 3, p. 168–173, 1929. 5

JANSSEN, M. A.; GULKIS, S.; BENNETT, C. L.; KOGUT, A. J. Design and results of differential microwave radiometers (dmr) on coBE. In: SPIE. **Infrared spaceborne remote sensing**. [S.l.], 1993. v. 2019, p. 211–221. 48

KASSIM, N.; PERLEY, R.; ERICKSON, W.; DWARAKANATH, K. Subarcminute resolution imaging of radio sources at 74 mhz with the very large array. **Astronomical Journal**, v. 106, p. 2218–2228, 1993. 24

KIM, A.; PADMANABHAN, N.; ALDERING, G.; ALLEN, S.; BALTAY, C.; CAHN, R.; D'ANDREA, C.; DALAL, N.; DAWSON, K.; DENNEY, K.; EISENSTEIN, D.; FINLEY, D.; FREEDMAN, W.; HO, S.; HOLZ, D.; KENT, A.; KASEN, D.; KESSLER, R.; KUHLMANN, S.; LINDER, E.; MARTINI, P.; NUGENT, P.; PERLMUTTER, S.; PETERSON, B.; RIESS, A.; RUBIN, D.; SAKO, M.; SUNTZEFF, N.; SUZUKI, N.; THOMAS, R.; WOOD-VASEY, W. M.; WOOSLEY, S. **Distance probes of dark energy**. 2014. Disponível em: <<<https://arxiv.org/abs/1309.5382>>>. 9

KOVETZ, E. D.; VIERO, M. P.; LIDZ, A.; NEWBURGH, L.; RAHMAN, M.; SWITZER, E.; KAMIONKOWSKI, M.; AGUIRRE, J.; ALVAREZ, M.; BOCK, J. et al. Line-intensity mapping: 2017 status report. **arXiv preprint arXiv:1709.09066**, 2017. 12, 14

LANGE, A. et al. The boomerang experiment. **Space Science Reviews**, v. 74, p. 145–150, 1995. 48, 56, 127

LICCARDO, V. et al. The bingo project-iv. simulations for mission performance assessment and preliminary component separation steps. **Astronomy & Astrophysics**, v. 664, p. A17, 2022. 66, 67, 129

LIDDLE, A. R.; LYTH, D. H. The cold dark matter density perturbation. **Physics Reports**, v. 231, n. 1-2, p. 1–105, 1993. 6, 7

LOEB, A.; WYITHE, J. S. B. Possibility of precise measurement of the cosmological power spectrum with a dedicated survey of 21 cm emission after reionization. **Physical Review Letters**, v. 100, n. 16, p. 161301, 2008. 17

LUPTON, R. **Statistics in theory and practice**. [S.l.]: Princeton University Press, 1993. ISBN 9780691074290. 49

MATHER, J. C. et al. Early results from the Cosmic Background Explorer (COBE). **Advances in Space Research**, v. 11, n. 2, p. 181–191, jan. 1991. 71

MERÍCIA, E. J. de. **Simulações do processo de separação de componentes e recuperação do sinal de 21 cm do HI aplicadas ao radiotelescópio BINGO**. Dissertação (Mestrado) — (Mestrado em Astrofísica) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2019. 60, 62, 67

MERÍCIA, E. J. de. **Testing the recovery of the Bingo 21cm cosmological signal from different mission simulations**. Tese (Doutorado) — Instituto Nacional de Pesquisas Espaciais, São José dos Campo, 2023. 97

MESINGER, A.; FERRARA, A.; SPIEGEL, D. S. Signatures of X-rays in the early Universe. **Monthly Notices of the Royal Astronomical Society**, v. 431, n. 1, p. 621–637, 03 2013. ISSN 0035-8711. Disponível em: <<<https://doi.org/10.1093/mnras/stt198>>>. 15

MILOTTI, E. 1/f noise: a pedagogical review. **arXiv preprint physics/0204033**, 2002. 30, 35

MORIWAKI, K.; YOSHIDA, N.; SHIMIZU, I.; HARIKANE, Y.; MATSUDA, Y.; MATSUO, H.; HASHIMOTO, T.; INOUE, A. K.; TAMURA, Y.; NAGAO, T. The distribution and physical properties of high-redshift [o<scp>iii</scp>] emitters in a cosmological hydrodynamics simulation. **Monthly Notices of the Royal Astronomical Society: Letters**, v. 481, n. 1, p. L84–L88, set. 2018. ISSN 1745-3933. Disponível em: <<<http://dx.doi.org/10.1093/mnrasl/sly167>>>. 13

NAN, R.; LI, D.; JIN, C.; WANG, Q.; ZHU, L.; ZHU, W.; ZHANG, H.; YUE, Y.; QIAN, L. The five-hundred-meter aperture spherical radio telescope (fast) project. **International Journal of Modern Physics D**, v. 20, n. 06, p. 989–1024, 2011. 24

NAPIER, P. J.; THOMPSON, A. R.; EKERS, R. D. The very large array: design and performance of a modern synthesis radio telescope. **Proceedings of the IEEE**, v. 71, n. 11, p. 1295–1320, 1983. 13

NATOLI, P.; GASPERIS, G. D.; GHELLER, C.; VITTORIO, N. A map-making algorithm for the planck surveyor. **Astronomy & Astrophysics**, EDP Sciences, v. 372, n. 1, p. 346–356, 2001. 42, 48, 56, 57, 58, 127, 128

NEWBURGH, L. et al. Hirax: a probe of dark energy and radio transients. In: SPIE. **Ground-based and Airborne Telescopes VI**. [S.l.], 2016. v. 9906, p. 2039–2049. 24

NYQUIST, H. Thermal agitation of electric charge in conductors. **Physical Review**, v. 32, n. 1, p. 110, 1928. 32

- OH, S. P.; SPERGEL, D. N.; HINSHAW, G. An efficient technique to determine the power spectrum from cosmic microwave background sky maps. **The Astrophysical Journal**, v. 510, n. 2, p. 551, 1999. 46, 51
- OLIVARI, L.; REMAZEILLES, M.; DICKINSON, C. Extracting h i cosmological signal with generalized needlet internal linear combination. **Monthly Notices of the Royal Astronomical Society**, v. 456, n. 3, p. 2749–2765, 2016. 58, 60, 61, 113
- OLIVEIRA-COSTA, A. de; TEGMARK, M.; GAENSLER, B.; JONAS, J.; LANDECKER, T.; REICH, P. A model of diffuse galactic radio emission from 10 mhz to 100 ghz. **Monthly Notices of the Royal Astronomical Society**, v. 388, n. 1, p. 247–260, 2008. 18, 19
- PADMANABHAN, T. Cosmological constant—the weight of the vacuum. **Physics Reports**, v. 380, n. 5-6, p. 235–320, 2003. 5, 6
- PERLMUTTER, S. Supernovae, dark energy, and the accelerating universe: the status of the cosmological parameters. **Proceedings of the International Symposium on Lepton and Photon Interactions at High Energies**, 19, 1999. 1
- PETERSON, J. B. et al. 21 cm intensity mapping. **arXiv preprint arXiv:0902.3091**, 2009. 1, 12
- POUTANEN, T. et al. Comparison of map-making algorithms for cmb experiments. **Astronomy and Astrophysics**, v. 449, n. 3, p. 1311–1322, 2006. 42
- PRESTAGE, R. M.; CONSTANTIKES, K. T.; HUNTER, T. R.; KING, L. J.; LACASSE, R. J.; LOCKMAN, F. J.; NORROD, R. D. The green bank telescope. **Proceedings of the IEEE**, v. 97, n. 8, p. 1382–1390, 2009. 24
- PRITCHARD, J. R.; LOEB, A. 21 cm cosmology in the 21st century. **Reports on Progress in Physics**, v. 75, n. 8, p. 086901, 2012. 12, 14, 15, 17
- PULLEN, A. R.; DORÉ, O.; BOCK, J. Intensity mapping across cosmic times with the $1y\alpha$ line. **The Astrophysical Journal**, v. 786, n. 2, p. 111, abr. 2014. ISSN 1538-4357. Disponível em: <<<http://dx.doi.org/10.1088/0004-637X/786/2/111>>>. 13
- REMAZEILLES, M.; DELABROUILLE, J.; CARDOSO, J.-F. Foreground component separation with generalized internal linear combination. **Monthly Notices of the Royal Astronomical Society**, v. 418, n. 1, p. 467–476, 2011. 60, 61, 76
- RIESS, A. G. et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. **The Astronomical Journal**, v. 116, n. 3, p. 1009, 1998. 1
- RÖTTGERING, H. Lofar, a new low frequency radio telescope. **New Astronomy Reviews**, v. 47, n. 4-5, p. 405–409, 2003. 24

RYBICKI, G. B.; LIGHTMAN, A. P. **Radiative processes in astrophysics**. [S.l.]: John Wiley & Sons, 2024. 16

RYDEN, B. **Introduction to cosmology**. [S.l.]: Cambridge University Press, 2017. 5

SEIFFERT, M.; MENNELLA, A.; BURIGANA, C.; MANDOLESINI, N.; BERSANELLI, M.; MEINHOLD, P.; LUBIN, P. noise and other systematic effects in the planck-lfi radiometers. **Astronomy & Astrophysics**, v. 391, n. 3, p. 1185–1197, 2002. 30, 35

SILVA, B. M.; ZAROUBI, S.; KOOISTRA, R.; COORAY, A. Tomographic intensity mapping versus galaxy surveys: observing the universe in $h\alpha$ emission with new generation instruments. **Monthly Notices of the Royal Astronomical Society**, v. 475, n. 2, p. 1587–1608, dez. 2017. ISSN 1365-2966. Disponível em: <<<http://dx.doi.org/10.1093/mnras/stx3265>>>. 13

SUTTON, D.; ZUNTZ, J.; FERREIRA, P.; BROWN, M.; ERIKSEN, H.; JOHNSON, B.; KUSAKA, A.; NÆSS, S.; WEHUS, I. Fast and precise map-making for massively multi-detector cmb experiments. **Monthly Notices of the Royal Astronomical Society**, v. 407, n. 3, p. 1387–1402, 2010. 43

TAUBER, J. et al. The planck mission. **Advances in Space Research**, v. 34, n. 3, p. 491–496, 2004. 1, 48, 71, 121, 127, 128

TEGMARK, M. How to make maps from cosmic microwave background data without losing information. **The Astrophysical Journal**, v. 480, n. 2, p. L87, 1997. 39, 42, 43, 45, 46, 48

_____. How to measure cmb power spectra without losing information. **Physical Review D**, v. 55, n. 10, p. 5895, 1997. 39, 40

TRISTRAM, M.; FILLIARD, C.; PERDEREAU, O.; PLASZCZYNSKI, S.; STOMPOR, R.; TOUZE, F. Iterative destriping and photometric calibration for planck-hfi, polarized, multi-detector map-making. **Astronomy & Astrophysics**, v. 534, p. A88, 2011. 42

TYSON, J. A. Large synoptic survey telescope: overview. **Survey and Other Telescope Technologies and Discoveries**, v. 4836, p. 10–20, 2002. 1

VILLAESCUSA-NAVARRO, F. et al. Ingredients for 21 cm intensity mapping. **The Astrophysical Journal**, v. 866, n. 2, p. 135, 2018. 14, 18

WANG, J. et al. H i intensity mapping with meerkat: calibration pipeline for multidish autocorrelation observations. **Monthly Notices of the Royal Astronomical Society**, v. 505, n. 3, p. 3698–3721, 2021. 121

WANG, Y. Dark energy constraints from baryon acoustic oscillations. **The Astrophysical Journal**, v. 647, n. 1, p. 1–7, ago. 2006. ISSN 1538-4357. Disponível em: <<<http://dx.doi.org/10.1086/505384>>>. 9

- WEINBERG, D. H.; MORTONSON, M. J.; EISENSTEIN, D. J.; HIRATA, C.; RIESS, A. G.; ROZO, E. Observational probes of cosmic acceleration. **Physics Reports**, v. 530, n. 2, p. 87–255, 2013. 1, 5, 9, 10
- WRIGHT, E. L.; HINSHAW, G.; BENNETT, C. L. Producing megapixel cosmic microwave background maps from differential radiometer data. **The Astrophysical Journal**, v. 458, n. 2, p. L53, 1996. 42
- WUENSCHÉ, C.; COLLABORATION, B. The bingo telescope: a new instrument exploring the new 21 cm cosmology window. **Journal of Physics, Conference Series**, v. 1269, n. 1, p. 012002, 2019. 1, 55
- WUENSCHÉ, C. A. et al. The bingo project-ii. instrument description. **Astronomy & Astrophysics**, v. 664, p. A15, 2022. 2, 25, 26, 28, 29, 30, 31, 32, 34, 63, 64, 71, 83, 129
- YOO, J.; WATANABE, Y. Theoretical models of dark energy. **International Journal of Modern Physics D**, v. 21, n. 12, p. 1230002, 2012. 5
- YORK, D. G.; ADELMAN, J.; JR, J. E. A.; ANDERSON, S. F.; ANNIS, J.; BAHCALL, N. A.; BAKKEN, J.; BARKHOUSER, R.; BASTIAN, S.; BERMAN, E. et al. The sloan digital sky survey: Technical summary. **The Astronomical Journal**, v. 120, n. 3, p. 1579, 2000. 11
- ZHUANG, Y.; STRIGARI, L. E.; JIN, L.; SINHA, S. Prospects for measuring the time variation of astrophysical neutrino sources at dark matter detectors. **Physical Review D**, v. 110, n. 4, p. 043037, 2024. 12

PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

Teses e Dissertações (TDI)

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

Manuais Técnicos (MAN)

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

Notas Técnico-Científicas (NTC)

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programas de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

Relatórios de Pesquisa (RPQ)

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

Propostas e Relatórios de Projetos (PRP)

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

Publicações Didáticas (PUD)

Incluem apostilas, notas de aula e manuais didáticos.

Publicações Seriadas

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Contam destas publicações o Internacional Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

Programas de Computador (PDC)

São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. Aceitam-se tanto programas fonte quanto os executáveis.

Pré-publicações (PRE)

Todos os artigos publicados em periódicos, anais e como capítulos de livros.