

Universidade de São Paulo  
Instituto de Física

# Cálculo de covariância de mapas cosmológicos na pipeline de dados do telescópio BINGO

Felipe Silva Caldas

Orientador: Prof. Dr. Elcio Abdalla

Dissertação de mestrado apresentada ao Instituto  
de Física da Universidade de São Paulo, como re-  
quisito parcial para a obtenção do título de Mestre  
em Ciências.

Banca Examinadora:  
Prof. Dr. Elcio Abdalla - Orientador (USP)

São Paulo  
2024

**FICHA CATALOGRÁFICA**  
**Preparada pelo Serviço de Biblioteca e Informação**  
**do Instituto de Física da Universidade de São Paulo**

Caldas, Felipe Silva

**Cálculo de covariância de mapas cosmológicos na pipeline de dados do telescópio BINGO.** São Paulo, 2024.

Dissertação (Mestrado) - Universidade de São Paulo.  
Instituto de Física. Depto. de Física Geral.

Orientador: Prof. Dr. Elcio Abdalla.

Área de Concentração: Cosmologia.

**Unitermos:** 1. Cosmologia; 2. Processamento de dados;  
3. Telescópios.

USP/IF/SBI-098/2024

University of São Paulo  
Physics Institute

Covariance calculation of cosmological maps in the  
BINGO telescope data pipeline

Felipe Silva Caldas

Supervisor: Prof. Dr. Elcio Abdalla

Dissertation submitted to the Physics Institute of  
the University of São Paulo in partial fulfillment  
of the requirements for the degree of Master of  
Science.

Examining Committee  
Prof. Dr. Elcio Abdalla - Supervisor (USP)

São Paulo  
2024

## **Agradecimentos**

Sou grato *àquele que é*, Princípio de todas as criaturas, por me mostrar o caminho do mistério.

Agradeço à minha Família por estar sempre ao meu lado.

Sou grato aos meus orientadores Elcio Abdalla e Filipe Batoni Abdalla por acreditarem em mim e por me oferecerem suporte nos momentos mais complicados.

Sou grato aos meus companheiros de equipe, pois só podemos realizar o que estamos realizando porque somos um grupo.

## Resumo

O Radiotelescópio BINGO (*Baryon acoustic oscillations from Integrated Neutral Gas Observations*; oscilações acústicas bariônicas em observações integradas de gás neutro) usará a técnica de *intensity mapping* para realizar observações da linha de 21cm do hidrogênio neutro para detectar oscilações acústicas bariônicas (BAO) em baixos *redshifts* (na faixa de 0.127 a 0.449, correspondendo a frequências de 980 MHz a 1260 MHz) a fim de investigar o setor escuro da cosmologia (matéria escura e energia escura).

Na pipeline do BINGO as observações do céu se traduzem em mapas cosmológicos e então são extraídas as médias dos espectros de potência. Estas, por sua vez, são utilizadas para gerar a matriz de covariância dos mapas, e, sequencialmente, esta matriz se torna a base para o cálculo dos parâmetros cosmológicos.

Em cosmologia observational as medições estão sujeitas a erros de vários tipos. Alguns destes erros são efeitos sistemáticos e mapeados. Realiza-se então uma análise para verificar se os mapas cosmológicos sofrem alguma influência destes efeitos através do espectro de potência cruzado entre mapas e mapas sistemáticos, para garantir que não sejam fontes de viés nas estimativas dos parâmetros cosmológicos.

Este trabalho tem o propósito de estruturar um software com algoritmo paralelizado capaz de automatizar o cálculo da matriz de covariância de dados cosmológicos na pipeline do BINGO utilizando um método de reamostragem, que consiste em gerar n-simulações dos mapas cosmológicos, mensurar o espectro de potência angular destas simulações através do estimador Pseudo- $C\ell$  e a partir destas medidas gerar a matriz de covariância.

**Palavras-chave:** Cosmologia, BINGO, Pseudo- $C\ell$ , covariância.

## Abstract

The BINGO Radiotelescope (*Baryon acoustic oscillations from Integrated Neutral Gas Observations*; baryonic acoustic oscillations in integrated neutral gas observations) will use the technique of *intensity mapping* to observe the 21cm line of neutral hydrogen to detect baryonic acoustic oscillations (BAO) at low *redshifts* (in the range of 0.127 to 0.449, corresponding to frequencies from 980 MHz to 1260 MHz) in order to investigate the dark sector of cosmology (dark matter and dark energy).

In the BINGO pipeline, sky observations are translated into cosmological maps, from which the average power spectra are extracted. These, in turn, are used to generate the covariance matrix of the maps, which sequentially becomes the basis for calculating cosmological parameters.

In observational cosmology, measurements are subject to various types of errors. Some of these errors are systematic effects and mapped. An analysis is then conducted to check if the cosmological maps are influenced by these effects through the cross-power spectrum between maps and systematic maps, to ensure they do not introduce bias in the estimates of cosmological parameters.

This work aims to structure software with a parallelized algorithm capable of automating the calculation of the covariance matrix of cosmological data in the BINGO pipeline using a resampling method, which consists of generating n-simulations of cosmological maps, measuring the angular power spectrum of these simulations through the Pseudo- $C\ell$  estimator and from these measurements generating the covariance matrix.

**Keywords:** Cosmology, BINGO, Pseudo- $C\ell$ , covariance.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>6</b>
<b>2</b>	<b>Dados cosmológicos</b>	<b>12</b>
2.0.1	Mapas cosmológicos . . . . .	12
2.0.2	Mapas sistemáticos . . . . .	18
<b>3</b>	<b>Espectro de potência</b>	<b>24</b>
3.1	Função de correlação de 2-pontos . . . . .	24
3.2	Espectro de potência . . . . .	25
3.3	Poisson shot noise . . . . .	27
3.4	Estimador Pseudo- $C_\ell$ . . . . .	28
3.5	Resultados: Espectro de potência dos mapas cosmológicos . . . . .	30
3.5.1	Pseudo-Power . . . . .	30
3.5.2	Healpy . . . . .	33
3.5.3	UCLCL . . . . .	35
<b>4</b>	<b>Métodos de reamostragem</b>	<b>42</b>
4.1	FLASK: Simulações de mapas cosmológicos . . . . .	43
4.2	Automatização do processo . . . . .	48
4.3	Resultados: Espectro de potência das simulações . . . . .	51
4.3.1	$C_\ell$ . . . . .	51
4.3.2	$S_\ell$ . . . . .	58
4.4	Resultados: $C_\ell$ 's entre mapas cosmológicos e sistemáticos . . . . .	66
<b>5</b>	<b>Matriz de covariância</b>	<b>73</b>
5.1	Resultados . . . . .	74
5.2	Binning: Compactando informação . . . . .	78
<b>6</b>	<b>Conclusão</b>	<b>79</b>
<b>7</b>	<b>Referências</b>	<b>80</b>
<b>8</b>	<b>Anexo I - Códigos</b>	<b>83</b>
8.1	dados_mascara.py . . . . .	83
8.2	flask.config . . . . .	84
8.3	SDSS_14Bins.dat . . . . .	88

8.4	SDSS_14Bins_radial-selection-function-f1.dat	89
8.5	esy.ini	92
8.6	UCLCL.ini	95
<b>9</b>	<b>Anexo II - Mapas cosmológicos</b>	<b>100</b>
9.1	Escala normal	100
9.2	Escala logarítmica	108
<b>10</b>	<b>Anexo III - Mapas sistemáticos</b>	<b>116</b>
<b>11</b>	<b>Anexo IV - <math>C_\ell</math> e <math>S_\ell</math> dos mapas cosmológicos</b>	<b>121</b>
<b>12</b>	<b>Anexo V - Matrizes de covariância: <math>C\ell</math></b>	<b>126</b>
<b>13</b>	<b>Anexo VI - Matrizes de covariância: <math>S\ell</math></b>	<b>140</b>

# 1 Introdução

A compreensão das estruturas de larga escala do universo e sua evolução é objeto de discussão central na cosmologia. Para que haja uma descrição coerente é necessário mensurar os parâmetros que caracterizam determinado modelo cosmológico do Universo. O modo como o ajuste dos parâmetros da teoria se dá é principalmente através da análise da distribuição de matéria do Universo, que é convenientemente descrita pelo espectro de potência angular (*Angular Power Spectrum*; APS) da matéria. A maneira mais natural de iniciar a discussão sobre o espectro de potência é introduzindo o conceito de **campo de densidade**. Para descrever a distribuição de matéria no universo, usamos uma função que expressa a diferença entre a densidade local de matéria e a densidade média do universo. Isso é feito através do campo de densidade  $\delta(\vec{x})$ , definido como:

$$\delta(\vec{x}) \equiv \frac{\rho(\vec{x}) - \langle \rho(\vec{x}) \rangle}{\langle \rho(\vec{x}) \rangle}, \quad (1)$$

onde  $\rho(\vec{x})$  é a densidade de matéria em um ponto específico  $\vec{x}$ ,  $\langle \cdot \rangle$  denota o valor esperado (média estatística) e  $\langle \rho(\vec{x}) \rangle$  é a densidade média de matéria no universo. Esse campo de densidade nos informa o quanto a densidade de matéria em uma posição específica  $\vec{x}$  se desvia da densidade média  $\langle \rho(\vec{x}) \rangle$ .

Para investigar as relações entre diferentes regiões no universo, utilizamos a **função de correlação de dois pontos**  $\xi(\vec{x}_1, \vec{x}_2)$ , que quantifica a distribuição de pares de pontos em um espaço, medindo a probabilidade de encontrar esses pares separados por uma distância  $\vec{r} \equiv \vec{x}_1 - \vec{x}_2$ . A função permite identificar padrões espaciais e pode capturar dependências não lineares na distribuição dos dados. A função de correlação é dada por:

$$\xi(\vec{x}_1, \vec{x}_2) \equiv \langle \delta(\vec{x}_1) \delta(\vec{x}_2) \rangle. \quad (2)$$

onde  $\delta(\vec{x}_1)$  e  $\delta(\vec{x}_2)$  são as flutuações de densidade em dois pontos  $\vec{x}_1$  e  $\vec{x}_2$  separados por uma distância  $|\vec{r}|$ . Essa função nos revela quão relacionadas estão as distribuições de matéria em diferentes pontos no espaço.

Para facilitar o estudo dessas correlações, é comum realizar a **transformada de Fourier** do campo de densidade, passando do espaço real para o **espaço de Fourier**. A transformada de Fourier do campo de densidade é dada por:

$$\delta(\vec{k}) = \int d^3x \delta(\vec{x}) e^{-i\vec{k}\cdot\vec{x}}, \quad (3)$$

onde  $\delta(\vec{k})$  é o campo de densidade no espaço de Fourier,  $\vec{k}$  é o vetor de onda que corresponde a uma determinada escala de flutuação,  $e^{-i\vec{k}\cdot\vec{r}}$  é o fator oscilatório da transformada de Fourier, e  $d^3x$  é o elemento de volume tridimensional no espaço real. Ao transformar o campo de densidade para o espaço de Fourier, podemos examinar as flutuações de densidade em termos das suas frequências, ou seja, nas diferentes escalas de comprimento.

O próximo passo é definir o **espectro de potência**  $P_m(\vec{k})$ , que mede a amplitude dessas flutuações em diferentes escalas espaciais. O espectro de potência é dado por:

$$\langle \delta(\vec{k}) \delta^\dagger(\vec{k}') \rangle = (2\pi)^3 \delta_D(\vec{k} - \vec{k}') P_m(\vec{k}), \quad (4)$$

onde  $\delta^\dagger(\vec{k}')$  é o complexo conjugado de  $\delta(\vec{k}')$ ,  $\delta_D(\vec{k} - \vec{k}')$  é a função delta de Dirac, que impõe a condição de que  $\vec{k} = \vec{k}'$  para termos a correlação e  $P_m(\vec{k})$  é o espectro de potência

da matéria, que nos informa a quantidade de flutuações de densidade em diferentes escalas, ou seja, nos permite quantificar como a matéria está distribuída no universo em diferentes escalas.

Para estimar o espectro de potência a partir de dados observacionais, usamos o **estimador Pseudo-Cl** ([Peebles-1973](#)), que nos permite calcular o espectro de potência angular mesmo em cenários com cobertura parcial do céu ou com dados contaminados.

Para esse propósito é necessária uma extensa observação dos objetos no céu profundo, onde se tornam distinguíveis estas estruturas. A determinação da distribuição de matéria do universo através do que se observa é uma medida indireta e necessita da hipótese de que tal observável representa bem a distribuição de matéria. Isto é, se observarmos a distribuição de galáxias, então supõe-se que esta distribuição seja equivalente à distribuição de matéria a menos de um viés (*bias*), que comumente se supõe que seja constante em casos simples ou como função do *redshift*  $z$ , de modo que em regimes lineares podemos escrever:

$$P_g(\vec{k}) = b(z)^2 P_m(\vec{k}). \quad (5)$$

Onde  $P(\vec{k})$  representa o espectro de potência no espaço de Fourier, com vetor de onda  $\vec{k}$ , os índices  $g$  e  $m$  representam galáxias e matéria respectivamente e  $b(z)$  é o viés em função do *redshift*. Neste trabalho, o estimador que será utilizado para o espectro de potência será o Pseudo-Cl .

A criação de bancos de dados desses objetos celestes (*surveys*) tem sido um grande desafio instrumental, uma vez que a técnica empregada até agora pelos principais rádio telescópios é a medição espectroscópica individual dos objetos celestes para então realizar a classificação da sua respectiva distância cosmológica (*redshift*).

Uma técnica alternativa para a medida dessas estruturas é o mapeamento de intensidade (*Intensity Mapping*), que consiste na observação da intensidade de regiões inteiras em vez de objetos individuais, tornando o mapeamento muito mais rápido. Essa técnica diminui a resolução das observações, mas de modo que seja suficiente para o propósito de observar como estruturas de larga escala estão distribuídas.

Ainda na busca por otimização, deve existir um direcionamento sobre o que é observado, ou seja, quais frequências devem ser observadas. Sabidamente, o hidrogênio neutro é o elemento mais abundante do universo e está presente no meio interestelar na forma de gás difuso. A hipótese que se faz é que a distribuição do hidrogênio é uma boa representação da distribuição de matéria, de modo que

$$P_{21cm}(\vec{k}) = b(z)^2 P_m(\vec{k}). \quad (6)$$

Em 1945, H. C. Van De Hulst ([Hulst-1945](#)) propôs a existência de uma transição na estrutura hiperfina do átomo de hidrogênio resultando na emissão de um fóton cujo comprimento de onda é de 21cm. O fenômeno é comumente nomeado *spin-flip*, porque se trata de uma transição do estado onde o spin do próton e do elétron estão alinhados paralelamente para um estado onde o alinhamento é antiparalelo, liberando o fóton no processo. Essa transição só foi detectada em 1951 ([Ewen,Purcell-1951](#)). Uma descrição mais detalhada do fenômeno é elaborada por Griffiths ([Griffiths-1982](#)).

O telescópio BINGO (*Baryon acoustic oscillations from Integrated Neutral Gas Observations*; oscilações acústicas bariônicas em observações integradas de gás neutro) usará a

técnica de *intensity mapping* para realizar observações da linha de 21cm do hidrogênio neutro para detectar oscilações acústicas bariônicas (BAO) em baixos *redshifts* (na faixa de 0.127 a 0.449, correspondendo a frequências de 980 MHz a 1260 MHz) a fim de investigar o setor escuro da cosmologia (matéria escura e energia escura), tendo também a possibilidade de realizar observações pertinentes no estudo da astrofísica, com a observação de rajadas rápidas de rádio (*Fast Radio Bursts*; FRB) e outros fenômenos ([BINGO-project-I](#)).

O objetivo deste trabalho é a criação de um processo (pipeline) automatizado de avaliação dos mapas cosmológicos com a finalidade de estimar o espectro de potência dos dados fornecidos (mapas cosmológicos) e sequencialmente calcular a respectiva matriz de covariância. A matriz de covariância servirá para etapas posteriores ao fluxo de trabalho aqui implementado, onde será base para o cálculo dos parâmetros cosmológicos utilizando estatística Bayesiana.

O produto final desta tese será um software capaz de automatizar todo o processo que aqui será descrito e que possa realizar todos os cálculos em múltiplos processadores de uma única vez (processamento paralelizado).

O processo que será desenvolvido pode ser dividido em cinco grandes blocos, a serem detalhados nos capítulos seguintes:

1. Cálculo do Pseudo- $C_\ell$  entre os mapas cosmológicos e efeitos sistemáticos.
2. Cálculo do Pseudo- $C_\ell$  dos mapas cosmológicos.
3. Geração de n-simulações de mapas cosmológicos gerados pelo software [FLASK](#) utilizando os espectros de potência da etapa anterior como parâmetros de entrada.
4. Cálculo do Pseudo- $C_\ell$  das simulações.
5. Cálculo da matriz de covariância dos mapas simulados.

Cada uma destas etapas possuem seus próprios aspectos a serem detalhados ao longo dos capítulos. Para os quatro primeiros pontos é necessário estimar o espectro de potência dos mapas cosmológicos, para isso usaremos a equação

$$S_\ell^{ij} = \frac{1}{w_\ell^2} \left[ \left( \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} C_{\ell m}^{ij} \right) - N_\ell \delta_{ij} \right], \quad (7)$$

onde  $w_\ell^2$  é a *pixel window function*,  $\ell$ 's caracterizam os harmônicos,  $C_{\ell m}^{ij}$  é o espectro de potência entre dois campos de sobredensidade de índice  $i$  e  $j$  e o termo  $N_\ell$  é o shotnoise

$$C_\ell = S_\ell + N_\ell, \quad (8)$$

onde  $N_\ell$  é escrito como

$$N_\ell = \frac{\Delta\Omega_{tot}}{n_{tot}} = \frac{1}{\bar{n}}, \quad (9)$$

sendo que  $\Delta\Omega_{tot}$  é o angulo sólido total da região observada e  $n_{tot}$  é o número de objetos observados (galáxias no nosso caso) e  $\bar{n}$  é a densidade média de objetos por ângulo sólido.

Como forma de verificação e afim gerar confiabilidade nos resultados, os dados serão comparados com as medidas extraídas pela biblioteca do Healpy<sup>1</sup>, que tem como parâmetros de entrada os mapas cosmológicos e pelo código UCLCL (*Unified Cosmological Library for Cℓ’s*, Loureiro-2019) que utiliza as distribuições dos objetos em relação ao redshift no cálculo, a fim de realizar um teste de consistência.

Para assegurar a qualidade dos dados utilizados, é fundamental incluir no processo a verificação dos mapas cosmológicos, garantindo que não sejam afetados por interferências externas ao objeto de estudo, conhecidas como efeitos sistemáticos. Para isso, serão calculados na primeira etapa os espectros de potência entre os mapas cosmológicos e os mapas sistemáticos, que representam efeitos indesejados e conhecidos, com o intuito de confirmar que os dados não estão sendo comprometidos por fenômenos externos prejudiciais.

A terceira etapa envolve a realização de simulações dos mapas cosmológicos, utilizando o espectro de potência dos mapas originais como parâmetro inicial. O espectro de potência descreve a distribuição dos objetos no universo, e essa informação é utilizada para gerar novas distribuições, introduzindo variabilidade aleatória, mas preservando a estrutura fundamental dos dados (usando a decomposição de Cholesky). Essa técnica é conhecida genericamente como simulação de Monte Carlo, na qual as soluções dos problemas são representadas como parâmetros de uma população hipotética. A partir de uma sequência de números aleatórios, constrói-se uma amostra dessa população, permitindo a obtenção de estimativas estatísticas dos parâmetros estudados.

A partir dos espectros de potência gerados por cada conjunto de simulações, será possível calcular a matriz de covariância correspondente, que quantifica a incerteza associada às medidas e captura a correlação entre diferentes escalas. Será avaliado a qualidade dessas simulações comparando a matriz de covariância obtida com a matriz derivada dos dados reais, garantindo que os efeitos aleatórios introduzidos respeitem a variabilidade estatística esperada. Além disso, para cada conjunto de simulações, avaliaremos o erro entre a medida real  $X$  e uma medida estimada das simulações  $\hat{X}$ ,

$$Erro = X - \hat{X}. \quad (10)$$

Além disso medimos a variância da amostra,  $\sigma^2$ , que servirá tanto para determinar um intervalo de confiança quanto para avaliar a qualidade das simulações,

$$\sigma^2 = \frac{\sum (X_i - \bar{X})^2}{n - 1}, \quad (11)$$

onde  $\bar{X}$  é a média da medida em questão. Neste estudo, o erro e a variância serão mensurados ponto a ponto. Ou seja, dada uma curva medida dos dados  $C(\ell)$  e  $n$ -curvas  $\hat{C}(\ell)$  provenientes das simulações, cada métrica  $\mathbf{M}$  será um  $M(\ell)$  levando em consideração um número  $n$  de medidas  $X_i$  na sua composição. Deste modo, em cada ponto teremos  $n$  medidas de erro e uma variância associados. Essas medidas trarão mais informação se forem normalizadas pela medida em questão. Neste caso, serão nomeadas medidas relativas (e.g.: erro relativo, variância relativa), do contrário, serão chamadas de absolutas.

O conjunto de medidas do espectro de potência proveniente das simulações servirá para calcular a matriz de covariância através da seguinte equação

$$S_{\ell\ell'}^{ij} = \frac{1}{N_s - 1} \sum_{s=0}^{N_s} (S_\ell^{ij,s} - \langle S_\ell^{ij,s} \rangle) (S_{\ell'}^{ij,s} - \langle S_{\ell'}^{ij,s} \rangle)^T. \quad (12)$$

---

<sup>1</sup><https://healpy.readthedocs.io/en/latest/generated/healpy.sphtfunc.anafast.html>

Desse modo,  $S_{\ell\ell'}^{ij}$  refere-se à matriz de covariância das medidas de Pseudo- $C\ell$  sem shotnoise,  $N_s$  é o número de simulações,  $s$  diz respeito às simulações,  $\ell$  e  $\ell'$  são os harmônicos de cada conjunto de dados,  $\langle S \rangle$  denota que os objetos estão sendo subtraídos de seus respectivos valores médios e  $T$  representa a transposta dos dados.

O projeto BINGO possui muitas etapas, uma descrição instrumental, incluindo o design óptico, construção dos receptores, calibração e estratégia de observação pode ser encontrada no [BINGO-project-II](#). O artigo também descreve as características da localização do projeto, justificando sua escolha. Um aprofundamento detalhado do design óptico da superfície focal do telescópio pode ser encontrado no [BINGO-project-III](#), neste estudo é descrito a análise que conduziu a escolha do arranjo óptico do BINGO, denominado duplo-retangular.

Uma simulação dos fenômenos a serem estudados e da pipeline de trabalho do BINGO pode ser encontrada em [BINGO-project-IV](#). O artigo simula os sinais cosmológicos utilizando o [FLASK](#), realiza uma modelagem de ruídos cosmológicos, ruídos instrumentais e realiza um estudo sobre máscaras das diferentes propostas de estratégia de observação. O artigo [BINGO-project-V](#) realiza avanços em casos específicos quanto a separação de componentes e utiliza o bispectrum como meio de análise e checagem deste procedimento. Mais detalhes a respeito das simulações dos catálogos utilizados estão dispostos em [BINGO-project-VI](#).

Um estudo que explora a capacidade de estimar parâmetros cosmológicos do BINGO e compara seus resultados projetados com os de outros telescópios como o Plank por exemplo, pode ser encontrado no [BINGO-project-VII](#). Este estudo conclui que embora o BINGO sozinho não forneça restrições tão competitivas quanto o Plank no modelo  $\Lambda$ CDM, sua combinação melhora significativamente a precisão dos parâmetros estimados. Além disso o estudo mostra como o BINGO pode complementar futuros levantamentos cosmológicos.

Na busca por previsões sobre a medição do sinal BAO (*Baryon Acoustic Oscillations*) temos o [BINGO-project-VII](#). Este trabalho utiliza dois estimadores de agrupamento, a função de correlação angular (ACF) e o espectro de potência angular (APS), para modelar e extraer informações sobre o sinal BAO a partir de simulações, com o APS se destacando por fornecer estimativas ligeiramente mais precisas e uma maior probabilidade de detecção do sinal BAO, especialmente em redshifts mais elevados. O estudo também avalia o impacto do ruído instrumental e sinais residuais.

O telescópio BINGO também possuirá capacidade de contribuir para o estudo dos fenômenos de FRB's (Fast Radio Burst). Um estudo que propõe o uso de um sistema de interferometria como extensão para o telescópio BINGO (BIS - BINGO Interferometry System), incluindo novos radiotelescópios auxiliares menores (outiggers), pode ser encontrado no [BINGO-project-IX](#). A abordagem interferométrica visa localizar as fontes de FRBs no céu. O estudo gera catálogos sintéticos de FRBs e avalia a taxa de detecção e a capacidade de localização das diferentes configurações do BIS. Os resultados indicam que, na melhor configuração, o BIS pode localizar 23 FRBs por ano, utilizando 10 outiggers com espelhos de 6 m e redshift  $z \leq 0.96$ .

A próxima fase do projeto BINGO, denominada Advanced Bingo Dark Universe Studies (ABDUS), pode ser encontrada em [BINGO-ABDUS](#). O projeto ABDUS visa ampliar as capacidades do BINGO através da inclusão de arranjos em fases (PAs - Phased Arrays) e estações de outiggers, permitindo uma maior cobertura do céu e melhorando a velocidade do levantamento. Arranjos de fases (PAs) são um conjunto de elementos de antenas

estreitamente distribuídos, separados por cerca de meio comprimento de onda, que podem sintetizar múltiplos feixes independentes, garantindo uma cobertura uniforme. O número de feixes simultaneamente sintetizados depende da capacidade de processamento digital. Com essa nova configuração, espera-se que o BINGO-ABDUS cubra até 50% do céu, alcançando um redshift de aproximadamente 2.1, com uma resolução angular entre 27'-40', e atingindo até 2,5" de resolução angular em fontes individuais via interferometria com estações localizadas a 20 km do telescópio principal. O ABDUS abrirá possibilidades para sinergias com outros grandes radiotelescópios, como o FAST e o Tianlai, o que pode resultar em uma das maiores e mais precisas redes de radiotelescópios do mundo.

## 2 Dados cosmológicos

### 2.0.1 Mapas cosmológicos

Os mapas cosmológicos estudados são medidas espectroscópicas que pertencem ao catálogo de dados BOSS DR12 (*Baryon Oscillation Spectroscopic Survey Data Release 12*)<sup>2</sup>, que é um dos dados disponíveis no SDSS DR12 (*Sloan Digital Sky Survey Data Release 12*). O BOSS DR12 se divide em dois subconjuntos principais, o LOWZ que possui um intervalo de redshift limitado a  $z < 0.45$  e o CMASS que possui um intervalo entre  $0.45 < z < 0.80$ . Para mais detalhes de como esse catálogo é construído pode-se consultar [Alem et all-2015](#). Neste estudo serão usados apenas os dados relacionados ao LOWZ e faremos um corte no redshift  $z = 0.40$ , dividindo os dados em quatorze intervalos distintos (eventualmente denominados *bins*). A tabela 1 mostra a forma como foram divididos os dados por intervalos de redshift.

Mapa	Redshift	Nº Objetos
0	0.000 - 0.050	127965
1	0.050 - 0.075	253058
2	0.075 - 0.100	460207
3	0.100 - 0.125	575762
4	0.125 - 0.150	629783
5	0.150 - 0.175	639377
6	0.175 - 0.200	593565
7	0.200 - 0.225	541703
8	0.225 - 0.250	456173
9	0.250 - 0.275	364873
10	0.275 - 0.300	289499
11	0.300 - 0.325	245367
12	0.325 - 0.350	207064
13	0.350 - 0.400	49410

Tabela 1: Divisão dos quatorze mapas por intervalo de redshift e o número de objetos totais de cada intervalo.

Do ponto de vista instrumental, a divisão dos mapas é feita de forma a manter uma quantidade mínima de objetos, de modo que seja possível realizar simulações cosmológicas destes mapas no FLASK, que será abordado mais adiante. A densidade de objetos dos mapas é um dos parâmetros de input na simulação. Quando essa densidade for muito baixa no mapa de referência, o mapa simulado acaba por ser gerado com uma variância muito

<sup>2</sup><https://data.sdss.org//sas/dr12/boss/lss/>

grande de densidade entre as simulações, perdendo a precisão desejada como veremos na sessão de resultados.

Há dois tipos de mapas cosmológicos que iremos utilizar neste estudo, são eles os mapas de sobredensidade (*overdensity*) e de densidade (*density*). Os mapas de densidade  $\rho$  são construídos de forma que o valor de cada pixel é o número de objetos contabilizados naquela região, e.g. se um pixel possuir valor dez, significa que existem dez objetos naquela região. Os mapas de sobredensidade  $\delta$  por sua vez medem a variação de densidade com base na densidade média  $\bar{\rho}$  de objetos. Estas grandezas se relacionam através de

$$\delta = \frac{\rho - \bar{\rho}}{\bar{\rho}} . \quad (13)$$

Os mapas de sobredensidade são de interesse, pois temos o objetivo de avaliar a forma como a densidade de matéria está distribuída relativamente no universo.

Existem dois ordenamentos principais de construir estes mapas cosmológicos no que se refere a posição do pixel baseado nas suas coordenadas cosmológicas, são eles o RING e o NESTED e a diferença entre eles é apenas uma questão de formato organização dos dados<sup>3</sup>. Nesta pipeline de dados utilizaremos o ordenamento RING, além disso, determinaremos uma resolução específica para os mapas, que neste contexto denomina-se NSIDE e no nosso caso será igual a 512. Para esta resolução, o numero de pixels totais do mapa é  $NPIX = 12 * 512^2$ .

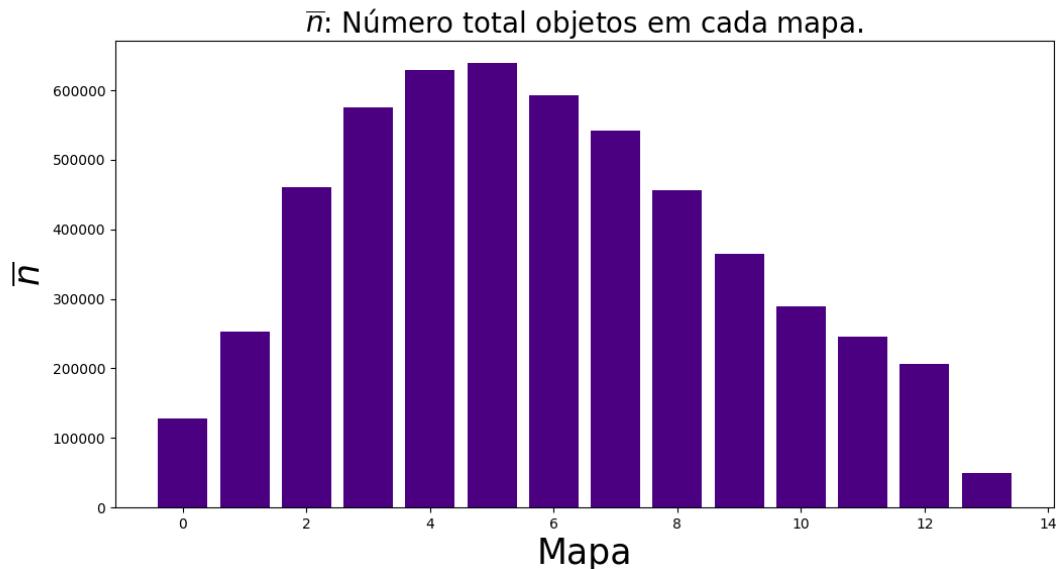


Figura 1: Número de galáxias  $\bar{n}$  dos mapas cosmológicos.

---

<sup>3</sup><https://healpy.readthedocs.io/en/1.16.6/tutorial.html>

A máscara dos dados provem do catálogo do BOSS para os dados do LOWZ<sup>4</sup>. Ela é uma máscara binária em resolução  $NSIDE = 512$  cobrindo aproximadamente 24% do céu, esse valor é escrito como a variável  $f_{sky} = 0.2479810$  e é uma informação necessária no código [CovAnalysis](#). Em posse da máscara binária, ambas as informações podem ser obtidas utilizando biblioteca do python [Healpy](#) e [NumPy](#), em anexo [dados\\_mascara.py](#). Para obter o número de pixels total de um mapa:

$$NPIX = 12 * NSIDE^2. \quad (14)$$

A área total da máscara em esferorradianos:

$$Area = f_{sky} \cdot 4\pi \simeq 3.116 [sr], \quad (15)$$

convertendo para  $deg^2$ :

$$Area = f_{sky} \cdot 4\pi \left( \frac{180}{\pi} \right)^2 \simeq 10229.95 [deg^2]. \quad (16)$$

A máscara binária é utilizada tanto para os mapas cosmológicos quanto para os mapas sistemáticos.

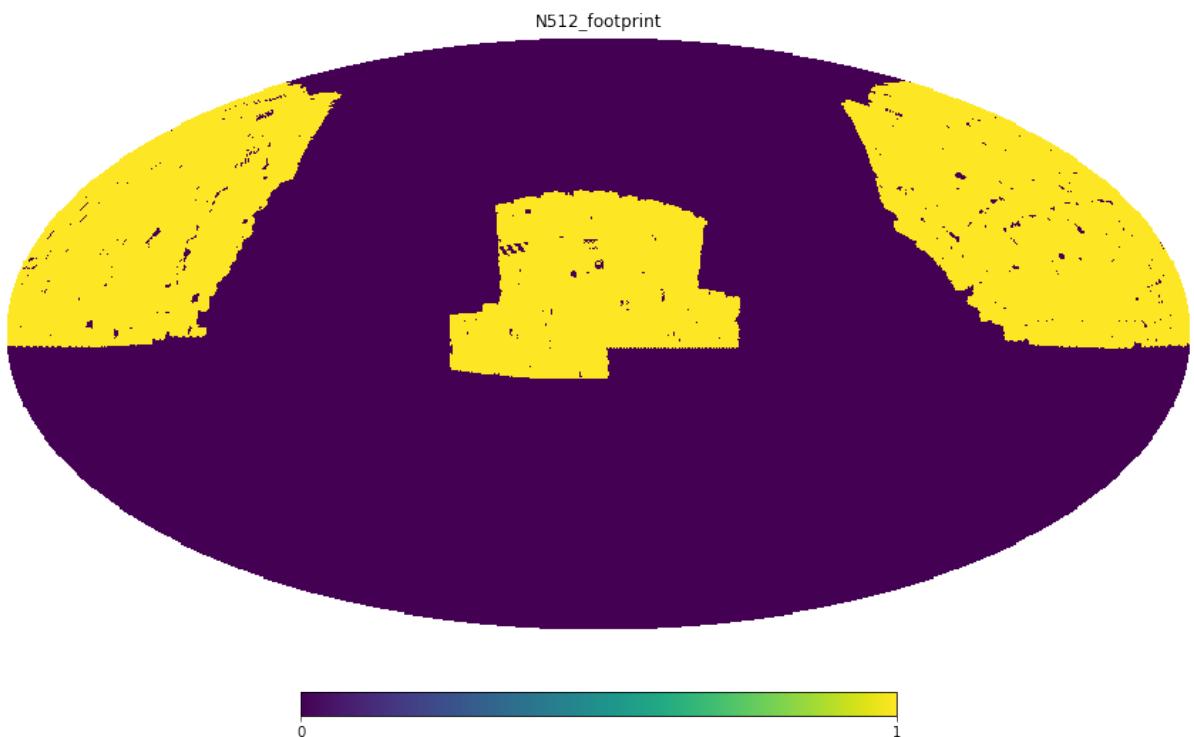


Figura 2: Máscara binária utilizada no estudo na resolução  $NSIDE = 512$  e cobrindo uma porcentagem de aproximadamente 24% do céu. Os valores de cada pixel podem ser zero ou um e correspondem aos dados do LOWZ no norte e no sul do catálogo BOSS.

---

<sup>4</sup><https://data.sdss.org/sas/dr12/boss/lss/>

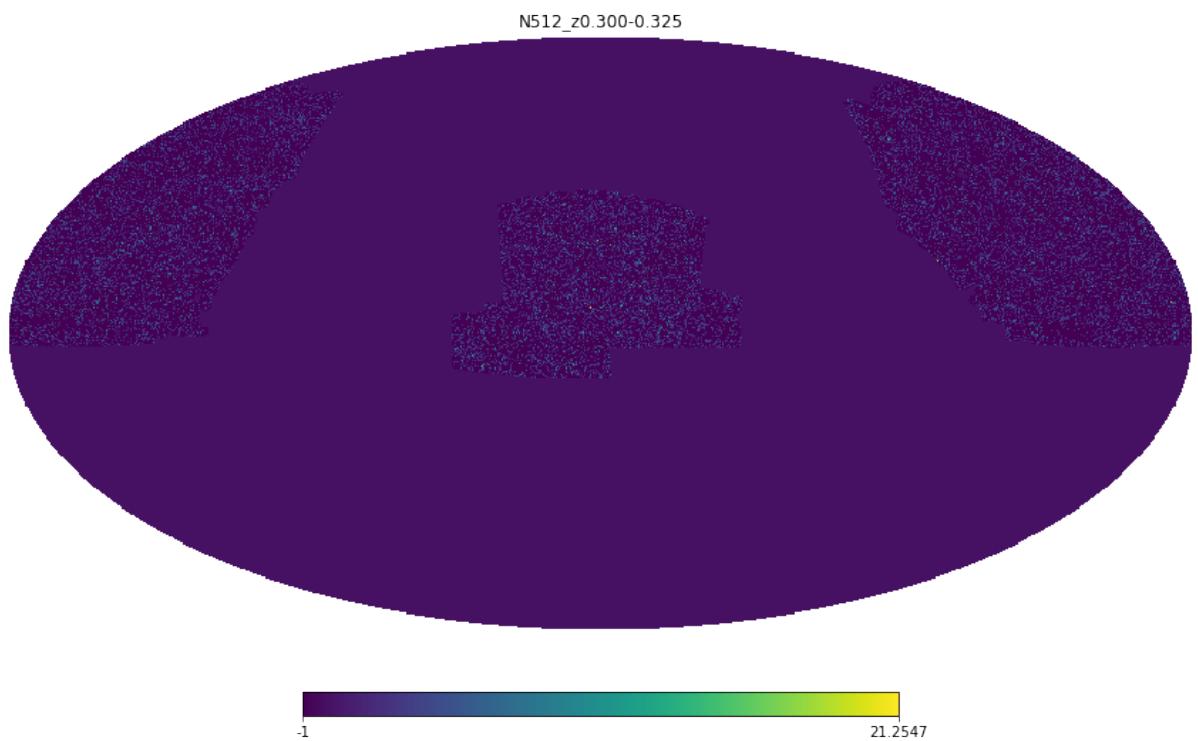


Figura 3: Mapa de sobredensidade do intervalo de redshift  $0.300 < z < 0.325$ . Estes são os mapas que serão mensurados.

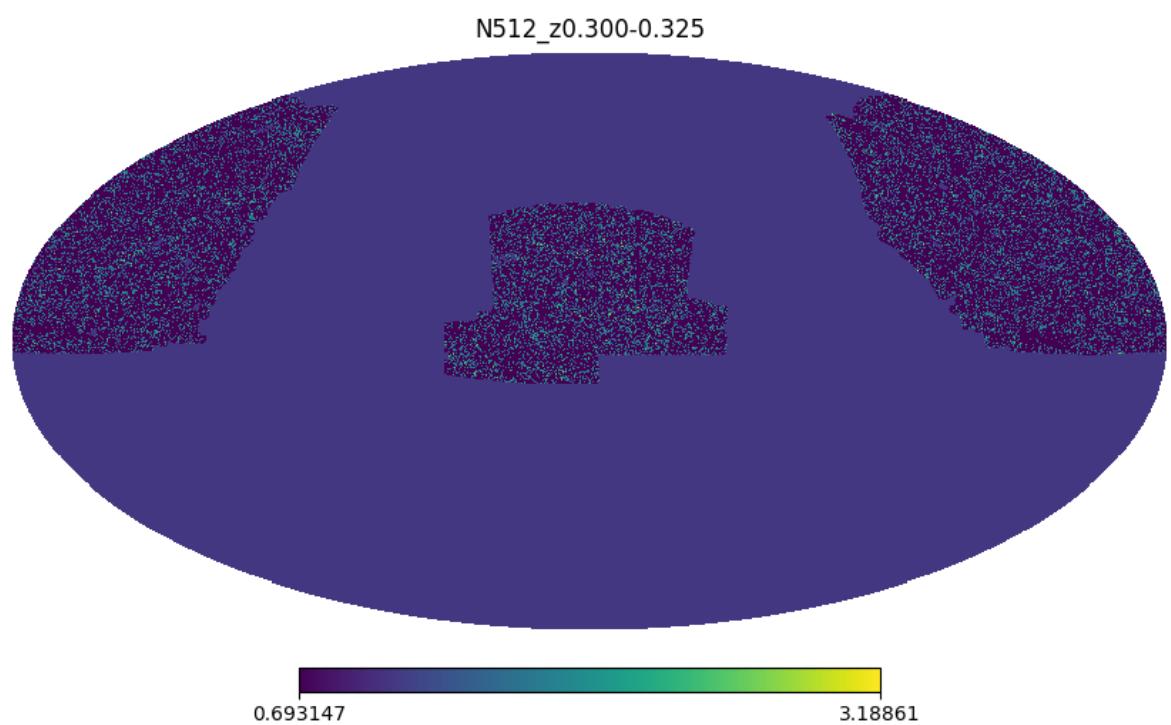


Figura 4: Mapa de sobredensidade do intervalo de redshift  $0.300 < z < 0.325$  em escala logarítmica para melhor visualização da distribuição de objetos.

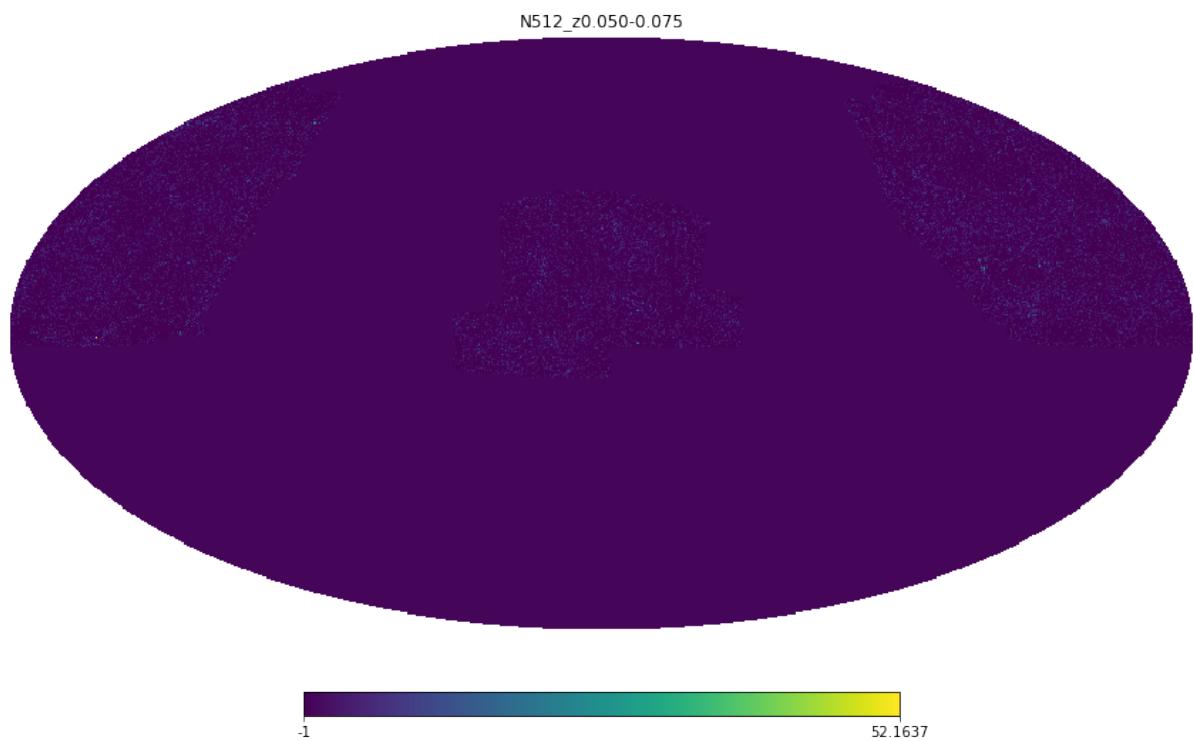


Figura 5: Mapa de sobredensidade do intervalo de redshift  $0.050 < z < 0.075$ .

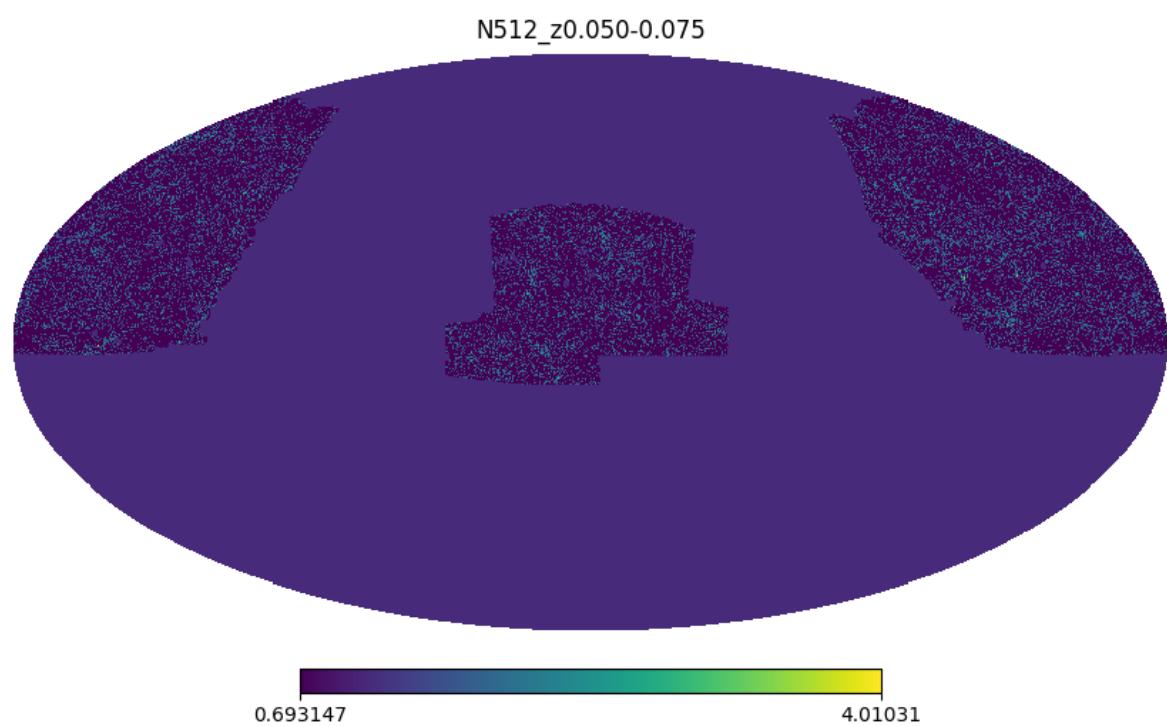


Figura 6: Mapa de sobredensidade do intervalo de redshift  $0.050 < z < 0.075$  em escala logarítmica (para enfatizar a distribuição de objetos).

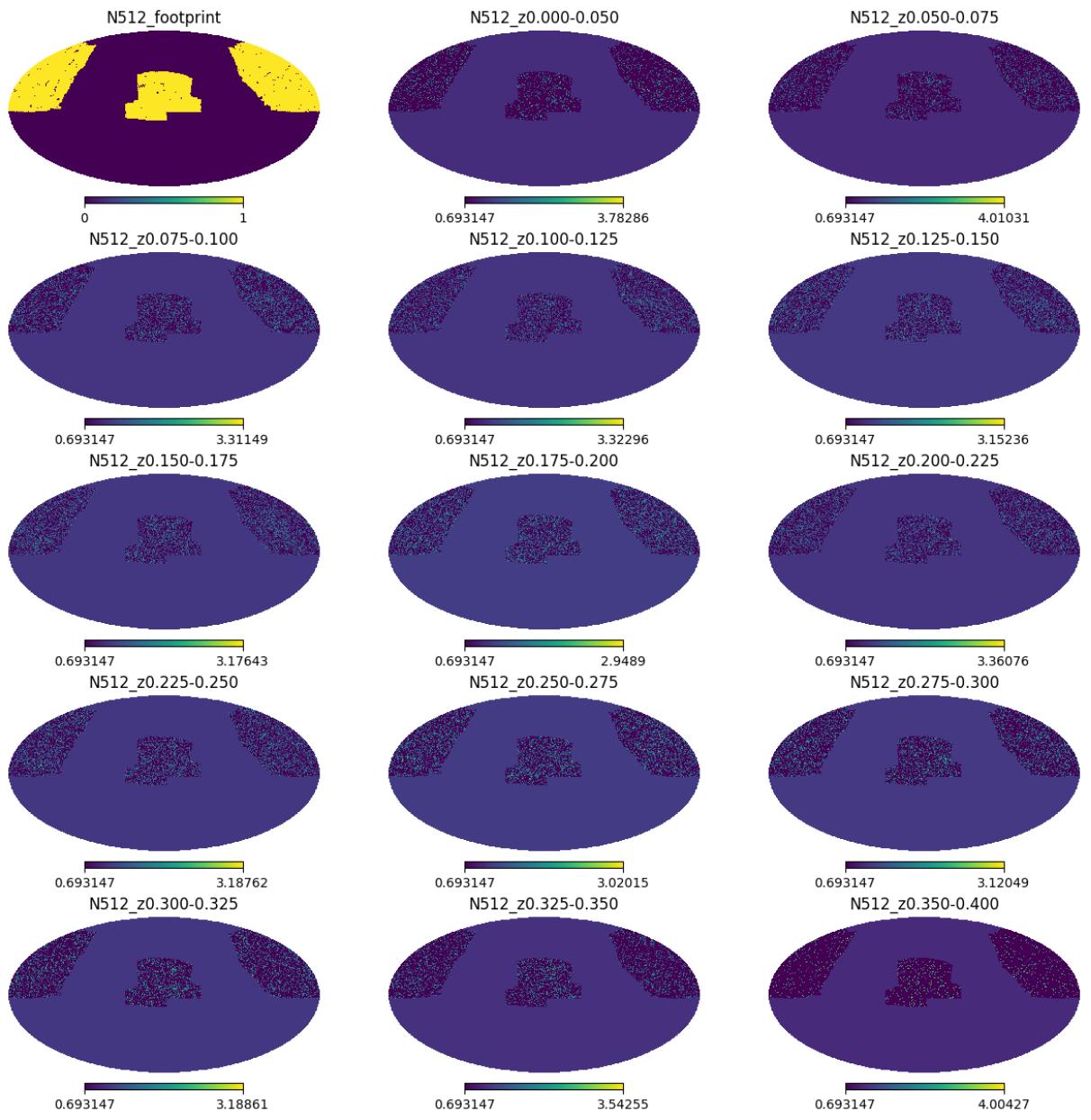


Figura 7: Máscara binária utilizada neste estudo, provida pelo catálogo do BOSS e ajustada por equipes de etapas anteriores na pipeline de dados do BINGO e os mapas de sobredensidade dos quatorze intervalos de redshift escolhidos em escala log para facilitar a visualização da distribuição de objetos.

## 2.0.2 Mapas sistemáticos

Os efeitos sistemáticos são erros inerentes às observações do céu que podem influenciar significativamente os resultados na estimativa de parâmetros cosmológicos, alguns destes são conhecidos e mapeados (mapas sistemáticos). Esses efeitos têm diversas origens, como interferências eletromagnéticas, condições climáticas adversas e limitações dos instrumentos de medição. É crucial mapear e compreender esses efeitos para minimizar seu impacto nas análises.

No contexto do SDSS, alguns dos principais efeitos sistemáticos mapeados incluem:

- - PSF-FWHM (Point Spread Function Full Width at Half Maximum): Mede a largura a meia altura do perfil de espalhamento da imagem, que pode variar devido a aberrações ópticas e turbulência atmosférica. É calculado separadamente para cada filtro fotométrico ( $r, i, u, z$ ).
- - SKYFLUX: Quantifica o nível médio de brilho do céu em cada campo observado, dependendo das condições de iluminação e transparência atmosférica. É calculado separadamente para cada filtro fotométrico ( $g, r, i, u, z$ ).
- - SCORE: É uma métrica combinada que avalia a qualidade geral da imagem, levando em conta parâmetros como a calibração fotométrica, o nível de brilho do céu e a PSF-FWHM. Quanto maior o SCORE, melhor a qualidade da imagem.

Onde  $g, i, r, u, z$  são cores cujos centros da sua distribuição no espectro são descritas na tabela 2.

Esses mapas sistemáticos permitem identificar regiões do céu com observações de qualidade inferior, devido a efeitos como nebulosidade, poluição luminosa ou problemas instrumentais. Ao considerar esses mapas nas análises, é possível obter resultados cosmológicos mais precisos e robustos, minimizando a contaminação por erros sistemáticos.

Portanto, o estudo e mapeamento dos efeitos sistemáticos é uma etapa crucial no processamento de dados astronômicos de grandes levantamentos, como o SDSS. Esses mapas permitem identificar e corrigir potenciais fontes de erro, melhorando a confiabilidade das medidas cosmológicas derivadas desses dados. Informações mais detalhadas podem ser obtidas na documentação do SDSS <sup>5</sup>.

---

<sup>5</sup><https://www.sdss4.org/instruments/camera/#Filters>

<b>Cor</b>	<b>Comprimento de onda (Angstrons)</b>
u	3551
g	4686
r	6166
i	7480
z	8932

Tabela 2: Comprimento de onda da luz, em Angstrons, associado ao o centro da distribuição de cada faixa de frequência nos mapas do SDSS.

### PSF-FWHM (i,r,u,z):

A imagem de um objeto considerado pontual (e.g. uma estrela) não é medido como ponto pelo telescópio, o que se observa é uma distribuição de energia na forma de uma difração circular de uma fonte pontual em uma abertura clara e livre de aberrações do telescópio. A descrição matemática desse efeito é a Point Spread Function (PSF)<sup>6</sup>, que aqui descreve a distribuição normalizada da intensidade medida; para estimar o limite de utilização do que foi medido pode-se utilizar a aproximação Full Width at Half-Maximum (FWHM), que estabelece uma largura que vai da intensidade máxima (centro da distribuição) até a metade do seu valor. Este mapa mede a qualidade das medidas dos objetos, quanto maior o valor de FWHM, menor a qualidade da medida da fonte pontual. Mais detalhes sobre o formalismo podem ser encontrado em Bramich-2012.

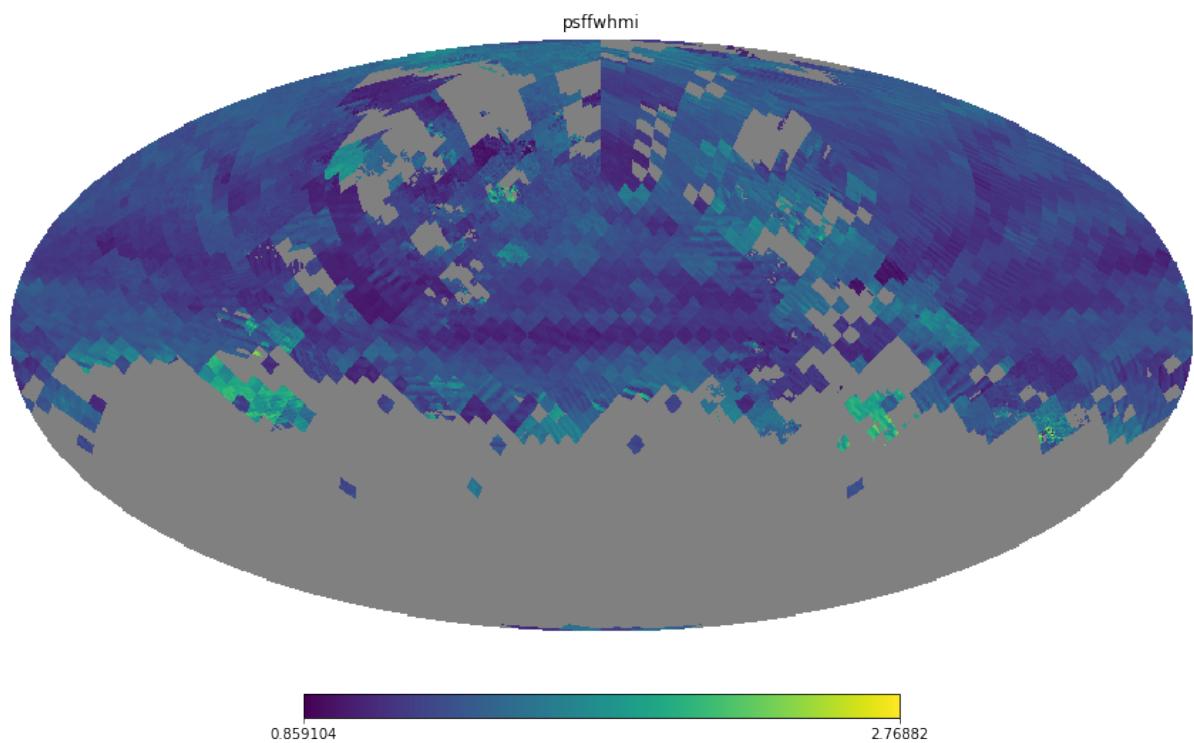


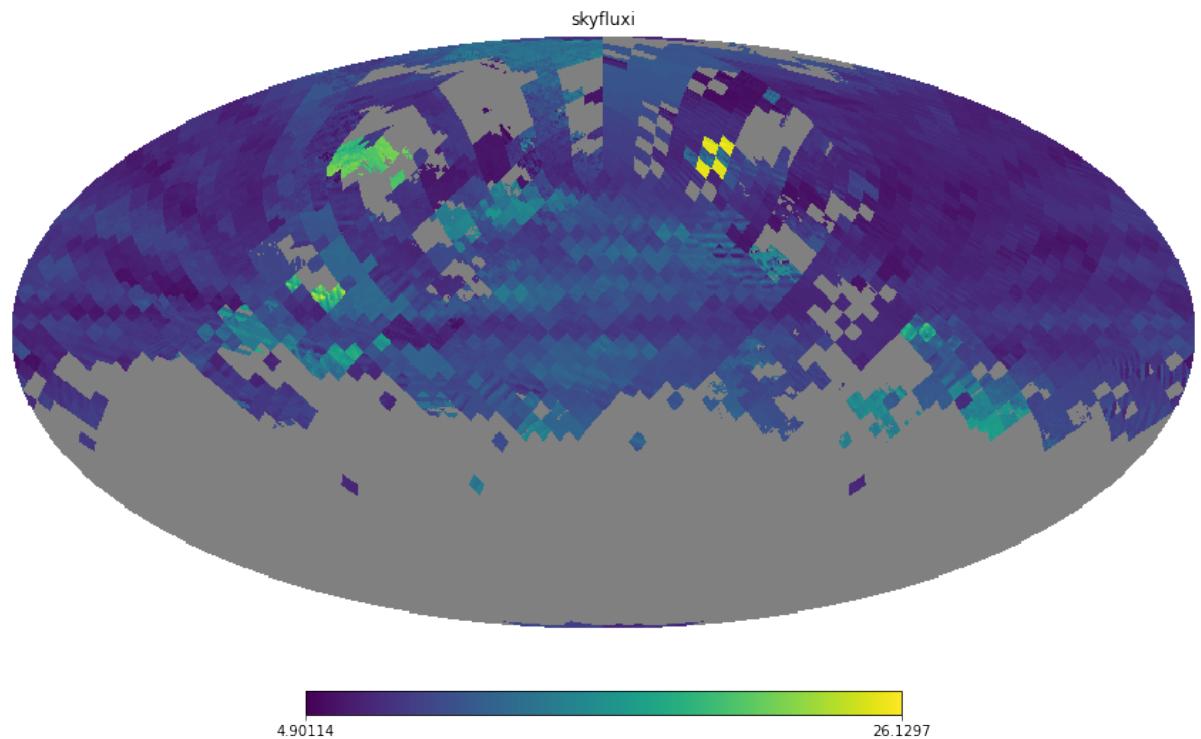
Figura 8: Mapa sistemático PSF-FWHM na banda i. Relaciona a medida óptica do fluxo total na banda i fitando o modelo Point Spread Function (PSF) com a aproximação *Full Width at Half-Maximum* (FWHM) para o objeto medido. A região em cinza não possuem dados de efeitos sistemáticos. O valor de FWHM indica a qualidade da medida das fontes pontuais, quanto maior este valor, mais difusa foi a medida da fonte pontual.

<sup>6</sup><https://www.sdss4.org/dr17/imaging/images/#psf>

## SKYFLUX (g,i,r,u,z):

O mapa sistemático SKYFLUX é uma medida da qualidade das imagens do SDSS (Sloan Digital Sky Survey) <sup>7</sup>. Ele indica o nível mediano de brilho do céu em cada campo observado, dado em nanomaggies por segundo de arco quadrado. O nível de brilho do céu é um fator importante para determinar o limite de fluxo fotométrico em cada campo. Nos filtros ugriz, os níveis medianos de brilho do céu em todos os campos fotométricos primários da pesquisa são, respectivamente, 1.5, 1.7, 4.2, 8.1 e 24.9 nanomaggies <sup>8</sup>.

O SKYFLUX é um dos vários indicadores de qualidade de imagem usados no SDSS, junto com PSF-FWHM e SCORE. Esses valores são armazenados nos arquivos photoField no SAS (Sloan Digital Sky Survey Science Archive Server) e na tabela field no CAS (Catalog Archive Server).



Mapa Skyflux na banda i do SDSS mostra o brilho mediano do céu em cada campo observado, com regiões mais brilhantes em amarelo e áreas menos brilhantes em azul. Essa informação sobre o nível de brilho do céu de fundo é essencial para entender os limites de detecção fotométrica e avaliar a qualidade das medições de fluxo e magnitude dos objetos celestes.

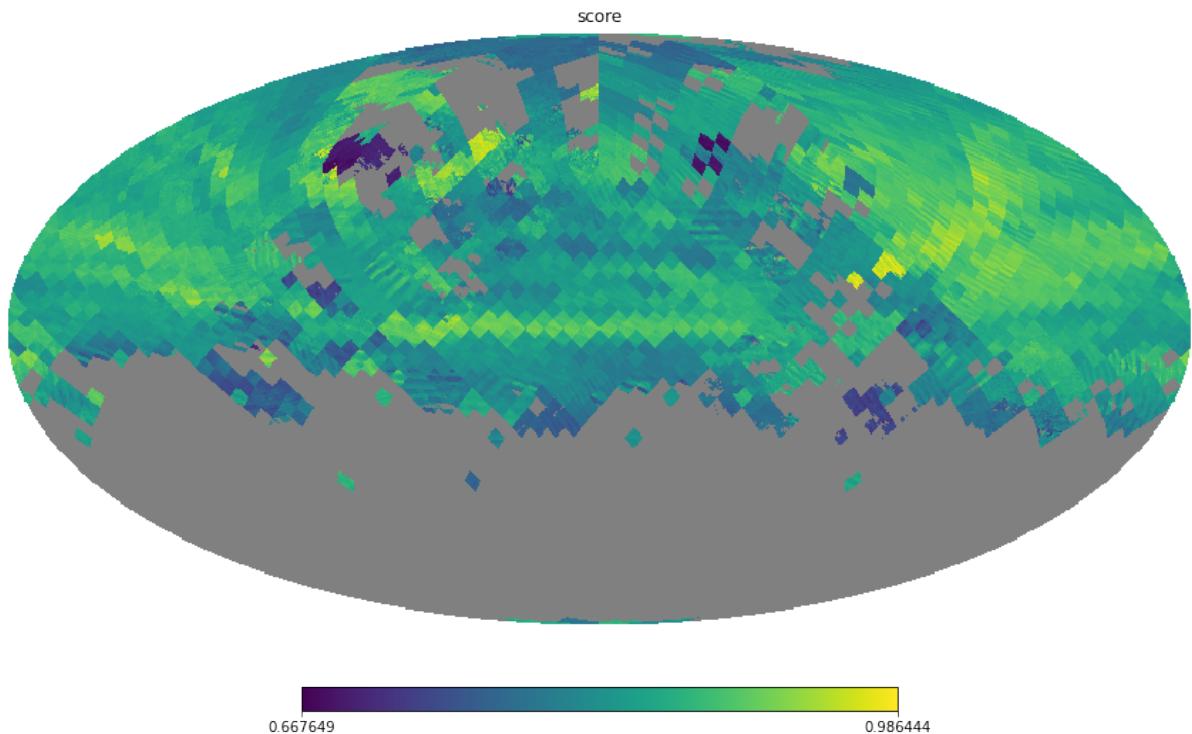
<sup>7</sup><https://live-sdss4org-dr12.pantheonsite.io/algorithms/fluxcal>

<sup>8</sup>[https://www.sdss4.org/dr17/algorithms/image\\_quality/#SKYFLUX](https://www.sdss4.org/dr17/algorithms/image_quality/#SKYFLUX)

## SCORE:

O SCORE é uma métrica combinada que indica a qualidade geral da imagem de cada campo do SDSS<sup>9</sup>. Ele é calculado a partir de diversos parâmetros de qualidade da imagem, como Skyflux que mede o nível médio de brilho do céu no campo; o PSF-FWHM mede a largura a meia altura (FWHM) do perfil de espalhamento da imagem (PSF).

Esses parâmetros são combinados em um único valor, o SCORE, que é usado para determinar qual campo é considerado primário em uma determinada região do céu. Quanto maior o SCORE, melhor a qualidade da imagem.



Mapa sistemático SCORE. Expressa a qualidade observational de determinada região. O valor do score varia entre 0-1. Esta medida é definida pelo SDSS como “medida de qualidade” e engloba diversas métricas em sua composição. As regiões amarelas são onde a métrica indica maior qualidade nas observações e a região em azul indica menor qualidade. A região cinza é onde não há valores associados ao filtro.

<sup>9</sup>[https://www.sdss4.org/dr17/algorithms/image\\_quality/#SCORE](https://www.sdss4.org/dr17/algorithms/image_quality/#SCORE)

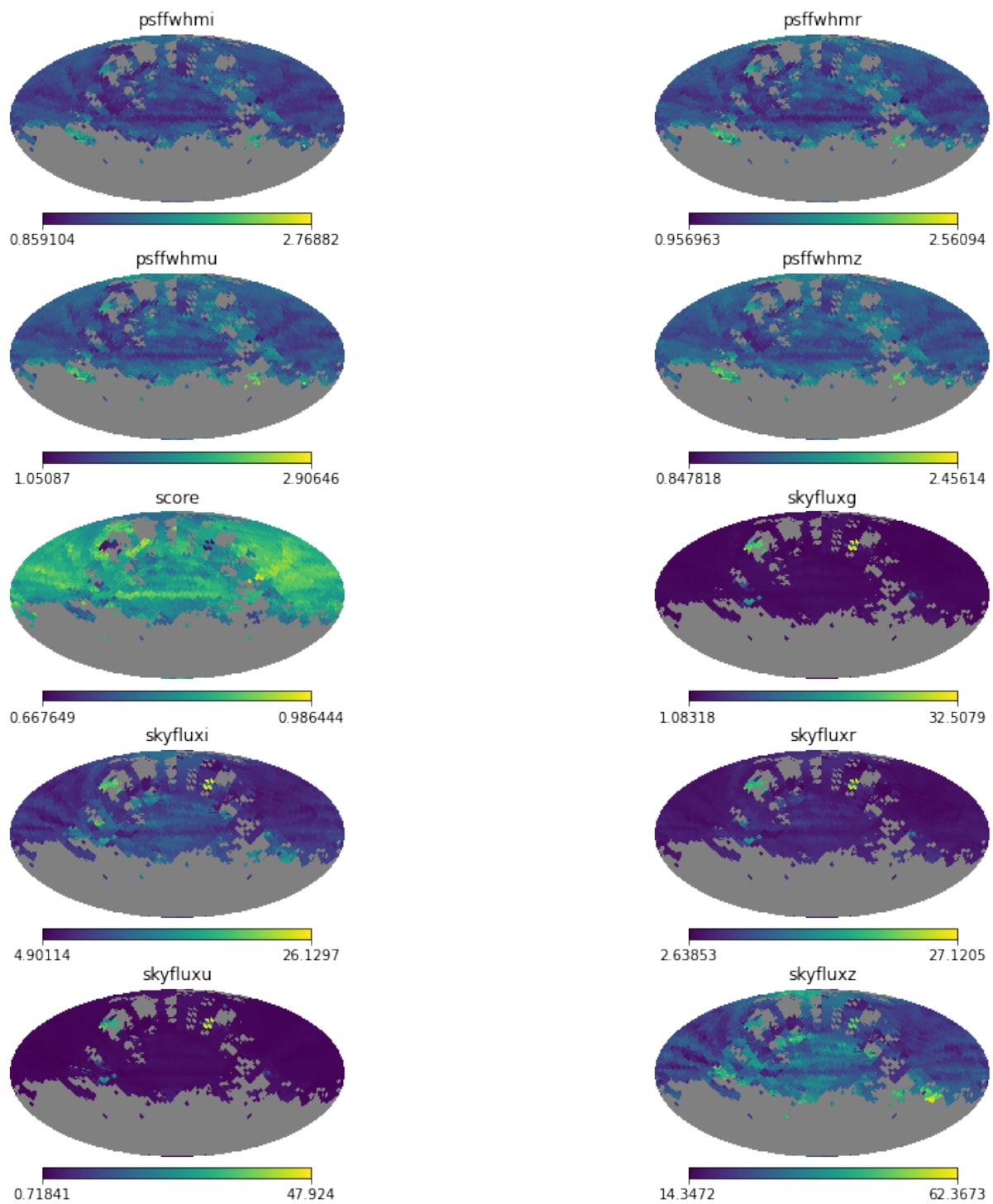


Figura 9: Mapas dos efeitos sistemáticos. PSF-FWHM nas bandas de frequência i,r,u,z; SKYFLUX nas bandas de frequência g,i,r,u,z e SCORE.

### 3 Espectro de potência

O espectro de potência  $P(k)$ , ou ainda, densidade espectral de energia, faz a conexão entre cosmologia observacional e os modelos teóricos. Os espectros de potência (*Power Spectrum, PS*) dos mapas contém a informação sobre como os objetos estão distribuídos e como essa distribuição evolui em função do redshift. Permite-nos também tirar conclusões a respeito das Leis Físicas de formação das estruturas observadas através de modelos cosmológicos. No nosso caso, estamos interessados em medir o espectro de potência da matéria. Mais especificamente, das galáxias. Para esse fim é necessário medir a distribuição de objetos em mapas cosmológicos, o que significa, do ponto de vista técnico, estimar o espectro de potência de uma esfera celeste discreta (pixelada) bidimensional.

Com o objetivo de obter um estimador para o espectro de potência, que no nosso caso será o Pseudo Espectro de Potência (Pseudo-Power Spectrum ou Pseudo- $C\ell$ ), de um modo inteligível, vamos primeiro descrever a função de correlação e fazer uma conexão com espectro de potência para finalmente abordar o estimador Pseudo- $C\ell$  para o caso discreto.

#### 3.1 Função de correlação de 2-pontos

Com o propósito de modelar as flutuações na homogeneidade do universo em largas escalas, acrescentamos uma perturbação na densidade total de matéria  $\rho(\vec{x})$  da forma

$$\begin{aligned}\rho(\vec{x}) &= \langle \rho(\vec{x}) \rangle + \delta\rho(\vec{x}) , \\ \rho(\vec{x}) &= \langle \rho(\vec{x}) \rangle + \langle \rho(\vec{x}) \rangle \delta(\vec{x}) , \\ \rho(\vec{x}) &= \langle \rho(\vec{x}) \rangle [1 + \delta(\vec{x})] ,\end{aligned}\tag{17}$$

onde  $\langle \rho(\vec{x}) \rangle$  é a função de 1-ponto, que, em particular, é o valor esperado da densidade de matéria, ou ainda, o valor médio.  $\delta(\vec{x})$  é a sobredensidade, definida como

$$\delta(\vec{x}) \equiv \frac{\rho(\vec{x}) - \langle \rho(\vec{x}) \rangle}{\langle \rho(\vec{x}) \rangle} = \frac{\rho(\vec{x})}{\langle \rho(\vec{x}) \rangle} - 1 .\tag{18}$$

Deste modo, o valor esperado para  $\delta(\vec{x})$  será

$$\langle \delta(\vec{x}) \rangle = \frac{\langle \rho(\vec{x}) \rangle}{\langle \rho(\vec{x}) \rangle} - 1 = 0 .\tag{19}$$

Podemos, então, calcular a função de 2-pontos, ou função de correlação cruzada (*cross-correlation*) da densidade de matéria em duas regiões distintas determinadas pelas coordenadas  $\vec{x}_1$  e  $\vec{x}_2$ ,

$$\langle \rho(\vec{x}_1) \rho(\vec{x}_2) \rangle = \langle \{ \langle \rho(\vec{x}_1) \rangle [1 + \delta(\vec{x}_1)] \} \{ \langle \rho(\vec{x}_2) \rangle [1 + \delta(\vec{x}_2)] \} \rangle .\tag{20}$$

Evocando o princípio cosmológico de homogeneidade e isotropia, no caso em que a distância entre  $\vec{x}_1$  e  $\vec{x}_2$  aumenta, estes valores se tornam independentes devido à invariância por translação nas devidas escalas e convergem para o mesmo valor esperado  $\langle \rho(\vec{x}) \rangle$ , de modo que

$$\langle \rho(\vec{x}_1) \rho(\vec{x}_2) \rangle \Rightarrow \langle \rho(\vec{x}_1) \rangle \langle \rho(\vec{x}_2) \rangle = \bar{\rho}^2 .\tag{21}$$

A partir desta hipótese, podemos simplificar a equação (21),

$$\begin{aligned}\langle \rho(\vec{x}_1)\rho(\vec{x}_2) \rangle &= \bar{\rho}^2 \langle 1 + \delta(\vec{x}_1)\delta(\vec{x}_2) + \delta(\vec{x}_1) + \delta(\vec{x}_2) \rangle , \\ &= \bar{\rho}^2 [1 + \langle \delta(\vec{x}_1)\delta(\vec{x}_2) \rangle + \langle \delta(\vec{x}_1) \rangle + \langle \delta(\vec{x}_2) \rangle] , \\ &= \bar{\rho}^2 [1 + \langle \delta(\vec{x}_1)\delta(\vec{x}_2) \rangle] , \\ \langle \rho(\vec{x}_1)\rho(\vec{x}_2) \rangle &= \bar{\rho}^2 [1 + \xi(\vec{x}_1, \vec{x}_2)] ,\end{aligned}\quad (22)$$

onde definimos a função de correlação de 2-pontos da sobredensidade  $\xi(\vec{x}_1, \vec{x}_2)$  (comumente nomeada de *reduced ou connected 2-point correlation function*),

$$\xi(\vec{x}_1, \vec{x}_2) \equiv \langle \delta(\vec{x}_1)\delta(\vec{x}_2) \rangle . \quad (23)$$

Deste modo, a informação sobre as flutuações na distribuição de densidade passam a estar contidas na função de correlação de 2-pontos da sobredensidade, que será nomeada simplesmente de função de correlação daqui em diante. Novamente nos munindo do princípio cosmológico, a hipótese de homogeneidade nos fornece uma importante informação: a função de correlação deve depender apenas da distância entre  $\vec{x}_1$  e  $\vec{x}_2$ ,

$$\xi(\vec{x}_1, \vec{x}_2) = \xi(\vec{x}_1 - \vec{x}_2) . \quad (24)$$

Assumindo a hipótese da isotropia, de que não há direções preferidas, a função de correlação deve ser invariante por rotações globais, de modo que podemos afirmar que

$$\xi(\vec{x}_1, \vec{x}_2) = \xi(|\vec{x}_1 - \vec{x}_2|) . \quad (25)$$

Fazendo a mudança de variáveis

$$\vec{r} \equiv \vec{x}_1 - \vec{x}_2 , \quad (26)$$

$$\vec{x} \equiv \vec{x}_1 , \quad (27)$$

podemos escrever a função de correlação de uma maneira mais adequada,

$$\xi(\vec{r}) = \langle \delta(\vec{x})\delta(\vec{x} + \vec{r}) \rangle . \quad (28)$$

### 3.2 Espectro de potência

Podemos representar a função de correlação (perturbação) no espaço de Fourier, deste modo podemos definir o espectro de potência  $P(\vec{k})$ . Iniciando pela representação da sobredensidade no espaço de Fourier,

$$\delta(\vec{k}) = \frac{1}{(2\pi)^3} \int d^3\vec{x} \delta(\vec{x}) e^{-i\vec{k}\cdot\vec{x}} , \quad (29)$$

$$\delta(\vec{x}) = \frac{1}{(2\pi)^3} \int d^3\vec{k} \delta(\vec{k}) e^{i\vec{k}\cdot\vec{x}} . \quad (30)$$

Considerando que a sobredensidade  $\delta(\vec{x})$  seja real, temos que  $\delta(\vec{x}) = \delta^\dagger(\vec{x})$ , o que implica que  $\delta^\dagger(\vec{k}) = \delta(-\vec{k})$ , onde o operador dagger ( $\dagger$ ) representa o conjugado complexo. A fim

de definir o espectro de potência  $P(\vec{k})$ , podemos escrever

$$\begin{aligned}\left\langle \delta(\vec{k})\delta^\dagger(\vec{k}') \right\rangle &= \left\langle \int d^3\vec{r} e^{-i\vec{k}\cdot\vec{r}} \delta(\vec{r}) \int d^3\vec{x} e^{i\vec{k}'\cdot(\vec{r}+\vec{x})} \delta(\vec{r}+\vec{x}) \right\rangle , \\ &= \int d^3\vec{r} e^{-i(\vec{k}-\vec{k}')\cdot\vec{r}} \int d^3\vec{x} e^{i\vec{k}'\cdot\vec{x}} \langle \delta(\vec{r})\delta(\vec{r}+\vec{x}) \rangle , \\ \left\langle \delta(\vec{k})\delta^\dagger(\vec{k}') \right\rangle &= \int d^3\vec{r} e^{-i(\vec{k}-\vec{k}')\cdot\vec{r}} \int d^3\vec{x} e^{i\vec{k}'\cdot\vec{x}} \xi(\vec{x}) .\end{aligned}\quad (31)$$

Definindo a delta de Dirac como

$$(2\pi)^3\delta_D(\vec{k}-\vec{k}') = \int d^3\vec{r} e^{-i(\vec{k}-\vec{k}')\cdot\vec{r}} , \quad (32)$$

temos que

$$\left\langle \delta(\vec{k})\delta^\dagger(\vec{k}') \right\rangle = \int d^3\vec{x} \xi(\vec{x}) e^{i\vec{k}'\cdot\vec{x}} (2\pi)^3\delta_D(\vec{k}-\vec{k}') , \quad (33)$$

$$\left\langle \delta(\vec{k})\delta^\dagger(\vec{k}') \right\rangle = P(\vec{k}') (2\pi)^3\delta_D(\vec{k}-\vec{k}') . \quad (34)$$

onde  $P(\vec{k})$  é definido como o espectro de potência, que em essência é a transformada de Fourier da função de dois pontos da sobredensidade,

$$P(\vec{k}) = \int d^3\vec{x} \xi(\vec{x}) e^{-i\vec{k}\cdot\vec{x}} . \quad (35)$$

Tomando a transformada inversa, obtemos novamente a função de correlação,

$$\xi(\vec{x}) = \int d^3\vec{k} P(\vec{k}) e^{i\vec{k}\cdot\vec{x}} . \quad (36)$$

Até agora estamos descrevendo o caso  $\delta(\vec{k}, z = 0)$ . Em escalas lineares, podemos supor que a função de crescimento  $D(z)$  pode ser escrita de forma separável e, portanto,

$$\delta(\vec{k}, z) = D(z) \delta(\vec{k}, 0) , \quad (37)$$

de modo que a função de 2-pontos se torna

$$\left\langle \delta(\vec{k}, z)\delta^\dagger(\vec{k}', z') \right\rangle = (2\pi)^3\delta(\vec{k}-\vec{k}') D(z)D^\dagger(z) P(\vec{k}') . \quad (38)$$

É importante notar que o campo  $\delta(\vec{k})$  mede a sobredensidade de matéria total, no entanto, sob a hipótese da matéria escura, a sobredensidade precisa ser considerada como a contribuição de dois termos,

$$\delta = \delta_{dm} + \delta_b , \quad (39)$$

onde  $\delta_{dm}$  é a contribuição relacionada à matéria escura e  $\delta_b$  é a contribuição relacionada à matéria bariônica. Para efetuar essa correção, define-se a função *bias*  $b(z)$ , como sendo a razão entre a flutuação bariônica e a flutuação total, que no regime linear depende apenas do redshift,

$$\delta_b(\vec{k}, z) = b(z)\delta(\vec{k}, z) . \quad (40)$$

O espectro de potência da matéria bariônica pode ser escrito como

$$P_b(\vec{k}, z) = b(z)^2 P(\vec{k}, z) . \quad (41)$$

### 3.3 Poisson shot noise

O ruido denominado *shot noise* é uma particularidade do campo discretizado. Iniciando pelo campo de sobredensidade continuo definido como

$$\delta(\vec{x}) = \frac{n(\vec{x}) - \langle n(\vec{x}) \rangle}{\langle n(\vec{x}) \rangle}. \quad (42)$$

Podemos representar um conjunto discreto de objetos  $n(\vec{x})$  como somas de deltas de Dirac em cada posição  $\vec{r}_i$  com  $n_i$  partículas,

$$n(\vec{x}) = \sum_i^N n_i \delta_D(\vec{x} - \vec{r}_i). \quad (43)$$

Substituindo este resultado na definição do campo de sobredensidade  $\delta(\vec{x})$  e definindo  $\langle n(\vec{x}) \rangle = \bar{n}$  temos

$$\delta(\vec{x}) = \frac{\sum_i n_i \delta_D(\vec{x} - \vec{r}_i) - \bar{n}}{\bar{n}}. \quad (44)$$

realizamos agora a transformada de Fourier para o caso discreto,

$$\delta(\vec{k}) = \int d^3x \left[ \frac{\sum_i^N n_i \delta_D(\vec{x} - \vec{r}_i) - \bar{n}}{\bar{n}} \right] e^{i\vec{k}\cdot\vec{x}}, \quad (45)$$

$$\delta(\vec{k}) = \frac{1}{\bar{n}} \int d^3x \left[ \sum_i^N n_i \delta_D(\vec{x} - \vec{r}_i) \right] e^{i\vec{k}\cdot\vec{x}} - \frac{1}{\bar{n}} \int d^3x \bar{n} e^{i\vec{k}\cdot\vec{x}}. \quad (46)$$

A primeira integral é resolvida utilizando-se a propriedade de filtragem da delta de Dirac, enquanto o segundo termo a direita integrado por todo espaço resulta em uma delta de Dirac  $\delta^D(\vec{k})$ , mas como estamos no caso discreto, o resultado é descrito por um delta de Kronecker,  $\delta_{\vec{k},0}^K$ , de modo que no espaço de Fourier discreto o campo de sobredensidade é descrito por

$$\delta(\vec{k}) = \frac{1}{\bar{n}} \sum_i^N n_i e^{i\vec{k}\cdot\vec{r}_i} - \delta_{\vec{k},0}^K. \quad (47)$$

Acima,  $\bar{n}$  é o número de objetos por esferorradiano e  $\delta^K$  é um delta de Kronecker, pois agora estamos trabalhando espaço de Fourier discreto, enquanto a delta de Dirac é apropriada ao espaço continuo. Está sendo utilizada a hipótese de que o volume onde a expansão é considerada é de tal modo que  $n_i$  possa assumir apenas dois valores  $n_i \in \{0, 1\}$  e, portanto,  $\langle n_i \rangle = \langle n_i^2 \rangle$ . Podemos então calcular a função de 2-pontos no espaço de Fourier,

$$\begin{aligned} \langle \delta_d(\vec{k}) \delta_d^\dagger(\vec{k}') \rangle &= \frac{1}{\bar{n}^2} \left\langle \left( \sum_i n_i e^{i\vec{k}\cdot\vec{r}_i} - \delta_{\vec{k},0}^K \right) \left( \sum_j n_j e^{-i\vec{k}'\cdot\vec{r}_j} - \delta_{\vec{k}',0}^K \right) \right\rangle, \\ &= \frac{1}{\bar{n}^2} \left\langle \sum_i \sum_j n_i n_j e^{i\vec{k}\cdot\vec{r}_i} e^{-i\vec{k}'\cdot\vec{r}_j} - \sum_i n_i e^{i\vec{k}\cdot\vec{r}_i} \delta_{\vec{k}',0}^K \right\rangle - \\ &\quad \frac{1}{\bar{n}^2} \left\langle \sum_j n_j e^{-i\vec{k}'\cdot\vec{r}_j} \delta_{\vec{k},0}^K + \delta_{\vec{k},0}^K \delta_{\vec{k}',0}^K \right\rangle, \\ \langle \delta_d(\vec{k}) \delta_d^\dagger(\vec{k}') \rangle &= \frac{1}{\bar{n}^2} \left\langle \sum_i \sum_j n_i n_j e^{i\vec{k}\cdot\vec{r}_i} e^{-i\vec{k}'\cdot\vec{r}_j} - \delta_{\vec{k},0}^K \delta_{\vec{k}',0}^K \right\rangle. \end{aligned} \quad (48)$$

Supondo que  $\mathbf{k}$  ou  $\mathbf{k}' \neq 0$  o produto das deltas de Kronecker se anulam. Feito isso, separamos o somatório para o caso  $i \neq j$  e  $i = j$ , para este ultimo, toma-se o limite para o volume, desta forma,

$$\begin{aligned} \langle \delta_d(\vec{k})\delta_d^\dagger(\vec{k}') \rangle &= \frac{1}{\bar{n}^2} \sum_{i \neq j} \langle n_i n_j \rangle e^{i\vec{k} \cdot \vec{r}_i} e^{-i\vec{k}' \cdot \vec{r}_j} + \frac{1}{\bar{n}^2} \sum_i \langle n_i^2 \rangle e^{i(\vec{k} - \vec{k}') \cdot \vec{r}_i}, \\ \langle \delta_d(\vec{k})\delta_d^\dagger(\vec{k}') \rangle &= \langle \delta(\vec{k})\delta^\dagger(\vec{k}') \rangle + \frac{\delta^K_{\vec{k}, \vec{k}'}}{\bar{n}}. \end{aligned} \quad (49)$$

Para o caso discreto, o espectro de potência  $P_d(\vec{k})$  possui dois termos, o primeiro termo das equações mede as flutuações de densidade em pontos distintos, ou seja,  $i \neq j$ . O segundo termo é o caso  $i = j$  e está relacionado à contagem de elementos proveniente do processo de discretização. Esse termo é um ruido de contagem e é denominado *shot noise*,

$$P_d(\vec{k}) = P(\vec{k}) + \frac{\delta^K_{\vec{k}, \vec{k}'}}{\bar{n}}. \quad (50)$$

### 3.4 Estimador Pseudo- $C_\ell$

Pseudo- $C_\ell$  (PCL) é um estimador para o espectro de potência introduzido por [Peebles-1973](#) e descrita com detalhes por [Loureiro-2019](#). O PCL não é o estimador com a menor variância disponível, entretanto por ser computacionalmente mais eficiente é a escolha adequada para nossa finalidade.

O espectro de potência continuará sendo descrito pelo campo de sobredensidade, expandido agora em termos dos harmônicos esféricicos, em seguida será considerado o caso onde a observação do sinal é parcial no céu e finalmente o formalismo será refeito para o caso em que o campo de sobredensidade é um sinal discreto (esfera pixelada).

Primeiro, expande-se o campo de sobredensidade em coordenadas esféricicas  $\delta(\theta, \phi)$  em termos dos harmônicos esféricicos  $Y_{\ell m}(\theta, \phi)$ ,

$$\delta(\theta, \phi) = \sum_{\ell=0}^{\ell_{max}} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\theta, \phi). \quad (51)$$

Os coeficientes  $a_{\ell m}$  são descritos pela integral

$$a_{\ell m} = \int d\Omega \delta(\theta, \phi) Y_{\ell m}^\dagger(\theta, \phi). \quad (52)$$

Acima,  $\Omega = (\theta, \phi)$  é a integral nas coordenadas esféricicas e o operador dagger ( $\dagger$ ) representa o complexo conjugado. O estimador normalizado para o espectro de potência  $C_\ell$  é escrito como

$$C_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} a_{\ell m} a_{\ell m}^\dagger. \quad (53)$$

Para o caso em que o sinal faz parte de uma observação parcial do céu (vide Fig(100)), precisamos corrigir a medida estimada para o espectro de potência. Para isso integramos sobre a área  $\Omega_{tot}$  referente,

$$J_{\ell m} = \int_{\Omega_{tot}} d\Omega |Y_{\ell m}|^2. \quad (54)$$

Deste modo, o valor de  $J_{\ell m} = 1$  é o caso de uma observação de céu total, e  $J_{\ell m} < 1$  uma observação parcial.

Podemos introduzir o espectro de potência entre dois campos de sobredensidade  $(i, j)$  com índices distintos ( $i \neq j$ ) ou iguais ( $i = j$ ),

$$C_{\ell}^{ij} = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} C_{\ell m}^{ij}. \quad (55)$$

O valor  $C_{\ell m}^{ij}$  é o produto de dois coeficientes  $a_{\ell m}$  de índices de sobredensidade quaisquer corrigidos por  $J_{\ell m}$ ,

$$C_{\ell m}^{ij} = \frac{\mathbb{R}(a_{\ell m}^i a_{\ell m}^{j\dagger})}{J_{\ell m}}. \quad (56)$$

Considerando agora o caso em que o campo de sobredensidade é discreto, ou seja, o caso da esfera pixelada (p), escrevemos a sobredensidade  $\delta_p$  como

$$\delta_p = \frac{n_p}{\Delta\Omega_p} \frac{\Delta\Omega_{tot}}{n_{tot}} - 1. \quad (57)$$

Aqui,  $n_p$  é o número de objetos por pixel  $p$ ,  $n_{tot}$  é o total de objetos contados no mapa,  $\Delta\Omega_p$  é o ângulo sólido para um pixel  $p$  e  $\Delta\Omega_{tot}$  é o angulo sólido total do mapa em questão. Esse mapa, pixelado e com resolução menor comparado ao campo de sobredensidade real (inacessível) é nomeado suavizado.

Os coeficientes  $a_{\ell m}$  são escritos como a soma sobre todos os pixels  $p$  do mapa,

$$a_{\ell m} = \sum_p \delta_p Y_{\ell m}^\dagger(\theta_p, \phi_p) \Delta\Omega_p. \quad (58)$$

Aqui, as coordenadas  $\theta_p$  e  $\phi_p$  se referem ao centro do pixel  $p$  em questão e  $\Delta\Omega_p$  é a área de  $p$ .

A princípio, o formalismo não restringe um máximo para os multipolos da expansão, no entanto quando se trata de um sinal discreto essa limitação existe pela própria natureza dos dados, o que adiciona uma distorção no estimador do espectro de potência, que é diretamente dependente da resolução do mapa suavizado em questão.

A distorção em relação ao espectro de potência não pixelado  $C_\ell$  e o espectro com efeito de suavização  $C_\ell^p$  é escrito como <sup>10</sup>

$$C_\ell^p = w_\ell^2 C_\ell. \quad (59)$$

Deste modo,  $w_\ell$  é escrito como

$$w_\ell^2 = \frac{1}{N_{pixel}(2\ell+1)} \sum_{p,m} |w_{\ell m}^p|^2. \quad (60)$$

O valor  $w_{\ell m}^p$  é a transformada em harmônicos esféricos para cada pixel,

$$w_{\ell m}^p = \frac{\int_p d\Omega Y_{\ell m}(\theta, \phi)}{Y_{\ell m}(\theta_p, \phi_p) \Delta\Omega_p}. \quad (61)$$

---

<sup>10</sup><https://healpix.jpl.nasa.gov/html/intronode14.htm>

Uma dedução mais detalhada pode ser encontrada em [Leistedt-2013](#).

Separando o estimador  $C_\ell$  em uma parcela do espectro de potência  $S_\ell$  e uma de ruido (shotnoise)  $N_\ell$ ,

$$C_\ell = S_\ell + N_\ell , \quad (62)$$

onde  $N_\ell$  é dado por

$$N_\ell = \frac{\Delta\Omega_{tot}}{n_{tot}} = \frac{1}{\bar{n}}, \quad (63)$$

onde  $\bar{n}$  é o número médio de galáxias por esterradianos. O estimador para o espectro de potência sem o shotnoise é dado pela expressão

$$S_\ell^{ij} = \frac{1}{w_\ell^2} \left[ \left( \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} C_{\ell m}^{ij} \right) - N_\ell \delta_{ij} \right] . \quad (64)$$

### 3.5 Resultados: Espectro de potência dos mapas cosmológicos

Serão realizadas três estimativas para o espectro de potência dos mapas cosmológicos afim de comparar os resultados e aumentar a confiança nos mesmos.

Serão utilizados os códigos [Pseudo-Power](#), a biblioteca [Healpy](#)<sup>11</sup> e o [UCLCL](#) ([Loureiro-2019](#)); para os dois primeiros mediremos o espectro de potência diretamente dos mapas pixelados e o UCLCL estimará um espectro de potência teórico, baseado na distribuição dos dados em função do redshift.

Nas próximas sessões o Pseudo-Power será utilizado como base para a construção de uma pipeline que automatizará essas médias de correlação e funcionará juntamente com o [FLASK](#) (simulador de mapas cosmológicos) para estimar um intervalo de confiança para os dados e produzir uma matriz de covariância; esta pipeline é um código escrito em Python e é a principal contribuição deste trabalho na pipeline do projeto BINGO.

#### 3.5.1 Pseudo-Power

[Pseudo-Power](#) é um código multifuncional que, nesse estágio do trabalho, será usado para calcular o estimador Pseudo- $C_\ell$  dos mapas cosmológicos, realizando todas as etapas descritas nos capítulos anteriores. De forma prática, as funções do Pseudo-Power serão executadas na seguinte ordem:

```
1 $ cd PATH/Pseudo-Power # Pasta de instalacao do Pseudo-Power.
2 $ ./IImJlm args*
3 $ ./Map2Alm args*
4 $ ./Alm2Cl args*
```

Listing 1: Ordem de execução das funções do Pseudo-Power para o cálculo do espectro de potência.

A linha 2 está relacionada com a equação 54, a linha 3 com a equação 58 e a linha 4 com as equações 56 e 64. O termo args\* refere-se aos argumentos necessários para cada etapa, podendo ser obtidos usando o comando -h, como por exemplo,

---

<sup>11</sup>Healpy docs: <https://healpy.readthedocs.io/en/latest/>

```

1 $ ./IlmJlm -h
2 # OUTPUT:
3 # Options:
4 # -O output: name of the 'outputfile' in ascii format
5 # -m mask: 'inputfile' of mask in fits format
6 # -N nside:'nside' of the mask and map
7 # -l Lmin: 'lmin'; start with 0 for complete runs;
8 # -L Lmax: 'lmax'

```

Listing 2: Obtendo argumentos necessários para cada funcionalidade do Pseudo-Power.

O que ficaria da seguinte forma,

```

1 $ ./IlmJlm -O "PATH/IlmJlm.dat" -m "PATH/mask.fits" -N 512 -l 0 -L 512

```

Cada etapa segue a mesma lógica de execução.

O espectro de potência para cada um dos quatorze mapas está expresso na Figura (10), que mostra os resultados com e sem shotnoise ( $C_\ell$  e  $S_\ell$  respectivamente, de acordo com a eq (62)).

O primeiro resultado notório é que os objetos em larga escala não são aleatoriamente distribuídos. É muito visível a interferência do shotnoise nos  $\ell$ 's maiores. Nos bins com menos objetos (últimos bins) a retirada do shotnoise tem efeitos mais drásticos, levando os valores a zero ou muito próximo disso.

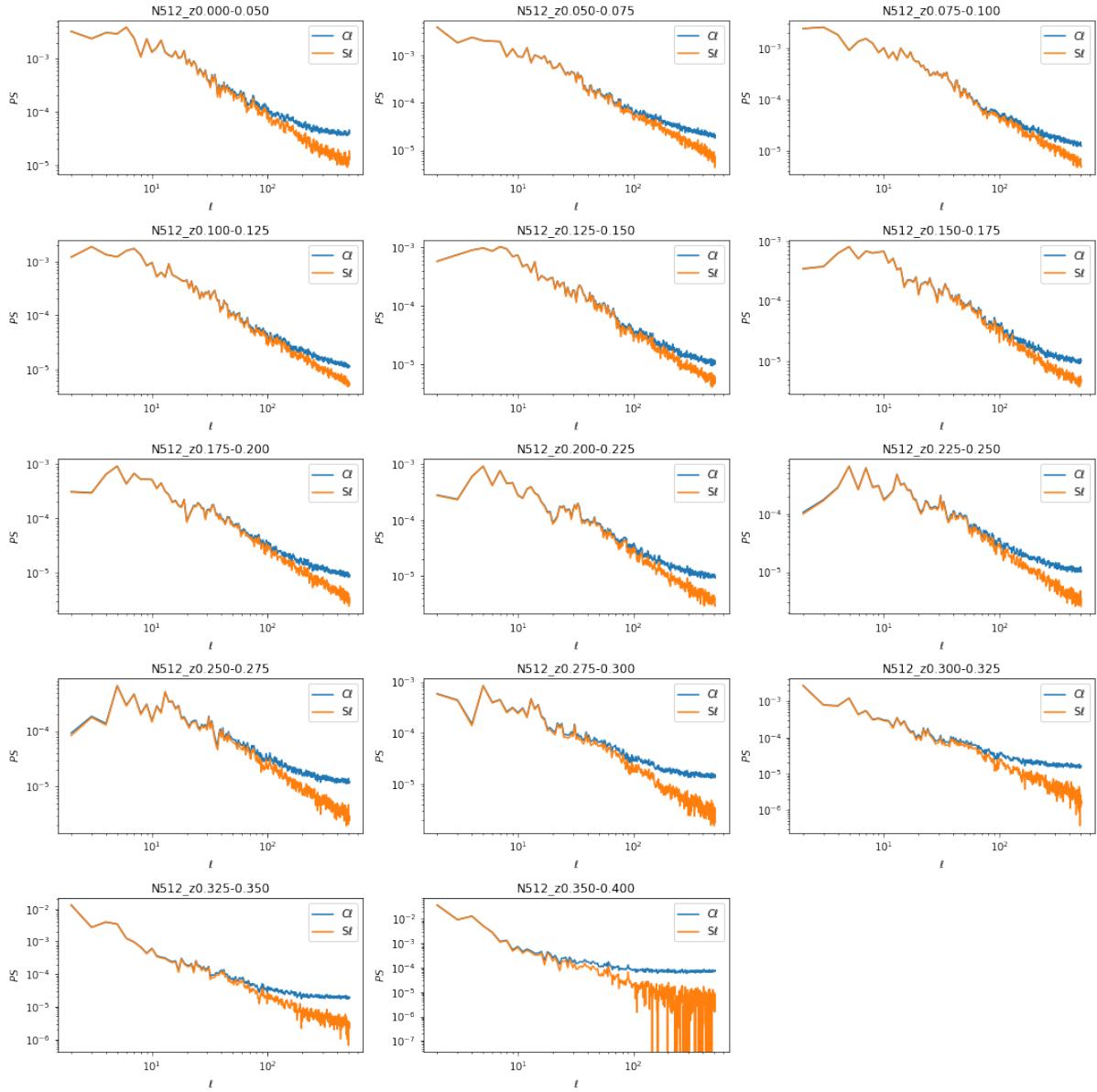


Figura 10: Espectro de potência  $PS$  em função do  $\ell$  calculado pelo Pseudo Power para cada intervalo de redshift na escala log x log. A linha azul é o  $C_\ell$ , o espectro de potência sem retirar o shotnoise e a linha laranja é  $S_\ell$  onde está sendo realizada esta correção.

### 3.5.2 Healpy

A função do Anafast do framework Healpy<sup>12</sup> aqui será usada para calcular o espectro de potência do mapa cosmológico afim de verificar a concordância entre os métodos e obter maior confiança na medida.

Para estimar o espectro de potência corretamente é necessário fazer algumas correções análogas às descritas na subseção 3.4. O espectro de potência será calculado como

$$S_\ell = \left( \frac{C_\ell}{f_{sky}} - Shotnoise \right) \frac{1}{w_\ell^2}, \quad (65)$$

onde  $f_{sky}$  é a parcela do céu que os dados abrangem. Neste caso, será a parcela do céu coberta pela máscara da Figura 100;  $w_\ell^2$  é o termo descrito na equação 59 e o *shotnoise* é o termo que aparece devido à discretização do sinal, descrito na equação 50, considerando a correção de céu parcial.

Os resultados que obtemos utilizando o Anafast foram extremamente satisfatórios. As medidas foram efetuadas com a correção do shotnoise. O espectro de potência  $S_\ell$ , mensurado para o mapa no intervalo de redshift  $0.000 < z \leq 0.050$  está disposto na figura 11. Os demais resultados podem ser visualizados na figura 12.

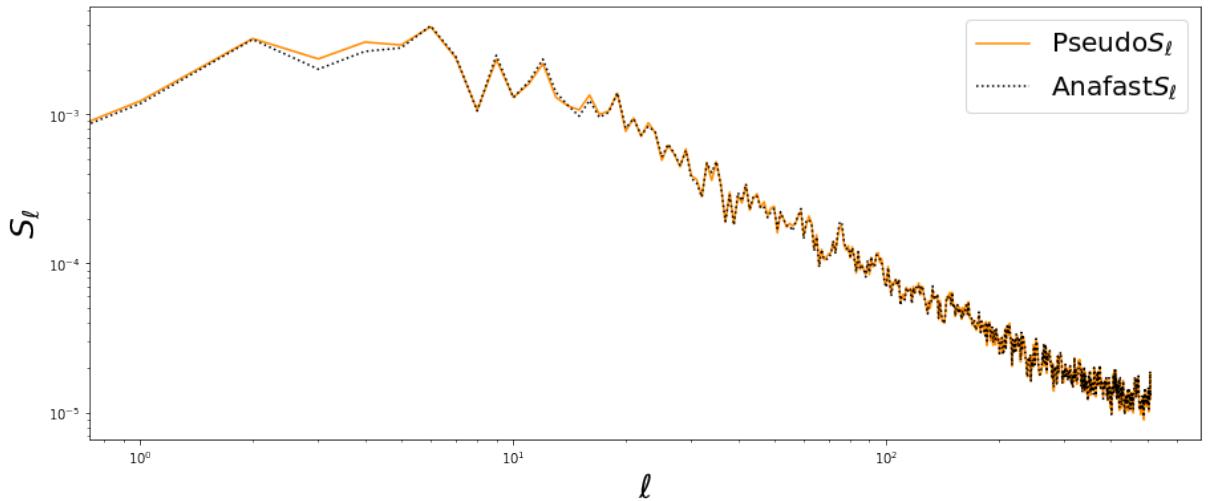


Figura 11: Espectros de potência para o intervalo  $0.000 < z \leq 0.050$ , calculados pelo PseudoPower e pela biblioteca Healpy, com as devidas correções.

---

<sup>12</sup>Healpy docs: <https://healpy.readthedocs.io/en/latest/>

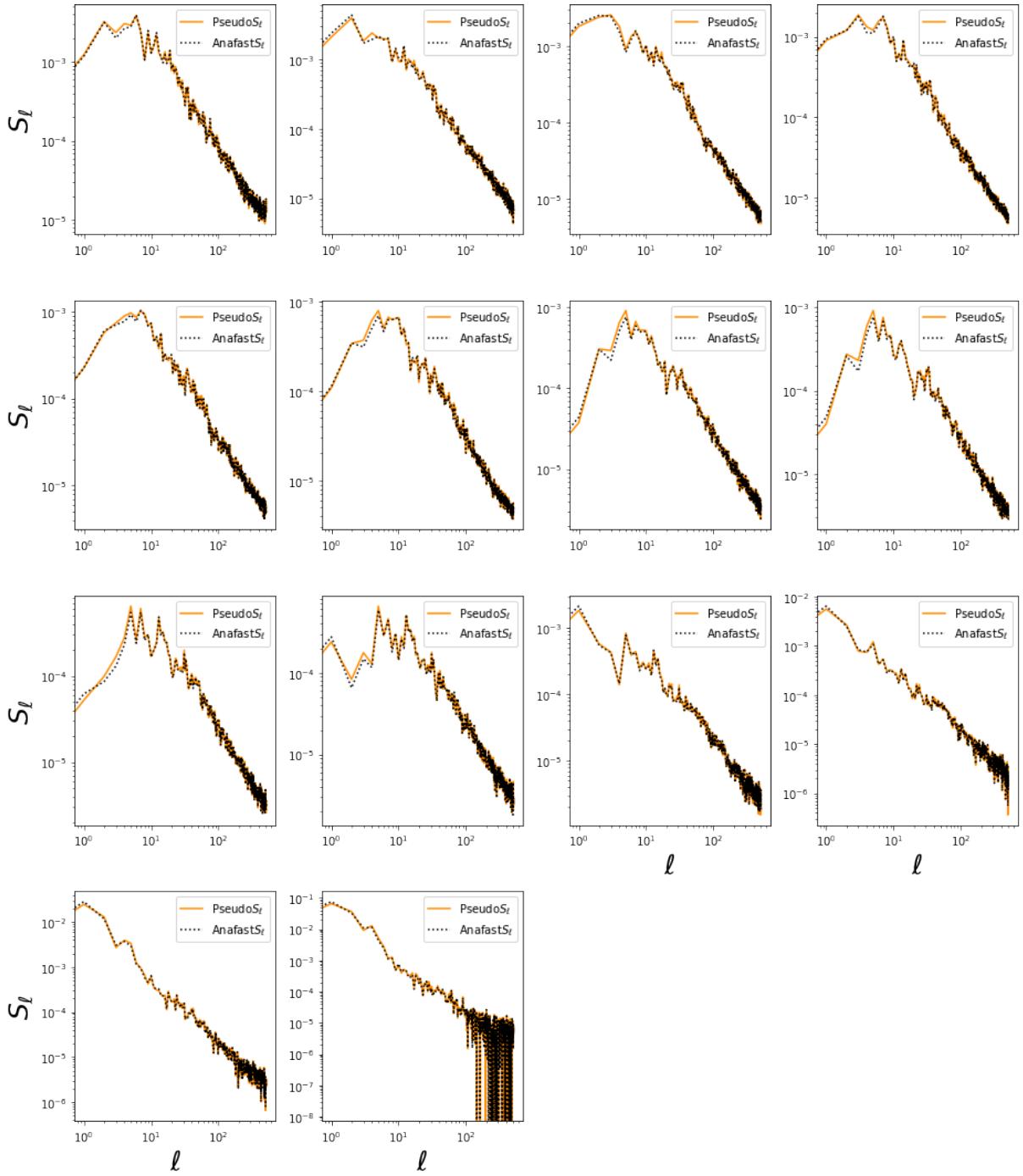


Figura 12: Espectros de potência para todos os intervalos de redshift, calculados pelo PseudoPower e pela biblioteca Healpy utilizando a equação 65. Não há divergência significativa entre os dois métodos de obtenção do espetro de potência.

### 3.5.3 UCLCL

Será utilizado o código UCLCL (*Unified Cosmological Library for  $C_\ell$ 's*, Loureiro-2019) para estimar o espectro de potência teórico dos dados utilizando a distribuição dos dados em relação ao redshift. O UCLCL é uma ferramenta idealizada especialmente para o cálculo da função de dois pontos relacionados a levantamentos cosmológicos de larga escala. Esta biblioteca, proposta por S.T. Balan, revisa o formalismo das correlações de potência angular e deriva equações analíticas exatas para o espectro de potência angular dos observáveis cosmológicos.

A proposta do UCLCL parte da generalização do cálculo do espectro de potência angular dos campos de densidade cosmológicos relacionados a estruturas de larga escala, que de forma simples pode ser expresso da seguinte forma:

$$C_\ell^{ij} = 4\pi \int \frac{dk}{k} W_\ell^i(k) W_\ell^j(k) \Delta^2(k), \quad (66)$$

onde  $W_\ell(k)$  é a função janela (window-function),  $k$  é o número de onda e os índices  $i$  e  $j$  denotam os diferentes intervalos de redshift estudados, representando na mesma equação as auto-correlações ( $i = j$ ) e as correlações cruzadas ( $i \neq j$ ). Neste termo são absorvidos todos os efeitos de evolução, convolução e projeção associados aos objetos de estudo. Incluindo efeitos associados à geometria, viés (*bias*)  $b(z)$  e à função de crescimento  $D(z)$ . A função  $\Delta^2(k)$  é o espectro de potência das flutuações da densidade de matéria expresso de forma adimensional, definido como variação da variância do campo por unidade de  $\ln(k)$ :

$$\Delta^2(k) \equiv \frac{d \langle \delta^2 \rangle}{d \ln(k)}. \quad (67)$$

Que no nosso caso é escrito como:

$$\Delta^2(k) = \frac{1}{(2\pi)^3} 4\pi k^3 P(k, 0). \quad (68)$$

A função janela depende fortemente das características experimentais dos objetos estudados, para o caso de um campo de galáxias, a função janela  $W_\ell^g(k)$  pode ser expresso por:

$$W_\ell^g(k) = \int dz f_g(z) j_\ell(kz). \quad (69)$$

Onde  $j_\ell(kz)$  são as funções esféricas de Bessel do primeiro tipo, de ordem  $\ell$ . A função de projeção  $f_g(z)$  por sua vez é dada por

$$f_g(z) = n(z) b^2(z) D(z) \frac{dz}{d\chi}, \quad (70)$$

onde  $b(z)$  é o viés,  $D(z)$  é a função de crescimento e  $\chi$  é a distância comóvel. A função  $n(z)$  introduzida aqui carrega a informação da probabilidade de encontrar as galáxias em um determinado intervalo de redshift, esta probabilidade é normalizada:

$$\int dz n(z) = 1. \quad (71)$$

Para calcular o  $n(z)$  neste estudo iremos considerar os dados de redshift espectroscópicos medidos e os dados de redshift fotométricos estimados por modelagem estatística. Em posse dos dados estimados, podemos realizar o caminho inverso, que consiste em selecionar os dados com a condição de que o redshift fotométrico esteja dentro de um determinado intervalo específico de redshift (que no nosso caso serão os intervalos dos nossos mapas cosmológicos especificados na tabela 1) e então observando seu respectivo redshift espectroscópico. Essa distribuição está expressa na figura 13. No UCLCL podemos informar tanto a distribuição  $n(z)$  quanto podemos ajustar a curva normal e informar apenas os parâmetros da distribuição. No primeiro caso o UCLCL realizará este procedimento de ajustar uma curva aos dados internamente. Como evidenciado na figura 13, os ajustes das curvas podem divergir, então se houver a necessidade de um controle maior neste ponto é indicado informar os parâmetros exatos. Neste estudo foi escolhido informar apenas a distribuição dos dados.

Todas as configurações relacionadas aos dados vão ser dispostas em um arquivo de configuração do UCLCL, como por exemplo o arquivo em anexo [UCLCL.ini](#). Para executar o UCLCL basta executar o seguinte comando no terminal,

```
1 $ PATH/uclcl/build/interface./uclcl /PATH/ini/UCLCL.ini
```

onde PATH são os caminhos da instalação do UCLCL e do arquivo de configuração, respectivamente.

Os resultados para os espectros de potência estimados pelo UCLCL para estes dados estão dispostos na figura 14, onde estão sendo comparados com os espectros mensurados pelo PseudoPower, para fins de validação. Note que para os espectros do UCLCL foram considerados dois casos, o primeiro considerando apenas termos lineares e o segundo considerando termos não lineares usando o modelo halofit (publicado de forma independente por [Peacock, J. A. and Smith, R. E.-2000](#) e [Seljak, U-2000](#)). Eventualmente nas etapas subsequentes pode ser de interesse utilizar os espectros onde os valores lineares e não lineares sejam coincidentes. Para alterar entre um e outro no UCLCL basta mudar o parâmetro nonlinear = none/halofit no arquivo de configurações [UCLCL.ini](#).

Os espectros comparados são da mesma ordem de grandeza, sendo este um resultado positivo e suficiente para o nosso propósito, porém divergem ligeiramente em forma.

Os  $\ell$ 's maiores sofrem efeitos não lineares que necessitam ser contabilizados no modelo teórico, então é natural que haja divergência nestes extremos. O ajuste do bias para cada mapa no modelo teórico do UCLCL é uma forma simples de ajustar os espectros que estão muito abaixo ou muito acima do Pseudo- $C\ell$ .

Para medir a qualidade do ajuste é utilizada uma combinação entre a comparação com os espectros estimados pelo PseudoPower e um gráfico de erro, definido neste contexto como a subtração entre os dois espectros. É esperado que o erro tenha distribuição constante e esteja no entorno de zero. Na figura 15 está plotado o erro para o caso linear, onde as distribuições não são constantes, o que denota que existem componentes que não estão sendo modeladas (sendo mais intenso nos primeiros mapas) e não estão centradas em zero. Na figura 16 estão plotados os erros para o caso não linear, onde a distribuição tem uma tendência de constante a partir de determinados  $\ell$ 's e estão em torno de zero majoritariamente.

Desvio padrão	Média
0.019	0.0337
0.025	0.0618
0.031	0.0863
0.030	0.112
0.025	0.136
0.031	0.161
0.031	0.186
0.034	0.207
0.042	0.237
0.034	0.262
0.027	0.282
0.035	0.309
0.044	0.334
0.045	0.361

Tabela 3: Valores descritivos das curvas normais associadas a seleção do redshift espetroscópico a partir dos intervalos redshift fotométrico. Estes são os dados das curvas ”Normal 2” da figura 13, que melhor se ajustou aos dados.

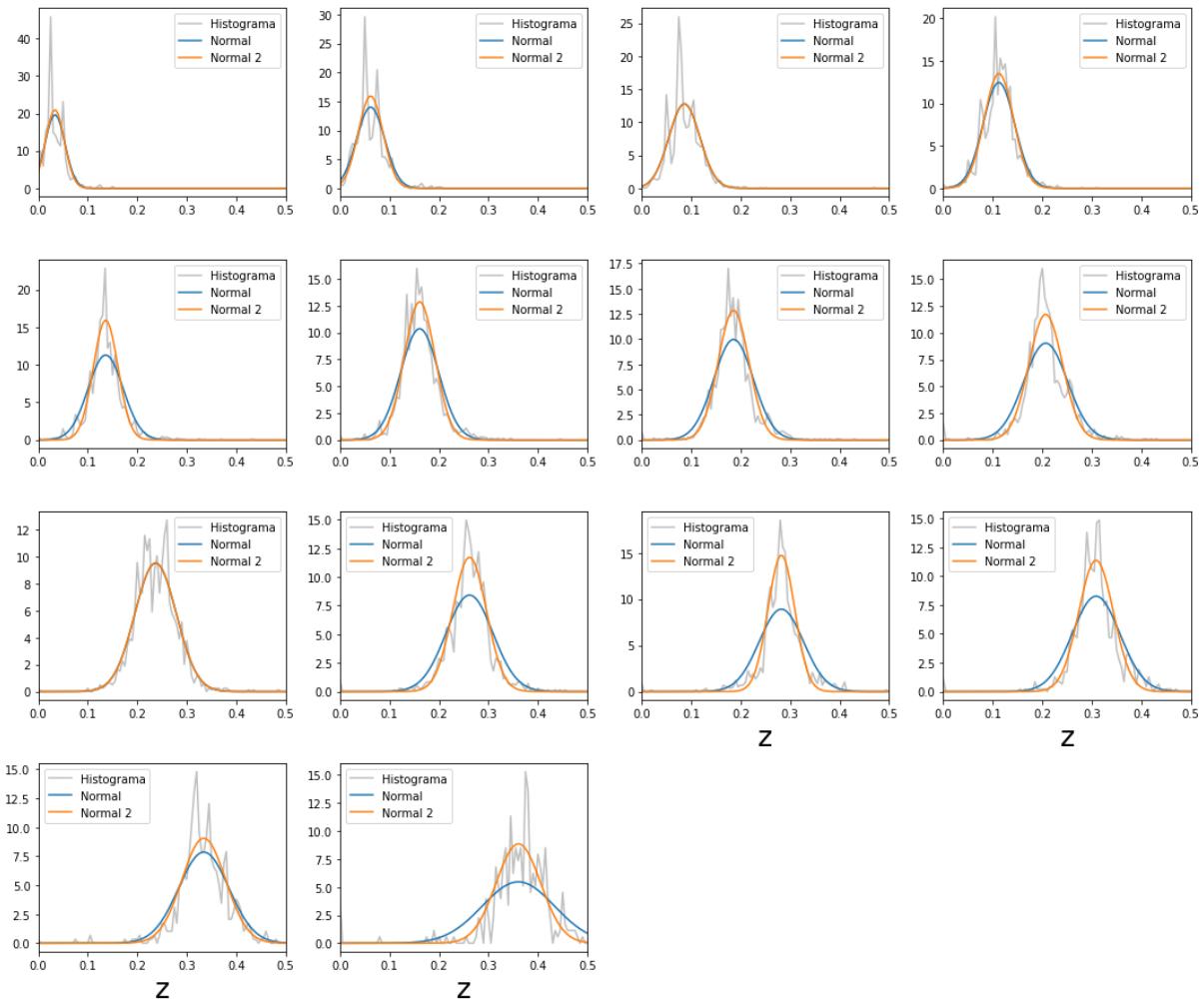


Figura 13: Histograma dos redshifts espectroscópicos para cada intervalo de redshift fotométrico selecionados conforme a Tabela (1).

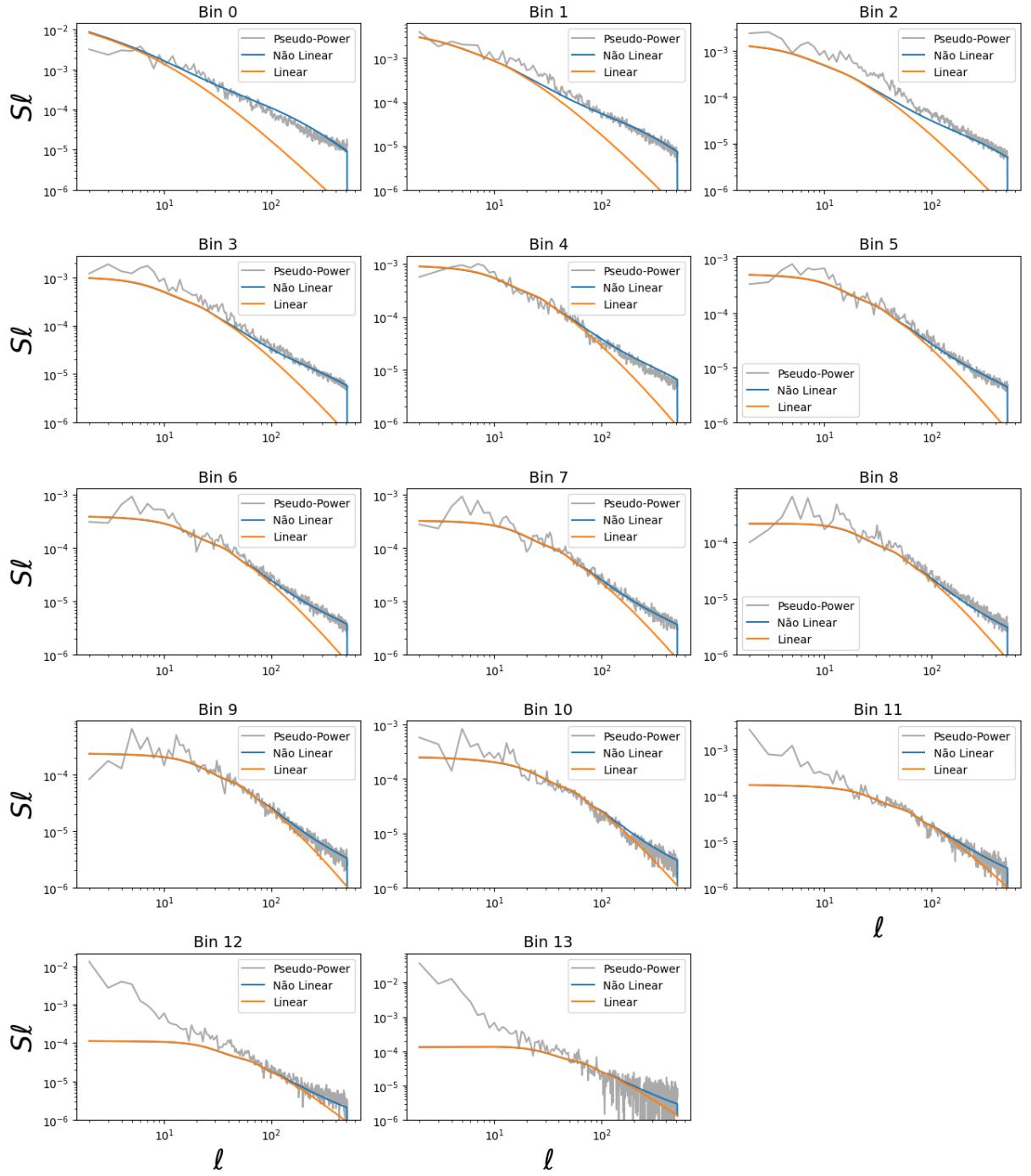


Figura 14: Comparação entre os espectros de potência  $S_\ell$  (sem shotnoise) calculados pelo PseudoPower em cinza e os espectros calculados pelo UCLCL. Na cor azul temos o cálculo considerando efeitos não lineares e na cor laranja apenas os efeitos lineares.

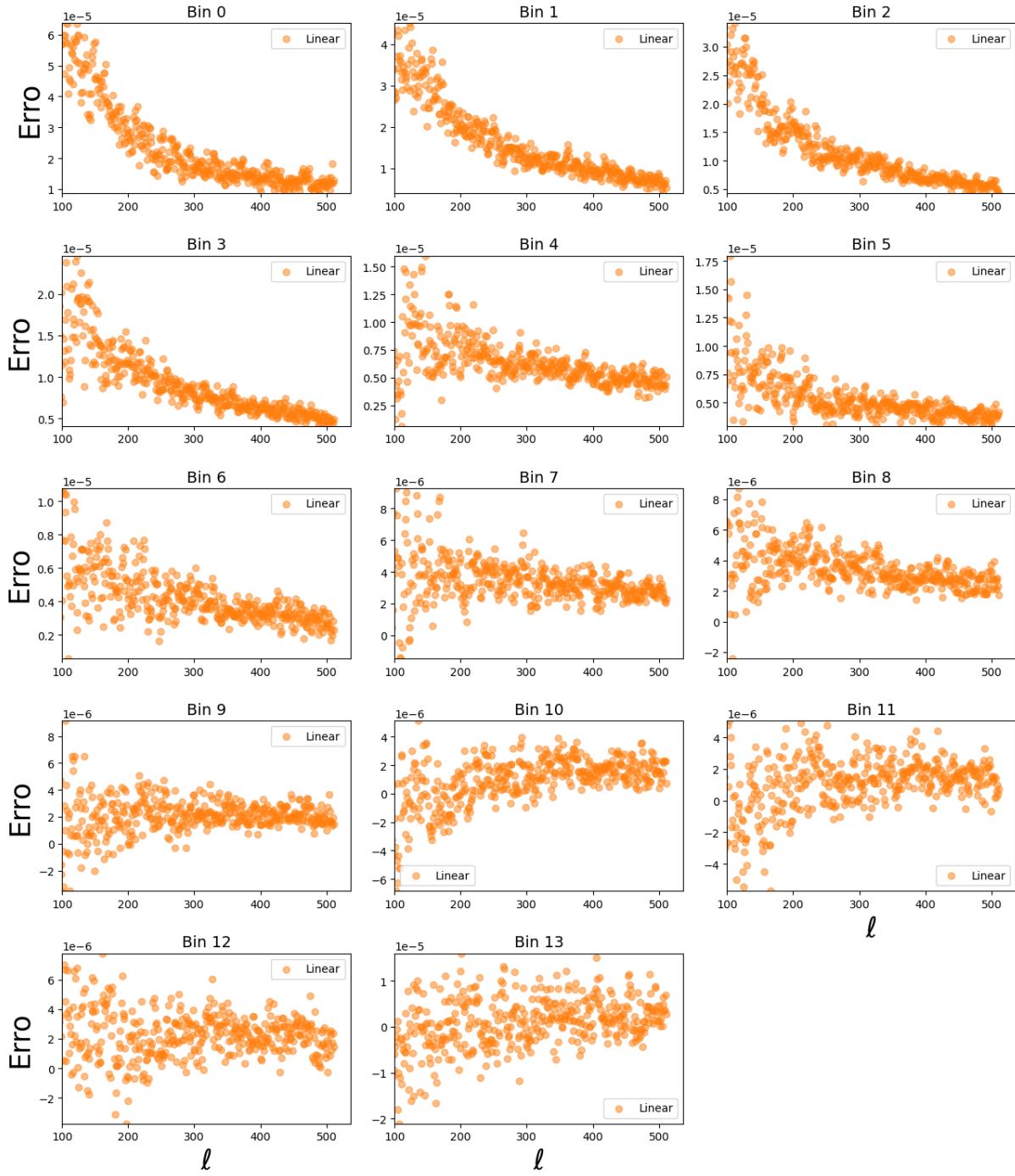


Figura 15: Erro entre o espectro de potência inferido pelo PseudoPower e o espectro de potência inferido pelo UCLCL considerando apenas interações lineares.

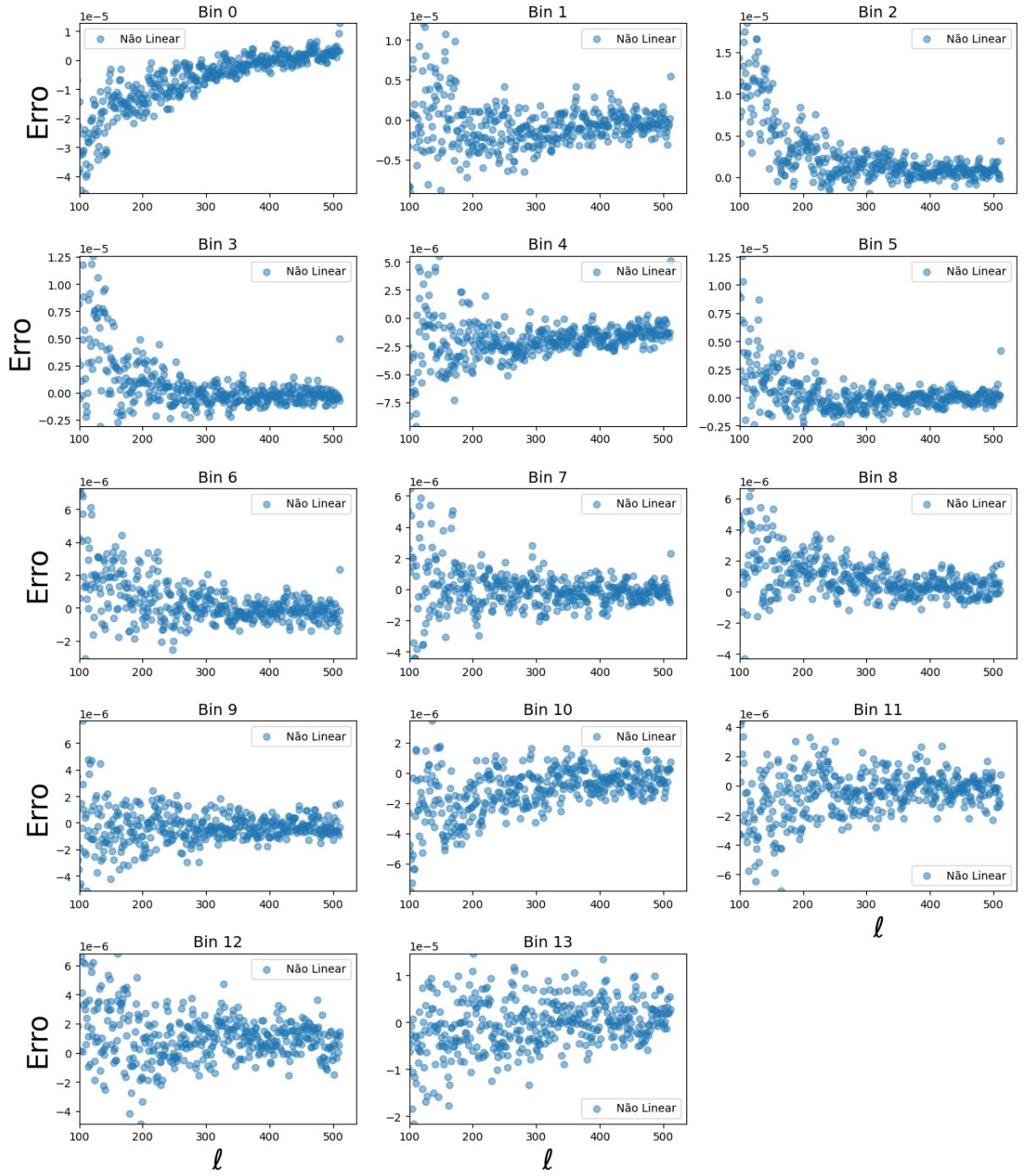


Figura 16: Erro entre o espectro de potência inferido pelo PseudoPower e o espectro de potência inferido pelo UCLCL considerando apenas interações não lineares.

## 4 Métodos de reamostragem

Os métodos de reamostragem, no contexto deste trabalho, envolvem a geração de múltiplas simulações de amostras a partir de um fenômeno específico, com o objetivo de estimar medidas de interesse. As amostras são geradas de acordo com um modelo predeterminado, que incorpora elementos de aleatoriedade. A utilização de números aleatórios é essencial nesse processo de simulação para introduzir a variabilidade inerente ao fenômeno analisado.

Como ilustração da ideia da técnica utilizada, suponha que nosso sinal fundamental seja da forma cossenoide, como ilustrado no primeiro gráfico da figura 17. Em problemas em que as medidas possuem muito ruído e podem ser realizadas apenas em um número pequeno de vezes, podem-se gerar N-simulações do fenômeno supondo uma determinada lei de formação afim de se separar o ruído do sinal. Após poucas interações, o que temos como resultado é muito próximo da nossa lei de formação, conforme ilustra o gráfico da figura 18.

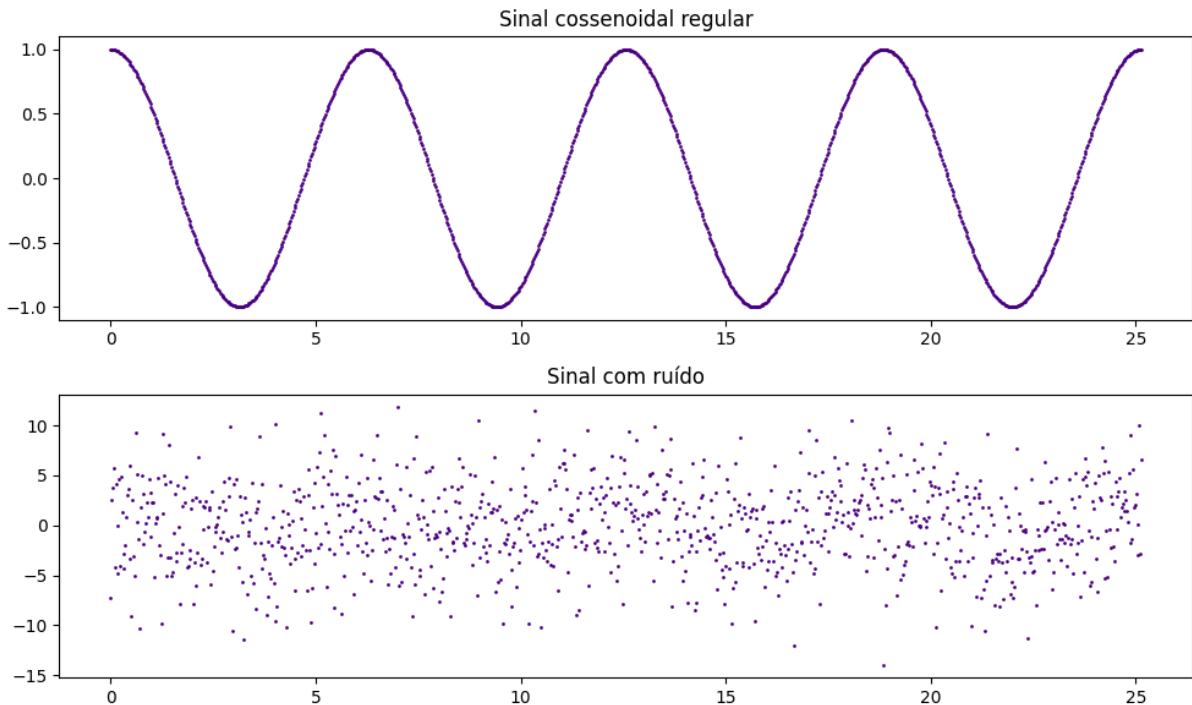


Figura 17: O primeiro gráfico é uma função cosseno que será base para o sinal. O segundo gráfico é a função cosseno mais um ruído branco.

Nesta sessão será apresentada a primeira etapa deste processo, que consiste em gerar n-mapas cosmológicos usando o [FLASK](#) e em extrair as estimativas para os respectivos espectros de potência. Para mais detalhes sobre como o FLASK funciona, pode-se consultar [Xavier-2016](#). Uma vez realizado esse procedimento pode-se inferir um intervalo de confiança para os dados e posteriormente gerar sua respectiva matriz de covariância.

Um segundo estudo em paralelo consiste em medir os Pseudo- $C_\ell$ 's cruzados entre os mapas cosmológicos simulados e os mapas de efeitos sistemáticos, com o objetivo de investigar a existência de correlação entre ambos, sendo o método de reamostragem significativo para estimar um intervalo de confiança para os resultados.

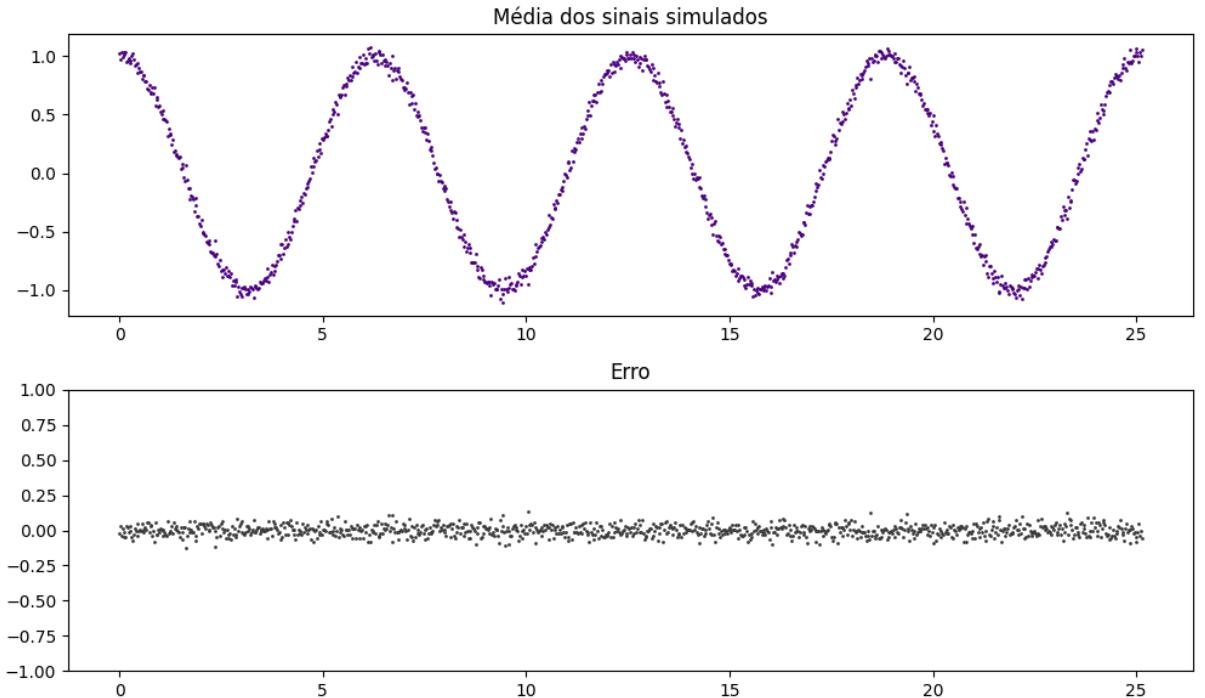


Figura 18: O primeiro gráfico é a média das  $N$ -simulações realizadas e o segundo gráfico é o erro tomando como parâmetro a função cosseno.

#### 4.1 FLASK: Simulações de mapas cosmológicos

*Full-sky Lognormal Astro-fields Simulation Kit* ([FLASK](#)) é a ferramenta que irá gerar as simulações dos mapas cosmológicos, ela foi projetada para modelar campos astrofísicos em larga escala. O FLASK possibilita a geração de simulações de um número arbitrário de campos aleatórios log-normais ou gaussianos correlacionados, permitindo simulações conjuntas precisas de todo o céu em geometria esférica, podendo ser de aglomerados ou lentes fracas. A distribuição dos objetos no mapa é definida *a priori*, a partir de espectros de potência fornecidos como parâmetros de entrada. Nesta sessão será feita uma introdução de como o FLASK gera os mapas cosmológicos e dos problemas envolvidos nesse processo, e como são as estratégias de resolução adotadas.

Na cosmologia, os campos log-normais são comumente usados para modelar observáveis da estrutura em larga escala devido à sua capacidade de representar com maior precisão a densidade de matéria, excluindo a necessidade de se tratarem valores não físicos, como densidades negativas.

O FLASK opera primeiro gerando campos gaussianos com base nos espectros de potência dos mapas cosmológicos fornecidos e depois os transformando em campos log-normais que serão base para a criação de mapas do céu inteiro.

Dado um conjunto de variáveis aleatórias  $Z_i$  que seguem uma distribuição gaussiana multivariada com o vetor associado às médias  $\mu_i$  e uma matriz de covariância  $\xi_g^{ij}$ , podemos estabelecer a transformação

$$X_i = e^{Z_i} - \lambda_i , \quad (72)$$

onde  $X_i$  são as variáveis aleatórias associadas com distribuição lognormal multivariada e  $\lambda_i$  são os parâmetros de corte, que estabelecem os valores mínimos na distribuição. A partir

disso, pode-se estabelecer uma relação entre a matriz de covariância de uma distribuição gaussiana  $\xi_g^{ij}$  e uma matriz de covariância lognormal  $\xi_{ln}^{ij}$  associada,

$$\xi_{ln}^{ij} = \langle X_i X_j \rangle - \langle X_i \rangle \langle X_j \rangle , \quad (73)$$

$$\xi_{ln}^{ij} = \alpha_i \alpha_j (e^{\xi_g^{ij}} - 1) . \quad (74)$$

$$\xi_g^{ij} = \ln \left( \frac{\xi_{ln}^{ij}}{\alpha_i \alpha_j} + 1 \right) , \quad (75)$$

onde  $\alpha_i \equiv \langle X_i \rangle + \lambda_i$ . Podemos estabelecer os coeficientes de correlação de Pearson das variáveis log-normais  $\rho_{ln}$  e seus pares gaussianos associados  $\rho_g$ ,

$$\rho_{ln} = \frac{e^{\rho_g \sigma_1 \sigma_2} - 1}{\sqrt{(e^{\sigma_1^2} - 1)(e^{\sigma_2^2} - 1)}} . \quad (76)$$

Acima,  $\sigma_i^2$  são as variâncias associadas às variáveis gaussianas. Neste momento, surge o primeiro problema a ser resolvido. Mesmo que haja correlação perfeita entre as variáveis gaussianas associadas, isso pode não resultar em uma correlação perfeita entre as variáveis lognormais associadas. Em outras palavras, se  $\rho_g = 1$  podemos obter  $\rho_{ln} \neq 1$ . Isso ocorre porque não é possível estabelecer uma relação linear entre duas variáveis se suas distribuições são diferentes.

Eventualmente, a transformação da matriz de covariância lognormal para gaussiana pode resultar em uma matriz não positivo definida. Para lidar com essa situação é possível recorrer a duas estratégias: a primeira consiste em computar as mudanças mínimas necessárias para que a matriz se torne positivo definida. Para isso é utilizado um gradiente multidimensional para minimizar a soma dos valores absolutos dos autovalores negativos, o que permite encontrar uma direção preferencial para distorcer a matriz. Aplicamos uma pequena mudança nesta direção. Repetimos esse processo até que todos os autovalores se tornem positivos. Outro método para regularizar uma matriz de covariância é realizar uma decomposição em autovalores da matriz de  $\xi_g = Q \Lambda Q^{-1}$ , onde  $Q$  é uma matriz formada pelos autovetores de  $\xi_g$  e  $\Lambda$  é uma matriz diagonal formada pelos autovalores de  $\xi_g$ . Em seguida, quaisquer autovalores negativos são definidos como zero.

São fornecidos para o FLASK os espectros de potência angular mensurados a partir das observações cosmológicas. A partir disso, calcula-se a matriz de covariância, que é regularizada para que esteja de acordo com a distribuição normal. A matriz de covariância pode ser expressa em termos do espectro de potência angular e vice-versa,

$$C^{ij}(\ell) = 2\pi \int_0^\pi \xi^{ij}(\theta) P_\ell(\cos(\theta)) \sin(\theta) d\theta, \quad (77)$$

$$\xi^{ij}(\theta) = \frac{1}{4\pi} \sum_{\ell=0}^{\infty} (2\ell+1) P_\ell(\cos(\theta)) C^{ij}(\ell) , \quad (78)$$

onde  $P_\ell$  são os polinômios de Legendre.

A matriz de covariância é base para criação de mapas, mas ainda é necessário adicionar uma componente de aleatoriedade para que os novos mapas sejam diferentes dos originais. Para isso, realiza-se uma decomposição de Cholesky, que é uma técnica de fatoração de uma matriz simétrica e definida positiva em álgebra linear. Dada uma matriz  $M$  simétrica

e definida positiva, a decomposição de Cholesky permite escrevê-la na forma

$$M = T T^T, \quad (79)$$

onde  $T$  é uma matriz triangular inferior e  $T^T$  é sua transposta e é nomeada matriz de Cholesky. O FLASK utiliza a biblioteca GSL - GNU Scientific Library\* para efetuar esse processo sobre a matriz de covariância,

$$C_g^{ij}(\ell) = \sum_k T_{ik}(\ell) T_{jk}(\ell) . \quad (80)$$

As simulações devem obedecer o *a priori* determinado pelos espectros de potência na sua lei de formação. Ao mesmo tempo, é necessário haver uma componente de aleatoriedade. Para esta finalidade multiplicamos uma das matrizes triangulares  $T(\ell)$  da decomposição de Cholesky por um conjunto de variáveis aleatórias Gaussianas de média zero e variância unitária,

$$Z_{i,\ell m} = \sum_k T_{ik}(\ell) Z_{k,\ell m}^0(\ell). \quad (81)$$

A partir dos  $Z_{i,\ell m}$  podem-se produzir os mapas  $Z_i(\hat{\theta})$  utilizando a função ALM2MAP<sup>13</sup> da biblioteca HEALPIX, isso produzirá os mapas com distribuição gaussiana. Para obter os mapas lognormais, exponencia-se o campo gaussiano, conforme a transformação de variáveis previamente estabelecida,

$$X(\hat{\theta}) = e^{\mu_i} e^{Z(\hat{\theta})} - \lambda_i , \quad (82)$$

onde  $X(\hat{\theta})$  representa o campo lognormal associado. O fator exponencial que multiplica é descrito como

$$e^{\mu_i} = (\langle X_i \rangle + \lambda_i) e^{-\sigma_i^2/2} . \quad (83)$$

O fator  $\sigma_i^2$  é a variância associada ao campo Gaussiano  $Z_i(\hat{\theta})$ , expressa por

$$\sigma_i^2 = \sum_{\ell=\ell_{min}}^{\ell_{max}} \frac{2\ell+1}{4\pi} C_g^{ii}(\ell) . \quad (84)$$

Ao final do processo, se utilizamos a mesma máscara e medirmos os respectivos espectros de potência dos mapas simulados, devemos recuperar os espectros de potência originais com certa variação devido à aleatoriedade proposta que está intrínseca no processo de geração de mapas simulados.

A execução do FLASK pode ser feita manualmente através do terminal, como exemplo, basta navegar pelo terminal até onde o projeto foi instalado (supondo aqui PATH/FLASK) e executar o comando a seguir:

```
1 $ cd PATH/FLASK
2 $ ./bin/flask example.config
```

Listing 3: Executando o exemplo padrão do FLASK para geração de mapas simulados.

---

<sup>13</sup>HEALPIX - ALM2MAP: [https://healpix.sourceforge.io/html/Healpix\\_cxx/group\\_\\_alm\\_healpix\\_\\_group.html](https://healpix.sourceforge.io/html/Healpix_cxx/group__alm_healpix__group.html)

Desta forma, serão geradas simulações de mapas cosmológicos. Para cada projeto, o arquivo de configuração deve ser personalizado conforme a necessidade. Para este projeto, o arquivo de configuração está disposto tanto no GIT do projeto<sup>14</sup> quanto no anexo `flask.config`. O arquivo de configuração está vinculado a mais dois arquivos necessários, um deles para definir os tipos de campos que serão gerados, juntamente com os intervalos de redshift `SDSS_14Bins.dat` e o outro é referente à função radial de seleção (descrita a seguir) `SDSS_14Bins_radial-selection-function-f1.dat`; ambos os arquivos devem ser referenciados no arquivo de configuração do FLASK. Todo esse processo é abstraído no aplicativo final deste projeto e o uso se torna facilitado, bastando o usuário instalar todas as dependências do `CovAnalysis` e selecionar o número de mapas que deseja gerar no processo dentro do arquivo `esy.ini` (arquivo com as configurações do programa), na variável `NFLASKMAPS`.

Neste estudo, a técnica que iremos empregar consiste em gerar n-mapas cosmológicos utilizando o FLASK e destas simulações extraír os respectivos espectros de potência. Serão mensurados tanto as autocorrelações quanto as correlações cruzadas dos mapas simulados. A partir das n-simulações poderemos gerar um intervalo de confiança para os espectros de potência. A finalidade principal deste estudo é estimar uma matriz de covariância, que será gerada a partir dos espectros de potência dos mapas cosmológicos simulados.

Para que o FLASK realize as simulações corretas dos mapas cosmológicos é necessário informar os  $S_\ell$ 's estimados entre todos os mapas, ou seja, tanto os auto  $S_\ell$ 's quantos os  $S_\ell$ 's cruzados, estes dados servirão aqui como "lei de formação" para os mapas simulados. O FLASK pode gerar tanto os mapas de sobredensidade quantos os mapas de densidade e com distribuição normal ou lognormal, conforme a necessidade. Neste caso, o FLASK irá simular mapas considerando uma distribuição do tipo lognormal que esteja de acordo com os espectros de potência sem shotnoise fornecidos como input. Nesta configuração, as simulações ainda contêm um número de objetos incorreto se comparado com o que é esperado para cada intervalo. Para que nossa simulação seja fidedigna, é necessário corrigir a densidade de objetos simulados para cada um dos intervalos de redshift. Isso é possível ajustando-se a função radial de seleção fornecida para o FLASK,

$$W(i, j) = A(j) \int_{z_{min}}^{z_{max}} dz f(z) , \quad (85)$$

para cada pixel  $j$  pertencente ao intervalo de redshift  $i$ , onde  $A(j)$  é uma função arbitrária e  $f(z)$  é a função radial de seleção. O número de objetos esperados  $\bar{n}(i, j)$  será proporcional à função

$$\bar{n}(i, j) \propto W(i, j). \quad (86)$$

Uma descrição detalhada deste processo está descrita na documentação do FLASK, na seção 4.1.7 do manual<sup>15</sup>.

---

<sup>14</sup>[https://gitlab.com/FSCPhysis/Covariance\\_Analysis](https://gitlab.com/FSCPhysis/Covariance_Analysis)

<sup>15</sup>Manual do FLASK: [https://github.com/ucl-cosmoparticles/flask/blob/master/doc/flask\\_manual.pdf](https://github.com/ucl-cosmoparticles/flask/blob/master/doc/flask_manual.pdf)

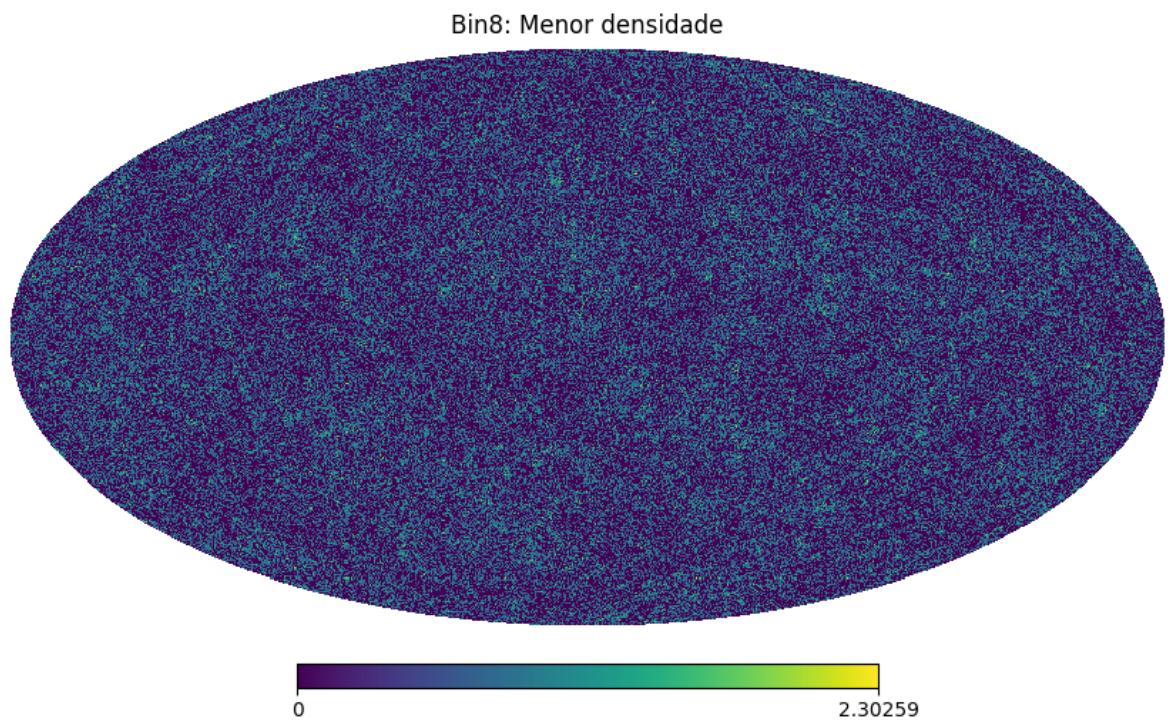


Figura 19: Mapa de densidade gerado pelo FLASK em escala log com um valor igual a 0.495 para a função radial de seleção, gerando menor densidade de objetos. Objetos totais: 2.159.531

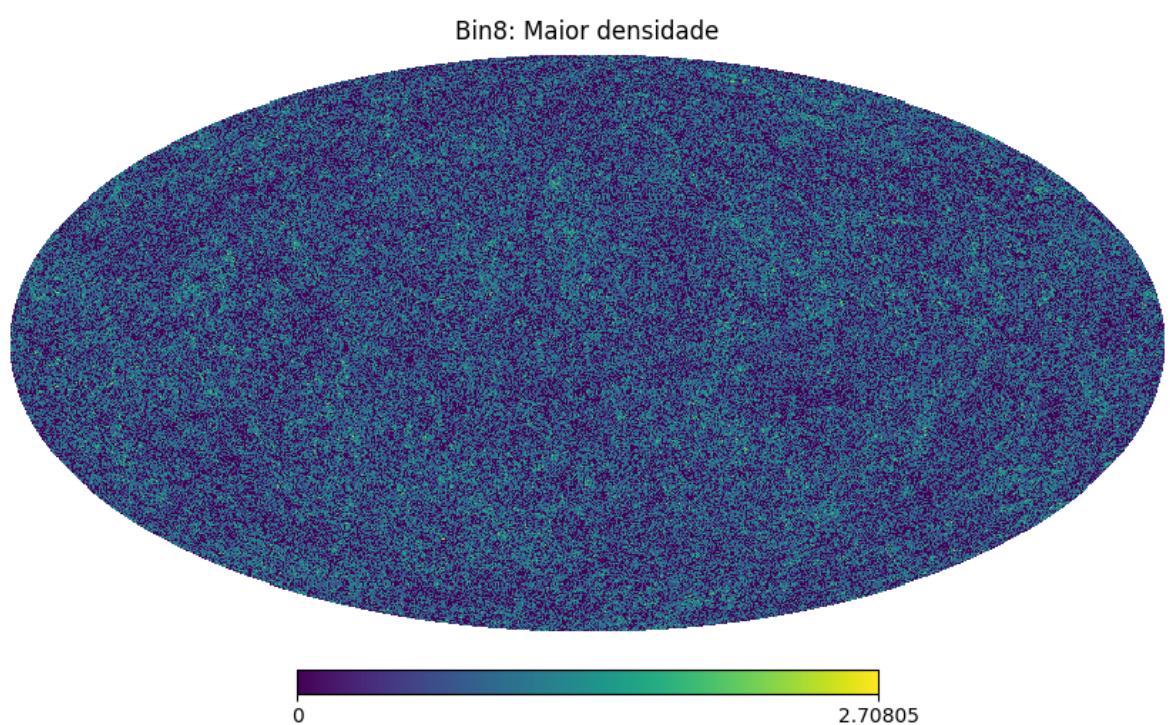


Figura 20: Mapa de densidade gerado pelo FLASK em escala log com um valor igual a 1.00 para a função radial de seleção, gerando maior densidade de objetos. Objetos totais: 3.712.252

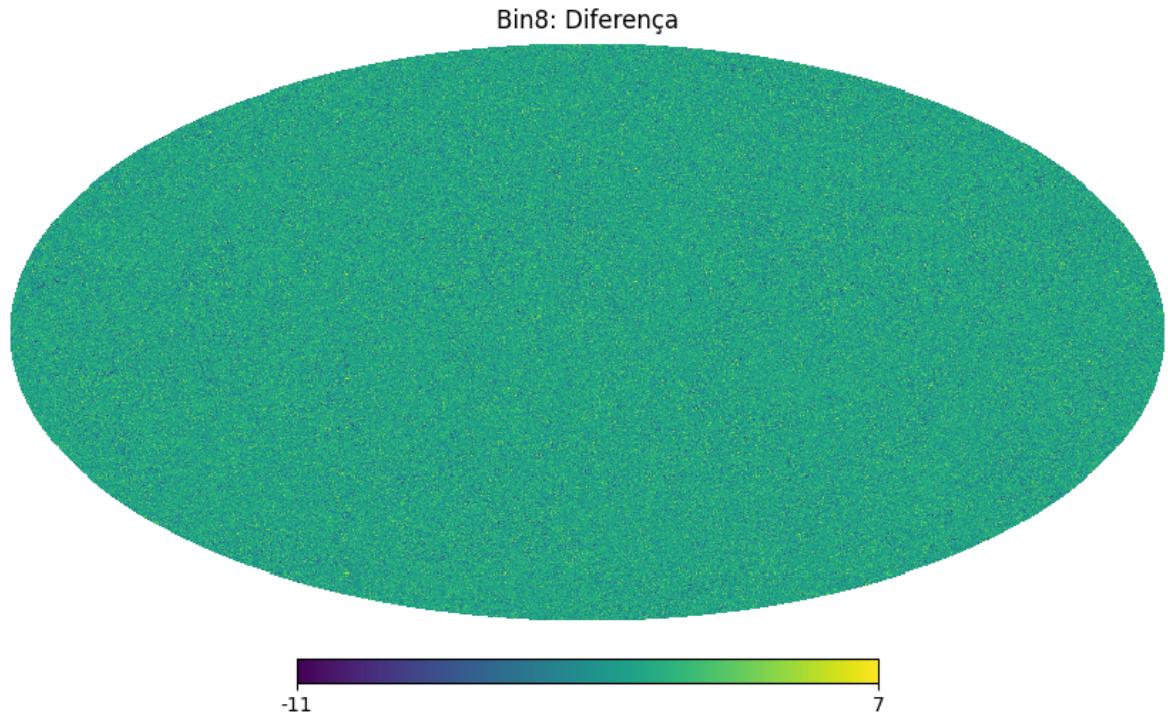


Figura 21: Diferença entre o mapa simulado pelo FLASK com densidade maior e o mapa com densidade menor.

## 4.2 Automatização do processo

Um dos principais produtos deste trabalho é a automatização dos processos através do código [CovAnalysis](#), escrito em Python. O código possui um arquivo de configuração no formato de texto junto com o seu executável na raiz do projeto onde se colocam todas as informações necessárias e personalizações. São necessários tanto o PseudoPower quanto o FLASK para o correto funcionamento, além das dependências que constam no repositório do projeto. Abaixo estão abstraídas as etapas do processo bem como a estratégia de paralelização do código.

Na base do processo está o diagrama da figura 22, onde são calculados os espectros de potências dos mapas de sobredensidade. Nesta etapa, a máscara da figura 100 é usada para calcular a medida  $IlmJlm$ , referente à equação 54, além de ser aplicada ao mapa em questão. A partir do mapa de sobredensidade é calculada a quantidade  $Alm$ , referente à equação 58. A partir destes dois elementos é estimado o espectro de potência e retirado o shotnoise, conforme a equação 64.

A máscara binária (onde cada pixel é zero ou um) é aplicada ao mapa de densidade, onde cada pixel é o numero de objetos computados na equivalente região espacial, para só então o mapa ser convertido para um mapa de sobredensidade através da equação 57 . O motivo é que no mapa de densidade, cada pixel está no intervalo  $[0, \mathbb{R}]$ , portanto a região vazia (zero) ao ser multiplicada pela região vazia da máscara (zero) permanece inalterada. Por outro lado, no mapa de sobredensidade, cada pixel está no intervalo  $[-1, \mathbb{R}]$ , onde a região vazia agora é a unidade negativa, que ao ser multiplicada por zero se altera. Em suma, o valor zero possui significado diferente em cada um dos mapas, para o mapa de

densidade, o zero é ausência de objetos, para o mapa de sobredensidade, o zero significa que a região possui uma densidade de objetos equivalente a densidade média de objetos do mapa.

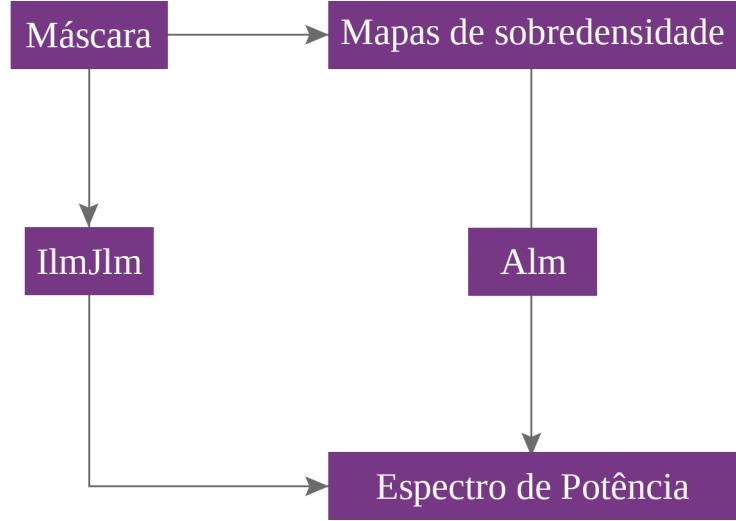


Figura 22: Fluxograma de como o espectro de potência é calculado através do código PseudoPower. O IlmJlm é calculado apenas uma vez para cada máscara e junto com o Alm produz o espectro de potência referente.

O método de reamostragem em particular consistem em realizar N-simulações dos mapas cosmológicos através do FLASK. Essas simulações podem ser geradas como mapas de sobredensidade diretamente ou como mapas de densidade. Na versão do FLASK usada, caso seja desejado um ajuste na densidade de objetos o código resultará em um mapa de densidade, que deve ser devidamente mascarado e então convertido para um mapa de sobredensidade, onde terá seu espectro de potência estimado de acordo com o fluxo da figura 22. O ajuste de densidade será feito através da função radial de seleção, que deve ser fornecida ao FLASK. Esta função dependerá da quantidade de objetos desejados para cada mapa.

O processo deve ser repetido N vezes para cada mapa analisado. A partir das N-simulações realizadas será gerada a matriz de covariância dos dados, que será detalhada na próxima sessão.

Os cálculos envolvidos neste processo são muito custosos, então para que haja maior velocidade no processamento dos dados foi realizada a paralelização do código. A paralelização refere-se à técnica de dividir uma tarefa ou problema computacional em partes menores, que podem ser executadas simultaneamente por diferentes processadores, núcleos de CPU ou até mesmo em diferentes máquinas. O objetivo é acelerar a execução do código, aproveitando a capacidade de processamento de múltiplos recursos de forma simultânea.

A estratégia adotada foi processar cada rodada da simulação e obtenção dos respectivos espectros de potência por núcleo, conforme o diagrama expresso na figura 23. O código reserva os núcleos de processamento pré determinados nas configurações do código (*pool de processamento*) e inicia a distribuição das tarefas de modo assíncrono, ou seja, quando um núcleo termina o tratamento de uma simulação, a próxima tarefa da fila lhe é atribuída, conforme o diagrama da figura 24, independentemente dos outros núcleos ainda estarem realizando suas atividades pré-determinadas.

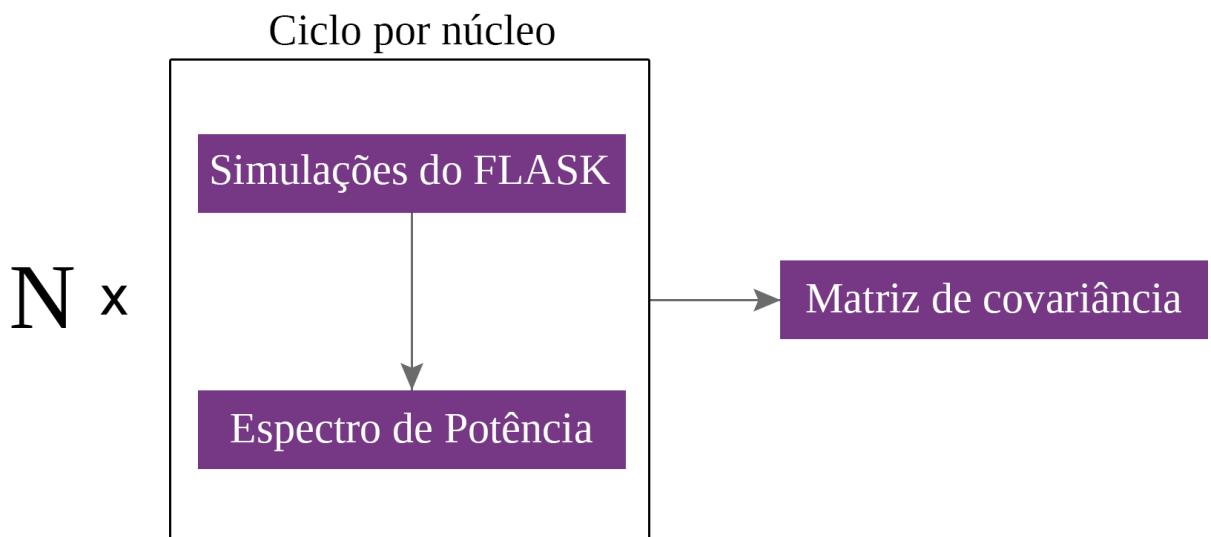


Figura 23: Estratégia de processamento em paralelo adotado na automatização do processo.  $N$  é o número de simulações desejadas. Cada simulação vai ser processada em um núcleo de processamento em específico. Esses resultados por fim irão gerar a matriz de covariância.

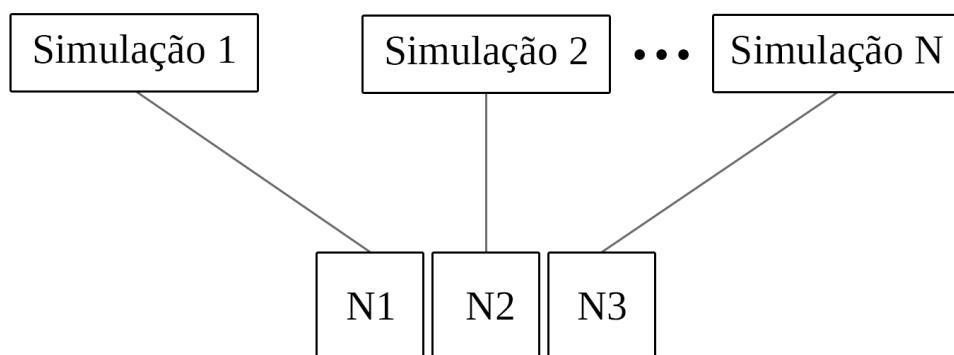


Figura 24: A fila de processos a serem realizados vão sendo alocados em um núcleo de processamento disponível de maneira assíncrona.

## 4.3 Resultados: Espectro de potência das simulações

### 4.3.1 $C_\ell$

A medida do Pseudo- $C_\ell$  representa um formalismo estatístico fundamental na análise de dados observacionais em cosmologia, permitindo a estimativa do espectro de potência dos mapas de sobredensidade. No contexto deste estudo, utilizamos o Pseudo- $C_\ell$  para calcular os espectros de potência dos mapas gerados pelo FLASK. No gráfico da figura 25, os espectros de potência simulados são apresentados em tons de cinza, enquanto o espectro de potência obtido a partir do mapa cosmológico de sobredensidade é destacado em violeta. As barras verticais representam a variância entre as medidas de espectro de potência dos mapas simulados em cada ponto, refletindo a incerteza associada a essas estimativas. Conforme o número de harmônicos  $\ell$  aumenta, as escalas angulares consideradas no Pseudo- $C_\ell$  diminuem, resultando em uma redução nas estimativas do Pseudo- $C_\ell$ .

A técnica de reamostragem servirá para mitigar ruídos nos sinais simulados e melhorar a confiabilidade das nossas estimativas dos espectros de potência contribuindo para a robustez da análise. Realizamos para este propósito duas mil simulações.

As métricas de avaliação das simulações serão métricas normalizadas pelo espectro de potência medido nos mapas, estas métricas aqui serão nomeadas métricas relativas (erro relativo, variância relativa etc). Como as curvas são descendentes na forma como foi escolhido trabalhar, as métricas também diminuem de escala, por esta razão a normalização é tão útil, ela permite uma avaliação adequada das métricas em toda a curva, independente da escala.

O erro relativo é calculado como a diferença entre os valores de espectro de potência medidos nos mapas cosmológicos e os valores medidos nas simulações do FLASK, dividida pelo valor medido nos mapas. O erro e o erro relativo podem ser expressos pelas seguintes equações:

$$\text{Erro} = C_{\ell \text{mapas}} - C_{\ell \text{FLASK}}, \quad (87)$$

$$\text{Erro relativo} = \frac{C_{\ell \text{mapas}} - C_{\ell \text{FLASK}}}{C_{\ell \text{mapas}}}. \quad (88)$$

Essa métrica (erro relativo) fornece uma medida quantitativa da discrepância entre os resultados observados e os resultados simulados, permitindo uma avaliação da precisão e da qualidade das simulações realizadas pelo FLASK. A análise do erro relativo é fundamental para identificar possíveis desvios sistemáticos nas simulações por ter a característica de ajustar a escala. É desejável que se tenha uma distribuição uniforme.

Da mesma forma, pode-se escrever a variância relativa para as simulações do FLASK, para cada valor de  $\ell$ ,

$$\sigma_\ell^2 = \frac{\sum_s (C_\ell^s - \bar{C}_\ell)^2}{n-1}, \quad (89)$$

onde  $s$  no somatório refere-se à contagem da simulação em específico,  $C_\ell$  são os espectros de potência dos mapas simulados pelo FLASK,  $n$  é o número de simulações e  $\bar{C}_\ell$  é a média de todos os espectros de potência para cada valor de  $\ell$ , definido como

$$\bar{C}_\ell \equiv \frac{1}{n} \sum_s C_\ell^s. \quad (90)$$

A figura 25 representa o resultado que obtemos no processo, neste caso para o mapa no primeiro intervalo de redshift (bin 0). Neste gráfico estão presentes os valores estimados diretamente no mapa cosmológico (em violeta) e todas as simulações realizadas pelo FLASK (em cinza), bem como os valores da variância ponto a ponto dos mapas simulados. O gráfico da figura 26 apresenta o erro relativo entre os valores do espectro de potência  $C_\ell$  obtidos a partir dos mapas cosmológicos e os  $C_\ell$ 's das simulações realizadas pelo FLASK. Este é um gráfico de dispersão, então para cada  $\ell$  haverá um número  $n$  de pontos, um para cada simulação realizada.

Procedemos à análise da variância normalizada dos dados simulados, visando investigar o comportamento dessas simulações. A figura 27 ilustra graficamente este estudo, apresentando a variância normalizada dos dados em relação aos valores do espectro de potência medidos a partir dos mapas cosmológicos. Notavelmente, observa-se que a variância relativa dos dados é mais pronunciada para os valores menores de  $\ell$  e exibe uma queda assintótica à medida que  $\ell$  aumenta.

A distribuição das variâncias está representada na figura 28, onde se observa que a maioria dos valores é pequena, indicando uma consistência geral nas simulações. No entanto, há uma pequena quantidade de valores discrepantes, especialmente nos valores de  $\ell$  mais baixos (de acordo com a figura 27).

A figura 29 é o erro (simples, não normalizado) dos  $C_\ell$  dos mapas simulados em comparação a cada um dos mapas cosmológicos. Em todos os casos ele possui um comportamento assintótico a medida que  $\ell$  aumenta. Como já comentado, a escala do erro diminui para valores mais altos de  $\ell$ , impedindo uma avaliação consistente do comportamento das simulações.

O erro simples pode não ser suficiente para avaliar os resultados, para isso utiliza-se o erro relativo, que é o erro normalizado pelo  $C_\ell$  do mapa cosmológico. Essa medida trás maior consciência sobre as divergências em  $\ell$  maiores, onde o valor do erro pode ser pequeno, mas ainda sim muito significativo em relação à escala trabalhada. Analisando desta forma nota-se que apenas os valores iniciais divergem para o caso dos  $C_\ell$ 's, com o erro convergindo para uma faixa de oscilação mais bem comportada a partir de um determinado valor de  $\ell$ .

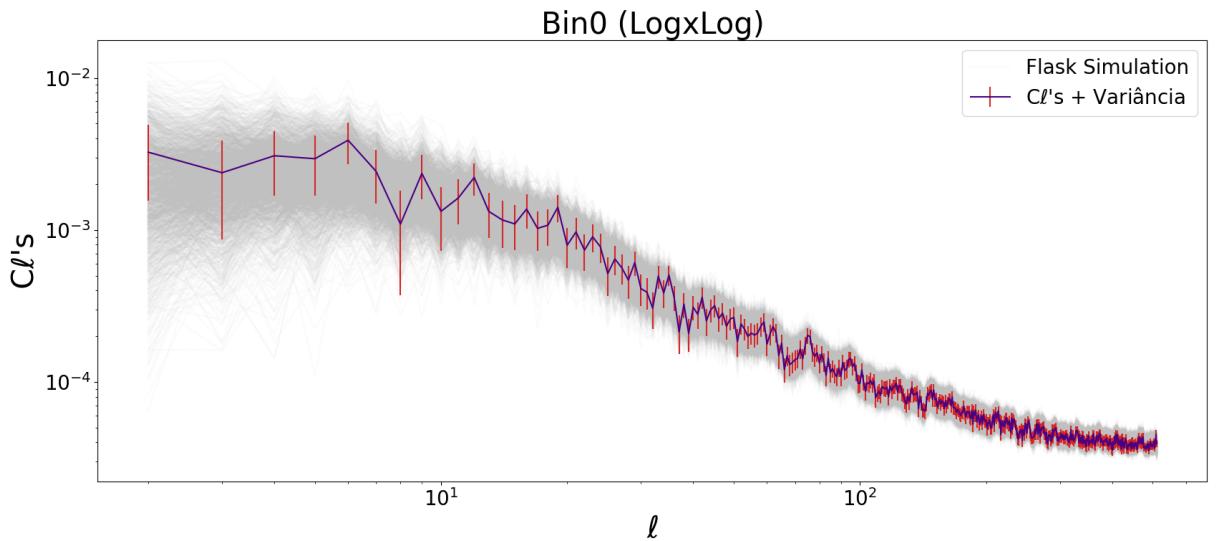


Figura 25:  $C_\ell$ 's dos mapas simulados pelo FLASK em cinza comparados com o  $C_\ell$  medido no mapa cosmológico de sobredensidade em violeta. As barras verticais são as medidas de variância entre as medidas de  $C_\ell$  dos mapas simulados naquele ponto.

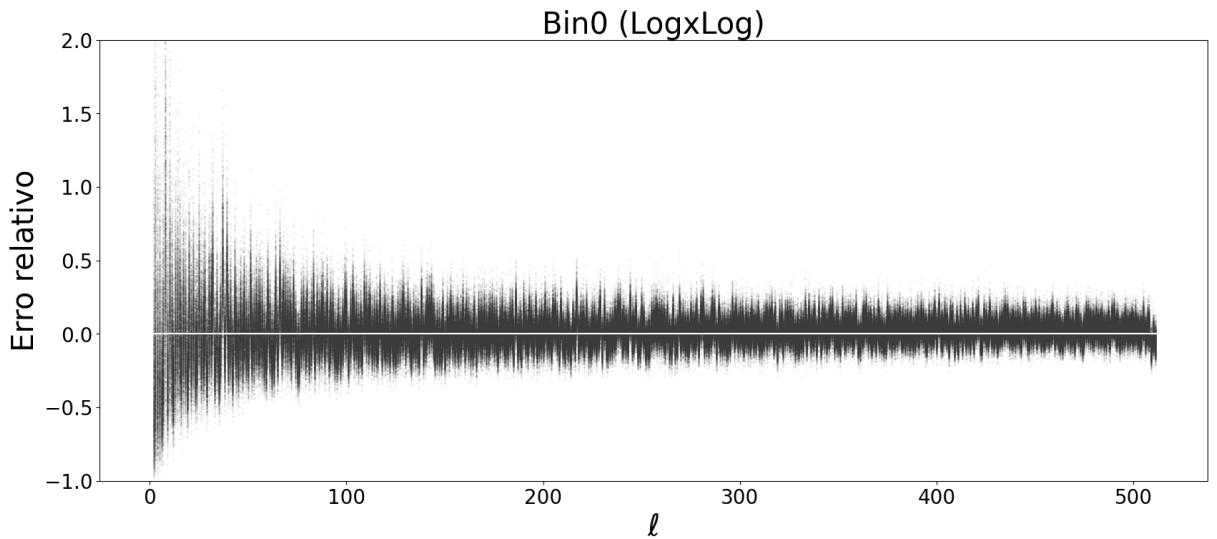


Figura 26: Erro relativo entre o  $C_\ell$  medidos através do mapa cosmológico e os  $C_\ell$ 's das simulações realizadas pelo FLASK.

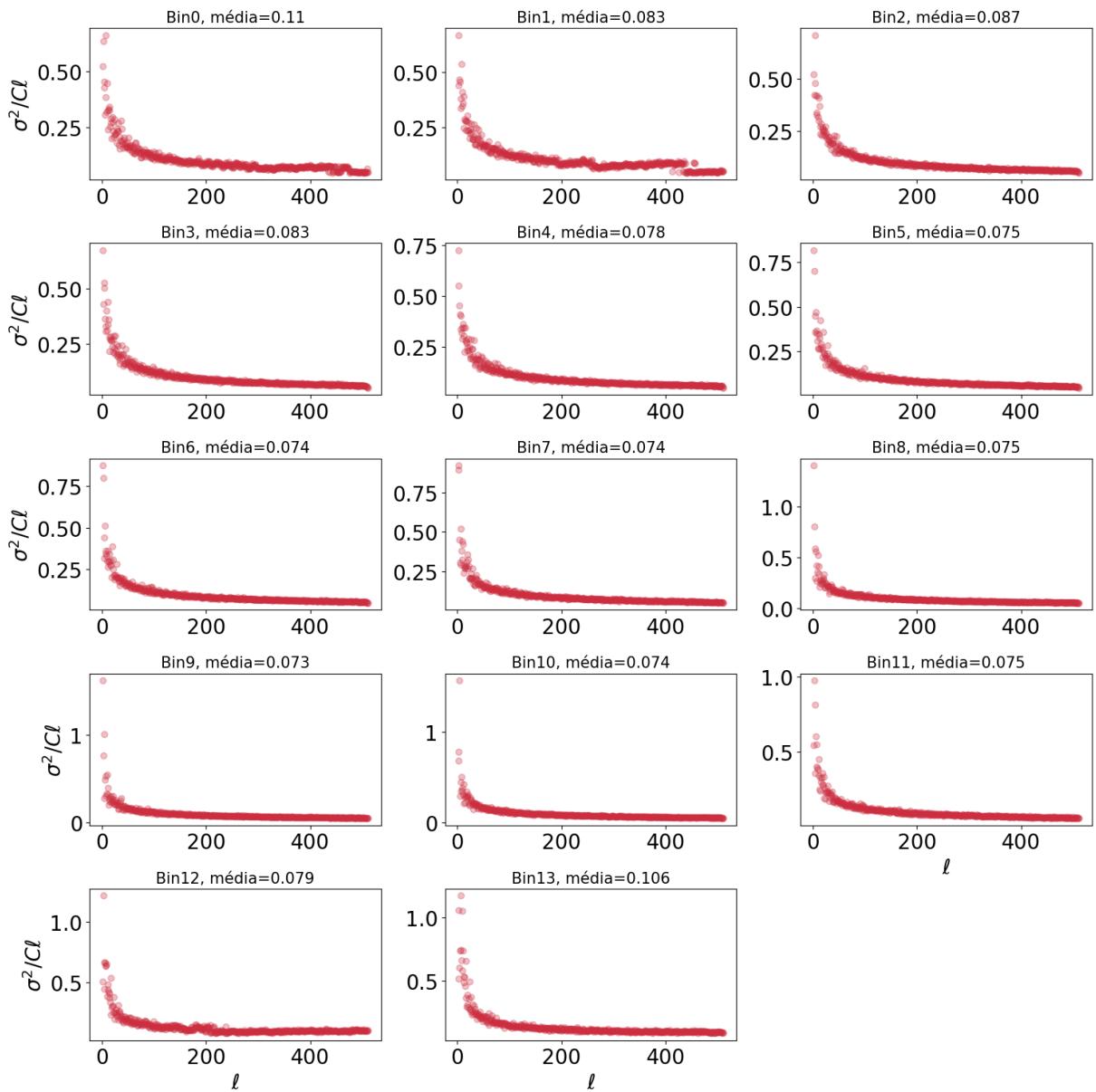


Figura 27: Variância normalizada pelo  $C_\ell$  medido a partir dos mapas cosmológicos.

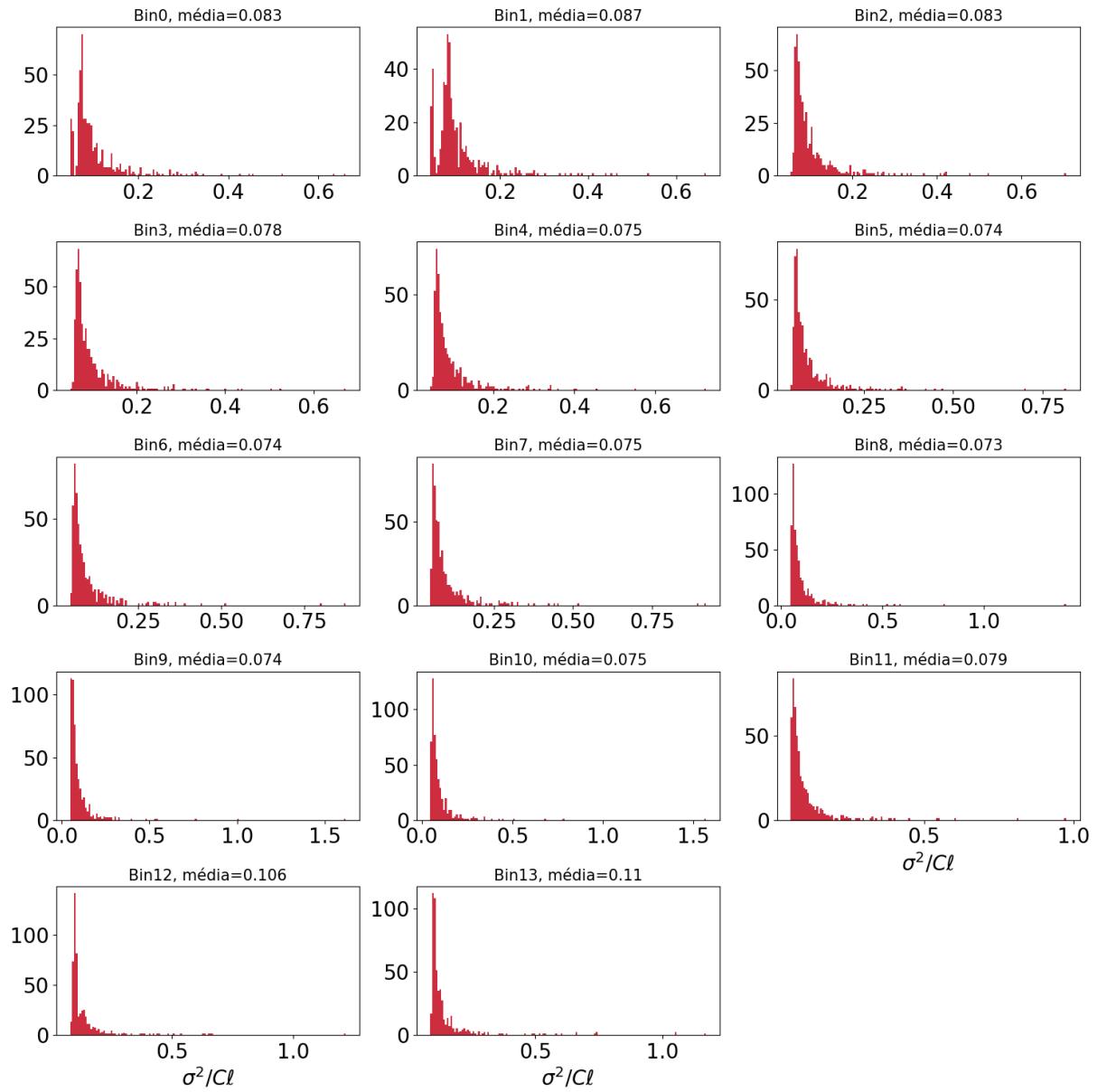


Figura 28: Distribuição da variância normalizada pelo  $C_\ell$  medido a partir dos mapas cosmológicos.

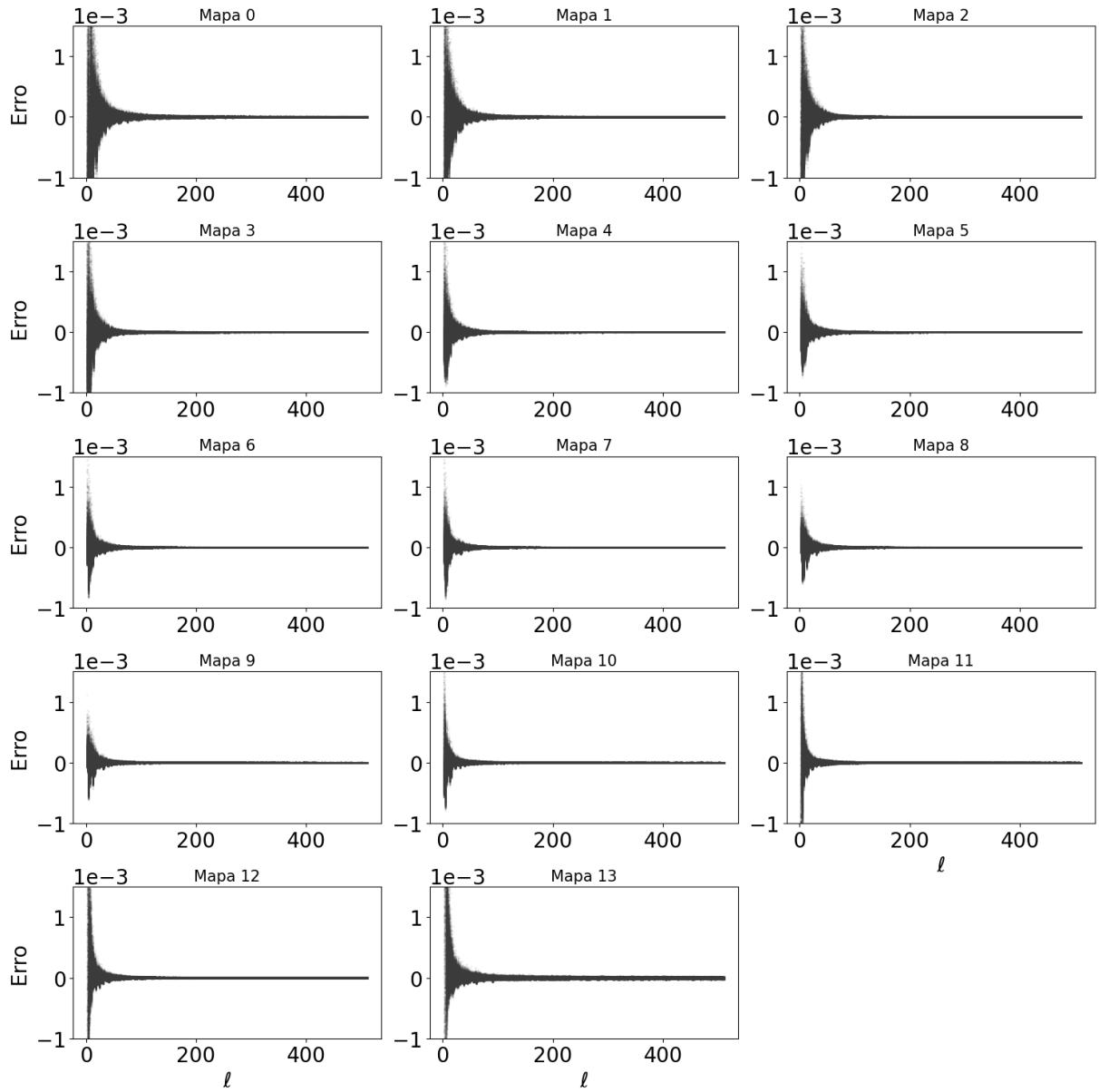


Figura 29: Erro entre os  $C_\ell$ 's medidos através dos mapas cosmológicos e os  $C_\ell$  das simulações realizadas pelo FLASK.

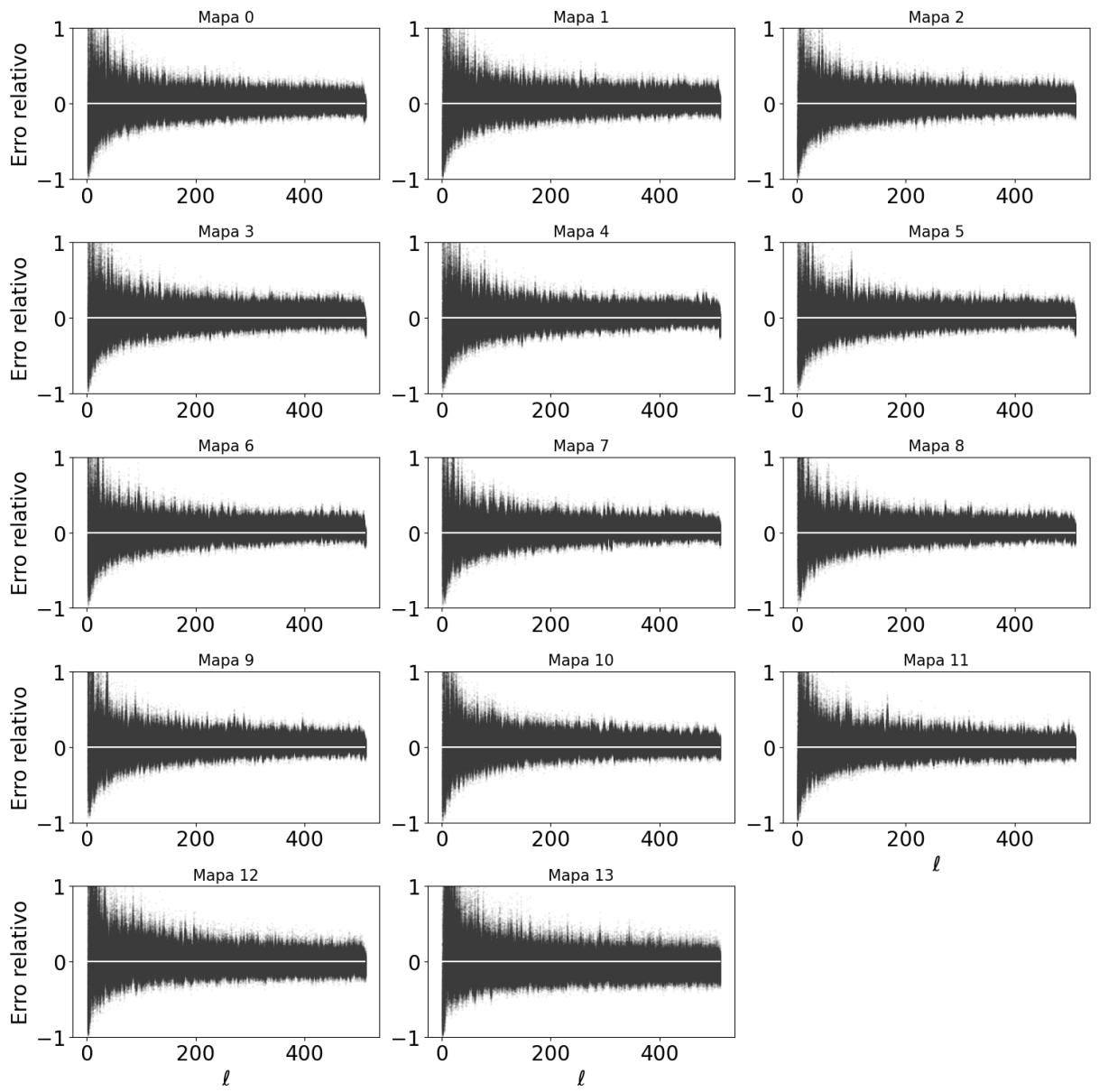


Figura 30: Erro relativo entre os  $C_\ell$ 's medidos através dos mapas cosmológicos e os  $C_\ell$  das simulações realizadas pelo FLASK.

### 4.3.2 $S_\ell$

Nesta etapa serão calculados os valores  $S_\ell$  dos mapas simulados. Para realizar essa medida é necessário obter o valor do shotnoise e subtrai-lo do  $C_\ell$ :

$$C_\ell = S_\ell + N_\ell. \quad (91)$$

Onde  $N_\ell$  é o shotnoise. O valor de shotnoise é dependente do número de galáxias da simulação. No processo aqui desenvolvido, os mapas simulados são gerados pelo FLASK primeiro como mapas de densidade de objetos (*density*), onde cada região no mapa vai ter um valor maior que zero associado, que representa o número de objetos na região; em seguida o mapa de densidade é transformado em um mapa de sobredensidade (*overdensity*) subtraindo cada objeto do vetor pela média total do mapa de densidade e dividindo pela média:

$$\text{sobredensidade} = \frac{\text{densidade} - \text{média}}{\text{média}}. \quad (92)$$

A partir do mapa de sobredensidade é mensurado o espectro de potência. Antes desta transformação, a informação do número de objetos do mapa é extraída do mapa de densidade simplesmente realizando a soma dos objetos e em seguida é calculado o valor do shotnoise conforme a equação 63. O número de objetos do mapa é controlado no FLASK pela função radial de seleção, como mencionado anteriormente. Esta função é fornecida no arquivo de configurações do FLASK , no nosso caso a função específica é o arquivo , que objetiva valores de acordo com a tabela 1 para os mapas simulados pelo FLASK. A função radial de seleção é ajustada de modo tal que o valor do número de galáxias não tenha divergência maior que 1% em relação aos mapas originais, este nível de precisão foi estabelecido arbitrariamente.

Para o caso dos  $S_\ell$  os erros envolvidos são mais proeminentes e tem maior sensibilidade ao número de objetos do mapa, começamos com o exemplo do primeiro intervalo de redshift (bin 0), com os valores de plotados na figura 32, com os valores mensurados do mapa cosmológico em violeta, a variância em vermelho e os mapas simulados do FLASK em cinza. Para esta simulação o erro relativo associado está plotado na figura 32. Os valores do erro relativo associados são muito mais protuberantes do que a medida do mesmo intervalo considerando o  $C_\ell$ . Esta medida também possui uma leve divergência nos valores para  $\ell$ 's mais altos, ao contrário do caso anterior em que os valores dos erros relativos eram convergentes (diminuíam cada vez mais conforme aumentava o valor de  $\ell$ ). Esta ultima característica é ainda mais proeminente em outros intervalos.

Através da variância relativa entre os  $S_\ell$  dos mapas simulados nota-se um ruido maior em gráficos que possuem menos objetos, como é o caso dos dois intervalos de redshift iniciais e dos três últimos. A medida que o  $\ell$  aumenta nota-se um aumento da variância relativa nestes casos.

O formato das distribuição das variâncias relativas se mantém semelhante ao caso em que foi medido o  $C_\ell$ , se assemelhando a uma distribuição lambda, porem neste caso o valor de lambda seria maior. Nota-se um aumento na variância relativa, principalmente nos quatro últimos mapas, onde anteriormente tínhamos valores majoritariamente abaixo de 0.3.

O erro simples apresentado na figura 35 também se mostra ineficaz neste caso. Observa-se no gráfico da variância relativa na figura 33 um aumento expressivo desta métrica para

valores maiores de  $\ell$ , um problema que fica explícito nos gráficos do erro relativo. Esse fenômeno, contudo, passa despercebido na métrica do erro simples.

O processo de mensuração do  $S_\ell$  envolve a subtração de valores em uma escala problemática em mapas em intervalos de redshift com pouca quantidade de objetos, o que em alguns casos leva a medidas que divergem muito do aceitável, que é o caso do intervalo de redshift treze por exemplo. Na figura 36 estão dispostos os erros relativos de todos os intervalos de redshift, o fenômeno da divergência começa com mais intensidade nos três últimos mapas, com a divergência crescendo conforme os valores de  $\ell$  aumentam.

O intervalo de redshift treze por exemplo apresenta um comportamento especialmente ruim devido principalmente a sua baixa quantidade de objetos e a forma como estes estão distribuídos. É importante ressaltar que este problema é causado pela forma como o código está lidando com valores pequenos e não pelo mapa em si.

O intervalo de redshift treze possui um erro relativo muito grande, pode ser visto com maiores detalhes na figura 38. Esse distúrbio trás consequências para a matriz de covariância, como por exemplo uma maior dificuldade em gerar uma matriz que possa ser inversível, um processo necessário para realização do cálculo dos parâmetros cosmológicos nas etapas subsequentes da pipeline.

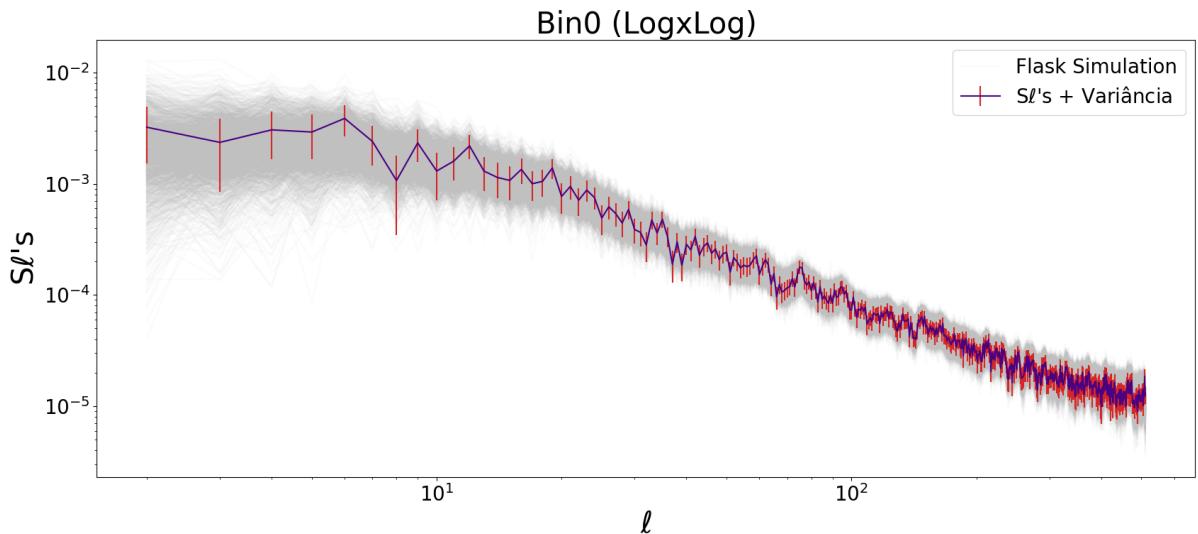


Figura 31:  $S_\ell$ 's dos mapas simulados pelo FLASK em cinza comparados com o  $S_\ell$  medido no mapa cosmológico de sobredensidade em violeta. As barras verticais são as medidas de variância entre as medidas de  $S_\ell$  dos mapas simulados naquele ponto.

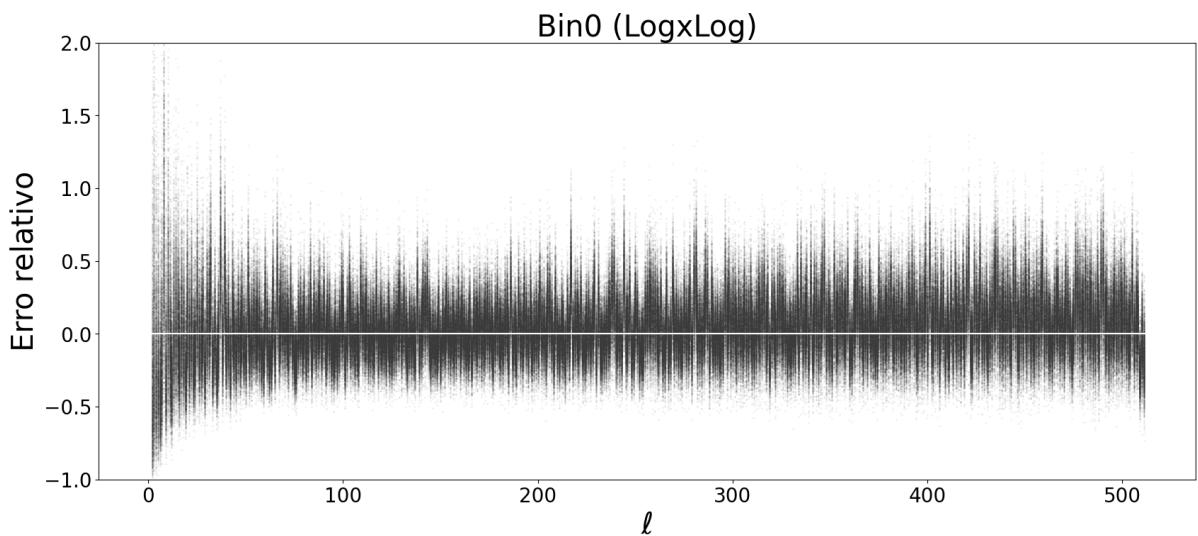


Figura 32: Erro relativo entre o  $S_\ell$  medido através do mapas cosmológico e os  $S_\ell$ 's das simulações realizadas pelo FLASK. Cada valor de  $\ell$  possui n-pontos, correspondentes a n-simulações geradas.

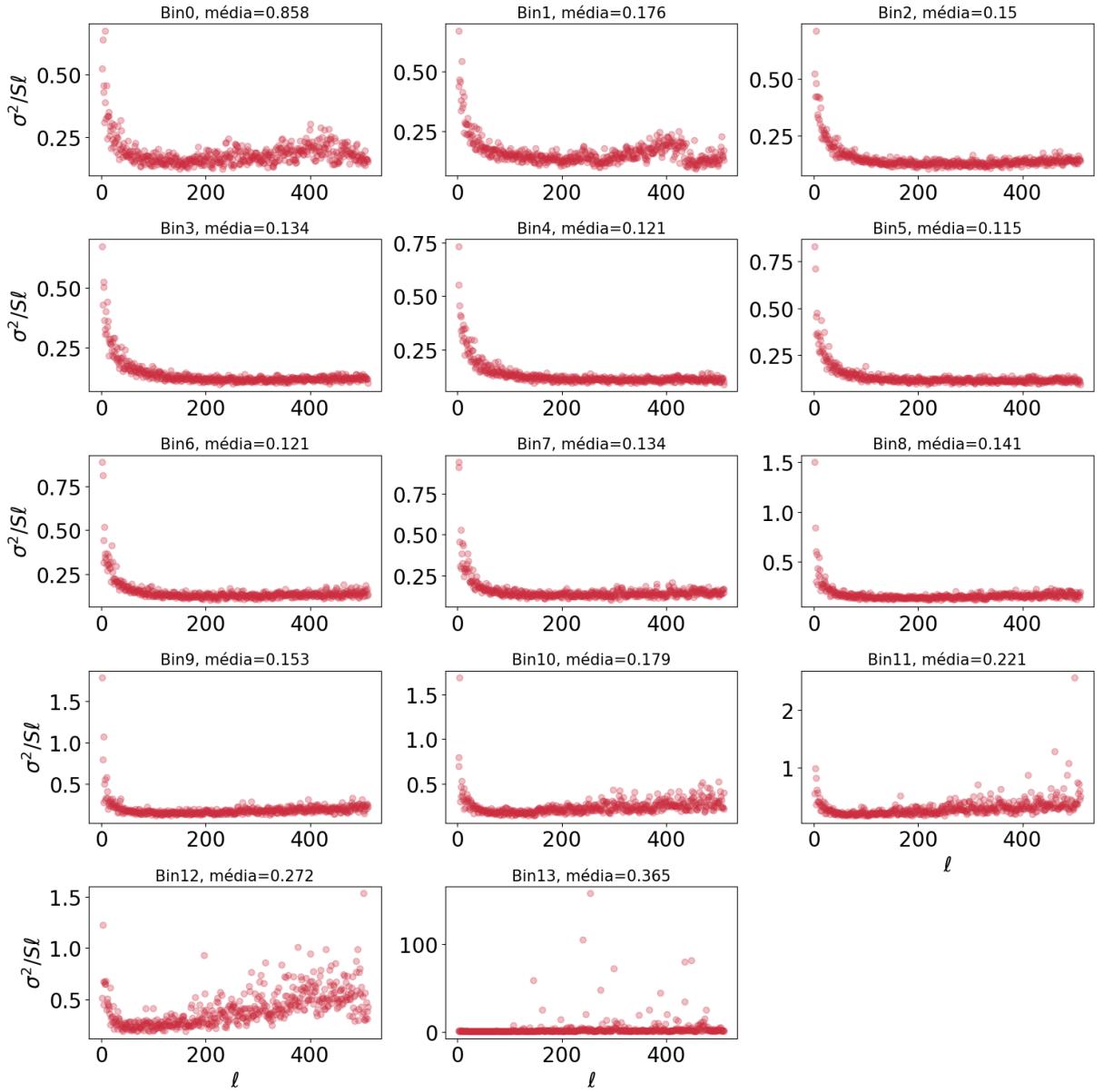


Figura 33: Variância normalizada pelo  $S_\ell$  medido a partir dos mapas cosmológicos para todos os intervalos de redshift

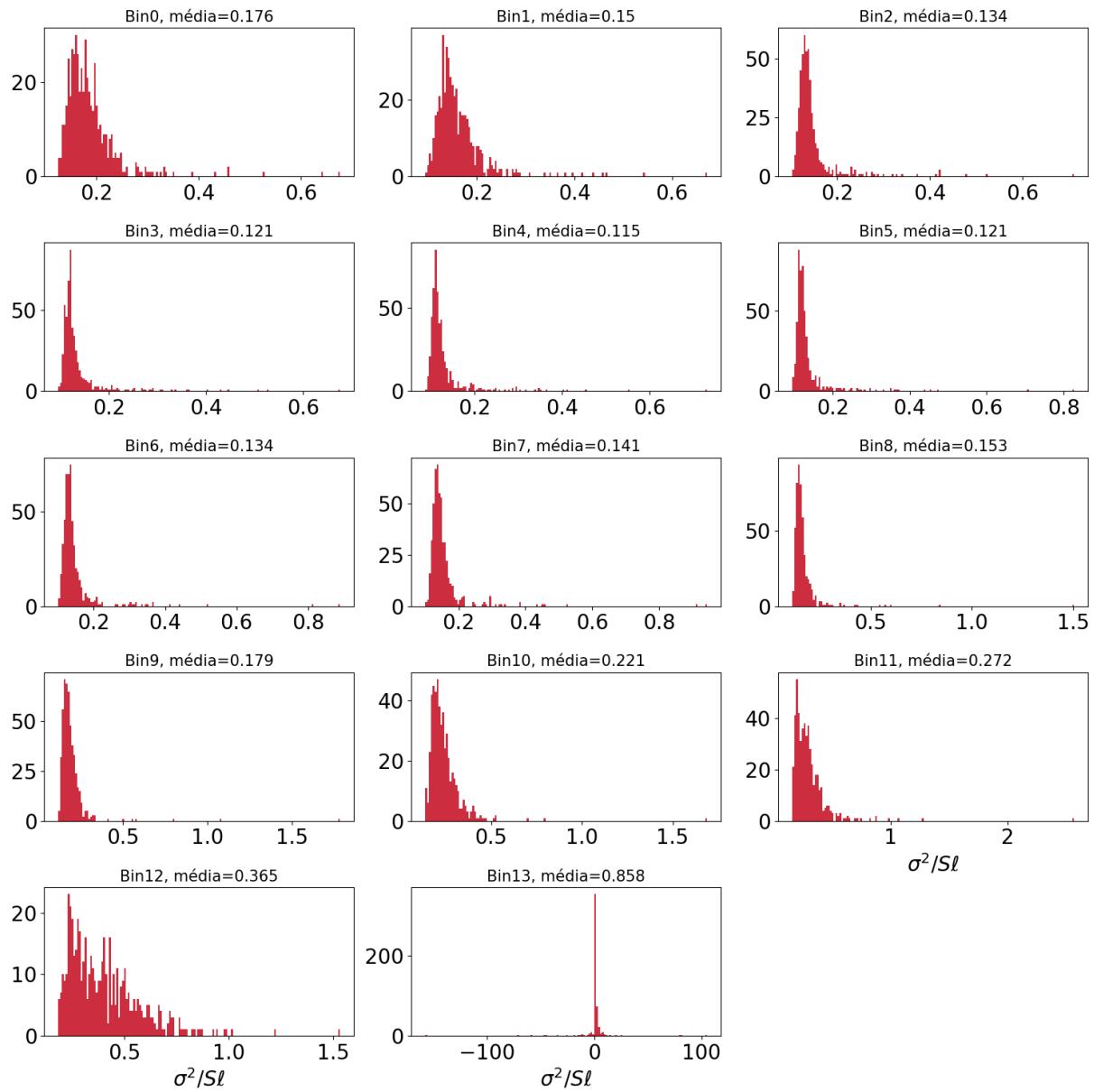


Figura 34: Distribuição da variância normalizada pelo  $S_\ell$  medido a partir dos mapas cosmológicos.

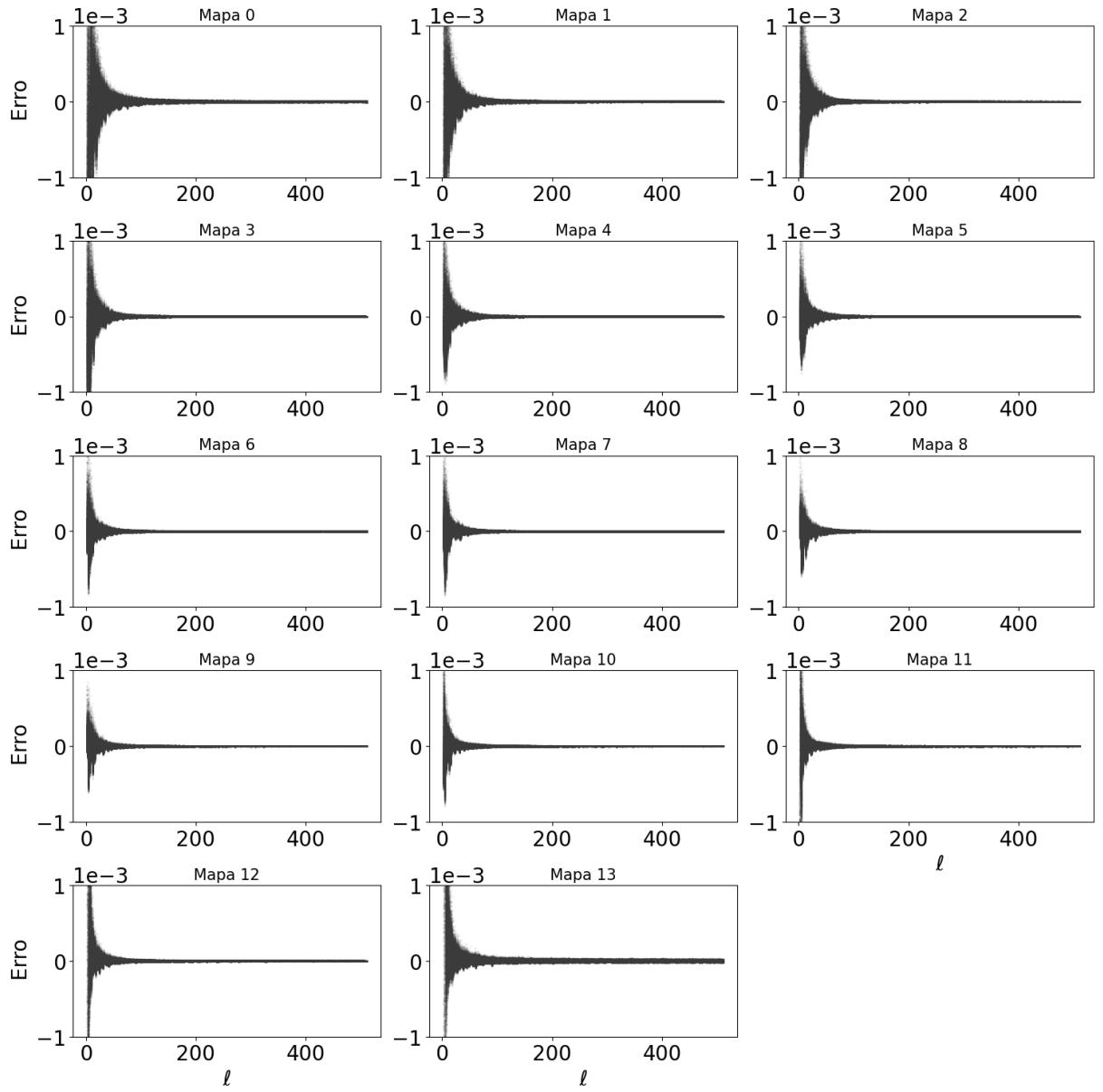


Figura 35: Erro entre os  $S_\ell$ 's medidas através dos mapas cosmológicos e os  $S_\ell$  das simulações realizadas pelo FLASK.

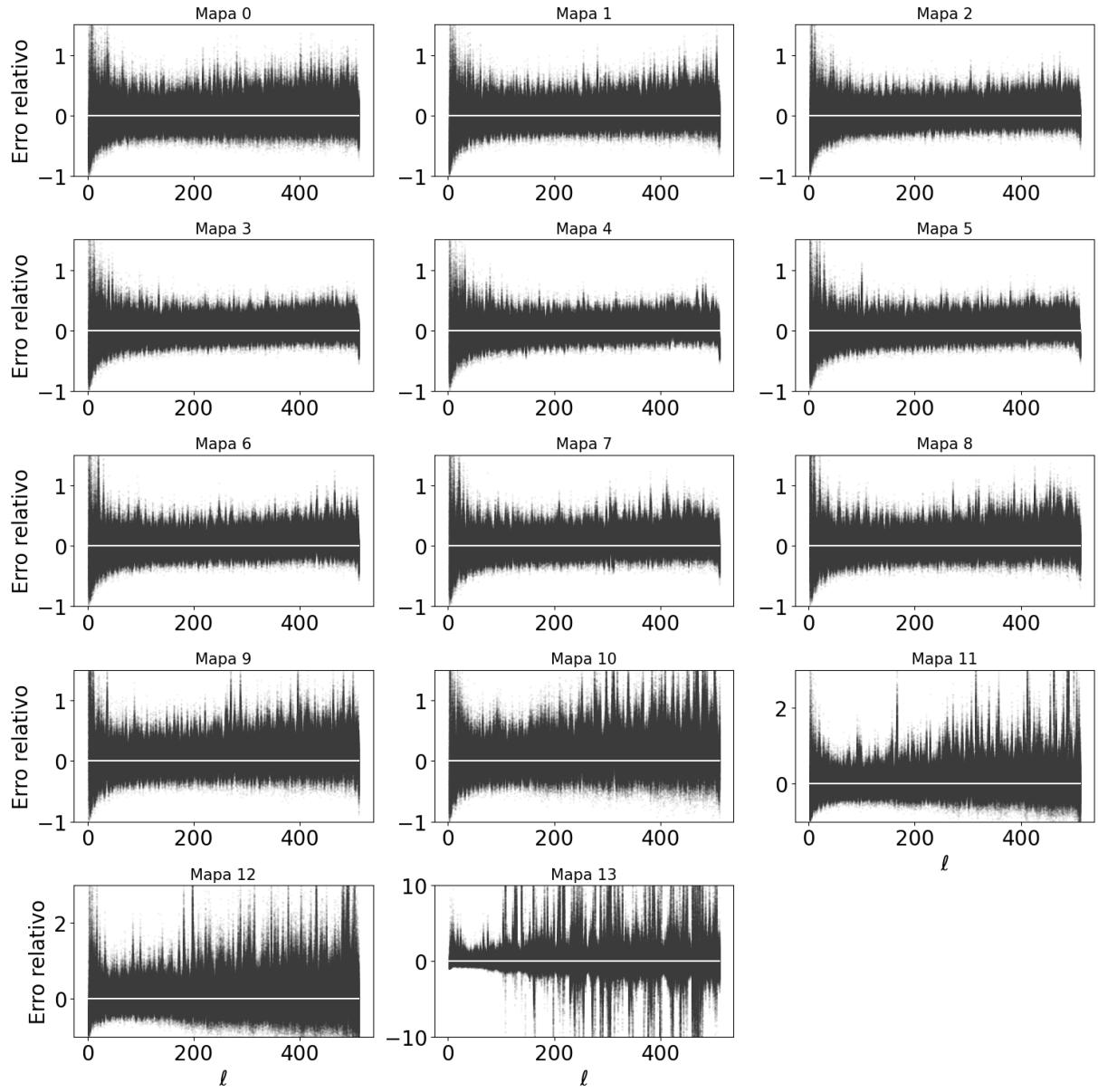


Figura 36: Erro relativo entre os  $S_\ell$ 's medidos através dos mapas cosmológicos e os  $S_\ell$  das simulações realizadas pelo FLASK.

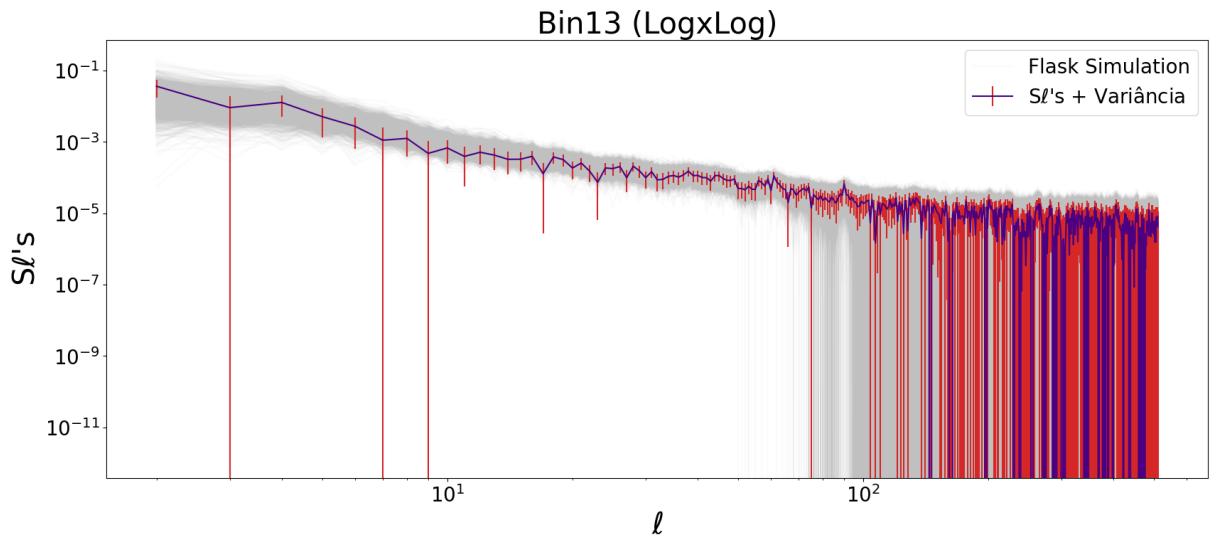


Figura 37:  $S_\ell$ 's dos mapas simulados pelo FLASK em cinza comparados com o  $S_\ell$  medido no mapa cosmológico de sobredensidade em violeta em uma situação onde o mapa base para as simulações possui poucos objetos no céu.

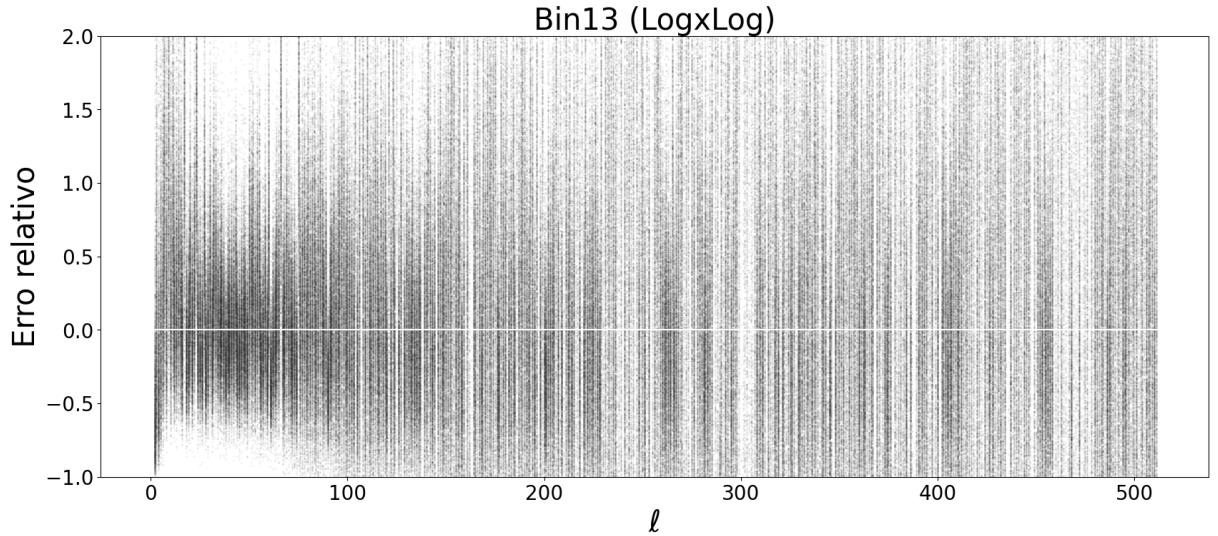


Figura 38: Erro relativo entre o  $S_\ell$  medidos através do mapa cosmológico e os  $S_\ell$ 's das simulações realizadas pelo FLASK em um mapa com poucos objetos no céu.

## 4.4 Resultados: $C_\ell$ 's entre mapas cosmológicos e sistemáticos

Para medir se algum desses efeitos sistemáticos está interferindo significativamente nos mapas de interesse, a correlação cruzada entre os mapas cosmológicos e os mapas sistemáticos é calculada. É esperado que não haja correlação entre o mapa do céu e um efeito sistemático, o que significa que nenhum destes efeitos está enviesando nossas medidas e portanto não irão interferir na estimativa de parâmetros cosmológicos nas etapas seguintes da pipeline. Para realizar este cálculo será utilizado o PseudoPower em conjunto com o [CovAnalysis](#) como ferramenta.

Outro aspecto pertinente é conhecer também qual o intervalo de confiança dessa medida, para tal, utilizamos as medidas de correlação extraídas dos mapas do céu para gerar n-mapas simulados do céu utilizando o FLASK, para cada mapa simulado calcula-se novamente as correlações cruzadas com os mapas sistemáticos, de modo a gerar um conjunto de medidas de correlação a qual poderemos medir a variância ponto a ponto para estimar um intervalo de confiança.

Para cada gráfico na figura 39 temos todas as correlações cruzadas entre os mapas cosmológicos e os mapas sistemáticos, com o eixo horizontal em escala logarítmica, apenas para evidenciar as múltiplas curvas. O comportamento de todos os gráficos é muito semelhante, sendo característicos de um sinal aleatório, isto é, um pico intenso no inicio e rapidamente decaindo para próximo de zero. Portanto, não temos evidências para aceitar a hipótese de que os mapas sistemáticos são correlacionados com os mapas cosmológicos. Evidenciando a característica de sinal aleatório, as correlações da figura 40 estão plotadas em escala normal.

Individualmente, as correlações podem ser avaliadas com as mesmas métricas utilizadas até então. Na figura 41 foi plotado a correlação cruzada do primeiro intervalo de redshift (bin 0, ou mapa 0) com os mapas sistemáticos em escala normal, para evidenciar o caráter de sinal aleatório. Na figura 42 foi adicionada a variância das simulações. Na figura 43 temos a análise da variância relativa, que apresenta um comportamento excelente, possuindo alguns ponto fora da curva, mas ainda sim com uma distribuição uniforme nos valores de  $\ell$ . A figura 44 representa a distribuição das variâncias relativas, se assemelhando também a uma distribuição lambda, assim como nos outros casos estudados.

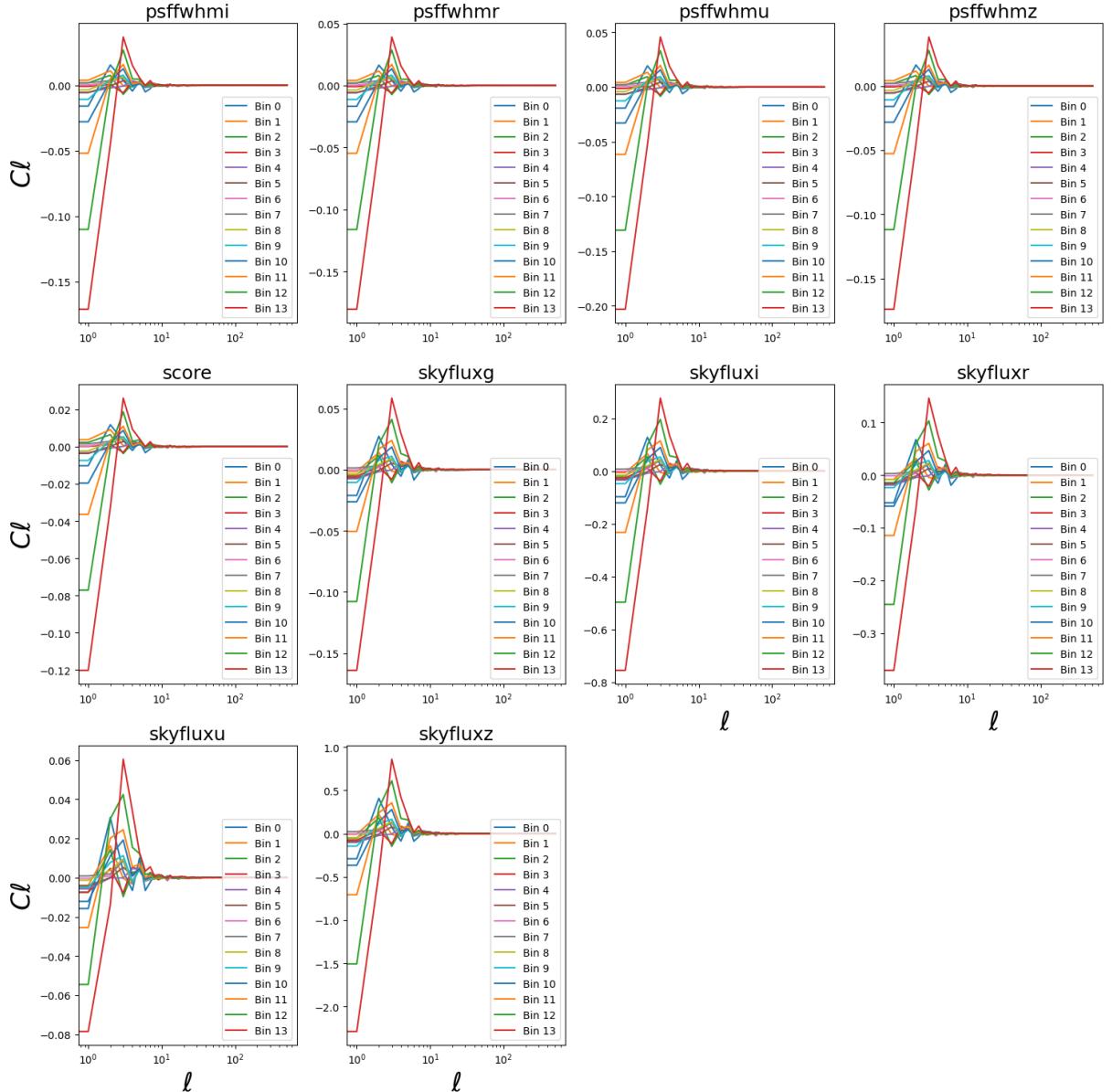


Figura 39: Pseudo- $C_\ell$  entre os mapas cosmológicos e os mapas sistemáticos em log no eixo X.

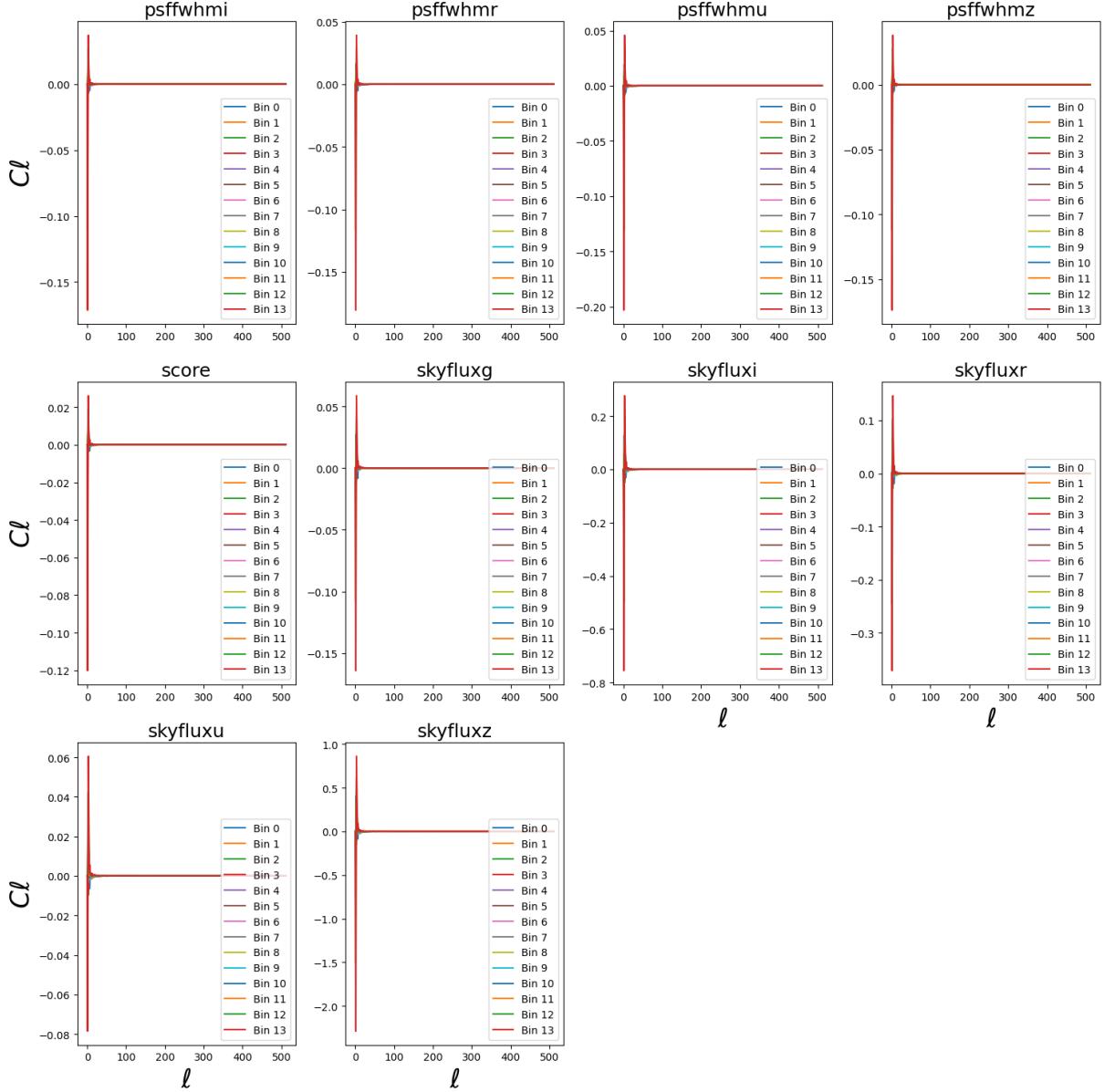


Figura 40: Pseudo- $C_\ell$  entre os mapas cosmológicos e os mapas sistemáticos em escala normal.

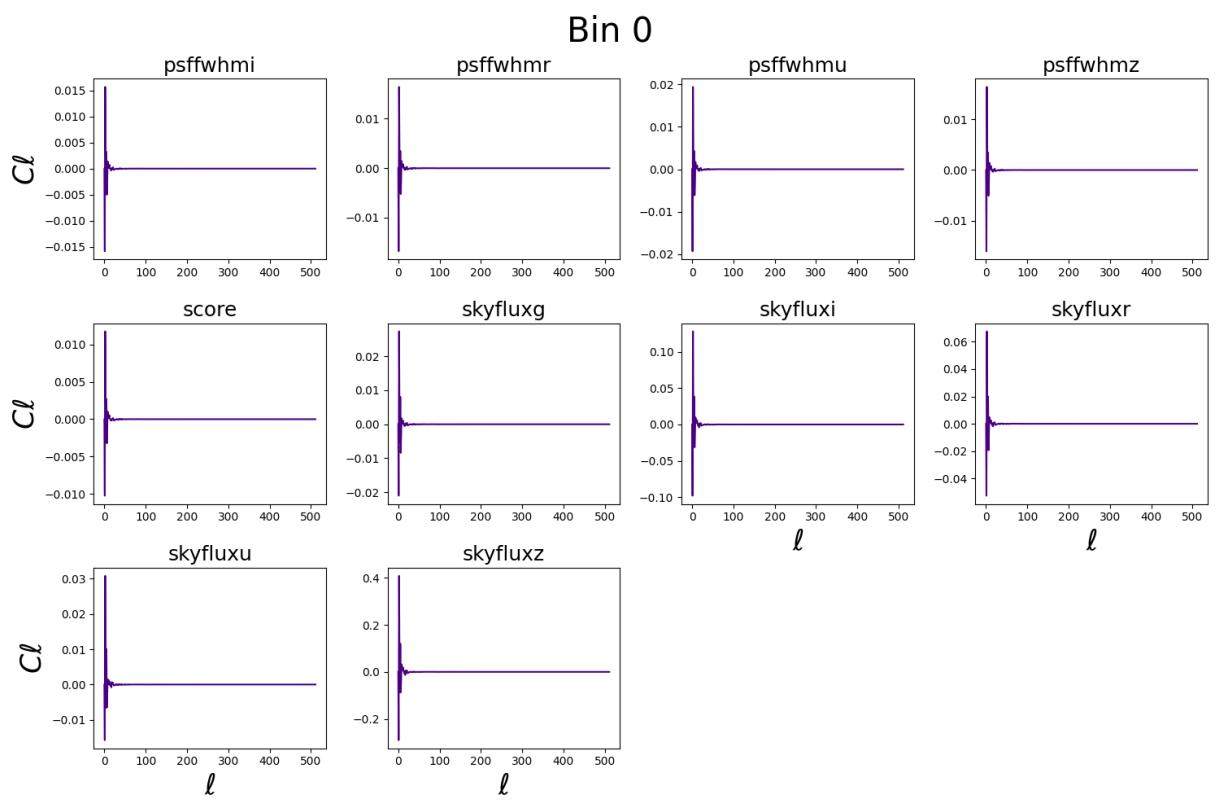


Figura 41: Pseudo- $C_\ell$  entre o mapa cosmológicos e os mapas sistemáticos em escala normal.

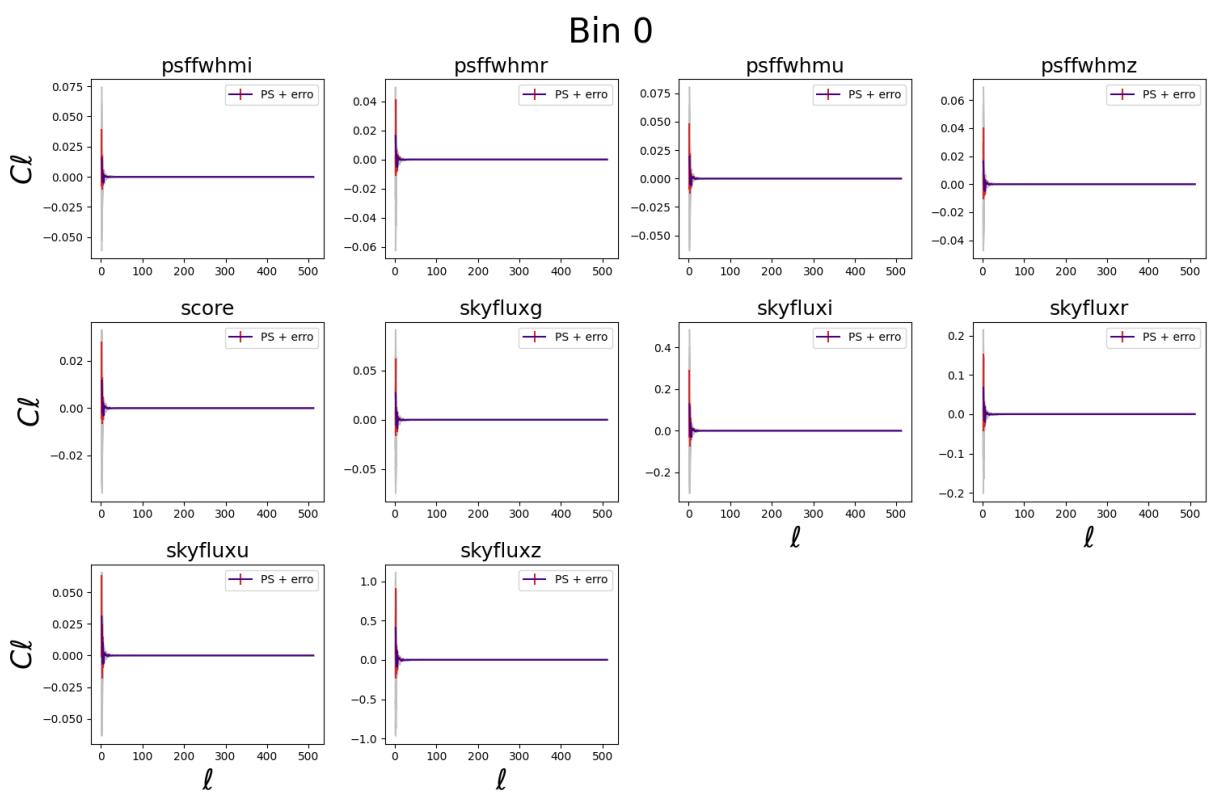


Figura 42: Pseudo- $C_\ell$  entre o mapa simulados e os mapas sistemáticos em escala normal.

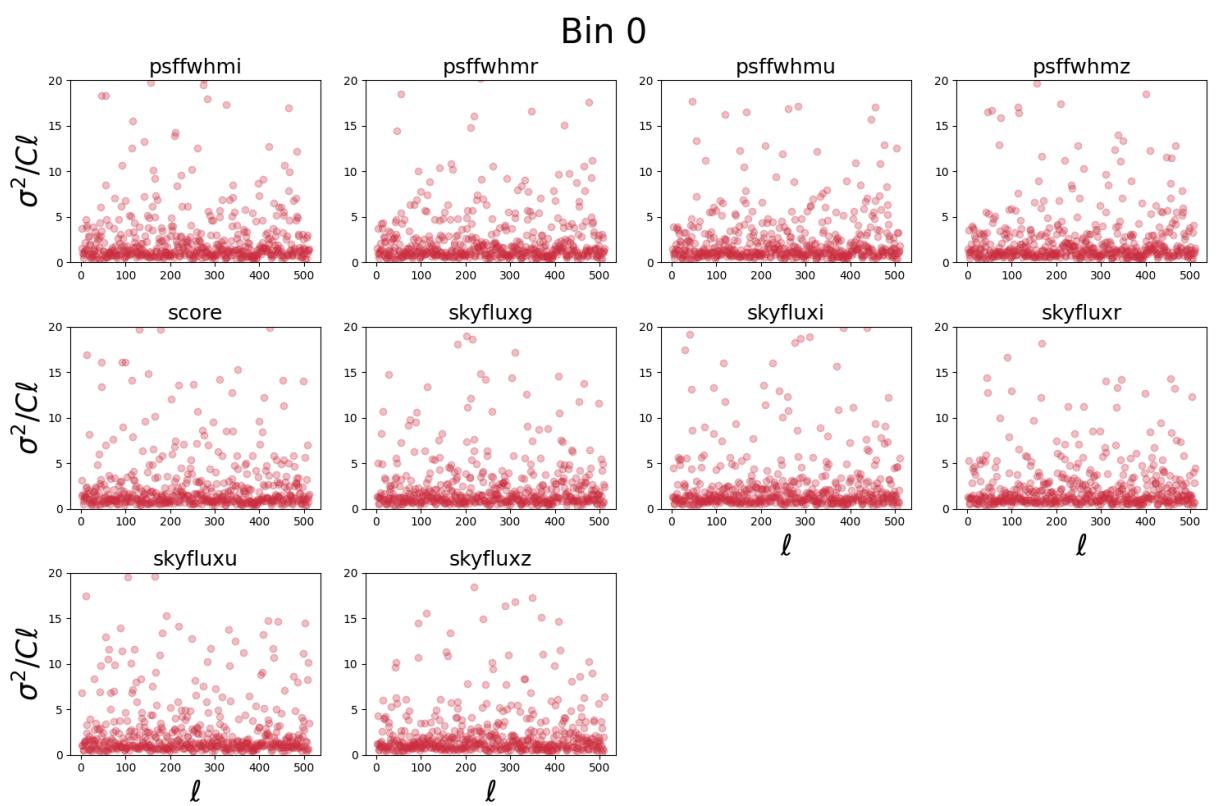


Figura 43: Variância entre as medidas de  $C_\ell$ 's dos mapa simulados e os mapas sistemáticos.

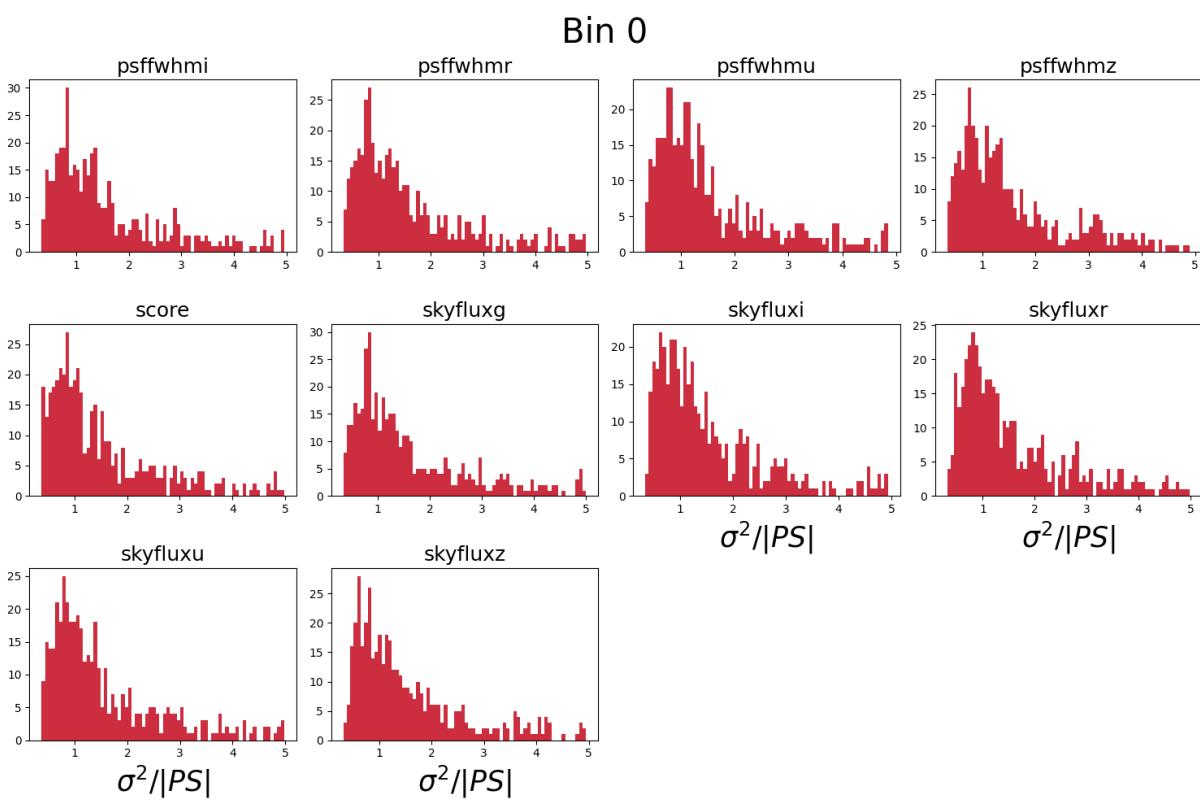


Figura 44: Distribuição da variância entre as medidas de  $C_\ell$ 's dos mapa simulados e os mapas sistemáticos.

## 5 Matriz de covariância

A covariância entre duas variáveis aleatórias  $\vec{X}$  e  $\vec{Y}$  com N elementos é definida como:

$$cov(\vec{X}, \vec{Y}) = E \left[ (\vec{X} - E(\vec{X})) (\vec{Y} - E(\vec{Y}))^T \right]. \quad (93)$$

Onde  $E$  é o operador valor esperado. Essa equação resulta em uma matriz de dimensão N x N.

A partir das medidas de  $C_\ell$  e  $S_\ell$  calculadas através das N-simulações do FLASK são estimadas as matrizes de covariância. Para tal estabelece-se um somatório sobre todas as amostras com uma normalização amostral utilizando a mesma estrutura da definição de covariância da equação 93. A covariância é uma matriz de dimensão N x N para cada par de correlações, a matriz completa leva em conta todos os  $S\ell$  (ou  $C_\ell$ ) cruzados, (e.g.:  $S_\ell^{1,2}$  = mapa-1 correlação cruzada com mapa-2), onde cada elemento da matriz completa é formado por soma de quatro termos da forma

$$S_{\ell\ell'}^{ij} = \frac{1}{N_s - 1} \sum_{s=0}^{N_s} (S_\ell^{ij,s} - \langle S_\ell^{ij,s} \rangle) (S_{\ell'}^{ij,s} - \langle S_{\ell'}^{ij,s} \rangle)^T, \quad (94)$$

$$C_{\ell\ell'}^{ij} = \frac{1}{N_s - 1} \sum_{s=0}^{N_s} (C_\ell^{ij,s} - \langle C_\ell^{ij,s} \rangle) (C_{\ell'}^{ij,s} - \langle C_{\ell'}^{ij,s} \rangle)^T, \quad (95)$$

onde  $S_{\ell\ell'}^{ij}$  e  $C_{\ell\ell'}^{ij}$  referem-se aos dados sem shotnoise e com shotnoise respectivamente,  $N_s$  é o número de simulações,  $s$  diz respeito as simulações,  $\ell$  e  $\ell'$  são os harmônicos de cada conjunto de dados,  $\langle S \rangle$  denota que os objetos estão sendo subtraídos de seus respectivos valores médios e  $T$  representa a transposta dos dados. O resultado é uma matriz quadrada do mesmo tamanho da extensão dos harmônicos. A matriz final é triangular simétrica e contém todas essas matrizes calculadas através de todas as correlações. Um exemplo da matriz final para o caso em que se tem apenas três mapas pode ser visto na tabela 4, onde cada elemento é uma matriz de acordo com a equação 95. A matriz resultante é triangular simétrica e leva em consideração todos as correlações cruzadas.

	$S_\ell^{0,0}$	$S_\ell^{0,1}$	$S_\ell^{0,2}$	$S_\ell^{1,1}$	$S_\ell^{1,2}$	$S_\ell^{2,2}$
$S_{\ell'}^{0,0}$	<b>00,00</b>	00,01	00,02	00,11	00,12	00,22
$S_{\ell'}^{0,1}$	01,00	<b>01,01</b>	01,02	01,11	01,12	01,22
$S_{\ell'}^{0,2}$	02,00	02,01	<b>02,02</b>	02,11	02,12	02,22
$S_{\ell'}^{1,1}$	11,00	11,01	11,02	<b>11,11</b>	11,12	11,22
$S_{\ell'}^{1,2}$	12,00	12,01	12,02	12,11	<b>12,12</b>	12,22
$S_{\ell'}^{2,2}$	22,00	22,01	22,02	22,11	22,12	<b>22,22</b>

Tabela 4: Matriz de covariância completa considerando apenas três mapas, onde cada elemento é uma matriz com o índice combinado das correlações. A matriz é triangular simétrica e considera todas as correlações cruzadas.

Com a finalidade de avaliar os resultados da matriz de covariância das N-simulações pode-se estabelecer uma matriz de covariância teórica para efeitos de comparação e validação. Seguindo os resultados de Loureiro-2019, a equação da matriz de covariância

teórica é dada por:

$$\Sigma_{\ell\ell'}^{ij} = \frac{R_{\ell\ell'}}{f_{sky}(2\ell'+1)} \left[ (w_\ell^2 S_\ell^{ij} + N_\ell \delta_{ij}) (w_{\ell'}^2 S_{\ell'}^{ij} + N_{\ell'} \delta_{ij}) + (w_\ell^2 S_\ell^{ii} + N_\ell \delta_{ii}) (w_{\ell'}^2 S_{\ell'}^{jj} + N_{\ell'} \delta_{jj}) \right] , \quad (96)$$

sendo  $w_\ell^2$  definido pela equação 60,  $f_{sky}$  a porcentagem do céu que a máscara (figura 100) dos dados abrange,  $N_\ell$  o shotnoise e  $R_{\ell\ell'}$  é a matriz mistura, expressa por

$$R_{\ell\ell'} = \frac{2\ell'+1}{4\pi} \sum_{\ell''} (2\ell''+1) W_{\ell''} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix} , \quad (97)$$

onde a matriz 2x3 é a função Wigner 3j e possui uma descrição mais detalhada em [Dahlen-2008](#). O fator  $W_{\ell''}$  depende apenas da máscara dos dados,

$$W_\ell = \sum_{m=-\ell}^{\ell} \frac{|I_{\ell m}|^2}{(2\ell+1)}, \quad (98)$$

$$I_{\ell m} = \sum_p^{N_{pix}} Y_{\ell m}^\dagger(\theta_p, \phi_p) \Delta\Omega_p . \quad (99)$$

(100)

Deste modo,  $\Delta\Omega_p$  representa o angulo sólido de um pixel  $p$ . O gráfico da matriz mistura referente a máscara da figura 100 está plotado na figura 47 para diversos valores de  $\ell'$ .

## 5.1 Resultados

Na figura 45 estão plotados os dados obtidos através dos  $C_\ell$ 's referentes ao primeiro intervalo de redshift. O primeiro gráfico é a comparação do  $C_\ell$  obtido através dos mapas cosmológicos comparado com a média dos  $C_\ell$ 's mensurados a partir dos mapas simulados do FLASK. É esperado que estes valores sejam semelhantes, pois aqui o espectro de potência é base para a criação dos mapas, que vão diferir entre si de modo aleatório (aleatoriedade gaussiana e com distribuição com média zero), conforme explicitado pela equação 81.

O segundo e o terceiro gráfico (gráficos centrais) são os elementos matriciais da diagonal principal da matriz de covariância, neste caso seria o elemento 00,00 no exemplo da tabela 4. A esquerda temos a matriz de covariância teórica, mensurada diretamente dos mapas cosmológicos. A direita temos a matriz de covariância extraída das simulações do FLASK.

O quarto gráfico é a comparação entre as diagonais das matrizes estimadas pelas simulações e as matrizes teóricas. Essa comparação nos fornece um parâmetro de comparação, que aqui será nomeado erro teórico. Quanto mais próximas as diagonais menor será o erro.

A figura 46 realiza as mesmas medidas para o caso onde se medem os  $S_\ell$ 's. Os gráficos de todos os intervalos de redshift para os  $C_\ell$ 's e os  $S_\ell$ 's estão dispostos no [Anexo V](#) e [Anexo VI](#) respectivamente.

Para os casos em que se mede as matrizes a partir do  $C_\ell$ , as diagonais principais das matrizes estiveram muito próximas uma da outra, aumentando sua distância levemente

nas situações em que os mapas possuem menos objetos, mantendo consistência com as medidas de erro e variância relativos referentes.

Nos casos em que se medem as matrizes a partir do  $S\ell$ , as diagonais principais das matrizes estiveram mais afastadas comparativamente, com a distância entre elas ficando especialmente grande para os mapas com poucos elementos, também mantendo a consistência com a análise do erro e variância relativos, onde se podia observar um comportamento inadequado dos dados.

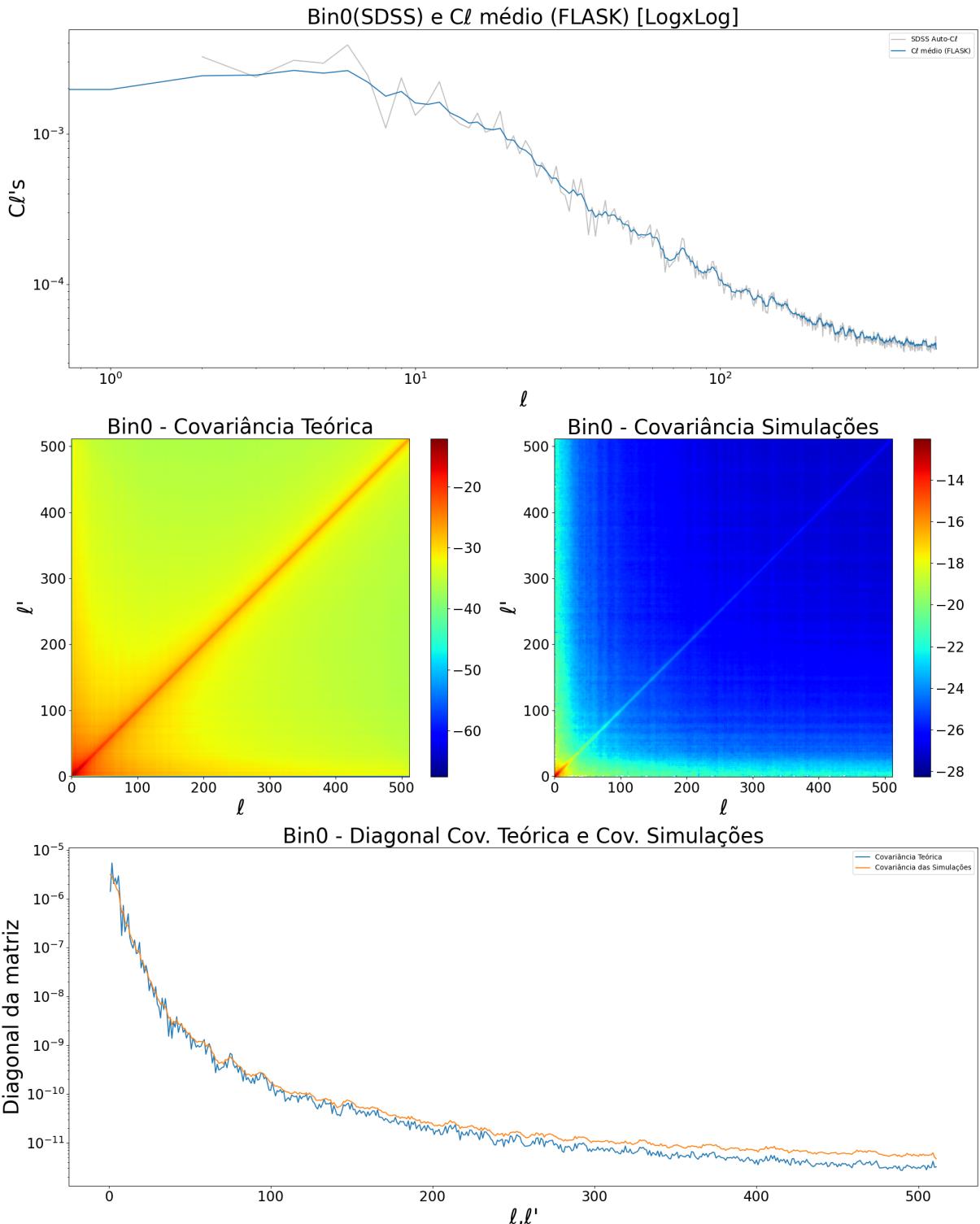


Figura 45: No primeiro gráfico há uma comparação entre o  $C\ell$  médio das simulações do FLASK e o  $C\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados e a matriz de covariância extraída das simulações. No ultimo gráfico temos a comparação das diagonais das duas matrizes.

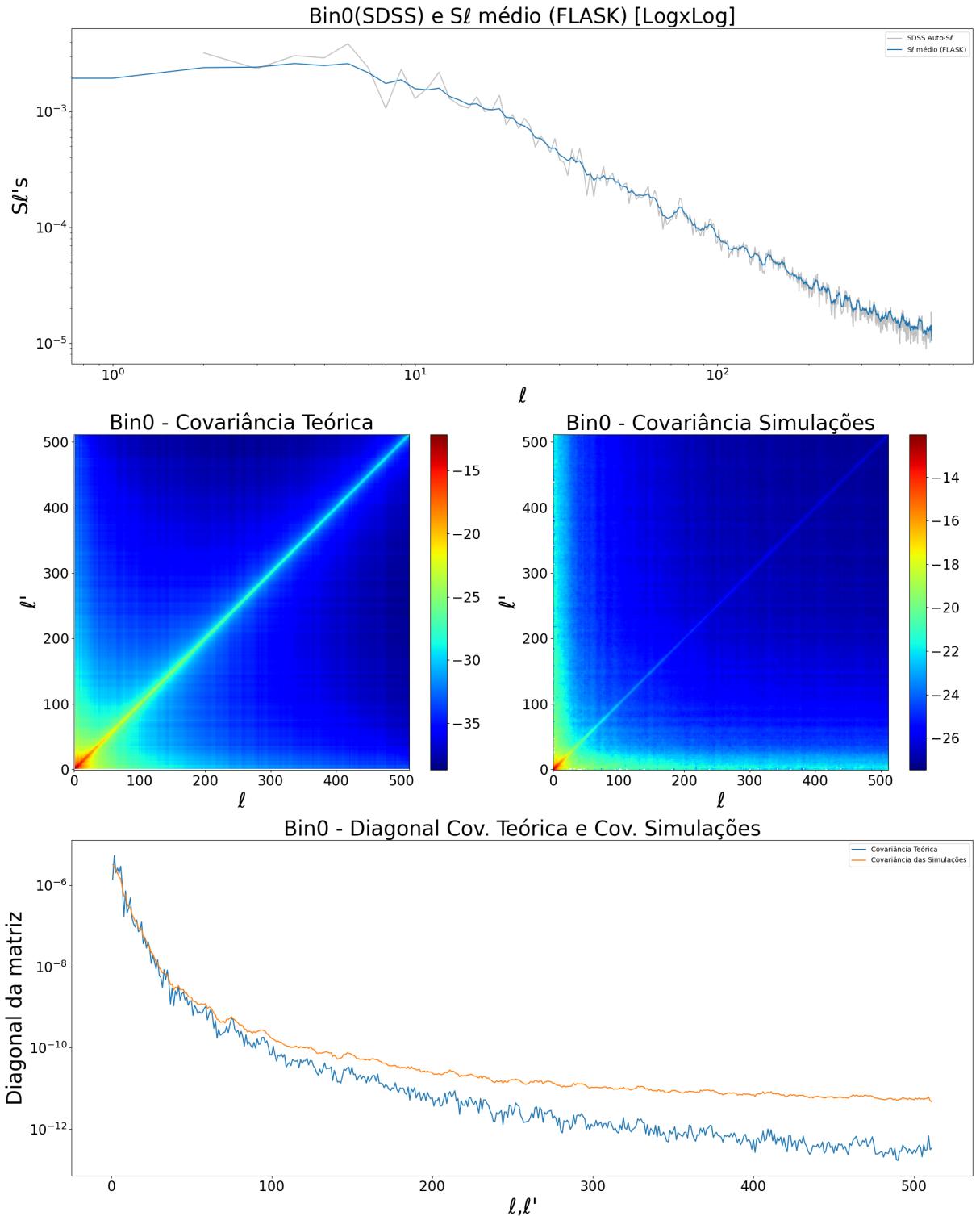


Figura 46: No primeiro gráfico há uma comparação entre o  $S_l$  médio das simulações do FLASK e o  $S_l$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

## 5.2 Binning: Compactando informação

O custo computacional para realizar todos os cálculos para estimar a matriz de covariância e posteriormente estimar os parâmetros cosmológicos é muito alto. Com a finalidade de reduzi-lo realiza-se um processo denominado *binning*, que consiste em escolher um valor do espectro de potência para representar um determinado intervalo  $\Delta\ell$ , reduzindo assim a quantidade de dados totais. A operação é ordenada pela equação

$$S_{\Delta\ell}^{ij} = \frac{\sum_{\ell \in \Delta\ell} (2\ell + 1) S_{\ell}^{ij}}{\sum_{\ell \in \Delta\ell} (2\ell + 1)}. \quad (101)$$

Para escolher o intervalo  $\Delta\ell$  para redução dos dados, utiliza-se a matriz mistura, plotada na figura 47. Ela fornece a informação referente a que tamanho de intervalo não causa perda excessiva de informações, dada a geometria da máscara dos dados. No nosso caso escolhemos o valor de  $R_{\ell\ell'} = 0.0105$ , sendo o valor associado de  $\Delta\ell = 9$  nesta condição. Outros cuidados também devem ser levados em conta para a diminuição da quantidade de dados, como por exemplo a quantidade correta de casas decimais armazenadas nos elementos das matrizes.

A compactação de dados é de extrema importância na parte operacional, pois para alguns sistemas o número de simulações  $S$  necessárias depende do número  $n$  de mapas e da quantidade de linhas  $L$  do espectro de potência (que ditará a dimensão das matrizes de covariância),

$$S = 3L \frac{n(n+1)}{2} \quad . \quad (102)$$

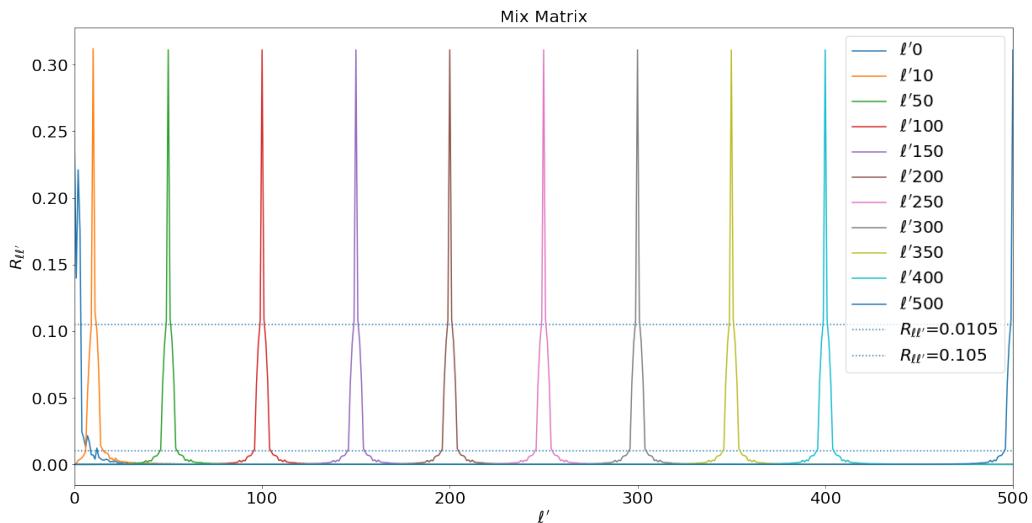


Figura 47: Matriz mistura para diversos valores de  $\ell'$ .

## 6 Conclusão

O espectro de potência foi estimado por três códigos distintos, o Pseudo-Power, Healpy (com os devidos ajustes) e o UCLCL, de modo a haver concordância completa entre os resultados do Pseudo-Power, Healpy e os resultados o UCLCL.

Os métodos de reamostragem são uma ferramenta adequada para gerar intervalos de confiança para os dados cosmológicos, dada a natureza das dificuldades de se realizarem as respectivas medidas experimentais neste campo de conhecimento. Avaliamos a qualidade das medidas do espectro de potência utilizando a variância relativa e o erro relativo, que se mostram mais adequados. As medidas de Pseudo- $C\ell$  possuem erro relativo majoritariamente abaixo de um e possui um comportamento adequado de variância relativa, sendo maior nos momentos menores ( $\ell$ 's baixos) e com comportamento assintótico. Para as medidas de Pseudo- $C\ell$  o comportamento é semelhante para os mapas que possuem números suficientes de objetos, mas em caso contrário o erro relativo aumenta demasiadamente e a variância começa a aumentar em  $\ell$ 's maiores, deixando de decair assintoticamente nestes casos.

Foram efetuadas as medidas de correlação entre os mapas cosmológicos e os mapas sistemáticos, os resultados encontrados são característicos de um sinal aleatório (não correlacionado). Não há evidências para rejeitar a hipótese de que os efeitos sistemáticos não estão enviesando os dados em análise. Os mapas SDSS não possuem correlação com os mapas sistemáticos, portanto não carregam informações indesejáveis para nossa análise de correlação.

As matrizes de covariância foram estimadas e apresentaram um comportamento melhor para os dados com maior número de objetos, que eram justamente aqueles cuja variância relativa dos espectros de potência variavam menos proporcionalmente, visto que a média das simulações continuam muito próximas dos  $C\ell$ 's mensurados dos mapas cosmológicos.

A utilização do código desenvolvido requer um cuidado quanto à separação dos dados cosmológicos em intervalos de redshift para que eles tenham a quantidade mínima de objetos para realizar o processo. Os valores dependerão de quanta variância será tolerada nas medidas de  $C\ell$  das simulações pela equipe de pesquisadores. O número de simulações necessárias variará de acordo com cada sistema, então é necessário fazer ajustes para cada projeto com base na experimentação.

## 7 Referências

- [1] F.S.C., CovAnalysis, GIT, 2023. Disponível em:  
[https://gitlab.com/FSCPhysis/Covariance\\_Analysis](https://gitlab.com/FSCPhysis/Covariance_Analysis). Acesso em: 23 dez. 2023.
- [2] WOLZ, Laura et al, Pseudo-Power, GIT, 2023. Disponível em: <https://github.com/Astrophysics-UCL/PseudoPower>. Acesso em: 22 jun. 2023.
- [3] H. S. Xavier et all, FLASK, GIT, 2023. Disponível em: <https://github.com/ucl-cosmoparticles/flask>. Acesso em: 28 dez. 2023.
- [4] Python Software Foundation, Python Language Reference, version 3.9, 2001. disponível em <https://www.python.org/>.
- [5] Zonca, A., Singer, L., Lenz, D., Reinecke, M., Rosset, C., Hivon, E., & Gorski, K., healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in Python. *Journal of Open Source Software*, 4(35), 1298, 2019. doi: doi.org/10.21105/joss.01298.
- [6] Górski, K. M., Hivon, E., Banday, A. J., Wandelt, 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*, 622(1), 759–7712005, 2005. <https://doi.org/10.1086/427976>
- [7] Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., Lim, P. L., Crawford, S. M., ... & Deil, C., The astropy project: Building an open-science project and status of the v2.0 core package. *The Astronomical Journal*, 156(3), 123, 2018. DOI: 10.3847/1538-3881/aabc4f
- [8] Harris, C. R. et all, Array programming with NumPy. *Nature*, 585(7825), 357-362, 2020. DOI: 10.1038/s41586-020-2649-2
- [9] The pandas development team. pandas-dev/pandas: Pandas (Version latest), Zenodo, 2020. doi: doi.org/10.5281/zenodo.3509134
- [10] H. S. Xavier et all, Improving lognormal models for cosmological fields, *Monthly Notices of the Royal Astronomical Society*, V. 459, 3693–3710, 2016. doi: dx.doi.org/10.1093/mnras/stw874.
- [11] Loureiro A. et all, Cosmological Measurements from Angular Power Spectra Analysis of BOSS DR12 Tomography , *Monthly Notices of the Royal Astronomical Society*, V. 485, 326-355, 2019. doi: doi.org/10.48550/arXiv.1809.07204.
- [12] F. A. Dahlen, Frederik J. Simons, Spectral estimation on a sphere in geophysics and cosmology, *Geophysical Journal International*, Volume 174, Issue 3, September 2008, Pages 774–807, doi: doi.org/10.1111/j.1365-246X.2008.03854.x
- [13] Van De Hulst, H. C., Origin of the Radio Waves from Space, *Nederlandsch Tijdschrift voor Natuurkunde*, Jrg. XI, No. 12, December 1945.
- [14] Griffiths, David J., Hyperfine splitting in the ground state of hydrogen, *American Journal of Physics*, 50(8), 698–, 1982. doi:10.1119/1.12733
- [15] Ewen, H.I., Purcell, E.M. Observation of a Line in the Galactic Radio Spectrum: Radiation from Galactic Hydrogen at 1,420 Mc./sec.. *Nature* 168, 356, 1951. doi:10.1038/168356a0

- [16] Abdalla, E. et all, The BINGO project I: Baryon acoustic oscillations from integrated neutral gas observations, *Astronomy and Astrophysics*, vol. 664, 2022. doi:10.1051/0004-6361/202140883.
- [17] Wuensche, Carlos A. et all, The BINGO project: II. Instrument description, *Astronomy and Astrophysics*, vol. 664, 2022. doi: 10.1051/0004-6361/202039962.
- [18] Abdalla, Filipe B. et all, The BINGO Project: III. Optical design and optimization of the focal plane, *Astronomy and Astrophysics*, vol. 664, 2022. doi: 10.1051/0004-6361/202141382.
- [19] Liccardo et all, The BINGO project: IV. Simulations for mission performance assessment and preliminary component separation steps, *Astronomy and Astrophysics*, vol. 664, 2022. doi: 10.1051/0004-6361/202140886.
- [20] Fornazier, Karin S. F. et all, The BINGO project: V. Further steps in component separation and bispectrum analysis, *Astronomy and Astrophysics*, vol. 664, 2022. doi: 10.1051/0004-6361/202141707.
- [21] Zhang, Jiajun et all, The BINGO project: VI. HI halo occupation distribution and mock building, *Astronomy and Astrophysics*, vol. 664, 2022. doi: 10.1051/0004-6361/202140887.
- [22] Costa, Andre A. et all, The BINGO project: VII. Cosmological forecasts from 21 cm intensity mapping, *Astronomy and Astrophysics*, vol. 664, 2022. doi: 10.1051/0004-6361/202140888.
- [23] Novaes et all, The BINGO project: VIII. Recovering the BAO signal in HI intensity mapping simulations, *Astronomy and Astrophysics*, vol. 666, 2022. doi: 10.1051/0004-6361/202243158.
- [24] dos Santos, Marcelo V. et all, The BINGO Project: IX. Search for fast radio bursts – A forecast for the BINGO interferometry system, *Astronomy and Astrophysics*, vol. 681, 2024. doi: 10.1051/0004-6361/202346924.
- [25] Elcio Abdalla et all, BINGO-ABDUS: a radiotelescope to unveil the dark sector of the Universe, *arXiv*, vol. N<sub>Volume</sub>, 2023. doi : doi.org/10.48550/arXiv.2309.05099.
- [26] Alam et all. The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III, *The Astrophysical Journal Supplement Series*, vol. 219, no. 1, 2015. doi:10.1088/0067-0049/219/1/12.
- [27] Peebles, P. J. E., Statistical Analysis of Catalogs of Extragalactic Objects. I. Theory, *The Astrophysical Journal*, vol. 185, pp. 413–440, 1973. doi:10.1086/152431.
- [28] Desjacques, V., Jeong, D., and Schmidt, F., Large-scale galaxy bias, *Physics Reports*, vol. 733, pp. 1–193, 2018. doi:10.1016/j.physrep.2017.12.002.
- [29] Leistedt, B. et all, Estimating the large-scale angular power spectrum in the presence of systematics: a case study of Sloan Digital Sky Survey quasars, *Monthly Notices of the Royal Astronomical Society*, vol. 435, pp. 1857–1873, 2013. doi:doi.org/10.1093/mnras/stt1359
- [30] D. M. Bramich, W. Freudling, Systematic trends in Sloan Digital Sky Survey photometric data, *Monthly Notices of the Royal Astronomical Society*, vol. 424, pp. 1584–1599, 2012. doi:doi.org/10.1111/j.1365-2966.2012.21385.x

- [31] Seljak, U., Analytic model for galaxy and dark matter clustering, *Monthly Notices of the Royal Astronomical Society*, vol. 318, no. 1, OUP, pp. 203–213, 2000. doi:10.1046/j.1365-8711.2000.03715.x.
- [32] Peacock, J. A. and Smith, R. E., Halo occupation numbers and galaxy bias, *Monthly Notices of the Royal Astronomical Society*, vol. 318, no. 4, OUP, pp. 1144–1156, 2000. doi:10.1046/j.1365-8711.2000.03779.x.

## 8 Anexo I - Códigos

### 8.1 dados\_mascara.py

```
1 import healpy as hp
2 import numpy as np
3
4 # Obter NSIDE:
5 path = "./data/mask/N512_footprint.fits"
6 mask = hp.read_map(path)
7 NSIDE = hp.get_nside(mask)
8 # output: NSIDE = 512
9
10 # Obter FSKY
11 condition = np.where(mask==1)
12 masked = mask[condition]
13 fsky = len(masked)/(12*NSIDE*NSIDE)
14 # output: fsky = 0.24798107147216797
```

Listing 4: Obtendo NSIDE e FSKY da máscara binária.

## 8.2 flask.config

```
1 ##########
2 # - This is a config file for the program FLASK.                      #
3 # - Program parameters are identified by the colon that follows it      #
4 #   (so do not use it for anything else!).                                #
5 # - Parameters must be followed by their values which can be a          #
6 #   string or a sequence of up to 3 int, double or char variables.        #
7 #   Such sequences cannot mix different types.                            #
8 # - In case of a string with spaces, nothing can appear after it.       #
9 # - Another input type is the full line as a string.                     #
10 # - Any other symbol or position of keywords and text are               #
11 #   completely arbitrary and chosen for aesthetic reasons.                #
12 #
13 # - All parameters used by the program should be listed here since     #
14 #   there are no hard-wired default values.                            #
15 # - Addition/removal/modification of parameters in this file must     #
16 #   be followed by a modification of ParDef namespace in the             #
17 #   ParameterList.hpp file.                                              #
18 #####
19
20
21 ##### Simulation basics #####
22
23 DIST: LOGNORMAL # Field type, either LOGNORMAL, GAUSSIAN or HOMOGENEOUS
# (no structure).
24 RNDSEED: 137      # Seed for random number generator.
25 POISSON: 1         # 1-Poisson (2-Gaussian) sample galaxy fields; 0-Do not
# sample (use expected number).
26
27 ##### Cosmology #####
28
29 OMEGA_m: 0.3067 # Matter (DM+Baryon) density parameter.
30 OMEGA_L: 0.6933 # Dark Energy (DE) density parameter.
31 W_de: -1.0       # DE equation of state (constant).
32 # The above are only used if SELEC_TYPE=1or3, DENS2KAPPA=1 or
# CATALOG_COLS contains r.
33
34 ELLIP_SIGMA: 0   # Std.dev. of each component of sources ellipticities.
# Set to <0 for circular sources.
35 GALDENSITY: 30  # Comoving galaxy density in (h^-1 Mpc)^-3, only used if
# SELEC_TYPE = 1 or 3.
36
37
38 ##### Input data #####
39
40 FIELDS_INFO: data/SDSS_14Bins.dat      # Means, shifts and z range of
# fields. Shift column is only relevant for LOGNORMAL.
41
42 CHOL_IN_PREFIX: 0                      # Mixing matrices prefix (skip Cl trafo
# and Cholesky decomp., 0 to do it).
43 CL_PREFIX: ./data/output/maps/cls/Cl-  # Prefix of the angular power
# spectrums C_ij(l) to use as input.
44 ALLOW_MISS_CL: 1                       # 0 - Return error if required
# Cl is missing; 1 - Set missing Cls to zero.
45
46 SCALE_CLS: 1.0                         # Constant re-scaling factor for all Cls.
47 WINFUNC_SIGMA: -1                      # Std. Dev. (arcmin) of field Gaussian smoothing,
```

```

        applied to C(l)s, set <0 for none.
48 APPLY_PIXWIN: 0      # Apply (1) or not (0) Healpix pixel window function
        to C(l)s.
49 SUPPRESS_L: -1      # l scale for generic exponential suppression of C(l)
        , set <0 for none.
50 SUP_INDEX: -1      # Index for the exponential suppression, exp(-(l/
        SUPPRESS_L)^SUP_INDEX), <0 for none.

51
52
53 ##### Survey selection function #####
54
55 SELEC_SEPARABLE: 1  # Non-sep. sel. func. (0), or separable with common
        (1) or distinct (2) angular part?
56 SELEC_PREFIX: 0      # Fields selection functions FITS files prefix; one
        file if separable; 0 for full sky.
57 SELEC_Z_PREFIX: data/SDSS_14Bins_radial-selection-function-    # Prefix
        for radial selection functions f(z), one for each galaxy field.
58 SELEC_SCALE: 1      # Overall factor to be applied to selection function
        (1 for no effect).
59 SELEC_TYPE: 0      # 0 - d_gal/dz/arcmin2; 1 - fraction of gals; +2 -
        angular part for bookkeeping only.

60
61 STARMASK: 0      # Mask over bright stars (Healpix FITS file, write 0
        for none).

62
63
64 ##### Multipole calculations, etc. #####
65
66 EXTRAP_DIPOLE: 1  # If dipole not specified by input Cls; 0 - set
        dipoles to zero. 1 - extrapolate.
67 LRANGE: 0 511      # Minimum and maximum l for which covariance
        matrices will be generated.
68 CROP_CL: 1      # Transf. lognormal to Gaussian Cl using all available L
        (0) or up to LMAX above (1).
69 SHEAR_LMAX: 512    # Max. l for kappa->shear computations, should be <=
        NSIDE, only affects lognormal sims.
70 NSIDE: 512      # Healpix Nside (Npixels = 12*NSIDE^2).
71 USE_HEALPIX_WGTS: 1 # 0 - Use equal (1.0) weights in map2alm; 1 - Use
        Healpix weights in map2alm.

72
73
74 ##### Covariance matrix regularization #####
75
76 MINDIAG_FRAC: 1e-12  # Null diagonal elements are set to this*(smallest
        diagonal element in all ells).
77 BADCORR_FRAC: 0      # Fraction added to variances if aux. Cov. matrices
        lead to |correlation|>1.
78 REGULARIZE_METHOD: 1  # Correcting Cov. matrices for pos. def. 0-None;
        1-Min. chi-sq; 2-Sampler with steps.
79 NEW_EVAL: 1e-18      # Eigenvalue to replace negative ones (only for
        method 1).
80 REGULARIZE_STEP: 0.0001 # Step size for successive approximations to pos
        . def.ness (only for method 2).
81 REG_MAXSTEPS: 1000    # Maximum number of iterations (only for method 2)
        .
82 ADD_FRAC: 1e-10      # If Eigenvalues>0 & Cholesky fail, add this*(
        smallest diagonal element) to diagonal.

83
```

```

84
85 ##### Other precision parameters #####
86
87 ZSEARCH_TOL: 0.0001      # Precision for finding radial selection
   function maximum location.
88
89
90 ##### Output #####
91
92 EXIT_AT: 0              # Write name of last output to be created (program
   will stop there). 0 for full run.
93 FITS2TGA: 0             # 0 - Only FITS; 1 - FITS and TGA; 2 - Only TGA.
94 USE_UNSEEN: 1            # Fill masked regions of output Healpix maps with healpy
   's UNSEEN (1) or zero (0).
95 LRANGE_OUT: 1 1024      # Inclusive l interval for alm and Cls output.
   Irrelevant if there is no such output.
96 MMAX_OUT: -1            # Truncate m for alm output at this value. Set it to <0
   for m<=l.
97 ANGULAR_COORD: 2         # 0 - ThetaPhi in radians 1 - ThetaPhi in degrees 2 -
   RADEC in degrees.
98 DENS2KAPPA: 0           # Integrate (1) or not (0) density to obtain convergence
   .
99
100 FLIST_OUT:            0 ./field-list.dat    # File to write the order of
   fields in covariance matrices. Write 0 for no output.
101 SMOOTH_CL_PREFIX:     0 example/pixwinCl-    # Prefix for C(l)s smoothed
   by Gaussian window function with WINFUNC_SIGMA.
102 XIOUT_PREFIX:          0 example/Xi-        # Prefix of output files for the
   correlation functions. Write 0 for no output.
103 GXIOUT_PREFIX:         0 example/gXi-        # Prefix of the auxiliary
   gaussian correlation functions. Write 0 for no output.
104 GCLOUT_PREFIX:         0 example/gCl-        # Prefix of the auxiliary
   gaussian C(l)s. Write 0 for no output.
105 COVL_PREFIX:           0 example/covl-       # Prefix for auxiliary
   covariance matrices for each l.
106 REG_COVL_PREFIX:       0 example/regl-       # Output prefix for the
   regularized aux. cov. matrices.
107 REG_CL_PREFIX:          0 example/regCl-      # Prefix for lognormal Cls,
   regularized.
108 CHOLESKY_PREFIX:       0 example/chol-       # Prefix for the GSL triangular
   matrices derived from Cholesky decomposition.
109 AUXALM_OUT:            0 example/aux-alm.dat # File output for auxiliary
   alm's. Write 0 for no output.
110 RECOVAUXCLS_OUT:       0 example/recov-gCl.dat # Cls estimated from
   auxiliary alm's. Write 0 for no output.
111 AUXMAP_OUT:             0 example/aux-map.dat # Text file output for pixel
   values before exponentiating or adding mean.
112 DENS2KAPPA_STAT:        0                  # Statistics of integrated densities
   , 0 - Ignore; 1 - Print to screen; or <filename>.
113 MAP_OUT:                0 example/map.dat    # Text file output of pixel
   coordinates and field values, after exponentiation.
114 MAPFITS_PREFIX:          0 ./data/FLASKmaps/map- # Prefix for the map .fits
   files. Write 0 for no output.
115 RECOVALM_OUT:            0 example/recov-alm.dat # File output for recovered
   alm's from MAP_OUT. Write 0 for no output.
116 RECOVCLS_OUT:             0 example/recovCls.dat # File output for Cl's
   calculated from the recovered alm's. Write 0 for no output.
117 SHEAR_ALM_PREFIX:        0 example/shear-Elm-  # Prefix for shear (E-mode)

```

```

    alm's output, calculated from convergence.
118 SHEAR_FITS_PREFIX: 0 example/kappa-gamma- # Prefix for lensing map (
    kappa, gamma1, gamma2) FITS files.
119 SHEAR_MAP_OUT:      0 example/shearmap.dat # File output of shear as a
    TEXT table.
120 MAPWER_OUT:        0 ./data/FLASKmaps/poisson-map.dat   # Text file
    output of pixel coordinates and field values, after mask & Poisson
    noise.
121 MAPWERFITS_PREFIX: 0 /data/FLASKmaps/map-   # Prefix for the map .fits
    files, same data as above. Write 0 for no output.
122 ELLIP_MAP_OUT:     0 example/ellip-map.dat # Text file output for
    galaxies ellipticities (reduced shear + shape noise).
123 ELLIPFITS_PREFIX:  0 ./data/FLASKmaps/ellip-flaskmap-simulation-0   #
    Prefix for Healpix map .fits files of galaxies ellipticities. Write 0
    for no output.
124 CATALOG_OUT:       0 example/catalog.dat # Sampled galaxy catalog w/
    errors (possible extensions are .dat, .fits, .fits.gz).
125
126
127 # Catalog columns, options are > (theta or dec) (phi or ra) z r galtype
    kappa gamma1 gamma2 ellip1 ellip2 pixel maskbit
128 CATALOG_COLS: ra dec z kappa gamma1 gamma2 ellip1 ellip2

```

Listing 5: Configurações do FLASK para este projeto.

### 8.3 SDSS\_14Bins.dat

```
1 # Field number, z bin number, mean, shift, field type, zmin, zmax
2 # Types: 1-galaxies 2-shear
3
4     1      1  0.0000  1.0000      1  0.00  0.05
5     1      2  0.0000  1.0000      1  0.05  0.075
6     1      3  0.0000  1.0000      1  0.075 0.100
7     1      4  0.0000  1.0000      1  0.100 0.125
8     1      5  0.0000  1.0000      1  0.125 0.150
9     1      6  0.0000  1.0000      1  0.150 0.175
10    1      7  0.0000  1.0000      1  0.175 0.200
11    1      8  0.0000  1.0000      1  0.200 0.225
12    1      9  0.0000  1.0000      1  0.225 0.250
13    1     10  0.0000  1.0000      1  0.250 0.275
14    1     11  0.0000  1.0000      1  0.275 0.300
15    1     12  0.0000  1.0000      1  0.300 0.325
16    1     13  0.0000  1.0000      1  0.325 0.350
17    1     14  0.0000  1.0000      1  0.350 0.400
```

Listing 6: Arquivo indicando os campos e os intervalos de redshift a serem computados.

## 8.4 SDSS\_14Bins\_radial-selection-function-f1.dat

```
1 0.000 0.0635000
2 0.006 0.0635000
3 0.011 0.0635000
4 0.017 0.0635000
5 0.022 0.0635000
6 0.028 0.0635000
7 0.033 0.0635000
8 0.039 0.0635000
9 0.044 0.0635000
10 0.050 0.1657500
11 0.050 0.2680000
12 0.053 0.2680000
13 0.056 0.2680000
14 0.058 0.2680000
15 0.061 0.2680000
16 0.064 0.2680000
17 0.067 0.2680000
18 0.069 0.2680000
19 0.072 0.2680000
20 0.075 0.3830000
21 0.075 0.4980000
22 0.078 0.4980000
23 0.081 0.4980000
24 0.083 0.4980000
25 0.086 0.4980000
26 0.089 0.4980000
27 0.092 0.4980000
28 0.094 0.4980000
29 0.097 0.4980000
30 0.100 0.5615000
31 0.100 0.6250000
32 0.103 0.6250000
33 0.106 0.6250000
34 0.108 0.6250000
35 0.111 0.6250000
36 0.114 0.6250000
37 0.117 0.6250000
38 0.119 0.6250000
39 0.122 0.6250000
40 0.125 0.6535000
41 0.125 0.6820000
42 0.128 0.6820000
43 0.131 0.6820000
44 0.133 0.6820000
45 0.136 0.6820000
46 0.139 0.6820000
47 0.142 0.6820000
48 0.144 0.6820000
49 0.147 0.6820000
50 0.150 0.6885000
51 0.150 0.6950000
52 0.153 0.6950000
53 0.156 0.6950000
54 0.158 0.6950000
55 0.161 0.6950000
56 0.164 0.6950000
```

57	0.167	0.6950000
58	0.169	0.6950000
59	0.172	0.6950000
60	0.175	0.6675000
61	0.175	0.6400000
62	0.178	0.6400000
63	0.181	0.6400000
64	0.183	0.6400000
65	0.186	0.6400000
66	0.189	0.6400000
67	0.192	0.6400000
68	0.194	0.6400000
69	0.197	0.6400000
70	0.200	0.6120000
71	0.200	0.5840000
72	0.203	0.5840000
73	0.206	0.5840000
74	0.208	0.5840000
75	0.211	0.5840000
76	0.214	0.5840000
77	0.217	0.5840000
78	0.219	0.5840000
79	0.222	0.5840000
80	0.225	0.5395000
81	0.225	0.4950000
82	0.228	0.4950000
83	0.231	0.4950000
84	0.233	0.4950000
85	0.236	0.4950000
86	0.239	0.4950000
87	0.242	0.4950000
88	0.244	0.4950000
89	0.247	0.4950000
90	0.250	0.4450500
91	0.250	0.3951000
92	0.253	0.3951000
93	0.256	0.3951000
94	0.258	0.3951000
95	0.261	0.3951000
96	0.264	0.3951000
97	0.267	0.3951000
98	0.269	0.3951000
99	0.272	0.3951000
100	0.275	0.3565500
101	0.275	0.3180000
102	0.278	0.3180000
103	0.281	0.3180000
104	0.283	0.3180000
105	0.286	0.3180000
106	0.289	0.3180000
107	0.292	0.3180000
108	0.294	0.3180000
109	0.297	0.3180000
110	0.300	0.2955000
111	0.300	0.2730000
112	0.303	0.2730000
113	0.306	0.2730000
114	0.308	0.2730000

```
115 0.311 0.2730000
116 0.314 0.2730000
117 0.317 0.2730000
118 0.319 0.2730000
119 0.322 0.2730000
120 0.325 0.2580000
121 0.325 0.2430000
122 0.328 0.2430000
123 0.331 0.2430000
124 0.333 0.2430000
125 0.336 0.2430000
126 0.339 0.2430000
127 0.342 0.2430000
128 0.344 0.2430000
129 0.347 0.2430000
130 0.350 0.1361000
131 0.350 0.0292000
132 0.356 0.0292000
133 0.361 0.0292000
134 0.367 0.0292000
135 0.372 0.0292000
136 0.378 0.0292000
137 0.383 0.0292000
138 0.389 0.0292000
139 0.394 0.0292000
140 0.400 0.0292000
```

Listing 7: Arquivo de dados do FLASK indicando os campos e os intervalos de redshift a serem computados.

## 8.5 esy.ini

```
1 [PARAMETERS]
2 # =====
3 # POWER SPECTRUM CALCULATIONS
4 # =====
5 NCORES = 7          # Number of cores for parallel processes >1.
6
7 #-----
8 # SKY-MAPS Cl's
9 #-----
10 NSIDE = 512         # Map resolution.
11 Lmax = 512          # 'L' max of spherical harmonics.
12 NBINS = 14           # Number of bins (maps).
13 FSKY = 0.2479810
14
15 CALCCROS = 0        # Make cros-correlation of bins map_i-map_j. [1]=True,
16 # [0]=False.
17 DEL_HEADER = 0       # To make the Cl be the input of Flask, we need to
18 # delete the fist line of PseudoPower output if it is the first time we
19 # generate the Cls: del_header = 2.
20
21 #-----
22 # FLASK SIMULATIONS Cl's
23 #-----
24 NFLASKMAPS = 0       # NFLASKMAPS: Number of flask simulations per bin.
25 START_IN = 0          # Number to start the names of the simulations:
26 # Initial nameclature.
27
28 #-----
29 # Flask-Systematics Cl's
30 #-----
31 FLASKMAP_PSEUDO_CL_CALC = 0      # Cl's of Flask maps. [1]=True, [0]=
32 # False.
33 FLASKMAP_PSEUDO_SL_CALC = 0      # Sl's of Flask maps. [1]=True, [0]=
34 # False.
35
36 #-----
37 # Standard variations
38 #-----
39 FLASKMAP_CL_STDV_CALC = 0        # Calc of standard variation of
40 # FlaskMap_i-FlaskMap_j Cl's. [1]=True, [0]=False. Requires NFLASKMAPS
41 # > 1.
42 FLASKMAP_SL_STDV_CALC = 0        # Calc of standard variation of
43 # FlaskMap_i-FlaskMap_i Sl's. [1]=True, [0]=False. Requires NFLASKMAPS
44 # > 1.
45
46 FLASKMAP_CROSS_SYS_STDV_CALC = 0 # Calc of standard variation of
47 # FlaskMaps-Systematics Cl's. [1]=True, [0]=False. Requires NFLASKMAPS
48 # > 1.
49
50 # =====
```

```

43 # COVARIANCE CALCULATION (not fully implemented)
44 # =====
45 BINING = 0 # The delta L to be considered will be 2.BINING+1. [>0]=True
        , [0]=False.
46
47 COVARIANCE THEORY = 0          # (Not implemented yet) Covariance of the
        real data Cl's. [1]=True, [0]=False.
48
49 # Get all folders with "Bined" in the name. If you don't want to
        calculate some folder, put '_' at the beginning of the name
50 COVARIANCE_CL_CALC = 0          # Covariance Matrix calculation of Cl's
        simulations. [1]=True, [0]=False.
51 COVARIANCE_SL_CALC = 0          # Covariance Matrix calculation of Sl's
        simulations. [1]=True, [0]=False.
52
53
54 # =====
55 # DIRECTORY SETTINGS
56 # =====
57 BASEDIR = /media/BINGODATA0/CovAnalysis/Covariance_Analysis-
        bug_densidade      # Current directory. Required to fill in this
        field.
58
59 FLASKDIR: /home/fcaldas/codes/flask           # FLASK code directory.
60 FLASKCFGNAME: /flask_felipe.config             # Name.config of the flask
        settings file inside FLASKDIR.
61 PSEUDOPOWERDIR: /home/fcaldas/codes/PseudoPower/PseudoPower-pmotta-
        version-funcional1    # PseudoPower code dir.
62
63
64
65
66
67
68 # Organization of the folders, you don't need to modify this.
69 #-----
70 # Real datas:
71 #-----
72 MAPSDIR: %(BASEDIR)s/data/maps                 # Sky maps folder. e
        .g: Sky-map_1.fits, sky-map_2.fits..
73 FOOTPRINTMASK: %(BASEDIR)s/data/mask/N512_footprint.fits      # Mask
        file. mask.fits.
74 SYSTEMATICSMAPS: %(BASEDIR)s/data/systematics     # Systematic maps
        folder. e.g: sdss_dr12_systematics_NAME1_16384_to_nside512_RING.fits
        ..
75 SYS_PREFIX: sdss_dr12_systematics_                # Prefix of
        systematics maps. e.g: Prefix_NAME1_sufix.fits
76 SYS_SUFFIX: _16384_to_nside512_RING.fits         # Sufix of
        systematics maps.
77 CL_PREFIX: %(BASEDIR)s/data/output/maps/cls/Cl-    # Prefix of Cl's of
        the real-maps. Required for FLASK to read all Cl's.
78 SL_PREFIX: %(BASEDIR)s/data/output/maps/sls/Sl-    # Prefix of Cl's of
        the real-maps. Required for FLASK to read all Cl's.
79
80 ## Simulation datas:
81 FLASKMAPSDIR: %(BASEDIR)s/data/FLASKmaps          #
        Specify the directory where the FLASK simulated maps will be
        generated (and deleted in the end of process). These files will be

```

```

    temporary.

82 MAPWERFITS_PREFIX: %(BASEDIR)s/data/FLASKmaps/FLASK_map-          # Prefix
    of Density Flask simulated maps. MAPWERFITS_PREFIX it's Flask config
    notation to density maps (which will be converted by us to
    overdensity map).

83
84 ## PseudoPower real-maps outputs:
85 FOOTPRINT_IlmJlm: %(BASEDIR)s/data/output/N512_footprint_IlmJlm_10_L512
    # IlmJlm.dat file.
86 MAPSAlmDIR: %(BASEDIR)s/data/output/maps/alm
    # Alm of real maps folder.
87 ClmapsDIR: %(BASEDIR)s/data/output/maps
    # Output folder for real-maps Cls.
88 SYS_DIR: %(BASEDIR)s/data/output/systematics
    # Systematics output dir.
89 SYS_AlmDIR: %(BASEDIR)s/data/output/systematics/alm
    # Output folder for alm of the systematics-maps.

90
91 ## PseudoPower FLASK-maps outputs:
92 FLASKMAPS_OUTPUT_DIR: %(BASEDIR)s/data/output/FLASKmaps           #
    Flaskmaps outputs dir.
93 FLASKMAPS_CL_DIR: %(BASEDIR)s/data/output/FLASKmaps/Cl            #
    Outputs of Cl's of the flask simulations.
94 FLASKMAPS_SL_DIR: %(BASEDIR)s/data/output/FLASKmaps/S1            #
    Outputs of S1's of the flask simulations.
95 FLASKMAPS_Alm_DIR: %(BASEDIR)s/data/output/FLASKmaps/alm          #
    Alm of simulated flask-maps.
96 FLASKMAPS_CROSS_SYS_DIR: %(BASEDIR)s/data/output/systematics/FLASK  #
    Output folder of cross correlation between flask maps and systematics
    + erro.dat

```

Listing 8: Parametros de configuração para o código CovAnalysis

## 8.6 UCLCL.ini

```
1 [cosmology]
2 engine = CLASS
3 h = 0.6766
4 ;100*theta_s = 1.041085
5 T_cmb = 2.7255
6 Omega_b = 0.04897
7 Omega_cdm = 0.2606
8 ;Omega_Lambda = 0
9 Omega_k = 0.0
10 Omega_fld = 0
11 Omega_scf = 0
12 w0_fld = -1.0
13 wa_fld = 0.
14 cs2_fld = 1
15 YHe = 0.245
16 tau_reio = 0.0561
17 k_pivot = 0.05
18 ln10^{10}A_s = 3.044
19 n_s = 0.9649
20 alpha_s = 0.
21 z_pk = 0.
22 nonlinear = none
23 lensing = no
24
25 [neutrino]
26 sum_m_ncdm = 0.060
27 hierarchy_ncdm = -1
28 N_ur = 2.03280
29 N_ncdm = 1
30 T_ncdm = 0.7137658555036082
31
32 [spectra]
33 output = nCl
34 P_k_max_h/Mpc = 1000.
35 l_min_lss = 2
36 l_max_lss = 511
37 l_switch_limber = 511
38 number count contributions = rsd
39
40 [experiment]
41 n_bins = 14
42
43 [accuracy]
44 ;The number of samples in Chi or k for the window function / integrand
45 ;Calculation time will of course scale linearly with nSteps in the case
        of spline integrations.
46 nSteps = 40000
47 KVALcut = 1000.
48 Nint = 5000
49 ;The minimum z for the window function integrand. A low min z --> high
        max k and therefore lower resolution in k space (since only k <~ 1
        matter.)
50 integrand_zmin = 1e-15
51 integrand_zmax = 0.99
52 ;Nint = 10000
53 ;gsl switch and accuracy parameters
```

```

54 gsl_int = false
55 gsl_limit = 8192
56 gsl_epsRel = 1e-4
57 gsl_epsAbs = 0.0
58 gsl_mDeg = 20
59
60 [bin_0]
61 form = file
62 file = PATH/bin0.dat
63 fsky = 0.24798107147216797
64 Ng_bar = 127965.8345203514
65 noise_nuisance = 2.4351984739876455e-05
66 probe = nClrsd
67 bias_form = const
68 bias_params = 0.7
69 lmin = 2
70 lmax = 512
71
72 [bin_1]
73 form = file
74 file = PATH/bin1.dat
75 fsky = 0.24798107147216797
76 Ng_bar = 253058.4309778168
77 noise_nuisance = 1.2314239195367185e-05
78 probe = nClrsd
79 bias_form = const
80 bias_params = 0.8
81 lmin = 2
82 lmax = 512
83
84 [bin_2]
85 form = file
86 file = PATH/bin2.dat
87 fsky = 0.24798107147216797
88 Ng_bar = 460207.1356194952
89 noise_nuisance = 6.77134665734884e-06
90 probe = nClrsd
91 bias_form = const
92 bias_params = 0.8
93 lmin = 2
94 lmax = 512
95
96 [bin_3]
97 form = file
98 file = PATH/bin3.dat
99 fsky = 0.24798107147216797
100 Ng_bar = 575762.3607539533
101 noise_nuisance = 5.412340684070596e-06
102 probe = nClrsd
103 bias_form = const
104 bias_params = 0.925
105 lmin = 2
106 lmax = 512
107
108 [bin_4]
109 form = file
110 file = PATH/bin4.dat
111 fsky = 0.24798107147216797

```

```

112 Ng_bar = 629783.8138515665
113 noise_nuisance = 4.948082152837949e-06
114 probe = nClrsd
115 bias_form = const
116 bias_params = 1.0
117 lmin = 2
118 lmax = 512
119
120 [bin_5]
121 form = file
122 file = PATH/bin5.dat
123 fsky = 0.24798107147216797
124 Ng_bar = 639377.8797019095
125 noise_nuisance = 4.873834626430925e-06
126 probe = nClrsd
127 bias_form = const
128 bias_params = 1.0
129 lmin = 2
130 lmax = 512
131
132 [bin_6]
133 form = file
134 file = PATH/bin6.dat
135 fsky = 0.24798107147216797
136 Ng_bar = 593565.1129589592
137 noise_nuisance = 5.250008771456583e-06
138 probe = nClrsd
139 bias_form = const
140 bias_params = 1.0
141 lmin = 2
142 lmax = 512
143
144 [bin_7]
145 form = file
146 file = PATH/bin7.dat
147 fsky = 0.24798107147216797
148 Ng_bar = 541703.6634712955
149 noise_nuisance = 5.752632406980722e-06
150 probe = nClrsd
151 bias_form = const
152 bias_params = 1.1
153 lmin = 2
154 lmax = 512
155
156 [bin_8]
157 form = file
158 file = PATH/bin8.dat
159 fsky = 0.24798107147216797
160 Ng_bar = 456173.5018747653
161 noise_nuisance = 6.831221096048361e-06
162 probe = nClrsd
163 bias_form = const
164 bias_params = 1.2
165 lmin = 2
166 lmax = 512
167
168 [bin_9]
169 form = file

```

```

170 file =PATH/bin9.dat
171 fsky = 0.24798107147216797
172 Ng_bar = 364873.6792312067
173 noise_nuisance = 8.540550406461412e-06
174 probe = nClrsd
175 bias_form = const
176 bias_params = 1.2
177 lmin = 2
178 lmax = 512
179
180 [bin_10]
181 form = file
182 file =PATH/bin10.dat
183 fsky = 0.24798107147216797
184 Ng_bar = 289499.0929090777
185 noise_nuisance = 1.0764185884492073e-05
186 probe = nClrsd
187 bias_form = const
188 bias_params = 1.1
189 lmin = 2
190 lmax = 512
191
192 [bin_11]
193 form = file
194 file =PATH/bin11.dat
195 fsky = 0.24798107147216797
196 Ng_bar = 245367.4018230989
197 noise_nuisance = 1.2700228417920968e-05
198 probe = nClrsd
199 bias_form = const
200 bias_params = 1.2
201 lmin = 2
202 lmax = 512
203
204 [bin_12]
205 form = file
206 file =PATH/bin12.dat
207 fsky = 0.24798107147216797
208 Ng_bar = 207064.8406644868
209 noise_nuisance = 1.5049498695505041e-05
210 probe = nClrsd
211 bias_form = const
212 bias_params = 1.25
213 lmin = 2
214 lmax = 512
215
216 [bin_13]
217 form = file
218 file =PATH/bin13.dat
219 fsky = 0.24798107147216797
220 Ng_bar = 49410.74272316132
221 noise_nuisance = 6.306770304839847e-05
222 probe = nClrsd
223 bias_form = const
224 bias_params = 1.55
225 lmin = 2
226 lmax = 512
227

```

```
228 [output]
229 root = PATH/output/nClrsd
230 write_output = true
231 lssCls = true
232 headers = true
233 format = class
234 stdout_verbosity_level = 5
```

Listing 9: Arquivo de configurações de input para o UCLCL.

## 9 Anexo II - Mapas cosmológicos

### 9.1 Escala normal

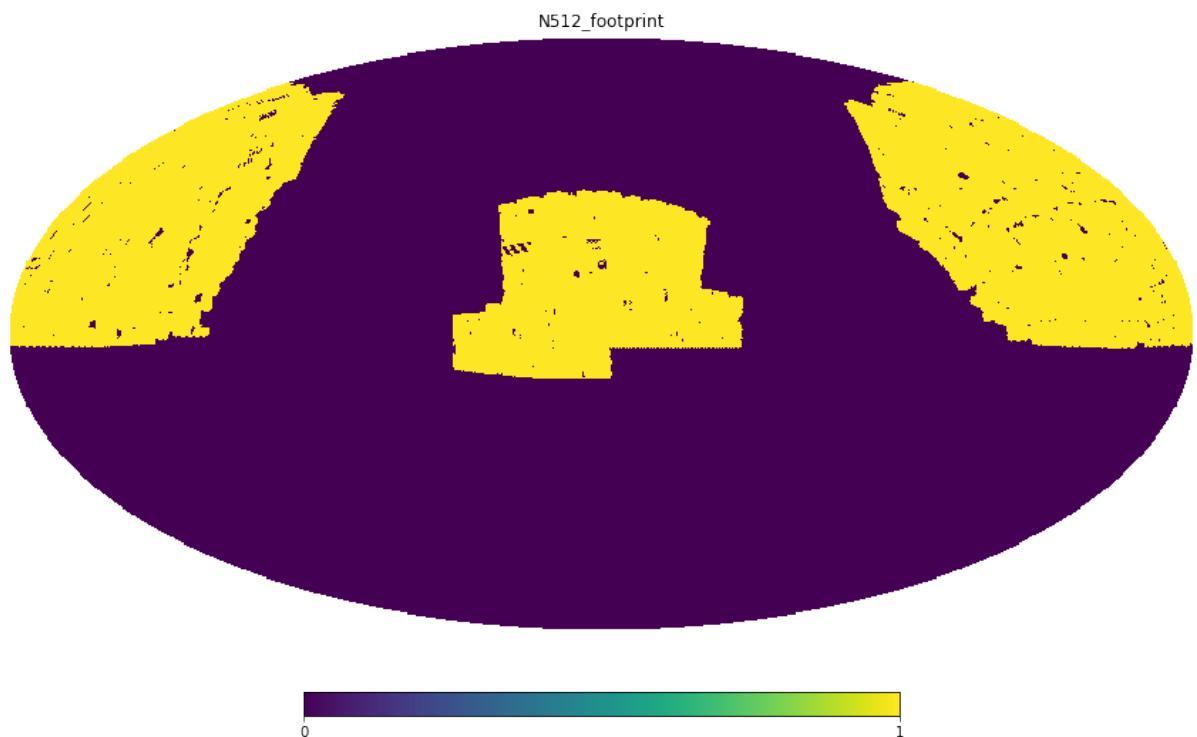


Figura 48: Máscara binária utilizada com NSIDE 512.

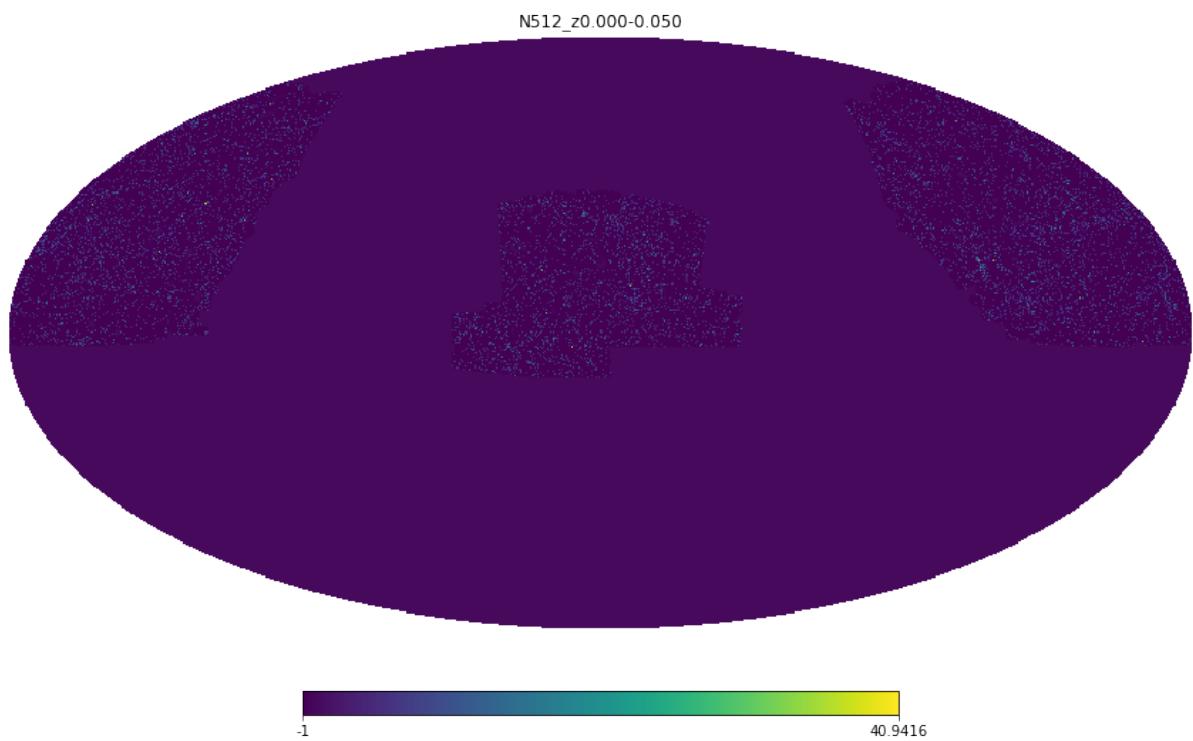


Figura 49: Mapa do intervalo de redshift 0.000-0.050 com NSIDE 512.

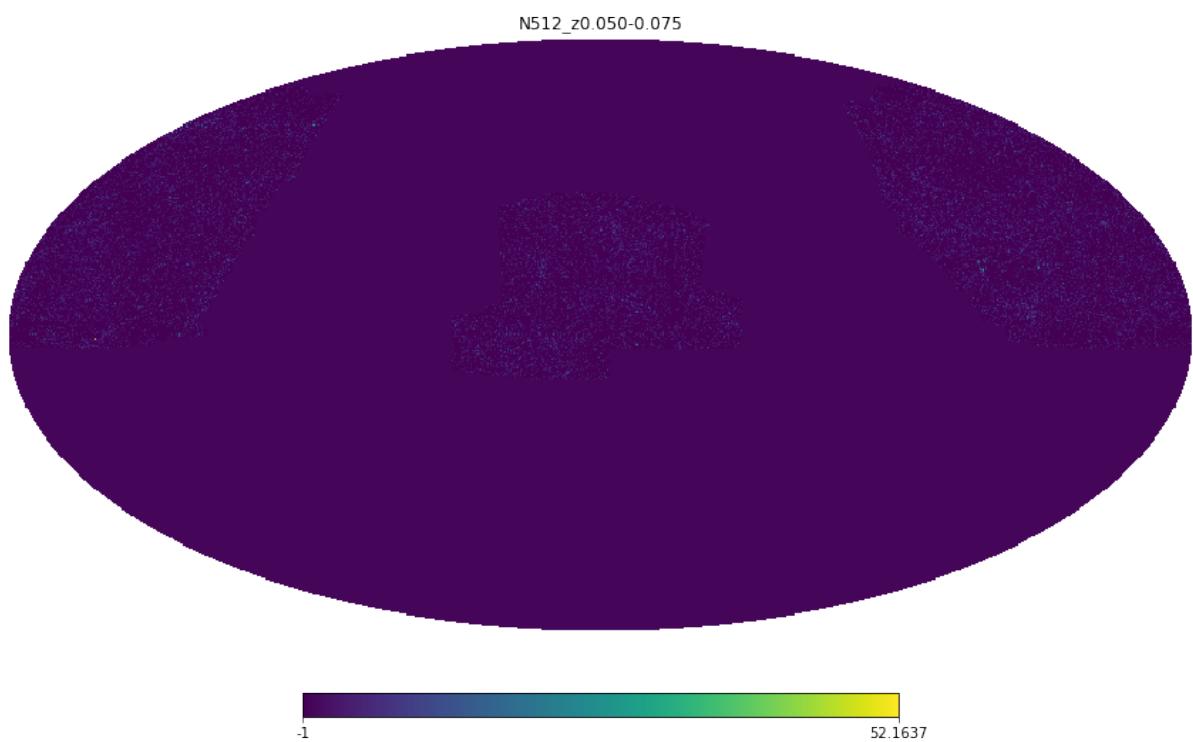


Figura 50: Mapa do intervalo de redshift 0.050-0.075 com NSIDE 512.

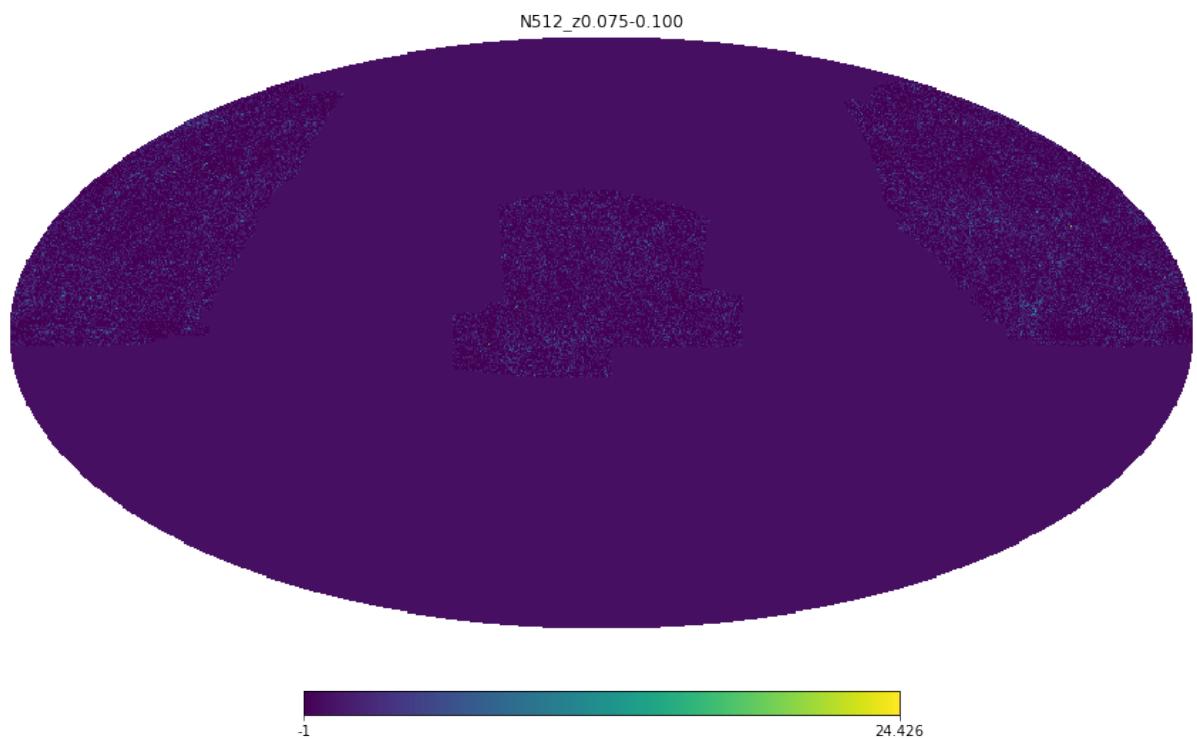


Figura 51: Mapa do intervalo de redshift 0.075-0.100 com NSIDE 512.

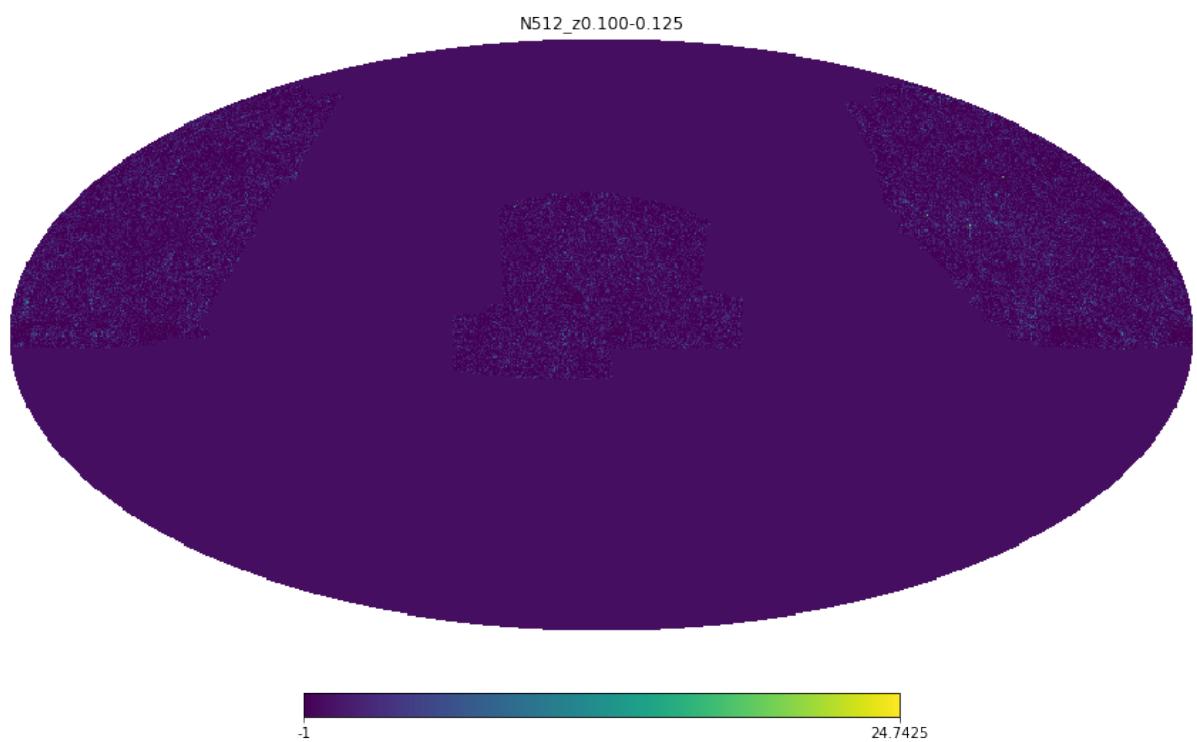


Figura 52: Mapa do intervalo de redshift 0.100-0.125 com NSIDE 512.

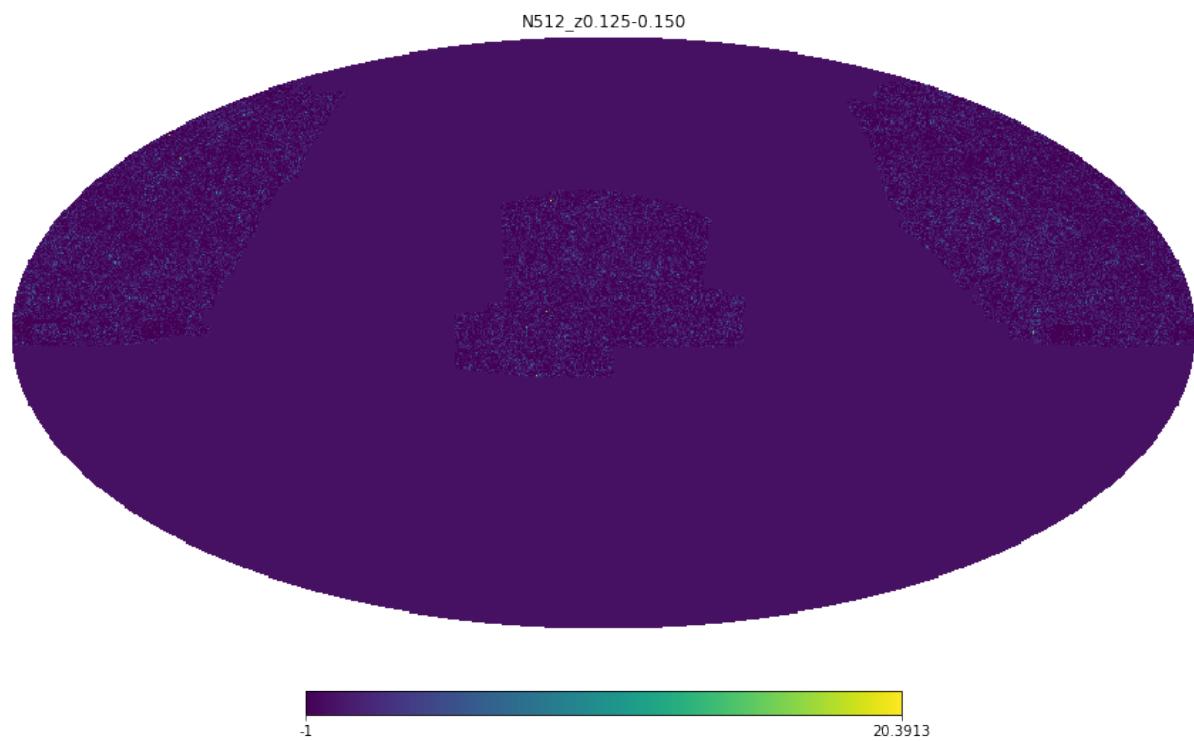


Figura 53: Mapa do intervalo de redshift 0.125-0.150 com NSIDE 512.

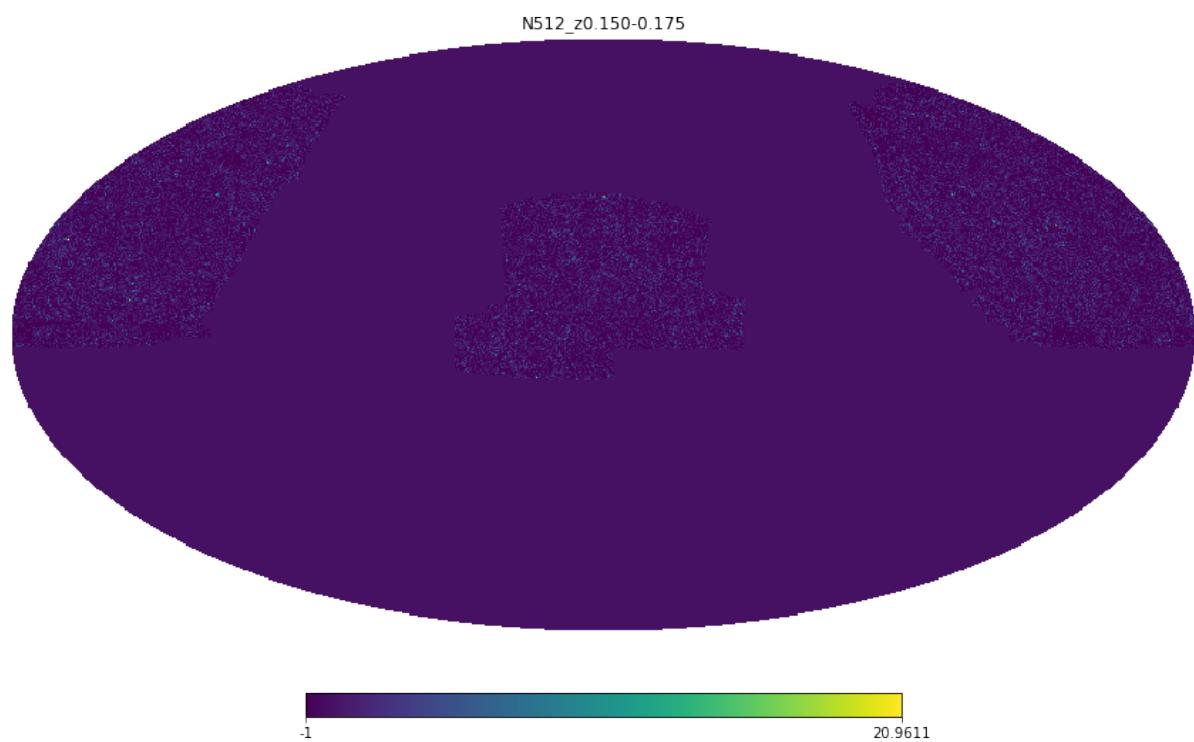


Figura 54: Mapa do intervalo de redshift 0.150-0.175 com NSIDE 512.

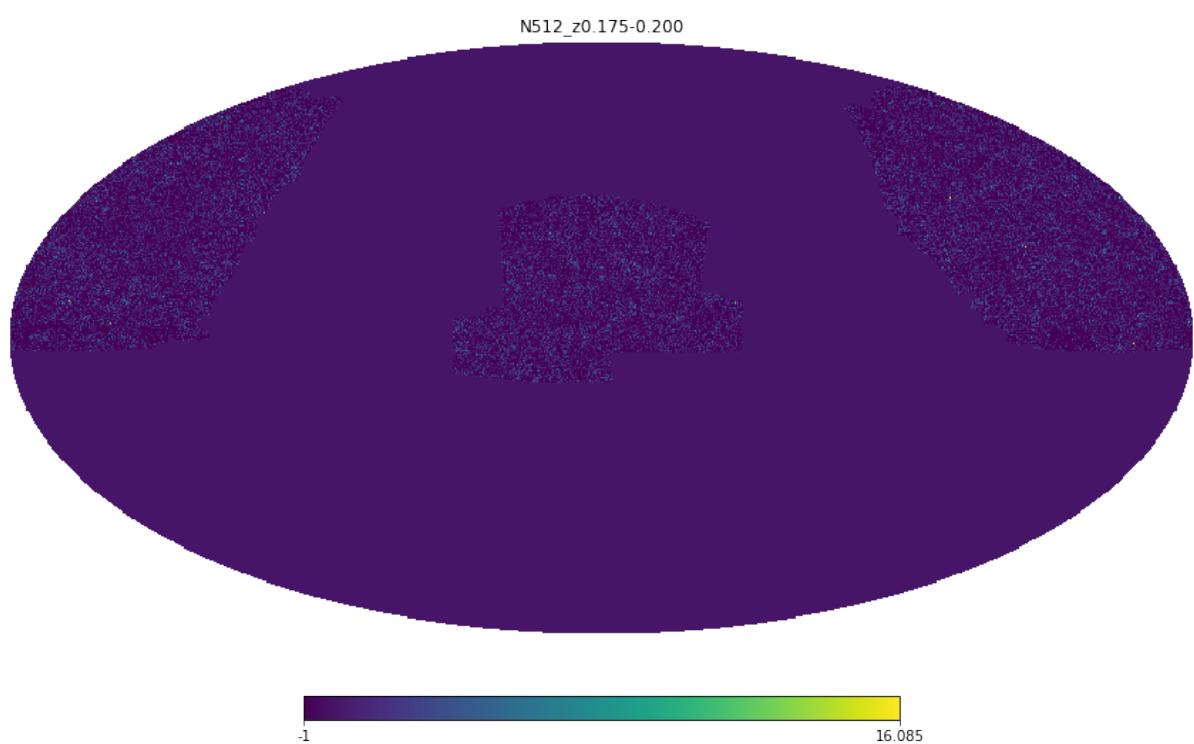


Figura 55: Mapa do intervalo de redshift 0.175-0.200 com NSIDE 512.

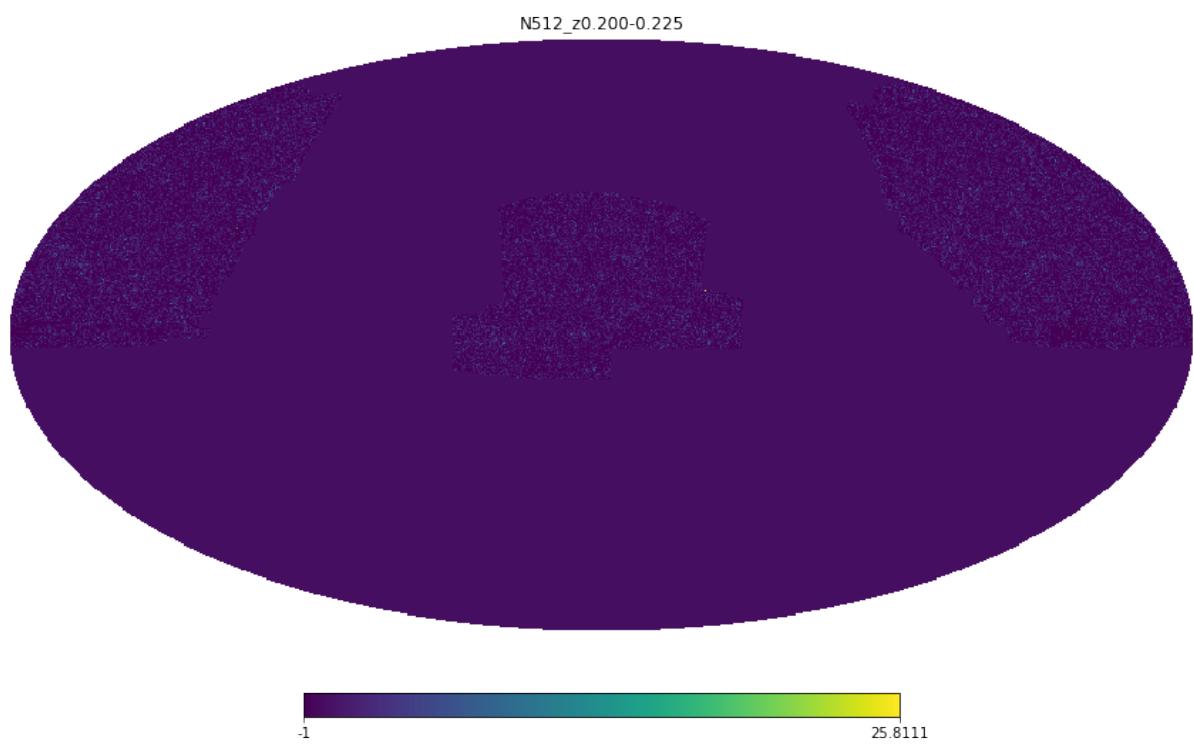


Figura 56: Mapa do intervalo de redshift 0.200-0.225 com NSIDE 512.

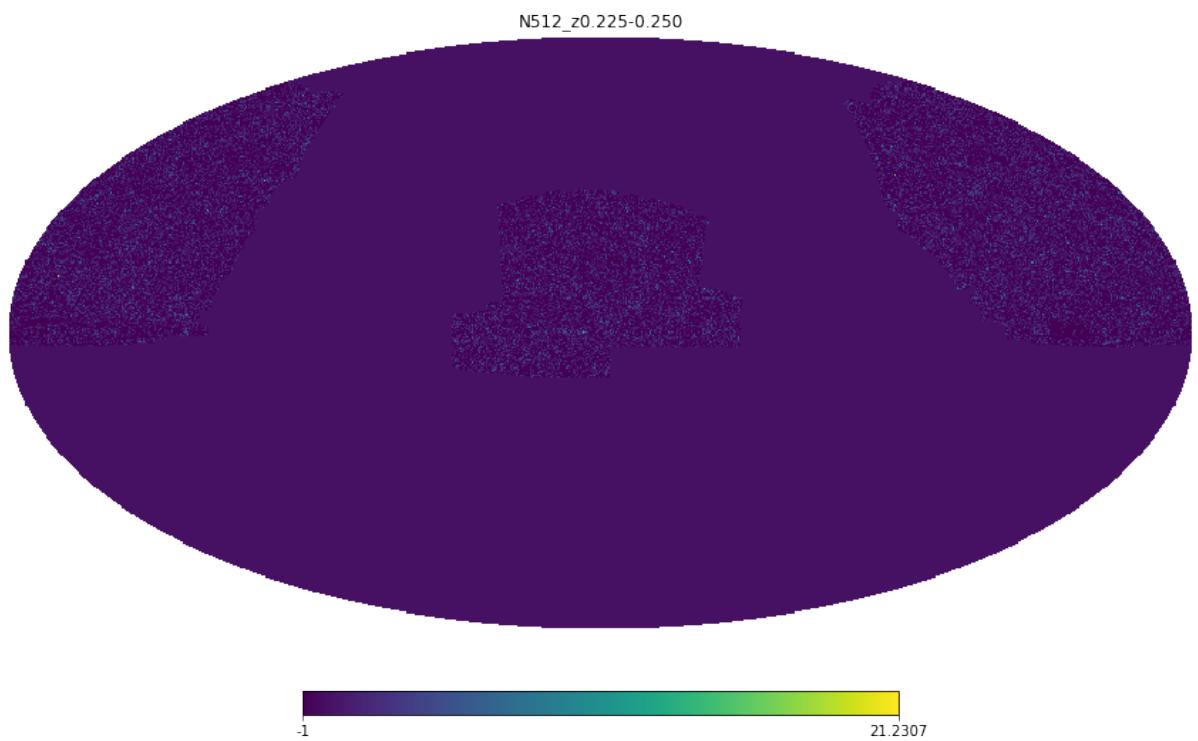


Figura 57: Mapa do intervalo de redshift 0.225-0.250 com NSIDE 512.

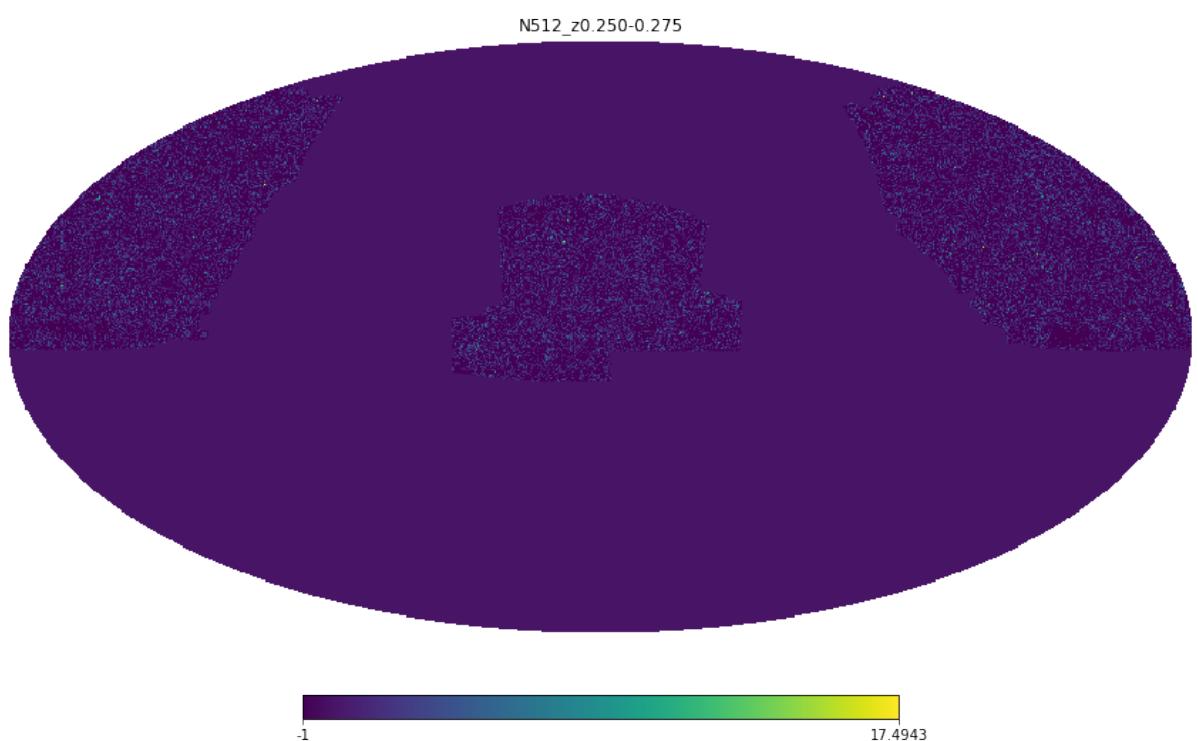


Figura 58: Mapa do intervalo de redshift 0.250-0.275 com NSIDE 512.

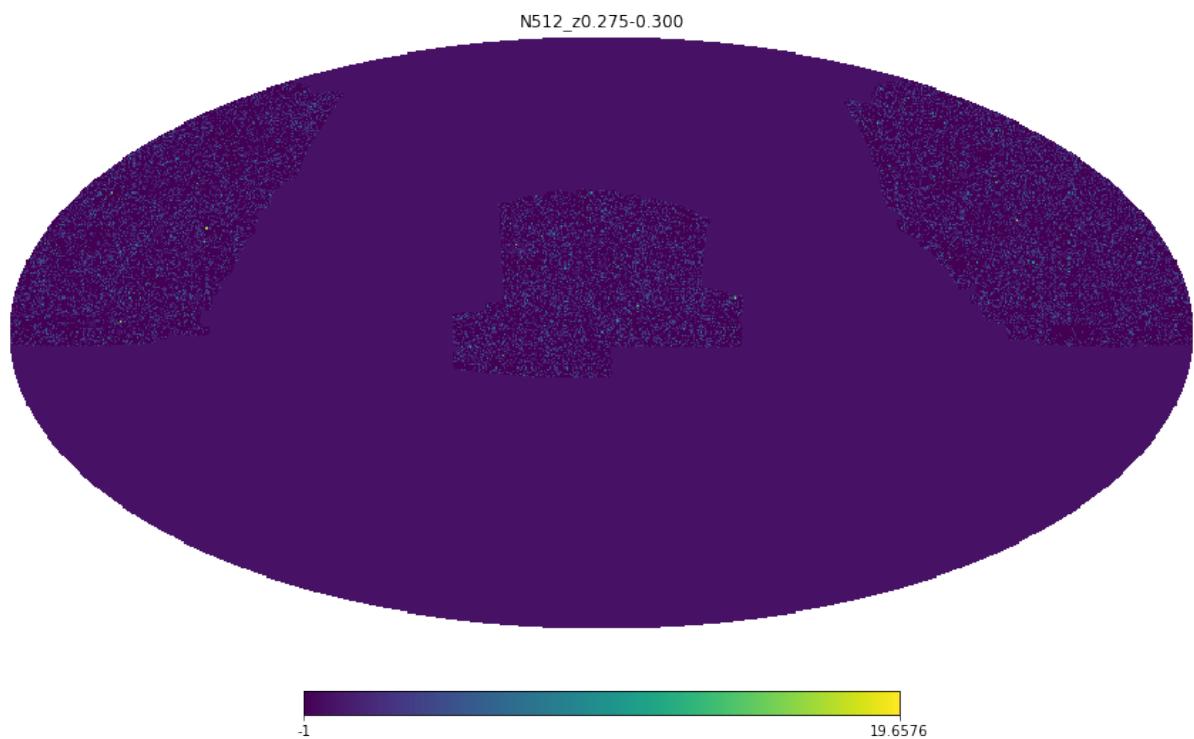


Figura 59: Mapa do intervalo de redshift 0.275-0.300 com NSIDE 512.

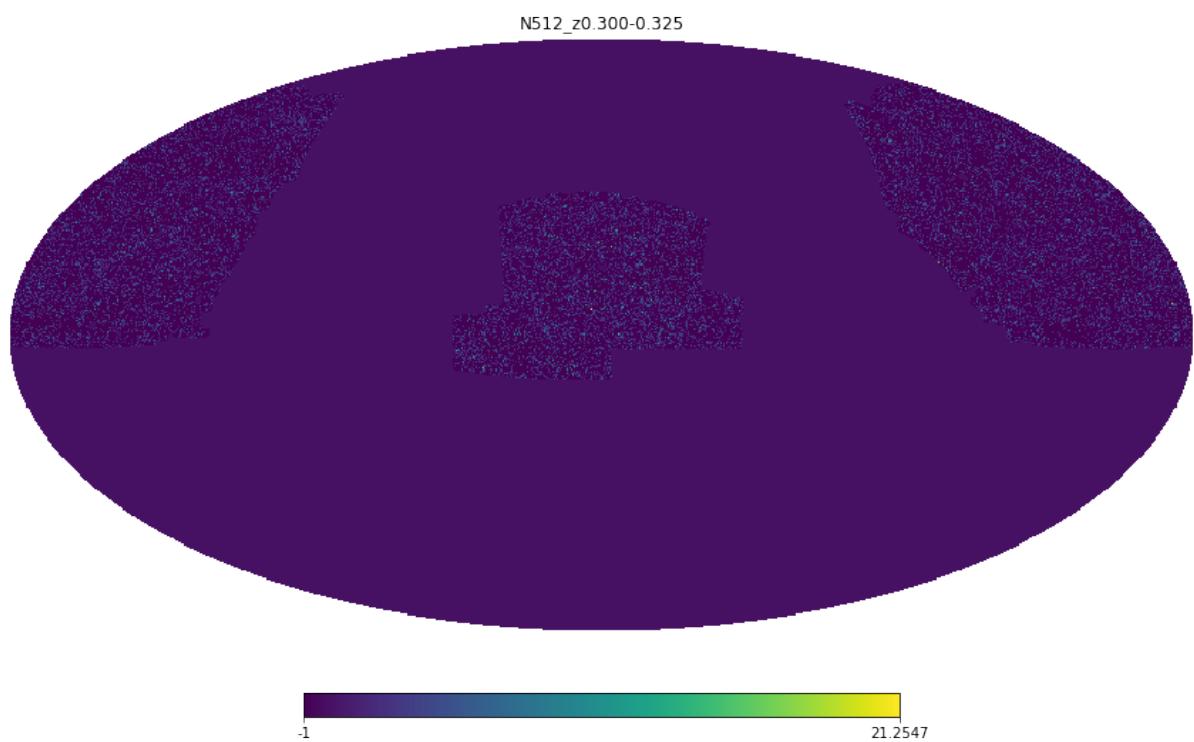


Figura 60: Mapa do intervalo de redshift 0.300-0.325 com NSIDE 512.

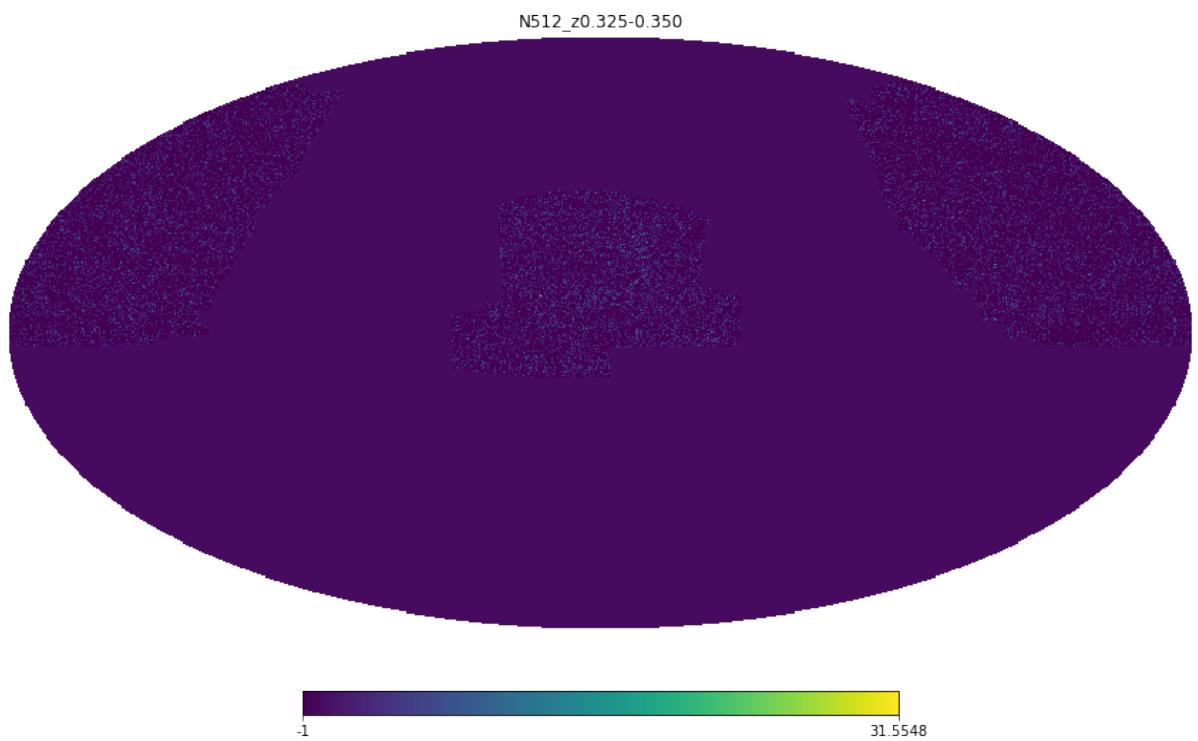


Figura 61: Mapa do intervalo de redshift 0.325-0.350 com NSIDE 512.

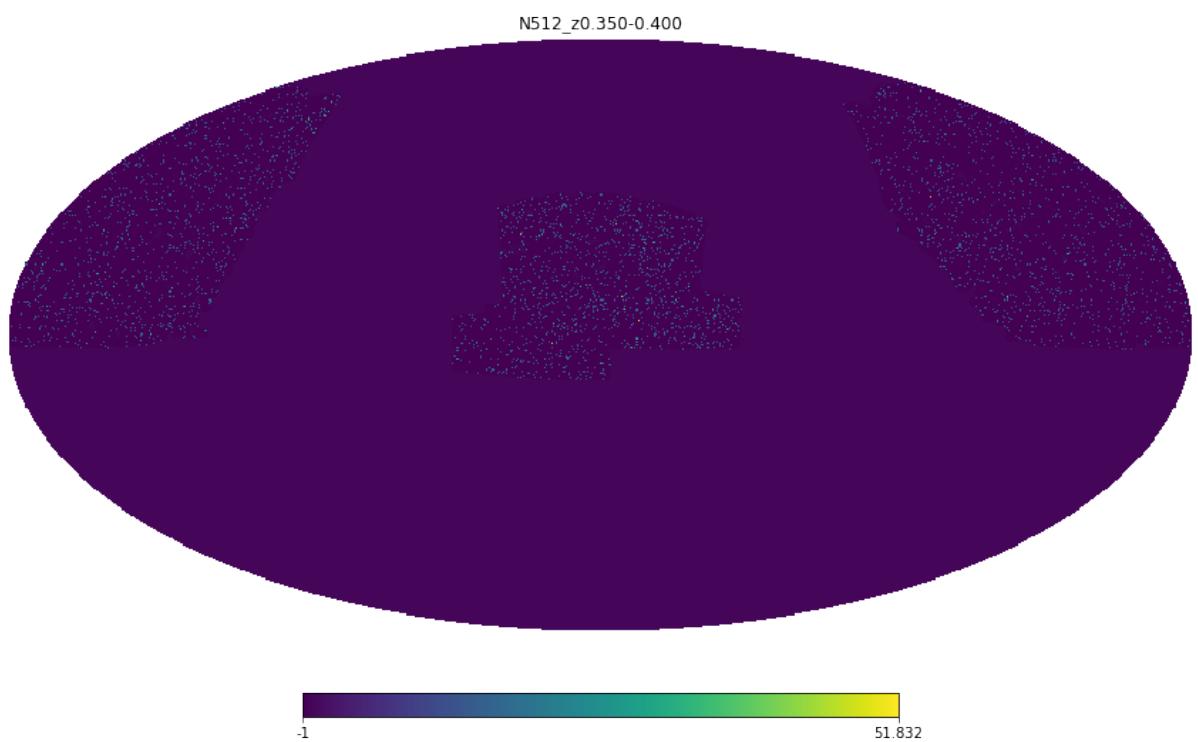


Figura 62: Mapa do intervalo de redshift 0.350-0.400 com NSIDE 512.

## 9.2 Escala logarítmica

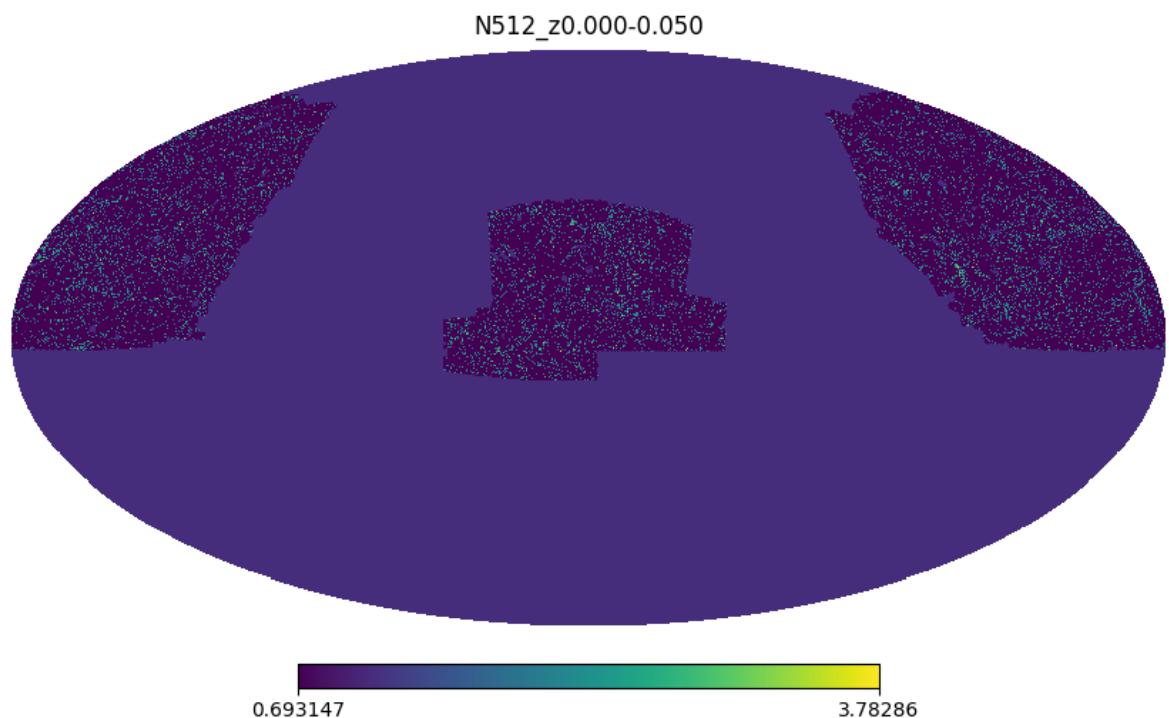


Figura 63: Mapa do intervalo de redshift 0.000-0.050 com NSIDE 512, na escala logarítmica.

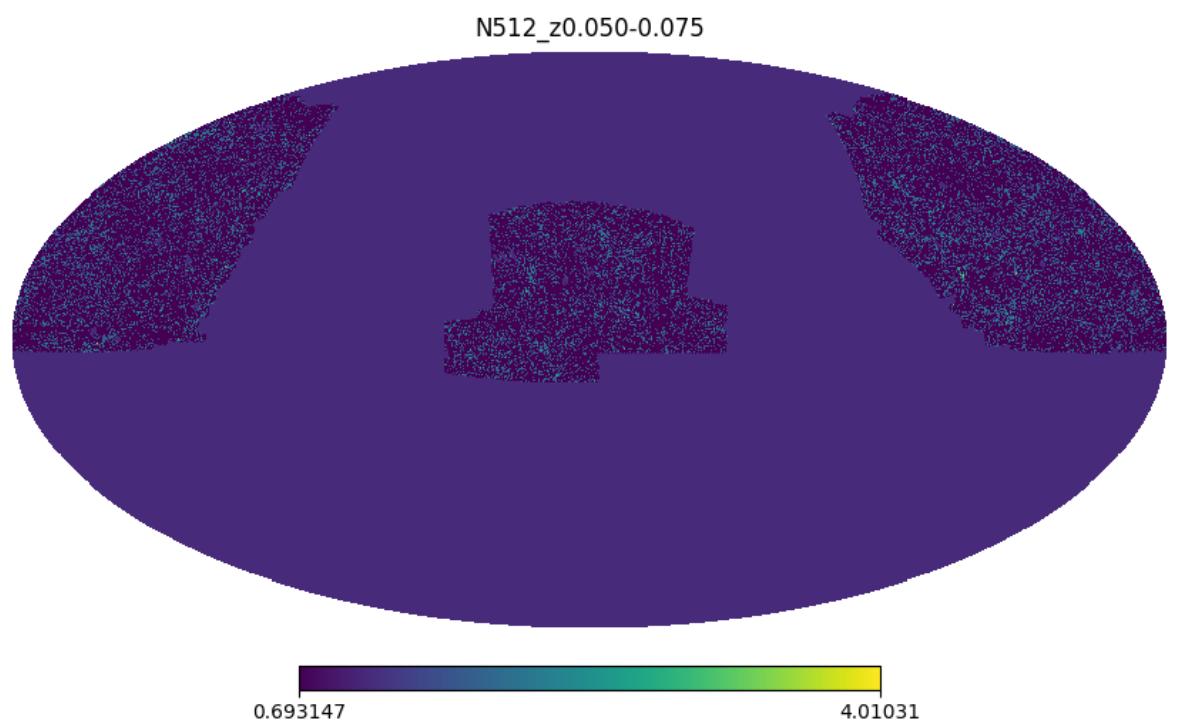


Figura 64: Mapa do intervalo de redshift 0.050-0.075 com NSIDE 512, na escala logarítmica.

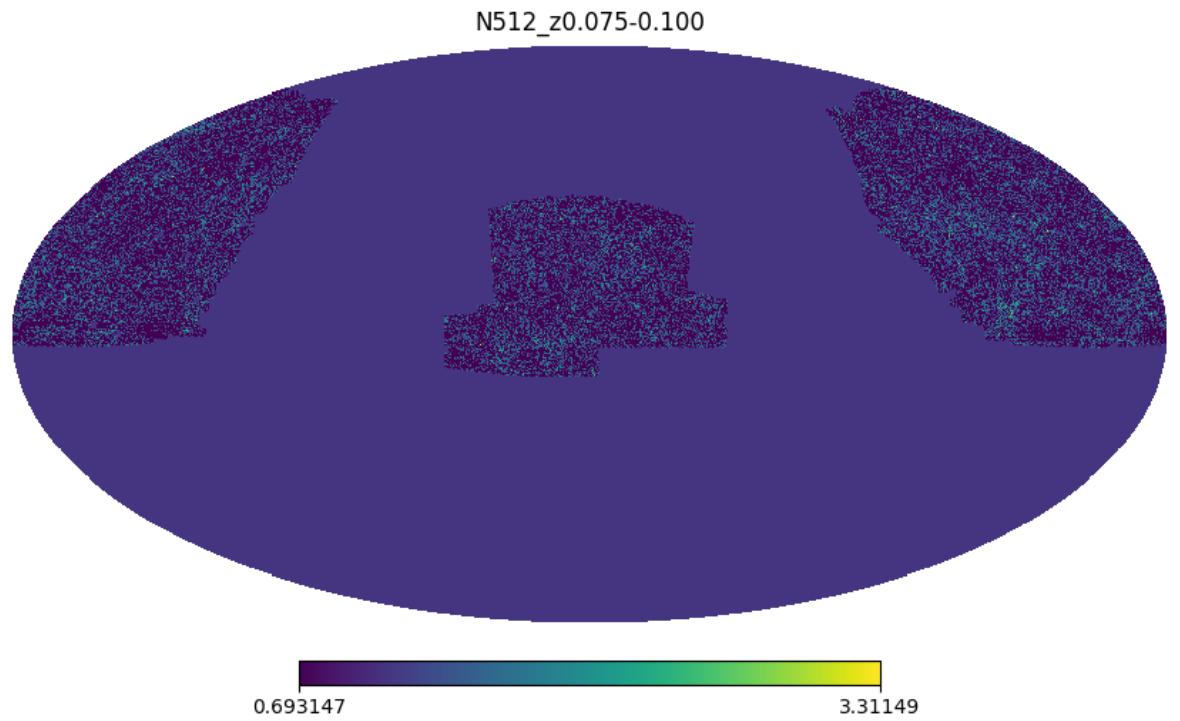


Figura 65: Mapa do intervalo de redshift 0.075-0.100 com NSIDE 512, na escala logarítmica.

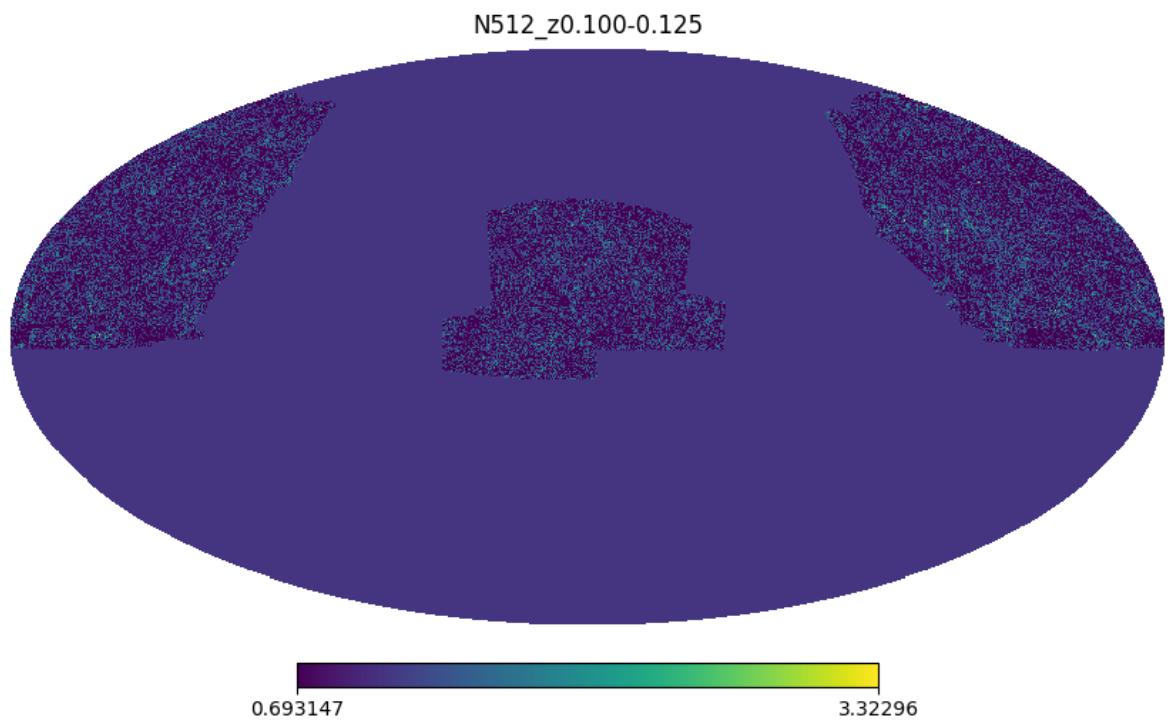


Figura 66: Mapa do intervalo de redshift 0.100-0.125 com NSIDE 512, na escala logarítmica.

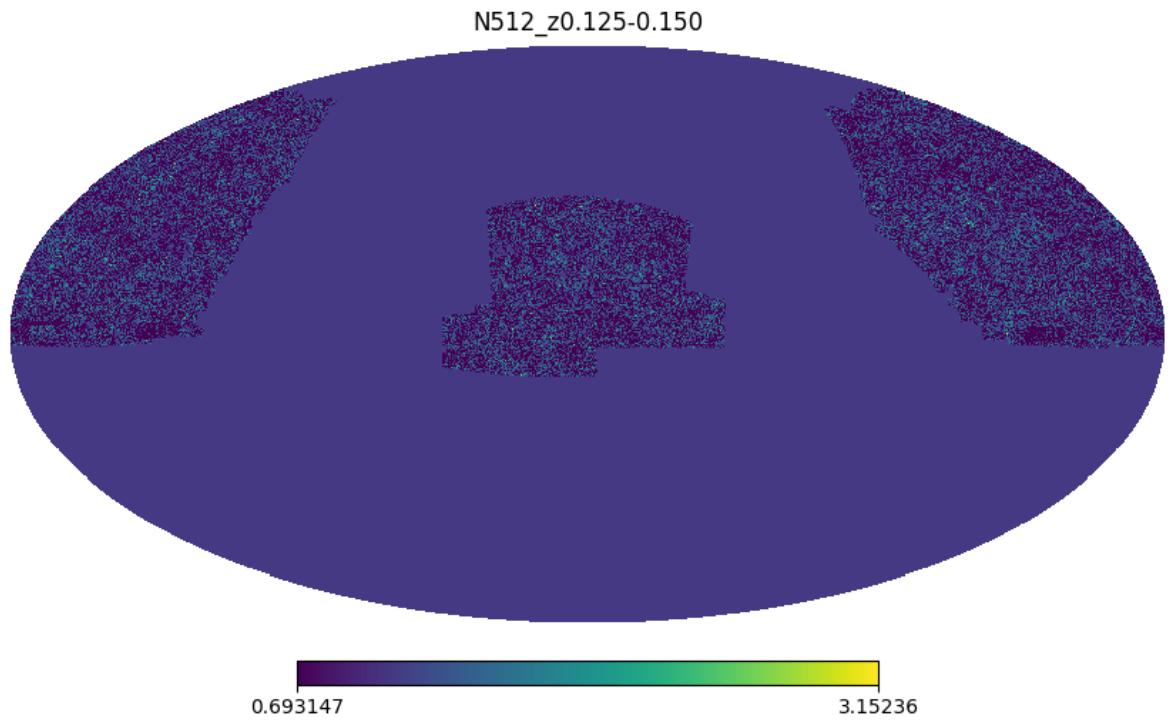


Figura 67: Mapa do intervalo de redshift 0.125-0.150 com NSIDE 512, na escala logarítmica.

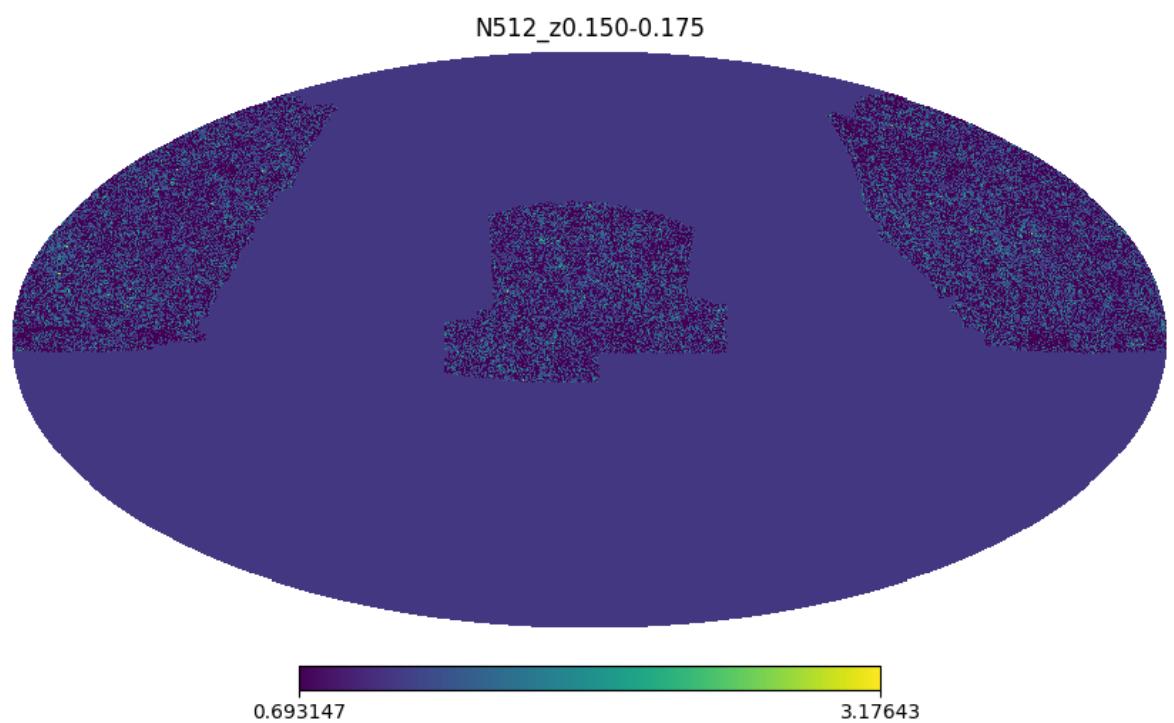


Figura 68: Mapa do intervalo de redshift 0.150-0.175 com NSIDE 512, na escala logarítmica.

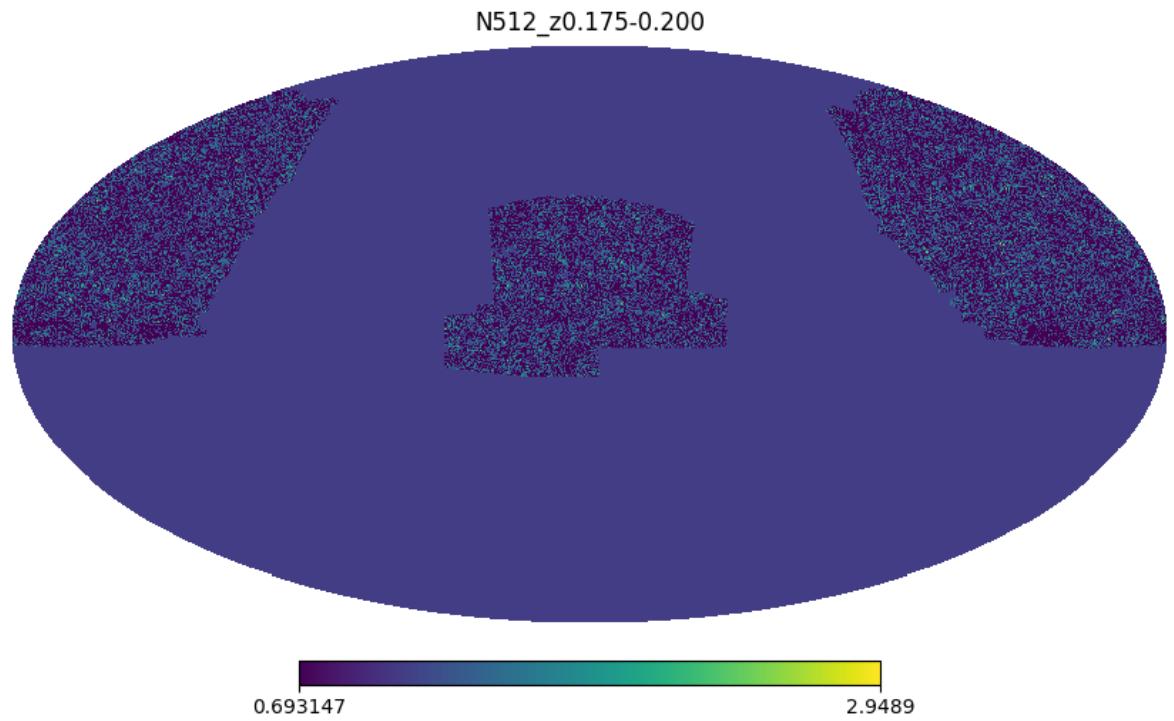


Figura 69: Mapa do intervalo de redshift 0.175-0.200 com NSIDE 512, na escala logarítmica.

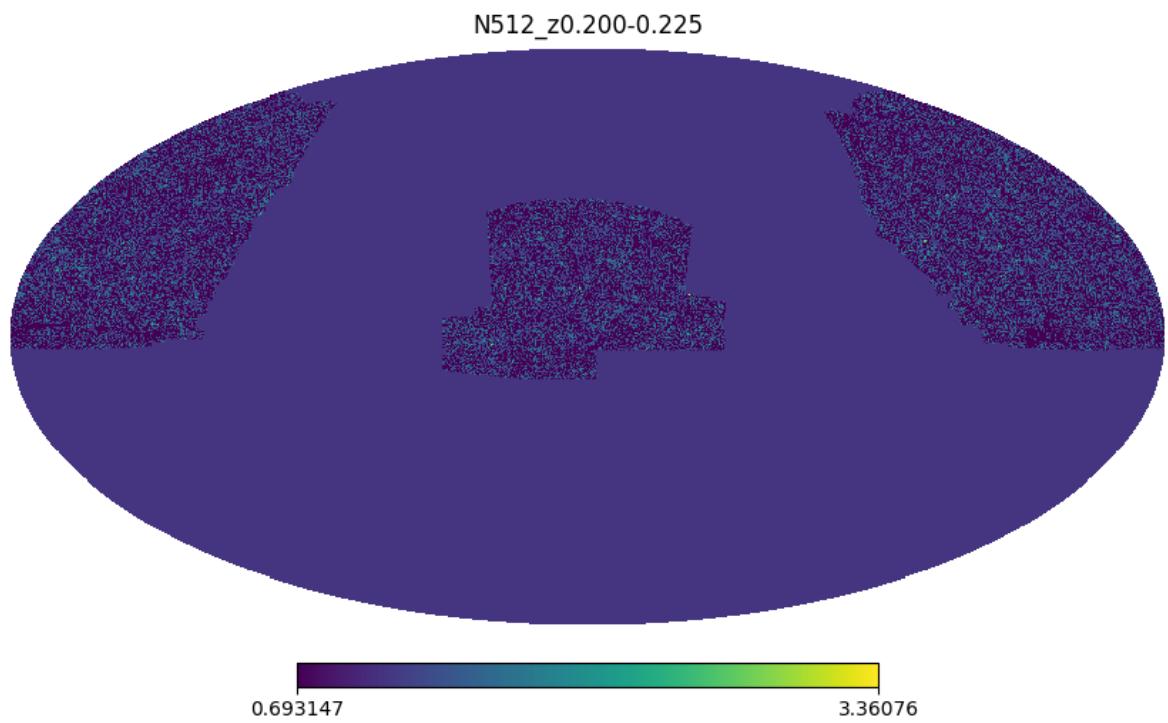


Figura 70: Mapa do intervalo de redshift 0.200-0.225 com NSIDE 512, na escala logarítmica.

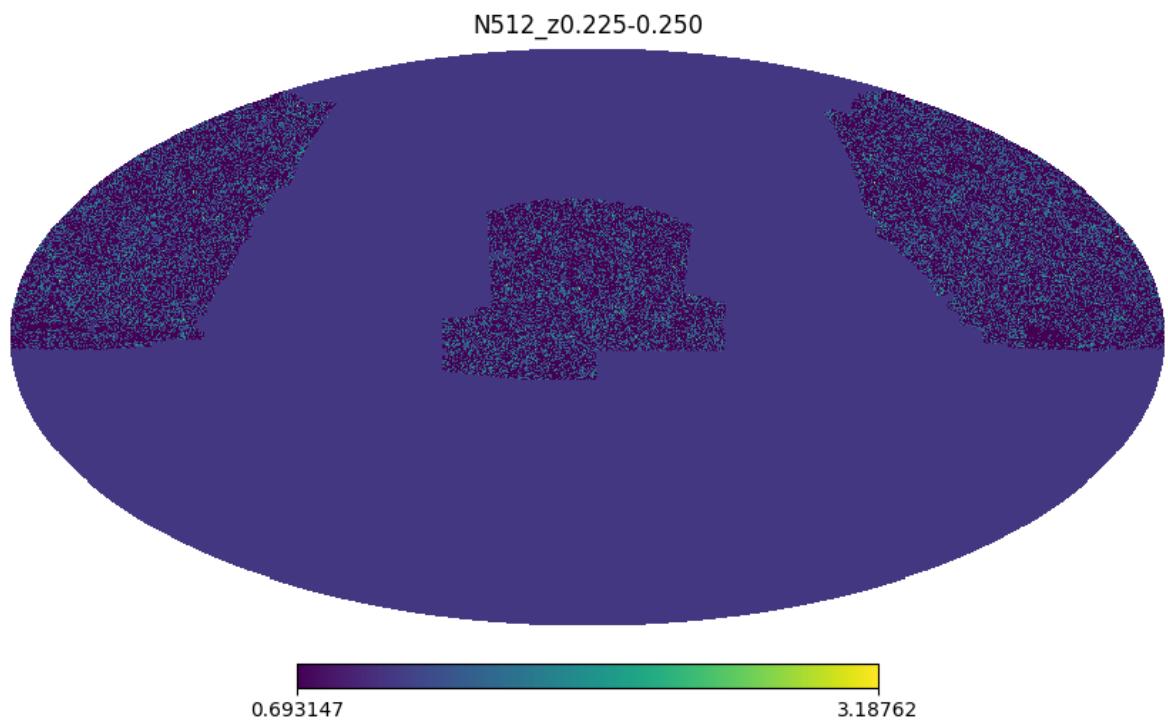


Figura 71: Mapa do intervalo de redshift 0.225-0.250 com NSIDE 512, na escala logarítmica.

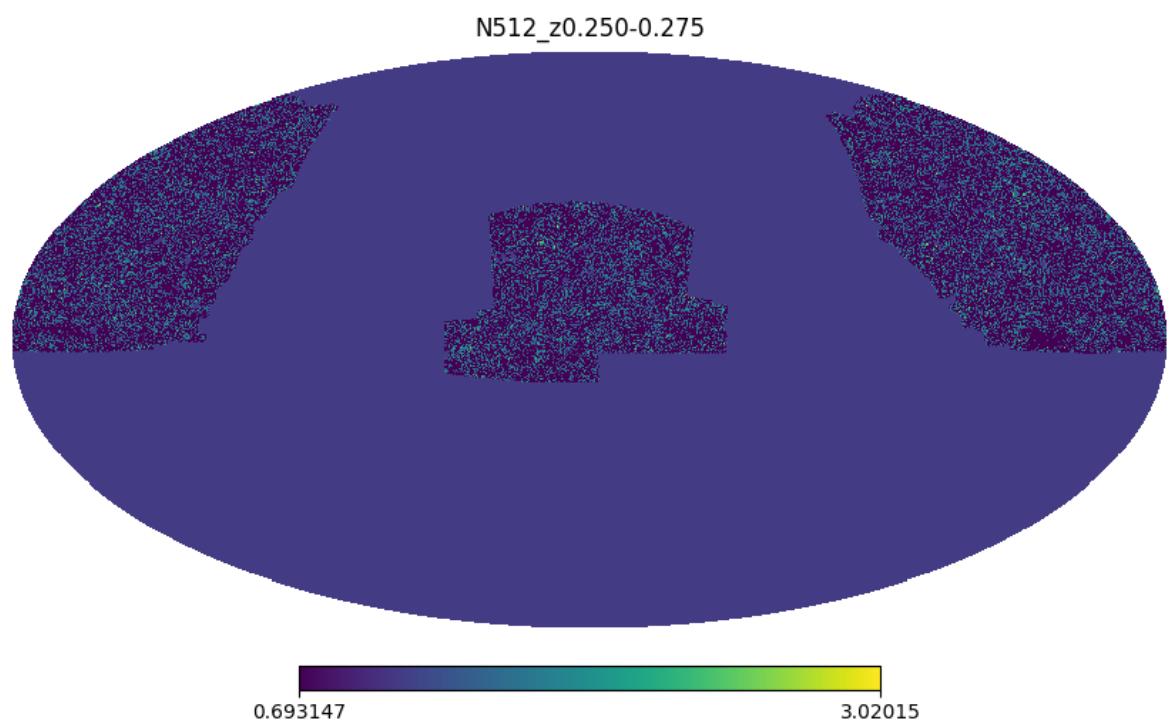


Figura 72: Mapa do intervalo de redshift 0.250-0.275 com NSIDE 512, na escala logarítmica.

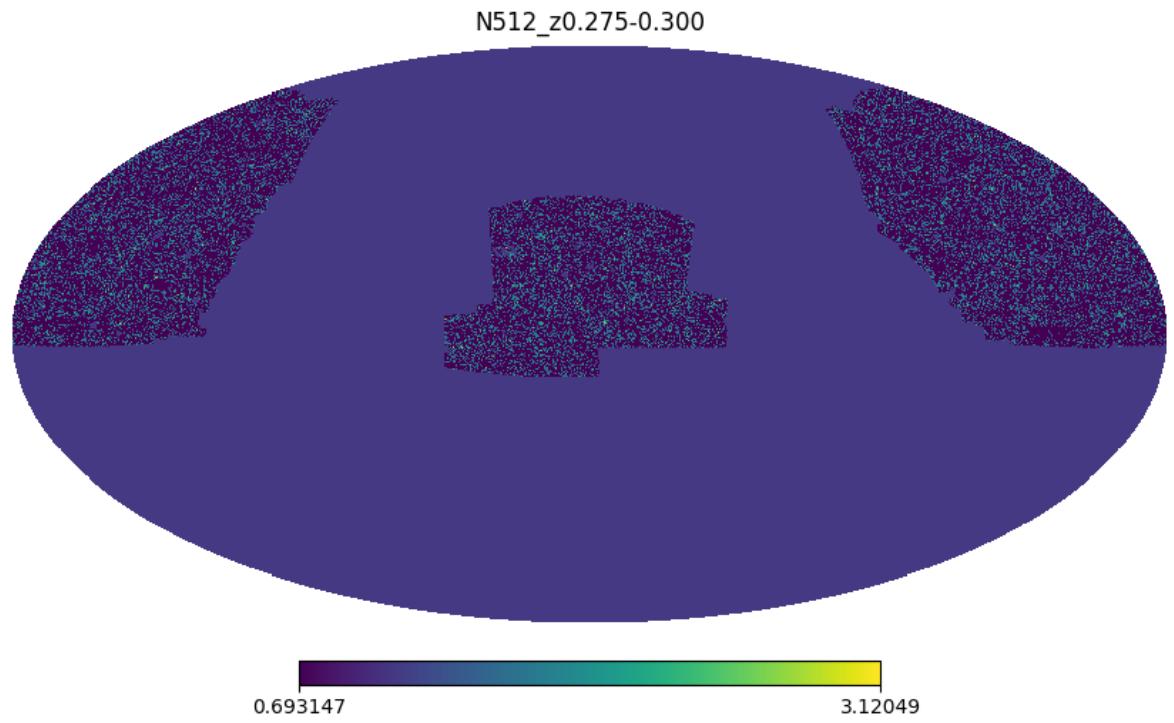


Figura 73: Mapa do intervalo de redshift 0.275-0.300 com NSIDE 512, na escala logarítmica.

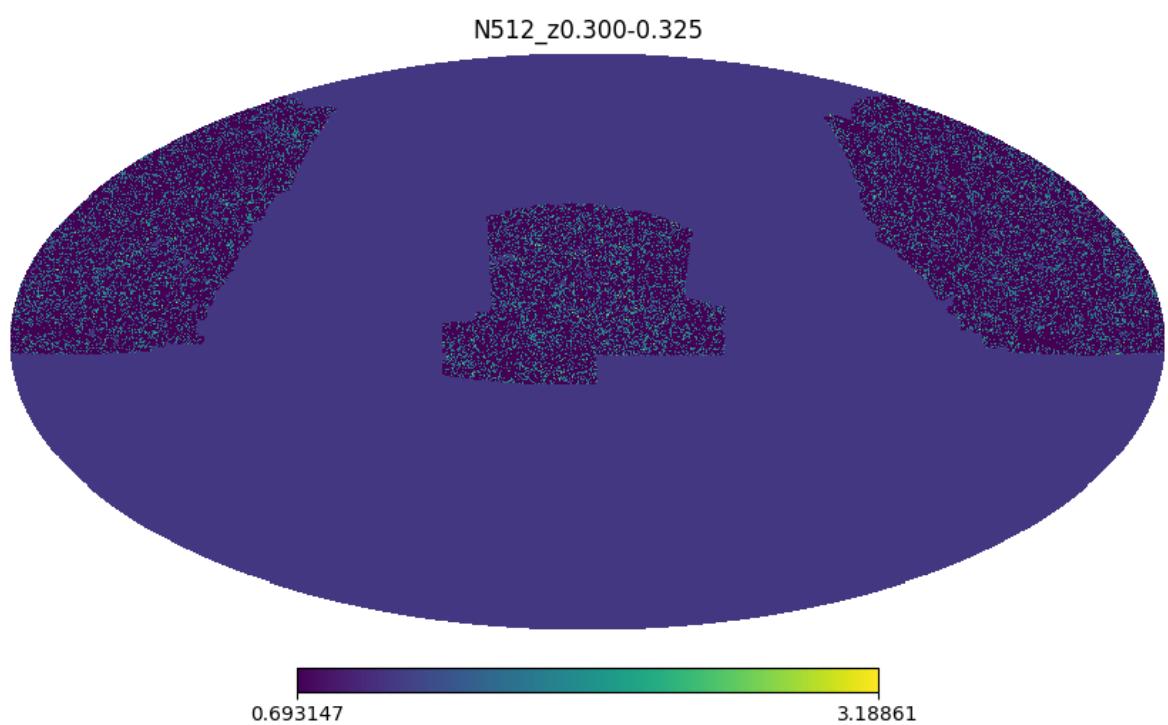


Figura 74: Mapa do intervalo de redshift 0.300-0.325 com NSIDE 512, na escala logarítmica.

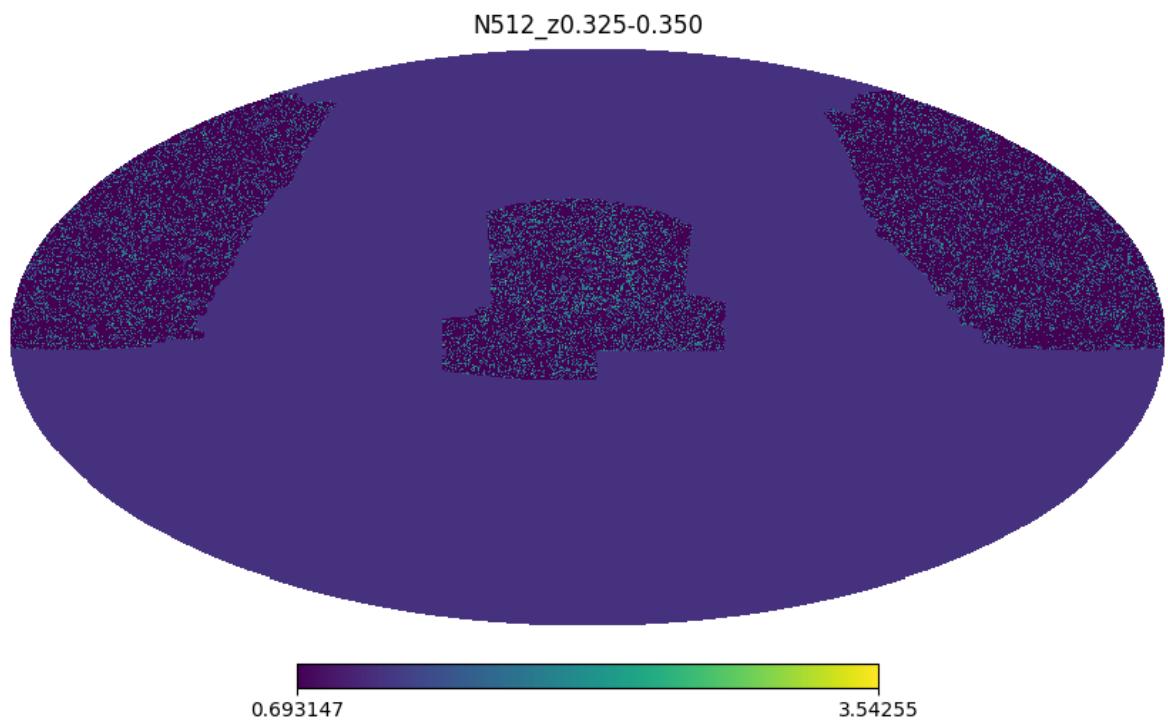


Figura 75: Mapa do intervalo de redshift 0.325-0.350 com NSIDE 512, na escala logarítmica.

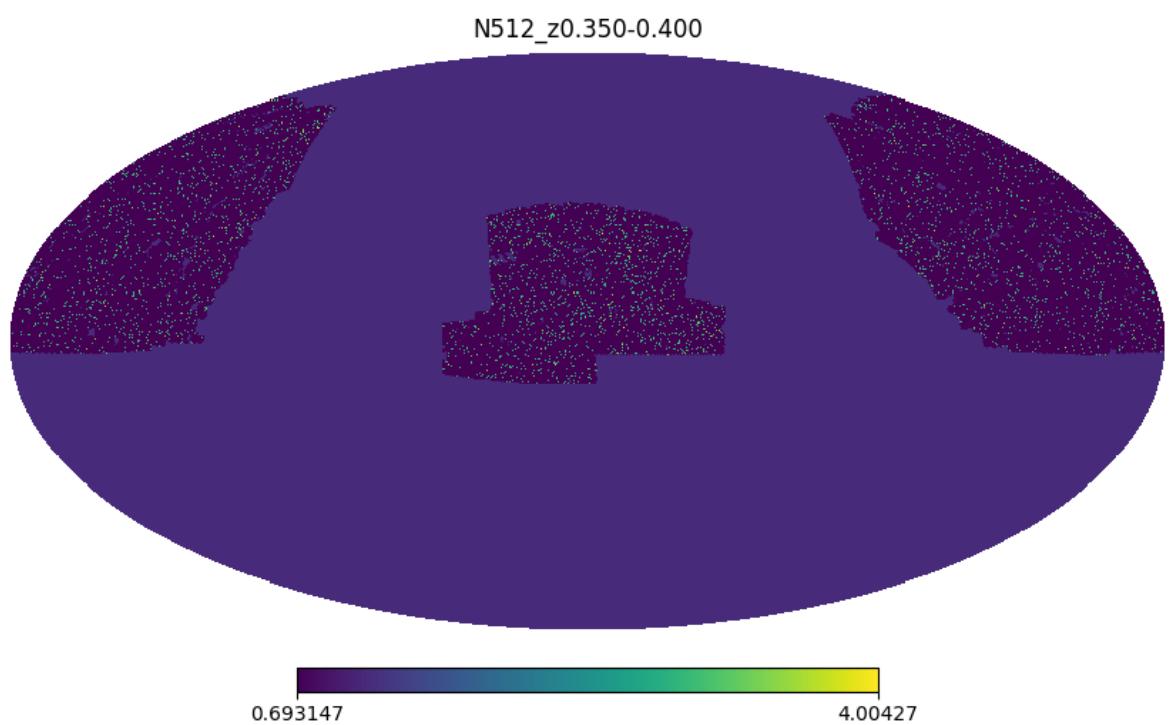


Figura 76: Mapa do intervalo de redshift 0.350-0.400 com NSIDE 512, na escala logarítmica.

## 10 Anexo III - Mapas sistemáticos

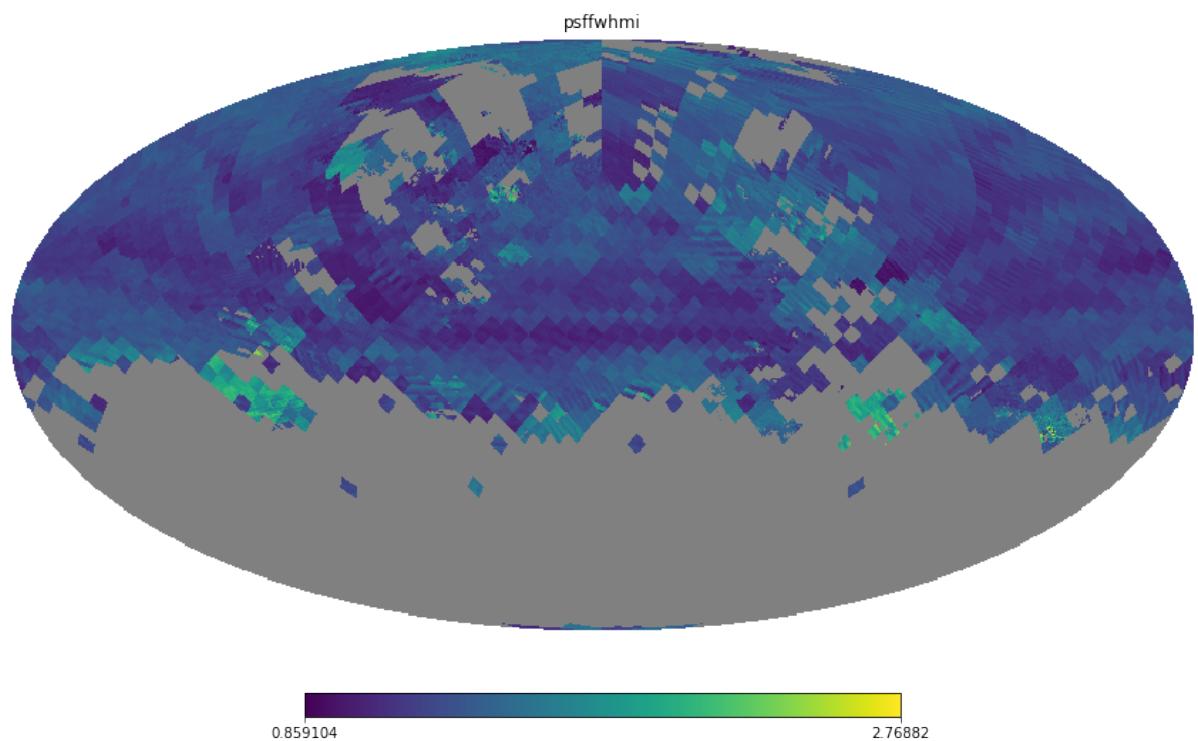


Figura 77: Mapa sistemático PSF-FWHM na banda *i*.

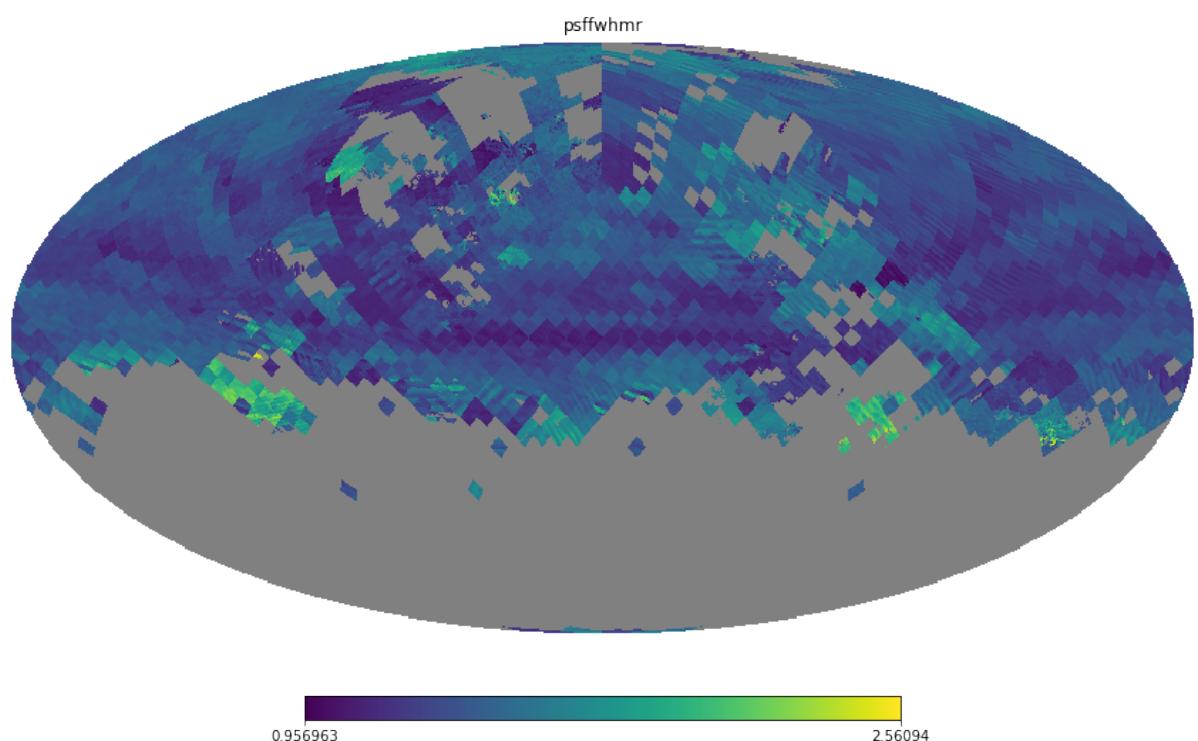


Figura 78: Mapa sistemático PSF-FWHM na banda *r*.

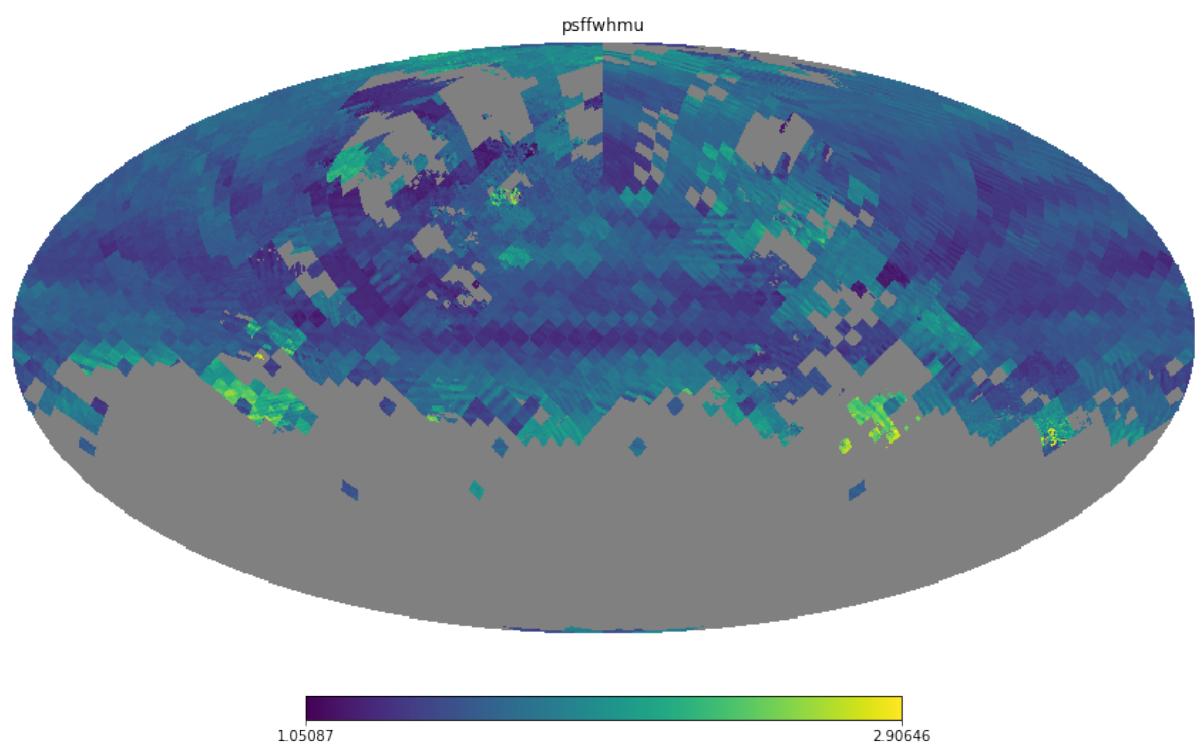


Figura 79: Mapa sistemático PSF-FWHM na banda u.

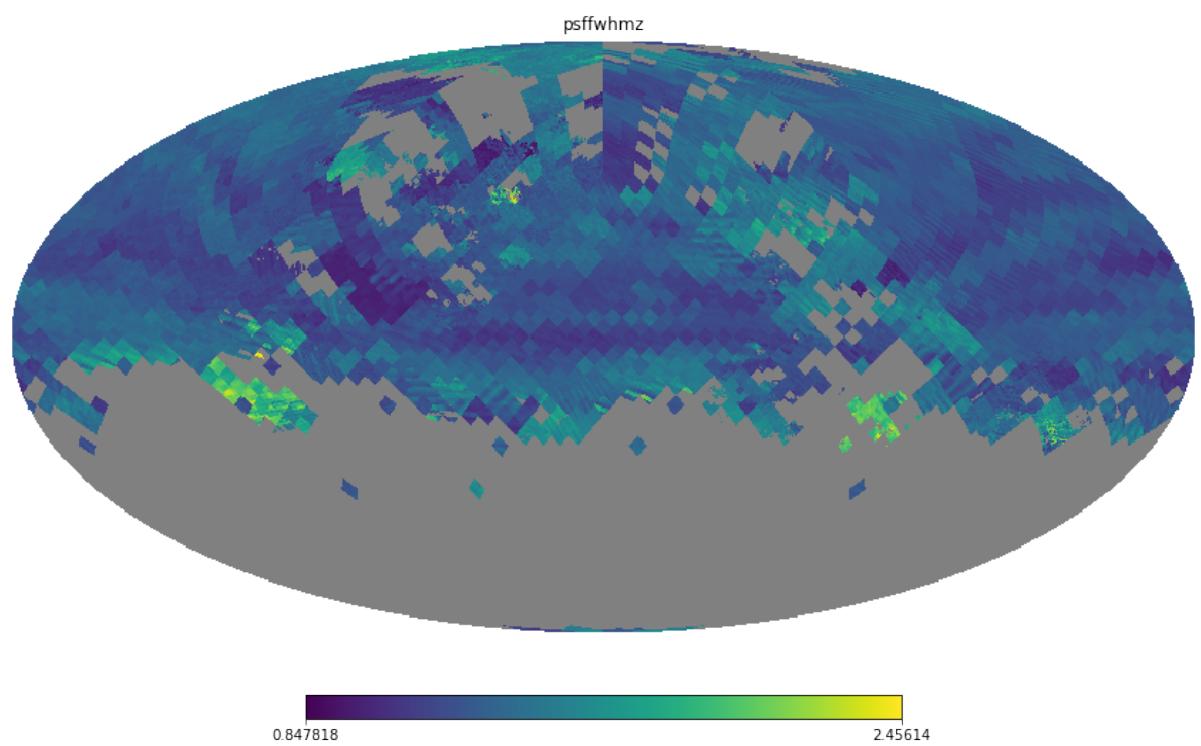


Figura 80: Mapa sistemático PSF-FWHM na banda z.

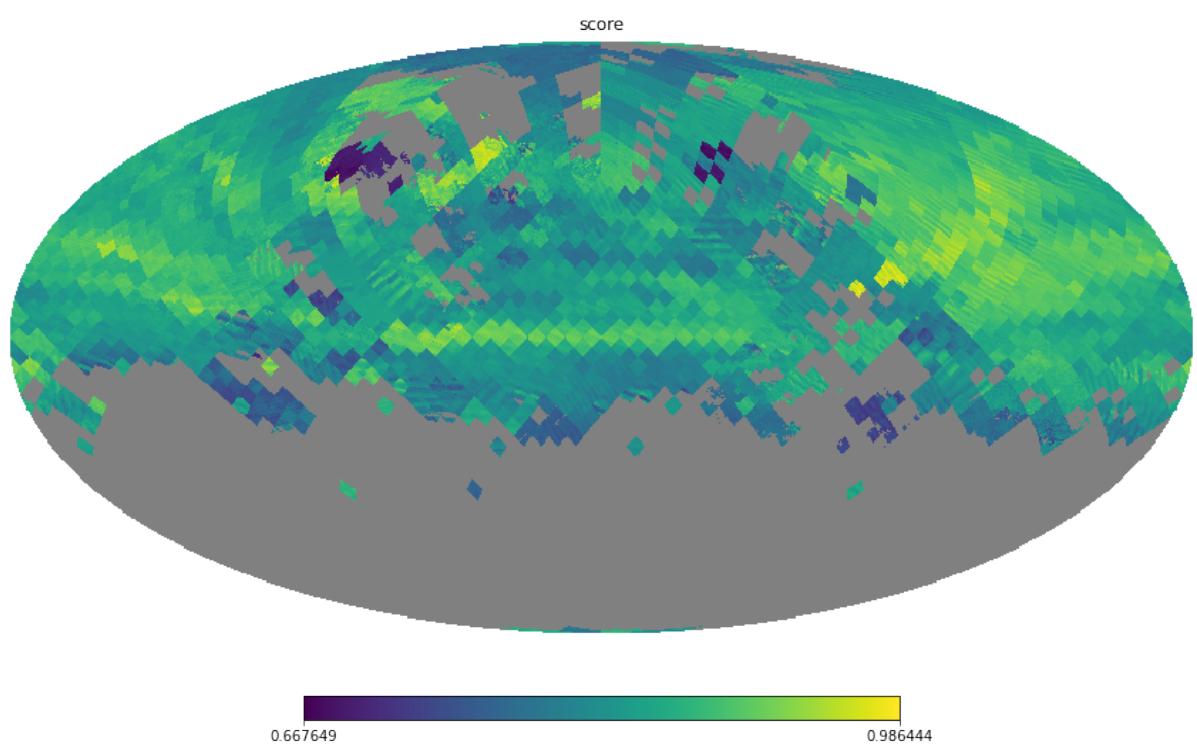


Figura 81: Mapa sistemático SCORE. Expressa a qualidade observacional de determinada região. O valor do score varia entre 0-1.

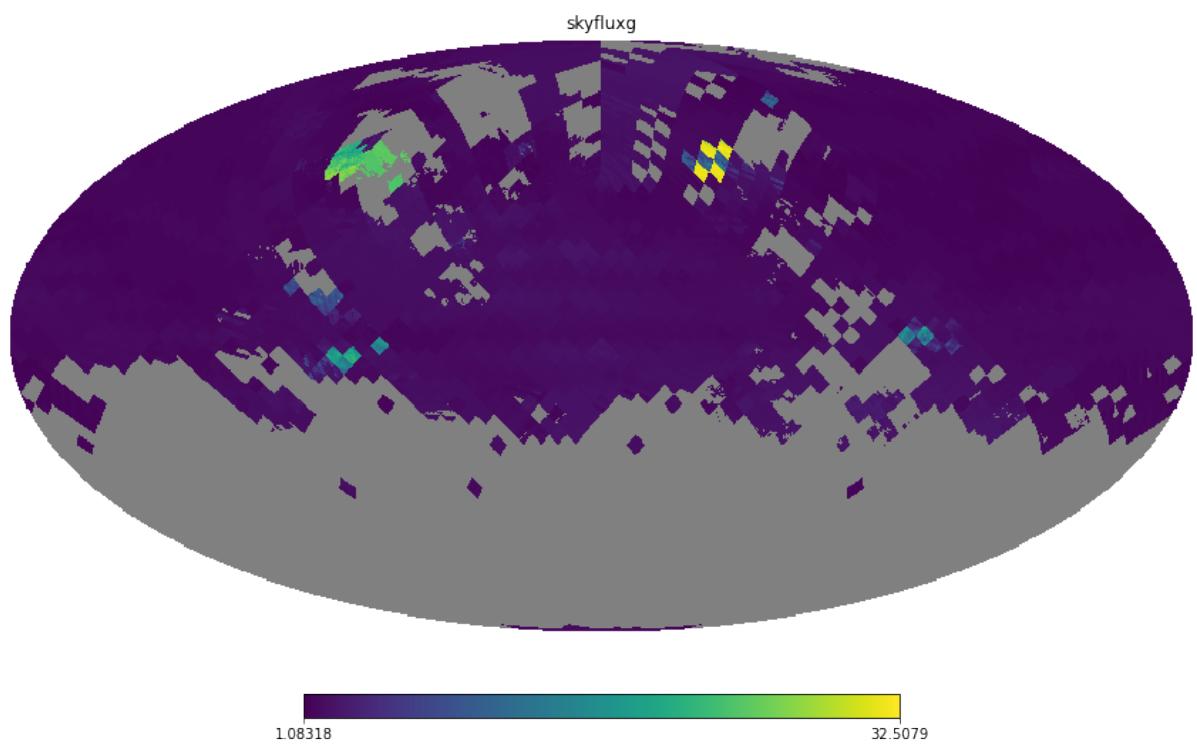


Figura 82: Mapa Skyflux na banda g.

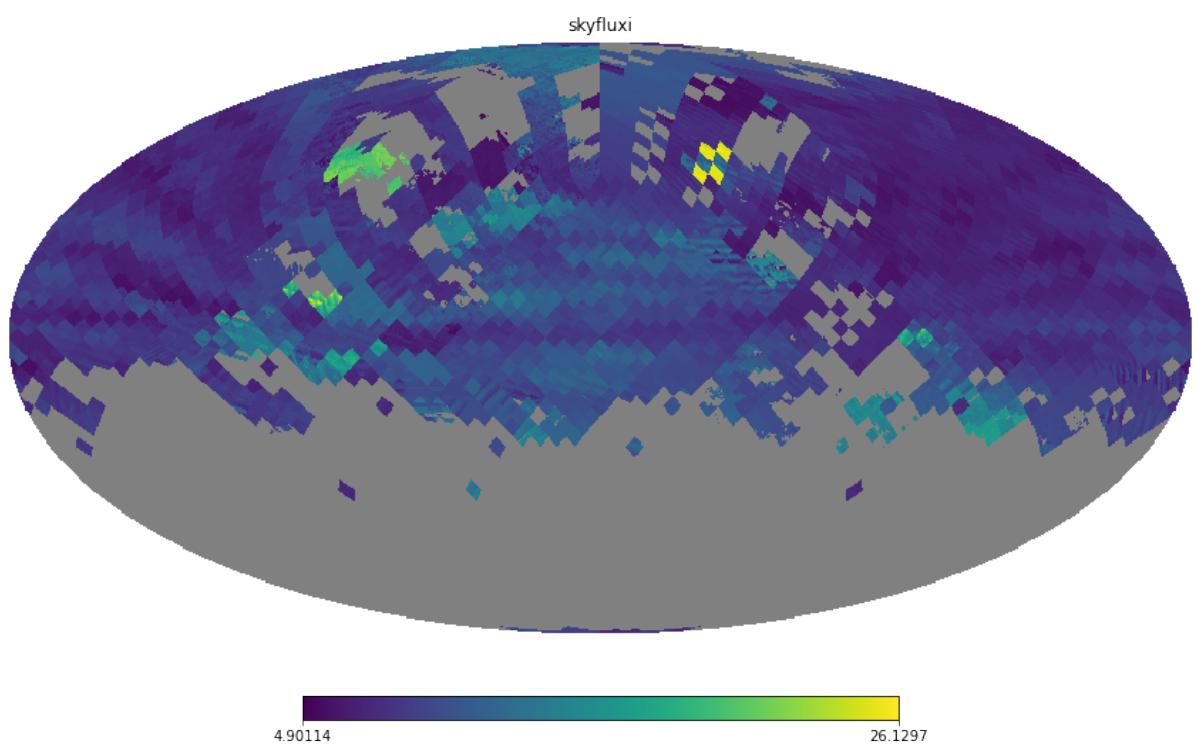


Figura 83: Mapa Skyflux na banda i.

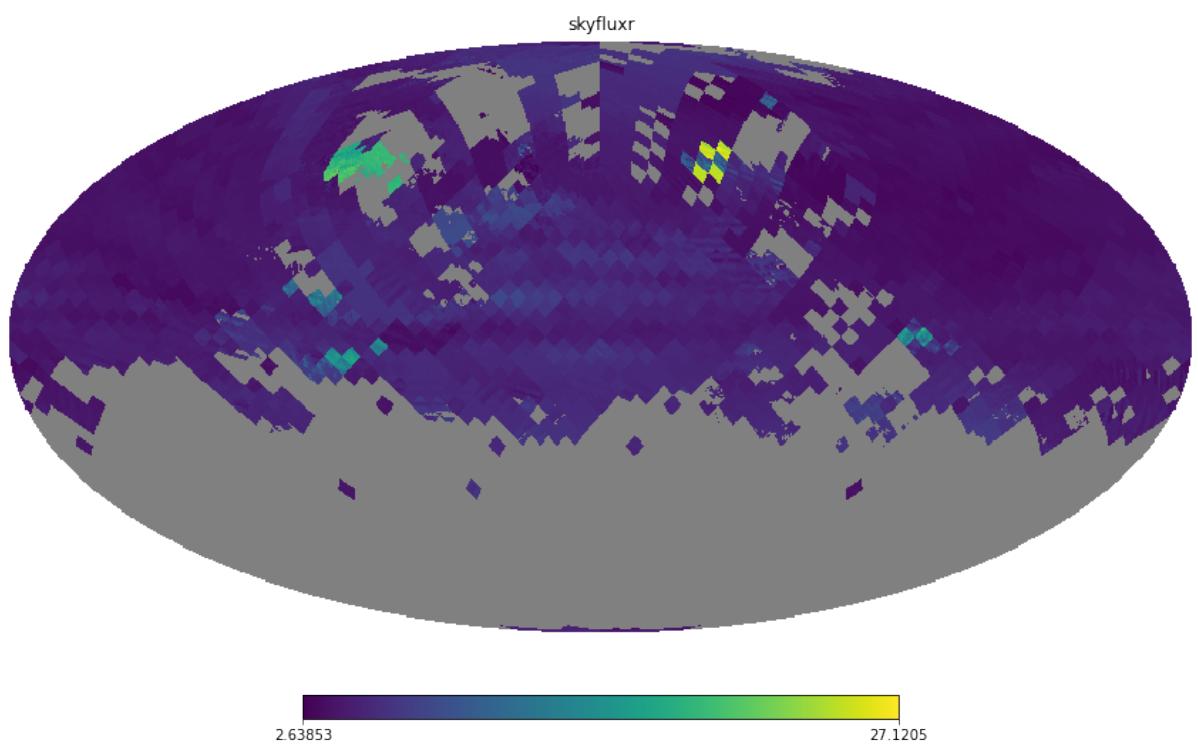


Figura 84: Mapa Skyflux na banda r.

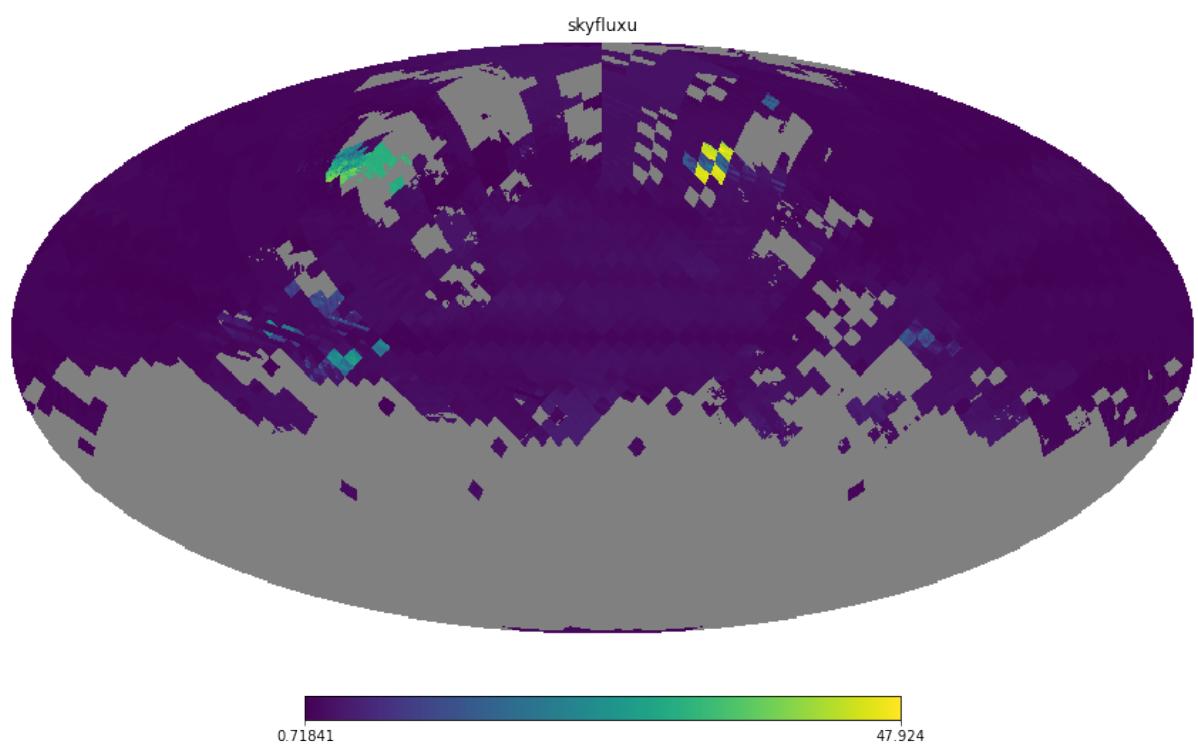


Figura 85: Mapa Skyflux na banda u.

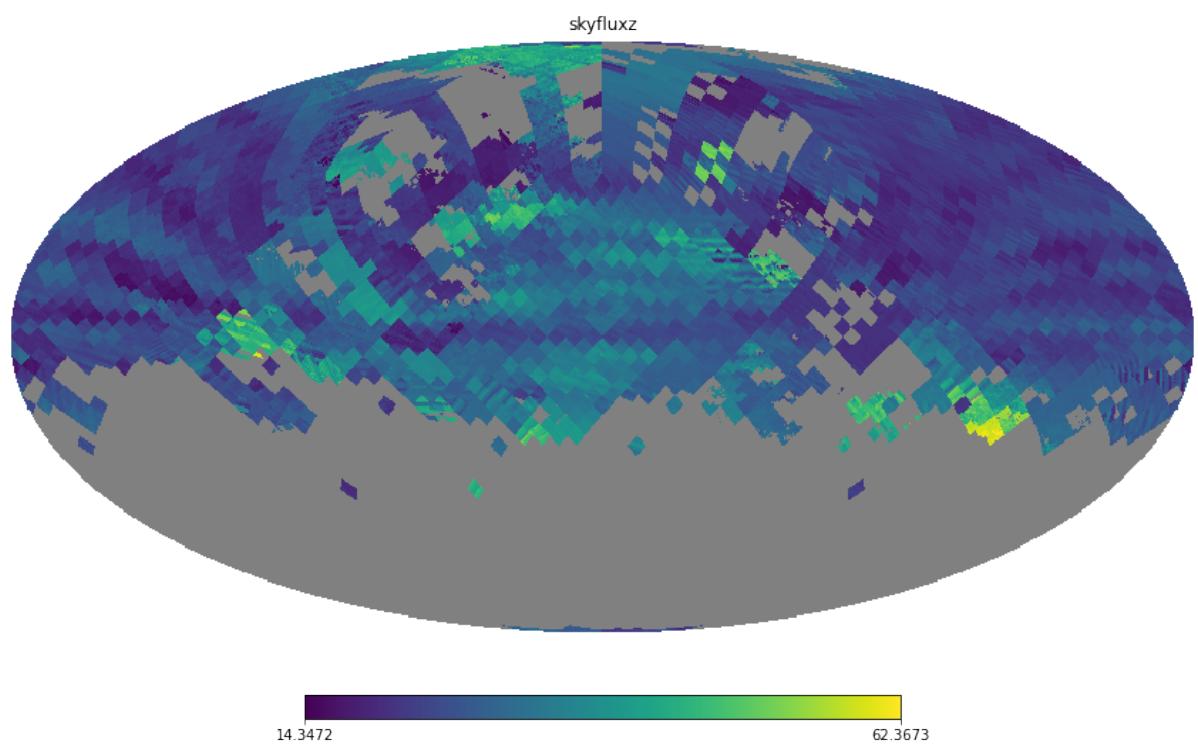


Figura 86: Mapa Skyflux na banda z.

## 11 Anexo IV - $C_\ell$ e $S_\ell$ dos mapas cosmológicos

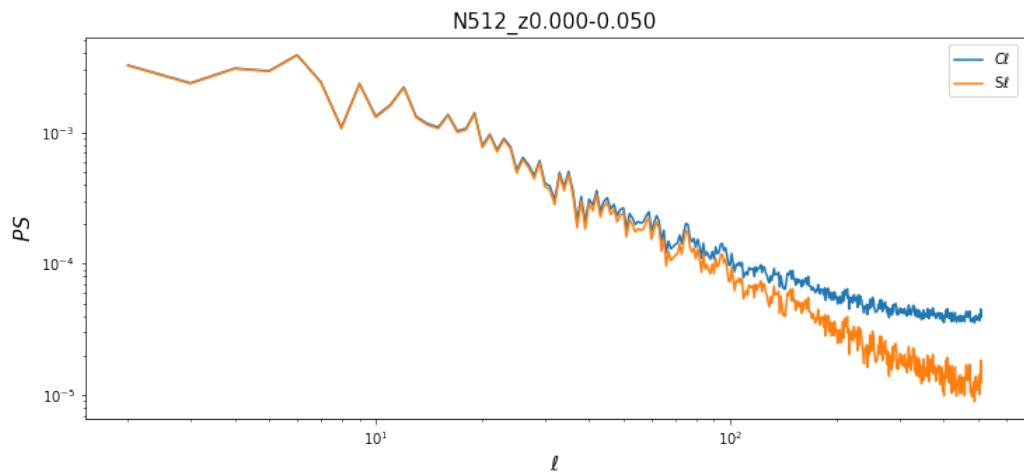


Figura 87:  $C_\ell$  e  $S_\ell$  do mapa com redshift entre 0.000-0.050.

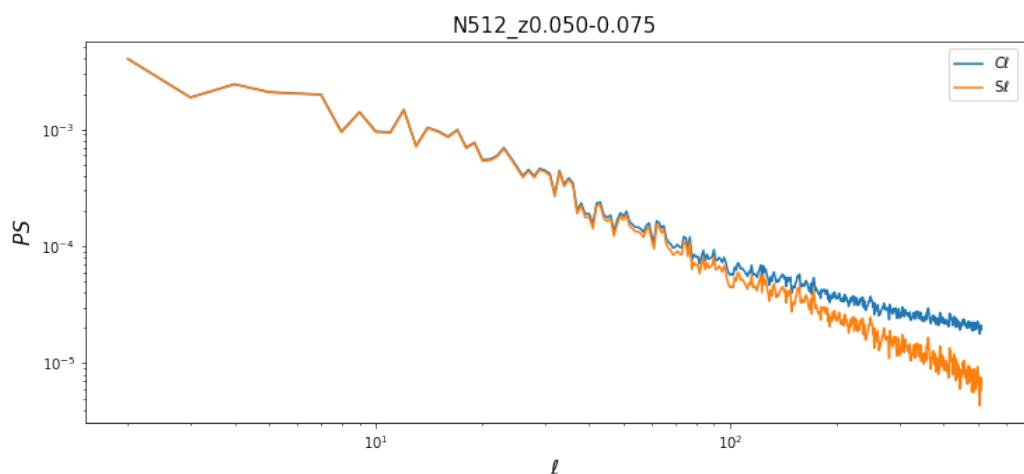


Figura 88:  $C_\ell$  e  $S_\ell$  do mapa com redshift entre 0.050-0.075.

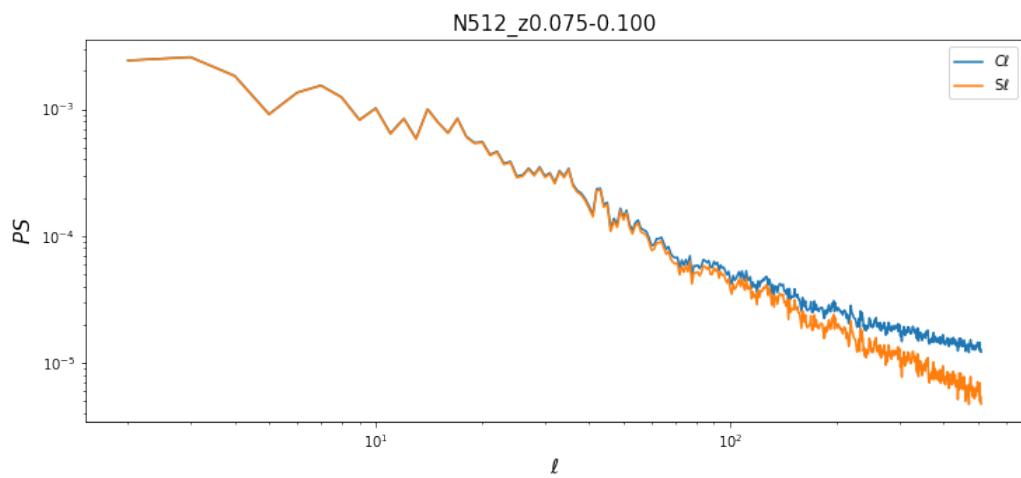


Figura 89:  $C_\ell$  e  $S_\ell$  do mapa com redshift entre 0.075-0.100.

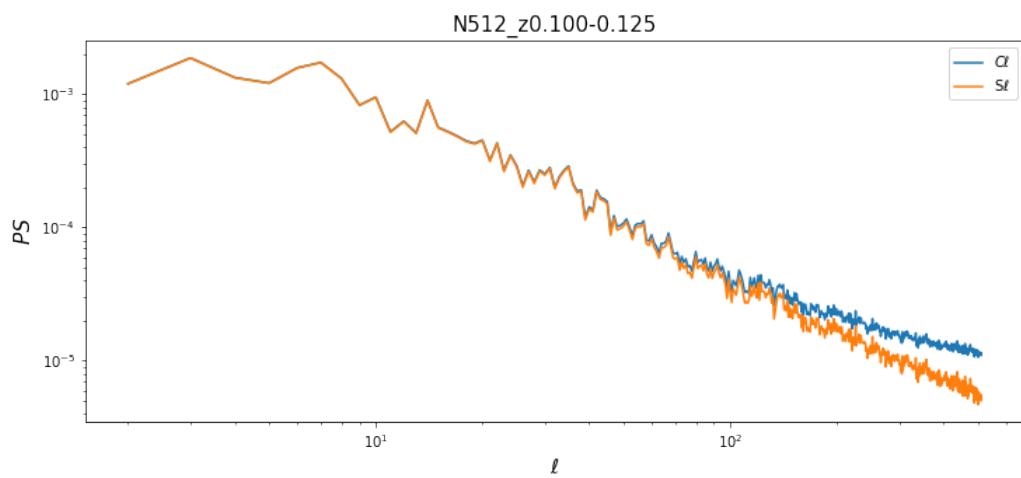


Figura 90:  $C_\ell$  e  $S_\ell$  do mapa com redshift entre 0.100-0.125.

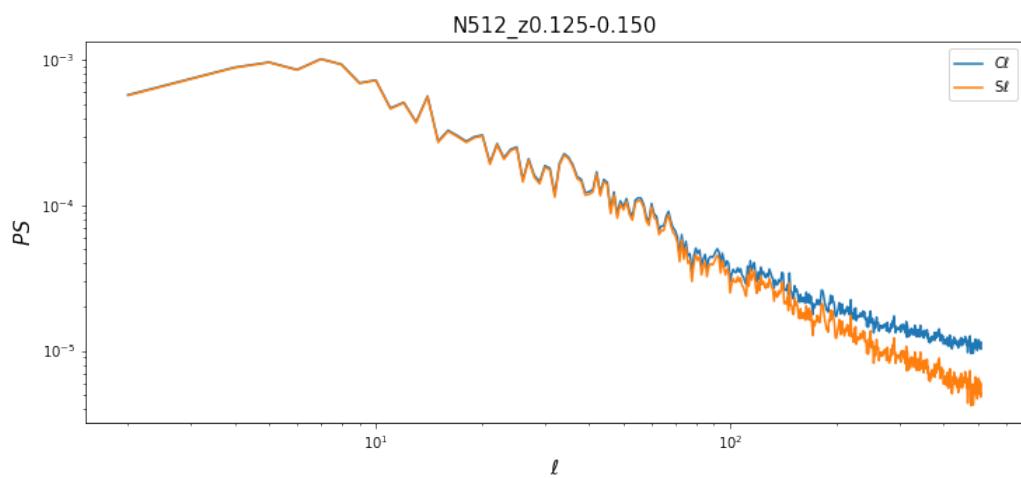


Figura 91:  $C_\ell$  e  $S_\ell$  do mapa com redshift entre 0.125-0.150.

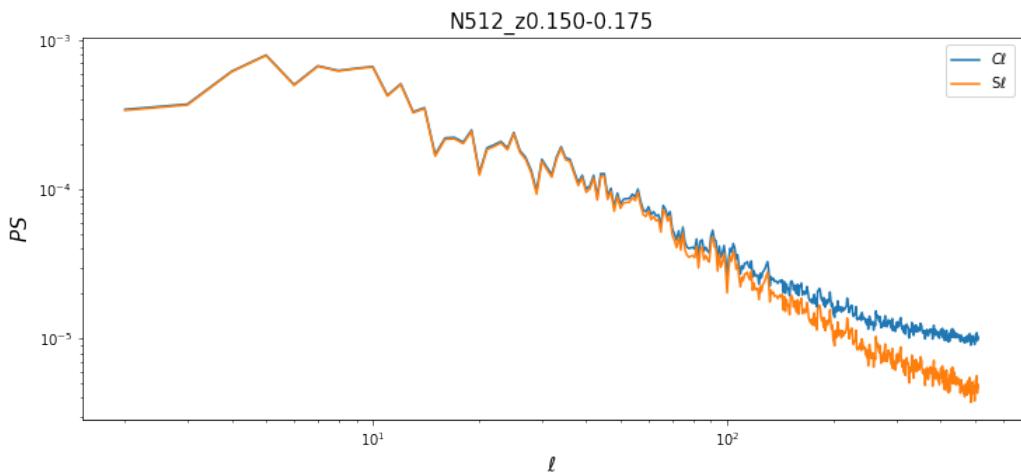


Figura 92:  $C_\ell$  e  $S_\ell$  do mapa com redshift entre 0.150-0.175.

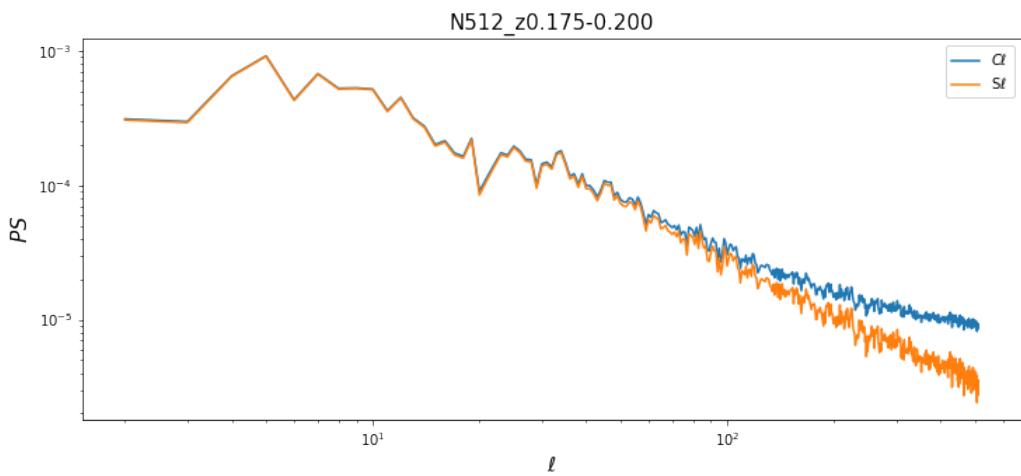


Figura 93:  $C_\ell$  e  $S_\ell$  do mapa com redshift entre 0.175-0.200.

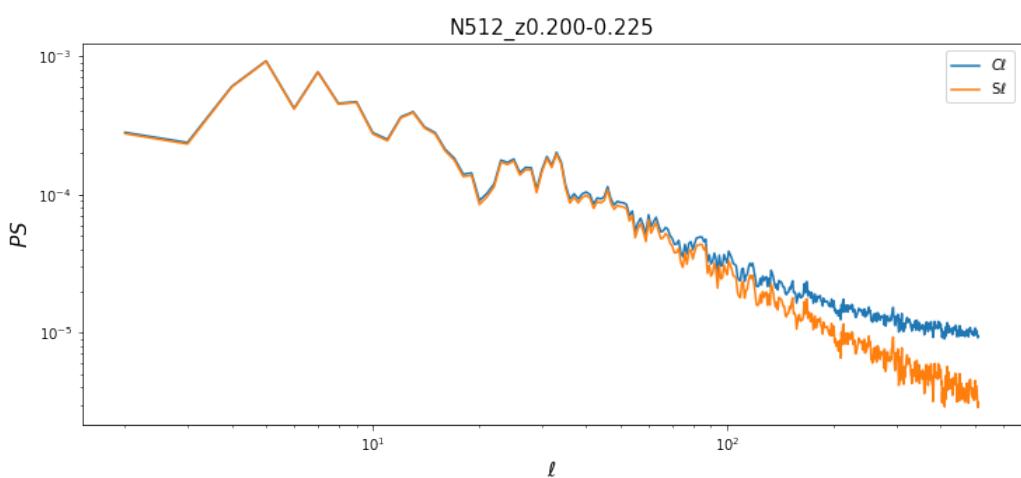


Figura 94:  $C_\ell$  e  $S_\ell$  do mapa com redshift entre 0.200-0.225.

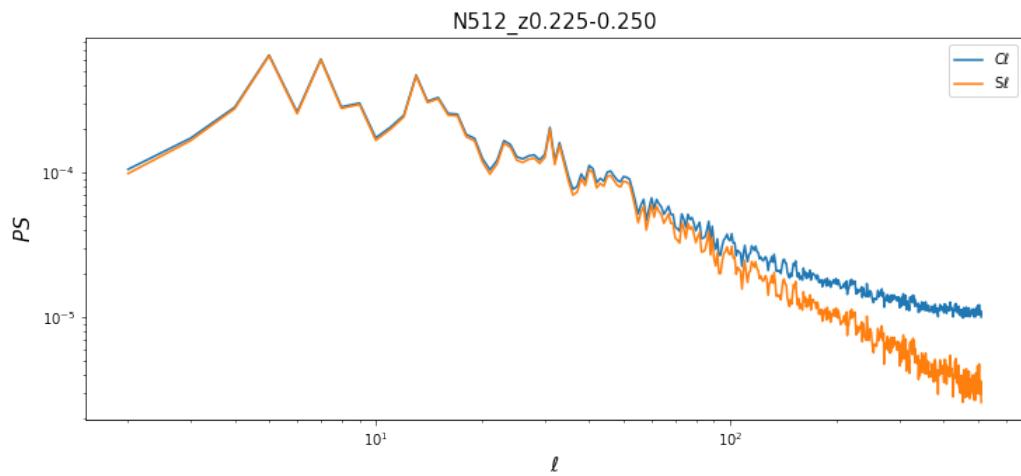


Figura 95:  $C_\ell$  e  $S_\ell$  do mapa com redshift entre 0.225-0.250.

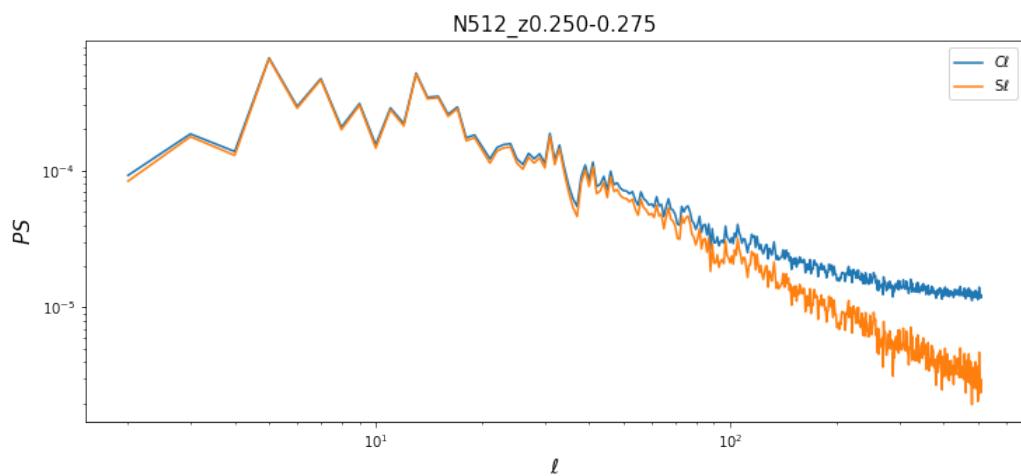


Figura 96:  $C_\ell$  e  $S_\ell$  do mapa com redshift entre 0.250-0.275.

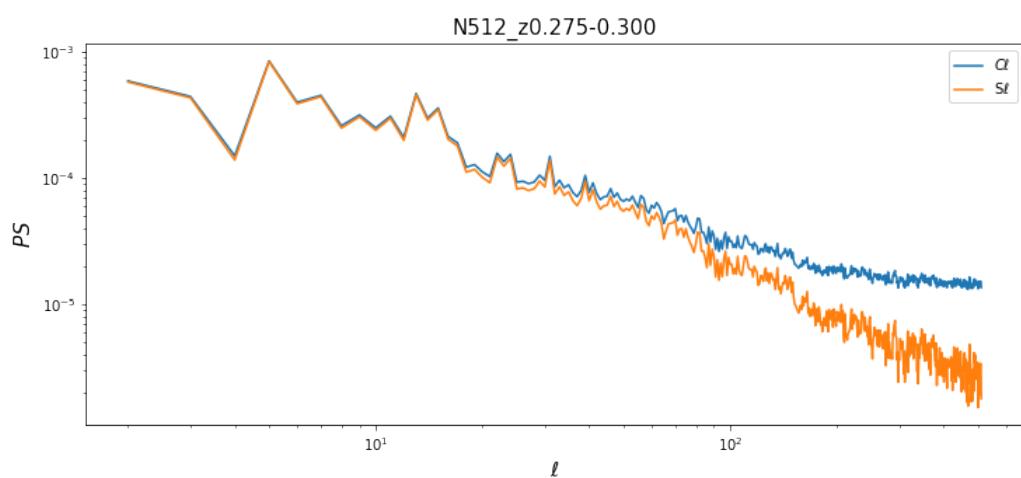


Figura 97:  $C_\ell$  e  $S_\ell$  do mapa com redshift entre 0.275-0.300.

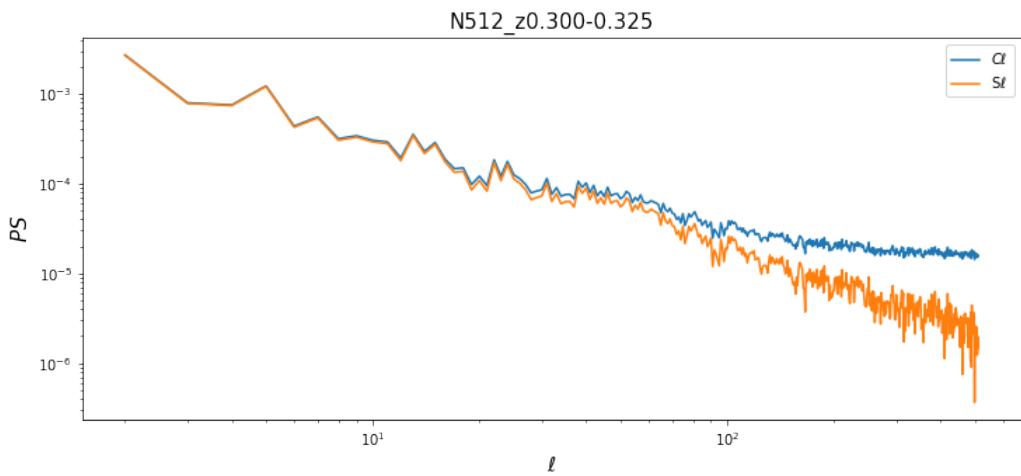


Figura 98:  $C_\ell$  e  $S_\ell$  do mapa com redshift entre 0.300-0.325.

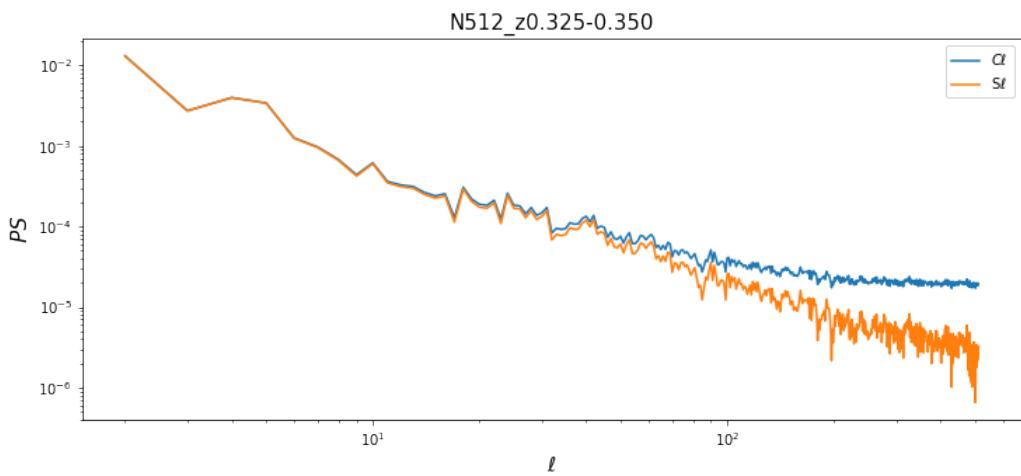


Figura 99:  $C_\ell$  e  $S_\ell$  do mapa com redshift entre 0.325-0.350.

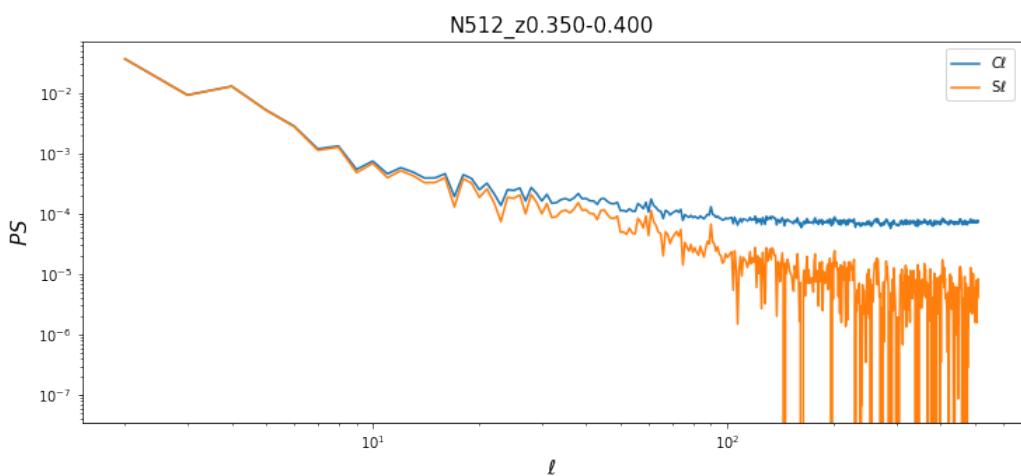


Figura 100:  $C_\ell$  e  $S_\ell$  do mapa com redshift entre 0.350-0.400.

## 12 Anexo V - Matrizes de covariância: $C\ell$

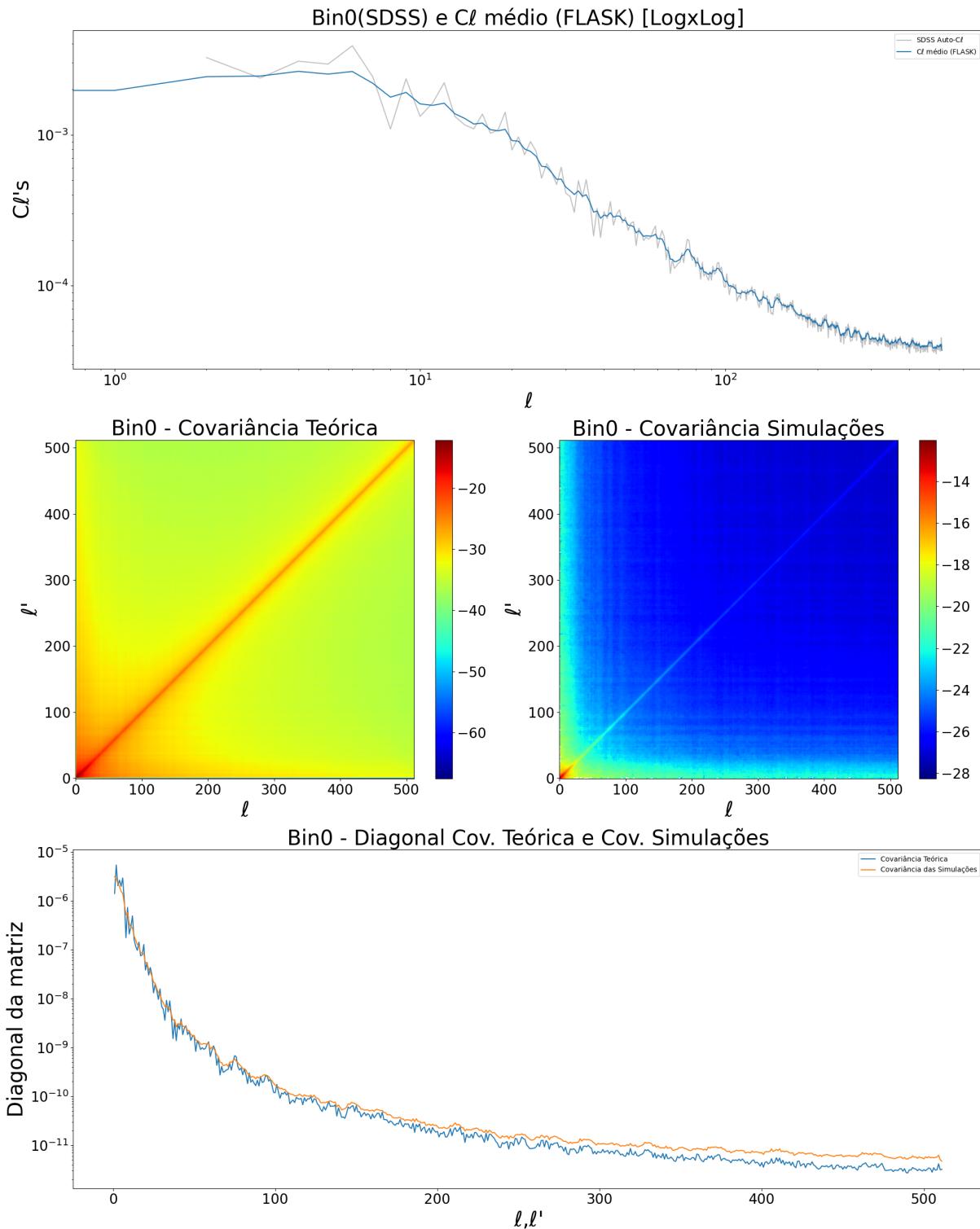


Figura 101: No primeiro gráfico há uma comparação entre o  $C\ell$  médio das simulações do FLASK e o  $C\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

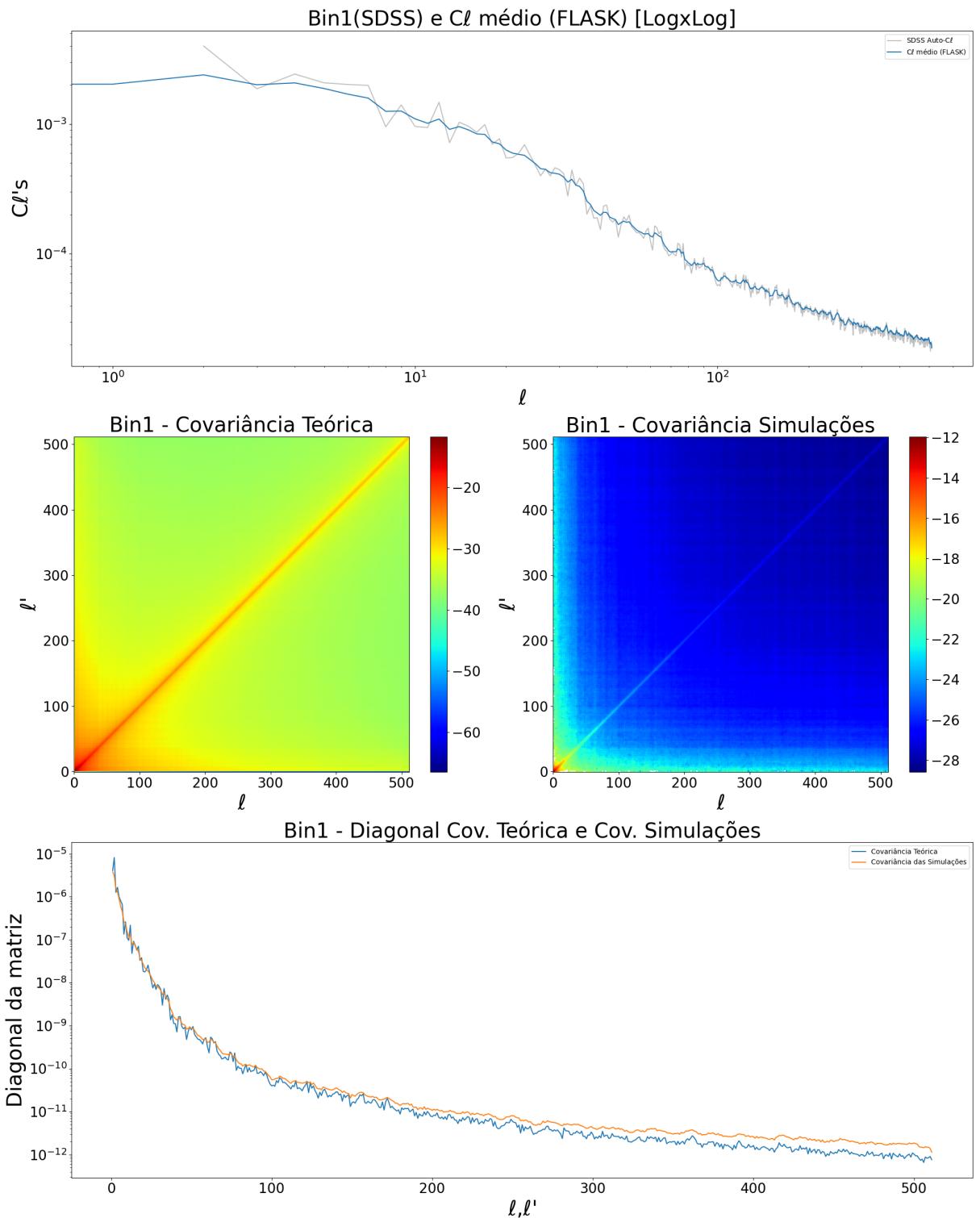


Figura 102: No primeiro gráfico há uma comparação entre o  $C\ell$  médio das simulações do FLASK e o  $C\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

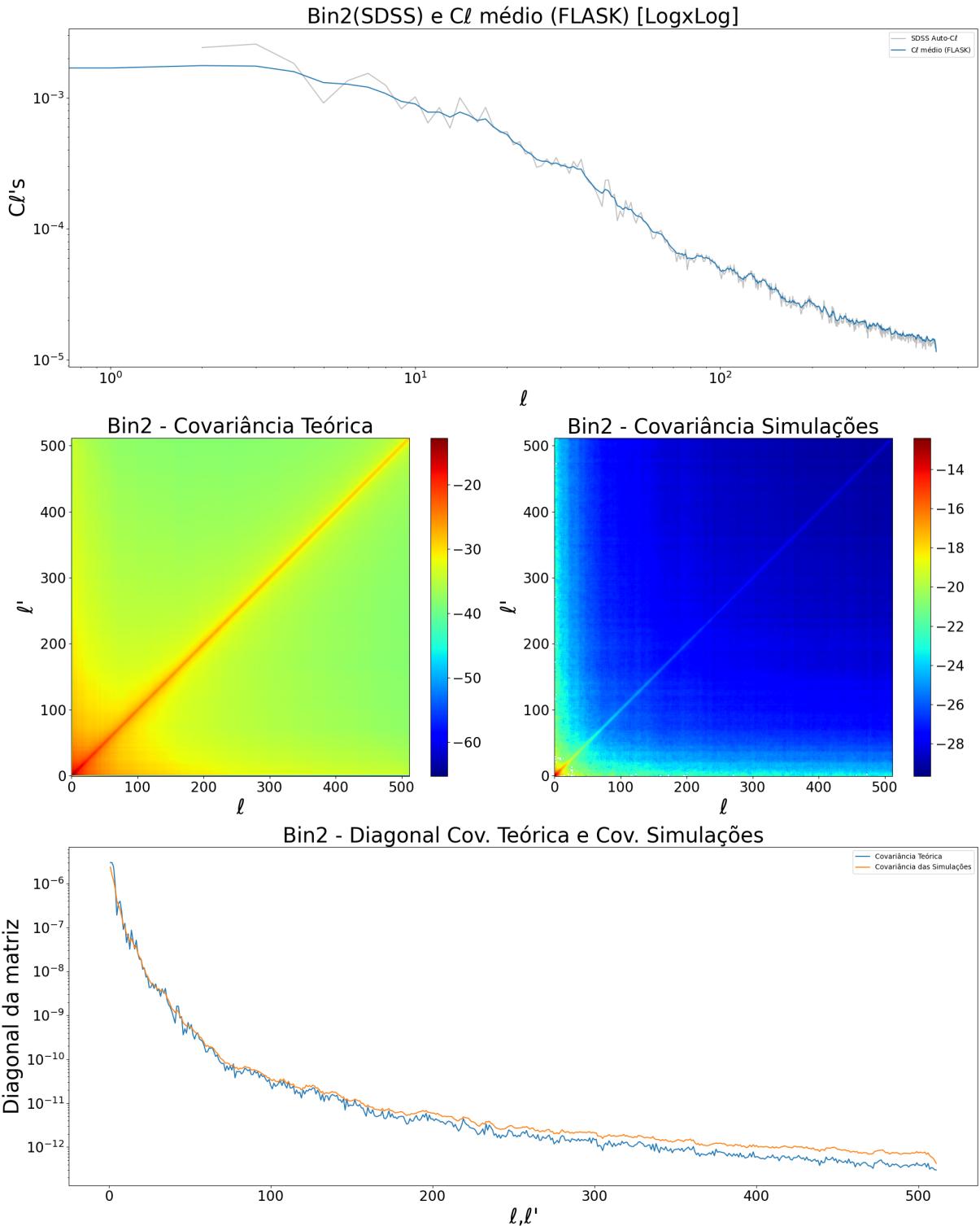


Figura 103: No primeiro gráfico há uma comparação entre o  $C\ell$  médio das simulações do FLASK e o  $C\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

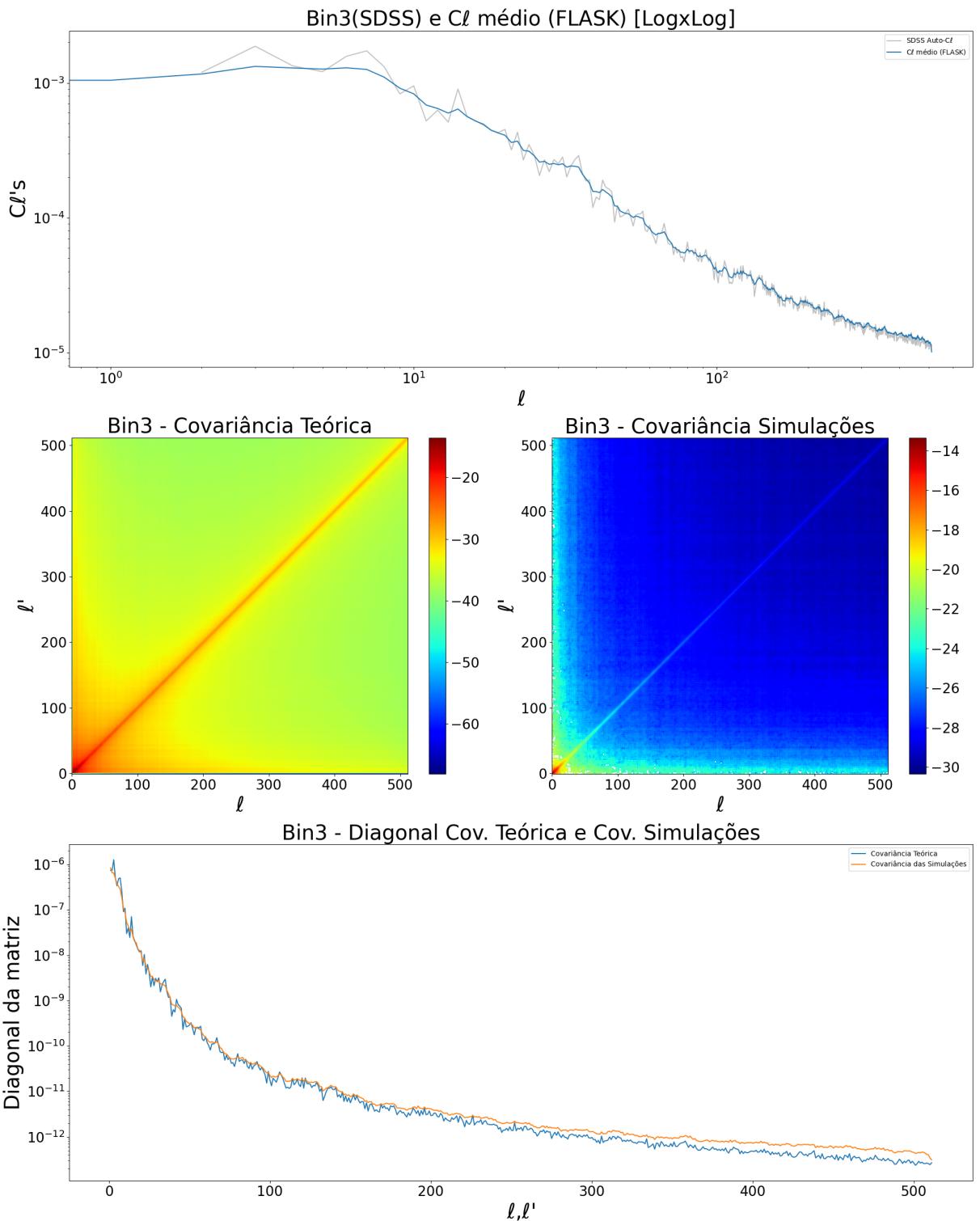


Figura 104: No primeiro gráfico há uma comparação entre o  $C\ell$  médio das simulações do FLASK e o  $C\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

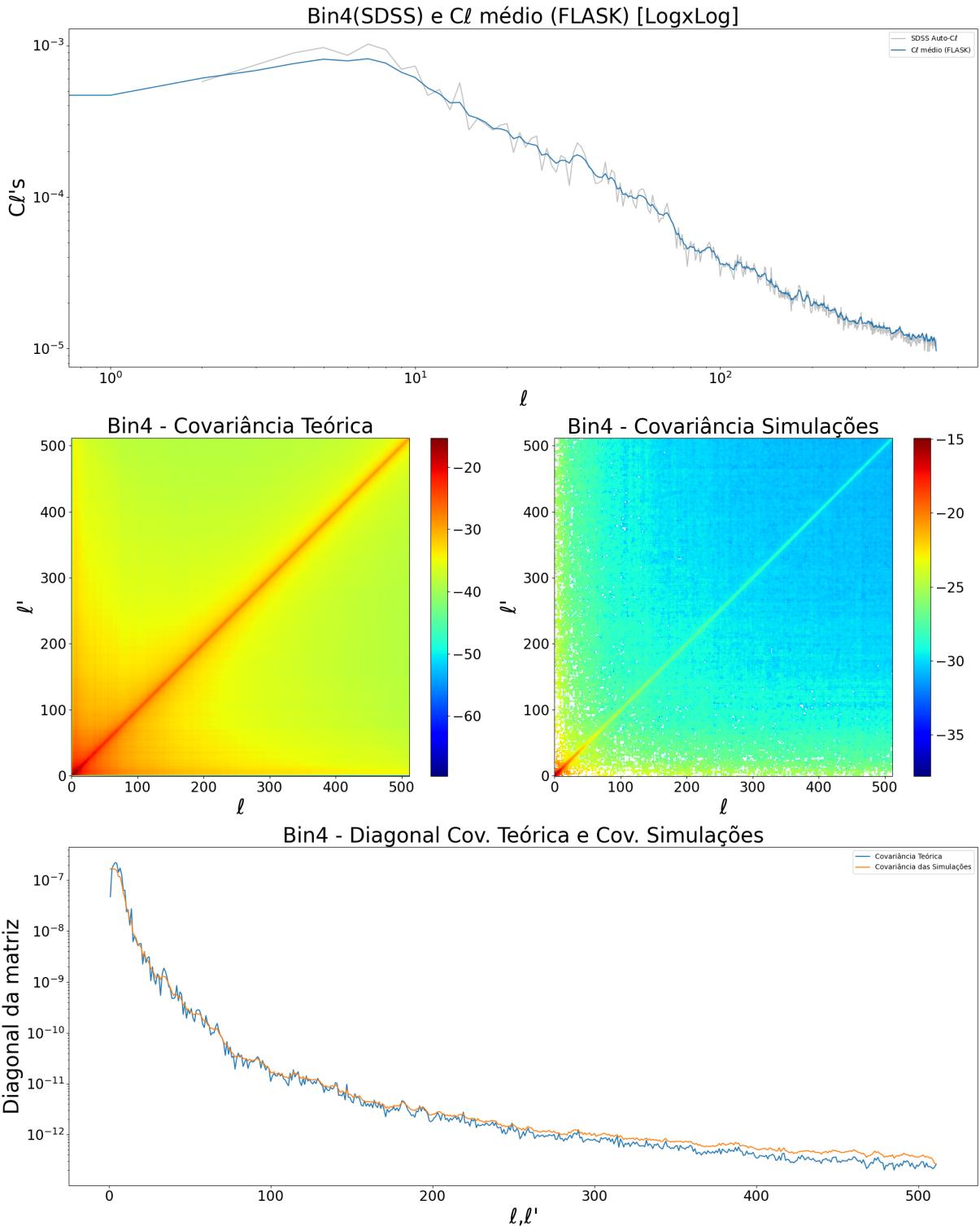


Figura 105: No primeiro gráfico há uma comparação entre o  $C\ell$  médio das simulações do FLASK e o  $C\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

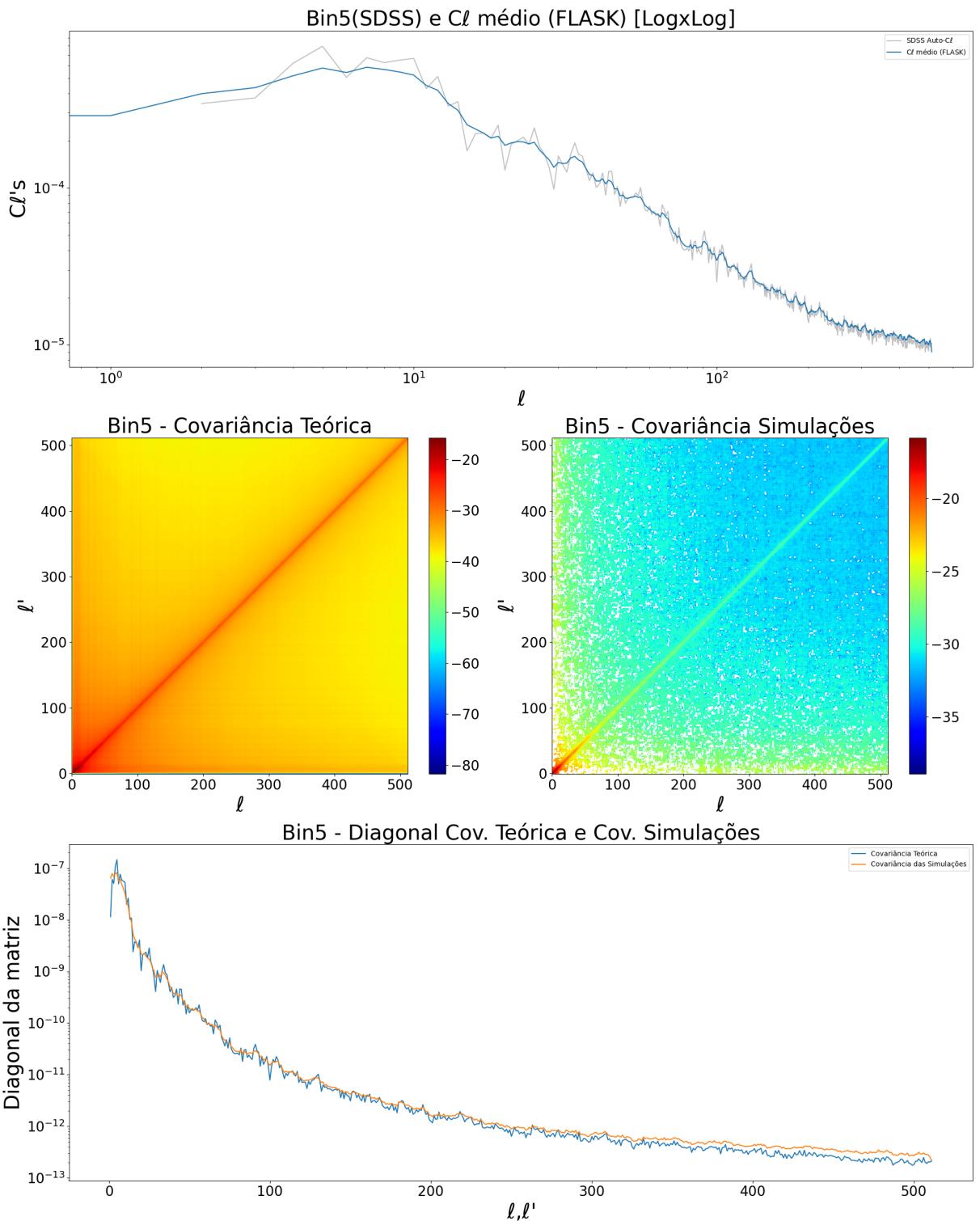


Figura 106: No primeiro gráfico há uma comparação entre o  $C\ell$  médio das simulações do FLASK e o  $C\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

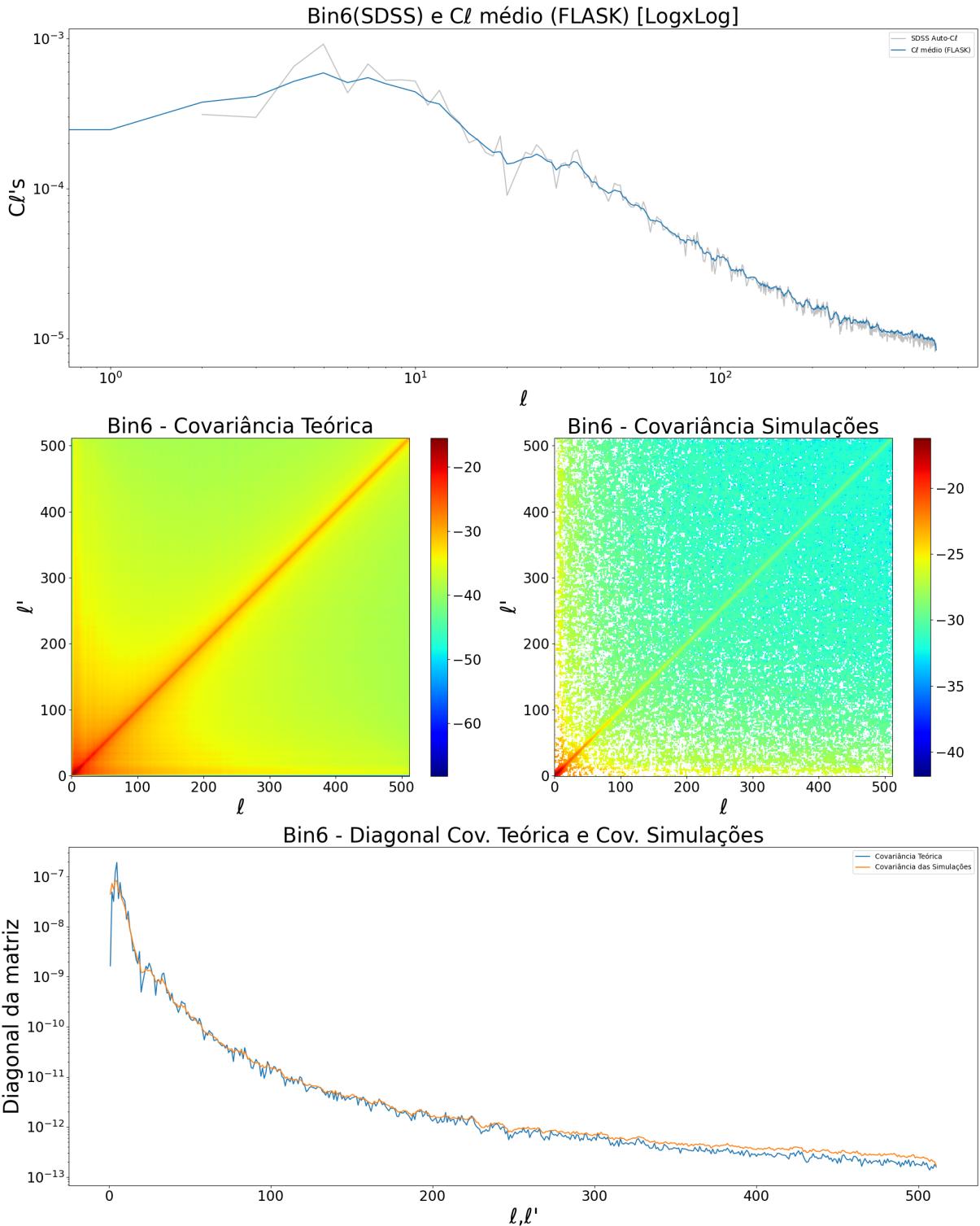


Figura 107: No primeiro gráfico há uma comparação entre o  $C\ell$  médio das simulações do FLASK e o  $C\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

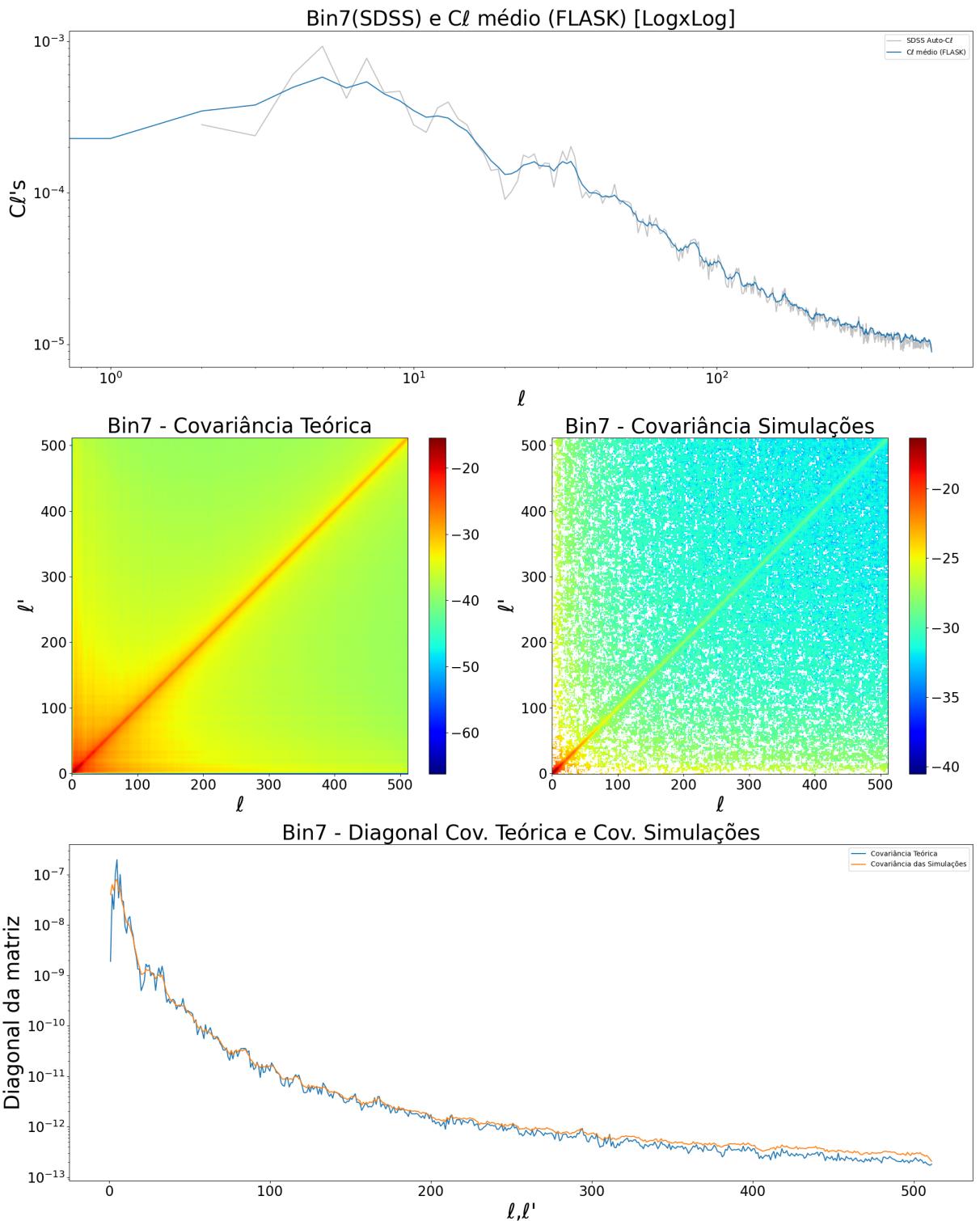


Figura 108: No primeiro gráfico há uma comparação entre o  $C\ell$  médio das simulações do FLASK e o  $C\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

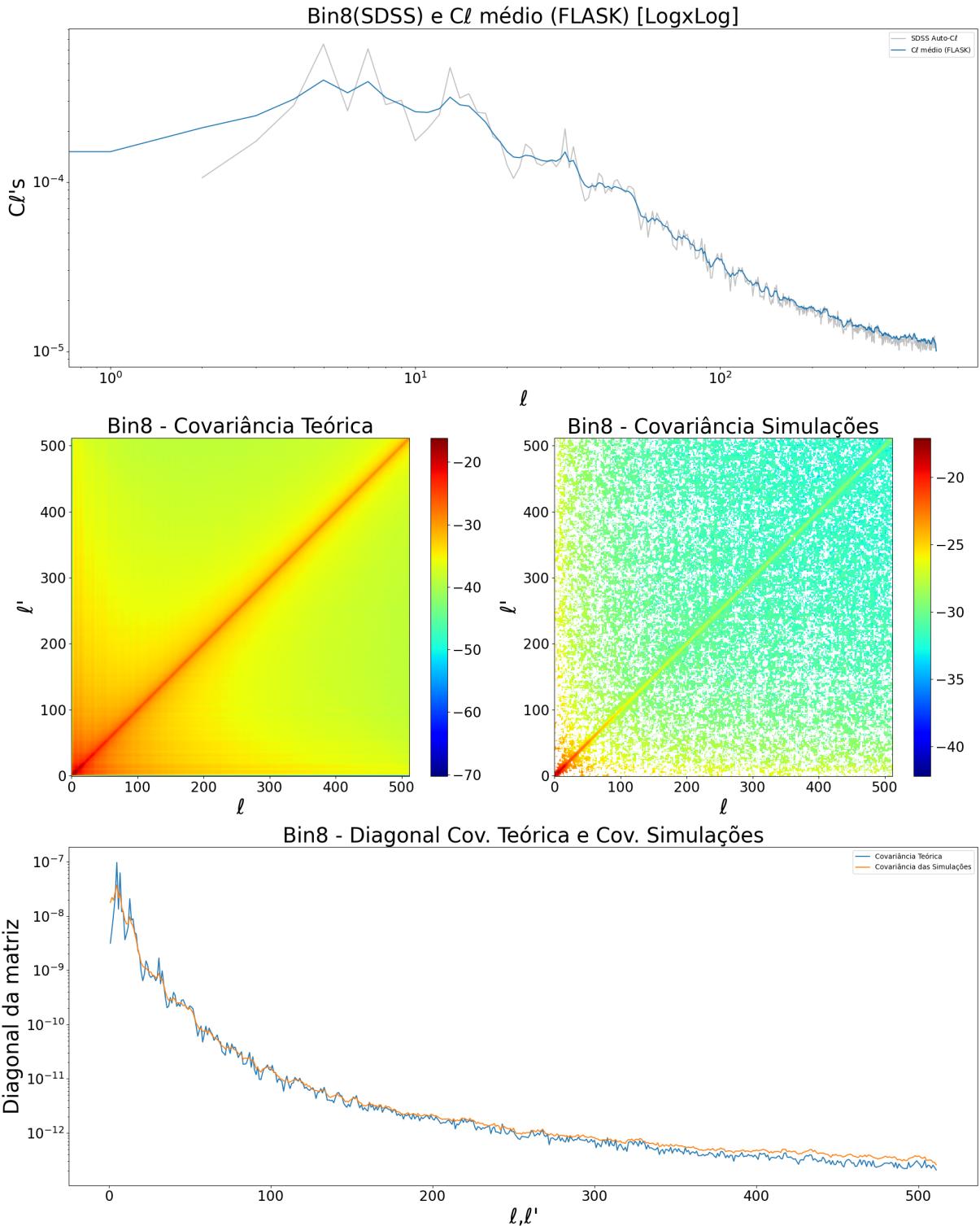


Figura 109: No primeiro gráfico há uma comparação entre o  $C\ell$  médio das simulações do FLASK e o  $C\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

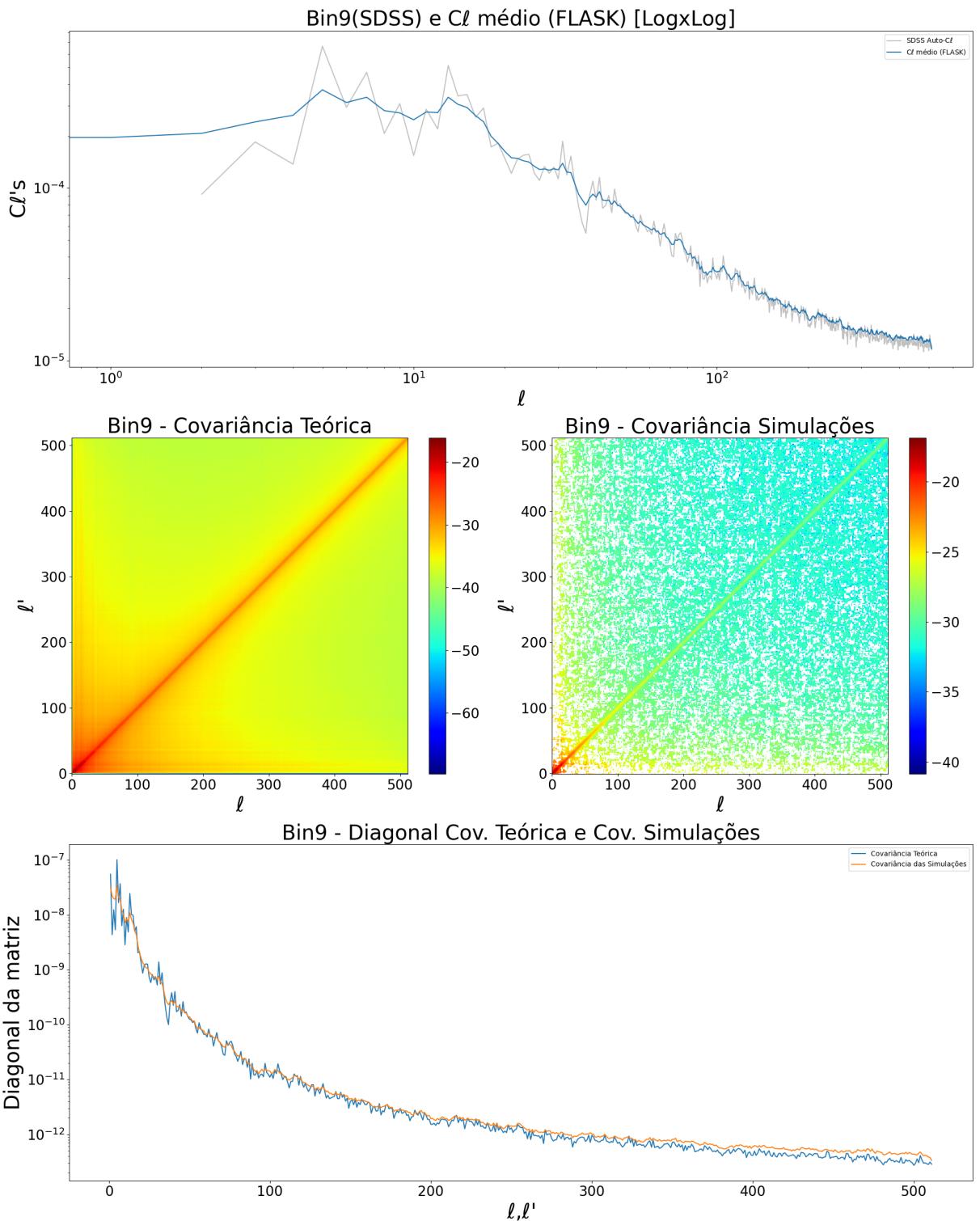


Figura 110: No primeiro gráfico há uma comparação entre o  $C\ell$  médio das simulações do FLASK e o  $C\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

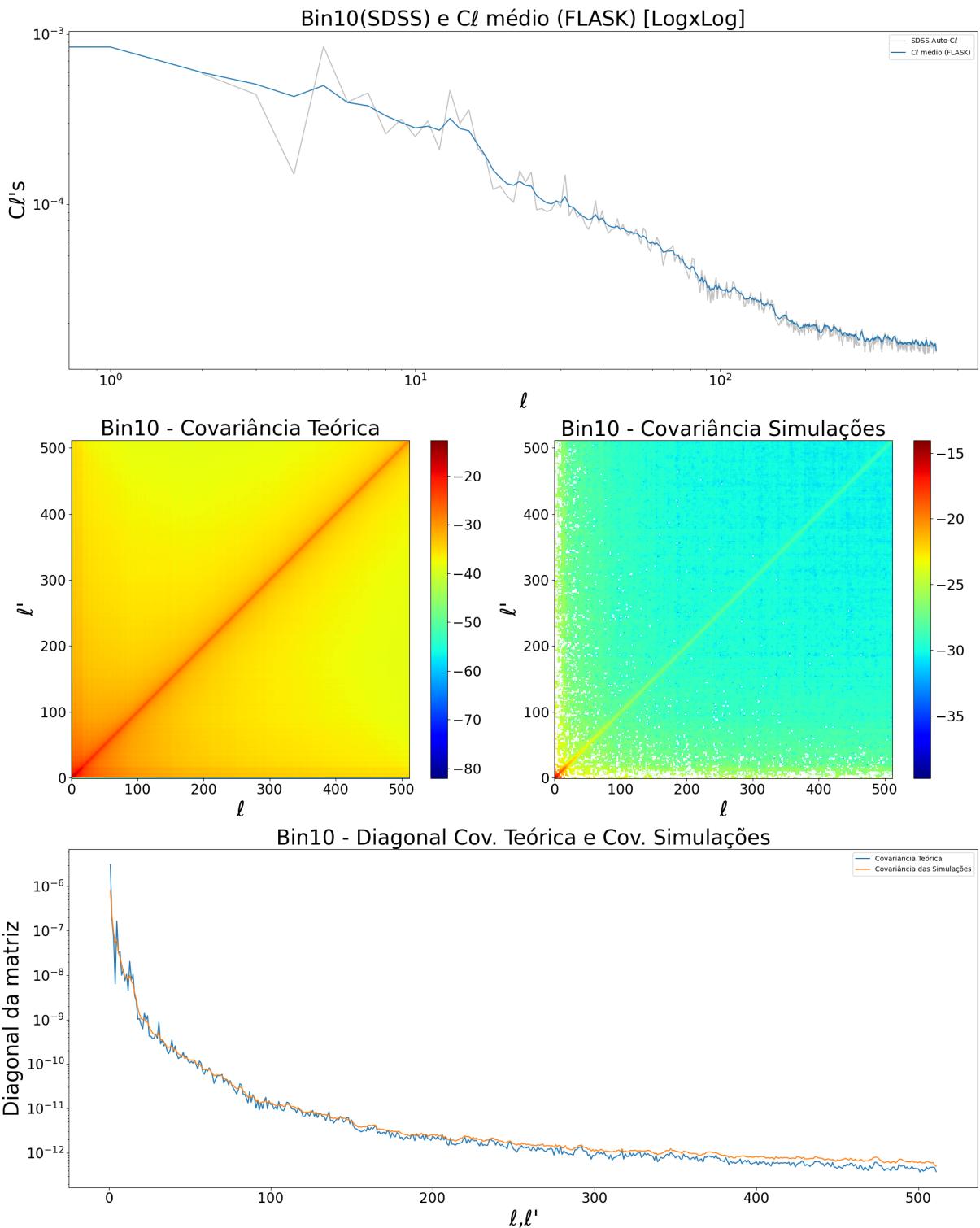


Figura 111: No primeiro gráfico há uma comparação entre o  $C\ell$  médio das simulações do FLASK e o  $C\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

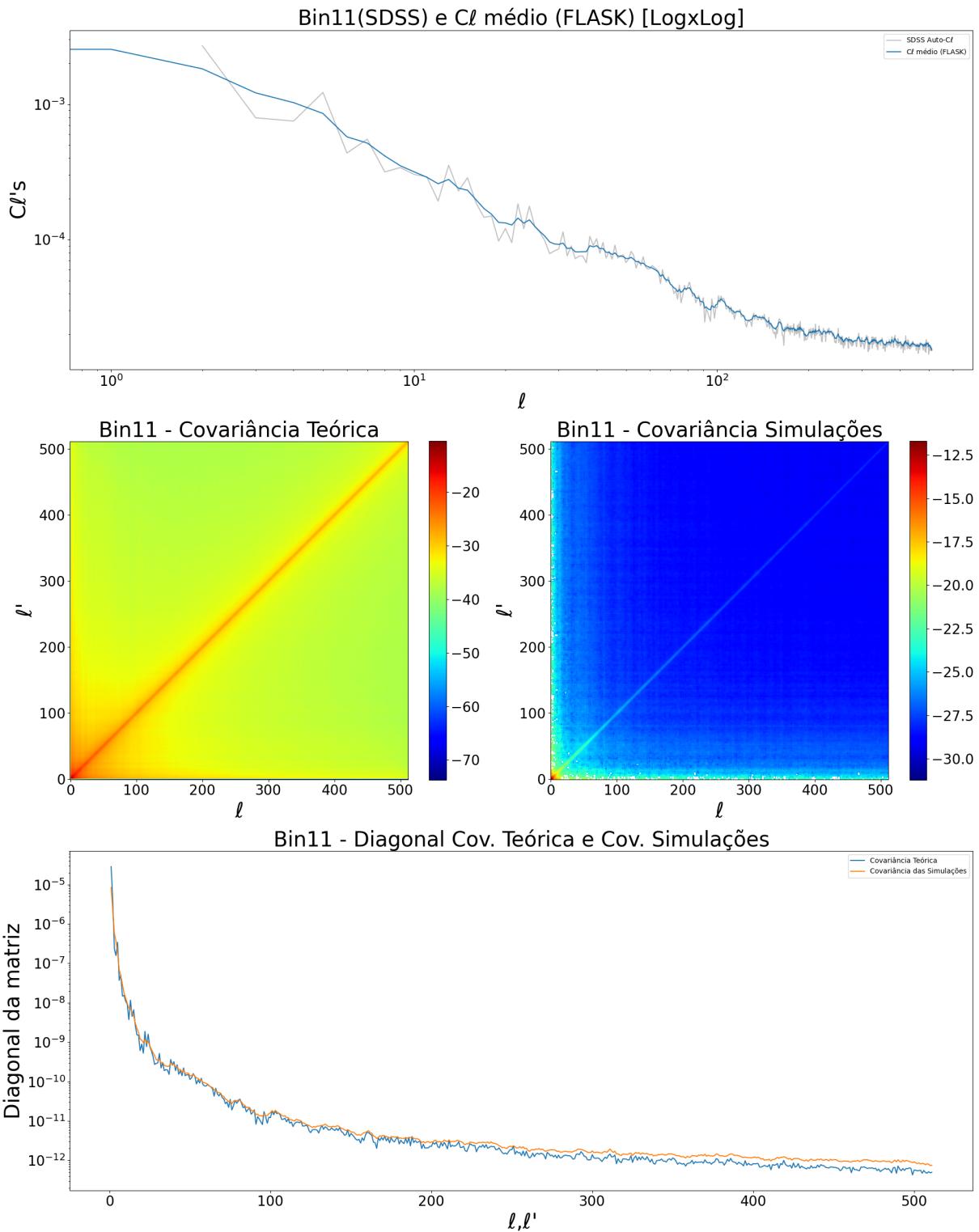


Figura 112: No primeiro gráfico há uma comparação entre o  $C\ell$  médio das simulações do FLASK e o  $C\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

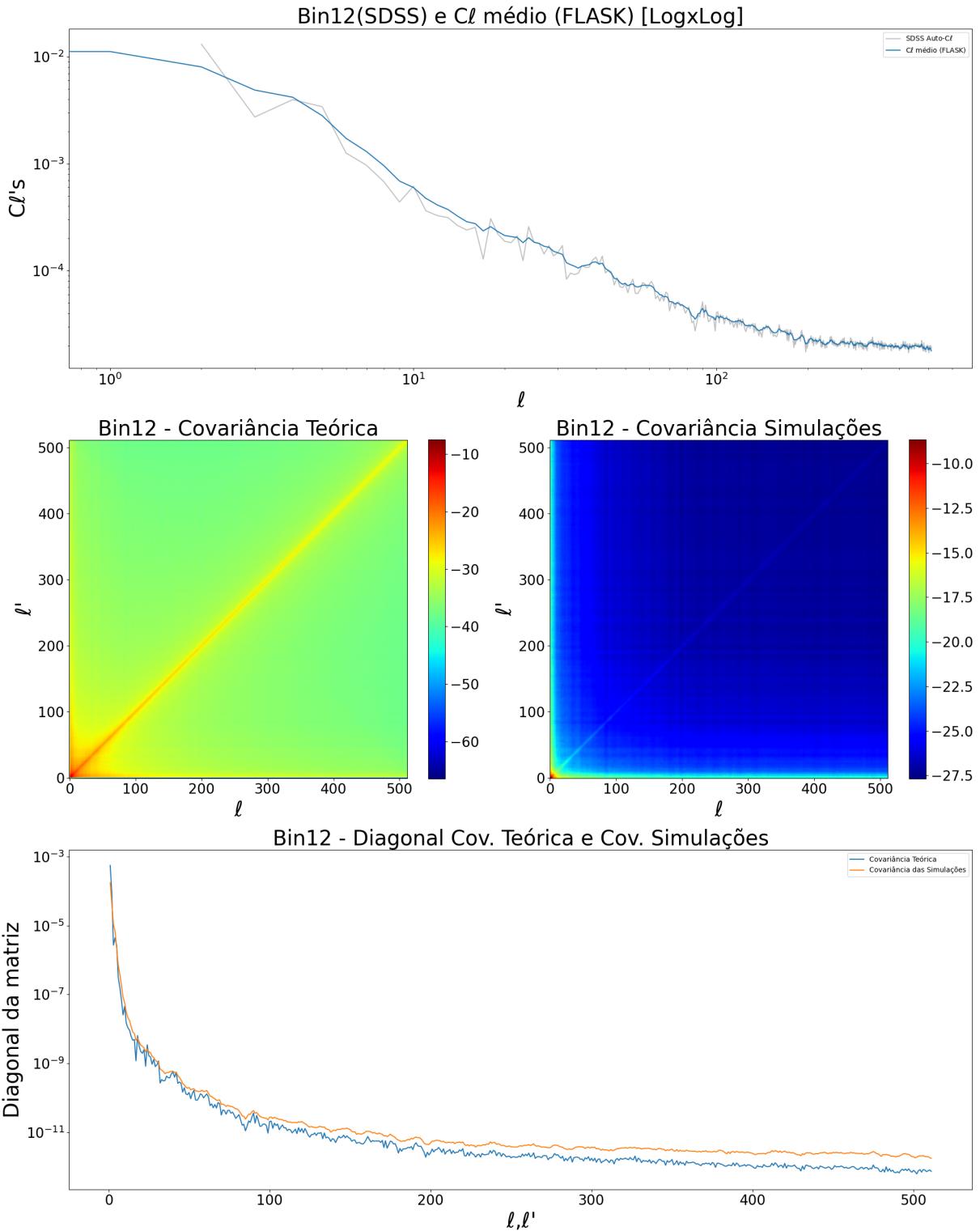


Figura 113: No primeiro gráfico há uma comparação entre o  $C\ell$  médio das simulações do FLASK e o  $C\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

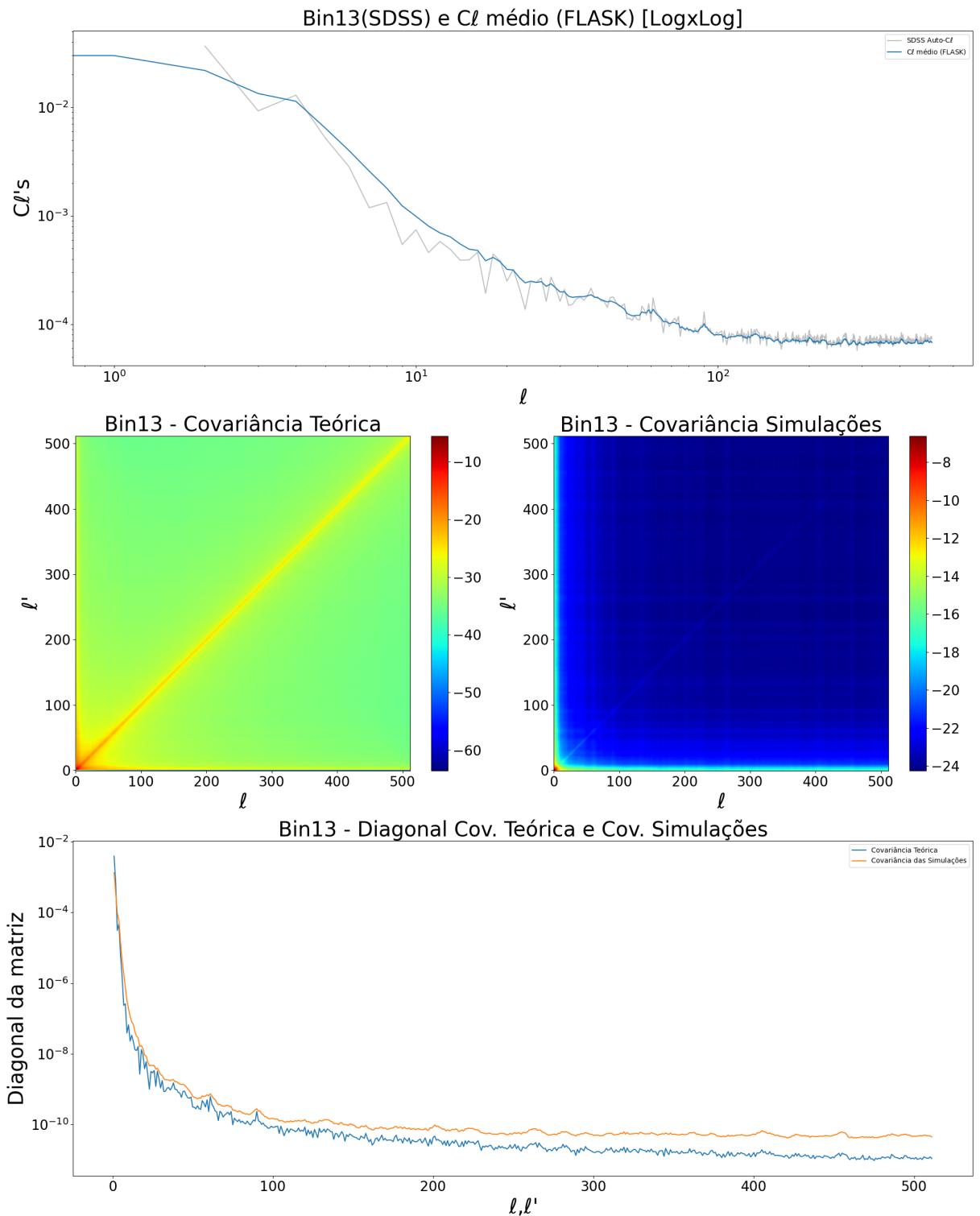


Figura 114: No primeiro gráfico há uma comparação entre o  $C\ell$  médio das simulações do FLASK e o  $C\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

## 13 Anexo VI - Matrizes de covariância: $S\ell$

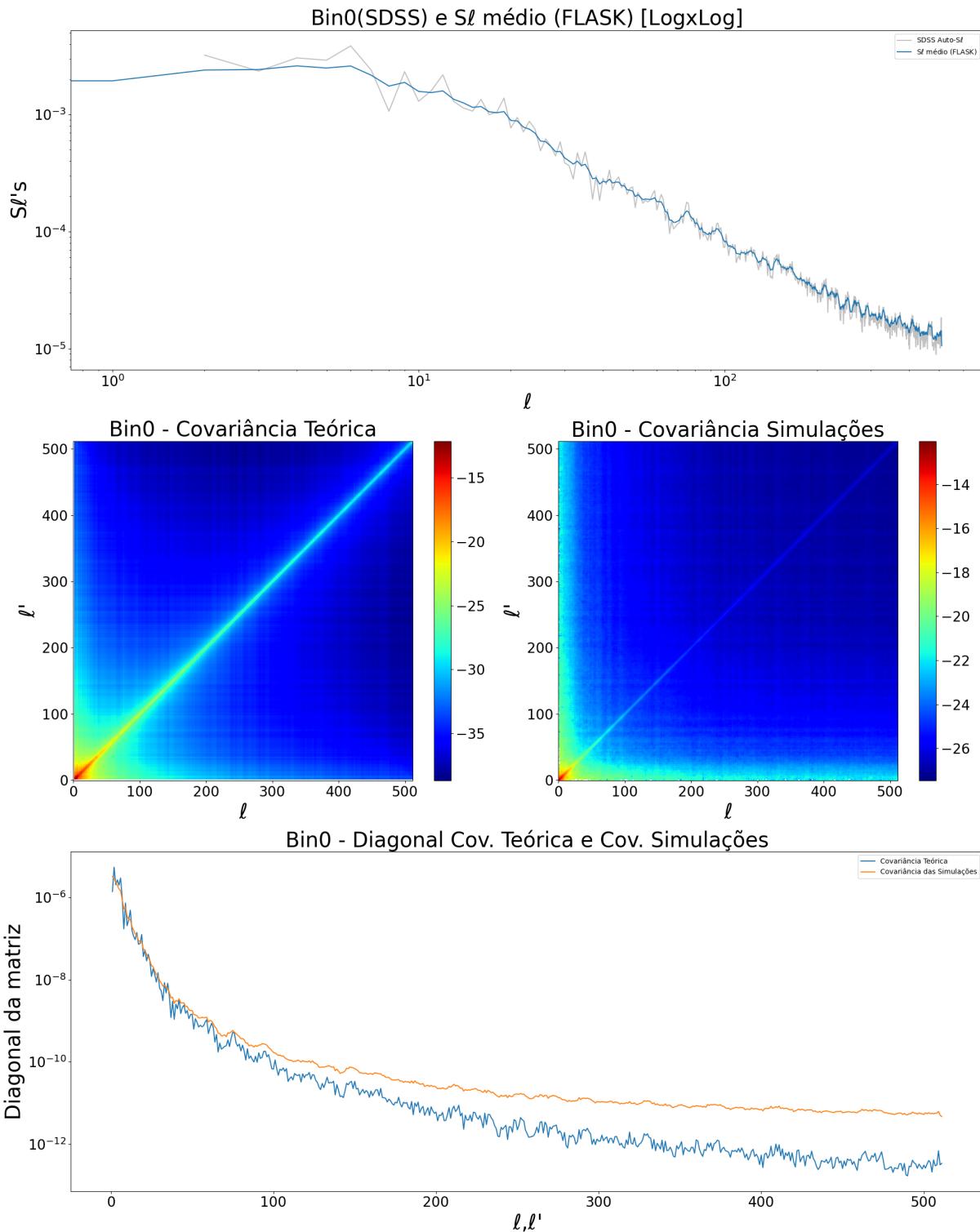


Figura 115: No primeiro gráfico há uma comparação entre o  $S\ell$  médio das simulações do FLASK e o  $S\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

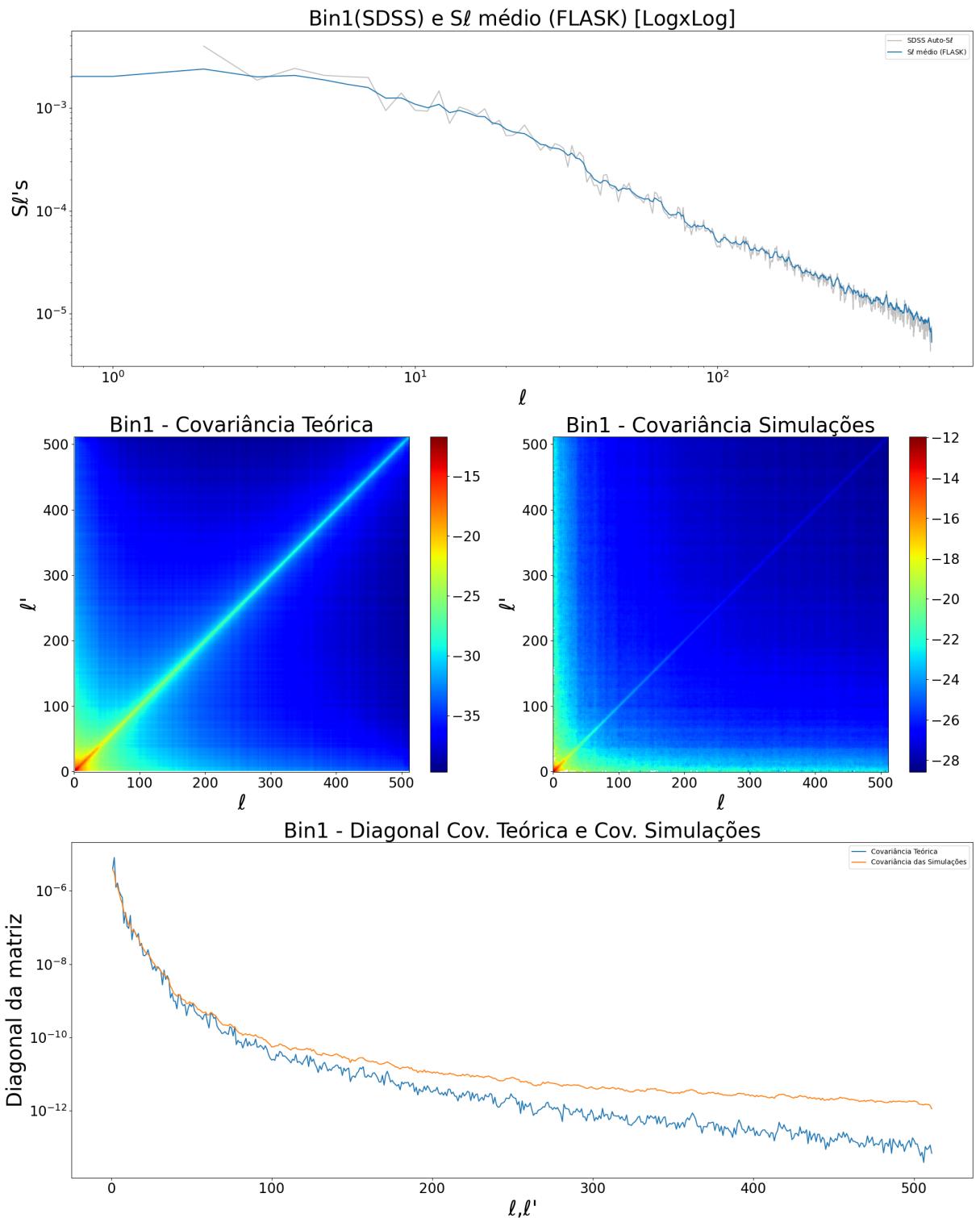


Figura 116: No primeiro gráfico há uma comparação entre o  $S\ell$  médio das simulações do FLASK e o  $S\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

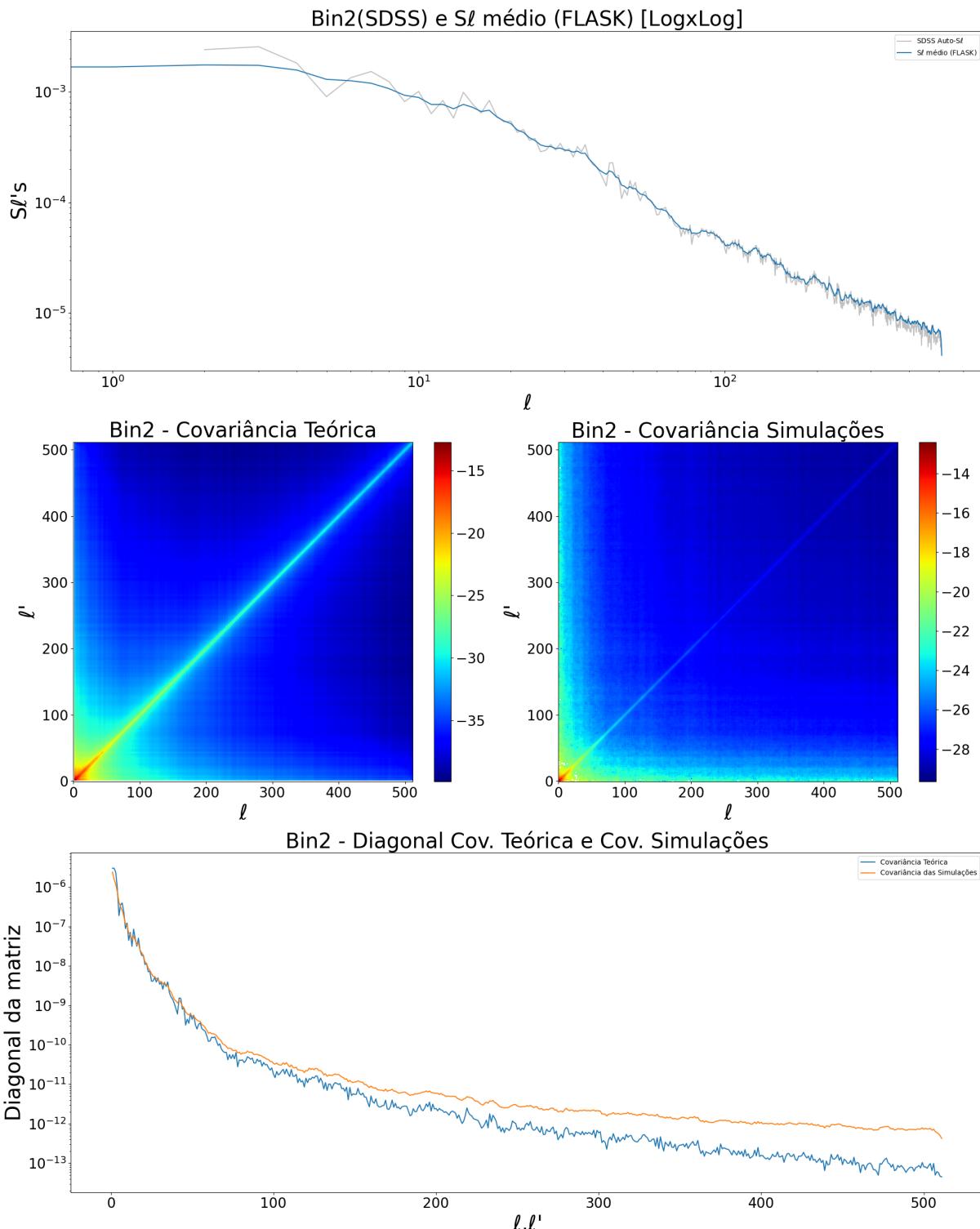


Figura 117: No primeiro gráfico há uma comparação entre o  $S\ell$  médio das simulações do FLASK e o  $S\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

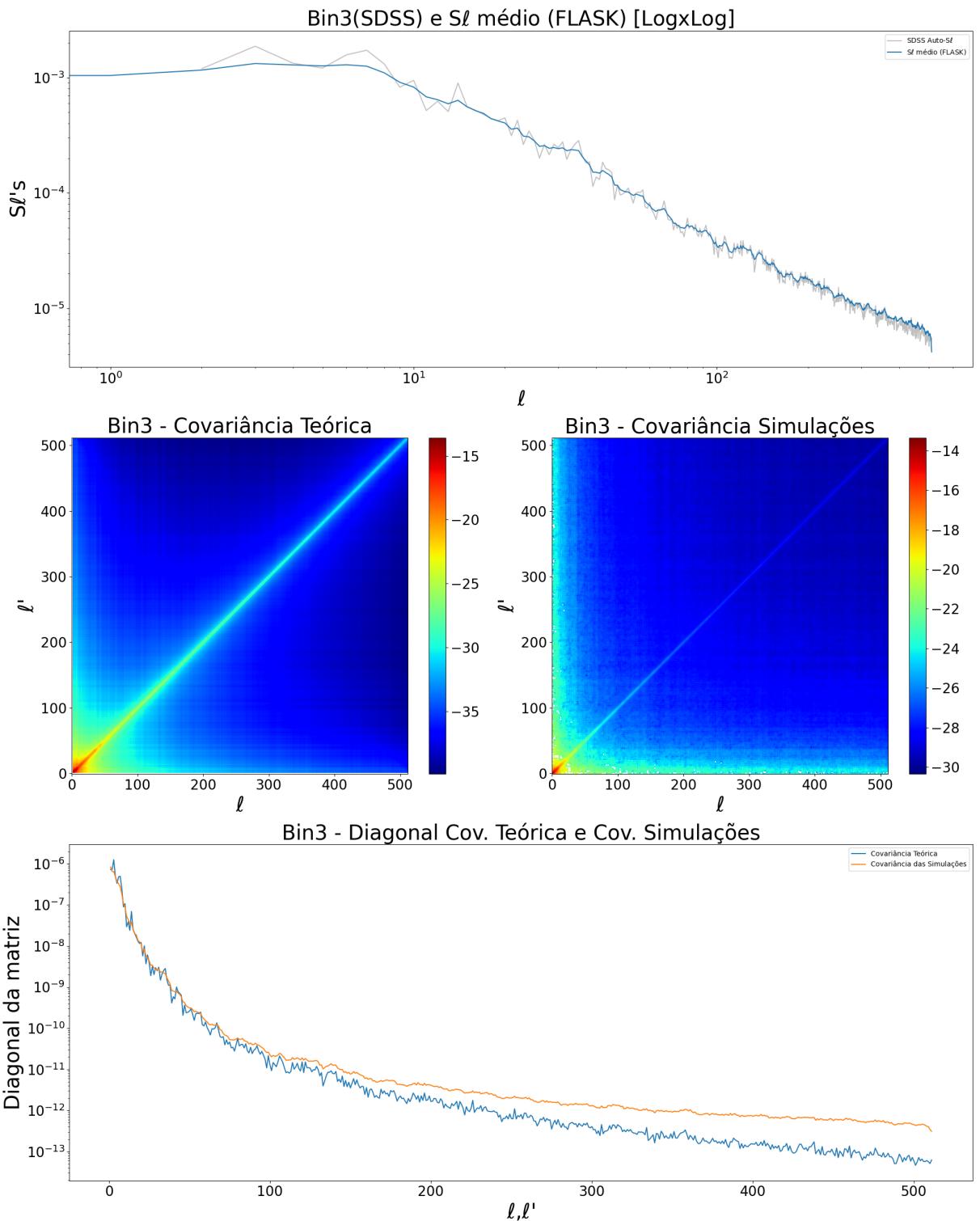


Figura 118: No primeiro gráfico há uma comparação entre o  $S_l$  médio das simulações do FLASK e o  $S_l$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

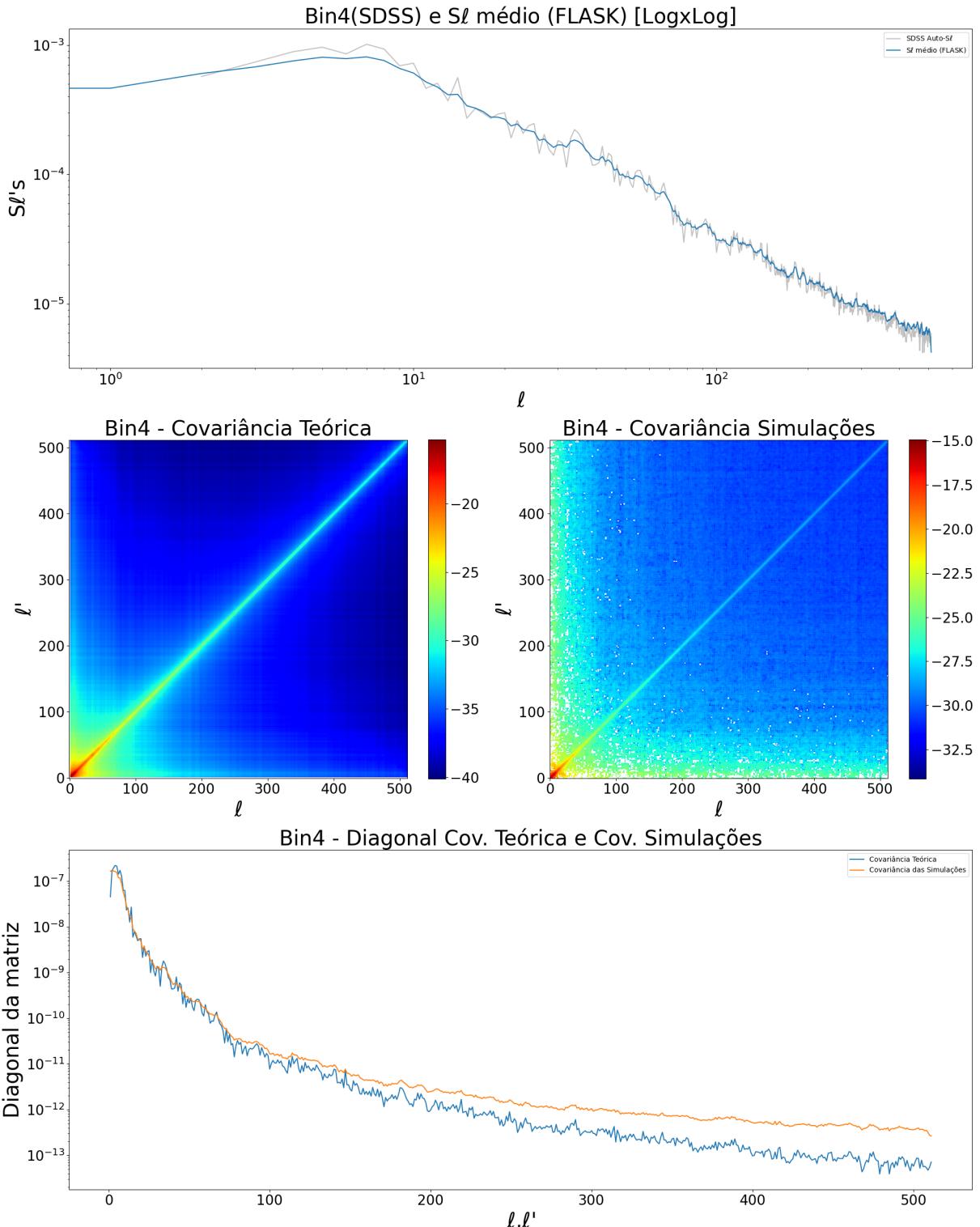


Figura 119: No primeiro gráfico há uma comparação entre o  $S\ell$  médio das simulações do FLASK e o  $S\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

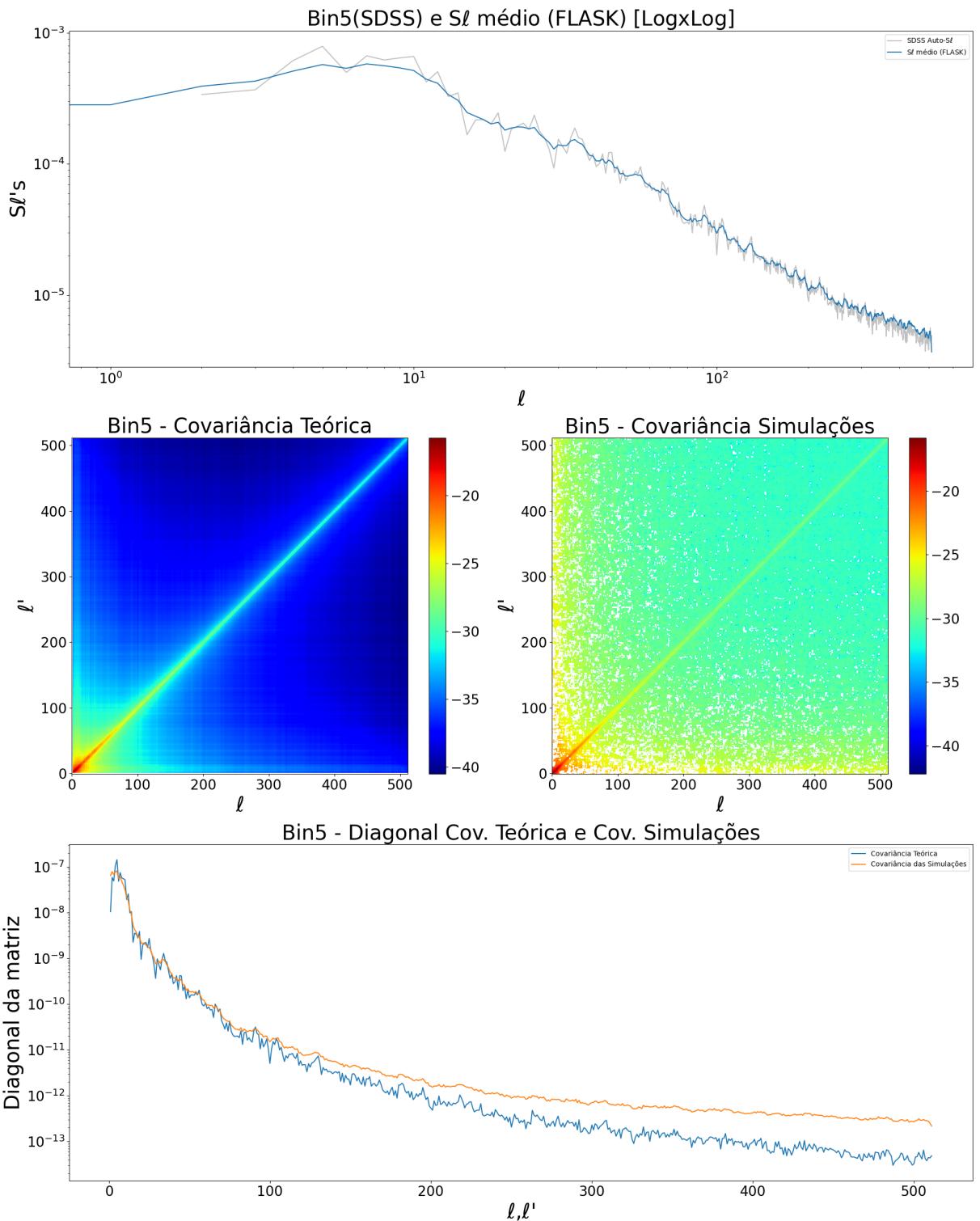


Figura 120: No primeiro gráfico há uma comparação entre o  $S\ell$  médio das simulações do FLASK e o  $S\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

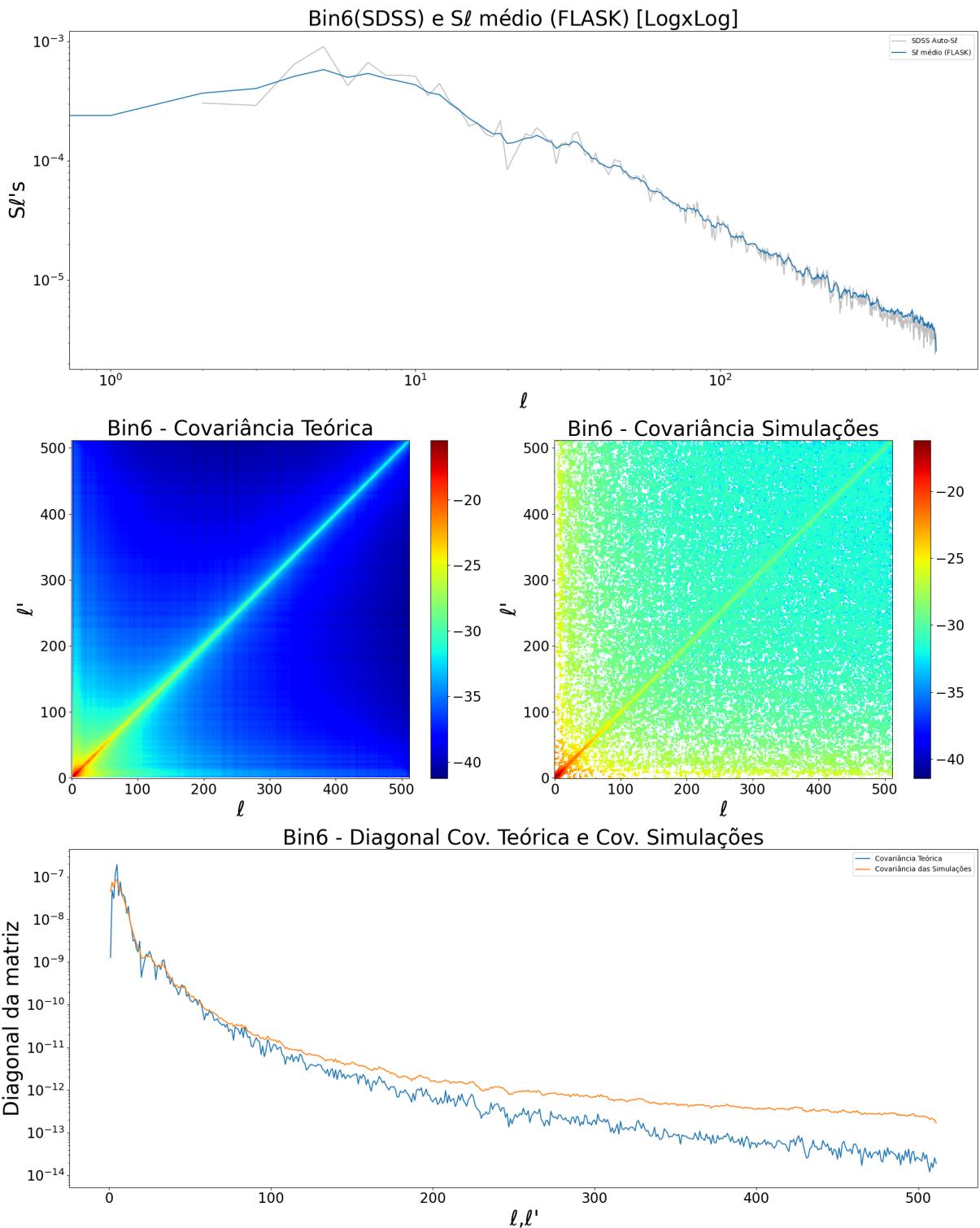


Figura 121: No primeiro gráfico há uma comparação entre o  $S\ell$  médio das simulações do FLASK e o  $S\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

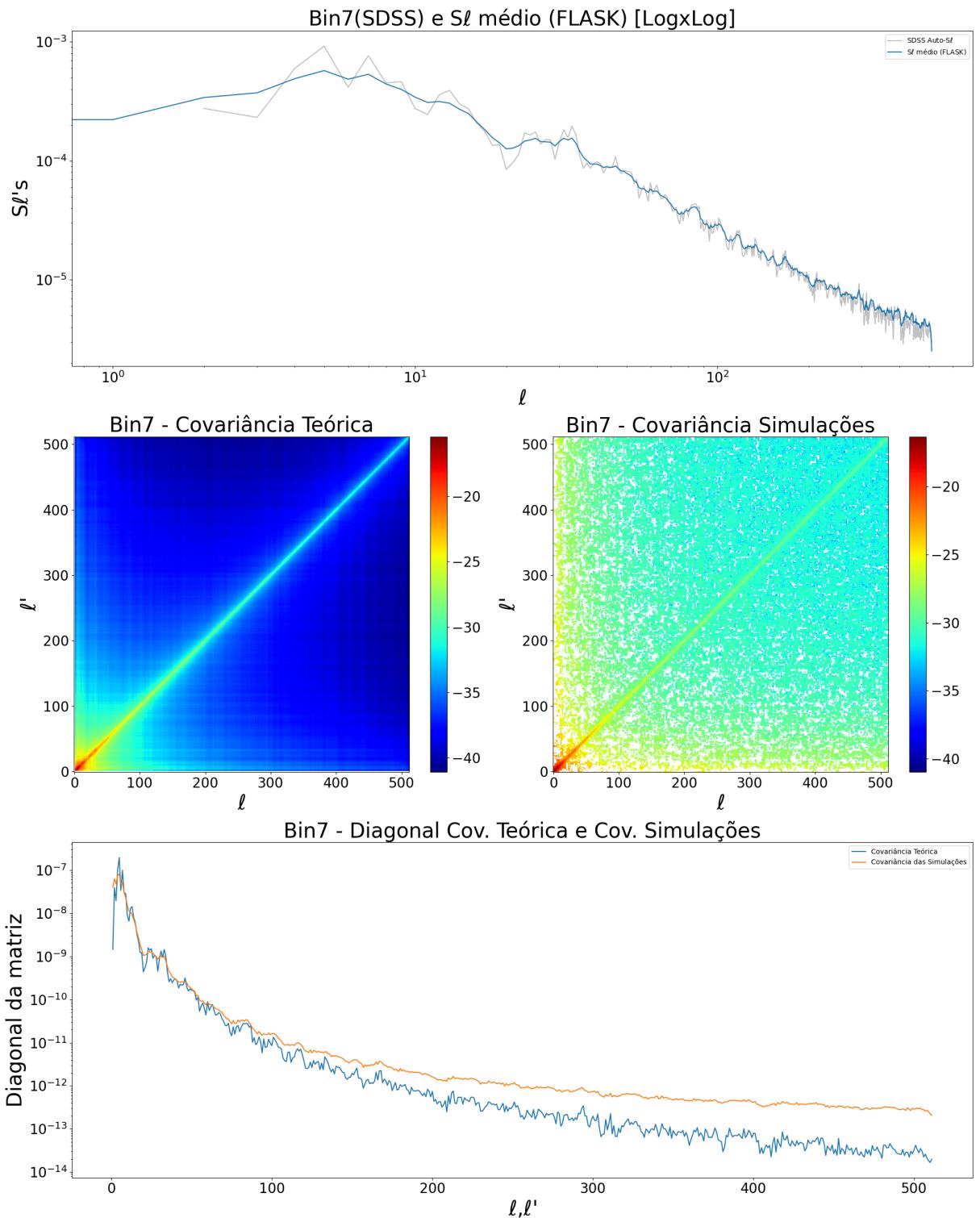


Figura 122: No primeiro gráfico há uma comparação entre o  $S\ell$  médio das simulações do FLASK e o  $S\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

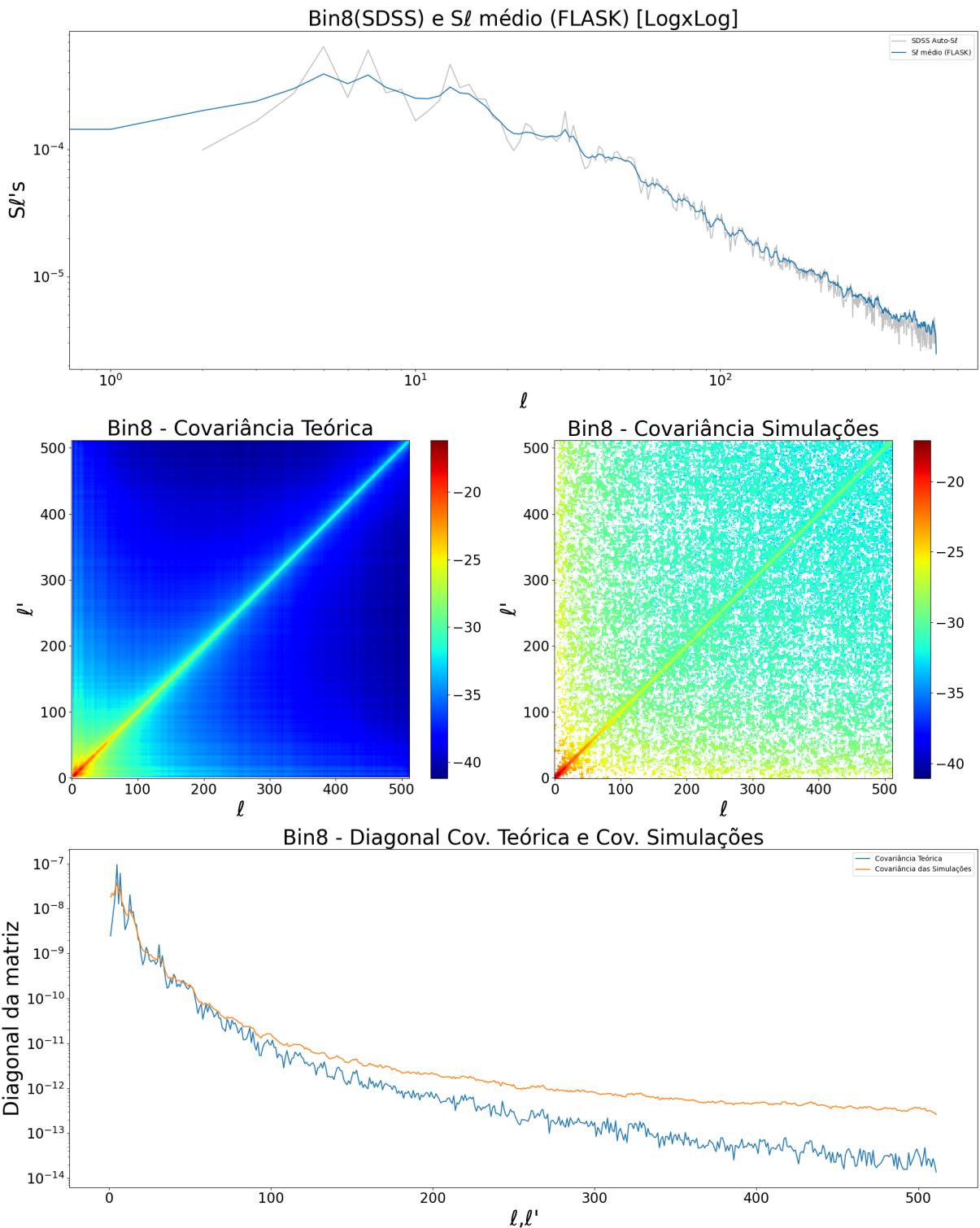


Figura 123: No primeiro gráfico há uma comparação entre o  $S\ell$  médio das simulações do FLASK e o  $S\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

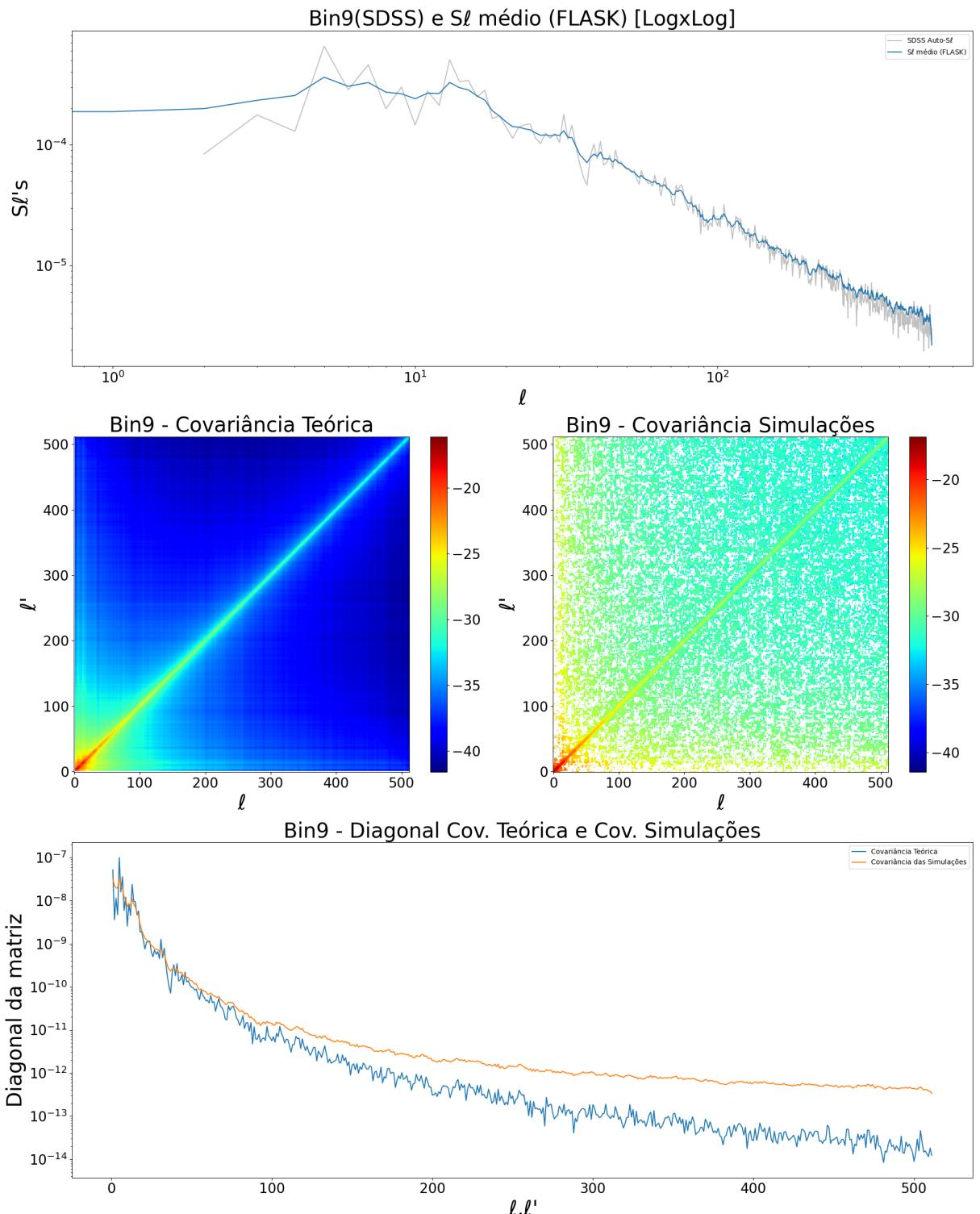


Figura 124: No primeiro gráfico há uma comparação entre o  $S_l$  médio das simulações do FLASK e o  $S_l$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

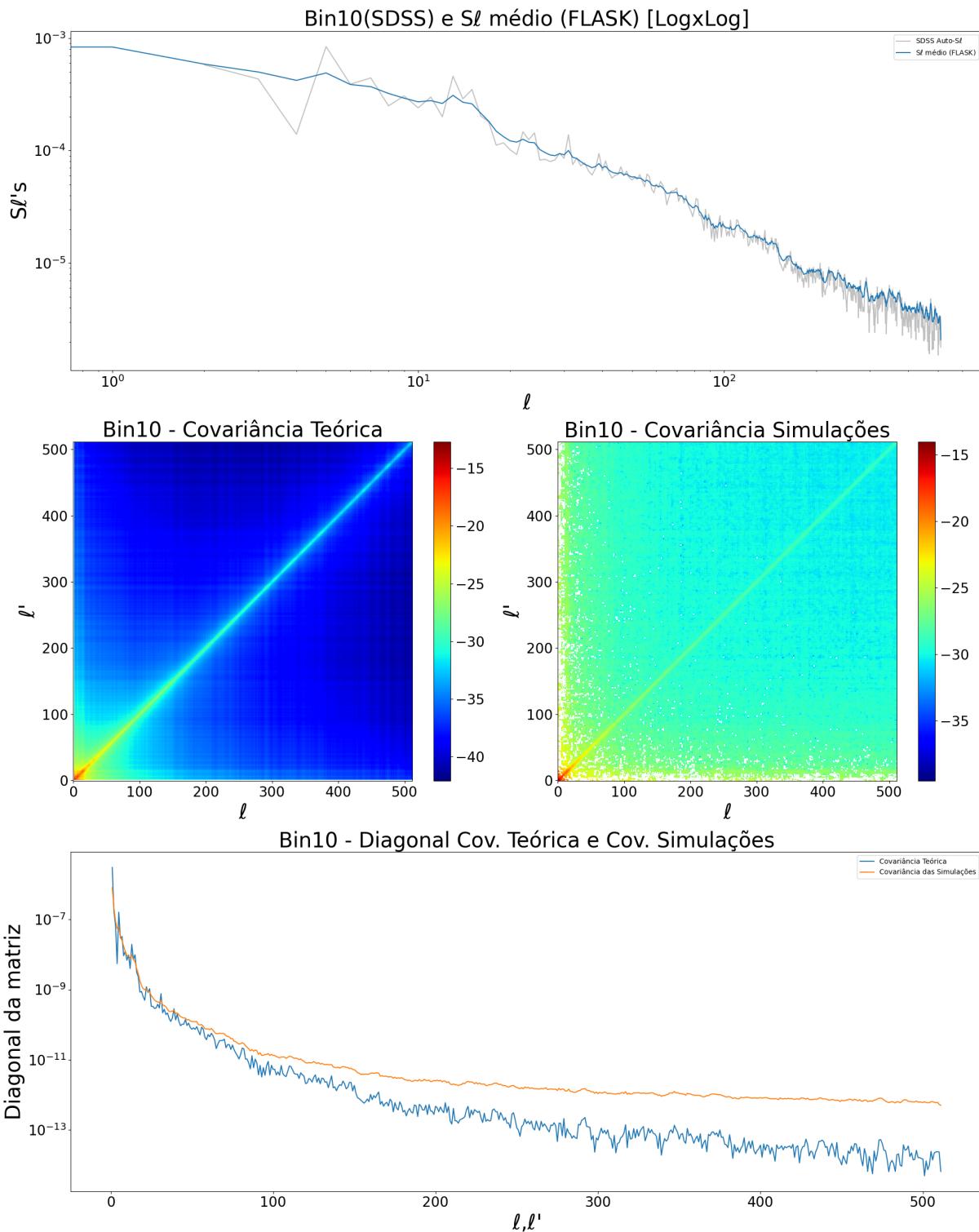


Figura 125: No primeiro gráfico há uma comparação entre o  $S\ell$  médio das simulações do FLASK e o  $S\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

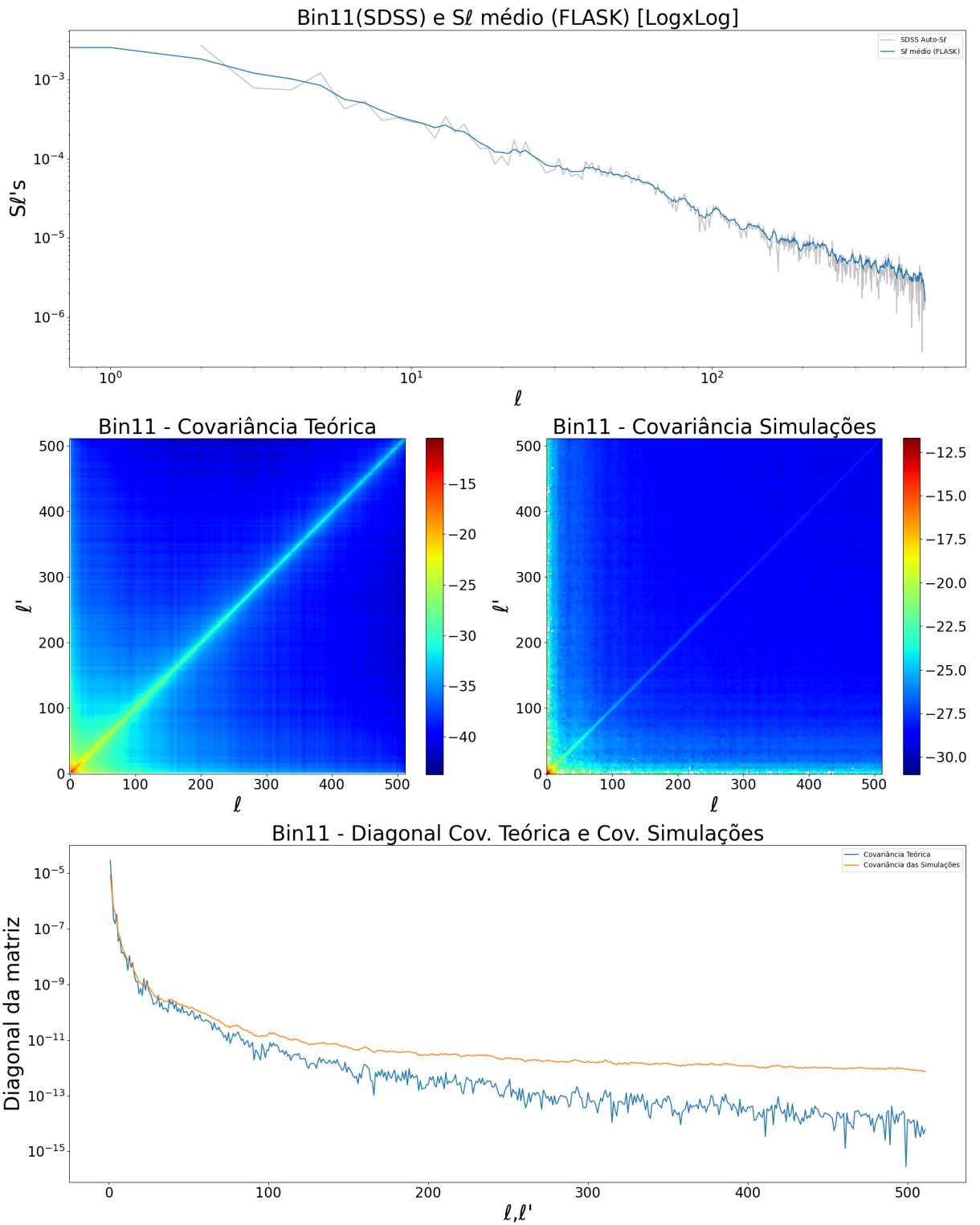


Figura 126: No primeiro gráfico há uma comparação entre o  $S\ell$  médio das simulações do FLASK e o  $S\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

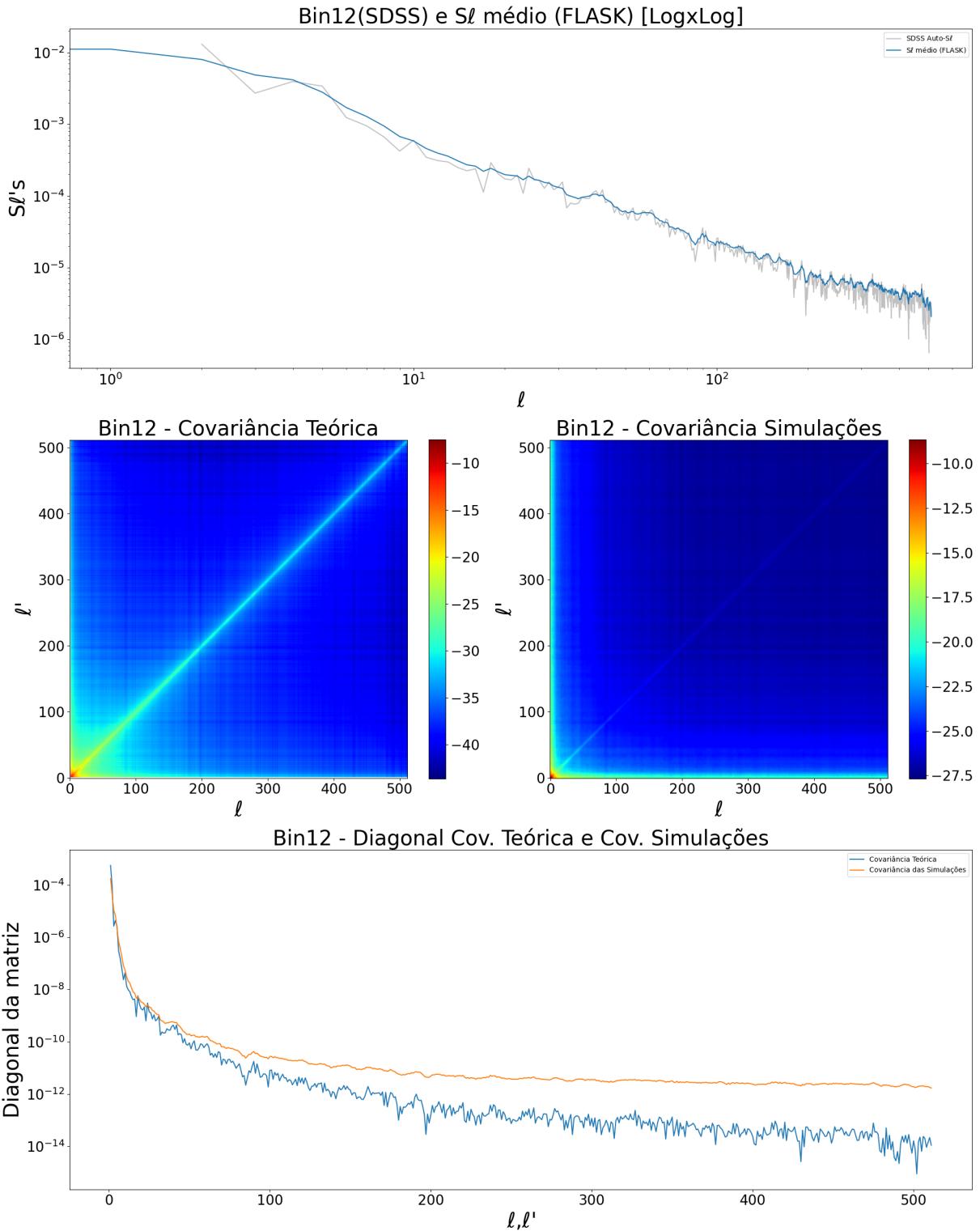


Figura 127: No primeiro gráfico há uma comparação entre o  $S_l$  médio das simulações do FLASK e o  $S_l$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.

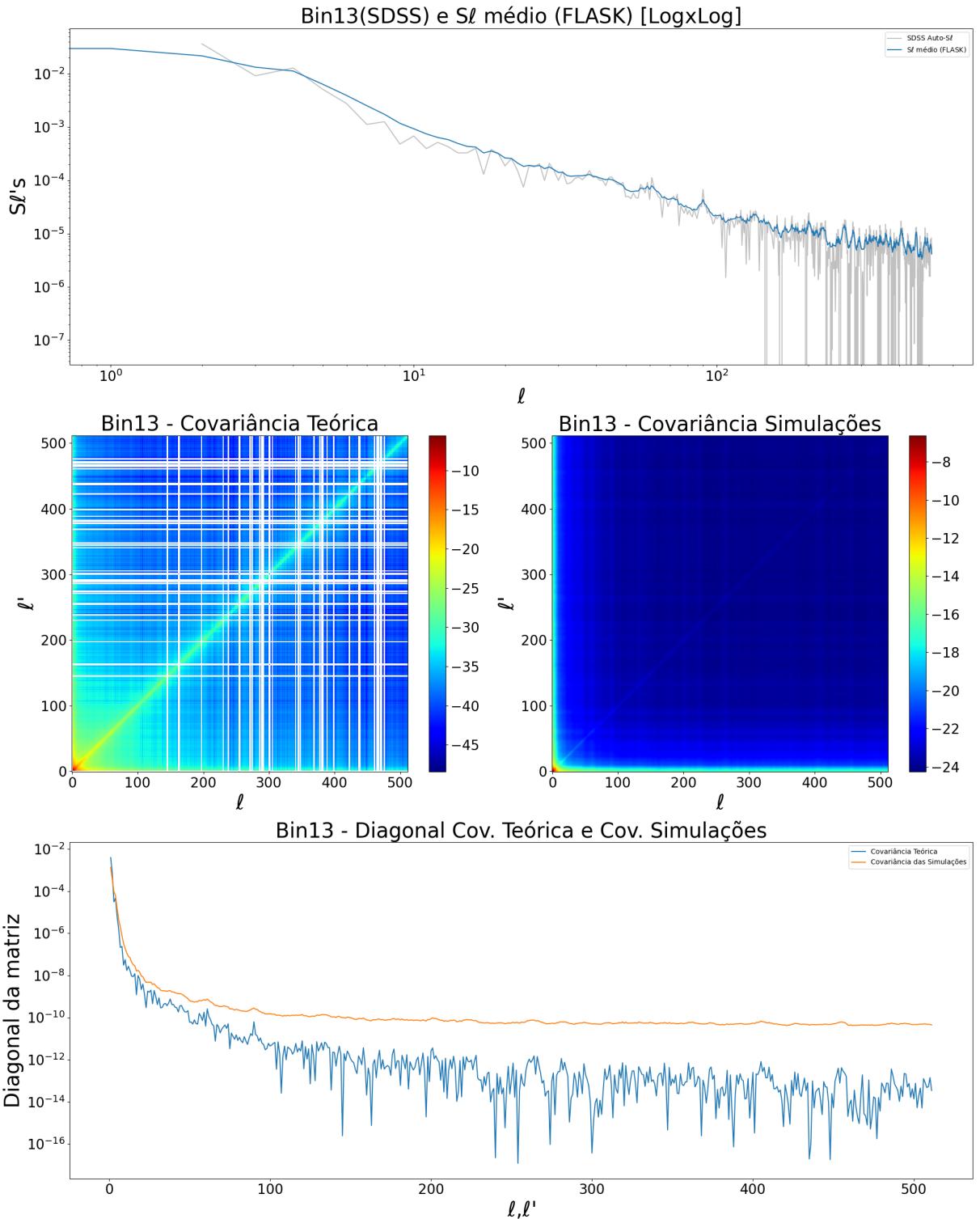


Figura 128: No primeiro gráfico há uma comparação entre o  $S\ell$  médio das simulações do FLASK e o  $S\ell$  dos mapas cosmológicos. No segundo mapa uma comparação entre a matriz de covariância extraída diretamente dos dados (à esquerda) e a matriz de covariância extraída das simulações (à direita). No ultimo gráfico temos a comparação das diagonais das duas matrizes.