

Doctoral Dissertation

博士論文

Effects of Replicated N-body Simulation Boxes in Simulating Weak Lensing Observables

弱重力レンズ統計量の模擬データ生成における
有限体積のN体シミュレーションの影響の研究

A Dissertation Submitted for the Degree of Doctor of Philosophy

December 2024

令和 6 年 1 月 博士（理学）申請

Department of Physics, Graduate School of Science,

The University of Tokyo

東京大学大学院理学系研究科

物理学専攻

Akira Tokiwa

常盤 晟

Contents

| | |
|--|-----------|
| 1 Weak Lensing | 10 |
| 1.1 Lens Equation | 10 |
| 1.1.1 Derivation | 10 |
| 1.1.2 Lensing Matrix | 12 |
| 1.2 Convergence | 13 |
| 1.2.1 Definition | 13 |
| 1.2.2 Convergence from Density Contrast | 14 |
| 1.3 Shear | 14 |
| 1.3.1 Definition | 15 |
| 1.3.2 E-mode and B-mode | 15 |
| 2 Power Spectrum | 16 |
| 2.1 Overview of Power Spectrum | 16 |
| 2.2 Two-Point Correlation Function | 17 |
| 2.3 Power Spectrum | 18 |
| 2.3.1 Matter Power Spectrum | 18 |
| 2.3.2 Convergence Power Spectrum | 18 |
| 2.4 Estimator | 20 |
| 3 Higher-Order Statistics | 21 |
| 3.1 Bispectrum | 21 |
| 3.2 Probability Density Functions | 22 |
| 3.2.1 Definition | 22 |
| 3.2.2 Normalization | 23 |
| 3.2.3 Moments and Cumulants | 23 |
| 3.3 Peak and Minimum Counts | 24 |
| 3.3.1 Identification of Peaks and Minima | 24 |
| 3.3.2 Peaks of Gaussian Random Fields | 25 |
| 3.4 Minkowski Functionals | 26 |

| | | |
|----------|---|-----------|
| 3.4.1 | Definition | 26 |
| 3.4.2 | Computation of Minkowski Functionals | 27 |
| 3.4.3 | Minkowski Functionals in Gaussian Random Fields | 27 |
| 4 | Covariance | 29 |
| 4.1 | Covariance of Matter Power Spectrum | 29 |
| 4.2 | Covariance of Angular Power Spectrum | 31 |
| 4.3 | Covariance of Higher-Order Statistics | 32 |
| 5 | Surveys(Make it shorter) | 34 |
| 5.1 | Introduction to Astronomical Surveys | 34 |
| 5.1.1 | Purpose and Scientific Objectives | 34 |
| 5.1.2 | Imaging vs. Spectroscopic Surveys | 35 |
| 5.1.3 | Ground-Based vs. Space-Based Surveys | 36 |
| 5.1.4 | Stage-III vs. Stage-IV Surveys | 36 |
| 5.2 | Subaru Hyper Suprime-Cam | 37 |
| 5.2.1 | Instrument Overview | 37 |
| 5.2.2 | Survey Design: The Hyper Suprime-Cam Subaru Strategic Program | 38 |
| 5.2.3 | Data Releases and Processing | 38 |
| 5.3 | Dark Energy Survey | 39 |
| 5.3.1 | Overview and Objectives | 39 |
| 5.3.2 | Survey Components | 39 |
| 5.3.3 | Instrumentation and Data Processing | 40 |
| 5.3.4 | Scientific Contributions | 40 |
| 5.4 | Legacy Survey of Space and Time | 40 |
| 5.4.1 | Overview and Scientific Objectives | 40 |
| 5.4.2 | Instrumentation and Survey Strategy | 41 |
| 5.4.3 | Scientific Impact | 41 |
| 5.5 | Nancy Grace Roman Space Telescope | 42 |
| 5.5.1 | Overview and Mission Objectives | 42 |
| 5.5.2 | Surveys and Scientific Goals | 42 |
| 5.5.3 | Instrumentation and Capabilities | 43 |
| 5.5.4 | Expected Scientific Contributions | 43 |
| 6 | Simulations | 45 |
| 6.1 | Initial condition | 45 |
| 6.1.1 | Primordial Power Spectrum | 45 |

| | | |
|----------|--|-----------|
| 6.1.2 | Initial Conditions | 46 |
| 6.2 | Simulation Basics | 47 |
| 6.2.1 | Direct Summation | 48 |
| 6.2.2 | Particle-Mesh (PM) Method | 49 |
| 6.2.3 | Particle-Particle Particle-Mesh (P3M) Method | 49 |
| 6.2.4 | Tree-Particle-Mesh (Tree-PM) Method | 50 |
| 6.3 | Tools for Fast Computation | 51 |
| 6.3.1 | Fast Fourier Transform | 51 |
| 6.3.2 | Mass Assignment Schemes | 52 |
| 6.3.3 | Parallelization Techniques | 53 |
| 6.3.4 | Adaptive Mesh Refinement | 53 |
| 6.3.5 | Tree Construction | 54 |
| 6.4 | FASTPM | 55 |
| 6.4.1 | Modified Kick and Drift Factors | 55 |
| 6.4.2 | Algorithm Steps | 57 |
| 7 | Methods | 59 |
| 7.1 | Dataset Overview | 59 |
| 7.2 | Generating Convergence Maps | 60 |
| 7.3 | Incorporating Noise | 62 |
| 7.4 | Patch Extraction for Analysis | 63 |
| 7.5 | Gaussian Smoothing | 64 |
| 7.6 | Measurements | 65 |
| 7.6.1 | Statistical Measures and Computational Methods | 65 |
| 7.6.2 | Covariance Matrix Estimation | 66 |
| 8 | Results | 67 |
| 8.1 | Overview | 67 |
| 8.2 | Effects of Noise | 68 |
| 8.3 | Effects of Smoothing Scale | 68 |
| 9 | Discussion | 78 |
| 9.1 | Possible Effects | 78 |
| 9.1.1 | Finite Support Effects | 78 |
| 9.1.2 | Box Replication Effect | 78 |
| 9.2 | Validation | 78 |
| 9.3 | Check if the gnomview matters | 78 |

| | |
|--|----|
| 9.4 Why higher-order statistics are less affected? | 78 |
| 9.5 correlation between different smoothing scales | 78 |

List of Figures

| | | |
|-----|---|----|
| 1.1 | Schematic representation of the lensing geometry. The source is located at β , while the observed image is at θ . The deflection angle α is the difference between the observed and true angular positions. | 12 |
| 1.2 | Illustration of the distortion of background sources due to gravitational lensing. The left panel depict the effect of the convergence κ and the middle and right panels show the components of the shear $\gamma = \gamma_1 + i\gamma_2$ on circular background sources. Positive and negative values of κ cause isotropic magnification or demagnification, while γ_1 and γ_2 introduce anisotropic distortions, stretching the sources along or at an angle to the principal axes. | 13 |
| 3.1 | Identification of peaks and minima in a smoothed convergence map. The left panel shows the smoothed convergence field $\kappa_{\text{smooth}}(\hat{\mathbf{n}})$ with peaks (red circles) satisfying the peak condition (Equation (3.19)), and the right panel highlights the minima (blue circles) satisfying the minimum condition (Equation (3.20)). | 25 |
| 3.2 | Visualization of excursion sets $\Sigma(\nu) = \{\kappa > \nu\sigma_0\}$ of a two-dimensional random field $\kappa(\hat{\mathbf{n}})$, where κ represents the convergence field with zero mean and variance σ_0^2 . Black regions correspond to regions where the field exceeds the threshold value $\nu\sigma_0$. The panels correspond to increasing threshold values ($\nu = 0.5, 1, 1.5, 2$), illustrating how the size and connectivity of the excursion sets change as the threshold increases. | 26 |
| 6.1 | Evolution of the number of particles used in N -body simulations as a function of the year of publication (Leclercq, 2020). The symbols and colors indicate the gravitational solver employed: P^3M and adaptive P^3M (AP^3M); parallel or vectorized P^3M ; Tree codes; TreePM; and particle-mesh methods with adaptive mesh refinement (PM AMR). Hydrodynamic simulations are represented by black squares. | 48 |
| 6.2 | Illustration of three mass assignment schemes—Nearest Grid Point (NGP), Cloud-In-Cell (CIC), and Triangular-Shaped Cloud (TSC)—used to map a particle’s mass onto a 1D grid. | 52 |

- 6.3 Illustration of adaptive mesh refinement (AMR) applied to a 2D image with two Gaussian kernels. The left panel shows the initial coarse grid structure over the image. The middle and right panels demonstrate progressively finer levels of mesh refinement in regions of higher intensity, where the Gaussian kernels are located. The red grid outlines indicate the adaptively refined mesh hierarchy, ensuring higher resolution where needed while maintaining computational efficiency in lower-intensity regions. 54
- 6.4 Illustration of an Octree decomposition for a 3D volume containing four particles. The left panel showcases the spatial subdivision of the volume into hierarchical grid cells, with color-coding indicating different levels of refinement—red for coarse, blue for intermediate, and orange for fine cells where particles reside. The right panel presents the corresponding Octree data structure, highlighting the hierarchical relationships between nodes. The root node (purple) represents the entire simulation volume, while the second layer comprises various node types: empty nodes are depicted in white, single-particle nodes in orange-pink with a black dot, and multi-particle nodes in orange-pink without a dot, which require further subdivision. This hierarchical decomposition enables efficient computation by concentrating resolution in regions with higher particle density while conserving computational resources in less populated areas. Credit by Powell et al. (2023) 55
- 7.1 Spatial and redshift setup for the **BIGBOX** and **TILED** simulations. The left side of the figure features red lines delineating the light cone boundaries, covering a 10° field of view on the sky. Overlaid on this are dashed blue grids that partition the overall simulation volume into smaller, manageable tiling regions. Within these grids, the "small box" highlights a specific tiled region that resides inside the extensive BIGBOX volume. The horizontal axis represents the line-of-sight distance measured in comoving megaparsecs per h , with the corresponding redshift values displayed on the top axis. 60
- 7.2 Normalized lensing efficiency as a function of comoving distance for multiple source redshifts. The lensing efficiency peaks at intermediate comoving distances, indicating regions where the distribution of matter enhances the gravitational lensing signal. 62
- 7.3 Visualization of the Fibonacci grid with $N_{\text{patches}} = 273$ patches, each covering approximately $10 \times 10 \text{ deg}^2$. After the optimization and masking, the number of patches is reduced to $N_{\text{patches}} = 262$, effectively covering 64% of the sky. Each panel show the patches distribution on the Front, Top and Side view. 63

| | |
|---|----|
| 7.4 Effect of Gaussian smoothing on a noisy convergence map. Each panel shows the result of applying a Gaussian filter with a different smoothing scale $\theta_G = 2', 5', 8', \text{ and } 10'$. As the smoothing scale increases, small-scale noise is progressively suppressed, and large-scale structures become more prominent. This demonstrates how Gaussian smoothing effectively reduces shape noise while enhancing the detection of the underlying lensing signal. | 65 |
| 8.1 Comparison of the mean values of the angular power spectrum ($C_\ell^{\kappa\kappa}$) for different source redshifts ($z_s = 0.5, 1.0, 1.5, 2.0, 2.5$) obtained from the BIGBOX (solid lines) and TILED (dashed lines) simulations. The lower subplots show the ratio of the TILED to BIGBOX mean values, with a reference line at unity to facilitate the assessment of agreement between the two simulations. | 69 |
| 8.2 Comparison of the covariance matrices and correlation matrices of the angular power spectrum ($C_\ell^{\kappa\kappa}$) between the BIGBOX and TILED simulations for various source redshifts ($z_s = 0.5, 1.0, 1.5, 2.0, 2.5$). The displayed ratios represent the element-wise division of the covariance and correlation matrices from the TILED simulations by those from the BIGBOX simulations. The "avg" denotes the average ratio of the considered matrix elements. | 69 |
| 8.3 Same as Figure 8.1, but for the bispectrum. | 70 |
| 8.4 Similar to Figure 8.2, but for the covariance matrices of the bispectrum. The noisy nature of the bispectrum covariance makes it challenging to discern clear trends between the simulations. | 70 |
| 8.5 Same as Figure 8.1, but for the probability density function (PDF) of the convergence field. The comparison highlights the agreement in mean PDF values between the simulations across different redshifts. | 71 |
| 8.6 Same as Figure 8.2, but for the covariance matrices of the PDF. The covariance ratios indicate higher covariance in the BIGBOX simulations, particularly at higher redshifts. | 71 |
| 8.7 Same as Figure 8.1, but for peak counts in the convergence maps. The analysis reveals deviations at low ν values due to resolution limitations affecting low-density regions. | 72 |
| 8.8 Same as Figure 8.2, but for the covariance matrices of peak counts. The covariance ratios suggest increased covariance in the BIGBOX simulations, with pronounced effects at higher redshifts. | 72 |
| 8.9 Same as Figure 8.1, but for minima in the convergence maps. The comparison underscores the simulation's limitations at resolving low-density minima accurately. | 73 |
| 8.10 Same as Figure 8.2, but for the covariance matrices of minima. The covariance ratios reflect higher values in the BIGBOX simulations, consistent with other statistical measures. | 73 |

| | |
|---|----|
| 8.11 Same as Figure 8.1, but for Minkowski Functionals (area V_0 , perimeter V_1 , and genus V_2). The agreement in mean values between simulations is generally good, with some discrepancies at extreme density thresholds. | 74 |
| 8.12 Similar to Figure 8.2, but for the covariance matrices of Minkowski Functionals. | 74 |
| 8.13 Average ratio of covariance matrices of statistical measures between the BIGBOX and TILED simulations for different shape noise levels (see Table 7.2). The increasing trend indicates does not affected by the noise level. | 75 |
| 8.14 Same as Figure 8.13, but for the correlation matrices. The off-diagonal elements compared to the diagonal elements do not show a clear trend with noise levels. | 75 |
| 8.15 Average ratio of covariance matrices of statistical measures between the BIGBOX and TILED simulations for different smoothing scales. Larger smoothing scales lead to increased discrepancies in covariance estimates due to the loss of small-scale information. | 75 |
| 8.16 Same as Figure 8.15, but for the correlation matrices. The instability at larger smoothing scales reflects the challenges in capturing correlations at reduced resolutions. | 75 |
| 8.17 Ratio of covariance matrices of the angular power spectrum ($C_\ell^{\kappa\kappa}$) between the BIGBOX and TILED simulations for different shape noise levels (see Table 7.2). The sensitivity of the power spectrum to noise is evident from the fluctuating covariance ratios with higher noise levels. | 76 |
| 8.18 Same as Figure 8.17, but for the non-Gaussian statistical measures. The robustness of these measures against noise variations is reflected in the relatively stable covariance ratios. | 76 |
| 8.19 Same as Figure 8.18, but showing the impact of different smoothing scales on the covariance matrices of non-Gaussian statistical measures. The results emphasize how increased smoothing affects the detection and characterization of small-scale features. | 77 |

List of Tables

| | | |
|-----|--|----|
| 7.1 | Cosmological parameters used in the N-body simulations. | 59 |
| 7.2 | Galaxy number densities for different surveys used to model shape noise levels. | 62 |
| 7.3 | Summary of the statistical measures employed in this investigation, including their respective value ranges and computational subroutines utilized for both full-sky and sky-patch analyses. | 65 |

Chapter 1

Weak Lensing

Weak gravitational lensing (WL), also known as cosmic shear, refers to the subtle distortions in the images of distant source caused by the gravitational fields of intervening mass distributions. Unlike strong lensing, which produces noticeable effects such as multiple images or arcs, weak lensing induces small, coherent distortions that require statistical analysis to detect and interpret. For the standard approach to lensing, we refer to [Schneider et al. \(1992\)](#), [Bartelmann & Schneider \(2001\)](#) and [Bartelmann \(2010\)](#).

1.1 Lens Equation

1.1.1 Derivation

To derive the lens equation, we consider a perturbed FLRW metric, which incorporates gravitational potential perturbations. The metric is expressed as

$$ds^2 = - \left(1 + \frac{2\Phi}{c^2}\right) c^2 dt^2 + a^2(t) \left(1 - \frac{2\Psi}{c^2}\right) [d\chi^2 + f_K^2(\chi) \omega_{ab} dx^a dx^b] \quad (a, b = 2, 3), \quad (1.1)$$

where the angular part of the metric is defined by

$$\omega_{ab} dx^a dx^b := d\theta^2 + \sin^2 \theta d\phi^2. \quad (1.2)$$

In this context, Φ and Ψ represent the scalar gravitational potentials, and $f_K(\chi)$ encodes the spatial curvature as previously defined in Eq. (??).

The trajectory of light within this spacetime is governed by the geodesic equation, which is given by

$$\frac{d^2 x^\mu}{d\lambda^2} + \Gamma_{\alpha\beta}^\mu \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} = 0, \quad (1.3)$$

where λ is an affine parameter and $\Gamma_{\alpha\beta}^\mu$ are the Christoffel symbols corresponding to the metric in Equation (1.1). To facilitate the derivation, we reparametrize the geodesic equation by substituting the affine parameter λ with the comoving radial distance χ . Applying the chain rule, the geodesic equation trans-

forms to

$$\frac{d^2x^\mu}{d\chi^2} + \Gamma_{\alpha\beta}^\mu \frac{dx^\alpha}{d\chi} \frac{dx^\beta}{d\chi} - \frac{d^2\lambda}{d\chi^2} \left(\frac{d\lambda}{d\chi} \right)^{-1} \frac{dx^\mu}{d\chi} = 0. \quad (1.4)$$

Setting $\mu = 1$ (where $x^1 = \chi$) in Equation (1.4) and simplifying, we obtain

$$\frac{d^2x^\mu}{d\chi^2} + (\Gamma_{\alpha\beta}^\mu - \Gamma_{\alpha\beta}^1) \frac{dx^\alpha}{d\chi} \frac{dx^\beta}{d\chi} = 0. \quad (1.5)$$

The evaluation of Equation (1.5) requires the computation of the Christoffel symbols. Additionally, the derivative $c dt/d\chi$ is derived from the null condition

$$g_{\mu\nu} \frac{dx^\mu}{d\chi} \frac{dx^\nu}{d\chi} = 0, \quad (1.6)$$

yielding

$$\frac{c dt}{d\chi} = -a(t) \left[1 - \frac{\Phi}{c^2} - \frac{\Psi}{c^2} + \frac{f_K^2(\chi)}{2} \omega_{ab} \frac{dx^a}{d\chi} \frac{dx^b}{d\chi} \right], \quad (1.7)$$

valid to first order in Φ and second order in $dx^a/d\chi$. Notably, for the evaluation of Equation (1.5), only the zeroth-order term $c dt/d\chi = -a(t)$ is required. However, the inclusion of perturbative terms in Equation (1.7) is essential for subsequent derivations of the lens equation.

Focusing on the angular components ($\mu = a$) of Equation (1.5), we derive the following differential equation:

$$\frac{d^2x^a}{d\chi^2} + 2 \frac{f'_K(\chi)}{f_K(\chi)} \frac{dx^a}{d\chi} + \omega^{ab} \frac{\Phi_b + \Psi_b}{c^2 f_K^2(\chi)} = 0, \quad (1.8)$$

where Φ_b and Ψ_b denote the derivatives of the gravitational potentials with respect to the angular coordinates x^a . Integrating Equation (1.8) twice with respect to χ , we obtain

$$\begin{aligned} x^a(\chi_s) - x^a(0) &= -\frac{1}{c^2} \int_0^{\chi_s} d\chi' \int_0^{\chi'} d\chi \omega_{ab} [\Phi_b(\chi, \theta(\chi)) + \Psi_b(\chi, \theta(\chi))] \\ &= -\frac{1}{c^2} \int_0^{\chi_s} d\chi \omega_{ab} [\Phi_b(\chi, \theta(\chi)) + \Psi_b(\chi, \theta(\chi))] \int_\chi^{\chi_s} d\chi' \frac{1}{f_K^2(\chi')} \\ &= -\frac{1}{c^2} \int_0^{\chi_s} d\chi \omega_{ab} [\Phi_b(\chi, \theta(\chi)) + \Psi_b(\chi, \theta(\chi))] \frac{f_K(\chi_s - \chi)}{f_K(\chi_s) f_K(\chi)}, \end{aligned} \quad (1.9)$$

where χ_s denotes the comoving radial distance to the source, and $x^a(0)$ is the angular position at the observer's location. Defining $(\nabla_\theta)^a = \omega^{ab} \partial_b$, the angular position of the source galaxy β is related to the observed angular position θ by

$$\theta(\chi_s) = \theta(0) - \frac{1}{c^2} \int_0^{\chi_s} d\chi \nabla_\theta [\Phi(\chi, \theta(\chi)) + \Psi(\chi, \theta(\chi))] \frac{f_K(\chi_s - \chi)}{f_K(\chi_s) f_K(\chi)}. \quad (1.10)$$

Applying the Born approximation (Born, 1926), and assume that $\Phi = \Psi$, we can simplify the lens

equation to:

$$\theta(\chi_s) = \theta(0) - \frac{2}{c^2} \int_0^{\chi_s} d\chi \nabla_\theta [\Phi(\chi, \theta(0))] \frac{f_K(\chi_s - \chi)}{f_K(\chi_s) f_K(\chi)}. \quad (1.11)$$

This final expression constitutes the lens equation, encapsulating the deflection of light due to the gravitational potentials Φ and Ψ along the line of sight.

1.1.2 Lensing Matrix

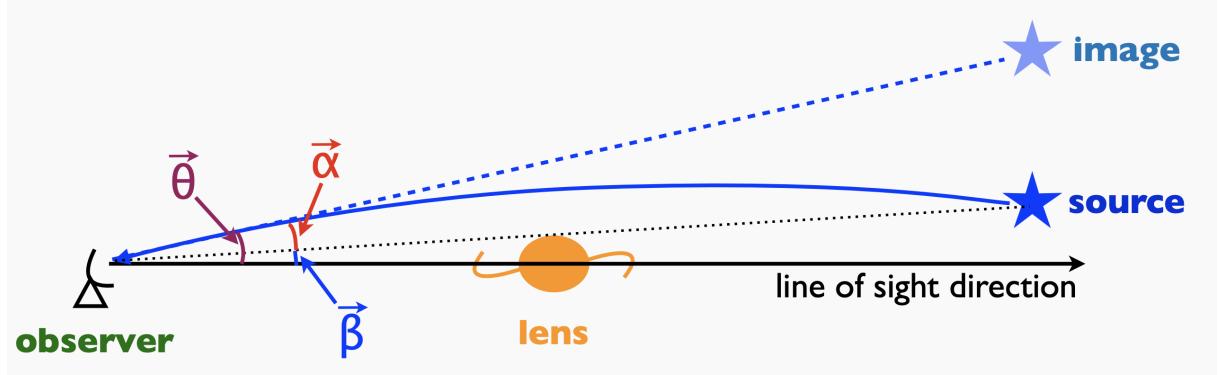


Figure 1.1: Schematic representation of the lensing geometry. The source is located at β , while the observed image is at θ . The deflection angle α is the difference between the observed and true angular positions.

Redefining the notation in Eq. (1.11) and considering the angular position of the source $\beta = \theta(\chi_s)$ and the observed angular position $\theta = \theta(0)$ (see Fig. 1.1), we can express the lens equation as (Bartelmann & Schneider, 2001; Hilbert et al., 2009; Kilbinger, 2015):

$$\beta = \theta - \alpha(\theta), \quad (1.12)$$

where the deflection angle $\alpha(\theta)$ is defined by:

$$\alpha(\theta) = \nabla_\theta \psi(\theta), \quad \psi(\theta) = \frac{2}{c^2} \int_0^{\chi_s} d\chi \frac{f_K(\chi_s - \chi)}{f_K(\chi_s) f_K(\chi)} \Phi(f_K(\chi)\theta, \chi). \quad (1.13)$$

The mapping between the source plane and the image plane can be described by the Jacobian matrix \mathcal{A} , which relates infinitesimal displacements in the source position to displacements in the image position:

$$\mathcal{A} := \frac{\partial \beta}{\partial \theta} = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix} = (1 - \kappa) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - |\gamma| \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix}, \quad (1.14)$$

where κ is the convergence and $\gamma = \gamma_1 + i\gamma_2 = |\gamma|e^{2i\phi}$ is the shear. The quantities κ and γ will be discussed in detail in the subsequent sections. Figure 1.2 illustrates the effects of gravitational lensing on the shapes of background sources through the lensing matrix \mathcal{A} . The panels demonstrate how the combined effects of convergence and shear components in the Jacobian matrix \mathcal{A} lead to complex distortions

of background sources.

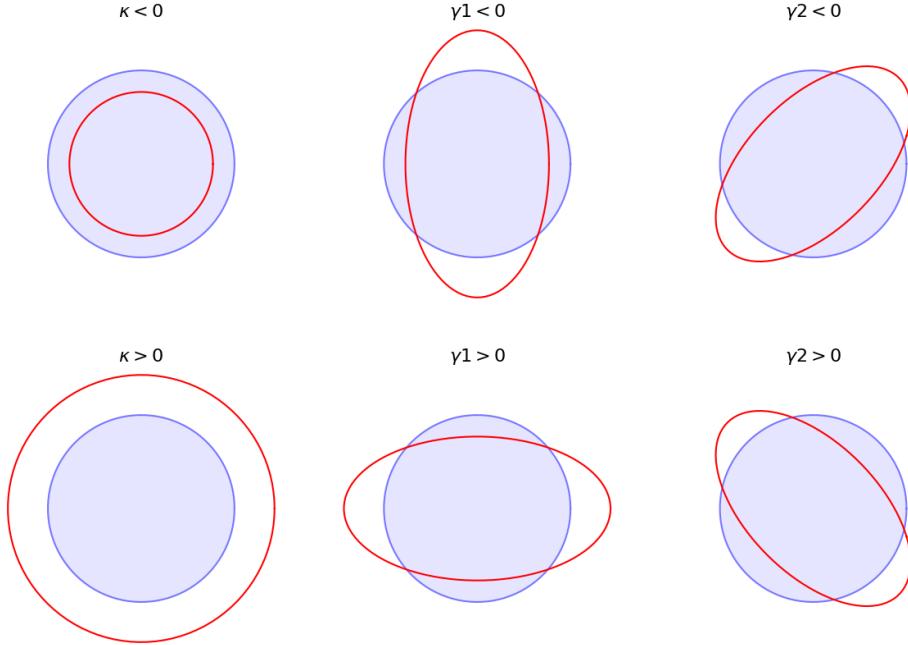


Figure 1.2: Illustration of the distortion of background sources due to gravitational lensing. The left panel depict the effect of the convergence κ and the middle and right panels show the components of the shear $\gamma = \gamma_1 + i\gamma_2$ on circular background sources. Positive and negative values of κ cause isotropic magnification or demagnification, while γ_1 and γ_2 introduce anisotropic distortions, stretching the sources along or at an angle to the principal axes.

1.2 Convergence

1.2.1 Definition

From the lensing matrix in Eq. (1.14), the convergence κ is defined as:

$$\kappa(\theta) := \frac{1}{2} \left(\frac{\partial^2 \psi}{\partial \theta_1^2} + \frac{\partial^2 \psi}{\partial \theta_2^2} \right) = \frac{1}{2} \nabla_\theta^2 \psi(\theta) \quad (1.15)$$

with θ_1 and θ_2 representing the angular coordinates on the sky. In Fourier space, the convergence field could be expressed as:

$$\tilde{\kappa}(\ell) = \int d^2\theta e^{-i\ell \cdot \theta} \kappa(\theta) = \frac{1}{2} \ell^2 \tilde{\psi}(\ell), \quad (1.16)$$

where $(\tilde{})$ denotes the Fourier transform of the corresponding quantity and $\ell = |\ell|$ is the Fourier counterpart to the angular position θ .

1.2.2 Convergence from Density Contrast

In a flat universe, the Poisson equation in comoving coordinates is expressed as

$$\begin{aligned}\nabla_{\mathbf{x}}^2 \Phi(\mathbf{x}, \chi) &= 4\pi G a^2(\chi) \bar{\rho}_m(\chi) \delta(\mathbf{x}, \chi) \\ &= 4\pi G a^2(\chi) \left[\frac{3H_0^2 \Omega_m}{8\pi G} a^{-3}(\chi) \right] \delta(\mathbf{x}, \chi) \\ &= \frac{3}{2} \Omega_m H_0^2 a^{-1}(\chi) \delta(\mathbf{x}, \chi),\end{aligned}\quad (1.17)$$

where we utilized Eq. (??) and Eq. (??). Substituting the expression for Φ from Eq. (1.17) into the lensing potential (Eq. (1.13)) and subsequently into the convergence (Eq. (1.15)), we derive:

$$\begin{aligned}\kappa(\theta) &= \frac{1}{2} \nabla_{\theta}^2 \psi(\theta) \\ &= \frac{1}{c^2} \int_0^{\chi_s} d\chi \frac{f_K(\chi_s - \chi)}{f_K(\chi_s) f_K(\chi)} \nabla_{\theta}^2 \Phi(f_K(\chi)\theta, \chi) \\ &= \frac{1}{c^2} \int_0^{\chi_s} d\chi \frac{f_K(\chi_s - \chi)}{f_K(\chi_s) f_K(\chi)} \left[\frac{1}{f_K^2(\chi)} \nabla_{\mathbf{x}}^2 \Phi(\mathbf{x}, \chi) \right] \\ &= \int_0^{\chi_s} d\chi \frac{3\Omega_m H_0^2}{2c^2} \frac{f_K(\chi_s - \chi) f_K(\chi)}{f_K(\chi_s)} \frac{a^{-1}(\chi)}{f_K^3(\chi)} \delta(\mathbf{x}, \chi) \\ &= \int_0^{\chi_s} d\chi W(\chi, \chi_s) \delta(\mathbf{x}, \chi),\end{aligned}\quad (1.18)$$

where the lensing efficiency function $W(\chi, \chi_s)$ is defined by:

$$W(\chi, \chi_s) := \frac{3\Omega_m H_0^2}{2c^2} \frac{f_K(\chi) f_K(\chi_s - \chi)}{f_K(\chi_s)} \frac{a^{-1}(\chi)}{f_K^2(\chi)}. \quad (1.19)$$

In a flat universe ($f_K(\chi) = \chi$), this simplifies to:

$$W(\chi, \chi_s) = \frac{3\Omega_m H_0^2}{2c^2} a^{-1}(\chi) \frac{\chi(\chi_s - \chi)}{\chi_s}. \quad (1.20)$$

1.3 Shear

The shear γ encapsulates the anisotropic stretching of galaxy images induced by gravitational lensing. Unlike convergence, which affects the size and brightness of images isotropically, shear induces distortions that alter the shapes of background galaxies coherently.

1.3.1 Definition

The shear components γ_1 and γ_2 describe distortions along different axes and are related to the lensing potential ψ by:

$$\gamma_1 := \frac{1}{2} \left(\frac{\partial^2 \psi}{\partial \theta_1^2} - \frac{\partial^2 \psi}{\partial \theta_2^2} \right), \quad \gamma_2 := \frac{\partial^2 \psi}{\partial \theta_1 \partial \theta_2}. \quad (1.21)$$

In Fourier space, the shear field can be expressed as:

$$\tilde{\gamma}_1(\ell) = \frac{1}{2} (\ell_1^2 - \ell_2^2) \tilde{\psi}(\ell) \quad \tilde{\gamma}_2(\ell) = \ell_1 \ell_2 \tilde{\psi}(\ell), \quad (1.22)$$

where $\tilde{\gamma}_1(\ell)$ and $\tilde{\gamma}_2(\ell)$ are the Fourier transforms of $\gamma_1(\theta)$ and $\gamma_2(\theta)$, respectively. Similar as Eq. (1.14), the shear field can be expressed in complex form as:

$$\tilde{\gamma}(\ell) := \tilde{\gamma}_1 + i\tilde{\gamma}_2 = |\tilde{\gamma}(\ell)| e^{2i\phi_\ell}, \quad \tan(2\phi_\ell) = \tilde{\gamma}_2/\tilde{\gamma}_1. \quad (1.23)$$

Therefore, the shear field in Fourier space is directly related to the convergence field combining Eq.(1.16) and Eq.(1.22):

$$\tilde{\kappa}(\ell) = \frac{\ell_1^2 - \ell_2^2 + 2i\ell_1 \ell_2}{\ell^2} \tilde{\gamma}(\ell). \quad (1.24)$$

1.3.2 E-mode and B-mode

The shear field can be decomposed into two distinct modes: the **E-mode** (gradient component) and the **B-mode** (curl component). By rotating the complex shear field to align with the principal axes of the shear $\tilde{\gamma}_{EB} = e^{-2i\phi_l} \tilde{\gamma}$, we can express the shear field in terms of the E-mode and B-mode components:

$$\begin{pmatrix} \tilde{\gamma}_E \\ \tilde{\gamma}_B \end{pmatrix} := \begin{pmatrix} \cos 2\phi_l & \sin 2\phi_l \\ -\sin 2\phi_l & \cos 2\phi_l \end{pmatrix} \begin{pmatrix} \tilde{\gamma}_1 \\ \tilde{\gamma}_2 \end{pmatrix} \quad (1.25)$$

where $\tilde{\gamma}_E$ and $\tilde{\gamma}_B$ are the E-mode and B-mode components of the shear, respectively. The E-mode represents the gradient component of the shear field, while the B-mode describes the curl component.

For standard gravitational lensing by density fluctuations, the B-mode component $\tilde{\gamma}_B(\ell)$ is expected to vanish in the absence of systematics or additional physical effects. This implies that all the shear signal is contained within the E-mode.such that:

$$\tilde{\gamma}_E(\ell) = \tilde{\kappa}(\ell), \quad \tilde{\gamma}_B(\ell) = 0. \quad (1.26)$$

Chapter 2

Power Spectrum

The power spectrum constitutes a fundamental second-order statistical measure that characterizes the distribution of fluctuations across varying spatial scales. In this chapter, we provide an overview of the power spectrum, focusing on its definition, properties, and cosmological applications following [Hu \(2000\)](#) and [Kilbinger et al. \(2017\)](#).

2.1 Overview of Power Spectrum

Power Spectrum is defined as the Fourier transform of the two-point correlation function, $\xi(\mathbf{x})$, which quantifies the statistical correlation between density fluctuations at two distinct spatial positions. Formally, the two-point correlation function is expressed as:

$$\xi(\mathbf{x}) := \langle \delta(\mathbf{x}') \delta(\mathbf{x}' + \mathbf{x}) \rangle, \quad (2.1)$$

where $\delta(\mathbf{x})$ represents the density fluctuation at position \mathbf{x} , and the angular brackets $\langle \cdot \rangle$ denote the ensemble average over all realizations of the density field.

Subsequently, the correlation of these fluctuations in Fourier space is considered. By taking the Fourier transform of the two-point correlation function, we obtain:

$$\begin{aligned} \langle \tilde{\delta}(\mathbf{k}) \tilde{\delta}(\mathbf{k}') \rangle &= \int d\mathbf{x} e^{-i\mathbf{k}\cdot\mathbf{x}} \int d\mathbf{x}' e^{-i\mathbf{k}'\cdot\mathbf{x}'} \langle \delta(\mathbf{x}) \delta(\mathbf{x}') \rangle \\ &= (2\pi)^3 \delta^{(3)}(\mathbf{k} + \mathbf{k}') \int d\mathbf{x} e^{-i\mathbf{k}\cdot\mathbf{x}} \xi(\mathbf{x}), \end{aligned} \quad (2.2)$$

where $\tilde{\delta}(\mathbf{k})$ denotes the Fourier transform of the density fluctuation field $\delta(\mathbf{x})$, and $\delta^{(3)}$ represents the three-dimensional Dirac delta function.

From this formulation, the power spectrum $P(k)$ is defined as:

$$\langle \tilde{\delta}(\mathbf{k}) \tilde{\delta}(\mathbf{k}') \rangle := (2\pi)^3 \delta^{(3)}(\mathbf{k} + \mathbf{k}') P(k), \quad \text{where} \quad P(k) := \int d\mathbf{x} e^{-i\mathbf{k}\cdot\mathbf{x}} \xi(\mathbf{x}). \quad (2.3)$$

Here, $P(k)$ encapsulates the amplitude of density fluctuations as a function of the wavenumber k .

This formalism can be naturally extended to two-dimensional (2D) analyses by substituting the three-dimensional wavenumber vector \mathbf{k} with a two-dimensional counterpart, ℓ , and correspondingly replacing the three-dimensional power spectrum with the two-dimensional power spectrum, $C(\ell)$. The two-dimensional correlation function is defined analogously:

$$\omega(\theta) := \langle \delta(\theta') \delta(\theta' + \theta) \rangle, \quad (2.4)$$

where θ denotes the angular separation vector in the 2D plane.

Consequently, the two-dimensional power spectrum is expressed as:

$$\langle \tilde{\delta}(\ell) \tilde{\delta}(\ell') \rangle := (2\pi)^2 \delta^{(2)}(\ell + \ell') C(\ell), \quad (2.5)$$

where $\delta^{(2)}$ is the two-dimensional Dirac delta function, and $C(\ell)$ is defined by the Fourier transform of the 2D correlation function:

$$C(\ell) = \int d^2\theta e^{-i\ell\cdot\theta} \omega(\theta) = 2\pi \int \theta d\theta J_0(\ell\theta) \omega(\theta). \quad (2.6)$$

In the final equality, the integral has been converted to polar coordinates, where J_0 denotes the zeroth-order Bessel function of the first kind, and $\ell = |\ell|$ is the magnitude of the 2D wavenumber vector.

2.2 Two-Point Correlation Function

The real-space shear two-point correlation function is one of the most basic observables for cosmic shear. The shear can be decomposed into a tangential γ_+ and a cross-component γ_\times with respect to a reference point θ :

$$\begin{aligned} \gamma_+(\theta, \theta') &= -\Re[\gamma(\theta') e^{-2i\phi}], \\ \gamma_\times(\theta, \theta') &= -\Im[\gamma(\theta') e^{-2i\phi}], \end{aligned} \quad (2.7)$$

where \Re and \Im denote the real and imaginary parts, respectively, and ϕ is the polar angle between θ and θ' . Following Kilbinger (2015), the two-point correlators $\langle \gamma_+ \gamma_+ \rangle, \langle \gamma_\times \gamma_\times \rangle$ can be combined into the two components of the shear 2PCF:

$$\begin{aligned} \xi_+(\theta) &= \langle \gamma \gamma^* \rangle(\theta) = \langle \gamma_+ \gamma_+ \rangle(\theta) + \langle \gamma_\times \gamma_\times \rangle(\theta), \\ \xi_\times(\theta) &= \Re[\langle \gamma \gamma \rangle(\theta) e^{-4i\phi}] = \langle \gamma_+ \gamma_- \rangle(\theta) + \langle \gamma_\times \gamma_\times \rangle(\theta) \end{aligned} \quad (2.8)$$

The main reason for this decomposition is that they can be estimated from the measured ellipticities of galaxies ϵ_i as:

$$\gamma_{\pm}(\theta) = \frac{\sum_{i,j} w_i w_j \epsilon_{+,i} \epsilon_{+,j} \pm \epsilon_{\times,i} \epsilon_{\times,j}}{\sum_{i,j} w_i w_j}, \quad (2.9)$$

where the sum runs over all the galaxy pairs and the weight w_i of the ellipticity of the i -th galaxy can include the measurement error and the intrinsic ellipticity of the galaxy. By going to the Fourier space and then back to the real space, the shear 2PCF can be expressed as:

$$\xi_+(\theta) = \frac{1}{2\pi} \int [P_{\kappa}^E(\ell) + P_{\kappa}^B(\ell)] J_0(\ell\theta) \ell d\ell, \quad (2.10)$$

$$\xi_{\times}(\theta) = \frac{1}{2\pi} \int [P_{\kappa}^E(\ell) - P_{\kappa}^B(\ell)] J_4(\ell\theta) \ell d\ell, \quad (2.11)$$

where P_{κ}^E and P_{κ}^B are the E-mode and B-mode power spectra of the convergence field, respectively, and J_n denotes the n -th order Bessel function of the first kind. For the convergence case, the 2PCF can be expressed as:

$$\xi(\theta) = \frac{1}{2\pi} \int P_{\kappa}(\ell) J_0(\ell\theta) \ell d\ell. \quad (2.12)$$

which is exactly the same as the shear 2PCF $\xi_+(\theta)$ as $P_{\kappa}^E(\ell) = P_{\kappa}(\ell)$.

2.3 Power Spectrum

2.3.1 Matter Power Spectrum

The matter power spectrum, $P(k)$, is a Fourier transform of the two-point correlation function of the dark matter density field, $\delta(\mathbf{x})$. The matter power spectrum is defined through the relation:

$$\langle \tilde{\delta}(\mathbf{k}) \tilde{\delta}(\mathbf{k}') \rangle = (2\pi)^3 \delta^{(3)}(\mathbf{k} + \mathbf{k}') P(k), \quad (2.13)$$

2.3.2 Convergence Power Spectrum

However, in the context of weak lensing, the matter power spectrum is the quantity that is not directly observable. Instead, the angular power spectrum of the convergence field, $C_{\ell}^{\kappa\kappa}$, is the quantity that is directly observable. The convergence power spectrum, $P_{\kappa}(\ell)$, quantifies the statistical properties of $\kappa(\theta)$ in Fourier space and is defined as:

$$\langle \tilde{\kappa}(\ell) \tilde{\kappa}(\ell') \rangle = (2\pi)^2 \delta^{(2)}(\ell + \ell') P_{\kappa}(\ell), \quad (2.14)$$

As we have seen in Eq. (1.18), the convergence field $\kappa(\theta)$ can be expressed as a weighted projection of the matter density contrast along the line of sight:

$$\kappa(\theta) = \int_0^{\chi_s} d\chi W(\chi) \delta_m(\chi\theta, \chi), \quad (2.15)$$

Recognizing the Fourier transform of the density field $\delta_m(\chi\theta, \chi)$:

$$\delta_m(\chi\theta, \chi) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \tilde{\delta}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}, \quad (2.16)$$

we substitute into the Fourier transform of the convergence field:

$$\begin{aligned} \tilde{\kappa}(\ell) &= \int_0^{\chi_s} d\chi W(\chi) \int d\theta e^{-i\ell\cdot\theta} \delta(\chi\theta, \chi) \\ &= \int_0^{\chi_s} d\chi W(\chi) \int d\theta e^{-i\ell\cdot\theta} \int \frac{d^3\mathbf{k}}{(2\pi)^3} \tilde{\delta}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} \\ &= \int_0^{\chi_s} d\chi W(\chi) \int d\theta e^{-i\ell\cdot\theta} \int \frac{dk_{\parallel}}{2\pi} \frac{d\mathbf{k}_{\perp}}{(2\pi)^2} \tilde{\delta}_m(k_{\parallel}, \mathbf{k}_{\perp}) e^{i(k_{\parallel}\chi + \chi\mathbf{k}_{\perp}\cdot\theta)} \\ &= \int_0^{\chi_s} d\chi \frac{W(\chi)}{\chi^2} \int \frac{dk_{\parallel}}{2\pi} \tilde{\delta}_m\left(k_{\parallel}, \frac{\ell}{\chi}\right) e^{ik_{\parallel}\chi} \end{aligned} \quad (2.17)$$

where \mathbf{k}_{\perp} is the component of \mathbf{k} perpendicular to the line of sight, and k_{\parallel} is the component along the line of sight. Then the ensemble average of the Fourier transform of the convergence field is:

$$\begin{aligned} \langle \tilde{\kappa}(\ell) \tilde{\kappa}(\ell') \rangle &= \int_0^{\chi_s} d\chi \frac{W(\chi)}{\chi^2} \int_0^{\chi_s} d\chi' \frac{W(\chi')}{\chi'^2} \int \frac{dk_{\parallel}}{2\pi} \int \frac{dk'_{\parallel}}{2\pi} \langle \tilde{\delta}_m\left(k_{\parallel}, \frac{\ell}{\chi}\right) \tilde{\delta}_m\left(k'_{\parallel}, \frac{\ell'}{\chi'}\right) \rangle e^{ik_{\parallel}\chi} e^{ik'_{\parallel}\chi'} \\ &= \int_0^{\chi_s} d\chi \frac{W(\chi)}{\chi^2} \int_0^{\chi_s} d\chi' \frac{W(\chi')}{\chi'^2} \int \frac{dk_{\parallel}}{2\pi} e^{ik_{\parallel}(\chi-\chi')} (2\pi)^2 \delta^{(2)}\left(\frac{\ell}{\chi} + \frac{\ell'}{\chi'}\right) P_m\left(\sqrt{k_{\parallel}^2 + \frac{\ell^2}{\chi^2}}\right) \end{aligned}$$

Under Limber approximation ($k_{\parallel} \ll \ell/\chi$; Limber, 1954), the last integral can be approximated as:

$$\int \frac{dk_{\parallel}}{2\pi} e^{ik_{\parallel}(\chi-\chi')} P_m\left(\sqrt{k_{\parallel}^2 + \frac{\ell^2}{\chi^2}}\right) \approx P_m\left(\frac{\ell}{\chi}\right) \int \frac{dk_{\parallel}}{2\pi} e^{ik_{\parallel}(\chi-\chi')} = P_m\left(\frac{\ell}{\chi}\right) \delta(\chi - \chi') \quad (2.18)$$

Therefore, we obtain the simplified ensemble average of the Fourier transform of the convergence field:

$$\langle \tilde{\kappa}(\ell) \tilde{\kappa}(\ell') \rangle = (2\pi)^2 \delta^{(2)}(\ell + \ell') \int_0^{\chi_s} d\chi \frac{W^2(\chi)}{\chi^2} P_m\left(\frac{\ell}{\chi}; \chi\right) \quad (2.19)$$

Thereby, substituting this into the definition of the convergence power spectrum (Eq. (2.14)), we get:

$$C_{\ell}^{\kappa\kappa} = \int_0^{\chi_s} d\chi \frac{W^2(\chi)}{\chi^2} P_m\left(\frac{\ell}{\chi}; \chi\right) \quad (2.20)$$

2.4 Estimator

In practice, the power spectrum is estimated from the observed data. The angular power spectrum is estimated by averaging over the $2\ell + 1$ m -modes for each ℓ :

$$\hat{C}_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2. \quad (2.21)$$

where convergence field is expanded in spherical harmonics:

$$\kappa(\theta) = \sum_{\ell, m} a_{\ell m} Y_{\ell m}(\theta). \quad (2.22)$$

This estimator provides an unbiased estimate of the true angular power spectrum under the assumption of statistical isotropy.

Chapter 3

Higher-Order Statistics

In weak lensing analysis, statistical measures are essential for extracting cosmological information from the convergence field $\kappa(\theta)$. The angular power spectrum is a fundamental two-point statistic that quantifies the variance of κ across different angular scales, effectively capturing the Gaussian features of the matter distribution. However, due to the non-linear growth of cosmic structures, the matter distribution exhibits significant non-Gaussianity, necessitating the use of higher-order statistics for a more comprehensive description.

This section explores higher-order statistics commonly employed in weak lensing analyses, including the bispectrum, probability density functions (PDFs), peak and minimum counts, and Minkowski functionals.

3.1 Bispectrum

The bispectrum is a higher-order statistic that measures the phase correlations between different modes in the convergence field, providing sensitivity to the non-Gaussian features arising from the non-linear evolution of large-scale structures. We will review [Takada & Jain \(2004\)](#) the formalism for the convergence bispectrum.

The bispectrum $B_{\ell_1 \ell_2 \ell_3}$ is defined through the expectation value of the product of three spherical harmonic coefficients ($a_{\ell m}$; Eq. 2.22): :

$$\langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3} \rangle = \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} B_{\ell_1 \ell_2 \ell_3}^{\kappa}, \quad (3.1)$$

where the term in parentheses is the Wigner 3j-symbol, which arises due to rotational invariance and enforces the selection rules:

- **Triangle Condition:** $|\ell_i - \ell_j| \leq \ell_k \leq \ell_i + \ell_j$ for all permutations of (i, j, k) .
- **Parity Condition:** $\ell_1 + \ell_2 + \ell_3$ must be even.

- **Magnetic Quantum Number Sum:** $m_1 + m_2 + m_3 = 0$.

Similar to the angular power spectrum, the convergence bispectrum can be written as ensemble averages of three modes of Fourier transformed κ :

$$\langle \tilde{\kappa}(\ell_1) \tilde{\kappa}(\ell_2) \tilde{\kappa}(\ell_3) \rangle = (2\pi)^2 \delta_D(\ell_1 + \ell_2 + \ell_3) B_{\ell_1 \ell_2 \ell_3}^\kappa, \quad (3.2)$$

The full-sky bispectrum is then approximately related to the flat-sky bispectrum as:

$$B_{\ell_1 \ell_2 \ell_3}^\kappa \simeq \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix} \times \sqrt{\frac{(2\ell_1 + 1)(2\ell_2 + 1)(2\ell_3 + 1)}{4\pi}} B^\kappa(\ell_1, \ell_2, \ell_3) \quad (3.3)$$

According to [Takada & Jain \(2004\)](#), an approximate form of Wigner 3j symbol is given by:

$$\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix} \simeq (-1)^L \frac{e^{3/2}}{\sqrt{2\pi}} \left(\frac{2}{L+2} \right)^{1/4} \times \prod_{i=1}^3 (L - \ell_i + 1)^{-1/4} \times \left(\frac{L - \ell_i + 1/2}{L - \ell_i + 1} \right)^{L - \ell_i + 1/4} \quad (3.4)$$

where $L = (\ell_1 + \ell_2 + \ell_3)/2$. Then the flat-sky lensing bispectrum is expressed in terms of the three-dimensional matter bispectrum as:

$$B_{\ell_1 \ell_2 \ell_3}^\kappa = \int_0^{\chi_s} \frac{W^3(\chi)}{\chi^4} B_m \left(\frac{\ell_1}{\chi}, \frac{\ell_2}{\chi}, \frac{\ell_3}{\chi}, z(\chi) \right) d\chi. \quad (3.5)$$

where $B_m(k_1, k_2, k_3, z)$ is the matter bispectrum at redshift z .

3.2 Probability Density Functions

The Probability Density Function (PDF) of the convergence field, κ , offers a comprehensive statistical characterization of the field's one-point distribution. It encapsulates all moments and cumulants, thereby capturing both Gaussian and non-Gaussian features inherent in the field. We adopt the formalism presented in [Boyle et al. \(2021\)](#) and [Uhlemann et al. \(2023\)](#) to rigorously describe the PDF of the convergence field.

3.2.1 Definition

The PDF, $P(\kappa)$, is formally defined such that:

$$P(\kappa) d\kappa = \text{Prob}(\kappa \leq \kappa' \leq \kappa + d\kappa), \quad (3.6)$$

which represents the probability of the convergence κ' lying within the infinitesimal interval $[\kappa, \kappa + d\kappa]$.

For a discrete set of measurements $\{\kappa_i\}_{i=1}^{N_{\text{pix}}}$ across N_{pix} pixels, the PDF can be expressed using the Dirac delta function δ_D :

$$P(\kappa) = \frac{1}{N_{\text{pix}}} \sum_{i=1}^{N_{\text{pix}}} \delta_D(\kappa - \kappa_i). \quad (3.7)$$

In practice, the exact PDF is approximated by discretizing the convergence values into bins of width $\Delta\kappa$. This leads to a binned estimator of the PDF:

$$P(\kappa) \approx \frac{1}{N_{\text{pix}}} \sum_{i=1}^{N_{\text{pix}}} \frac{\Theta(|\kappa_i - \kappa| \leq \frac{\Delta\kappa}{2})}{\Delta\kappa}, \quad (3.8)$$

where $\Theta(x)$ is the Heaviside step function. This estimator effectively counts the number of convergence values κ_i that fall within each bin centered at κ , normalizing by the total number of pixels and the bin width.

3.2.2 Normalization

To enable comparisons across different datasets and to facilitate the analysis of statistical properties, the convergence values are often normalized by their standard deviation, σ_κ . The standardized convergence, $\tilde{\kappa}_i$, is defined as:

$$\tilde{\kappa}_i = \frac{\kappa_i - \langle \kappa \rangle}{\sigma_\kappa}, \quad (3.9)$$

where $\langle \kappa \rangle$ is the mean convergence, typically set to zero following mean-field subtraction:

$$\langle \kappa \rangle = \frac{1}{N_{\text{pix}}} \sum_{i=1}^{N_{\text{pix}}} \kappa_i \approx 0. \quad (3.10)$$

The normalized PDF, $P(\tilde{\kappa})$, then satisfies the normalization condition:

$$\int_{-\infty}^{\infty} P(\tilde{\kappa}) d\tilde{\kappa} = 1. \quad (3.11)$$

3.2.3 Moments and Cumulants

The moments of the PDF provide crucial statistical descriptors of the convergence field. The n -th moment about the mean is defined as:

$$\mu_n = \langle (\tilde{\kappa})^n \rangle = \int_{-\infty}^{\infty} \tilde{\kappa}^n P(\tilde{\kappa}) d\tilde{\kappa}. \quad (3.12)$$

While the first two moments (mean and variance) fully describe a Gaussian PDF, higher-order moments and cumulants are necessary to characterize non-Gaussian features. Skewness (γ_1) and kurtosis (γ_2) are the standardized third and fourth cumulants, respectively.

The cumulants, κ_n , are related to the moments and provide information about the shape of the PDF beyond the mean and variance. They can be derived using the generating function approach. The cu-

mulant generating function (CGF), $K(t)$, is defined as the logarithm of the moment generating function (MGF):

$$K(t) = \log(\langle e^{t\tilde{\kappa}} \rangle) = \log \left(\int_{-\infty}^{\infty} e^{t\tilde{\kappa}} P(\tilde{\kappa}) d\tilde{\kappa} \right). \quad (3.13)$$

Expanding $K(t)$ in a Taylor series around $t = 0$ yields:

$$K(t) = \sum_{n=1}^{\infty} \frac{\kappa_n}{n!} t^n, \quad (3.14)$$

These relations allow us to express the cumulants in terms of moments, capturing the deviations from Gaussianity ($\kappa_n = 0$ for all $n > 2$ in a Gaussian distribution).

3.3 Peak and Minimum Counts

Peaks and minima in the convergence field correspond to localized over-densities and under-densities, respectively. Counting these extrema provides valuable information about the non-Gaussian features of the matter distribution and can be used to constrain cosmological models. We adopt the formalism presented in [Bardeen et al. \(1986\)](#), [Kacprzak et al. \(2016\)](#) and [Martinet et al. \(2018\)](#) to rigorously describe the peak and minimum counts in the convergence field.

3.3.1 Identification of Peaks and Minima

To effectively identify peaks and minima while suppressing noise and small-scale fluctuations, the convergence map $\kappa(\hat{\mathbf{n}})$ is first smoothed with a Gaussian kernel. The smoothed convergence field, $\kappa_{\text{smooth}}(\hat{\mathbf{n}})$, is defined as:

$$\kappa_{\text{smooth}}(\hat{\mathbf{n}}) = \int_{S^2} \kappa(\hat{\mathbf{n}}') W(\hat{\mathbf{n}} - \hat{\mathbf{n}}') d\hat{\mathbf{n}}', \quad (3.15)$$

where $W(\theta)$ is the Gaussian smoothing kernel given by:

$$W(\theta) = \frac{1}{2\pi\sigma_\theta^2} \exp\left(-\frac{\theta^2}{2\sigma_\theta^2}\right), \quad (3.16)$$

with $\theta = \arccos(\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}')$ representing the angular separation between the points $\hat{\mathbf{n}}$ and $\hat{\mathbf{n}}'$ on the unit sphere S^2 , and σ_θ is the smoothing scale.

To standardize the statistical analysis, the smoothed convergence values are normalized by their standard deviation. The normalized smoothed convergence, $\tilde{\kappa}_{\text{smooth},i}$, is defined as:

$$\tilde{\kappa}_{\text{smooth},i} = \frac{\kappa_{\text{smooth},i} - \langle \kappa_{\text{smooth}} \rangle}{\sigma_{\text{smooth}}}, \quad (3.17)$$

where:

$$\langle \kappa_{\text{smooth}} \rangle = \frac{1}{N_{\text{pix}}} \sum_{i=1}^{N_{\text{pix}}} \kappa_{\text{smooth},i}, \quad \sigma_{\text{smooth}}^2 = \sigma_{\text{signal}}^2 + \sigma_{\text{noise}}^2. \quad (3.18)$$

A pixel i in the smoothed convergence map is identified as a peak or a minimum based on the comparison of its value with its neighboring pixels. Formally, let $\mathcal{N}(i)$ denote the set of neighboring pixels of pixel i . Then:

$$\text{Peak Condition: } \kappa_{\text{smooth},i} > \kappa_{\text{smooth},j} \quad \forall j \in \mathcal{N}(i), \quad (3.19)$$

$$\text{Minimum Condition: } \kappa_{\text{smooth},i} < \kappa_{\text{smooth},j} \quad \forall j \in \mathcal{N}(i). \quad (3.20)$$

These conditions ensure that peaks are local maxima and minima are local minima in the convergence field. Figure 3.1 illustrates the identification of peaks (red circles) and minima (blue circles) in the smoothed convergence map $\kappa_{\text{smooth}}(\hat{\mathbf{n}})$.

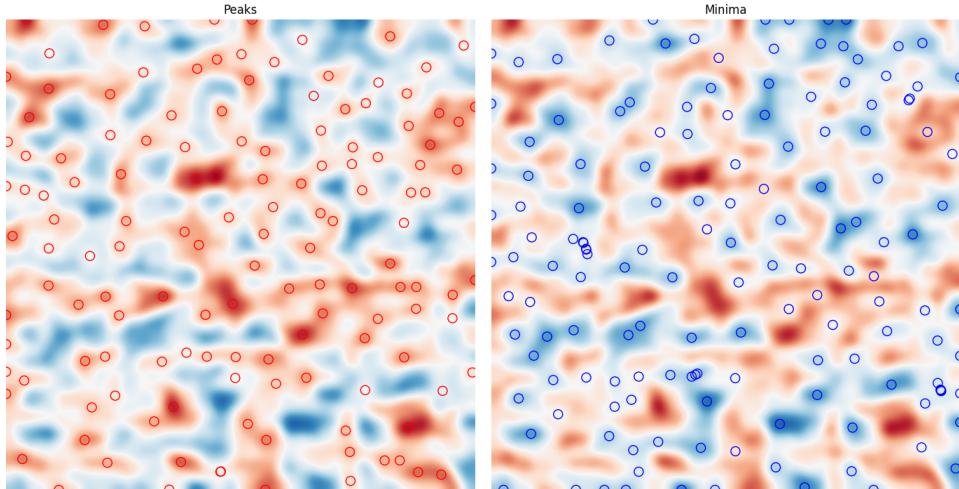


Figure 3.1: Identification of peaks and minima in a smoothed convergence map. The left panel shows the smoothed convergence field $\kappa_{\text{smooth}}(\hat{\mathbf{n}})$ with peaks (red circles) satisfying the peak condition (Equation (3.19)), and the right panel highlights the minima (blue circles) satisfying the minimum condition (Equation (3.20)).

3.3.2 Peaks of Gaussian Random Fields

For a Gaussian random field (GRF), the statistics of peaks and minima can be analytically derived. The expected number density of peaks above a threshold ν in a GRF is given by:

$$N_{\text{peak}}(\nu) = \frac{1}{(2\pi)^{3/2} R_*^3} e^{-\nu^2/2} [\nu^2 - 1], \quad (3.21)$$

where $\nu = \tilde{\kappa}_{\text{smooth}}$, and R_* is a characteristic scale defined by:

$$R_*^2 = \frac{\langle |\nabla \kappa_{\text{smooth}}|^2 \rangle}{\langle \kappa_{\text{smooth}}^2 \rangle}. \quad (3.22)$$

Similarly, the expected number density of minima below $-\nu$ is:

$$N_{\min}(-\nu) = N_{\text{peak}}(\nu). \quad (3.23)$$

3.4 Minkowski Functionals

Minkowski functionals are powerful morphological descriptors derived from integral geometry, widely used to quantify the geometry and topology of spatial structures in cosmological datasets. Here, we will review [Matsubara \(2010\)](#), [Kratochvil et al. \(2012\)](#) and [Petri et al. \(2013\)](#) for a comprehensive understanding of Minkowski functionals and their applications in cosmology.

3.4.1 Definition

Given a two-dimensional random field $\kappa(\hat{\mathbf{n}})$ (in our case, the convergence field) with zero mean and variance $\kappa^2 = \sigma_0^2$, we can consider its excursion sets $\Sigma(\nu) = \{\kappa > \nu\sigma_0\}$ which consist of all the points at which the field exceeds a particular threshold value $\nu\sigma_0$. Figure 3.2 depicts the excursion sets $\Sigma(\nu) = \{\kappa > \nu\sigma_0\}$ of a two-dimensional convergence field $\kappa(\hat{\mathbf{n}})$, which is characterized by a zero mean and variance σ_0^2 . The figure consists of four panels, each corresponding to an increasing threshold value of $\nu = 0.5, 1, 1.5$, and 2 . As the threshold ν increases, the excursion sets progressively diminish in size and connectivity. The three Minkowski functionals $V_0(\nu)$, $V_1(\nu)$, and $V_2(\nu)$ measure, respectively, the

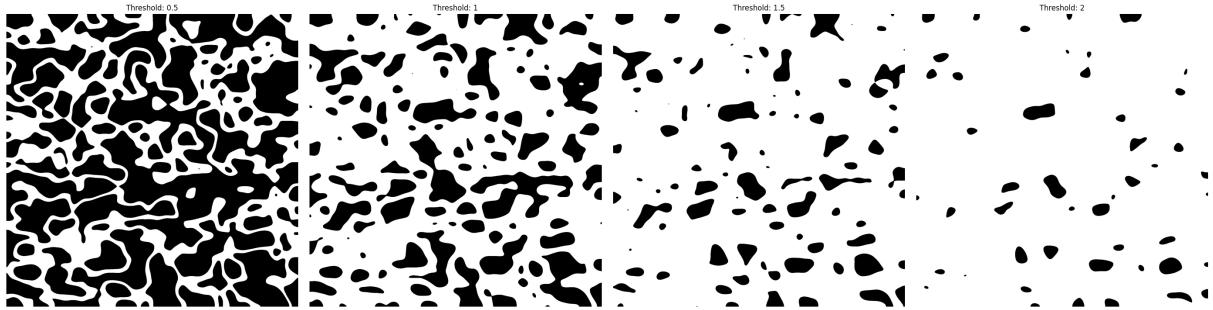


Figure 3.2: Visualization of excursion sets $\Sigma(\nu) = \{\kappa > \nu\sigma_0\}$ of a two-dimensional random field $\kappa(\hat{\mathbf{n}})$, where κ represents the convergence field with zero mean and variance σ_0^2 . Black regions correspond to regions where the field exceeds the threshold value $\nu\sigma_0$. The panels correspond to increasing threshold values ($\nu = 0.5, 1, 1.5, 2$), illustrating how the size and connectivity of the excursion sets change as the threshold increases.

area, the length of the boundary, and the genus characteristic of these excursion sets:

$$V_0(\nu) = \frac{1}{A} \int_{\Sigma(\nu)} da, \quad (3.24)$$

$$V_1(\nu) = \frac{1}{4A} \int_{\partial\Sigma(\nu)} dl, \quad (3.25)$$

$$V_2(\nu) = \frac{1}{2\pi A} \int_{\partial\Sigma(\nu)} \mathcal{K} dl, \quad (3.26)$$

where A is the total area of the field, da and dl are the area and length elements, and \mathcal{K} is the geodesic curvature along the boundary $\partial\Sigma(\nu)$.

3.4.2 Computation of Minkowski Functionals

In practical applications, Minkowski functionals are computed numerically from discrete convergence maps.

For a given threshold ν , the excursion set $\Sigma(\nu)$ is identified by:

$$\Sigma(\nu) = \{\hat{\mathbf{n}} \in S^2 \mid \tilde{\kappa}(\hat{\mathbf{n}}) > \nu\}, \quad (3.27)$$

where $\tilde{\kappa} = (\kappa - \langle \kappa \rangle)/\sigma_0$ is the normalized convergence field.

Given a pixelized convergence map, the continuous integrals in Equations (3.24)–(3.26) are approximated by discrete sums (Kratochvil et al., 2012):

$$V_0(\nu) \approx \frac{1}{N_{\text{pix}}} \sum_{i=1}^{N_{\text{pix}}} \Theta(\tilde{\kappa}_i - \nu), \quad (3.28)$$

$$V_1(\nu) \approx \frac{1}{N_{\text{pix}}} \sum_{i=1}^{N_{\text{pix}}} \delta_D(\tilde{\kappa}_i - \nu) \sqrt{\kappa_{,x}^2 + \kappa_{,y}^2}, \quad (3.29)$$

$$V_2(\nu) \approx \frac{1}{N_{\text{pix}}} \sum_{i=1}^{N_{\text{pix}}} \delta_D(\tilde{\kappa}_i - \nu) \frac{2\kappa_{,x}\kappa_{,y}\kappa_{,xy} - \kappa_{,x}^2\kappa_{,yy} - \kappa_{,y}^2}{\kappa_{,xx}\kappa_{,x}^2 + \kappa_{,y}^2} \quad (3.30)$$

where $\mathcal{N}(i)$ denotes the set of neighboring pixels to pixel i , and δ_D is the Dirac delta function. The first and second order derivatives $\kappa_{,x}$ etc., are approximated by finite differences.

3.4.3 Minkowski Functionals in Gaussian Random Fields

In the analysis of *Gaussian random fields* (GRFs), Minkowski functionals provide a robust analytical framework for quantifying the morphological characteristics of the field's excursion sets. For a two-dimensional GRF $\kappa(\hat{\mathbf{n}})$ with zero mean and variance σ_0^2 , the Minkowski functionals can be explicitly

calculated and are given by ([Matsubara, 2010](#)):

$$V_0(\nu) = \frac{1}{2} \left[1 - \operatorname{erf} \left(\frac{\nu}{\sqrt{2}\sigma_0} \right) \right], \quad (3.31)$$

$$V_1(\nu) = \frac{1}{8\sqrt{2}} \frac{\sigma_1}{\sigma_0} \exp \left(-\frac{\nu^2}{2\sigma_0^2} \right), \quad (3.32)$$

$$V_2(\nu) = \frac{\nu}{4\sqrt{2}} \frac{\sigma_1^2}{\sigma_0^3} \exp \left(-\frac{\nu^2}{2\sigma_0^2} \right), \quad (3.33)$$

where ν denotes the threshold level defining the excursion set, erf is the error function, and $\sigma_1^2 = \langle |\nabla \kappa|^2 \rangle = \langle \kappa_{,x}^2 + \kappa_{,y}^2 \rangle$ represents the variance of the gradient of the field. Here, $V_0(\nu)$ corresponds to the area fraction of the excursion set, $V_1(\nu)$ to the total boundary length per unit area, and $V_2(\nu)$ to the Euler characteristic per unit area.

Chapter 4

Covariance

Covariance matrices play a crucial role in weak lensing analyses, quantifying the uncertainties and correlations between different statistical measurements of the convergence field. Accurate estimation of covariance matrices is essential for reliable parameter estimation and for understanding the precision of cosmological constraints derived from weak lensing data.

The covariance matrix between two observables O_i and O_j is defined as:

$$\text{Cov}(O_i, O_j) = \langle (O_i - \langle O_i \rangle)(O_j - \langle O_j \rangle) \rangle, \quad (4.1)$$

where $\langle \cdot \rangle$ denotes the ensemble average over multiple realizations. For an unbiased estimator, the covariance matrix is given by:

$$\text{Cov}(O_i, O_j) = \frac{1}{N_{\text{sim}} - 1} \sum_{n=1}^{N_{\text{sim}}} (O_i^{(n)} - \langle O_i \rangle)(O_j^{(n)} - \langle O_j \rangle), \quad (4.2)$$

where N_{sim} is the number of simulations, and $O_i^{(n)}$ is the i -th realization of the statistic in the n -th simulation.

4.1 Covariance of Matter Power Spectrum

Before delving into two-dimensional weak lensing statistics, it is essential to understand the covariance matrix for the matter power spectrum. The covariance matrix for the matter power spectrum $P_m(k)$ is defined as:

$$\text{Cov}(k, k') = \langle P_m(k)P_m(k') \rangle - \langle P_m(k) \rangle \langle P_m(k') \rangle, \quad (4.3)$$

Ω where $P_m(k)$ is defined via the two-point correlation function in Fourier space:

$$\langle \tilde{\delta}(\mathbf{k})\tilde{\delta}(\mathbf{k}') \rangle = (2\pi)^3 \delta_D(\mathbf{k} + \mathbf{k}') P_m(|\mathbf{k}|), \quad (4.4)$$

where $\tilde{\delta}(\mathbf{k})$ is the Fourier transform of the matter overdensity field $\delta(\mathbf{x})$:

$$\tilde{\delta}(\mathbf{k}) = \int_V \delta(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} d^3x. \quad (4.5)$$

The matter power spectrum from a single realization is given by:

$$\hat{P}_m(k) = (2\pi)^3 \delta_D(\mathbf{k} + \mathbf{k}') \tilde{\delta}(\mathbf{k}) \tilde{\delta}(-\mathbf{k}) = V_f \int_{V_s(k)} \frac{d^3\mathbf{k}}{V_s(k)} \tilde{\delta}(\mathbf{k}) \tilde{\delta}(-\mathbf{k}), \quad (4.6)$$

where $V_f = (2\pi)^3/V$ is the volume of a Fourier cell where V is the total survey volume, and $V_s(k) = 4\pi k^2 \Delta k$ is the volume of the shell in Fourier space corresponding to wavenumber k . To derive the covariance matrix, we substitute the estimator for $P_m(k)$ into the covariance definition:

$$\text{Cov}(\mathbf{k}, \mathbf{k}') = V_f^2 \left(\langle \tilde{\delta}(\mathbf{k}) \tilde{\delta}(-\mathbf{k}) \tilde{\delta}(\mathbf{k}') \tilde{\delta}(-\mathbf{k}') \rangle - \langle \tilde{\delta}(\mathbf{k}) \tilde{\delta}(-\mathbf{k}) \rangle \langle \tilde{\delta}(\mathbf{k}') \tilde{\delta}(-\mathbf{k}') \rangle \right) \quad (4.7)$$

The first term on the right-hand side involves the four-point correlation function of the density field, which can be decomposed using Wick's theorem:

$$\begin{aligned} \langle \tilde{\delta}(\mathbf{k}) \tilde{\delta}(-\mathbf{k}) \tilde{\delta}(\mathbf{k}') \tilde{\delta}(-\mathbf{k}') \rangle &= \langle \tilde{\delta}(\mathbf{k}) \tilde{\delta}(-\mathbf{k}) \rangle \langle \tilde{\delta}(\mathbf{k}') \tilde{\delta}(-\mathbf{k}') \rangle \\ &+ \langle \tilde{\delta}(\mathbf{k}) \tilde{\delta}(\mathbf{k}') \rangle \langle \tilde{\delta}(-\mathbf{k}) \tilde{\delta}(-\mathbf{k}') \rangle \\ &+ \langle \tilde{\delta}(\mathbf{k}) \tilde{\delta}(-\mathbf{k}') \rangle \langle \tilde{\delta}(-\mathbf{k}) \tilde{\delta}(\mathbf{k}') \rangle \\ &+ \langle \tilde{\delta}(\mathbf{k}) \tilde{\delta}(-\mathbf{k}) \tilde{\delta}(\mathbf{k}') \tilde{\delta}(-\mathbf{k}') \rangle_c, \end{aligned} \quad (4.8)$$

where the subscript c denotes the connected part of the four-point function, also known as the **trispectrum** $T(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4)$. Substituting the decomposed four-point function back into the covariance expression, and utilize the Dirac delta function approximation, the covariance matrix simplifies to:

$$\text{Cov}(k, k') = \frac{P_m(k) P_m(k')}{V} (\delta_{\mathbf{k}, -\mathbf{k}'} + \delta_{\mathbf{k}, \mathbf{k}'}) + \frac{T(\mathbf{k}, -\mathbf{k}, \mathbf{k}', -\mathbf{k}')}{V}. \quad (4.9)$$

In practice, measurements are averaged over all modes within spherical shells around each k . Let $N(k) = V_s(k)/V_f$ be the number of independent modes in the shell at wavenumber k . The covariance simplifies to (Barreira & Schmidt, 2017):

$$\text{Cov}(k, k') = \frac{2P_m(k)^2}{N(k)} \delta_{k, k'} + \frac{T(k, k')}{V} \quad (4.10)$$

For the case with survey window effects, the covariance matrix is modified to include the super-sample covariance term:

$$\text{Cov}(k, k') = \frac{2P_m(k)^2}{N(k)} \delta_{k,k'} + \frac{T(k, k')}{V} + \frac{\partial P_m(k)}{\partial \delta_b} \frac{\partial P_m(k')}{\partial \delta_b} \sigma_b^2, \quad (4.11)$$

where

$$\sigma_b^2 = \int \frac{d^3 k}{(2\pi)^3} P_m(k, z) \tilde{W}^2(k), \quad (4.12)$$

and $\tilde{W}(k)$ is the Fourier transform of the survey window function.

4.2 Covariance of Angular Power Spectrum

We consider a cosmological survey characterized by a window function $W(\theta)$ and a survey area A . The survey area is defined as the integral of the window function over the sky ([Takada & Hu, 2013](#)):

$$A = \int d^2 \theta W(\theta), \quad (4.13)$$

The Fourier transform of the window function, $\tilde{W}(\mathbf{q})$, plays a pivotal role in relating the observed field to its underlying statistical properties. The estimator for angular power spectrum C_ℓ in the presence of a window function is given by:

$$\hat{C}_\ell = \frac{1}{A} \int_{A_\ell} \frac{d^2 \ell'}{A_\ell} \int \frac{d^2 q_1}{(2\pi)^2} \int \frac{d^2 q_2}{(2\pi)^2} \tilde{W}(\mathbf{q}_1) \tilde{W}(\mathbf{q}_2) \tilde{\kappa}(\ell' - \mathbf{q}_1) \tilde{\kappa}(-\ell' - \mathbf{q}_2), \quad (4.14)$$

where:

- $\tilde{\kappa}(\ell)$ is the Fourier transform of the convergence field $\kappa(\theta)$,
- A_ℓ is the area of the annulus in Fourier space corresponding to multipole ℓ , defined as:

$$A_\ell = \int_{|\ell'|=\ell} d^2 \ell' \approx 2\pi \ell \Delta \ell (\Delta \ell / \ell \ll 1), \quad (4.15)$$

- The integrals over \mathbf{q}_1 and \mathbf{q}_2 account for the effects of the survey window.

Following a procedure analogous to that of the matter power spectrum, the covariance matrix for the angular power spectrum C_ℓ is defined as:

$$\text{Cov}(\ell_1, \ell_2) = \langle \hat{C}_{\ell_1} \hat{C}_{\ell_2} \rangle - \langle \hat{C}_{\ell_1} \rangle \langle \hat{C}_{\ell_2} \rangle, \quad (4.16)$$

Substituting the estimator for \hat{C}_ℓ into the covariance definition and expanding the resulting expression leads to terms involving two-point and four-point correlation functions of the convergence field. Specif-

ically, the covariance can be expressed as:

$$\text{Cov}(\ell_1, \ell_2) = \frac{1}{A^2} \int_{A_{\ell_1}} \frac{d^2 \ell'_1}{A_{\ell_1}} \int_{A_{\ell_2}} \frac{d^2 \ell'_2}{A_{\ell_2}} [\langle \tilde{\kappa}(\ell'_1) \tilde{\kappa}(-\ell'_1) \tilde{\kappa}(\ell'_2) \tilde{\kappa}(-\ell'_2) \rangle - \langle \tilde{\kappa}(\ell'_1) \tilde{\kappa}(-\ell'_1) \rangle \langle \tilde{\kappa}(\ell'_2) \tilde{\kappa}(-\ell'_2) \rangle]. \quad (4.17)$$

After the same computation as that of matter power spectrum, the covariance matrix for the angular power spectrum C_ℓ can be expressed as:

$$\text{Cov}(\ell_1, \ell_2) = \frac{2}{A} \int_{A_{\ell_1}} \frac{d^2 \ell'_1}{A_{\ell_1}} \int_{A_{\ell_2}} \frac{d^2 \ell'_2}{A_{\ell_2}} C_{\ell'_1} C_{\ell'_2} \delta_{\ell_1 \ell_2} + \frac{1}{A} \int_{A_{\ell_1}} \frac{d^2 \ell'_1}{A_{\ell_1}} \int_{A_{\ell_2}} \frac{d^2 \ell'_2}{A_{\ell_2}} T_{\ell'_1, -\ell'_1, \ell'_2, -\ell'_2} + \text{Cov}_{\ell_1, \ell_2}^{\text{SSC}}, \quad (4.18)$$

By simplifying that we can obtain:

$$\text{Cov}(\ell_1, \ell_2) = \frac{1}{A} \left[\frac{(2\pi)^2}{A_\ell} C_{\ell_1}^2 \delta_{\ell_1, \ell_2} + \tilde{\mathcal{T}}_{\ell_1, \ell_2}^W \right] \quad (4.19)$$

where $\tilde{\mathcal{T}}_{\ell_1, \ell_2}^W$ is the windowed trispectrum, and the super-sample covariance term is included in the covariance matrix. Using the limber approximation, the covariance matrix can be related to the matter power spectrum covariance matrix and decomposed using tripectrum consistency relation:

$$\text{Cov}(\ell_1, \ell_2) = \text{Cov}^G(\ell_1, \ell_2) + \text{Cov}^{T0}(\ell_1, \ell_2) + \text{Cov}^{\text{SSC}}(\ell_1, \ell_2) \quad (4.20)$$

where:

$$\text{Cov}^G(\ell_1, \ell_2) = \frac{1}{A} \frac{(2\pi)^2}{A_\ell} C_{\ell_1}^2 \delta_{\ell_1, \ell_2} \quad (4.21)$$

$$\text{Cov}^{T0}(\ell_1, \ell_2) = \frac{1}{A} \int_{A_{\ell_1}} \frac{d^2 \ell'_1}{A_{\ell_1}} \int_{A_{\ell_2}} \frac{d^2 \ell'_2}{A_{\ell_2}} T_{\ell'_1, -\ell'_1, \ell'_2, -\ell'_2} \quad (4.22)$$

$$\text{Cov}^{\text{SSC}}(\ell_1, \ell_2) = \frac{1}{A^2} \int_0^{\chi_s} \frac{W^4(\chi)}{\chi^6} \frac{\partial P_m(k_i)}{\partial \delta_b} \frac{\partial P_m(k_j)}{\partial \delta_b} \sigma_b^2 \quad (4.23)$$

4.3 Covariance of Higher-Order Statistics

Most of the higher-order statistics, such as the bispectrum, peak counts, and Minkowski functionals, are difficult to compute their covariance matrices analytically. Therefore, researchers often rely on simulations to estimate the covariance matrices for these statistics. Assuming an analogy between the matter power spectrum and the higher-order statistics, the covariance matrix for the higher-order statistics can be expressed as:

$$\text{Cov}(O_i, O_j) = \text{Cov}^{\text{noSSC}}(O_i, O_j) + \text{Cov}^{\text{SSC}}(O_i, O_j), \quad (4.24)$$

Since in the higher-order statistics, we expect them to be non-gaussian so that it makes non-sense to separate them into gaussian and non-Gaussian. Therefore, we just separate the super-sample covariance

part apart. The super-sample covariance term with response function can be given by (Li et al., 2014):

$$\text{Cov}^{\text{SSC}}(O_i, O_j) = \frac{\partial O_i}{\partial \delta_b} \frac{\partial O_j}{\partial \delta_b} \sigma_b^2 \quad (4.25)$$

Chapter 5

Surveys(Make it shorter)

5.1 Introduction to Astronomical Surveys

5.1.1 Purpose and Scientific Objectives

Contemporary astronomical surveys constitute extensive and systematic observational endeavors designed to map and catalog vast regions of the sky with unparalleled depth and precision. These surveys yield comprehensive datasets that are crucial for addressing fundamental questions in astrophysics and cosmology. They facilitate investigations into the elusive nature of *dark energy*, provide rigorous tests of the *standard cosmological model*—the Λ CDM paradigm—and enable detailed studies of the formation and evolution of *cosmic structures*. By delivering high-fidelity observational data, these surveys significantly advance our understanding of the Universe and the fundamental physical laws that govern it.

Testing the Standard Cosmological Model One of the primary objectives of astronomical surveys is to rigorously test the standard cosmological model, known as the Λ CDM model (Lambda Cold Dark Matter). While the Λ CDM model successfully accounts for a wide range of cosmological observations ?, it encounters notable challenges such as the *Hubble tension*—a significant discrepancy between measurements of the Universe’s expansion rate derived from early-universe observations ? and those obtained from late-universe observations ?—and inconsistencies in parameters like the amplitude of matter clustering, denoted by S_8 ?. These surveys endeavor to produce precise measurements of cosmological parameters, aiming to either corroborate the Λ CDM model or uncover deviations that could indicate new physics beyond the current paradigm. By juxtaposing high-precision observational data with theoretical predictions, astronomical surveys enhance our understanding of the fundamental forces and constituents that shape the cosmos.

Studying the Growth and Evolution of Cosmic Structures Another fundamental objective of astronomical surveys is to investigate the formation and evolution of cosmic structures, including galax-

ies, galaxy clusters, and dark matter halos. By mapping the spatial distribution of millions of galaxies, these surveys construct detailed maps of dark matter over vast cosmological volumes. Techniques such as *weak gravitational lensing*—which measures subtle distortions in the shapes of background galaxies caused by the gravitational potential of intervening mass distributions—illuminate the role of dark matter in the process of structure formation [?](#). Moreover, *galaxy-galaxy lensing*, which correlates the lensing signal with the positions of foreground galaxies, provides insights into the connection between luminous matter and the underlying dark matter distribution [?](#). These comprehensive maps enable stringent tests of theoretical predictions from various cosmological models regarding the large-scale structure of the Universe.

5.1.2 Imaging vs. Spectroscopic Surveys

Astronomical surveys can be broadly categorized into *imaging surveys* and *spectroscopic surveys*, each employing distinct methodologies to capture and analyze celestial phenomena.

Imaging Surveys

Imaging surveys acquire wide-field images of the sky across multiple wavelengths, providing spatial data that enable astronomers to map cosmic structures, identify transient events, and analyze galaxy populations. By systematically scanning large regions of the sky, these surveys generate extensive catalogs documenting object positions, apparent magnitudes, and morphological characteristics. The varying depths and angular resolutions of these surveys facilitate the detection of both nearby and distant astronomical objects, thereby supporting detailed morphological and statistical studies across cosmic time. Prominent examples of imaging surveys include the *Sloan Digital Sky Survey* (SDSS; [Kollmeier et al. 2019](#)), the *Dark Energy Survey* (DES; [Abbott et al. 2018](#)), and the forthcoming *Legacy Survey of Space and Time* (LSST; [Ivezic et al. 2019](#)).

Spectroscopic Surveys

Spectroscopic surveys collect detailed spectral data from individual celestial objects by dispersing their emitted light into constituent wavelengths, thereby revealing redshifts, chemical compositions, and kinematic properties. These measurements are indispensable for studying galaxy dynamics, the distribution of dark matter, and the expansion history of the Universe. Unlike imaging surveys that capture broad swaths of the sky, spectroscopic surveys often target specific objects identified in imaging data, focusing on obtaining high-resolution spectral information. The spectral resolution and wavelength coverage are critical parameters that determine the precision of measurements such as redshift determinations and elemental abundances. Key spectroscopic surveys include the *Baryon Oscillation Spectroscopic Survey* (BOSS; [Dawson et al. 2013](#)), which investigates baryon acoustic oscillations, the *Dark Energy Spectro-*

scopic Instrument (DESI; [DESI Collaboration et al. 2016](#)), designed to map large-scale cosmic structure, and the *Kilo-Degree Survey* (KiDS; [de Jong et al. 2013](#)) with its spectroscopic extensions that enhance research on dark matter and dark energy.

5.1.3 Ground-Based vs. Space-Based Surveys

The operational platform of an astronomical survey—whether ground-based or space-based—fundamentally determines its observational capabilities, including wavelength coverage, angular resolution, and sensitivity.

Ground-Based Surveys

Ground-based surveys utilize telescopes located on Earth’s surface, capitalizing on accessible infrastructure and enabling routine maintenance and upgrades of large, high-power instruments. These observatories can host a diverse array of instruments designed for observations across multiple wavelengths and facilitate prompt follow-up observations. However, atmospheric effects such as absorption, scattering, and turbulence degrade image quality and restrict observations at certain wavelengths, particularly in the ultraviolet and infrared regions. Notable ground-based surveys include the *Hyper Suprime-Cam* (HSC; [Aihara et al. 2018a](#)), the *Dark Energy Survey* (DES; [Abbott et al. 2018](#)), and the *Kilo-Degree Survey* (KiDS; [de Jong et al. 2013](#)).

Space-Based Surveys

Space-based surveys operate from satellites or space telescopes positioned above Earth’s atmosphere, offering exceptional clarity and sensitivity across a wide range of wavelengths, especially in the ultraviolet and infrared bands that are largely inaccessible from the ground. Free from atmospheric distortions, these instruments achieve high angular resolution and can detect faint, distant objects through long-duration, stable observations. Despite these advantages, space missions are costly, require extensive international collaboration, and present limited opportunities for maintenance, upgrades, or repairs. Prominent examples include the *Hubble Space Telescope* (HST; [Freedman et al. 2001](#)), the forthcoming *Nancy Grace Roman Space Telescope* (*Roman*; [Spergel et al. 2015](#)), and the *Gaia* mission ([Gaia Collaboration et al., 2016](#)).

5.1.4 Stage-III vs. Stage-IV Surveys

The classification of astronomical surveys into *Stage-III* and *Stage-IV* categories serves to distinguish successive generations based on technological sophistication, scale, and scientific objectives. While this framework is particularly pertinent in dark energy research [Albrecht et al. \(2006\)](#), it is broadly applicable across cosmology and astrophysics.

Stage-III Surveys

Stage-III surveys represent the current generation of large-scale astronomical projects, integrating advanced yet moderately scaled instrumentation and methodologies. Their primary aims include refining cosmological parameters, mapping large-scale structures, and deepening our understanding of dark energy and dark matter. Utilizing multi-band imaging and spectroscopy, Stage-III surveys achieve moderate sky coverage and depth, supported by sophisticated data processing pipelines capable of handling substantial datasets. Notable examples of Stage-III surveys include the *Dark Energy Survey* (DES; [Abbott et al. 2018](#)), the *Kilo-Degree Survey* (KiDS; [de Jong et al. 2013](#)), and the *Hyper Suprime-Cam* survey (HSC; [Aihara et al. 2018a](#)), all of which have provided high-quality data instrumental in advancing cosmological models.

Stage-IV Surveys

Stage-IV surveys represent the forthcoming generation of astronomical initiatives, characterized by unparalleled scale, depth, and precision. Leveraging cutting-edge technologies, these surveys aim to achieve high-precision cosmological measurements, explore dark energy and dark matter with greater fidelity, and uncover new astrophysical phenomena. They are distinguished by ultra-wide sky coverage, deep imaging and spectroscopy, high angular resolution, and the integration of multi-wavelength data. Advanced data management techniques are employed to process petabyte-scale datasets efficiently. Prominent examples include the *Legacy Survey of Space and Time* (LSST) conducted by the Vera C. Rubin Observatory [Ivezić et al. \(2019\)](#), the *Dark Energy Spectroscopic Instrument* (DESI; [DESI Collaboration et al. 2016](#)), and the forthcoming *Nancy Grace Roman Space Telescope* (*Roman*; [Spergel et al. 2015](#)), all of which are poised to transform our understanding of the cosmos and address critical questions in astrophysics.

5.2 Subaru Hyper Suprime-Cam

5.2.1 Instrument Overview

The *Subaru Hyper Suprime-Cam* (HSC; [Miyazaki et al. 2018](#)) is a state-of-the-art wide-field imaging instrument mounted at the prime focus of the 8.2-meter Subaru Telescope located on Maunakea, Hawaii. The HSC features an expansive 1.5° diameter field of view, achieved through an array of 104 charge-coupled devices (CCDs), each with 2048×4096 pixels, culminating in a total of approximately 870 megapixels. This sophisticated configuration enables the acquisition of high-fidelity, wide-field optical data, augmented by a precision corrector lens system that preserves image quality across the entire observational field. The substantial field of view, combined with Subaru's significant light-collecting capabilities, renders the HSC exceptionally well-suited for deep and wide imaging surveys that are critical

for studies of *weak gravitational lensing* and the *large-scale structure* of the Universe.

The instrument is equipped with five broad-band filters— g, r, i, z , and y ([Kawanomoto et al. 2018](#))—alongside several narrow-band filters that cover optical to near-infrared wavelengths. Benefiting from the exceptional atmospheric conditions at Maunakea and the high-resolution optics of the Subaru Telescope, the HSC achieves a median seeing of 0.6 arcseconds in the i -band, facilitating detailed and precise astronomical analyses.

5.2.2 Survey Design: The Hyper Suprime-Cam Subaru Strategic Program

The *Hyper Suprime-Cam Subaru Strategic Program* (HSC-SSP; [Aihara et al. 2018a](#)) is a comprehensive, multi-band imaging survey that exploits the advanced capabilities of the HSC instrument. Initially allocated 300 observing nights over a five-year period starting in March 2014, the program’s duration was extended to 340 nights to compensate for observational losses due to weather conditions, technical issues, and seismic activity. The HSC-SSP is meticulously designed to address a broad spectrum of astrophysical inquiries, including but not limited to cosmology, galaxy evolution, and solar system science.

The survey is stratified into three distinct layers, each engineered to achieve specific scientific objectives through varying degrees of sky coverage and imaging depth:

- **Wide Layer:** Encompassing approximately 1,400 square degrees around the celestial equator, the Wide layer attains a depth of $i \sim 26$ magnitudes. This extensive coverage is optimized for mapping large-scale structures and conducting statistical studies of weak gravitational lensing, thereby contributing significantly to our understanding of dark matter distribution and cosmic acceleration.
- **Deep Layer:** Covering roughly 28 square degrees across four distinct fields, the Deep layer reaches a depth of $i \sim 27$ magnitudes. This intermediate configuration balances area and depth, supporting research on galaxy evolution, supernova detection, and the statistical properties of faint astronomical objects.
- **UltraDeep Layer:** Targeting two fields that collectively cover 3.5 square degrees, the UltraDeep layer achieves an impressive depth of $i \sim 28$ magnitudes. Utilizing both broad-band and narrow-band filters, this layer is dedicated to probing the high-redshift Universe, facilitating investigations into early galaxy formation, quasars, and the epoch of reionization, thereby providing critical insights into the early phases of cosmic evolution.

5.2.3 Data Releases and Processing

To date, the HSC-SSP has disseminated several public data releases—DR1, DR2, and DR3 ([Aihara et al. 2018b, 2019, 2022](#))—each incorporating progressively larger datasets, enhanced calibration procedures,

and refined data processing methodologies. The data processing pipeline is managed using the *Legacy Survey of Space and Time* (LSST) Science Pipelines, which provide high-quality analysis tools accessible to the astronomical community. These tools generate calibrated imaging data, comprehensive photometric catalogs, and advanced data products such as photometric redshifts and precise shape measurements. Consequently, the HSC-SSP data releases empower a wide array of scientific investigations and promote collaborative advancements within the field of astronomy.

5.3 Dark Energy Survey

5.3.1 Overview and Objectives

The *Dark Energy Survey* (DES; [The Dark Energy Survey Collaboration 2005](#); Abbott et al. 2018, 2021) was meticulously designed as a comprehensive, wide-field imaging survey operating in optical and near-infrared wavelengths, with the primary objective of probing the nature of cosmic acceleration and the underlying mechanism of *dark energy*. Spanning approximately 5,000 square degrees of the Southern Galactic Cap, DES was conducted over a six-year period from August 2013 to January 2019. The survey utilized the 570-megapixel *Dark Energy Camera* (DECam; [Flaugher et al. 2015](#)), which is mounted on the 4-meter *Blanco Telescope* at the Cerro Tololo Inter-American Observatory (CTIO) in Chile.

5.3.2 Survey Components

DES observations were strategically divided into two primary components: the *Wide-Area Survey* and the *Supernova Survey*, each meticulously tailored to address specific scientific goals within the overarching framework of DES.

Wide-Area Survey

The Wide-Area Survey constituted the principal observational effort of DES, covering approximately 5,000 square degrees across five photometric bands (g , r , i , z , Y). This extensive footprint was purposefully selected to overlap with other significant surveys, such as the *South Pole Telescope Survey* and the *Sloan Digital Sky Survey* (SDSS) Stripe 82 [Abazajian et al. \(2009\)](#), thereby enhancing calibration accuracy and enabling synergistic cross-survey analyses. The Wide-Area Survey aimed to map the large-scale structure of the Universe with high precision, facilitating studies on galaxy clustering, weak gravitational lensing, and the abundance of galaxy clusters, all of which are sensitive probes of dark energy and cosmological parameters.

Supernova Survey

Complementing the Wide-Area Survey, the Supernova Survey focused on a more concentrated region of approximately 27 square degrees. This component employed a higher observational cadence, with images acquired in the g , r , i , and z bands roughly every seven days. Such a cadence was specifically designed to capture transient phenomena, notably *Type Ia supernovae* Abbott et al. (2019), which serve as standard candles for cosmic distance measurements. By monitoring these supernovae over time, the Supernova Survey aimed to construct a Hubble diagram extending to high redshifts, thereby providing stringent constraints on the equation of state of dark energy.

5.3.3 Instrumentation and Data Processing

DECam, the instrument central to DES, features a focal plane array of 62 science CCDs, each with 2048×4096 pixels, providing a large field of view of 2.2 degrees in diameter. The camera's design includes specialized optics to achieve uniform image quality across the field, essential for weak lensing measurements that require precise shape determinations of distant galaxies.

Data processing for DES was handled by the *DES Data Management* (DESDM) system Mohr et al. (2008), which performed tasks ranging from image detrending and astrometric calibration to co-addition and catalog generation. Advanced algorithms were employed for photometric calibration, object detection, and classification, ensuring high-quality data products suitable for cosmological analyses.

5.3.4 Scientific Contributions

DES has made significant contributions to cosmology and astrophysics, including precise measurements of cosmic shear Amon et al. (2022) and galaxy clustering Abbott et al. (2022). The combination of these probes has led to improved constraints on cosmological parameters, particularly the matter density Ω_m and the amplitude of matter fluctuations σ_8 . Furthermore, DES data have been instrumental in addressing tensions in cosmological measurements, such as the S_8 discrepancy, and in exploring possible extensions to the Λ CDM model.

5.4 Legacy Survey of Space and Time

5.4.1 Overview and Scientific Objectives

The *Legacy Survey of Space and Time* (LSST; LSST Science Collaboration et al. 2009; Ivezić et al. 2019), conducted by the Vera C. Rubin Observatory, is a groundbreaking optical survey meticulously designed to perform a comprehensive imaging campaign covering approximately half of the celestial sphere. Operating with six photometric filters (*ugrizy*), LSST will systematically re-image its surveyed regions at

intervals of a few nights over a planned ten-year observational period. This extensive survey is anticipated to produce an unparalleled dataset, comprising approximately 3.2×10^{13} individual observations of around 2×10^{10} galaxies, along with a comparable number of stellar objects. This vast repository of observations will enable a broad spectrum of scientific investigations, including studies in cosmology, Galactic structure, and time-domain astrophysics.

5.4.2 Instrumentation and Survey Strategy

To achieve its ambitious scientific objectives, LSST has been constructed as an advanced, wide-field observatory centered around an 8.4-meter primary mirror and a field of view spanning 9.6 square degrees. The system's observational capabilities are enhanced by a 3.2-gigapixel camera, facilitating high-throughput data acquisition. The unique optical design employs a three-mirror system to deliver a wide field of view with excellent image quality across the entire focal plane.

The survey primarily employs a *deep-wide-fast* observing strategy, allocating approximately 90% of the telescope's operational time to this primary mode [LSST Science Collaboration et al. \(2009\)](#). This approach aims to maximize both spatial coverage and temporal cadence across the main survey region, enabling the detection of transient phenomena and the mapping of large-scale structures. The remaining 10% of LSST's observing time is dedicated to specialized programs designed to address specific astrophysical phenomena. These programs include *deep drilling fields*, which involve intensified observations in selected regions to achieve greater depth, and targeted surveys of dynamically complex areas, such as the Galactic plane. These tailored observational approaches are integral to fulfilling LSST's broader scientific mandate, offering critical data to complement the primary survey and enabling detailed analyses of both local and distant astronomical phenomena.

5.4.3 Scientific Impact

LSST is expected to make transformative contributions across multiple domains of astrophysics. In cosmology, it will provide precise measurements of dark energy parameters through studies of weak gravitational lensing, supernovae, baryon acoustic oscillations, and large-scale structure [LSST Dark Energy Science Collaboration \(2012\)](#). In the study of the Milky Way, LSST will map the distribution and kinematics of stars, facilitating investigations into the Galaxy's formation history and structure [Robertson et al. \(2019\)](#). The time-domain capabilities of LSST will enable the discovery and monitoring of transient and variable objects, such as supernovae, variable stars, and potentially hazardous near-Earth objects [Jones et al. \(2018\)](#). Furthermore, LSST's extensive dataset will be invaluable for probing fundamental physics, such as testing general relativity on cosmological scales and searching for signatures of new physics beyond the Standard Model.

5.5 Nancy Grace Roman Space Telescope

5.5.1 Overview and Mission Objectives

The *Nancy Grace Roman Space Telescope* (hereafter referred to as *Roman*; formerly designated as the *Wide-Field Infrared Survey Telescope*, *WFIRST*; [Spergel et al. 2015](#)) represents a seminal endeavor within NASA's astrophysics portfolio. This mission leverages a 2.4-meter aperture telescope, originally developed for national reconnaissance purposes, repurposing its advanced instrumentation and observational capabilities toward forefront astronomical research. The *Roman* mission is designed to address key questions in cosmology and exoplanet science, thereby significantly enhancing our understanding of the Universe's expansion history, the distribution of dark matter and dark energy, and the prevalence of planetary systems in our galaxy.

5.5.2 Surveys and Scientific Goals

The *Roman* Space Telescope's scientific program is anchored by two primary surveys: the *High Latitude Survey* and the *Galactic Bulge Time Domain Survey*, each meticulously designed to achieve specific scientific objectives.

High Latitude Survey

A cornerstone of the *Roman*'s scientific agenda is the **High Latitude Survey**, engineered to probe the enigmatic phenomenon of dark energy, which is hypothesized to drive the accelerated expansion of the Universe [Riess et al. \(1998\)](#); [Perlmutter et al. \(1999\)](#). This survey integrates a tripartite observational strategy to examine both the temporal evolution of cosmic expansion and the growth of large-scale structures. The employed methodologies encompass:

- **Supernova Surveys:** Utilizing Type Ia supernovae as standard candles to trace the expansion history of the Universe [Sullivan et al. \(2011\)](#).
- **Weak Gravitational Lensing Imaging:** Mapping the distribution of dark matter through the subtle distortions of background galaxies [Bartelmann & Schneider \(2001\)](#).
- **Spectroscopic Surveys of Baryon Acoustic Oscillations:** Measuring the scale of baryon acoustic oscillations to serve as a cosmological standard ruler [Eisenstein et al. \(2005\)](#).

These techniques collectively furnish a comprehensive understanding of the Universe's expansion history, thereby elucidating the properties and dynamics of dark matter and dark energy. The High Latitude Survey is anticipated to cover thousands of square degrees, providing high-resolution imaging and spectroscopy in the near-infrared wavelengths.

Galactic Bulge Time Domain Survey

Complementing the High Latitude Survey, the **Galactic Bulge Time Domain Survey** employs the microlensing technique to detect and characterize a substantial population of exoplanets, including those with masses as low as that of Mars [Penny et al. \(2019\)](#). Focusing on stellar populations within the Milky Way’s bulge, this survey enhances our comprehension of planetary formation and distribution mechanisms in environments with high stellar density. By monitoring millions of stars over time, the *Roman* Space Telescope aims to detect the gravitational microlensing events caused by planets orbiting both luminous and dark host stars, thus expanding the demographic census of exoplanets [Gaudi \(2012\)](#).

5.5.3 Instrumentation and Capabilities

The *Roman* Space Telescope is equipped with a 2.4-meter primary mirror, comparable in size to the *Hubble Space Telescope*, but with a significantly larger field of view—approximately 100 times greater [Spergel et al. \(2015\)](#). The primary instruments include:

- **Wide Field Instrument (WFI):** A near-infrared imager and spectrometer that provides high-resolution imaging and integral-field spectroscopy over a large field of view.
- **Coronagraph Instrument:** A technology demonstration instrument designed to perform high-contrast imaging and spectroscopy of exoplanets and circumstellar disks.

These advanced instruments enable the *Roman* Space Telescope to perform wide-field surveys with unprecedented sensitivity and resolution in the near-infrared regime.

5.5.4 Expected Scientific Contributions

The *Roman* Space Telescope is poised to make transformative contributions to astrophysics and cosmology. In addition to its primary surveys, it will offer opportunities for a *Guest Observer* program, allowing the broader scientific community to propose observations for a wide range of astrophysical phenomena [Akeson et al. \(2019\)](#). The mission is expected to:

- Provide precise measurements of cosmological parameters, refining our understanding of dark energy and the geometry of the Universe [Hounsell et al. \(2018\)](#).
- Expand the census of exoplanets, particularly those in the outer regions of planetary systems, thereby informing models of planetary formation and evolution [Penny et al. \(2019\)](#).
- Advance the study of galaxy formation and evolution through deep imaging and spectroscopy of distant galaxies [Behroozi & Silk \(2018\)](#).

- Enhance our understanding of the Milky Way's structure and stellar populations [Johnson et al. \(2020\)](#).

By addressing these fundamental questions, the *Roman* Space Telescope will significantly augment our knowledge of the cosmos and lay the groundwork for future astronomical discoveries.

Chapter 6

Simulations

In this section, we describe the N-body simulations used in this study, the process of generating convergence maps, the incorporation of shape and measurement noise, and the techniques employed for patch extraction in covariance analysis.

6.1 Initial condition

6.1.1 Primordial Power Spectrum

Primordial power spectrum is defined as:

$$P_p(k) = A \left(\frac{k}{k_*} \right)^{n_s - 1} \quad (6.1)$$

where A is the amplitude of the power spectrum, k is the wavenumber, k_* is the pivot scale, and n_s is the spectral index. The power spectrum describes the distribution of matter in the universe and is an essential ingredient in cosmological simulations. As the universe evolves, various physical processes (like radiation pressure, baryon-photon interactions, and dark matter dynamics) affect the growth of perturbations. These effects are encapsulated in the transfer function $T(k)$, which describes the evolution of the power spectrum from the early universe to the present day. The power spectrum at redshift z is given by

$$P(k; z) = P_p(k)T^2(k, z)D^2(z) \quad (6.2)$$

$$P(k; z = 0) = P_p(k)T^2(k, z = 0) \quad (6.3)$$

where $D(z)$ is the linear growth factor, which describes the growth of perturbations in the linear regime. The evolution of cosmological perturbations is governed by the Boltzmann equations for each species.

For example, for cold dark matter (CDM), the perturbation equation is given by

$$\ddot{\delta}_m + 2H\dot{\delta}_m - 4\pi G\rho_m\delta_m = 0 \quad (6.4)$$

where δ_m is the density contrast, H is the Hubble parameter, and ρ_m is the total matter density.

In a FLRW universe, the linear growth factor is given by:

$$D(a) = \frac{5\Omega_m a}{2} \int_0^1 \frac{da'}{a'^3 H(a')^3} \quad (6.5)$$

The transfer function $T(k)$ is computed using the CAMB code ([Lewis et al., 2000](#)).

6.1.2 Initial Conditions

Initial Density Field

Express the density contrast in terms of its Fourier components:

$$\delta(\mathbf{x}) = \int \frac{d^3 k}{(2\pi)^3} \tilde{\delta}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} \quad (6.6)$$

Assuming the density field is a Gaussian random field, each Fourier mode $\tilde{\delta}(\mathbf{k})$ is a complex Gaussian random variable with zero mean and variance $P(k)$:

$$\tilde{\delta}(\mathbf{k}) = A(\mathbf{k}) + iB(\mathbf{k}) \quad (6.7)$$

$$\langle A(\mathbf{k}) \rangle = \langle B(\mathbf{k}) \rangle = 0 \quad (6.8)$$

$$\langle A(\mathbf{k})A(\mathbf{k}') \rangle = \langle B(\mathbf{k})B(\mathbf{k}') \rangle = \frac{P(k)}{2} \delta_D(\mathbf{k} - \mathbf{k}') \quad (6.9)$$

$$\langle A(\mathbf{k})B(\mathbf{k}') \rangle = 0 \quad (6.10)$$

where δ_D is the Dirac delta function. The density field is then obtained by taking the inverse Fourier transform of $\delta(\mathbf{k})$.

Initial Displacement Field

Under the Zel'dovich approximation, the initial displacement field $\Psi(\mathbf{x})$ is linearly related to the initial density field $\delta(\mathbf{x})$:

$$\Psi(\mathbf{x}) = -\nabla\Phi(\mathbf{x}) \quad (6.11)$$

$$\nabla^2\Phi = \delta(\mathbf{x}) \quad (6.12)$$

In Fourier space, the displacement field is given by:

$$\begin{aligned} -k^2 \tilde{\Phi}(\mathbf{k}) &= \tilde{\delta}(\mathbf{k}) \\ \tilde{\Psi}(\mathbf{k}) &= -i\mathbf{k}\tilde{\Phi}(\mathbf{k}) = i\mathbf{k} \frac{\tilde{\delta}(\mathbf{k})}{k^2} \\ \Psi(\mathbf{x}) &= \int \frac{d^3 k}{(2\pi)^3} i\mathbf{k} \frac{\tilde{\delta}(\mathbf{k})}{k^2} e^{i\mathbf{k}\cdot\mathbf{x}} \end{aligned} \quad (6.13)$$

Initial Positions and Velocities

Assuming that particles start in a uniform grid in Lagrangian (initial) coordinates \mathbf{q} , the initial positions \mathbf{x} is given by:

$$\mathbf{x}(\mathbf{q}) = \mathbf{q} + \Psi(\mathbf{q}) \quad (6.14)$$

Under the Zel'dovich approximation, the initial velocities \mathbf{v} are proportional to the displacement field:

$$\begin{aligned} \mathbf{v}(\mathbf{x}) &= aHf\Psi(\mathbf{x}) \\ \tilde{\mathbf{v}}(\mathbf{k}) &= aHfi\mathbf{k} \frac{\tilde{\delta}(\mathbf{k})}{k^2} \\ \mathbf{v}(\mathbf{x}) &= iaHf \int \frac{d^3 k}{(2\pi)^3} \frac{\mathbf{k}}{k^2} \tilde{\delta}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} \end{aligned} \quad (6.15)$$

where $f = d \ln D / d \ln a$ is the growth rate of structure formation.

6.2 Simulation Basics

Numerical simulations are indispensable tools in physics and astronomy for investigating complex systems of interacting particles, such as galaxies, star clusters, and the large-scale structure of the Universe. The inherent complexity and nonlinearity of these systems render analytical solutions impractical or intractable, thereby necessitating the application of numerical methods. This section provides a comprehensive overview of N -body simulations commonly employed to study the large-scale structure of the cosmos.

To address the computational challenges posed by the long-range nature of gravitational interactions, significant efforts in numerical cosmology since the 1980s have focused on developing algorithms that reduce the need for global communication across the entire computational domain. These algorithms include mesh-based methods, tree codes, and multipole expansions ([Hockney & Eastwood, 1981](#)). Figure 6.1 illustrates the number of particles utilized in selected N -body simulations employing such techniques. The symbols and colors in the figure denote the gravitational solvers used: particle-particle-particle-mesh (P^3M) and adaptive P^3M (AP^3M); parallel or vectorized P^3M ; Tree codes; TreePM; and particle-mesh methods with adaptive mesh refinement (PM AMR).

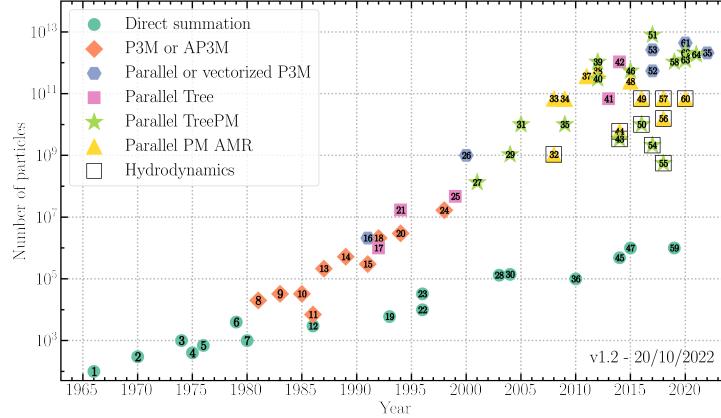


Figure 6.1: Evolution of the number of particles used in N -body simulations as a function of the year of publication (Leclercq, 2020). The symbols and colors indicate the gravitational solver employed: P^3M and adaptive P^3M (AP^3M); parallel or vectorized P^3M ; Tree codes; TreePM; and particle-mesh methods with adaptive mesh refinement (PM AMR). Hydrodynamic simulations are represented by black squares.

Owing to advancements in computational algorithms and software optimization, the number of particles in cosmological simulations has increased at a rate surpassing that achievable with direct summation methods. Notably, since 1990, a super-exponential growth trend has been observed for gravity-only simulations, as indicated by the quadratic regression in Figure 6.1. This acceleration cannot be solely attributed to improvements in computational hardware performance and reflects significant methodological innovations (Leclercq, 2020).

6.2.1 Direct Summation

Direct Summation computes gravitational forces between all pairs of particles directly. This method scales with $\mathcal{O}(N^2)$, making it computationally intensive for large N . Each particle i is characterized by its position \mathbf{r}_i , velocity \mathbf{v}_i , and mass m_i .

The algorithm proceeds through the following steps at each time step t :

1. Compute Forces:

$$\mathbf{F}_i = G m_i \sum_{\substack{j=1 \\ j \neq i}}^N \frac{m_j (\mathbf{r}_j - \mathbf{r}_i)}{\|\mathbf{r}_j - \mathbf{r}_i\|^3}$$

2. Update Particle States:

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i(t) + \frac{\mathbf{F}_i}{m_i} \Delta t, \quad \mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t + \Delta t) \Delta t$$

3. Advance Time:

$$t \leftarrow t + \Delta t$$

6.2.2 Particle-Mesh (PM) Method

Particle-Mesh (PM) methods approximate gravitational forces by mapping particles onto a grid and solving for the gravitational potential. This approach reduces the computational cost to $\mathcal{O}(N + M \log M)$, where M is the number of grid points. PM methods efficiently handle large-scale simulations but tend to smooth out small-scale forces, potentially sacrificing accuracy at smaller scales.

The algorithm proceeds through the following steps at each time step t :

1. **Assign Particles to Grid:** See Section 6.3.2.

2. **Compute Density Field:**

$$\rho(\mathbf{x}) = \sum_i m_i W(\mathbf{x} - \mathbf{r}_i) \quad (\text{where } W: \text{Interpolation Kernel})$$

3. **Solve Poisson's Equation:**

$$\nabla^2 \Phi = 4\pi G \rho$$

4. **Compute Gravitational Field:**

$$\mathbf{E} = -\nabla \Phi$$

5. **Update Particle States:**

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i(t) + \mathbf{E}_i \Delta t, \quad \mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t + \Delta t) \Delta t$$

6. **Advance Time:**

$$t \leftarrow t + \Delta t$$

6.2.3 Particle-Particle Particle-Mesh (P3M) Method

The Particle-Particle Particle-Mesh (P3M) method combines direct summation for short-range force calculations with the Particle-Mesh (PM) approach for long-range interactions. This hybrid technique maintains the $\mathcal{O}(N \log N)$ complexity of PM methods while achieving higher accuracy for close particle interactions by explicitly computing particle-particle (PP) forces. Key considerations in the P3M method include the selection of mesh size, the softening parameter ϵ , and the management of force resolution to balance computational efficiency and accuracy.

The algorithm proceeds through the following steps at each time step t :

1. **Long-Range Forces (PM):** Compute the long-range gravitational forces using the Particle-Mesh approach:

$$\mathbf{F}_{\text{long},i} = m_i \mathbf{E}_{\text{long}}(\mathbf{r}_i)$$

where \mathbf{E}_{long} is the gravitational field derived from the mesh-based potential.

2. Short-Range Forces (Direct Summation):

- (a) **Neighbor Search:** For each particle i , identify neighboring particles j within a cutoff radius r_{cut} .
- (b) **Force Calculation:** Compute the short-range gravitational force for each neighboring particle pair:

$$\mathbf{F}_{\text{short},i} = -Gm_i \sum_{j \in \text{neighbors}} \frac{m_j(\mathbf{r}_i - \mathbf{r}_j)}{\left(\|\mathbf{r}_i - \mathbf{r}_j\|^2 + \epsilon^2\right)^{3/2}}$$

3. Combine Forces:

Combine the long-range and short-range forces to obtain the total force on each particle:

$$\mathbf{F}_i = \mathbf{F}_{\text{long},i} + \mathbf{F}_{\text{short},i}$$

4. Update Particle States:

Update the velocity and position of each particle using the computed total force:

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i(t) + \frac{\mathbf{F}_i}{m_i} \Delta t, \quad \mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t + \Delta t) \Delta t$$

5. Advance Time:

Increment the simulation time by the time step Δt :

$$t \leftarrow t + \Delta t$$

6.2.4 Tree-Particle-Mesh (Tree-PM) Method

Tree-Particle-Mesh (Tree-PM) methods combine the Particle-Mesh (PM) approach for efficient long-range force computation with the tree algorithm for accurate short-range forces. This hybrid model reduces the computational complexity from $\mathcal{O}(N^2)$ to $\mathcal{O}(N \log N)$, enabling efficient, high-resolution simulations of large-scale structures. However, careful parameter tuning (e.g., grid size, softening length ϵ , and opening angle θ_{\max}) and adequate memory allocation for tree structures are crucial for optimal performance.

The algorithm proceeds through the following steps at each time step t :

1. Long-Range Forces (PM):

Compute the long-range gravitational forces using the Particle-Mesh approach:

$$\mathbf{F}_{\text{long},i} = m_i \mathbf{E}_{\text{long}}(\mathbf{r}_i)$$

where \mathbf{E}_{long} is the gravitational field derived from the mesh-based potential.

2. Short-Range Forces (Tree Algorithm):

- (a) **Tree Construction:** Build spatial cells (e.g., octree) and assign particles to the appropriate nodes.
- (b) **Multipole Moments:** Compute the mass M_j , center of mass $\mathbf{r}_{\text{cm},j}$, and higher-order multipole moments for each tree node j .
- (c) **Force Calculation:** For each particle i , traverse the tree to compute the short-range gravitational force:

$$\mathbf{F}_{\text{short},i} = -Gm_i \sum_{\text{nodes}} \frac{M_j(\mathbf{r}_i - \mathbf{r}_j)}{(\|\mathbf{r}_i - \mathbf{r}_j\|^2 + \epsilon^2)^{3/2}}$$

using the opening angle criterion:

$$\theta = \frac{l_j}{\|\mathbf{r}_i - \mathbf{r}_j\|} < \theta_{\max}$$

where l_j is the size of node j and θ_{\max} is the maximum allowed opening angle.

3. **Combine Forces:** Combine the long-range and short-range forces to obtain the total gravitational force on each particle:

$$\mathbf{F}_i = \mathbf{F}_{\text{long},i} + \mathbf{F}_{\text{short},i}$$

4. **Update Particle States:** Update the velocity and position of each particle using the computed total force:

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i(t) + \frac{\mathbf{F}_i}{m_i} \Delta t, \quad \mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t + \Delta t) \Delta t$$

5. **Advance Time:** Increment the simulation time by the time step Δt :

$$t \leftarrow t + \Delta t$$

6.3 Tools for Fast Computation

Efficient computational tools are essential for handling large-scale simulations and data analysis in scientific and engineering applications. This section provides an overview of key computational techniques and algorithms commonly employed in N -body simulations and large-scale structure studies.

6.3.1 Fast Fourier Transform

The Fast Fourier Transform (FFT) is a highly efficient algorithm for computing the Discrete Fourier Transform (DFT) of a sequence. Given a sequence of N complex numbers $\{x_n\}_{n=0}^{N-1}$, the DFT is defined as:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-2\pi i k n / N}, \quad k = 0, 1, \dots, N-1. \quad (6.16)$$

The naive computation of the DFT requires $\mathcal{O}(N^2)$ operations. The FFT reduces this complexity to $\mathcal{O}(N \log N)$ by exploiting the symmetry and periodicity properties of the exponential kernel. The most common FFT algorithm is the Cooley-Tukey radix-2 FFT ([Cooley & Tukey, 1965](#)), which recursively decomposes the DFT into smaller DFTs of even and odd-indexed elements:

$$X_k = \sum_{n=0}^{N/2-1} x_{2n} e^{-2\pi i k(2n)/N} + \sum_{n=0}^{N/2-1} x_{2n+1} e^{-2\pi i k(2n+1)/N} \quad (6.17)$$

$$= X_k^{\text{even}} + e^{-2\pi i k/N} X_k^{\text{odd}}, \quad (6.18)$$

where X_k^{even} and X_k^{odd} are the DFTs of the even and odd subsequences, respectively.

6.3.2 Mass Assignment Schemes

Mass assignment schemes are critical for mapping particle masses onto a computational grid to compute density fields and gravitational forces. These schemes ensure that the discretization process preserves important physical properties such as mass conservation and minimizes aliasing errors. Figure 6.2 illustrates the mass assignment process for a particle distribution on a 1D grid using different schemes.

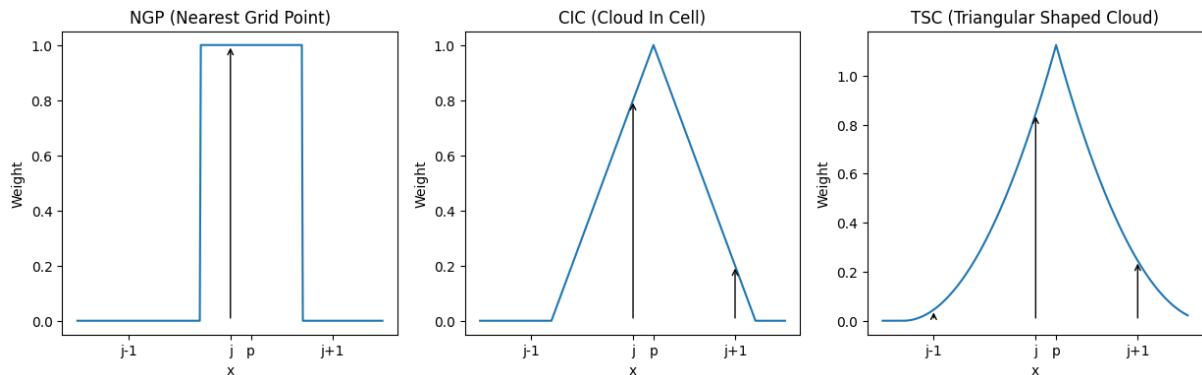


Figure 6.2: Illustration of three mass assignment schemes—Nearest Grid Point (NGP), Cloud-In-Cell (CIC), and Triangular-Shaped Cloud (TSC)—used to map a particle’s mass onto a 1D grid.

- **Nearest Grid Point (NGP):** Each particle is assigned entirely to the nearest grid point.
- **Cloud-In-Cell (CIC):** Mass is linearly interpolated to the nearest $2^3 = 8$ surrounding grid points.
- **Triangular-Shaped Cloud (TSC):** Mass is distributed to the nearest $3^3 = 27$ grid points using a quadratic interpolation function.

In Fourier space, these mass assignment window functions are represented as:

$$W(\mathbf{k}) = \prod_{i=1}^3 W(k_i), \quad (6.19)$$

where

$$W(k_i) = \left[\frac{\sin(\pi k_i / (2k_N))}{\pi k_i / (2k_N)} \right]^p, \quad (6.20)$$

with k_N being the Nyquist wavenumber, k_i the i -th component of the wavevector \mathbf{k} , and $p = 1$ for NGP, $p = 2$ for CIC, and $p = 3$ for TSC ([Hockney & Eastwood, 1981](#); [Cui et al., 2008](#)).

6.3.3 Parallelization Techniques

Parallelization techniques are pivotal for accelerating computations in large-scale simulations by leveraging multiple processors or computing nodes. The primary strategies include:

- **Domain Decomposition:** The computational domain is partitioned into smaller subdomains, each assigned to a separate processor ([Barnes & Hut, 1986](#)).
- **Task Parallelism:** Independent tasks are distributed across multiple processors.
- **Data Parallelism:** Identical operations are performed concurrently on different data elements, enabling SIMD (Single Instruction, Multiple Data) execution.

These parallelization strategies can be combined to maximize computational efficiency.

6.3.4 Adaptive Mesh Refinement

Adaptive Mesh Refinement (AMR) dynamically adjusts the resolution of the computational grid in regions requiring higher accuracy. The primary goal of AMR is to allocate computational resources efficiently by refining the mesh where necessary and coarsening it elsewhere. AMR generates a hierarchy of refined grids, where each level has finer resolution ([Berger & Colella, 1989](#)).

Refinement is typically based on physical quantities such as density gradients. For a scalar field $\phi(\mathbf{x})$, refinement is triggered where:

$$|\nabla \phi(\mathbf{x})| > \theta, \quad (6.21)$$

with θ being a predefined threshold. Figure 6.3 demonstrates the application of Adaptive Mesh Refinement (AMR) to a two-dimensional image containing two Gaussian kernels. Initially, a uniformly coarse grid overlays the entire image (left panel). As the refinement process progresses, the mesh becomes increasingly finer in regions with higher intensity, specifically around the Gaussian kernels (middle and right panels). The red grid lines represent the hierarchy of the refined meshes, enabling higher resolution where it is most needed and optimizing computational resources by keeping a coarser grid in less significant areas.

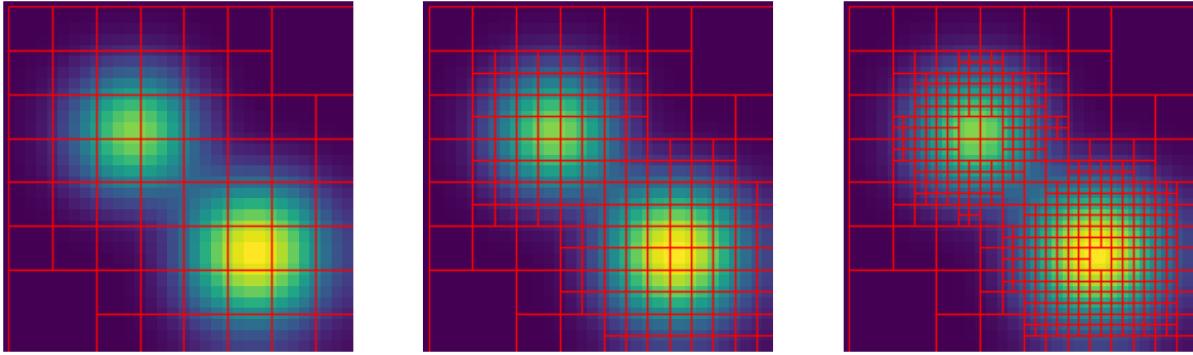


Figure 6.3: Illustration of adaptive mesh refinement (AMR) applied to a 2D image with two Gaussian kernels. The left panel shows the initial coarse grid structure over the image. The middle and right panels demonstrate progressively finer levels of mesh refinement in regions of higher intensity, where the Gaussian kernels are located. The red grid outlines indicate the adaptively refined mesh hierarchy, ensuring higher resolution where needed while maintaining computational efficiency in lower-intensity regions.

6.3.5 Tree Construction

Tree-based data structures are fundamental for efficiently organizing and querying hierarchical spatial data. In computational simulations, the Barnes-Hut Octree is commonly used for tasks like force calculations in N -body simulations ([Barnes & Hut, 1986](#)).

Barnes-Hut Octree

The Barnes-Hut algorithm employs an Octree to hierarchically partition the simulation space. Each node represents a cubic region, recursively subdivided into eight octants if it contains more than one particle.

The force on a particle is computed by approximating distant clusters of particles as single mass points. This approximation is controlled by the opening angle θ :

$$\frac{s}{d} < \theta, \quad (6.22)$$

where s is the size of the node, and d is the distance from the particle to the node's center of mass. This method reduces the computational complexity from $\mathcal{O}(N^2)$ to $\mathcal{O}(N \log N)$. Figure 6.4 illustrates the application of the Barnes-Hut Octree algorithm to partition a three-dimensional simulation space containing four particles.

Parallel Implementation

Parallel tree construction involves building local trees within each subdomain and integrating them for global computations ([Dubinski, 1996](#)). Efficient parallelization enhances scalability and performance in large-scale simulations.

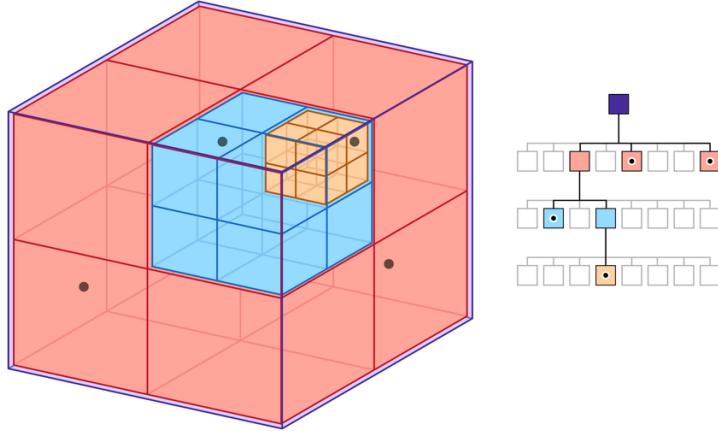


Figure 6.4: Illustration of an Octree decomposition for a 3D volume containing four particles. The left panel showcases the spatial subdivision of the volume into hierarchical grid cells, with color-coding indicating different levels of refinement—red for coarse, blue for intermediate, and orange for fine cells where particles reside. The right panel presents the corresponding Octree data structure, highlighting the hierarchical relationships between nodes. The root node (purple) represents the entire simulation volume, while the second layer comprises various node types: empty nodes are depicted in white, single-particle nodes in orange-pink with a black dot, and multi-particle nodes in orange-pink without a dot, which require further subdivision. This hierarchical decomposition enables efficient computation by concentrating resolution in regions with higher particle density while conserving computational resources in less populated areas. Credit by [Powell et al. \(2023\)](#)

6.4 FASTPM

FASTPM (Fast Particle Mesh; [Feng et al. 2016](#)) is an advanced N-body simulation code tailored for efficiently modeling the evolution of dark matter and halo structures on cosmological scales. Building upon the foundational Particle-Mesh (PM) approach, FASTPM integrates modified kick and drift factors derived from the Zel'dovich Approximation (ZA). This enhancement allows FASTPM to achieve high accuracy in large-scale structure formation while significantly reducing computational overhead. This subsection delineates the core methodology of FASTPM, incorporating the mathematical formalism of its modified kick and drift factors.

6.4.1 Modified Kick and Drift Factors

The cornerstone of FASTPM's enhanced performance lies in its **modified kick** (K_{FASTPM}) and **drift** (D_{FASTPM}) factors. These factors are meticulously derived from the Zel'dovich Approximation (ZA), a first-order Lagrangian perturbation theory (1LPT), to rectify inaccuracies in large-scale growth inherent in standard PM solvers, especially when operating with a limited number of time steps.

First, the Zel'dovich equation of motion to the first order is defined as:

$$\begin{aligned}\mathbf{x}_{\text{ZA}}(a) &= \mathbf{q} + D(a)\mathbf{s}_1, \\ \mathbf{p}_{\text{ZA}}(a) &= a^3 E(a) g_p(a) \mathbf{s}_1, \\ \mathbf{f}_{\text{ZA}}(a) &= a^2 E(a) g_f(a) \mathbf{s}_1,\end{aligned}\quad (6.23)$$

where $E(a) = \frac{H(a)}{H(a=1)}$ is the dimensionless Hubble parameter, and $g_p(a)$ and $g_f(a)$ are auxiliary factors defined as:

$$g_p(a) = \frac{dD}{da}, G_p(a) = D(a) \quad (6.24)$$

$$g_f(a) = \frac{d(a^3 E g_p)}{da}, G_f(a) = a^3 E g_p(a) \quad (6.25)$$

The ZA equations of motion are reformulated in terms of drift and kick operators by integrating over a time step from a_0 to a_1 and eliminating the ZA displacement \mathbf{s}_1 :

$$\begin{aligned}\Delta \mathbf{x}_{\text{ZA}} &= \mathbf{x}_{\text{ZA}}(a_1) - \mathbf{x}_{\text{ZA}}(a_0) \\ &= [D(a)]_{a_0}^{a_1} \mathbf{s}_1 \\ &= \frac{\mathbf{p}_{\text{ZA}}(a_r)}{a_r^3 E(a_r)} \left(\frac{\Delta G_p}{g_p(a_r)} \right),\end{aligned}\quad (6.26)$$

$$\begin{aligned}\Delta \mathbf{p}_{\text{ZA}} &= \mathbf{p}_{\text{ZA}}(a_1) - \mathbf{p}_{\text{ZA}}(a_0) \\ &= \frac{\mathbf{f}_{\text{ZA}}(a_r)}{a_r^2 E(a_r)} \left(\frac{\Delta G_f}{g_f(a_r)} \right),\end{aligned}\quad (6.27)$$

where $\Delta \mathbf{x}_{\text{ZA}}$ is the change in displacement over the time step, $\Delta \mathbf{p}_{\text{ZA}}$ is the change in momentum over the time step, a_r is a reference scale factor within the time step, $\Delta G_p = G_p(a_1) - G_p(a_0)$, and $\Delta G_f = G_f(a_1) - G_f(a_0)$. Therefore, the modified kick and drift factors in FASTPM are defined as:

$$\mathcal{D}_{\text{FASTPM}} = \frac{\Delta \mathbf{x}_{\text{ZA}}}{\mathbf{p}_{\text{ZA}}} = \frac{1}{a_r^3 E(a_r)} \left(\frac{\Delta G_p}{g_p(a_r)} \right) \quad (6.28)$$

$$\mathcal{K}_{\text{FASTPM}} = \frac{\Delta \mathbf{p}_{\text{ZA}}}{\mathbf{f}_{\text{ZA}}} = \frac{1}{a_r^2 E(a_r)} \left(\frac{\Delta G_f}{g_f(a_r)} \right) \quad (6.29)$$

These operators ensure the exact integration of the ZA equations of motion, thereby accurately capturing the linear growth of structures within each time step.

6.4.2 Algorithm Steps

For each simulation time step t , FASTPM executes the following sequence of computational procedures to update particle positions and velocities accurately:

1. **Assign Particles to Grid:** Particles are assigned to a computational grid using an interpolation kernel $W(\mathbf{x} - \mathbf{r}_i)$. This step transforms the discrete particle distribution into a continuous density field suitable for solving Poisson's equation.

$$\rho(\mathbf{x}) = \sum_i m_i W(\mathbf{x} - \mathbf{r}_i)$$

2. **Compute Density Field:** Utilizing the assigned grid, compute the density field $\rho(\mathbf{x})$ by summing the contributions of all particles through the interpolation kernel.
3. **Solve Poisson's Equation:** Solve Poisson's equation on the grid to obtain the gravitational potential $\Phi(\mathbf{x})$:

$$\nabla^2 \Phi = 4\pi G \rho$$

4. **Compute Gravitational Forces:** Calculate the gravitational acceleration $\mathbf{g}(\mathbf{x})$ by taking the gradient of the potential:

$$\mathbf{g}(\mathbf{x}) = -\nabla \Phi$$

5. **Apply Modified Operators:** Utilize the modified kick (K_{FASTPM}) and drift (D_{FASTPM}) factors to update particle velocities and positions. These factors, derived from the ZA, ensure accurate linear growth:

- (a) **Kick Step:** Update particle velocities by applying the gravitational acceleration scaled by the modified kick factor:

$$\mathbf{v}_i \left(t + \frac{\Delta t}{2} \right) = \mathbf{v}_i(t) + \mathbf{g}_i(t) \cdot K_{\text{FASTPM}} \cdot \Delta t$$

- (b) **Drift Step:** Update particle positions using the updated velocities and the modified drift factor:

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i \left(t + \frac{\Delta t}{2} \right) \cdot D_{\text{FASTPM}} \cdot \Delta t$$

- (c) **Second Kick Step:** Apply another kick to update velocities to the full time step:

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i \left(t + \frac{\Delta t}{2} \right) + \mathbf{g}_i(t + \Delta t) \cdot K_{\text{FASTPM}} \cdot \Delta t$$

6. **Update Particle States:** Finalize the update of particle velocities and positions after applying the modified kick and drift operators:

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i \left(t + \frac{\Delta t}{2} \right) + \mathbf{g}_i(t + \Delta t) \cdot K_{\text{FASTPM}} \cdot \Delta t$$

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i \left(t + \frac{\Delta t}{2} \right) \cdot D_{\text{FASTPM}} \cdot \Delta t$$

7. **Advance Time:** Increment the simulation time by the time step Δt :

$$t \leftarrow t + \Delta t$$

Chapter 7

Methods

7.1 Dataset Overview

To achieve high accuracy while minimizing computational time, we employed the FASTPM particle-mesh simulation code ([Feng et al., 2016](#)). This code enhances convergence by adjusting kick and drift factors to closely adhere to the Zel'dovich approximation, thereby improving the efficiency of structure formation modeling.

In our study, we quantify super-sample covariance in higher-order statistics using two N-body simulations: **BIGBOX** and **TILED**. The **BIGBOX** simulation, conducted as part of the HalfDome project ([Bayer et al., 2024](#)), encompasses a large cosmic volume, whereas the **TILED** simulation represents a smaller volume with the same resolution as BIGBOX, specifically excluding large-scale modes. The cosmological parameters for both simulations align with those of IllustrisTNG ([Nelson et al., 2019](#)), as listed in Table 7.1.

| Parameter | Symbol | Value |
|---------------------------|------------|---|
| Hubble constant | H_0 | 67.74 [km s ⁻¹ Mpc ⁻¹] |
| Matter density | Ω_m | 0.3089 |
| Baryon density | Ω_b | 0.0486 |
| Amplitude of fluctuations | σ_8 | 0.8159 |
| Spectral index | n_s | 0.9667 |
| Sum of neutrino masses | M_ν | 0.0 [eV] |

Table 7.1: Cosmological parameters used in the N-body simulations.

The **BIGBOX** simulation models an extensive cubic volume of $L = 3750$ Mpc/h with 6144^3 particles, enabling detailed capture of large-scale structures. From this data, a full-sky map was generated by first covering an octant of the sky and subsequently extending it to a full-sky projection.

The **TILED** simulation covers a smaller volume of $L = 625$ Mpc/h, populated with 1024^3 particles. To achieve resolution parity with the BIGBOX simulation, multiple smaller volumes were tiled together, forming a full-sky map comparable in detail to the BIGBOX simulation but without the inclusion of large-scale modes.

Both simulations commence at an initial redshift of $z = 9$, utilizing an initial linear matter power spectrum at $z = 0$ generated via the CLASS code (Blas et al., 2011). This setup ensures consistency with observational data of the early universe. We evolved the simulations over 60 time steps, reaching the present day ($z = 0$), thereby capturing the non-linear growth of cosmic structures—an aspect crucial for studies of weak gravitational lensing.

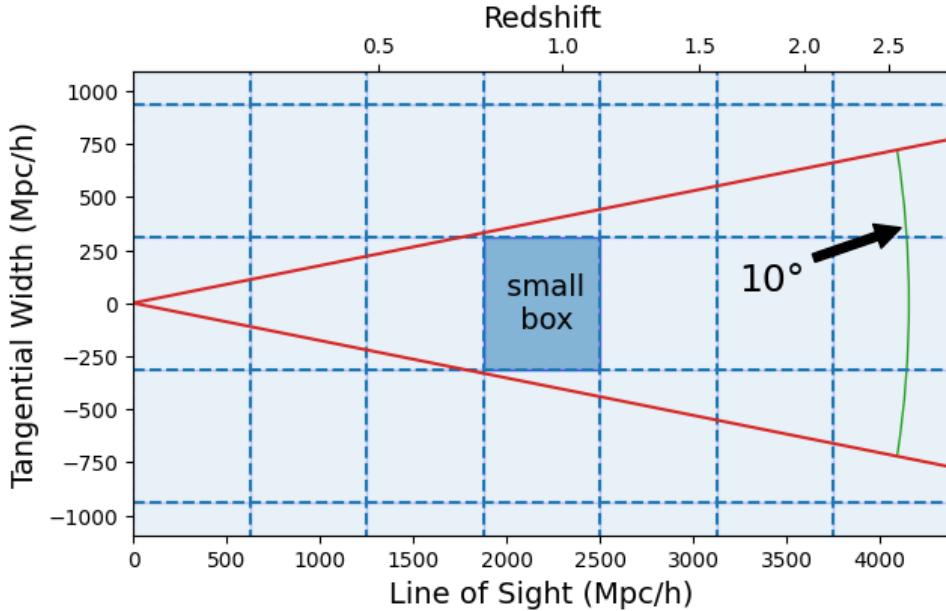


Figure 7.1: Spatial and redshift setup for the **BIGBOX** and **TILED** simulations. The left side of the figure features red lines delineating the light cone boundaries, covering a 10° field of view on the sky. Overlaid on this are dashed blue grids that partition the overall simulation volume into smaller, manageable tiling regions. Within these grids, the "small box" highlights a specific tiled region that resides inside the extensive BIGBOX volume. The horizontal axis represents the line-of-sight distance measured in comoving megaparsecs per h , with the corresponding redshift values displayed on the top axis.

Figure 7.1 showcases the spatial and redshift setup for the **BIGBOX** and **TILED** simulations used in cosmological studies. The left side of the figure features red lines delineating the light cone boundaries, covering a 10° field of view on the sky. Overlaid on this are dashed blue grids that partition the overall simulation volume into smaller, manageable tiling regions. Within these grids, the "small box" highlights a specific tiled region that resides inside the extensive BIGBOX volume. The horizontal axis represents the line-of-sight distance measured in comoving megaparsecs per h , with the corresponding redshift values displayed on the top axis.

7.2 Generating Convergence Maps

To simulate the weak gravitational lensing effect observed in surveys, we generated convergence maps from our N-body simulations. Light cones were constructed from these simulations to model the observable universe, with particles inserted on-the-fly at the appropriate redshifts through interpolation between

N-body time steps.

To balance computational efficiency with the need to accurately capture lensing effects, we set the width of each radial shell to $\Delta a = 0.01$, corresponding to a comoving distance of approximately $\Delta \chi \approx 100 h^{-1} \text{ Mpc}$.

For each shell, the three-dimensional matter density $\delta(\mathbf{x}, z_i)$ was projected onto a two-dimensional plane perpendicular to the line of sight. The projected surface density $\Sigma(\hat{\mathbf{n}}, \chi_i)$ at an angular position $\hat{\mathbf{n}}$ was computed by integrating the matter density within the shell along the radial direction:

$$\Sigma(\hat{\mathbf{n}}, \chi_i) = \int_{\chi_i}^{\chi_{i+1}} \delta(\chi \hat{\mathbf{n}}, z(\chi)) d\chi. \quad (7.1)$$

In practice, the surface density was mapped onto a HEALPix grid (Górski et al., 2005) to create a full-sky map $\Sigma(n_j, \chi_i)$, where n_j represents discretized angular positions. The HEALPix grid resolution was set to $N_{\text{side}} = 8192$, providing an angular resolution of approximately 0.43 arcminutes, which is sufficient to capture small-scale structures relevant to weak lensing studies.

The convergence $\kappa(n_j; z_s)$ at each pixel of the HEALPix grid was then obtained by summing contributions from all the shells up to the source redshift:

$$\kappa(n_j; z_s) = \sum_i W(\chi_i, \chi_s) \Sigma(n_j, \chi_i) \Delta \chi_i, \quad (7.2)$$

where $W(\chi, \chi_s)$ is the lensing efficiency function. The effective comoving distance to the center of the j -th shell, χ^j , is calculated as (Shirasaki et al., 2015):

$$\chi^j = \frac{\int_{\chi_{\min}}^{\chi_{\max}} \chi^3 d\chi}{\int_{\chi_{\min}}^{\chi_{\max}} \chi^2 d\chi} = \frac{3}{4} \frac{\chi_{\max}^4 - \chi_{\min}^4}{\chi_{\max}^3 - \chi_{\min}^3}, \quad (7.3)$$

The density contrast within the i -th pixel of the j -th shell, $\delta^j(\hat{n}_i)$, is determined by:

$$\delta^j(\hat{n}_i) = \frac{n_{\text{part},i}^j}{\bar{n}_{\text{part}}^j} - 1, \quad (7.4)$$

where $n_{\text{part},i}^j$ is the number of particles in the i -th pixel of the j -th shell, and \bar{n}_{part}^j is the average number of particles per pixel in that shell.

We considered source redshifts z_s from 0.5 to 2.5 in increments of 0.5, covering the range of distances relevant for current and future galaxy surveys, such as DES, LSST, *Euclid*, and *Roman*.

Figure 7.2 presents the normalized lensing efficiency as a function of comoving distance (measured in Mpc/h) for multiple source redshifts (z). Each curve corresponds to a specific source redshift, as detailed in the legend. The lensing efficiency curves exhibit peaks at intermediate comoving distances, indicating the regions where the distribution of matter along the line of sight most significantly enhances

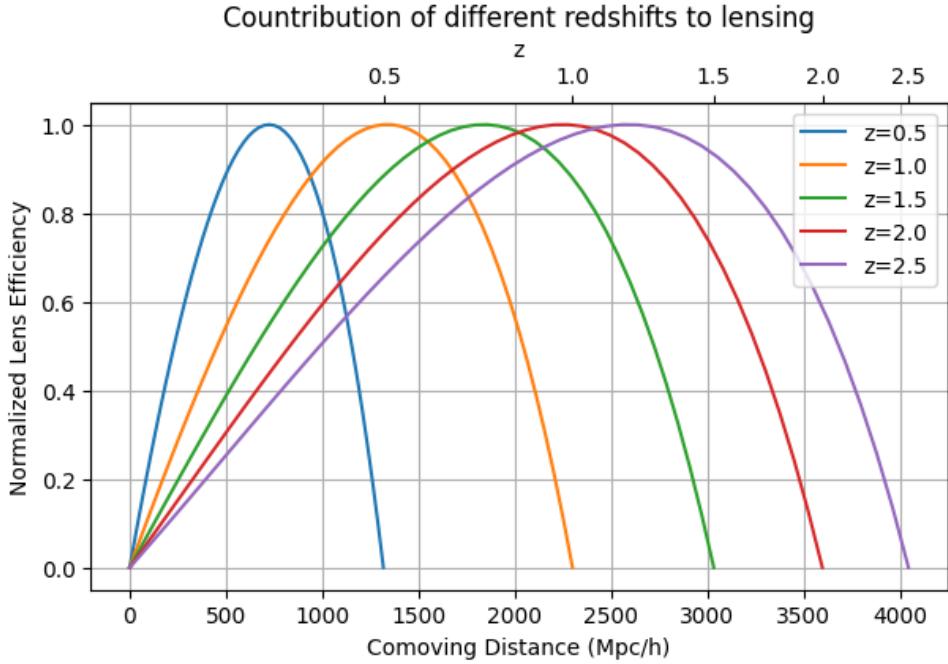


Figure 7.2: Normalized lensing efficiency as a function of comoving distance for multiple source redshifts. The lensing efficiency peaks at intermediate comoving distances, indicating regions where the distribution of matter enhances the gravitational lensing signal.

the gravitational lensing signal.

7.3 Incorporating Noise

In real observations, measurements of the lensing signal are contaminated by noise arising from the intrinsic shapes of galaxies and errors in shape measurements. This noise, referred to as shape noise, constitutes a significant source of uncertainty, particularly on small angular scales.

To simulate shape noise, we added Gaussian noise to our convergence maps. We considered four different surveys with varying galaxy number densities, as detailed in Table 7.2.

| Survey | Galaxy Number Density [arcmin $^{-2}$] |
|---------------------|---|
| DES/KiDS | 7 |
| HSC | 15 |
| <i>Euclid</i> /LSST | 30 |
| <i>Roman</i> | 50 |

Table 7.2: Galaxy number densities for different surveys used to model shape noise levels.

The variance of the shape noise per pixel was calculated as:

$$\sigma_{\kappa, \text{noise}}^2 = \frac{\sigma_e^2}{2n_{\text{gal}}A_{\text{pix}}}, \quad (7.5)$$

where σ_e is the intrinsic ellipticity dispersion of galaxies, set to $\sigma_e = 0.26$ ([Euclid Collaboration et al.](#),

2019), n_{gal} is the galaxy number density per square arcminute, and A_{pix} is the solid angle of a pixel, set to 0.43 arcminutes². We generated a Gaussian random field $n(\hat{\mathbf{n}})$ with the calculated variance and added it to the convergence maps:

$$\kappa_{\text{obs}}(\hat{\mathbf{n}}) = \kappa(\hat{\mathbf{n}}) + n(\hat{\mathbf{n}}). \quad (7.6)$$

7.4 Patch Extraction for Analysis

In order to simplify the analysis onto a flat patch, we extracted patches from the full-sky convergence maps. Each patch covers an area of $10^\circ \times 10^\circ$ and is uniformly distributed across the sky using a Fibonacci grid (Swinbank & James Purser, 2006; Ferlito et al., 2023). The center of each patch is positioned at the vertices of the Fibonacci grid defined by golden ratio spirals:

$$\sin \theta_i = \frac{2i}{2N+1}, \quad \phi_i = \frac{2\pi i}{\varphi}, \quad -N \leq i \leq N, \quad -\frac{\pi}{2} \leq \theta_i \leq \frac{\pi}{2}, \quad (7.7)$$

where N is the number of patches and $\varphi = (1 + \sqrt{5})/2$ is the golden ratio.

The number of patches, denoted N_{patches} , was optimized to ensure that individual patches do not overlap, except in regions near the poles where overlapping patches were subsequently discarded. The optimization process commenced with an initial count of $N_{\text{patches}} = 400$ and involved iteratively reducing this number until a configuration was achieved wherein the patches remained non-overlapping, except for centers located within $10\sqrt{2}^\circ$ deg of the poles, that is $|\theta_i| \geq 10\sqrt{2}^\circ$ and $|\phi_i| \leq \pi - 10\sqrt{2}^\circ$. Additionally, patches include points heavily tiled along with line of sight and near the equator are excluded to avoid severe Box Replication Effect (see Sec. ?? for further check). After optimization and masking, the number of patches was set to $N_{\text{patches}} = 273$, effectively reducing to $N_{\text{patches}} = 262$, effectively cover 64% of the sky. The visualization of the Fibonacci grid is shown in Figure 7.3.

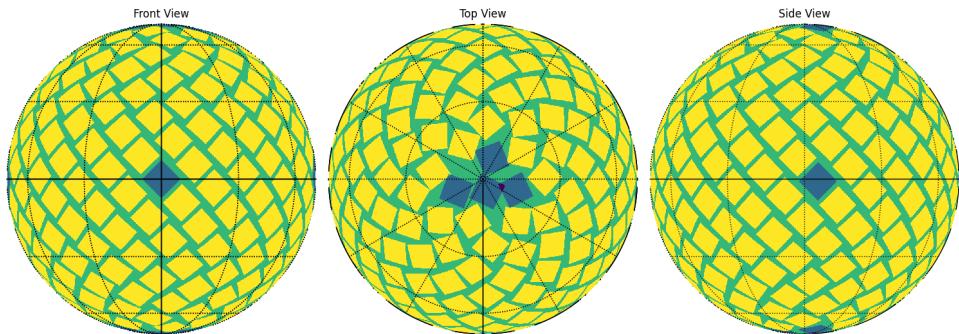


Figure 7.3: Visualization of the Fibonacci grid with $N_{\text{patches}} = 273$ patches, each covering approximately $10 \times 10 \text{ deg}^2$. After the optimization and masking, the number of patches is reduced to $N_{\text{patches}} = 262$, effectively covering 64% of the sky. Each panel show the patches distribution on the Front, Top and Side view.

To verify the absence of overlap between patches, we calculated the vertices of each patch and conducted pairwise overlap checks with all other patches. For a Fibonacci grid center characterized by

coordinates (θ_i, ϕ_i) , the vertices of the corresponding patch are defined as:

$$\begin{aligned} & (\theta_i - \Delta\theta, \phi_i - \Delta\phi), \quad (\theta_i - \Delta\theta, \phi_i + \Delta\phi), \\ & (\theta_i + \Delta\theta, \phi_i - \Delta\phi), \quad (\theta_i + \Delta\theta, \phi_i + \Delta\phi), \end{aligned} \quad (7.8)$$

where

$$\Delta\theta = 5\sqrt{2} \text{ deg}, \quad \Delta\phi = 5\sqrt{2} \sin \theta_i \text{ deg.} \quad (7.9)$$

Using the vertices of the Fibonacci grid as centers, we employed the `gnomview` function from the `healpy` library ([Zonca et al., 2019](#)) to project each spherical patch onto a flat plane via a gnomonic projection. Each patch is represented by a 2048×2048 grid of pixels, resulting in a pixel size of:

$$\Delta\theta = \frac{10^\circ}{2048} \approx 0.00488^\circ \approx 0.293' \quad \text{per pixel.} \quad (7.10)$$

7.5 Gaussian Smoothing

Shape noise predominantly affects small angular scales. To mitigate this noise and enhance the detection of the underlying lensing signal, we applied Gaussian smoothing to the noisy convergence maps. The Gaussian filter used is defined by:

$$W(\theta) = \frac{1}{\pi\theta_G^2} \exp\left(-\frac{\theta^2}{\theta_G^2}\right), \quad (7.11)$$

where θ is the angular distance from the center of the filter, and θ_G is the smoothing scale. For our analysis, we selected $\theta_G = 2'$, $5'$, $8'$, and $10'$.

By convolving the noisy convergence map with the Gaussian filter, we obtained the smoothed convergence map:

$$\kappa_{\text{smoothed}}(\hat{\mathbf{n}}) = \int d\Omega' W(|\hat{\mathbf{n}} - \hat{\mathbf{n}}'|) \kappa_{\text{obs}}(\hat{\mathbf{n}}'). \quad (7.12)$$

Figure 7.4 demonstrates the application of Gaussian smoothing to a noisy convergence map. The figure presents four panels, each corresponding to a different smoothing scale: $\theta_G = 2'$, $5'$, $8'$, and $10'$. As the smoothing scale increases, the convolution of the convergence map with the Gaussian filter effectively reduces small-scale noise, as evidenced by the diminishing high-frequency fluctuations in the map.

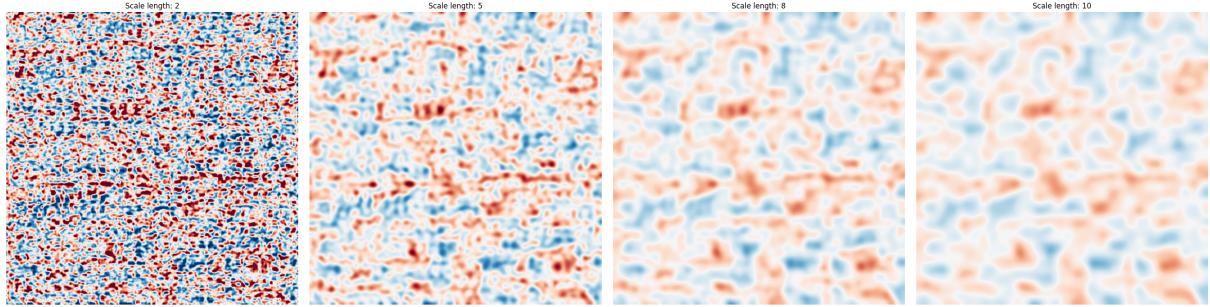


Figure 7.4: Effect of Gaussian smoothing on a noisy convergence map. Each panel shows the result of applying a Gaussian filter with a different smoothing scale $\theta_G = 2'$, $5'$, $8'$, and $10'$. As the smoothing scale increases, small-scale noise is progressively suppressed, and large-scale structures become more prominent. This demonstrates how Gaussian smoothing effectively reduces shape noise while enhancing the detection of the underlying lensing signal.

7.6 Measurements

In order to characterize the influence of super-sample covariance on higher-order statistics, this study concentrates on the bispectrum, probability distribution function (PDF), peak counts, minima counts, and Minkowski functionals. These statistical measures offer complementary insights into the underlying matter distribution and exhibit sensitivity to distinct features of the gravitational lensing signal.

7.6.1 Statistical Measures and Computational Methods

Table 7.3 delineates the range of values and the computational subroutines employed for each statistical measure. Each statistic is computed both for full-sky analyses and sky patches, utilizing the appropriate methodologies as specified. The angular power spectrum, denoted as $C_\ell^{\kappa\kappa}$, alongside three configurations

| Statistic | Range | Subroutine (Sky Patch) |
|---|---------------------------------------|---|
| Angular Power Spectrum | $300 \leq \ell \leq 3000$ | <code>lenstools.powerSpectrum</code> |
| Bispectrum | $300 \leq \ell \leq 3000$ | <code>lenstools.bispectrum</code> |
| Peak Counts | $-4 \leq \kappa/\sigma_\kappa \leq 4$ | <code>lenstools.peakCount</code> |
| Minima Counts | $-4 \leq \kappa/\sigma_\kappa \leq 4$ | <code>lenstools.peakCount</code> |
| Probability Distribution Function (PDF) | $-4 \leq \kappa/\sigma_\kappa \leq 4$ | <code>lenstools.pdf</code> |
| Minkowski Functionals | $-4 \leq \kappa/\sigma_\kappa \leq 4$ | <code>lenstools.minkowskiFunctionals</code> |

Table 7.3: Summary of the statistical measures employed in this investigation, including their respective value ranges and computational subroutines utilized for both full-sky and sky-patch analyses.

of the bispectrum, $B_{\ell_1 \ell_2 \ell_3}^{\kappa\kappa\kappa}$, are derived from unsmoothed convergence maps. The bispectrum calculations encompass three distinct configurations: equilateral ($\ell_1 = \ell_2 = \ell_3$), squeezed ($\ell_1 = \ell_2 = 10\ell_3$), and isosceles ($\ell_1 = \ell_2 = 2\ell_3$). All bispectrum and angular power spectrum computations are confined within the multipole range $\ell \in [300, 3000]$, consistent with the multipole selection in the HSC Y3 cosmic shear analysis (Dalal et al., 2023). We adopt a logarithmic binning approach to effectively sample the range of scales, dividing the multipole interval into 8 bins that are evenly spaced in logarithmic space.

Conversely, the PDF, peak counts, minima counts, and Minkowski functionals are derived from smoothed convergence maps, where the smoothing angle is fixed at 2 arcminutes for the primary results. These measurements are conducted within the normalized range $-4 \leq \kappa/\sigma_\kappa \leq 4$, linearly divided into 8 bins following [Martinet et al. \(2021\)](#). σ_κ denotes the standard deviation of each patch's convergence map.

All statistical computations are performed using the `lenstools` package ([Petri, 2016](#)).

7.6.2 Covariance Matrix Estimation

Following the measurement phase, this study examines the influence of super-sample covariance on the covariance matrices associated with the aforementioned statistical measures. To achieve this, we employ an unbiased estimator for the covariance matrix as previously defined in Equation 4.2. The analysis utilizes 11 realizations from the BIGBOX simulation and 20 realizations from the TILED simulation. For each realization, the covariance is computed using 262 patches extracted from the full-sky map of each simulation. Therefore, we obtain a total of 2882 patches from the BIGBOX simulation and 5240 patches from the TILED simulation.

Additionally, we also compute the correlation matrix for each statistical measure to investigate the interdependence between different scales and configurations. The correlation matrix is defined as:

$$\rho_{ij} = \frac{\text{Cov}(\mathcal{O}_i, \mathcal{O}_j)}{\sqrt{\text{Cov}(\mathcal{O}_i, \mathcal{O}_i)\text{Cov}(\mathcal{O}_j, \mathcal{O}_j)}}, \quad (7.13)$$

where \mathcal{O}_i and \mathcal{O}_j represent the i -th and j -th statistical measures, respectively.

After the covariance and correlation matrices are computed for both the BIGBOX and TILED simulations, we compare the matrices to quantify the impact of super-sample covariance on the statistical measures. The comparison is conducted by calculating the ratio between the covariance and correlation matrices of the BIGBOX and TILED simulations. This ratio provides a direct measure of the influence of super-sample covariance on the statistical measures, enabling a comprehensive assessment of the impact on the covariance and correlation matrices. In the case of correlation matrices, we exclude the diagonal elements—since they are always unity—when calculating the ratio. For covariance matrices, the ratio is calculated over the entire matrix. Regarding ℓ -binned statistics, we determine the ratio across all bins. For ν -binned statistics, we exclude the first and last bins from the ratio calculation due to their limited data points and unreliability.

Chapter 8

Results

8.1 Overview

In this chapter, we initiate a comprehensive analysis by systematically comparing each statistical measure derived from our simulations. We focus on evaluating the mean values, covariance matrices, and correlation matrices obtained from the *BIGBOX* and *TILED* simulations to assess their consistency and understand the underlying discrepancies.

Figures 8.1 through 8.12 provide detailed visualizations of the mean values, variances, covariance matrices, and correlation matrices for each statistical measure under consideration. For each statistic, one figure illustrates the comparison of mean values and variances, while another figure presents the comparison of covariance and correlation matrices. Due to limitations in space, we have included only the covariance matrix comparisons for the bispectrum and Minkowski Functionals.

From these figures, we observe that the mean values of most statistical measures exhibit excellent agreement between the *BIGBOX* and *TILED* simulations, with differences remaining below 1% across the majority of the studied range. However, notable deviations occur at low ν values for peak counts, minima, and the Minkowski Functionals V_1 and V_2 . These deviations are attributed to the limited resolution of the simulations, which affects the accurate detection of regions with the lowest density contrasts.

Analyzing the covariance matrices reveals that, except for the bispectrum, the ratios of covariance matrix elements between the *BIGBOX* and *TILED* simulations are consistently greater than unity. This indicates that the *BIGBOX* simulations yield higher covariance values compared to the *TILED* simulations, and this discrepancy becomes more pronounced at higher source redshifts. The bispectrum, on the other hand, exhibits noisy covariance matrices without a clear trend, making it challenging to draw definitive conclusions for this statistic.

Examining the correlation matrices further, we focus on the off-diagonal elements to assess the degree of inter-bin correlations. For statistical measures that are not inherently correlated, the off-diagonal elements remain close to unity, as expected. In contrast, the power spectrum shows off-diagonal elements that exceed unity, displaying a clear increasing trend with higher source redshifts. This behavior

Takada & Hu (2013), suggesting that larger-scale modes beyond the survey volume contribute to the observed correlations.

Overall, these findings support the hypothesis that super-sample covariance significantly impacts the statistical measures derived from our simulations. The discrepancies observed between the BIGBOX and TILED simulations emphasize the importance of considering super-sample effects in cosmological analyses. We will explore these effects in greater depth and seek further validation in the subsequent discussion chapter.

8.2 Effects of Noise

To assess the impact of observational noise, we have introduced five different shape noise levels into the simulations. Due to the significant influence of noise on higher-order statistics, the bispectrum has been excluded from this part of the analysis.

Figures [8.13](#) and [8.14](#) illustrate how the average ratios of covariance matrices and correlation matrices change with varying shape noise levels. Except for the angular power spectrum, the non-Correlation statistics exhibit stable covariance ratios across different noise levels.

Figures [8.17](#) and [8.18](#) demonstrate how the ratios of covariance matrices for the angular power spectrum and the non-correlation statistics change with different shape noise levels. The results indicate that the angular power spectrum and minima are particularly sensitive to the shape noise level, exhibiting significant variations in their covariance matrices. In contrast, other non-correlation statistics remain more robust against changes in the shape noise level, maintaining relatively stable off-diagonal elements in their covariance matrices.

8.3 Effects of Smoothing Scale

To evaluate the impact of smoothing on the statistical measures, we have applied four different smoothing scales to the simulations. Smoothing affects the resolution of the convergence maps and can influence the detection of small-scale structures.

Figures [8.15](#) and [8.16](#) show how the average ratios of covariance matrices and correlation matrices change with varying smoothing scales. The ratios become more unstable due to the smoothing effect washing out small-scale structures.

Figure [8.19](#) illustrates the effects of smoothing scale on non-Correlation statistical measures. As the smoothing scale increases, the finer structures in the convergence maps are blurred, leading to changes in the statistical properties. The blank bins that previously contained little or no signal begin to be filled due to the spread of signals from neighboring bins, while the overall signal intensity is redistributed.

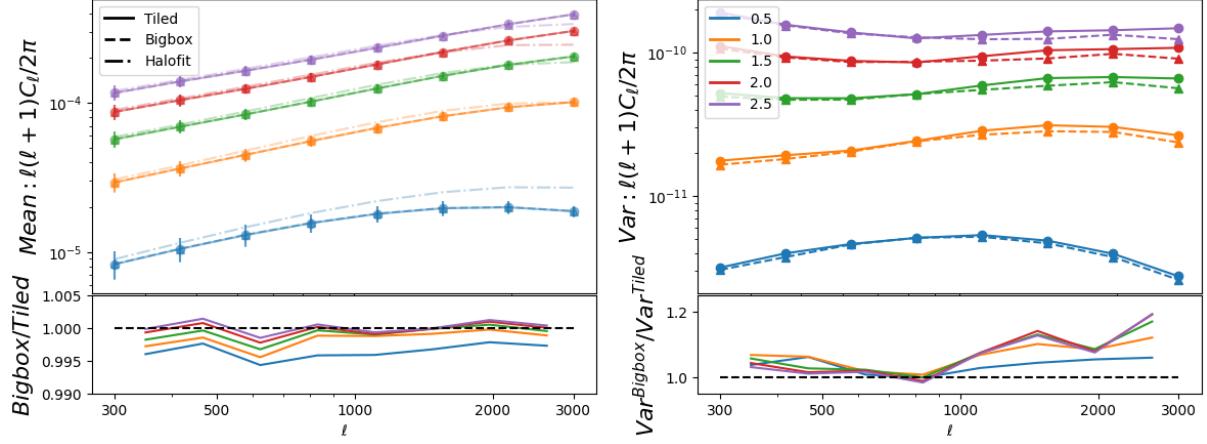


Figure 8.1: Comparison of the mean values of the angular power spectrum ($C_\ell^{\kappa\kappa}$) for different source redshifts ($z_s = 0.5, 1.0, 1.5, 2.0, 2.5$) obtained from the BIGBOX (solid lines) and TILED (dashed lines) simulations. The lower subplots show the ratio of the TILED to BIGBOX mean values, with a reference line at unity to facilitate the assessment of agreement between the two simulations.

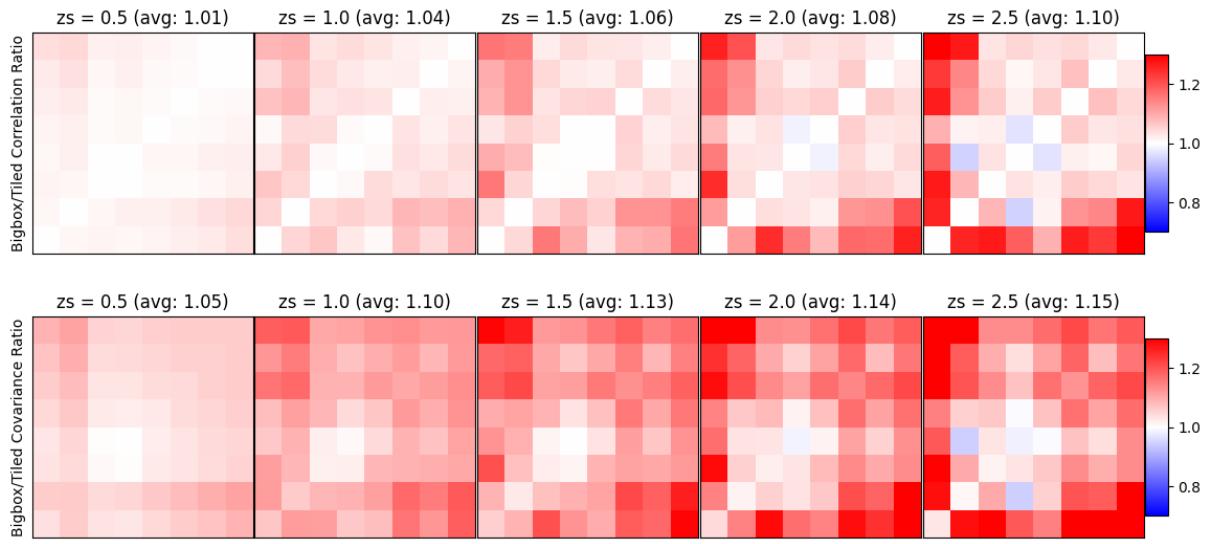


Figure 8.2: Comparison of the covariance matrices and correlation matrices of the angular power spectrum ($C_\ell^{\kappa\kappa}$) between the BIGBOX and TILED simulations for various source redshifts ($z_s = 0.5, 1.0, 1.5, 2.0, 2.5$). The displayed ratios represent the element-wise division of the covariance and correlation matrices from the TILED simulations by those from the BIGBOX simulations. The "avg" denotes the average ratio of the considered matrix elements.

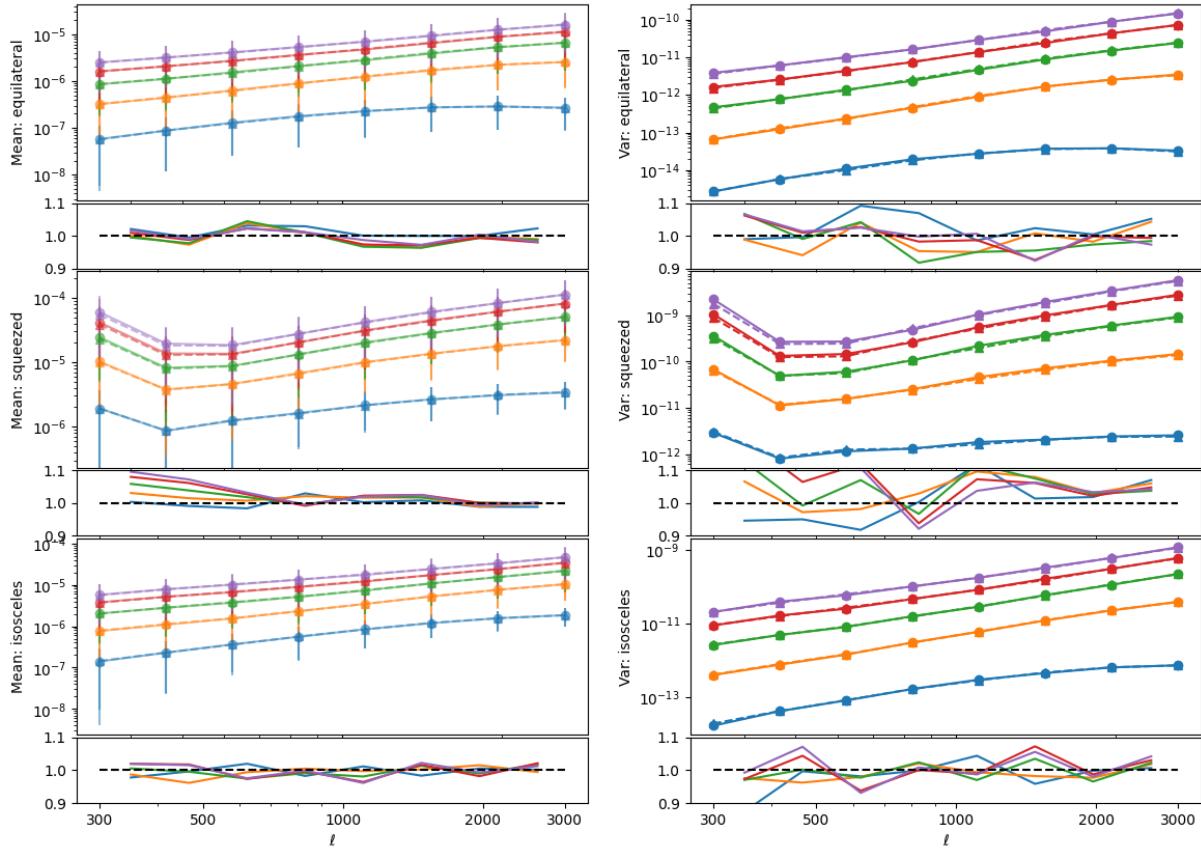


Figure 8.3: Same as Figure 8.1, but for the bispectrum.

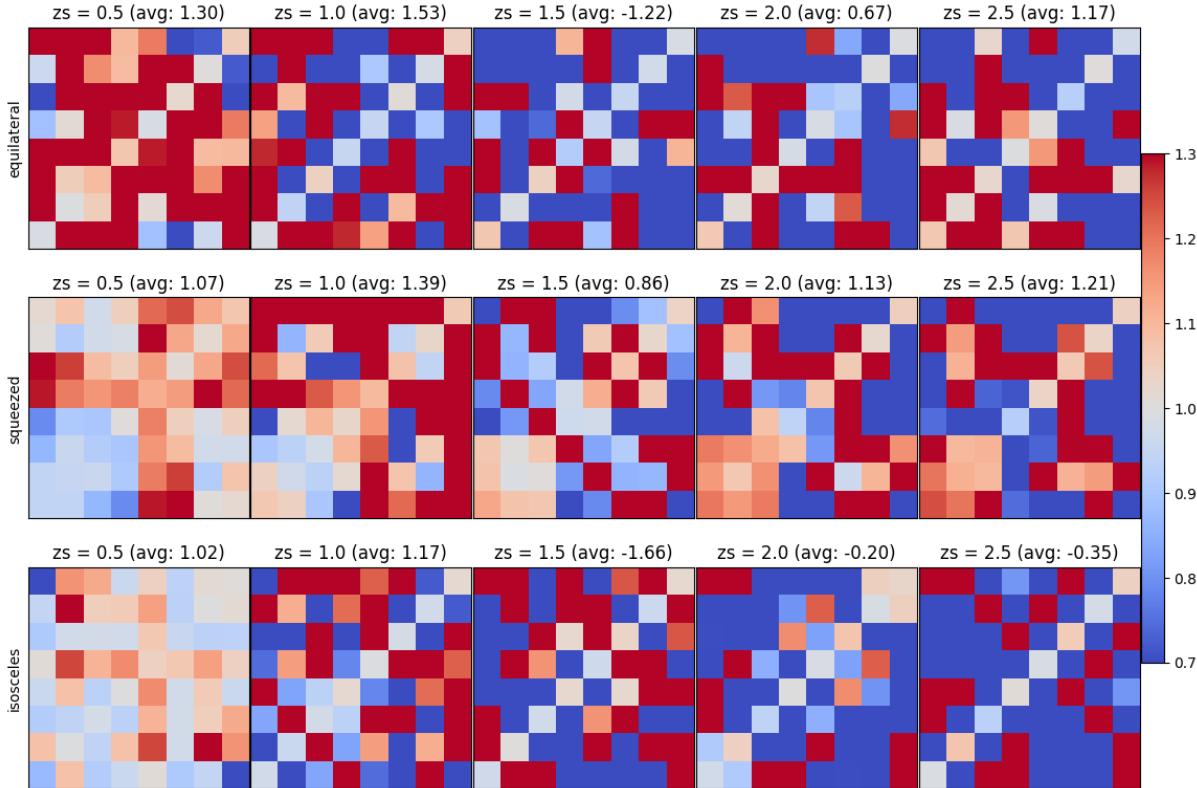


Figure 8.4: Similar to Figure 8.2, but for the covariance matrices of the bispectrum. The noisy nature of the bispectrum covariance makes it challenging to discern clear trends between the simulations.

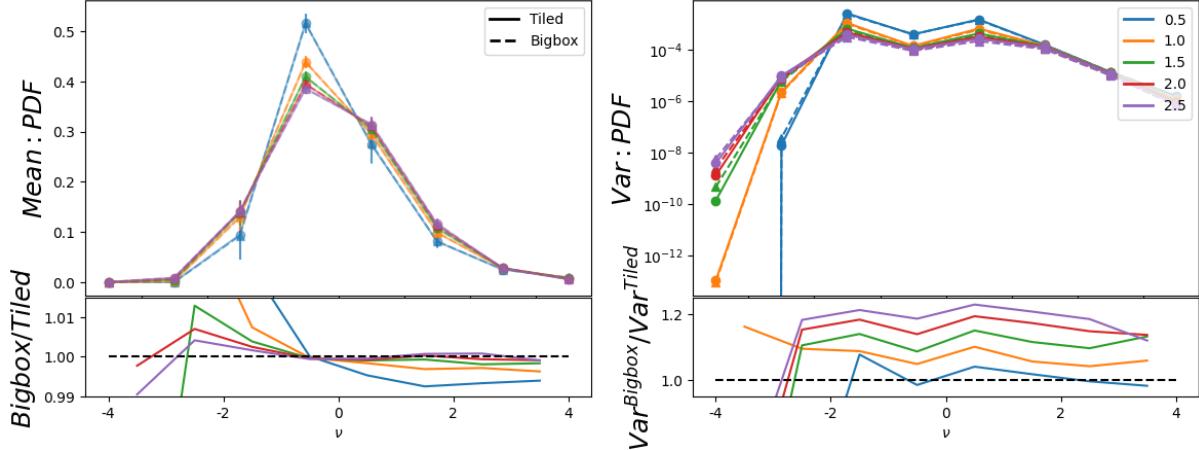


Figure 8.5: Same as Figure 8.1, but for the probability density function (PDF) of the convergence field. The comparison highlights the agreement in mean PDF values between the simulations across different redshifts.

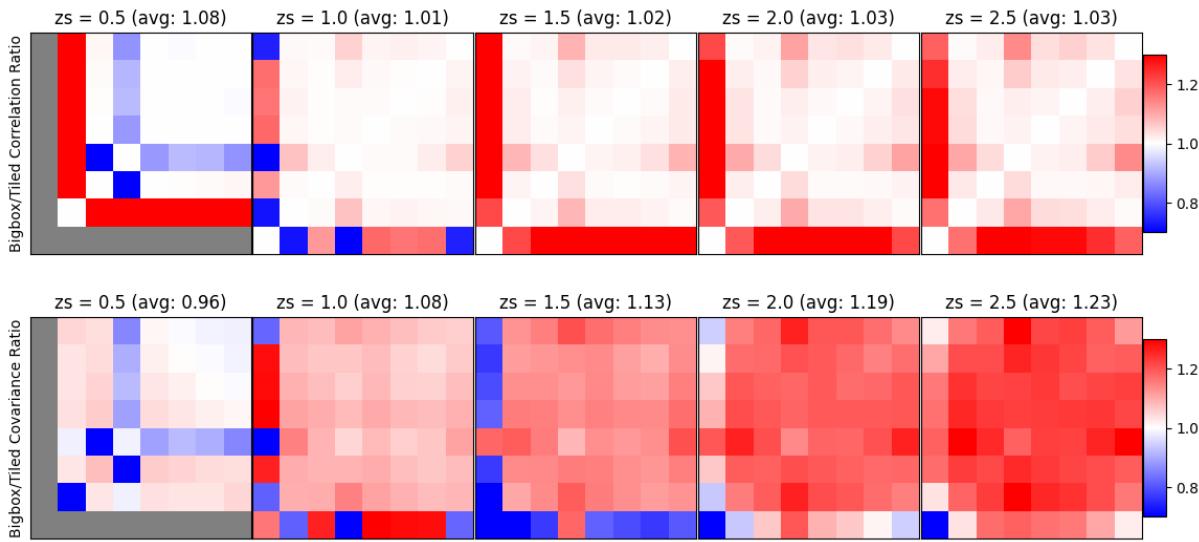


Figure 8.6: Same as Figure 8.2, but for the covariance matrices of the PDF. The covariance ratios indicate higher covariance in the BIGBOX simulations, particularly at higher redshifts.

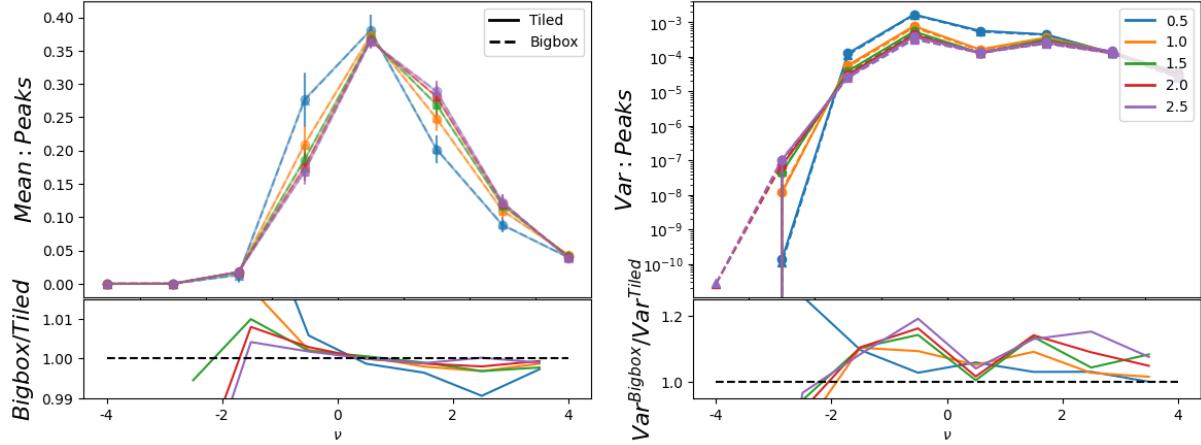


Figure 8.7: Same as Figure 8.1, but for peak counts in the convergence maps. The analysis reveals deviations at low ν values due to resolution limitations affecting low-density regions.

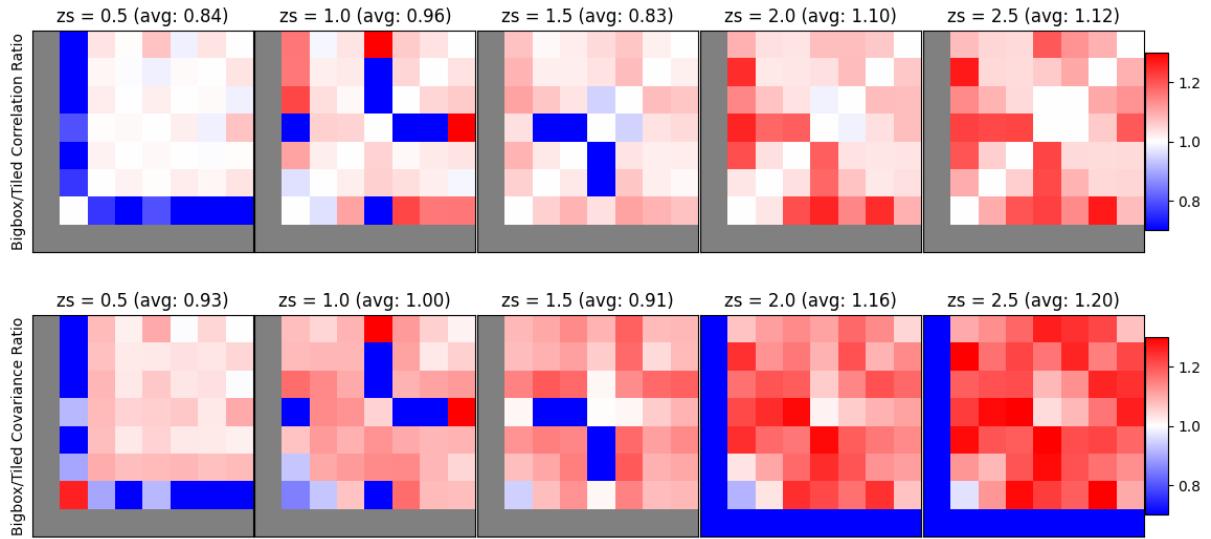


Figure 8.8: Same as Figure 8.2, but for the covariance matrices of peak counts. The covariance ratios suggest increased covariance in the BIGBOX simulations, with pronounced effects at higher redshifts.

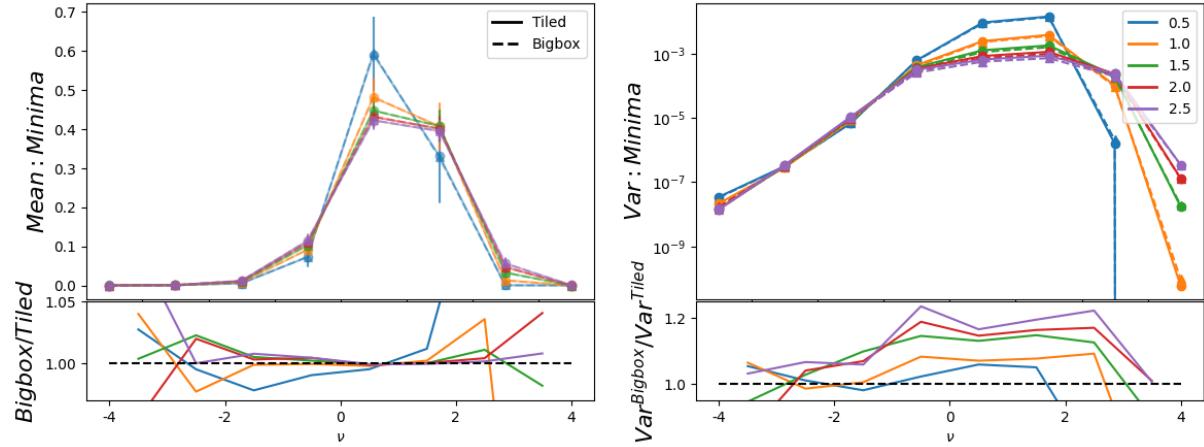


Figure 8.9: Same as Figure 8.1, but for minima in the convergence maps. The comparison underscores the simulation’s limitations at resolving low-density minima accurately.

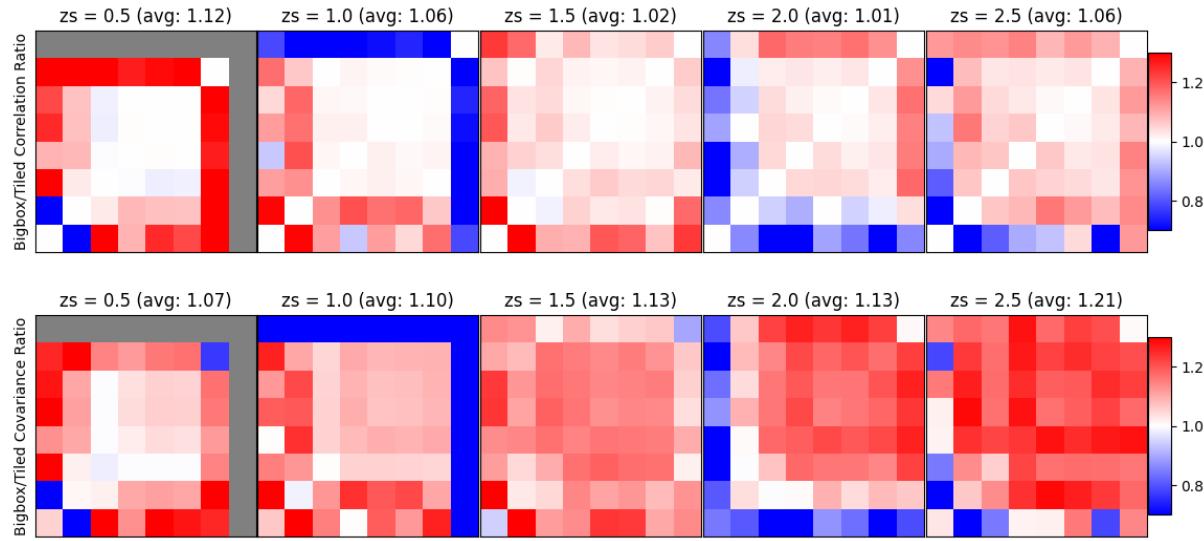


Figure 8.10: Same as Figure 8.2, but for the covariance matrices of minima. The covariance ratios reflect higher values in the BIGBOX simulations, consistent with other statistical measures.

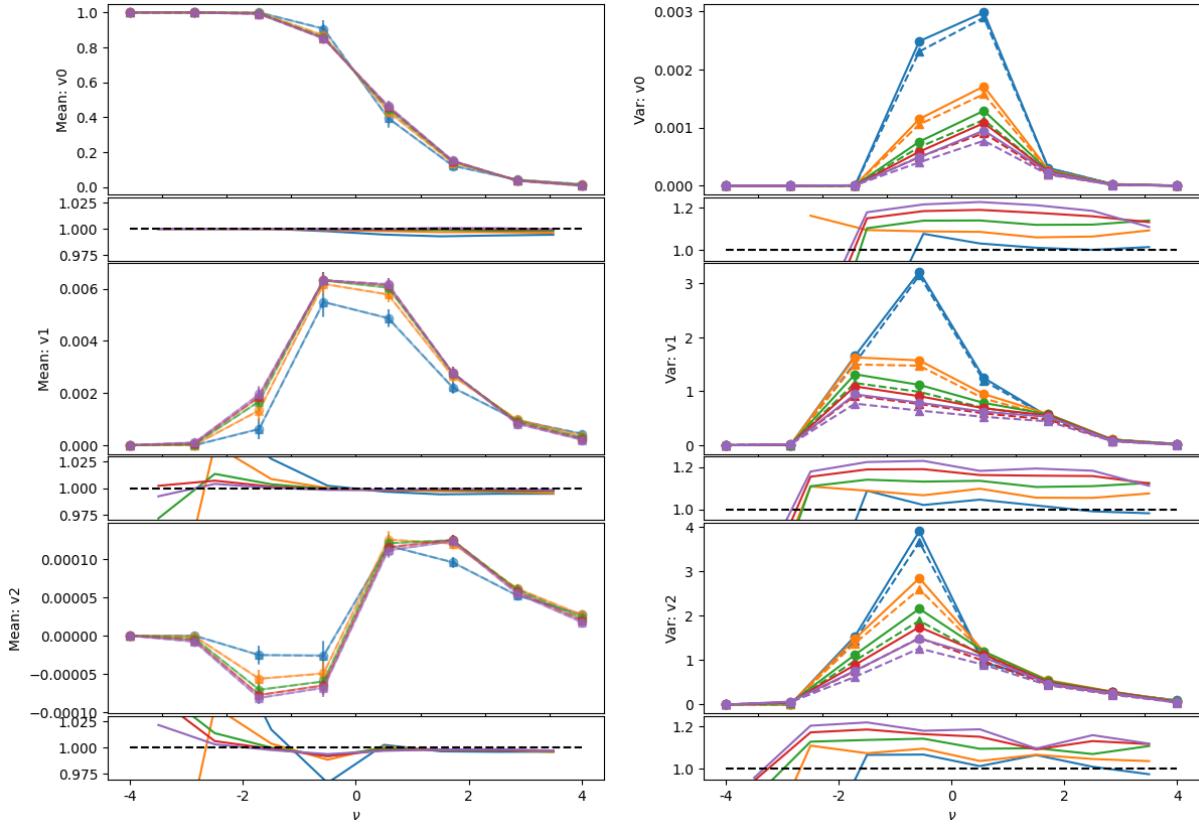


Figure 8.11: Same as Figure 8.1, but for Minkowski Functionals (area V_0 , perimeter V_1 , and genus V_2). The agreement in mean values between simulations is generally good, with some discrepancies at extreme density thresholds.

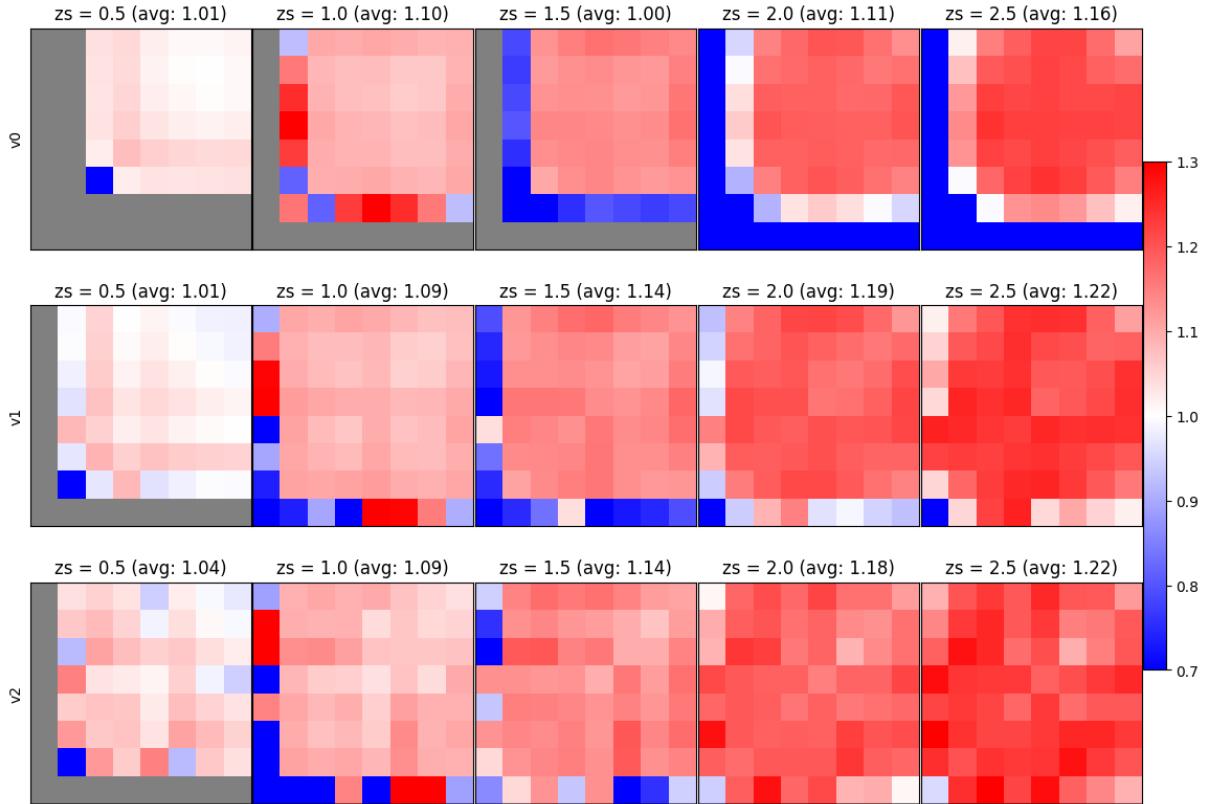


Figure 8.12: Similar to Figure 8.2, but for the covariance matrices of Minkowski Functionals.

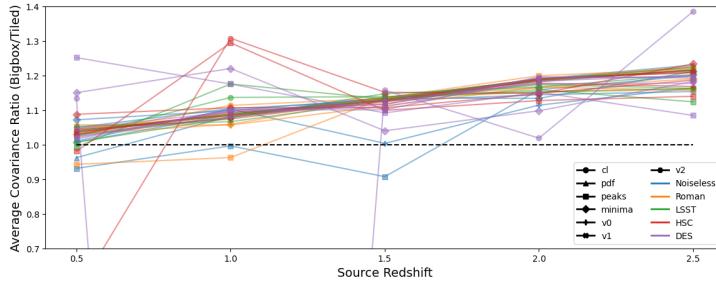


Figure 8.13: Average ratio of covariance matrices of statistical measures between the BIGBOX and TILED simulations for different shape noise levels (see Table 7.2). The increasing trend indicates does not affected by the noise level.

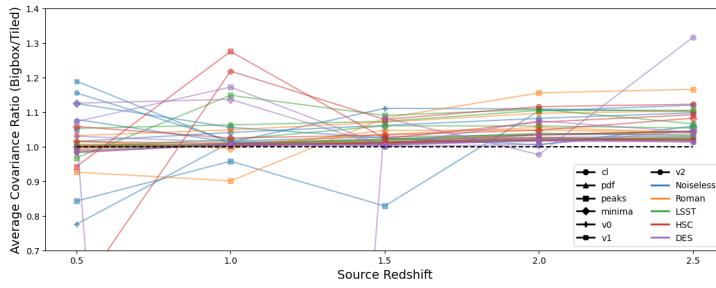


Figure 8.14: Same as Figure 8.13, but for the correlation matrices. The off-diagonal elements compared to the diagonal elements do not show a clear trend with noise levels.

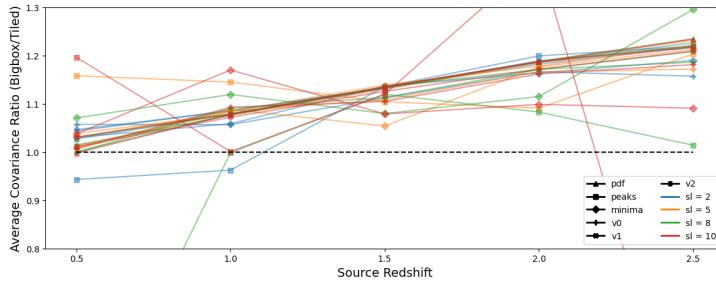


Figure 8.15: Average ratio of covariance matrices of statistical measures between the BIGBOX and TILED simulations for different smoothing scales. Larger smoothing scales lead to increased discrepancies in covariance estimates due to the loss of small-scale information.

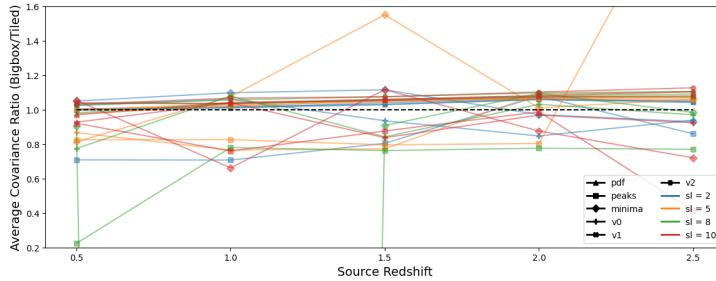


Figure 8.16: Same as Figure 8.15, but for the correlation matrices. The instability at larger smoothing scales reflects the challenges in capturing correlations at reduced resolutions.

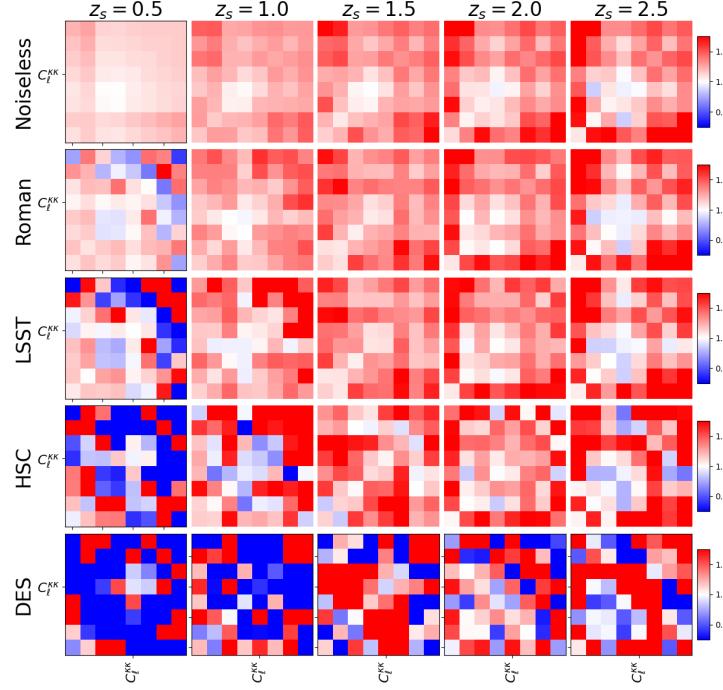


Figure 8.17: Ratio of covariance matrices of the angular power spectrum ($C_{\ell}^{\kappa\kappa}$) between the BIGBOX and TILED simulations for different shape noise levels (see Table 7.2). The sensitivity of the power spectrum to noise is evident from the fluctuating covariance ratios with higher noise levels.

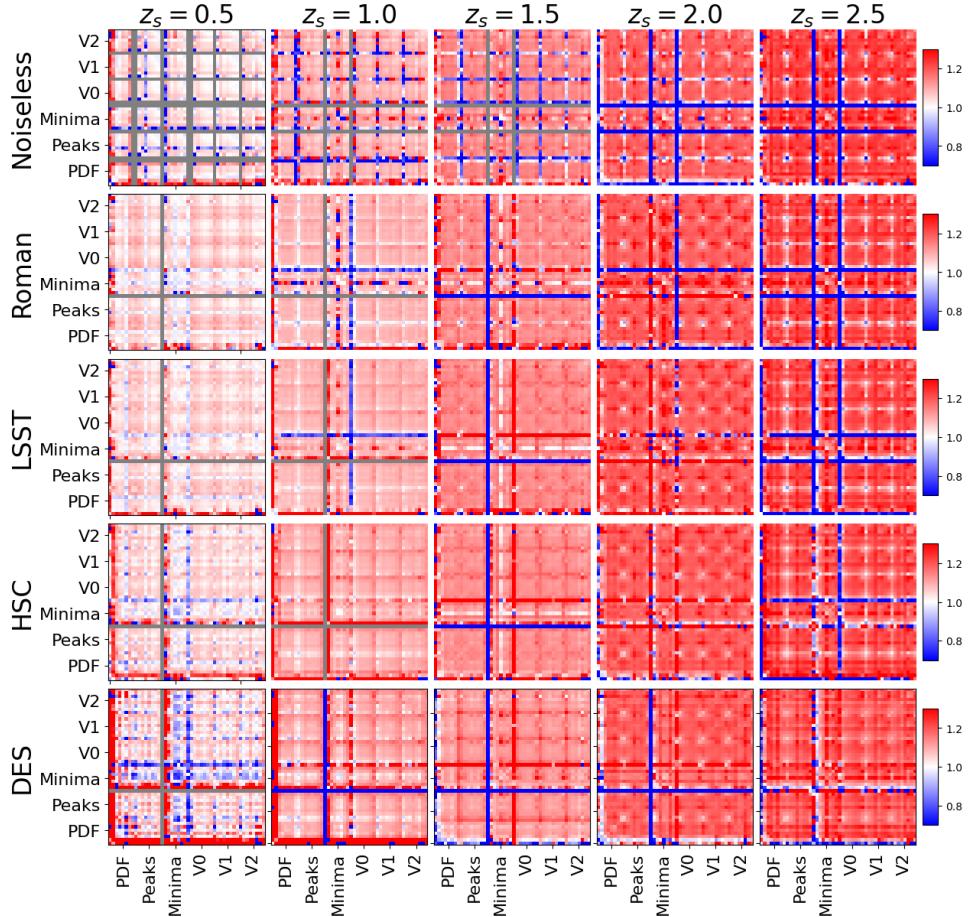


Figure 8.18: Same as Figure 8.17, but for the non-Gaussian statistical measures. The robustness of these measures against noise variations is reflected in the relatively stable covariance ratios.

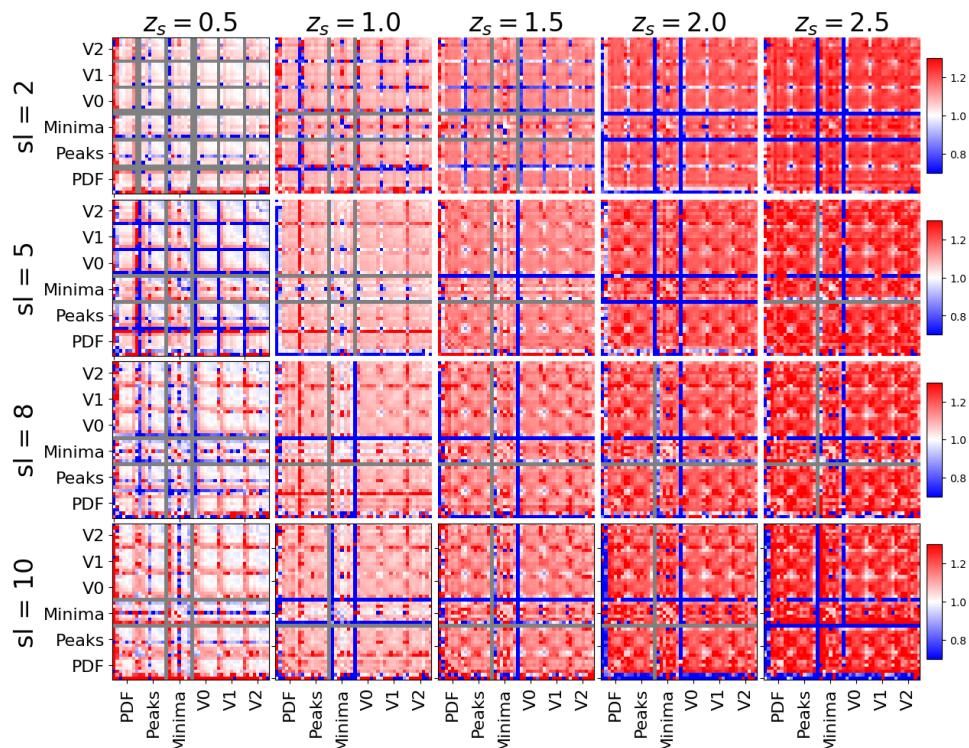


Figure 8.19: Same as Figure 8.18, but showing the impact of different smoothing scales on the covariance matrices of non-Gaussian statistical measures. The results emphasize how increased smoothing affects the detection and characterization of small-scale features.

Chapter 9

Discussion

9.1 Possible Effects

In this section, we will consider possible effects that could affect the covariance matrix, except for the super-sample covariance.

9.1.1 Finite Support Effects

9.1.2 Box Replication Effect

It is clear that the patches lying on the equator are more tiled compared to the rest. For a rough estimation, we check the statistics of the patches lying on the equator and compare them with the rest of the patches.

9.2 Validation

We conducted simulations to validate the effects of finite support and box replication. The simulations were performed with box sizes (L_{box} [Mpc/ h]) of 125, 250, 500, 1000, 2000, and 4000, corresponding to particle numbers (N_{part}) of 125^3 , 250^3 , 500^3 , 1000^3 , 2000^3 , and 4000^3 , respectively. The simulations cover redshifts from 0 to 3, and for each set of parameters, we generated 5 realizations.

We check the statistics for each simulation boxes and compare them each other.

9.3 Check if the gnomview matters

9.4 Why higher-order statistics are less affected?

9.5 correlation between different smoothing scales

Bibliography

Abazajian K. N., et al., 2009, [ApJS](#), **182**, 543

Abbott T. M. C., et al., 2018, [ApJS](#), **239**, 18

Abbott T. M. C., et al., 2019, [Phys. Rev. Lett.](#), **122**, 171301

Abbott T. M. C., et al., 2021, [ApJS](#), **255**, 20

Abbott T. M. C., et al., 2022, [Phys. Rev. D](#), **105**, 023520

Aihara H., et al., 2018a, [PASJ](#), **70**, S4

Aihara H., et al., 2018b, [PASJ](#), **70**, S8

Aihara H., et al., 2019, [PASJ](#), **71**, 114

Aihara H., et al., 2022, [PASJ](#), **74**, 247

Akeson R., et al., 2019, [arXiv e-prints](#), p. arXiv:1902.05569

Albrecht A., et al., 2006, [arXiv e-prints](#), pp astro-ph/0609591

Amon A., et al., 2022, [Phys. Rev. D](#), **105**, 023514

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, [ApJ](#), **304**, 15

Barnes J., Hut P., 1986, [Nature](#), **324**, 446

Barreira A., Schmidt F., 2017, [J. Cosmology Astropart. Phys.](#), **2017**, 051

Bartelmann M., 2010, [Classical and Quantum Gravity](#), **27**, 233001

Bartelmann M., Schneider P., 2001, [Phys. Rep.](#), **340**, 291

Bayer A. E., Zhong Y., Li Z., DeRose J., Feng Y., Liu J., 2024, [arXiv e-prints](#), p. arXiv:2407.17462

Behroozi P., Silk J., 2018, [MNRAS](#), **477**, 5382

Berger M. J., Colella P., 1989, [Journal of Computational Physics](#), **82**, 64

- Blas D., Lesgourgues J., Tram T., 2011, [J. Cosmology Astropart. Phys.](#), 2011, 034
- Born M., 1926, [Zeitschrift fur Physik](#), 38, 803
- Boyle A., Uhlemann C., Friedrich O., Barthelemy A., Codis S., Bernardeau F., Giocoli C., Baldi M., 2021, [MNRAS](#), 505, 2886
- Cooley J. W., Tukey J. W., 1965, [Mathematics of Computation](#), 19, 297
- Cui W., Liu L., Yang X., Wang Y., Feng L., Springel V., 2008, [ApJ](#), 687, 738
- DESI Collaboration et al., 2016, [arXiv e-prints](#), p. arXiv:1611.00036
- Dalal R., et al., 2023, [Phys. Rev. D](#), 108, 123519
- Dawson K. S., et al., 2013, [AJ](#), 145, 10
- Dubinski J., 1996, [New Astronomy](#), 1, 133
- Eisenstein D. J., et al., 2005, [ApJ](#), 633, 560
- Euclid Collaboration et al., 2019, [A&A](#), 627, A59
- Feng Y., Chu M.-Y., Seljak U., McDonald P., 2016, [MNRAS](#), 463, 2273
- Ferlito F., et al., 2023, [MNRAS](#), 524, 5591
- Flaugher B., et al., 2015, [AJ](#), 150, 150
- Freedman W. L., et al., 2001, [ApJ](#), 553, 47
- Gaia Collaboration et al., 2016, [A&A](#), 595, A2
- Gaudi B. S., 2012, [ARA&A](#), 50, 411
- Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, [The Astrophysical Journal](#), 622, 759
- Hilbert S., Hartlap J., White S. D. M., Schneider P., 2009, [A&A](#), 499, 31
- Hockney R. W., Eastwood J. W., 1981, Computer Simulation Using Particles. CRC Press
- Hounsell R., et al., 2018, [ApJ](#), 867, 23
- Hu W., 2000, [Phys. Rev. D](#), 62, 043007
- Ivezić Ž., et al., 2019, [ApJ](#), 873, 111

- Johnson S. A., Penny M., Gaudi B. S., Kerins E., Rattenbury N. J., Robin A. C., Calchi Novati S., Henderson C. B., 2020, [AJ, 160, 123](#)
- Jones R. L., et al., 2018, [Icarus, 303, 181](#)
- Kacprzak T., et al., 2016, [MNRAS, 463, 3653](#)
- Kawanomoto S., et al., 2018, [PASJ, 70, 66](#)
- Kilbinger M., 2015, [Reports on Progress in Physics, 78, 086901](#)
- Kilbinger M., et al., 2017, [MNRAS, 472, 2126](#)
- Kollmeier J., et al., 2019, in Bulletin of the American Astronomical Society. p. 274
- Kratochvil J. M., Lim E. A., Wang S., Haiman Z., May M., Huffenberger K., 2012, [Phys. Rev. D, 85, 103513](#)
- LSST Dark Energy Science Collaboration 2012, [arXiv e-prints, p. arXiv:1211.0310](#)
- LSST Science Collaboration et al., 2009, [arXiv e-prints, p. arXiv:0912.0201](#)
- Leclercq F., 2020, Evolution of cosmological simulations over the last 50 years, <http://florent-leclercq.eu/blog.php?page=2>
- Lewis A., Challinor A., Lasenby A., 2000, [ApJ, 538, 473](#)
- Li Y., Hu W., Takada M., 2014, [Phys. Rev. D, 89, 083519](#)
- Limber D. N., 1954, [ApJ, 119, 655](#)
- Martinet N., et al., 2018, [MNRAS, 474, 712](#)
- Martinet N., Castro T., Harnois-Déraps J., Jullo E., Giocoli C., Dolag K., 2021, [A&A, 648, A115](#)
- Matsubara T., 2010, [Phys. Rev. D, 81, 083505](#)
- Miyazaki S., et al., 2018, [PASJ, 70, S1](#)
- Mohr J. J., et al., 2008, in Brissenden R. J., Silva D. R., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 7016, Observatory Operations: Strategies, Processes, and Systems II. p. 70160L ([arXiv:0807.2515](#)), doi:10.1117/12.789550
- Nelson D., et al., 2019, [Computational Astrophysics and Cosmology, 6, 2](#)
- Penny M. T., Gaudi B. S., Kerins E., Rattenbury N. J., Mao S., Robin A. C., Calchi Novati S., 2019, [ApJS, 241, 3](#)

- Perlmutter S., et al., 1999, [ApJ, 517, 565](#)
- Petri A., 2016, [Astronomy and Computing, 17, 73](#)
- Petri A., Haiman Z., Hui L., May M., Kratochvil J. M., 2013, [Phys. Rev. D, 88, 123002](#)
- Powell J., Caudill L., Young O., 2023, [American Journal of Physics, 91, 478](#)
- Riess A. G., et al., 1998, [AJ, 116, 1009](#)
- Robertson B. E., et al., 2019, [Nature Reviews Physics, 1, 450](#)
- Schneider P., Ehlers J., Falco E. E., 1992, Gravitational Lenses, [doi:10.1007/978-3-662-03758-4.](#)
- Shirasaki M., Hamana T., Yoshida N., 2015, [MNRAS, 453, 3043](#)
- Spergel D., et al., 2015, [arXiv e-prints, p. arXiv:1503.03757](#)
- Sullivan M., et al., 2011, [ApJ, 737, 102](#)
- Swinbank R., James Purser R., 2006, [Quarterly Journal of the Royal Meteorological Society, 132, 1769](#)
- Takada M., Hu W., 2013, [Phys. Rev. D, 87, 123504](#)
- Takada M., Jain B., 2004, [MNRAS, 348, 897](#)
- The Dark Energy Survey Collaboration 2005, [arXiv e-prints, pp astro-ph/0510346](#)
- Uhlemann C., Friedrich O., Boyle A., Gough A., Barthelemy A., Bernardeau F., Codis S., 2023, [The Open Journal of Astrophysics, 6, 1](#)
- Zonca A., Singer L., Lenz D., Reinecke M., Rosset C., Hivon E., Gorski K., 2019, [Journal of Open Source Software, 4, 1298](#)
- de Jong J. T. A., et al., 2013, [The Messenger, 154, 44](#)