Alma Mater Studiorum – Università di Bologna

School of Science
Department of Physics and Astronomy
Master Degree Programme in Astrophysics and Cosmology

Inspecting signatures of parity violation with angular redshift fluctuations

Graduation Thesis

Presented by:
Matteo Santini

Supervisor:
Prof. Federico Marulli

Co-supervisor:
Dr. Raúl Esteban Angulo de la Fuente

Academic year 2023-2024
Graduation date IV

Abstract

Parity is the transformation that inverts the spatial coordinates of a physical system. Among the known fundamental interactions of nature, only the weak force does not obey parity symmetry. For examples, neutrinos, which interact solely via the weak force, are uniquely left-handed, as empirically demonstrated by Chien-Shiung Wu in 1956 [171]. The gravitational interaction, as described by Einstein's general theory of relativity in 1915, is understood as the intrinsic curvature of spacetime. This is the dominant interaction at cosmological scales and it is invariant under reflection of space through the origin of coordinates. For this reason, examining manifestations of parity asymmetries in the large-scale structure (LSS) of the Universe at late times allows us to probe for the existence of early-Universe parity violation signatures. Possible indirect evidences arises in baryogenesis [148] which generates asymmetry between baryons and anti-baryons, as first described by the Sakharov conditions [136]. Other candidates for parity-violating processes that could imprint late-time observables are related to modifications to general relativity, such as Chern-Simons gravity [12, 26, 40], generation of magnetic fields in the primordial Universe [62], reheating [37, 8], perturbations from cosmic strings [126] and inflationary scenarios such as particle exchange [77].

In general, determining whether the Universe is parity-symmetric is a non-trivial task. Data from large spectroscopic surveys have been used to investigate the mirror-symmetric properties of the Universe. In particular, [28, 120] claimed a detection of slight asymmetries in the distribution of chiral tetrahedra of galaxies. As pointed out in [38], these results may be due to systematic effects, though the possibility that they are instead physical cannot yet be ruled out.

In this work, a novel approach involving the use of angular redshift fluctuations (ARFs) as a probe for parity violation is adopted. ARFs are a powerful tool to retrieve cosmological information from the LSS of the Universe. These statistics of sky maps are built by projecting galaxy redshifts, weighted by the galaxy density field, onto a Gaussian window with a certain width and a fixed central redshift. One of the advantages of ARFs is their low expected sensitivity to systematic effects. Thus, measuring a parity-violating signal with ARFs could be a valuable step toward unveiling the symmetric properties of the Universe and understanding the nature of the mechanism that leads to its evolution.

To this end, this study conducts a test for parity violation using binned numerical estimators of the trispectra for ARFs. This choice is motivated by the limitations of lower-order statistics which are inherently insensitive to parity violations. Specifically, for power spectra and bispectra, the inversion of coordinates is indistinguishable from a 3-dimensional rotation, making the unsuitable for parity tests. Trispectra, therefore, represent the lowest-order correlation function useful to search for parity asymmetries.

For this work, we used the BOSS DR12 galaxy catalogue together with the MultiDark-`PATCHY` mocks, which are specifically designed to accurately replicate the observed clustering evolution and its dependency with stellar mass, thus providing robust error estimates and allowing us to a deep understanding of systematic uncertainties. The test for parity violation was conducted on the CMASS and LOWZ samples, covering both the North Galactic Cap and the South Galactic Cap. A new class, MapCalculator, was implemented in the `CosmoBolognaLib`, a suite of C++ libraries for cosmological calculations, and employed to generate ARF maps. This class includes methods to build sky maps by taking a catalogue and a mask as inputs. The code converts the angular coordinates of the objects in the catalogue into pixel indices of the map and assigns a value to each pixel based on the central redshift and the width of a Gaussian shell. The input mask is then applied to the map to exclude noise sources and focus on the regions of interest.

Parity violation was tested by comparing the difference between the parity-odd and parity-even components for each trispectrum estimation. This comparison determines whether the results are compatible or not with a null value, thus assessing the possible presence of parity violation. In this study, we modeled the distribution of the scale-averaged signal-to-noise ratio across different catalogues. The results of our analysis on the window-deconvolved trispectrum estimator yielded a value of 0.035 ± 0.89 for the CMASS sample and -0.013 ± 0.92 for the LOWZ sample. No evidence of parity violation was detected.

Additional tests were performed using parity-violating Quijote-ODD simulations, which are standard N-body simulations that evolve parity-violating initial conditions. The level of parity violation is quantified by the parameter p_{NL}, which defines two subtypes of these simulations: ODD_M (with $p_{NL} = -10^6$) and ODD_P (with $p_{NL} = +10^6$). We compared their odd-trispectrum estimations with those derived from fiducial LCDM Quijote simulations. The distribution of the scale-averaged signal-to-noise ratio for the difference between the odd-trispectra of the two simulations was fitted with Gaussian functions. At the reference redshift $z_{cen} = 0.8$, the best-fit mean and standard deviation for ODD_M are ($0.0030^{+0.00010}_{-0.00010}$, $0.0236^{0.0010}_{0.0009}$) and for ODD_P they are ($-0.0031^{+0.0011}_{-0.0010}$, $0.0240^{+0.0010}_{-0.0010}$. Similarly, at the reference redshift $z_{cen} = 1.25$ we obtained best-fit mean and standard deviation of ($0.0011^{+0.00008}_{-0.00008}$, $0.0227^{0.0008}_{0.0007}$) for ODD_M and ($0.0004^{+0.0008}_{-0.0008}$, $0.0229^{+0.0008}_{-0.0007}$) for ODD_P, showing no significant parity-violating features.

Contents

Introduction

Cosmology is the study of the Universe as a whole, exploring its origins, structure, and evolution over time. It seeks to understand the large-scale distribution of matter and the physical laws governing its dynamics. From the nearly homogeneous state of the early Universe, cosmologists have been modelling the intricate web of galaxies and cosmic structures we observe today. Modern cosmology is the culmination of a continuously evolving network of interconnected findings summarized in the Λ-Cold Dark Matter model (ΛCDM), the standard model of cosmology.

However, several critical questions remain unresolved, some of which may point to fundamental gaps in our current understanding. Although the ΛCDM model has been rigorously validated through numerous observational data sets, subtle and hidden flaws still exist. These uncertainties could stem from the unknown nature of the energy components that make up the majority of the Universe's content, the physical processes driving the very early Universe, and incomplete descriptions of gravity.

An intriguing approach to exploring the early Universe and testing gravity theories involves the study of parity symmetry, a transformation that inverts the spatial coordinates of a system. In the standard ΛCDM scenario, parity is conserved, meaning that the laws of physics remain unchanged under this transformation. While parity violation is well-established in the weak interaction, determining whether similar violations occur in cosmology remains an open question.

The goal of this work is to test whether the Universe's appearance is invariant under coordinate inversion. While this may seem counterintuitive, since coordinates are a human invention and the laws of physics should not depend on the chosen chart (principle of relativity), coordinate transformations play a fundamental role in physics. In particular, *Noether's theorem* establishes a profound connection between continuous coordinate transformations and conservation laws: whenever a physical system is invariant under a continuous coordinate transformation, a corresponding physical quantity is conserved. For instance, if we repeat an experiment multiple times, rotating the experimental apparatus before each measurement, and find that the results are consistent within experimental uncertainties, this demonstrates invariance under spatial rotation (isotropy), which is directly linked to the conservation of total angular momentum.

Parity, however, is not a continuous transformation and therefore does not correspond to a conserved quantity. Nevertheless, in a cosmological context, studying parity invariance could provide insights into the nature of dark matter, dark energy, and the primordial Universe. Determining whether the Universe violates parity symmetry is a challenging but significant question.

Efforts to detect parity-violating signals have primarily focused on studying the cosmic microwave background (CMB). In particular, parity tests have been performed on the CMB temperature fluctuations [121] and polarization modes [142, 124]. While these studies have not found definitive evidence for parity violation, they often lacked sufficient data to place strong constraints [172].

Other investigations of parity symmetries in the CMB focus on cosmic birefringence, which explores the interaction between the primordial electromagnetic field and the medium it traverses. This phenomenon can rotate the linear polarization plane of photons, potentially due to interactions with dark matter or dark energy. Under a parity transformation, the rotation angle flips direction (clockwise to counterclockwise). If the CMB exhibits a preferential direction of rotation, this would indicate parity violation [53, 51]. Such findings could provide evidence for new physics, possibly challenging the cosmological constant model [88] and favoring parity-violating particles in the dark matter sector, such as sterile neutrinos. Furthermore, evidence of B-mode polarization patterns could reveal physics at the grand unified theory (GUT) scale [126].

Observations of LSS provide an additional avenue to investigate parity violation. These structures retain information about the early Universe through indirect signatures from processes such as baryogenesis, reheating, and inflation [148, 136, 8, 37, 26, 77]. LSS studies also offer a platform to test modifications of standard gravitational theories, such as Chern-Simons gravity [12]. In these theories, parity violation could manifest in phenomena like the birefringence of gravitational waves [112], detectable with gravitational wave interferometers [2, 4], or in the primordial scalar trispectrum induced by graviton exchange, where the interaction is mediated by hypothetical graviton particles [40].

Studies like [28, 120] have proven that LSS is a valuable laboratory for parity tests, and they reported slight asymmetries in four-point correlations. However, these results may be affected by systematic uncertainties [38].

This thesis investigates parity symmetry in cosmology using ARFs, a novel observable that traces the underlying density and velocity fields of galaxies [69, 71, 70]. Unlike traditional statistics, ARFs are minimally affected by systematic uncertainties and do not depend on a fiducial cosmology to relate distances and redshifts. This makes them a robust and reliable tool for testing parity violations, especially in cases where other observables are hindered by biases or limited data.

To provide context, Chapter 1 reviews the theoretical framework of cosmology, discussing the spacetime description of the Universe, the Cosmological Principle, and General Relativity, while offering a comprehensive overview of the ΛCDM model. The mechanisms of clustering and structure formation are explored in Chapter 2, with an emphasis on their connection to primordial fluctuations and their role in shaping the observed Universe. Particular attention is given to how small initial perturbations evolved into the LSS we observe today, driven by gravitational interactions among the components of the Universe. Chapter 3 delves into parity violation, examining its role in fundamental physics, with a specific focus on gravity. We also discuss its significance in cosmology and present methods for detecting it, introducing higher-order correlations, with particular attention to the trispectrum. Chapter 4 details ARFs, focusing on their definition and implementation. By applying this tool to analyze data presented in Chapter 5, this work aims to detect subtle signals of parity violation in the large-scale distribution of matter. Finally, the results are presented in Chapter 6, where we discuss the insights this thesis provides on parity symmetry.

Chapter 1

Introduction to cosmology

In this chapter, a general overview of the fundamental building blocks of cosmology will be provided to understand the key aspects of this unique physical science.

Cosmology is an ancient subject that concerns the formation and evolution of the Universe as a whole. It seeks to make sense of the large-scale distribution of matter by leveraging the tools of physics and inheriting its tradition of understanding where the Universe comes from, why it appears as it does, and how it will evolve over time.

1.1 The axioms of cosmology

The field of cosmology originated as a branch of philosophy and began to evolve into a physical science in 1917, with Albert Einstein's development of General Relativity. Scientific inquiry in cosmology follows the general method of constructing models supported by observations. In this framework, General Relativity is elevated to the status of an axiom and serves as the foundation for building models with the fewest possible free parameters. Since the theory does not include any parity-violating features, understanding how it describes the real world and whether this description agrees with a parity-symmetric Universe are key aspects of this work. However, General Relativity is not the only assumption in cosmology as it must be complemented by the assumptions of isotropy and homogeneity of the Universe. These two assumptions are embodied in the Cosmological Principle, a symmetry principle that helps reduce the degrees of freedom in cosmological models.

The concept of isotropy states that the Universe looks the same in all directions of observation. It is closely related to the Copernican Principle, which asserts that there is no preferred or special location from which to observe the Universe. When isotropy is combined with the Copernican Principle, the result is the concept of homogeneity, which holds that the structural properties of the Universe remain invariant across different locations. Clearly, the requirements imposed by the Cosmological Principle are not always fulfilled. In any direction one points their gaze, the Universe reveals inhomogeneities in the form of e.g. planets, stars, galaxies, or clusters of galaxies. This points out a key aspect of the Cosmological Principle: it only holds on the largest scales (around 185 Mpc) [109] and it concerns average quantities. This characterization of the Universe is supported by studies of the LSS and of the CMB, and it is assumed to apply to the entire Universe, even beyond the observable one.

The Cosmological Principle, however, should not be interpreted in the same way as other principles in physics, such as Heisenberg's uncertainty principle in quantum mechanics. While the latter is a fundamental requirement of the physical theory, the Universe's homogeneity and isotropy in the large-scale average are not demanded by the theory of General Relativity [115].

1.1.1 General Relativity

The theory of General Relativity describes gravity as the intrinsic curvature of spacetime, generalizing the understanding of gravity compared to Newtonian physics. This theory, proposed by Albert Einstein in 1915, provides a framework for understanding gravitational interactions as geometric properties of a four-dimensional manifold.

A manifold is a mathematical structure that is locally topologically isomorphic to $\mathbb{R}^4$, meaning that the neighborhood of each point of the manifold can be assigned coordinates as in $\mathbb{R}^4$. This structure allows us to define a coordinate system locally, while the topological properties of the manifold ensure both the convergence of series and the continuity of functions across the manifold. The shape of a manifold can be arbitrary; it may be curved, and its intrinsic geometry may differ from the familiar (flat) Euclidean geometry. The metric is a crucial tool for describing the geometry of the manifold. Formally, it is defined as a symmetric and non-degenerate (0,2)-tensor that acts as a bilinear map from the tangent bundle of the manifold to the algebraic field $\mathbb{R}$, encoding the intrinsic geometry and determining the length of curves. A manifold equipped with a metric is called a metric manifold.

The length of an infinitesimal curve element relates to the metric through the following expression (in Cartesian coordinates):

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu, \tag{1.1}$$

where $\mathrm{d}s^2$ is the squared length of an infinitesimal line element, $g_{\mu\nu}$ are the components of the metric tensor and $\mathrm{d}x^\mu$ and $\mathrm{d}x^\nu$ are the coordinates of the two endpoints of the curve segment. Here and throughout the text, Greek indices will range from 0 to 3, where 0 corresponds to the time component, and the remaining indices, 1, 2 and 3, represent the spatial components.

The geometry of a manifold is encoded in the metric field g, but how may we understand the manifold is flat or curved? To this aim, let us introduce an important mathematical object used in differential geometry to classify the warping of manifolds. This is the Riemann's tensor

$$R^{\alpha}_{\mu\nu\rho} := \partial_\nu \Gamma^{\alpha}_{\mu\rho} - \partial_\rho \Gamma^{\alpha}_{\mu\nu} + \Gamma^{\alpha}_{\nu\beta}\Gamma^{\beta}_{\mu\rho} + \Gamma^{\alpha}_{\rho\beta}\Gamma^{\beta}_{\mu\nu}, \tag{1.2}$$

and it provides a way to define flatness: if $R^{\sigma}_{\mu\nu\rho} = 0$, then the manifold is flat, independently of the chart. Here, we used $\partial_\mu := \partial/\partial x^\mu$, while $\Gamma^{\alpha}_{\mu\nu}$ are the components of the Levi-Civita connection, defined as

$$\Gamma^{\alpha}_{\mu\nu} := \frac{1}{2} g^{\alpha\beta}(\partial_\mu g_{\nu\beta} + \partial_\nu g_{\mu\beta} - \partial_\beta g_{\mu\nu}). \tag{1.3}$$

The Levi-Civita connection is closely related to the metric and can be interpreted either as a spacetime geometrical effect or as a gravitational force affecting the motion of objects.

The contraction over one pair of indices in Eq.(1.2) results in the Ricci's tensor

$$R_{\mu\nu} := \partial_\alpha \Gamma^{\alpha}_{\mu\nu} - \partial_\nu \Gamma^{\alpha}_{\mu\alpha} + \Gamma^{\beta}_{\mu\nu}\Gamma^{\alpha}_{\beta\alpha} - \Gamma^{\beta}_{\mu\alpha}\Gamma^{\alpha}_{\beta\nu}. \tag{1.4}$$

At the core of General Relativity lie Einstein's field equations, which can be derived from the least action principle applied to the Einstein-Hilbert action. These equations, in natural units, take the form:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G \, T_{\mu\nu}, \tag{1.5}$$

where $R_{\mu\nu}$ is the Ricci curvature tensor, R is the Ricci scalar, G is the gravitational constant and $T_{\mu\nu}$ is the stress-energy tensor representing the energy content of the Universe. This set of equations encapsulates the fundamental relationship between the intrinsic geometry of spacetime, encoded in the metric tensor, and the energy (matter) content that influences this geometry.

1.1.2 Friedmann-Lemaître-Robertson-Walker metric

In cosmology, the selection of an appropriate metric $g_{\mu\nu}$ is crucial for defining distances between events in spacetime. From Eq.(1.1), we can explicit Einstein's summation convention, which implies summing over repeated indices, to obtain three distinct contributions:

$$\mathrm{d}s^2 = g_{00}\mathrm{d}t^2 + 2g_{0i}\mathrm{d}t\mathrm{d}x^i + g_{ij}\mathrm{d}x^i\mathrm{d}x^j, \tag{1.6}$$

where the Latin indices i and j refer to spatial components and range from 1 to 3. Throughout this work, we adopt the metric signature convention (-,+,+,+), where the negative sign corresponds to the time component and the positive signs correspond to the spatial components. Depending on the sign of $\mathrm{d}s^2$, we can distinguish three different types of spacetime intervals:

- $\mathrm{d}s^2 < 0$: *time-like*
- $\mathrm{d}s^2 = 0$: *light-like* (corresponding to the geodesics of photons)
- $\mathrm{d}s^2 > 0$: *space-like*

The Cosmological Principle can be formalized by assuming the existence of six killing vectors, three for spatial rotations (isotropy) and three for spatial translation (homogeneity), describing the spatial symmetry between spacetime intervals and guaranteeing that the mixed term $2g_{0i}\mathrm{d}t\mathrm{d}x^i$ in Eq.(1.1) vanishes. For a flat manifold, the metric simplifies to

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}\ell^2, \tag{1.7}$$

where $\mathrm{d}\ell^2 = g_{ij}\mathrm{d}x^i\mathrm{d}x^j$ is the spatial distance between two points.

Moving beyond flat space, consider a homogeneous and isotropic unit 3-sphere which is a geometric object that can be embedded in four-dimensional spacetime, constructed by gluing the boundaries of two 2-spheres. In Cartesian coordinates, the squared distance element is:

$$\mathrm{d}\ell^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 + \mathrm{d}u^2. \tag{1.8}$$

After adopting a suitable change of coordinates, we can express $\mathrm{d}\ell^2$ in polar coordinates and it takes the form

$$\mathrm{d}\ell^2 = \frac{\mathrm{d}r^2}{1-r^2} + r^2 d\Omega^2 = \frac{\mathrm{d}r^2}{1-r^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2). \tag{1.9}$$

This result can be trivially generalized to a 3-sphere with radius a by including the radius length as scaling in the formula. Such space has a positive curvature, but we could similarly describe spaces with negative curvature, like a 3-hyperboloid.

The most general metric satisfying the Cosmological Principle is the *Friedmann-Lemaître-Robertson-Walker* (FLRW) metric:

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a(t)^2\left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)\right]. \tag{1.10}$$

Here:

- *(r, θ, ϕ)* are the comoving polar coordinates.
- t is the proper time (or cosmic time) measured by observers at rest with respect to the comoving coordinates.
- $a(t)$ is the scale factor, which accounts for the expansion of the Universe over time.
- k is the curvature parameter, which can take three possible values (-1, 0, +1), corresponding to a Universe with negative, zero, or positive curvature, respectively.

1.1.3 Expansion of the Universe

In cosmology, multiple definitions of distance are used to describe separations between objects or events. A fundamental definition is the *proper distance*, which refers to the distance between two events at the same cosmic time (i.e., when $\mathrm{d}t = 0$). This type of distance gives the instantaneous separation between two points in space at a fixed time.

To derive the proper distance, we start with the general metric expressed along a specific line of sight, where the angular components vanish ($\mathrm{d}\theta = 0$ and $\mathrm{d}\phi = 0$):

$$d_{pr}(t) := \int_0^r \frac{\mathrm{d}r'^2 a(t)}{\sqrt{1-kr'^2}} = a(t)F(r). \tag{1.11}$$

In this definition, $F(r)$ is a function that accounts for the spatial part of the FLRW metric

$$F(r) := \int_0^r \frac{\mathrm{d}r'^2}{\sqrt{1-kr'^2}},$$

and it assumes different forms depending on the geometry of the Universe. In particular, if:

- $k = 0 \rightarrow F(r) = r$,
- $k = +1 \rightarrow F(r) = \arcsin(r)$,
- $k = -1 \rightarrow F(r) = \operatorname{arcsinh}(r)$.

The proper distance has a time dependence. If we set the time to the present time t_0, then the proper distance is called *comoving distance*

$$d_C := d_{pr}(t_0) = a(t_0)F(r) = \frac{a(t_0)}{a(t)} d_{pr}(t). \tag{1.12}$$

Since the proper distance depends on time, it can be derived with respect to t in order to find an expression for the radial velocity:

$$V_r(t) = \frac{\mathrm{d}[d_{pr}(t)]}{\mathrm{d}t} = \frac{\mathrm{d}[a(t)F]}{\mathrm{d}t} = \frac{\dot{a}}{a} aF = \frac{\dot{a}}{a} d_{pr} \tag{1.13}$$

This is the Hubble-Lemaître law and describes the radial velocity of an object due to the expansion of the Universe. $H := \dot{a}/a$ is called Hubble parameter and it encodes all the information about the expansion of the Universe. As different values for the Hubble constant were obtained from direct measurements, like Cepheids and Type Ia supernovae, and indirect measurements from anisotropies in the CMB, the evaluation of the Hubble parameter at the present time t_0 is still an open problem [131]. The most recent estimations of the Hubble constant are $H_0 \approx 67.4$ km s^{-1} Mpc^{-1} from

CMB observations [9] and $H_0 \approx 74$ km s^{-1} Mpc^{-1} from local standard candles [132], highlighting a Hubble tension of approximately 10% in the estimated value.

Eq.(1.13) does not indicate the existence of a special point as the center of the expansion. To understand why it is not so, consider an observer measuring the position $\vec{l}$ and the velocity $\vec{v}$ of a galaxy, assuming a non-relativistic regime ($|\vec{v}| \ll c$). The relation between these two quantities is expressed by the recession law $\vec{v} = H_0\vec{l}$, obtained from the Hubble-Lemaître law at the present time t_0. let us consider now another observer, being at the position $\vec{l'}$. This observer will have velocity $\vec{v'} = H_0\vec{l'}$ relative to the first one. From the non-relativistic composition law, the velocity of the galaxy measured by the second observer is

$$\vec{V} = \vec{v} - \vec{v'} = H_0(\vec{l} - \vec{l'}) = H_0\vec{L}, \tag{1.14}$$

where $\vec{L}$ denotes the position of the galaxy relative to the second observed. This highlight the fact that the linearity condition makes the Hubble's law to hold for all comoving observers, in agreement with the Cosmological Principle.

The gravitational constant G is not the only physical constant that appears in Eq.(1.5). In the standard cosmological framework, the cosmological constant Λ also contributes to the energy content of the Universe. This term was originally introduced in Einstein's field equations to prevent the dynamical behaviour of continuous expansion or contraction of the Universe. In this sense, the cosmological constant assumes the role of stabilizer, yet this equilibrium is inherently unstable. However, if it were found that this constant alone is inadequate to explain the observed acceleration, it would be necessary to explore new approaches, such as introducing a new source of energy known as dark energy, or modifying the laws of gravity.

As explained by the Hubble-Lemaître law Eq.(1.13), the Universe actually presents dynamical features, in particular it is continuously expanding. However this expansion does not happen with constant velocity, but rather it is accelerated. A very first evidence of an accelerated expansion of the Universe was obtained from Type Ia supernovae [130], [118] and it was later supported by baryonic acoustic oscillations [49]. More recent studies [78, 24, 66] provide strong evidence of an accelerating Universe. Other studies, such as those conducted with the *Planck satellite* [9], have measured fluctuations in the CMB radiation, revealing crucial insights about the nature of dark energy.

In summary, while Hubble-Lemaître law offers a clear view of the relationship between distance and recessional velocity of galaxies, the presence of the cosmological constant invites us to reflect on a continuously evolving Universe, where the acceleration of expansion presents a mystery yet to be solved.

1.1.4 Cosmological redshift

Consider a photon emitted at a distance r from an astrophysical source at time t_{em}. This photon is detected by an observer at the time t_{obs}. From the metric Eq.(1.10), setting the line of sight such that $\mathrm{d}\theta = \mathrm{d}\phi = 0$ and remembering that photons always follow null geodesics, we can obtain an expression for the time element:

$$\mathrm{d}t^2 = a^2(t)\frac{\mathrm{d}r^2}{1-kr^2}.$$

Integrating this equation between t_{em} and t_{obs} leads to

$$\int_{t_{em}}^{t_{obs}} \frac{\mathrm{d}t}{a(t)} = \int_0^r \frac{\mathrm{d}r'}{\sqrt{1-kr'^2}} = F(r).$$

Now suppose the astrophysical source emits another photon at time $t_{em} + \delta t_{em}$. This second photon will be detected by the observer at time $t_0 + \delta t_0$, and if the two photons travel across the same distance r, then

$$\int_{t_{em}}^{t_{obs}} \frac{\mathrm{d}t}{a(t)} = F(r) = \int_{t_{em}+\delta t_{em}}^{t_{obs}+\delta t_{obs}} \frac{\mathrm{d}t}{a(t)}.$$

This gives us the relation:

$$\frac{\delta t_{obs}}{a(t_{obs})} = \frac{\delta t_{em}}{t_{em}}. \tag{1.15}$$

Eq.(1.15) can be expressed both in terms of frequency

$$a(t_{obs})\nu_{obs} = a(t_{em})\nu_{em} \tag{1.16}$$

and in terms of wavelenghts

$$\frac{\lambda_{obs}}{\lambda_{em}} = \frac{a(t_{obs})}{a(t_{em})}. \tag{1.17}$$

These equations illustrate that the observed wavelength (or frequency) of a photon reflects the change in the scale factor a. We can build a function which is a proxy of the expansion of the Universe and that grows linearly with $a(t)$ back in time. This is the *cosmological redshift*

$$z := \frac{\lambda_{obs} - \lambda_{em}}{\lambda_{em}} = \frac{\Delta\lambda}{\lambda}. \tag{1.18}$$

By setting the observed time to t_0 as reference, we find a linear relation between the redshift and the scale factor

$$1 + z = \frac{a(t_0)}{a(t)}. \tag{1.19}$$

At then present time, $t = t_0$ and so $z = 0$. Since $a(t_0) > a(t)$ for $t > t_0$, observations indicate that the scale factor is monotonic and has been growing up to the present day, it follows that $z(t_0) < z(t)$.

1.2 Friedmann models

Although Einstein's idea of a static Universe was incorrect, a positive cosmological constant has been recently reintroduced into the Einstein equations to explain the late-time accelerated expansion of the Universe. As a result, the stress-energy tensor in Eq.(1.5) can be replaced by the effective stress-energy tensor, accounting for an additional term

$$\tilde{T}_{\mu\nu} := T_{\mu\nu} + \frac{\Lambda}{8\pi G} g_{\mu\nu}. \tag{1.20}$$

Friedmann models treat the Universe as an ideal fluid, neglecting viscosity and thermal conduction, so that the stress-energy tensor can be expressed solely in terms of the fluid's energy density ρ and pressure p:

$$T^{\mu\nu} = -pg^{\mu\nu} + (p+\rho)u^\mu u^\nu \tag{1.21}$$

where u^μ and u^ν are the components of the 4-velocity vector of the fluid. In the absence of anisotropic stress, the stress-energy tensor reduces to a diagonal matrix:

$$T^\mu_\nu = \begin{pmatrix} -\rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix}. \tag{1.22}$$

The tensorial equation Eq.(1.5) contains 16 equations due to the various combinations of indices μ and ν. However, due to the symmetry of the metric tensor g, this number reduces to 10 independent components. Furthermore, the freedom in choosing coordinates (see Sec.2.1.1) allows for an additional reduction, leaving 6 independent components. Finally, the assumptions of homogeneity and isotropy inherent in the FLRW metric reduce this to only 2 independent components. The metric for a flat Universe can be written in Cartesian coordinates as diagonal matrix thanks to the constraints of the Cosmological Principle:

$$g_{\mu\nu} = diag(-1, a^2(t), a^2(t), a^2(t)). \tag{1.23}$$

One may prove that, for the metric Eq.(1.23), the Levi-Civita connection Eq.(1.3) reduces to

$$\Gamma^0_{00} = \Gamma^0_{0i} = \Gamma^0_{i0} = \Gamma^i_{00} = \Gamma^i_{\alpha\beta} = 0 \tag{1.24}$$

$$\Gamma^0_{ij} = \delta_{ik}\dot{a}a = \delta_{ij}a^2 H \tag{1.25}$$

$$\Gamma^i_{0j} = \Gamma^i_{j0} = \delta_{ik}\frac{\dot{a}}{a} = \delta_{ik}H, \tag{1.26}$$

where $\delta_{ij} = diag(1,1,1)$ is the Kroneker delta. This is not surprising: all the terms off diagonal must vanish due to the symmetry of the metric as well as the $time - time$ g_{00}

component since it is a constant and all the partial derivatives ∂_μ are zero. Besides, all the *space* $-$ *space* components g_{ij} depend only on time, thus only the time derivation ∂_0 survives during the computation. To solve the Einstein's field equations Eq.(1.5), we need to compute the Ricci tensor and the Ricci scalar by plugging Eq.(1.3) in the definition of $R_{\mu\nu}$. From the *time-time* components

$$R_{00} = -3\frac{\ddot{a}}{a}, \tag{1.27}$$

while from the *space-space* components we obtain

$$R_{ij} = \delta_{ij}(2\dot{a}^2 + \ddot{a}a). \tag{1.28}$$

The Ricci's scalar is defined as the contraction of the Ricci's tensor with the metric:

$$R = R_{\mu\nu}g^{\mu\nu} = R_{00}g^{00} + R_{ij}g^{ij} = 6\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2\right]. \tag{1.29}$$

Finally, we can plug these results together with the expression of the stress-energy tensor for an ideal fluid in the Einstein's field equations. From the *time-time* component we obtain the first Friedmann equation for a flat Universe

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho, \tag{1.30}$$

while from the *space-space* components we obtain the second Friedmann equation

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(3p + \rho). \tag{1.31}$$

Even if the Friedmann equations are the two independent equations of the tensorial equation Eq.(1.5), they are not entirely independent since they are linked by the *adiabaticity condition* as the Universe is an isolated system ($\delta Q = 0$) and the energy must be conserved. So, from the first principle of thermodynamics

$$\mathrm{d}U = -p\mathrm{d}V \to \mathrm{d}(\rho a^3) = p\mathrm{d}a^3, \tag{1.32}$$

as the volume scales like the cube of the scale factor. This equation puts constraints on how the energy density of each fluid component evolves with the Universe's expansion.

At this point, we can introduce some useful parameters that embody the information of physical relevance in a compact way. We have already encountered the Hubble parameter

$$H(t) = \frac{\dot{a}(t)}{a(t)}, \tag{1.33}$$

which quantifies the expansion rate of the Universe. Another useful quantity is the density parameter of the fluid s

$$\Omega_s(t) = \frac{\rho_s(t)}{\rho_{crit}} = 8\pi G \frac{\rho_s(t)}{3H_0^2}, \tag{1.34}$$

describing the evolution of the fluid density. We can now express the first Friedmann equation in a more convenient form, taking into account the mixture of species present in the cosmological fluid

$$\frac{H^2(t)}{H_0^2} = \sum_s \Omega_s(t). \tag{1.35}$$

1.2.1 Equation of state

The equation of state is a function $p_s = p_s(\rho_s)$ that relates the pressure of a fluid species s to its energy density. In cosmology, the generic form for the equation of state is

$$p_s = w_s \rho_s, \tag{1.36}$$

with w_s dimensionless constant, whose value is set by the species. In particular:

- *non-relativistic matter*: $p = Nk_BT \approx 0 \rightarrow w_s = 0$,
- *radiation and relativistic matter*: $p = \frac{1}{3}\rho \rightarrow w_s = \frac{1}{3}$,
- *cosmological constant*: $p = -\rho \rightarrow w_s = -1$.

From the adiabaticity condition in Eq.(1.32), we can directly express the energy density of the fluid species s as a function of the scale factor a

$$\rho(t) \propto a^{-3(1+w_s)} \propto (1+z)^{3(1+w_s)}. \tag{1.37}$$

It is now possible to reconstruct the evolution of each fluid species through cosmic time Eq.(1.1). The present day value of each fluid species can be determined from Eq.(1.34):

$$\rho_{0,s} = \frac{3\Omega_{0,s} H_0^2}{8\pi G} = 1.88 \times 10^{-29} \Omega_{0,s} h^2 \text{ g cm}^{-1}, \tag{1.38}$$

where we used the dimensionless Hubble constant

$$h := H_0/(100 \text{ km s}^{-1} \text{ Mpc}^{-1}). \tag{1.39}$$

The dominance of one fluid component over others defines the three distinct epochs in the Universe's history, each marked by changes in the evolution of energy density with redshift, as shown in Fig.1.1:

- *radiation dominated*: the first cosmological epoch sees a prevalence of radiation in the cosmic mixture. This epoch ended at the *matter-radiation equivalence* around 10^3 years after the Big Bang. In terms of redshift it corresponds to

$$\rho_{0,m}(1+z_{eq}^{rm})^3 = \rho_{0,r}(1+z_{eq}^{rm})^4 \rightarrow z_{eq}^{rm} \approx 10^4. \tag{1.40}$$

- *matter dominated*: after the matter-radiation equivalence time, the new epoch of matter domination began. This epoch lasted up to 5 million years ago at the Λ-*matter equivalence*, when the last cosmic epoch dominated by the cosmological constant began.

$$\rho_{0,\Lambda} = \rho_{0,m}(1+z_{eq}^{DM\Lambda})^3 \rightarrow z_{eq}^{DM\Lambda} \approx 0.7. \tag{1.41}$$

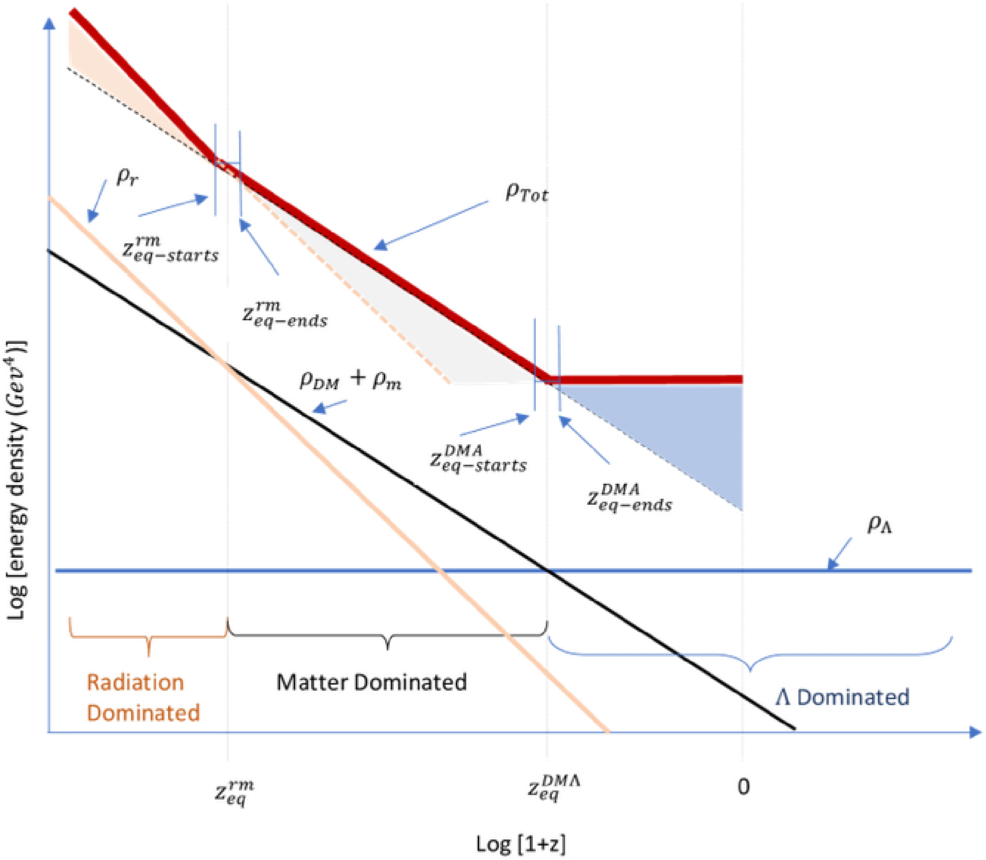


Figure 1.1: Evolution of the radiation, matter and cosmological constant energy densities. On the x-axis, the redshift corresponding to the radiation-matter equivalence and Λ-matter equivalence are reported. The solid red line describes the evolution of the total energy density of the Universe. Credits for the figure to [113].

Finally we can express the first Friedmann equation Eq.(1.35) using the content of the equation of state as follows:

$$H^2(z) = H_0^2 \sum_s \Omega_{0,s}(1+z)^{3(1+w_s)}. \tag{1.42}$$

1.2.2 Curved Universe

To generalize Eq.(1.35) for a non-flat Universe, we need to include an additional term that accounts for the contribution of the curvature parameter k

$$H^2(z) = H_0^2 \left(\sum_s \Omega_{0,s}(1+z)^{3(1+w_s)} + \Omega_{0,k}(1+z)^2 \right), \tag{1.43}$$

where $\Omega_k(z) = 1 - \sum_s \Omega_s(z)$ quantifies the deviation from flatness. Notice how, at high redshift, the dominant term is the energy density of the fluid because it appears with the highest power of $(1+z)$. Consequently, under such condition, the expansion rate $H(t)$ of Eq.(1.35) approaches the Einstein-de Sitter limit (single-component cosmological fluid and flat Universe) with solution

$$t = \frac{2}{3} \frac{1}{H_0 \Omega_0^{1/2} (1+z)^{3/2}}, \tag{1.44}$$

which is independent of the cosmological constant and the curvature $\Omega_k(z)$. This means that every curved Universe tends toward the flat one at very early times.

Depending on the value of $\Omega_0 = \sum_s \Omega_{0,s}$, we have different curved models:

- *closed models* ($\Omega_0 > 1$): in closed models, the Universe will eventually reach a maximum scale factor. By setting $\dot{a} = 0$, we can find the maximum scale factor corresponding to

  $$a_{max} = a_0 \left(\frac{\Omega_0}{1 - \Omega_0} \right)^{1/(1+3w)}. \tag{1.45}$$

  Since Eq.(1.35) is quadratic, the symmetry implies that the expansion will eventually reverse. Therefore, a closed Universe will reach a vanish scale factor at $2t(a_{max})$, corresponding to the *Big Crunch*.

- *open models* ($\Omega_0 < 1$): for open models, the scale factor $a(t)$ is monotonically growing and so it never reaches a maximum. Solving the Eq.(1.35), we can verify that during the radiation and matter era the growth of the scale factor goes linearly with time

  $$H(t) \propto \frac{1}{t}, \tag{1.46}$$

  and so the Universe expands forever.

The evolution of the scale factor $a(t)$ is reported in Fig.1.2 and underlines how the three possible geometries converge to the same one in the early Universe.

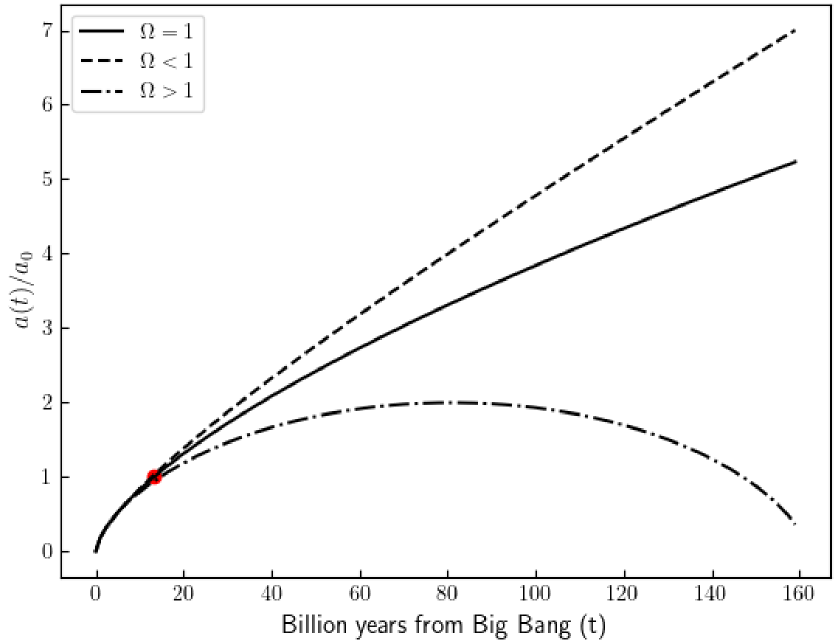


Figure 1.2: Evolution of the scale factor in Universes with different geometries. The dot-dashed line represents the evolution for a closed Universe ($\Omega > 1$) while the dashed black line for an open Universe ($\Omega < 1$). The solid line stands for a geometrically flat Universe ($\Omega = 1$). The red dot corresponds to the present time.

1.3 ΛCDM model

The concepts described so far lay the foundation for the concordance model of cosmology: the Λ-cold dark matter model. This model, which forms the cornerstone of modern cosmology, will be instrumental in conducting the parity violation test in the next chapter. Let's first break down the meaning of the acronym ΛCDM.

1.3.1 Λ - the cosmological constant

This symbol represents the cosmological constant, introduced in Einstein's field equations and responsible for the accelerated expansion of the Universe. It acts as a source of gravitational repulsive force, counteracting the smaller scales gravitational pull of matter. The value of Λ is derived from observational data and geometric constraints. Its current estimate is approximately: $\Lambda \approx 1.1 \times 10^{-48}\ cm^{-2}$, which is extremely small.

A widely accepted physical interpretation links the cosmological constant to vacuum energy. According to quantum physics, even "empty" space contains energy, arising from quantum fluctuations confined by the limits of detectability imposed by the Heisenberg uncertainty principle. However, there is a serious discrepancy between theoretical predictions and observations: quantum field theory suggests a vacuum energy density about 120 orders of magnitude larger than what cosmology infers from Λ. Resolving this enormous mismatch remains one of the biggest challenges in modern physics, known as the cosmological constant problem.

1.3.2 CDM - cold dark matter

Cold dark matter refers to a form of matter that does not emit, absorb, or interact with electromagnetic radiation, making it invisible to telescopes. The concept of dark matter was firstly suggested by Zwicky in 1933 [177], who discovered a new type of matter referred to as *Dunkle Materie* (missing matter) which is much more abundant than the ordinary matter. This hypothesis was then supported by several observational facts:

- Galactic rotation curves: Observations of the rotational velocity of stars within galaxies show that they rotate faster than can be explained by visible matter alone, implying the presence of an unseen mass component [169].

- Gravitational lensing: Light from distant objects is bent more strongly by galaxy clusters than would be expected from their visible mass, indicating additional mass in the form of dark matter [167].

- Cosmic Microwave Background: The temperature fluctuations in the CMB, along with their statistical properties, suggest the need for dark matter to explain the growth of structures from the early Universe to the present [46].

- LSS formation: Models of the growth of cosmic structures require dark matter to produce the observed distribution of galaxies and galaxy clusters [57].

It is "cold" at the decoupling in the sense that its particles move slowly compared to the speed of light, meaning their kinetic energy is low. This property ensures that dark matter clumps efficiently in the early Universe over large scales, driving the formation of large structures like galaxies and galaxy clusters through gravitational attraction. With only "hot" dark matter, virialized structures with masses less than $10^{16}\ M_{\odot}$ could not have formed. This rules out "hot" candidates like neutrinos as the dominant dark matter component.

The exact nature of cold dark matter remains one of the great unsolved mysteries of cosmology. While it is widely accepted that dark matter exists, its particle composition is still unknown. Several candidates have been proposed, such as weakly interacting massive particles (WIMPs), axions, sterile neutrinos, primordial black holes and white holes but none of them possess sufficient empirical support to be considered as a solution to the CDM problem. Indeed, direct detection experiments, as well as particle collider experiments like those conducted at the Large Hadron Collider, aim to detect or produce dark matter particles, but so far, no definitive detection has been made.

1.3.3 Inflation and the six ΛCDM model parameters

The ΛCDM model requires one additional ingredient for its complete description: *inflation.* This mechanism is responsible for generating the initial perturbations in the very early Universe that eventually grew into the structures we observe today. During the inflationary epoch, which lasted from 10^{-36} s to 10^{-32} s according to the standard inflationary scenarios, the scale factor a increased exponentially over time [65]. Specifically, the number of *e-foldings* is expected to be

$$N = \ln\left(\frac{a_f}{a_i}\right) \gg 60,$$

where a_f and a_i represent the scale factor at the end and beginning of inflation, respectively.

In a parity-violating scenario, inflation may also be responsible for breaking symmetry, leaving a statistical imprint on the primordial perturbations. In particular, particle exchanges during inflation can generate non-Gaussianities in these primordial fluctuations, as new particles are produced through inflaton decay [77, 17]. Additionally, other parity-violating inflationary models propose a phase of magnetogenesis as the driver of the symmetry breaking [62].

Finally, we can summarize the concordance model of cosmology as a geometrically flat Universe, dominated by CDM and Λ, with initial perturbations produced by the mechanism of inflation. This model is remarkably successful in predicting the evolution of the Universe while relying on only a few key parameters. Here is an overview of the primary parameters used in constructing the ΛCDM model:

- **Hubble constant $\mathbf{H_0}$**: As already introduced, this parameter describes the current expansion rate of the Universe. Its value from indirect measurements from CMB [9] is

  $$H_0 \approx 67.4 \text{ km s}^{-1} \text{ Mpc}^{-1},$$

  while the value derived from local distance ladders [132] is

  $$H_0 \approx 73 \text{ km s}^{-1} \text{ Mpc}^{-1}.$$

- **Matter density parameter $\mathbf{\Omega_{0,m}}$**: This parameter is defined as the current total matter density of the Universe, comprehensive of CDM and baryonic matter, normalized to the critical density. It rules the hierarchical formation of structures. Its typical value is:
$$\Omega_{0,m} \approx 0.31.$$
- **Baryonic density parameter $\mathbf{\Omega_{0,b}}$**: It refers to the fraction of Universe's energy density attributed to the ordinary matter (baryons).
$$\Omega_{0,b} \approx 0.049.$$
- **Dark energy density parameter $\mathbf{\Omega_{\Lambda}}$**: It represents the energy density associated to the cosmological constant, responsible for the accelerated expansion of the Universe.
$$\Omega_{\Lambda} \approx 0.69.$$
- **Curvature parameter $\mathbf{\Omega_{0,k}}$**: Associated to the current geometry of the Universe, $\Omega_k \approx 0$ in ΛCDM model, meaning that measures of angles fulfill the rules of Euclidean geometry.
- **Spectral index $\mathbf{n_s}$**: Describes the shape of the primordial power spectrum $P(k) \propto k^{n_s}$.
$$n_s \approx 0.96.$$
- **Amplitude of fluctuations $\boldsymbol{\sigma_8}$**: Measures the strength of matter clustering on scales of 8 h^{-1} Mpc,
$$\sigma_8 \approx 0.8.$$
- **Optical depth $\boldsymbol{\tau}$**: This parameter quantifies the Universe's transparency to photons during reionization, which occurred when the first luminous objects (likely Pop. III stars) ionized the surrounding gas. The free electrons produced during reionization scattered CMB photons via Compton scattering, leaving a distinct imprint on the CMB polarization. This helps us determine when the reionization epoch took place.
$$\tau \approx 0.054.$$

These results are obtained from *Planck 2018* data [9].

Chapter 2

Clustering

Clustering is a fundamental concept in cosmology, describing the tendency of matter to group together into cosmic structures, rather than being uniformly distributed throughout space.

As discussed in the previous chapter, the early Universe exhibits an almost homogeneous matter distribution, as evidenced by observations of the CMB. However, small density fluctuations present in the early stages of cosmic evolution grew over time, leading to the formation of galaxies, galaxy clusters, and the vast filamentary structures observed today. These perturbations are expected to form in the primordial Universe where small matter fluctuations are allowed from the energy-time uncertainty condition $\Delta E \Delta t \geq \hbar/2$. Their existence is proven by the small fluctuations of temperature and density in the CMB at z around 1100

$$\frac{\delta T}{T} \sim \frac{\delta \rho}{\rho} \sim 10^{-5}. \tag{2.1}$$

The process of structure formation is driven primarily by gravitational forces, which draw matter toward over-dense regions, thereby amplifying these initial inhomogeneities.

In this chapter, we will introduce tools and methods to study the spatial distribution of cosmic structures, a critical aspect for testing theories of parity violation.

2.1 Metric perturbations

The homogeneous Universe evolved trough a thermal history, started at the Big Bang and lasted up to the formation of the first atoms, Big Bang nucleosynthesis and production of dark matter. The structure formation is complicated and requires approximations, relying on the fact that perturbations are small in magnitude, to solve the equations that govern the evolution of cosmological species.

Let us start by adding small perturbations $|h_{\mu\nu}| \ll 1$ to the FLRW metric Eq.(1.10):

$$\begin{cases} g_{00}(t,\vec{x}) = -1 + h_{00}(t,\vec{x}) \\ g_{0i}(t,\vec{x}) = g_{i0}(t,\vec{x}) = a(t)h_{0i}(t,\vec{x}) \\ g_{ij}(t,\vec{x}) = a^2(t)[\delta_{ij} + h_{ij}(t,\vec{x})]. \end{cases} \tag{2.2}$$

While for the flat homogeneous Universe we had only one degree of freedom, that is the scale factor a, for the perturbed metric Eq.(2.2) we have 10 additional degrees of freedom given by the independent components of the symmetric four-dimensional perturbation tensor $h_{\mu\nu}$. To keep the discussion general, let us characterize each component of the perturbation.

- h_{00}: the *time-time* component is a 3-scalar, therefore it is invariant under any spatial rotation. Given an arbitrary scalar field A, this component can be written as:

  $$h_{00} = -2A \tag{2.3}$$

  where the prefactor -2 is a convention.

- h_{0i}: the *time-space* perturbation is a 3-vector and can be decomposed using the *Helmholtz decomposition theorem*, which states that any sufficiently smooth vector field that decays rapidly at infinity can be uniquely decomposed into the sum of an irrotational component, represented by a scalar potential, and a solenoidal component, represented by a vector potential. Introducing an arbitrary 3-scalar field $B(t,\vec{x})$ and a solenoidal 3-vector field $B_i(t,\vec{x})$, the perturbation can be expressed as the sum of a longitudinal and a transverse part:

  $$h_{0i} = -\frac{\partial B}{\partial x^i} - B_i. \tag{2.4}$$

- h_{ij}: the *space-space* component is a 3-dimensional symmetric tensor. Generalizing the Helmholtz decomposition theorem, we introduce two arbitrary 3-scalar fields $D(t,\vec{x})$ and $E(t,\vec{x})$, as well as a solenoidal vector field $V_i(t,\vec{x})$. However, this decomposition provides only 8 constraints (4 scalar functions + 2 transverse vectors) despite the fact that the perturbation tensor has 10 independent components. This means that we are missing two degrees of freedom, which cannot be expressed as combinations of scalars or vectors: these are the transverse-traceless components h_{ij}^{TT}. Thus, the full decomposition of the spatial perturbation is given by:

  $$h_{ij} = 2D\delta_{ij} - 2\frac{\partial^2 E}{\partial x^i \partial x^j} + \frac{\partial V_i}{\partial x^j} + \frac{\partial V_j}{\partial x^i} + h_{ij}^{TT}. \tag{2.5}$$

This decomposition is general and can be applied to any tensor. It leads to an important result, known as the *decomposition theorem*, which states that 3-scalar, 3-vector, and 3-tensor perturbations evolve independently at linear order, thanks to the symmetry of the FLRW metric. This result is crucial in cosmology, as it allows the independent study of scalar perturbations, which are responsible for the growth of structures.

2.1.1 Conformal Newtonian gauge

In General Relativity, a choice of coordinates is often referred to as *gauge* and it is a key aspect to look for gauge invariance under transformation of coordinates in order to reduce the number of free parameter in the equations and so allowing for the resolution. Consider a generic scalar field $\phi(\vec{x}, t)$ given by the application of a small scalar perturbation $\delta\phi(\vec{x}, t)$ to a background term $\bar{\phi}(t)$

$$\phi(\vec{x}, t) = \bar{\phi}(t) + \delta\phi(\vec{x}, t). \tag{2.6}$$

We aim to understand how this field changes under a generic small coordinates transformation $x^\mu \rightarrow \hat{x}^\mu(x^\mu)$. This requirement is necessary to keep the magnitude of perturbations small as well and to allow a Taylor expansion of the transformed coordinates:

$$\begin{cases} \hat{t} = t + \zeta(t, \vec{x}) \\ \hat{x}^i = x^i + \dfrac{\partial \xi(t, \vec{x})}{\partial x^i}, \end{cases} \tag{2.7}$$

where ζ and ξ represent, respectively, the time and space shifts and are treated as first order perturbations. By leveraging the scalar transformation law, we obtain the rule for scalar perturbations transformations:

$$\hat{\delta\phi}(\hat{t}, \vec{\hat{x}}) = \delta\phi(\hat{t}, \vec{\hat{x}}) - \frac{\mathrm{d}\bar{\phi}(\hat{t})}{\mathrm{d}\hat{t}} \zeta(t, \vec{\hat{x}}). \tag{2.8}$$

Taking the decomposition for scalar perturbations only, we focus on the scalar fields A, B, D, and E as defined in Eq.(2.3), Eq.(2.4), Eq.(2.5). By applying the tensor transformation rule to the metric components $g_{\mu\nu}$, we can derive the expressions for the transformed scalar fields:

$$\begin{cases} \hat{A} = A - \dot{\zeta} \\ \hat{B} = B - a^{-1}\zeta + a\dot{\xi} \\ \hat{D} = D - \dfrac{\dot{a}}{a}\zeta \\ \hat{E} = E + \xi. \end{cases} \tag{2.9}$$

This transformation is useful for reducing the degrees of freedom: the scalar perturbation transformation rule is determined by four scalar functions (A, B, D, E), which themselves depend on two spacetime shift functions (ζ, ξ). Consequently, we are left with two degrees of freedom. By appropriately choosing the shift functions, we can set two equations from Eq.(2.9) to zero. This approach underlies the conformal Newtonian gauge, where we select (ζ, ξ) such that B and E vanish.

So, we have seen how choosing specific coordinates reduces the number of degrees of freedom in the metric. This coordinate choice, known as *gauge freedom*, is responsible for the reduction from 10 to 6 degrees of freedom mentioned in Sec.1.2.

A common expression for the FLRW metric with scalar perturbations involves the standard Bardeen potentials Φ_A and Φ_H [18]:

$$\Phi_A := A + \frac{\partial}{\partial t}(a^2\dot{E} - aB) \tag{2.10}$$

$$\Phi_H := -D + \dot{a}(B - aE). \tag{2.11}$$

Thus, the perturbed metric can be expressed as:

$$\mathrm{d}s^2 = -(1 + 2\Phi_H)\mathrm{d}t^2 + a(t)^2[(1 - 2\Phi_A)\delta_{ij}]\mathrm{d}x^i\mathrm{d}x^j \ , \tag{2.12}$$

Finally, we can apply the conformal Newtonian gauge conditions $E = B = 0$ to the Bardeen potentials, yielding $\Phi_H = -\phi$ and $\Phi_A = \psi$. In this gauge, the perturbed FLRW metric takes the form:

$$\begin{cases} g_{00}(t, \vec{x}) = -1 - 2\psi(t, \vec{x}) \\ g_{i0}(t, \vec{x}) = g_{0i}(t, \vec{x}) = 0 \\ g_{ij}(t, \vec{x}) = a^2(t)\delta_{ij}[1 + 2\phi(t, \vec{x})], \end{cases} \tag{2.13}$$

where ψ is the well-known Newtonian potential and ϕ is the local perturbation of the scale factor. Both the scalar fields have magnitudes $< 10^{-4}$, making them suitable for linear theory.

2.1.2 Boltzmann equation

Now that we have made a gauge choice for the metric perturbations, we can proceed to derive the equations governing the evolution of perturbations in each fluid species of the ΛCDM model, excluding the perturbations of the cosmological constant.

In cosmology, we do not track the evolution of individual particles but rather study their statistical distributions. Let's consider a collection of particles in the phase space $(\vec{x}, \vec{p})$, where the position and momentum of each particle are sufficient to describe the system's dynamics. Due to the uncertainty principle, it is impossible to determine the exact position of a particle in phase space, limiting our knowledge of a particle's location to within an uncertainty of $\hbar$. Therefore, the phase space for each particle is not a collection of discrete points but is instead divided into six-dimensional cells with volumes given by quantum units, $\Delta V = \Delta\vec{x}\Delta\vec{p}/(2\pi\hbar)^3$, as illustrated in Fig.2.1. The number of particles within a single phase-space cell can be computed as,

$$N_{part} = f_1(\vec{x}, \vec{p}, t)dV = f_1(\vec{x}, \vec{p}, t)\frac{\Delta\vec{x}\Delta\vec{p}}{(2\pi\hbar)^3}, \tag{2.14}$$

where $f_1(\vec{x}, \vec{p}, t)$ is the 1-particle distribution function and describes the probability density to find a particle in a specific state. For a collisionless system, there are two well-known equilibrium distribution functions that describe particle evolution in phase space: the *Fermi-Dirac* distribution for fermions and the *Bose-Einstein* distribution for bosons.

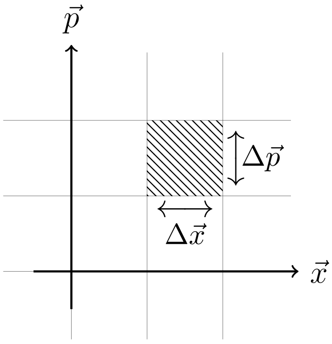


Figure 2.1: Discretization of the phase space.

In the context of cosmology, especially when studying the growth of structures, particles are often out of equilibrium, necessitating the calculation of distribution functions to understand the evolution of cosmological species.

For a generic system with a large number of interacting particles, the dynamics are described by the *Bogoliubov–Born–Green–Kirkwood–Yvon* (BBGKY) hierarchy. This hierarchy consists of a series of coupled equations that relate the n-particle distribution function to the $(n+1)$-particle distribution function:

$$f_n(\vec{x}^{\ (n)}, \vec{p}^{\ (n)}, t) \sim \int \mathrm{d}V f_{n+1}(\vec{x}^{\ (n+1)}, \vec{p}^{\ (n+1)}, t). \tag{2.15}$$

This chain of equations involves internal dependencies extending up to the total number of particles in the system. This means that the calculation of whatever distribution function reduces to the computation of the distribution function of all the particles of the system, which is infeasible. Thus, the problem becomes the closure of the hierarchy to some level. Approximations are necessary to truncate the sequence in Eq.(2.15) at a suitable order. Specifically, truncating at the first equation is permitted under the assumption of molecular chaos (*Stosszahlansatz*), leading to the Boltzmann equation.

For a collisionless system, the Boltzmann equation conserves the total number of particles and takes the form:

$$\frac{\mathrm{d}f_n(\vec{x}, \vec{p}, t)}{\mathrm{d}t} = 0 \rightarrow \nabla_x\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} + \nabla_p\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} + \frac{\partial f}{\partial t} = 0. \tag{2.16}$$

In this case, the equation describes the flux of particles in and out of a phase-space volume element. The collisionless nature of the system ensures conservation of the occupied phase-space volume, as described by *Liouville's theorem.* When interactions between particles are included, however, this conservation no longer holds, so the Boltzmann equation becomes non-zero. Specifically, it gains a collision term, $C[f]$, that accounts for all particle interactions (conserving four-momenta) and quantum effects such as differences between bosons and fermions, *Pauli blocking*, and *Bose enhancement*:

$$\frac{\mathrm{d}f_n(\vec{x},\vec{p},t)}{\mathrm{d}t} = 0 \rightarrow \nabla_x \frac{\mathrm{d}\vec{x}}{\mathrm{d}t} + \nabla_p \frac{\mathrm{d}\vec{p}}{\mathrm{d}t} + \frac{\partial f}{\partial t} = C[f]. \tag{2.17}$$

Perturbed Boltzmann equation

To complete the picture, the Boltzmann equation requires the inclusion of the perturbation of the metric described in the conformal Newtonian gauge in Eq.(2.13). Starting from the general form given in Eq.(2.17), we rewrite the equation by explicitly using the physical momentum $\vec{p} = p \cdot \hat{p}$:

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x^i}\frac{\mathrm{d}x^i}{\mathrm{d}t} + \frac{\partial f}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}t} + \frac{\partial f}{\partial \hat{p}^i}\frac{\mathrm{d}\hat{p}^i}{\mathrm{d}t}. \tag{2.18}$$

The four-momentum expression can be derived from the mass-shell condition and is written as

$$P^\mu = (P^0, P^i) = \left(E(1-\psi), \frac{p^i}{a}(1-\phi)\right), \tag{2.19}$$

where E is the total energy of the system. Using this expression, the Boltzmann equation can be reformulated to depend on energy, scalar potentials, and physical momenta. The perturbed Boltzmann equation then takes the form:

$$\begin{aligned}\frac{\mathrm{d}f}{\mathrm{d}t} =& \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x^i}\left[\frac{p}{E}\frac{\hat{p}^i}{a}(1+\psi-\phi)\right] - \frac{\partial f}{\partial p}\left[(H+\dot{\phi})p + \frac{E}{a}\hat{p}^i\psi_{,i}\right] \\ &+ \frac{\partial f}{\partial p^i}\frac{E}{ap}\left[\delta^{ik} - \hat{p}^i\hat{p}^k\right]\left(\frac{p^2}{E^2}\phi - \psi\right)_{,k} = C[f]\end{aligned} \tag{2.20}$$

where the notation $_{,i}$ denotes a partial derivative with respect to x^i. We are now prepared to analyze how the Boltzmann equation applies specifically to each cosmological species. Working in Fourier space, so substituting the wave vector $\vec{k}$ for the spatial coordinate $\vec{x}$, the Boltzmann equations take a simplified form where spatial derivatives are replaced by ik factors. The set of perturbed Boltzmann equations in ΛCDM cosmology is given by:

$$\theta' + ik\mu\theta + \phi' + ik\mu\psi + \tau'\left[\theta_0 - \theta + \mu u_b - \frac{1}{2}\mathcal{P}_2(\mu)\Pi\right] = 0, \tag{2.21}$$

$$\delta_c' + iku_c + 3\phi' = 0, \tag{2.22}$$

$$u_c' + \frac{a'}{a}u_c + ik\psi = 0, \tag{2.23}$$

$$\delta_b' + iku_b + 3\phi' = 0, \tag{2.24}$$

$$u_b' + \frac{a'}{a}u_b + ik\psi - \frac{4}{3}\tau'\frac{\rho_\gamma}{\rho_b}[u_b + 3i\theta_1] = 0, \tag{2.25}$$

$$\mathcal{N}' + ik\mu\frac{p}{E_\nu(p)}\mathcal{N} - Hp\frac{\partial\mathcal{N}}{\partial p} + \phi' + ik\mu\frac{E_\nu(p)}{p}\psi = 0. \tag{2.26}$$

Here and afterward, the prime apex denotes the derivative with respect to conformal time η which is related to the physical time through the relation $\mathrm{d}\eta = a^{-1}\mathrm{d}t$. For the full derivation, we refer to [43], which serves as the primary reference for most of the results and derivations in this chapter. This set of equations is essential, as it encapsulates the laws governing the dynamics of perturbations in the linear regime. Capturing the evolution of density, velocity, and other perturbations for different components of the Universe, such as CDM, baryons, and photons, each equation corresponds to a distinct fluid or species in the presence of metric perturbations ϕ and ψ in the conformal Newtonian gauge.

Generally, solutions to Eqs.(2.21)-(2.26) are computed numerically using specialized codes known as *Boltzmann solvers*. Widely used codes like `CLASS` [95] and `CAMB` [96] calculate matter power spectra, providing theoretical predictions crucial for various cosmological analyses. In particular, we will use `CAMB` in the next chapter to calculate theoretical power spectra, supporting trispectrum estimation and the outcome of the parity test.

Here's an overview of each equation:

- Eq.(2.21) describes the evolution of temperature perturbations, $\theta := \mathrm{d}T/T$, with respect to the Bose-Einstein photon distribution. The terms θ' and $ik\mu\theta$ account for the effects of photon motion, specifically the *free-streaming* behaviour of photons. Here, $\mu := \hat{p}^i\hat{k}^i$ captures the directional dependence between the perturbations and the momenta of the photons.

  The terms ϕ' and $ik\mu\psi$ introduce metric perturbations, representing how gravitational potentials affect photon energy. The scattering term, $\tau'[\theta_0 - \theta + \mu u_b - 1/2\ \mathcal{P}_2(\mu)\Pi]$, models interactions between photons and baryons through Compton scattering. Here, $\theta_0 := \frac{1}{2\pi}\int \mathrm{d}\Omega'\theta(\hat{p'}, \vec{x}, t)$ represents the monopole term, while $\vec{u}_b$ denotes the bulk velocity of the baryonic fluid. The last term, where $\Pi := \theta_2 + \theta_{p,2} + \theta_0$ and $\mathcal{P}_2$ is the Legendre polynomial, accounts for the angular dependence of the

Compton scattering, involving the quadrupole θ_2 and its photon polarization field $\theta_{p,2}$.

Notice that in the absence of bulk velocity, the collision terms drive θ toward the monopole term θ_0. In scenarios of very efficient scattering, only the monopole term survives, as the temperature anisotropies are effectively washed out. Intuitively, this implies that when the mean free path of a photon is very short, all photons tend to reach the same local temperature. Conversely, bulk velocity is necessary to generate multipole terms and consequently create anisotropies in the temperature distribution.

Finally, this equation is linear, indicating that perturbations grow linearly over time. This linear growth implies that perturbations do not increase significantly throughout cosmic time, allowing the equation to remain valid for CMB photons.

- In Eq.(2.22), δ_c denotes the number density contrast of CDM. The evolution of CDM plays a fundamental role in structures formation as it clumps efficiently and creates potential wells that gather baryonic matter (*baryonic drag*). Since CDM is non-relativistic and interacts only gravitationally, it is treated as an effective fluid so the collision term is set to zero and the Boltzmann equation assumes a simple form.

  However, since in the equation appears also the bulk velocity u_c, we need another equation to fix the hierarchy for CDM. This is the Euler equation Eq.(2.23) for the conservation of the velocity. Note that, since CDM is cold, the second moment of the Boltzmann distribution vanishes and the hierarchy is closed.

- In cosmology, baryons include electrons, protons, neutrons, helium, and trace amounts of heavier nuclei. Typically, the proton mass serves as a representative mass for baryons. Protons are coupled to electrons through Coulomb scattering, and the rate of these collisions is much higher than the Universe's expansion rate before recombination. This coupling allows us to use a common overdensity value for both protons and electrons. After recombination, baryonic matter behaves non-relativistically, and the continuity equation for baryons, Eq.(2.24), becomes identical to that of CDM as the collision term vanishes due to the conservation of the total electron and proton numbers.

  In contrast, the Euler equation for baryons, before recombination, includes the effects of Compton scattering between photons and electrons. Here, the dipole term, defined as $\theta_1(k,\eta) = i\int_0^1 d\mu\mu/2\ \theta(\mu,k,\eta)$, appears in the scattering term. This dipole term encapsulates the anisotropic motion of electrons in response to temperature gradients: electrons moving toward hotter regions encounter a "head-wind" effect, causing them to shift in the opposite direction. This phenomenon is referred to as *Compton drag*.

- Eq.(2.26) governs the evolution of temperature perturbations, $\mathcal{N} := \mathrm{d}T/T$, with respect to the Fermi-Dirac neutrino distribution. Since neutrinos are decoupled and poorly interacting at the times of interest, the collision term is set to zero. The term $ik\mu\mathcal{N}p/E_\nu(p)$ captures the neutrino free-streaming features, while $Hp\,\partial\mathcal{N}/\partial p$ accounts for redshifting effects on neutrino momenta.

  In this discussion, neutrinos are considered as massless; however, it is important to underline that neutrinos have a small mass (with current constraints indicating $\sum m_\nu < 0.072$ [35]), and several non-standard cosmological models account for neutrino mass. In fact, cosmological models themselves serve as powerful tools to constrain neutrino masses through accurate power spectra measurements.

2.1.3 Perturbed Einstein's field equations

In the previous section, we discussed how to address non-gravitational interactions in the presence of scalar perturbations in the metric Eq.(2.13). Now we turn to gravity, whose behaviour is governed by Eq.(1.5).

The first step is to incorporate scalar perturbations of the metric into the expressions for the Levi-Civita connections, enabling us to compute the left-hand side of Einstein's field equations to linear order. This procedure is similar to the one described in Sec.1.2, and the results yield the perturbed form of Γ:

$$\Gamma^0_{00} = \dot{\psi}, \tag{2.27}$$

$$\Gamma^0_{i0} = \partial_i\psi, \tag{2.28}$$

$$\Gamma^i_{00} = a^{-2}\partial_i\psi, \tag{2.29}$$

$$\Gamma^0_{ij} = \delta_{ij}a^2[H + 2H(\phi+\psi) + \dot{\phi}], \tag{2.30}$$

$$\Gamma^i_{0j} = \delta_{ij}(H + \dot{\phi}), \tag{2.31}$$

$$\Gamma^i_{jk} = \phi(\partial_k\delta_{ij} + \partial_j\delta_{ik} - \partial_i\delta_{jk}). \tag{2.32}$$

Recovering the expression for the Ricci tensor Eq.(1.4) restricted to the time components only

$$R_{00} = \partial_\alpha\Gamma^\alpha_{00} - \partial_0\Gamma^\alpha_{0\alpha} + \Gamma^\beta_{00}\Gamma^\alpha_{\beta\alpha} - \Gamma^\beta_{0\alpha}\Gamma^\alpha_{\beta 0}, \tag{2.33}$$

we note that, for $\alpha = 0$, it follows that $R_{00} = 0$, while for $\beta = 0$ there are second-order terms. Since we are working at the linear order, we discard terms at higher orders, implying that both α and β must be spatial indices. The resulting expression for the perturbed *time-time* component of the Ricci tensor in Fourier space is:

$$R_{00} = -3\frac{\ddot{a}}{a} - \frac{k^2}{a^2}\psi - 3\ddot{\phi} + 3H(\dot{\psi} - 2\dot{\phi}). \tag{2.34}$$

For the spatial components, the computation is more complex as it involves all indices. Starting with the *space-space* components of the Ricci tensor,

$$R_{ij} = \partial_\alpha \Gamma^\alpha_{ij} - \partial_j \Gamma^\alpha_{i\alpha} + \Gamma^\beta_{ij}\Gamma^\alpha_{\beta\alpha} - \Gamma^\beta_{i\alpha}\Gamma^\alpha_{\beta j}, \tag{2.35}$$

we substitute the results for the Levi-Civita connections to obtain the perturbed expression for the spatial Ricci tensor:

$$\begin{aligned} R_{ij} = \delta_{ij}[&(2a^2H^2 + a\ddot{a})(1 + 2\phi - 2\psi) \\ &+ a^2H(6\phi_{,0} - \psi_{,0}) + a^2\phi_{,00} + k^2\phi] + k_ik_j(\phi + \psi). \end{aligned} \tag{2.36}$$

By contracting the components $R_{\mu\nu}$ with the metric components $g_{\mu\nu}$ yields the Ricci scalar:

$$\begin{aligned} R =&(-1 + 2\psi)\left[-3\frac{\ddot{a}}{a} - \frac{k^2}{a^2}\psi - 3\phi_{,,00} + 3H(\psi_{,0} - 2\phi_{,0})\right] \\ &+ \frac{1 - 2\psi}{a^2}\{3[(2a^2H^2 + a\ddot{a})(1 + 2\phi - 2\psi) + a^2H(6\phi_{,0} - \psi_{,0}) \\ &+ a^2\phi_{,00} + k^2\phi] + k^2(\phi + \psi)\}. \end{aligned} \tag{2.37}$$

By discarding all nonlinear terms in ψ and ϕ, we obtain the first-order part of the Ricci scalar:

$$\delta R = -12\psi\left(H^2 + \frac{\ddot{a}}{a}\right) + 2\frac{k^2}{a^2}\psi + 6\phi_{,00} - 6H(\psi_{,0} - 4\psi_{,0}) + 4\frac{k^2}{a^2}\phi. \tag{2.38}$$

Now that we have completed the geometric side of Einstein's field equations by obtaining all necessary terms, let us turn to the energy side, which involves the stress-energy tensor. Based on the discretization of phase space illustrated in Fig.2.1, we can express the energy density of all particles in the Universe by summing the energy of each particle species s, weighted by the particle number and divided by the phase-space volume element:

$$T^0_0(\vec{x}, t) = -\rho(\vec{x}, t) = -\sum_s g_s \int \frac{\mathrm{d}^3p}{(2\pi)^3} E_s(p) f_s(\vec{x}, \vec{p}, t). \tag{2.39}$$

Here, g_s is known as the degeneracy parameter, representing the number of quantum states with the same energy that are accessible to a given particle. According to kinetic theory, macroscopic pressure relates to energy density by the expression $p = 1/3nm|v|^2$, where n is the particle number density in a given volume. By generalizing this relation to relativistic particles and applying it to Eq.(2.39), we derive the spatial components of the stress-energy tensor:

$$T^i_j(\vec{x}, t) = p(\vec{x}, t) = \sum_s g_s \int \frac{\mathrm{d}^3p}{(2\pi)^3} \frac{p^ip^j}{E_s(p)} f_s(\vec{x}, \vec{p}, t). \tag{2.40}$$

Finally, we have all the components needed to write the perturbed Einstein's field equations. For a complete and detailed derivation, refer to [43]. The scalar-perturbed Einstein's field equations take the form

$$k^2\phi + 3\frac{a'}{a}\left(\phi' - \psi\frac{a'}{a}\right) = 4\pi G a^2(\rho_c\delta_c + \rho_b\delta_b + 4\rho_\gamma\Theta_0 + 4\rho_\nu\mathcal{N}_0) \tag{2.41}$$

for the *time-time* component. This equation is particularly important for describing the evolution of modes larger than the Hubble radius H^{-1}. Notice that this equation reduces to the classical Poisson equation in the absence of expansion (a = constant). For the spatial component, we obtain

$$k^2(\psi + \phi) = -32\pi G a^2(\rho_\gamma\Theta_2 + \rho_\nu\mathcal{N}_2). \tag{2.42}$$

This equation is significant because, in the case of a vanishing quadrupole, such as under tight-coupling conditions, the perturbation potentials ϕ and ψ are equal in magnitude but opposite in sign.

2.2 Structure formation

We now explore the solutions to our equations Eqs.(2.21)-(2.26) and Eqs.(2.41)-(2.42), with initial conditions set by inflation. At late times, when the Universe is matter-dominated, gravity is mediated by the potentials ϕ and ψ. In contrast, during early radiation-dominated epochs, anisotropies are primarily related to monopole and dipole terms of radiation perturbations. This implies that while CDM is weakly coupled to radiation, radiation anisotropies still depend on CDM perturbations. We will focus on CDM perturbations, as CDM couples to other components solely through gravitational interactions and plays a dominant role in structure formation.

Overdensities on the order of $\delta\rho/\rho \sim 10^{-4}$ can accumulate enough matter over the age of the Universe to form observed cosmic structures. However, this growth process is counteracted by two opposing effects:

- Expansion of the Universe: The cosmic expansion tends to pull particles apart. In a static Universe, perturbations would grow exponentially, but with expansion, the growth of perturbations follows a power-law.

- Pressure from baryons and photons: While baryons and photons exert pressure that grows with density, CDM does not exhibit such behaviour.

2.2.1 Particle horizon

Before delving into structure formation, we must introduce the concept of the particle horizon, also known as the cosmological horizon. In Ch.1, we discussed the recession law and the tools necessary for calculating distances in spacetime. However, when using comoving coordinates, the speed of light is not constant in the usual sense. While the peculiar velocity of light remains c, the recession velocity, induced by the Universe's expansion, alters the total velocity of photons. Specifically, if a photon is emitted at time t_{em} and detected by an observer at time t_{obs}, then

$$V_{tot}(t_{em}) \neq V_{tot}(t_{obs}). \tag{2.43}$$

Since the trajectory of a photon (its geodesic) in spacetime represents the shortest possible path, everything contained within a light cone represents events that can have a causal connection with the observer. The distance traveled by a photon emitted at time t_{em} and observed at the present time t_0 (past light cone) can be calculated as:

$$r_{LC}(t_{em}) = \int_{t_{em}}^{t_0} \frac{dt'}{a(t')} = \int_0^z \frac{dz'}{H(z')}. \tag{2.44}$$

Extending the extremes of integration from the beginning of the Universe up to now leads to the definition of particle horizon

$$r_H(t) := \int_0^t \frac{dt'}{a(t')}, \tag{2.45}$$

which is equivalent to the definition of conformal time itself, $\eta(t)$. The particle horizon serves as a boundary for observable information, effectively defining the limits of the observable Universe. In this sense, it marks the boundary within which scientific inquiry can apply.

2.2.2 The evolution of perturbations

The evolution of cosmological perturbations is largely governed by the gravitational potential, ϕ. The behaviour of perturbation modes varies across cosmic epochs, primarily based on whether the modes are inside or outside the horizon. We can categorize the evolution of perturbations into three distinct stages:

- Early evolution ($k\eta \ll 1$): At this stage, the potential remains constant, and all perturbation modes are outside the horizon.
- Intermediate evolution ($k\eta \gtrsim 1$): As perturbations enter the horizon, they begin to evolve with the changing gravitational potential. Modes that enter the horizon before matter-radiation equality behave differently than those entering after, reflecting the impact of the evolving matter and radiation content.
- Late evolution ($k\eta \gg 1$): In this stage, the potential stabilizes, leading all modes to evolve similarly. However, as the Universe approaches Λ-matter equivalence, the gravitational potential declines due to the dominance of the cosmological constant.

To analyze CDM perturbations, we utilize the perturbed Einstein field equations Eq.(2.41) and Eq.(2.42) alongside the Boltzmann equations Eqs.(2.21)-(2.26). Given that photons are tightly coupled to electrons via Compton scattering during the radiation era, we can ignore terms above the dipole in the photon temperature perturbations. In the matter-dominated era, photons can be largely ignored as CDM dynamics prevail. A further simplification arises from the tight coupling condition, Eq.(2.42) leads to the relation $\phi = -\psi$. This allows us to express the gravitational potential in a straightforward manner, thereby closing the set of equations. These simplifications reduce the Boltzmann equations to the following core set:

$$\theta'_{r,0} + k\theta_{r,1} = -\phi', \tag{2.46}$$

$$\theta'_{r,1} - \frac{k}{3}\theta_{r,0} = -\frac{k}{3}\phi, \tag{2.47}$$

$$\delta'_c + iku_c = -3\phi', \tag{2.48}$$

$$u'_c + \frac{a'}{a}u_c = ik\phi. \tag{2.49}$$

Since there are no analytical solutions that are valid across all scales and times, we often resort to interpolating solutions when analytical derivation is not feasible. However, we can study specific limit cases where analytical solutions can be obtained:

- Super-horizon regime: In this regime, exact solutions can be derived that hold true throughout the entire evolution of the perturbations.
- Horizon entry: We can distinguish between early times, where in the small-scale approximation we can neglect the CDM density perturbation δ_c, and late times, where the large-scale approximation yields a constant gravitational potential ϕ.
- Sub-horizon regime: In this regime, solutions remain exact as we can neglect θ_r for small scales. Conversely, for large scales, the potential stabilizes to a constant value.

This framework helps us understand the dynamics of perturbations in different epochs of the Universe's evolution, offering insight into structure formation processes.

Super-horizon evolution

In the regime $k\eta \ll 1$, the wavelength dependence in Eq.(2.46), Eq.(2.47), and Eq.(2.41) can be ignored. This implies that $\delta_c - 3\theta'_{r,0}$ must remain constant, and by the adiabaticity condition, this constant is zero. To describe the evolution of perturbations, we introduce a parameter $y := \rho_m/\rho_r$. With this substitution, Eq.(2.41) can be rewritten as:

$$y + \frac{\mathrm{d}\phi}{\mathrm{d}y} + \phi = \frac{3y+4}{6(y+1)}\delta_c. \tag{2.50}$$

By differentiating both sides with respect to y, we obtain the following second-order differential equation

$$\frac{\mathrm{d}^2\phi}{\mathrm{d}y^2} + \frac{21y^2 + 54y + 32}{2y(y+1)(3y+4)}\frac{\mathrm{d}\phi}{\mathrm{d}y} + \frac{\phi}{y(y+1)(3y+4)} \tag{2.51}$$

which can be solved by introducing a new variable $u := \phi \; y^3/\sqrt{1+y}$. The analytical solution then becomes [82]:

$$\phi(\vec{k}, y) = \frac{1}{10y^3}\left[16\sqrt{1+y} + 9y^3 + 2y^2 - 8y - 16\right]\phi(\vec{k}, 0). \tag{2.52}$$

This solution reveals that, for small scales ($y \ll 1$), the potential remains constant at $\phi(0)$, while for large scales ($y \gg 1$), the potential asymptotes to $9/10\phi(0)$.

Horizon entry

When perturbations enter the horizon, their behaviour differs significantly depending on whether this occurs during the radiation-dominated era (small scales) or the matter-dominated era (large scales). We'll start by examining the latter scenario. Deep in the matter-dominated era, radiation contributions can be neglected. Under these conditions, Eq.(2.46) implies $\phi' = 0$, a constraint set by the initial conditions from super-horizon evolution. Next, we examine whether Eq.(2.48), Eq.(2.49), and Eq.(2.41) permit constant solutions for the potential.

In the matter-dominated era, the Hubble parameter H scales as $a^{-3/2}$. Using this fact, we can rewrite Eq.(2.48) as follows:

$$\frac{2k^2\phi'}{3a^2H^2} + \left[\frac{iu_c}{k} + \frac{2\phi}{3aH}\right]\left(\frac{9a^2H^2}{2} + k^2\right) = 0. \tag{2.53}$$

By differentiating this equation and neglecting all terms proportional to ϕ, we obtain a second-order equation for ϕ in the form

$$\alpha\phi'' + \beta\phi' = 0, \tag{2.54}$$

where α and β are two real constants. The fact that equations like Eq.(2.54) allow for constant solutions has significant implications for the evolution of the gravitational potential. Since the initial conditions set by super-horizon scales remain preserved as perturbations enter the horizon during the matter-dominated era, the gravitational potential ϕ remains constant.

This outcome is particularly important because it indicates that, in the matter era, the forces driving structure formation reach an equilibrium. The small-scale gravitational pull from overdense regions, which would typically cause matter to collapse and enhance the density contrast, is exactly counterbalanced by the large-scale gravitational expansion of the Universe, which tends to stretch and dilute structures. As a result, the gravitational potential stays constant over time, maintaining its initial value.

For small scales, perturbations enter the horizon during the radiation-dominated era, where the evolution of the potential is governed by radiation perturbations. In this regime, CDM perturbations do not significantly impact the potential but are instead affected by it. To proceed, we first compute the gravitational potential ϕ using Eq.(2.46), Eq.(2.47), and Eq.(2.41), discarding all matter source terms and utilizing the relation $aH = \eta^{-1}$, which holds in the radiation era. This yields a relationship between the gravitational potential and the dipole component of the radiation perturbation:

$$\phi' + \frac{1}{\eta}\phi = -\frac{6}{\eta^2 k}\theta_{r,1}. \tag{2.55}$$

To obtain a second-order equation, we differentiate this first-order equation, leading to

$$\phi'' + \frac{4}{\eta}\phi' + \frac{k^2}{3}\phi = 0, \tag{2.56}$$

which is the Fourier space form of a damped wave equation. Solving this requires transforming it into a spherical Bessel equation of the first order, with the solution:

$$\phi(\vec{k},\eta) = 2\left(\frac{\sin x - x\cos x}{x^3}\right)_{k\eta/\sqrt{3}} \mathcal{R}(\vec{k}). \tag{2.57}$$

Here, $\mathcal{R}(\vec{k})$ is the curvature perturbation predicted by inflation. This result indicates that, as Fourier modes enter the horizon, the potential oscillates and decays as η^{-2}.These oscillations reflect sound waves driven by the potential, where the decay is due to radiation pressure counteracting gravitational collapse, which prevents perturbation growth. This oscillatory process underlies the generation of BAOs.

Using the behaviour of ϕ, we can then solve Eq.(2.48) and Eq.(2.49) for CDM perturbations. By combining these equations and encapsulating terms containing the potential into a source term $S(k,\eta)$, we obtain the differential equation:

$$\delta_c'' + \frac{1}{\eta}\delta_c' = 0, \tag{2.58}$$

which, for the associated homogeneous equation ($S(k\eta) = 0$), has two solutions: $\delta_c =$ const, $\delta_c = \ln(\eta)$. The general solution is thus a linear combination of these homogeneous solutions, plus an additional term accounting for the source contribution via the Green's function. For $k\eta \sim 1$, the result is a function that grows logarithmically with the perturbation scale:

$$\delta_c(k,\eta) \propto \text{const} + \ln(k\eta). \tag{2.59}$$

This key result highlights that, although radiation modes are decaying, CDM perturbations can still grow during the radiation era, a direct consequence of CDM's lack of pressure.

Sub-horizon evolution

In the sub-horizon regime, large-scale modes have a constant potential if they enter the horizon after matter-radiation equality. For small scales, however, as $\rho_c \delta_c$ becomes dominant over $\rho_r \theta_{r,0}$, radiation can be neglected. In this context, the Eq.(2.48), Eq.(2.49) and Eq.(2.41) combine into the *Meszaros equation*:

$$\frac{\mathrm{d}\delta_c^2}{\mathrm{d}y^2} + \frac{2+3y}{2y(y+1)}\frac{\mathrm{d}\delta_c}{\mathrm{d}y} - \frac{3}{2y(y+1)}\delta_c = 0. \tag{2.60}$$

This hypergeometric equation admits two solutions: a growing mode

$$D_+(a) = a + \frac{2}{3}a_{eq}, \tag{2.61}$$

known as the growth factor, which drives structure formation, and a decaying mode

$$D_-(y) = \left(y + \frac{2}{3}\right) + \ln\left[\frac{\sqrt{1+y}+1}{\sqrt{1+y}-1}\right] - 2\sqrt{1+y}. \tag{2.62}$$

The general solution is a linear combination of these two modes, but at late times, the decaying mode becomes negligible.

Finally, we can express the general form of the observed gravitational potential, consolidating the results from the Boltzmann equation solutions. The late-time evolution of the gravitational potential is given by

$$\phi(\vec{k}, a) = \frac{3}{5}\mathcal{R}(\vec{k})T(k)\frac{D_+(a)}{a}, \tag{2.63}$$

where $T(k)$ is the *transfer function*, which captures the potential's decay as modes enter the horizon and pass the matter-radiation equality. The transfer function is normalized to 1 on large scales, such that

$$T(k) := \frac{\phi(\vec{k}, a_{\text{late}})}{\phi_{\substack{\text{large}\\\text{scale}}}(\vec{k}, a_{late})}, \tag{2.64}$$

where the large-scale potential equals the primordial value reduced by a factor of 9/10, as previously derived for large-scale modes. This transfer function is typically calculated using numerical codes, such as `CAMB` or `CLASS`, to propagate the primordial power spectrum through cosmic time. In the next section, we will introduce the concept of the power spectrum and explore its use in studying the distribution of galaxies.

2.3 Statistical properties of fields

Up to this point, we have established the formalism needed to describe the evolution of the cosmic fluid's energy and pressure, incorporating both the perturbed Boltzmann equation and Einstein's field equations within the framework of a scalar-perturbed FLRW metric in conformal Newtonian gauge. However, while we understand the laws governing structure formation in the linear regime, we lack a description of the spatial distribution of particles across the Universe. To address this, we must introduce statistical tools, such as N-point correlation functions, to understand how cosmic structures are arranged spatially and to compare the prediction of models to the numerical measurements. Correlation functions are designed to provide measurements of the mass distribution and its dynamical evolution. This aspect is critical for the objective of this work: if we aim to detect parity violation, we need to verify whether there is a statistically significant difference in the galaxy distribution under the inversion of spatial coordinates.

2.3.1 Two-point correlation function

We start by modelling the spatial distribution of galaxies as a realization of a statistically random stationary point process. This means that we consider the positions of galaxies in the Universe as distributed according to statistical rules, rather than following a completely deterministic model. In other words, we assume that, while galaxies tend to cluster in groups and clusters, their arrangement has a random component. Each possible configuration of galaxies in space that results from this probabilistic model is called a realization of the point process. If we could observe multiple "replicas" of the Universe, we would obtain different configurations of galaxies, each a unique realization of the same process. Clearly, the fact that the Universe is unique is a problem: we cannot average different samplings over different Universes. A common approach to overcome this issue is to assume the *ergodic hypothesis* that allows us to work with independent subdivisions of our Universe to obtain statistical relevance.

Consider two infinitesimal comoving volumes, $\mathrm{d}V_1$ and $\mathrm{d}V_2$. For a stationary point process, the probability of finding a galaxy in the center of an infinitesimal volume $\mathrm{d}V$ is given by the mean number density of galaxies, denoted by $\bar{n}$:

$$\mathrm{d}P = \bar{n}\mathrm{d}V. \tag{2.65}$$

By Bayes' theorem, the probability of two outcomes occurring can be expressed as the probability of the first outcome times the conditional probability of the second outcome, given the first:

$$\mathrm{d}P_{12} = \mathrm{d}P_1 \cdot \mathrm{d}P_{(2|1)} = \mathrm{d}P_2 \cdot \mathrm{d}P_{(1|2)}. \tag{2.66}$$

Thus, for two volume elements separated by a comoving distance r, we have the joint probability

$$\mathrm{d}P_{12} = \bar{n}^2\mathrm{d}V_1\mathrm{d}V_2[1 + \xi(r)] \tag{2.67}$$

where ξ is the two-point correlation function (2PCF). The value of ξ indicates the level of correlation between the two points:

- $\xi(r) = 0$: The objects are uncorrelated, meaning they are randomly distributed.
- $\xi(r) > 0$: The objects are positively correlated, so the probability of finding a pair of galaxies at a distance r is higher than in a random distribution.
- $-1 \leq \xi(r) < 0$: The objects are anti-correlated, meaning that the probability is lower at separation r than in a random distribution.

A practical method for estimating the 2PCF is to count galaxy pairs in observational data (denoted DD) and in a random distribution (denoted RR) at comoving separation r as follows [116]:

$$1 + \xi(r) = \frac{DD(r)}{RR(r)}. \tag{2.68}$$

Now, if we consider a sphere of radius r centered on a galaxy, we can calculate the mean number of galaxies within that sphere:

$$\langle N(< r)\rangle = \int_V \mathrm{d}V_2 \bar{n}[1 + \xi(r)] = \bar{n}V + \bar{n}\int_V \mathrm{d}V\xi(r) = \frac{4}{3}\pi r^3 \bar{n} + 4\pi \int_0^r \mathrm{d}r' r'^2 \xi(r'). \tag{2.69}$$

The second term represents the excess probability relative to a random distribution and tends to zero as $r \to \infty$, while the first term represents the mean galaxy number density in the Universe.

In a discrete model, if we partition the Universe into cells with at most one galaxy each, we can retrieve information about the distribution moments. Let $\langle n_i \rangle = \bar{n}\mathrm{d}V_i$ represents the mean density of the $i-th$ cell. Since there is at most one galaxy per cell, all moments of the galaxy count distribution in a single cell are equal:

$$\langle n_i \rangle = \left\langle n_i^2 \right\rangle = ... = \left\langle n_i^N \right\rangle. \tag{2.70}$$

The first moment of the ensemble of cells is simply the sum of first moments across all cells:

$$\langle N \rangle = \left\langle \sum_i n_i \right\rangle = \int_V \bar{n}\mathrm{d}V_i = \bar{n}V. \tag{2.71}$$

Similarly, the second moment is

$$\left\langle N^2 \right\rangle = \left\langle \sum_i n_i \sum_j n_j \right\rangle = \bar{n}^2 V^2 + \bar{n}^2 \int\int \mathrm{d}V_i \mathrm{d}V_j [1 + \xi(r)]. \tag{2.72}$$

We can express these results in terms of the normalized fluctuation around the expected number of objects:

$$\langle\Delta\rangle = \frac{N - \langle N\rangle}{\langle N\rangle}, \tag{2.73}$$

where the mean of this fluctuation distribution vanishes, and the variance is

$$\left\langle\Delta^2\right\rangle = \frac{1}{\bar{n}V} + \frac{1}{V^2}\int\int \mathrm{d}V_i \mathrm{d}V_j[1 + \xi(r)]. \tag{2.74}$$

In this equation, $1/(\bar{n}V)$ is termed *shot noise*, which represents statistical noise arising from discretization. This concept is common in counting experiments and will reappear during the construction of sky maps, where pixelization introduces noise that affects parity tests.

While the discussion above follows a discrete model, a continuous definition of the 2PCF can also be provided. If galaxies are viewed as probes of a continuous density field, correlation functions provide the average correlation level of density fluctuations. Considering the mass density at position $\vec{x}$ as $\rho(\vec{x}) := n(\vec{x})m$ with background density $\rho_b := \bar{n}m$, the probability of finding an object at $\vec{x}$ is

$$\mathrm{d}P(\vec{x}) = n(\vec{x})\mathrm{d}V = \frac{\rho(\vec{x})}{\rho_b}\bar{n}\mathrm{d}V. \tag{2.75}$$

Thus, the probability of finding two objects separated by $\vec{r}$ is:

$$\mathrm{d}^2P_{12} = \mathrm{d}V_1\mathrm{d}V_2 \left\langle n(\vec{x}+\vec{r})n(\vec{x})\right\rangle = \mathrm{d}V_1\mathrm{d}V_2\bar{n}^2\frac{\langle\rho(\vec{x}+\vec{r})\rho(\vec{x})\rangle}{\rho_b}. \tag{2.76}$$

By defining the density contrast

$$\delta(\vec{x}) := \delta\rho(\vec{x})/\rho = \frac{\rho(\vec{x}) - \rho}{\rho}, \tag{2.77}$$

and combining Eq.(2.67) and Eq.(2.76), we find that the 2PCF can be expressed as

$$\xi(|\vec{r}|) = \langle\delta(\vec{x}+\vec{r})\delta(\vec{x})\rangle \tag{2.78}$$

where $\vec{x}$ represents the position of a galaxy in space. An example of how the 2PCF appears at different redshifts is illustrated in Fig.2.2.

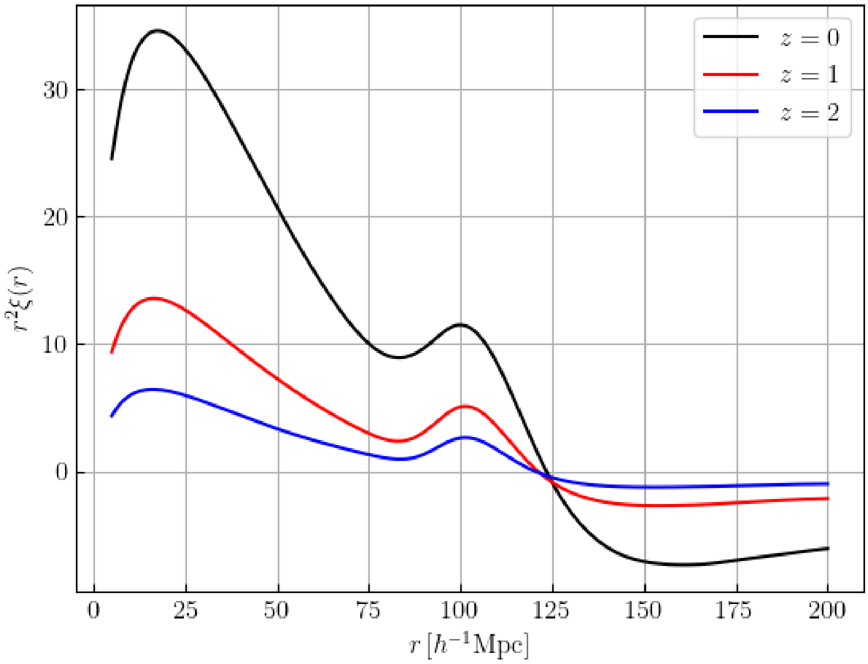


Figure 2.2: Nonlinear 2PCF monopole in real space, obtained with `CAMB` and halofit model [159], for a ΛCDM cosmology using parameters from *Planck 2018* (see definition in Sec.1.3.3). This plot highlights the BAO feature around 100 h^{-1} Mpc.

2.3.2 Power spectrum

Let's now delve into the concepts underlying the power spectrum. The latter is not new to us, as we encountered the primordial power spectrum in Sec.1.3.3 when discussing the six parameters of the ΛCDM model. According to standard inflationary models, the primordial power spectrum took shape in the early Universe under the assumption that no specific scale was favored. This leads to a scale-invariant distribution, meaning that fluctuations are statistically similar on all scales.

A power-law functional form satisfies this scale invariance, and thus the power spectrum at the end of inflation is typically parameterized by the gauge-invariant curvature perturbation, $\mathcal{R}$:

$$P_{\mathcal{R}}(k) := 2\pi^2 \mathcal{A}_s k^{-3} \left(\frac{k}{k_p}\right)^{n_s-1}, \tag{2.79}$$

where

- $\mathcal{A}_s$ is the scalar amplitude of perturbations, representing the variance of $\mathcal{R}$ around a reference scale k_p, known as the pivot scale.

- n_s is the spectral index, which modulates the distribution of power across different scales.

Scale invariance, which assumes a flat distribution of fluctuations, implies $n_s = 1$, giving a constant $P(k)$. This specific case is known as the *Harrison-Zel'dovich power spectrum*, introduced in [67, 175, 117]. The Harrison-Zel'dovich model reflects a Universe with a constant primordial gravitational potential, leading to a balanced distribution of power over all scales.

Considering a continuous density field in Fourier space, the matter power spectrum $P(k)$ is defined by the expectation value of the matter density contrast in Fourier space, denoted by $\tilde{\delta}(\vec{k})$:

$$\left\langle \tilde{\delta}(\vec{k})\tilde{\delta}^*(\vec{k}') \right\rangle := (2\pi)^3 P(k)\delta_D^{(3)}(\vec{k} - \vec{k}'). \tag{2.80}$$

Here, $\delta^{(3)}$ is the 3-dimensional Dirac delta defined as

$$\delta_D^{(3)}(\vec{k}) := \frac{1}{(2\pi)^3}\int_{\mathbb{R}^3} e^{i\vec{k}\cdot\vec{r}} \mathrm{d}^3\vec{r}, \tag{2.81}$$

and it enforces the independence of different modes in the expression above. From Eq.(2.78), it is clear that if we perform the Fourier transform of the 2PCF, we obtain the power spectrum. This important result is known as the *Wiener-Khinchin theorem*, explaining why the power spectrum can be used to study correlations in the distribution of objects.

The Fourier density contrast evolves as described by the growing solution of the linear perturbations:

$$\tilde{\delta}(\vec{k}, t) \propto D_+(t). \tag{2.82}$$

Suppose we have a set of particles with mass m_j and position $\vec{r_j}$. The Fourier amplitude is a sum over the positions [115]:

$$\tilde{\delta}(\vec{k}) \propto \sum_j m_j e^{-i\vec{k}\cdot\vec{r_j}}. \tag{2.83}$$

Let r_0 be the clustering length at which nonlinearity starts to be significant. Nonlinear interactions cause displacements $\Delta\vec{r_j}$ in the positions of the particles; we can estimate the effects on the Fourier amplitude by Taylor expanding the equation above:

$$\Delta\tilde{\delta}(\vec{k}) \propto \sum_j m_j \left[i\vec{k}\cdot\Delta\vec{r_j} - (\vec{k}\cdot\Delta\vec{r_j})^2/2 + \dots \right] e^{i\vec{k}\cdot\vec{r_j}}. \tag{2.84}$$

Since the interactions preserve momentum, we expect that each mass shift on the $j-th$ particle is balanced by an opposite shift of another particle. This means that the leading term $\vec{k} \cdot \Delta\vec{r}$ vanishes, so the perturbations to the power spectrum at scales $k^{-1} \gg r_0$ are of the order $(kr_0)^4$. Therefore, if the power spectrum approaches a zero value slower than k^4, then the nonlinear contribution is negligible [115].

At late times the evolution of baryons follow closely the one of CDM, so we can relate the total matter density contrast δ_m to the potential Eq.(2.63) through Poisson's equation in order to express the matter overdensity as:

$$\delta_m(\vec{k}, a) = \frac{2ka^2}{3\Omega_m H_0^2}\phi(\vec{k}, a). \tag{2.85}$$

Finally, we have an expression for the linear matter power spectrum at late times:

$$P_L(k, a) = \frac{8\pi^2}{25}\frac{\mathcal{A}_s}{\Omega_m^2}D_+^2(a)T^2(k)\frac{k^{n_s}}{H_0^4 k_p^{n_s-1}}, \tag{2.86}$$

whose evolution at different redshifts is shown in Fig.2.3.

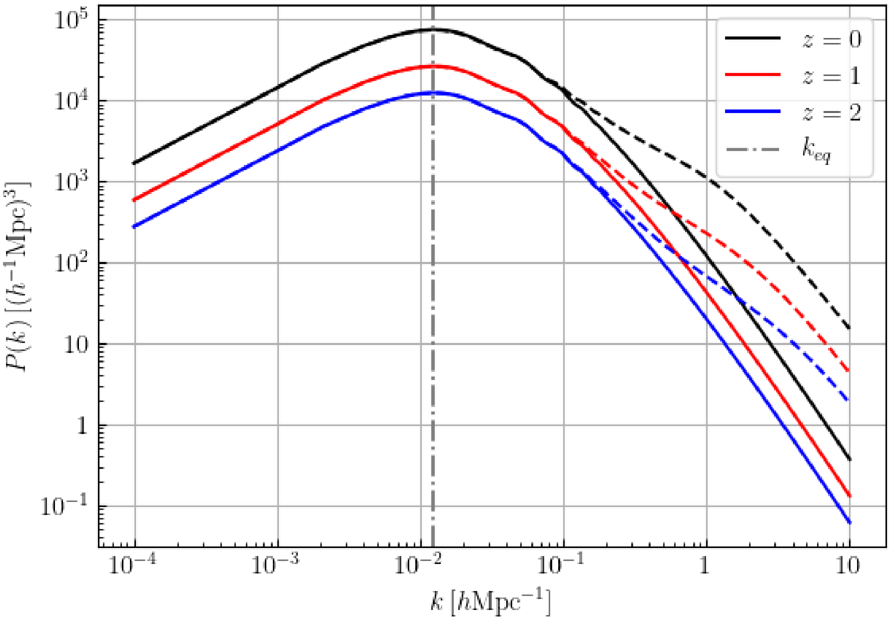


Figure 2.3: Real-space linear (solid lines) and nonlinear (dashed lines) matter power spectra for an ΛCDM cosmology with parameters from *Planck 2018* (see definition in Sec.1.3.3). The power spectra were computed using `CLASS`, with the nonlinear spectrum obtained via the halofit model [159]. The dot-dashed vertical line marks the scale at which perturbations enter the horizon at matter-radiation equivalence.

Whilst the power spectrum can be used to study the LSS of the Universe, it is not directly useful for investigating parity-violating features. This limitation is related to the concept of parity symmetry, which will be discussed in detail in the next chapter.

Chapter 3

Parity

Parity is the transformation that inverts the coordinates of a physical system. In 1956 two theoretical physicists, Tsung-Dao Lee and Chen-Ning Yang, noted that, while parity symmetry had been tested for strong and electromagnetic interactions, it had not been verified for weak interactions [92]. Chien-Shiung Wu provided empirical evidence of parity violation while studying the β-decay of cobalt-60 [171]:

$$^{60}_{27}\mathrm{Co} \rightarrow \, ^{60}_{28}\mathrm{Ni} + e^{-} + \bar{\nu}_{e} + \gamma.$$

The experiment involved measuring the helicity, which was obtained from spin orientation and the momentum of the emitted electron. To understand this phenomenon, it is helpful to introduce the concepts of *pseudo-vectora* and differential forms. A *1-form* is a function that maps an element of the tangent bundle of a smooth manifold M (denoted TM) to a real number (i.e., an element of the algebraic field $\mathbb{R}$):

$$\omega : TM \rightarrow \mathbb{R}, \tag{3.1}$$

with the property that ω is linear on each tangent space T_pM at each point $p \in M$. A *2-form* is a tensor field that can be seen as the section of the fiber of anti-symmetric (0,2) tensors. We can obtain a 2-form as the external product of two 1-forms. That is, if we have two vectors $\vec{v}$ and $\vec{w}$, we can build a 2-form $\omega(\vec{v}, \vec{w}) = \vec{v} \wedge \vec{w}$, which satisfies the following properties:

$$\omega(a\vec{v}, b\vec{w}) = a\omega(\vec{w}) + b\omega(\vec{v}) \quad \text{for each } a, b \in \mathbb{R} \text{ (bilinear)} \tag{3.2}$$

$$\omega(\vec{v}, \vec{w}) = -\omega(\vec{w}, \vec{v}) \quad \text{(anti-symmetric)}. \tag{3.3}$$

In the context of parity violation, 2-forms in 3-dimensional space can be associated with vectors through the Hodge dual operator. Specifically, when a 2-form is mapped via the Hodge dual, the resulting object behaves as a *pseudo-vector*; it is similar to a vector but transforms differently under coordinate reflections, not changing sign under parity inversion.

Spin is a *pseudo-vector*, meaning it does not change orientation under parity transformations, whereas momentum is a true vector and thus appears flipped when transformed. When measuring the alignment between spins and momenta, we categorize the events into two types: type A, where the spin and momentum are aligned, and type A', where they are anti-aligned. A statistically significant difference in the counts of type A and A' events indicates parity violation. This has been observed, as there is a directional preference in the emission of electrons, resulting in a greater number of type A events compared to type A' events. This important result is also connected to the fact that neutrinos, which interacts only through weak force, are uniquely left-handed. Right-handed neutrinos do not have weak charge and so they must interact only though gravitational interaction, and are called sterile neutrinos: these might be promising dark matter candidates, though being generally predicted to be of low mass, they constitute a component of the warm dark matter (WDM) [45].

Thus, the only known fundamental interaction in nature that does not preserve parity symmetry is the weak force. Therefore, examining signatures of parity violation is crucial for understanding processes that govern the early Universe and, eventually, for revealing unknown physical mechanisms that may cause breaks in parity symmetries in the distribution of galaxies.

3.1 Parity transformation in physics

Parity ($\mathbb{P}$) is a discrete transformation which satisfies the following conditions:

- $\mathbb{P}^2 = \mathbf{1}$: parity is a unitary operator, meaning it preserves the norm of the vector it is applied to. For a position vector $\vec{r}$, this implies that the resulting vector has the same magnitude as the original.
- $(\mathbb{P}^*)^{\mathrm{T}} = \mathbb{P}$: parity is a self-adjoint operator, which means that the parity operator is equal to its Hermitian adjoint (the conjugate transpose). This implies that $\mathbb{P}$ has real eigenvalues and can be diagonalized, as follows from the spectral theorem.

These properties imply that $\mathbb{P}$ can be represented by the following diagonal matrix:

$$\mathbb{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{3.4}$$

To illustrate the parity symmetry in gravitational theory, let us start with Newton's law for the gravitational force along $\vec{r}$ between two objects of masses m_1 and m_2:

$$\vec{F}_g = \frac{G m_1 m_2}{r^2}\hat{r}, \tag{3.5}$$

where $\hat{r} = (x\hat{\imath} + y\hat{\jmath} + z\hat{k})/|\vec{r}|$. Applying the parity transformation gives

$$\mathbb{P}\left[\vec{F}_g(x, y, z)\right] = \vec{F}_g(-x, -y, -z) = \vec{F}_g, \tag{3.6}$$

Therefore, both the magnitude and the direction of the force remain the same. This shows that Newton's gravitational law is invariant under parity reflection.

Now, let us examine why parity symmetry also holds in General Relativity by analyzing how the metric transforms under a parity transformation. The operation

$$\mathbb{P}g\mathbb{P}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \times \begin{pmatrix} g_{00} & g_{01} & g_{02} & g_{03} \\ g_{10} & g_{11} & g_{12} & g_{13} \\ g_{20} & g_{21} & g_{22} & g_{23} \\ g_{30} & g_{31} & g_{32} & g_{33} \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{3.7}$$

returns a transformed matrix g' in which the *time-time* and *space-space* components remain the same as in the original matrix, while the mixed *time-space* terms change sign:

$$g' = \begin{pmatrix} g_{00} & -g_{01} & -g_{02} & -g_{03} \\ -g_{10} & g_{11} & g_{12} & g_{13} \\ -g_{20} & g_{21} & g_{22} & g_{23} \\ -g_{30} & g_{31} & g_{32} & g_{33} \end{pmatrix}. \tag{3.8}$$

The determinant of g' equals that of g, so applying the parity transformation to the Einstein-Hilbert action yields the same form of the field equations, showing that Eq.(1.5) remains invariant under parity. Moreover, for the FLRW metric, which is diagonal, the parity transformation leaves the metric unchanged, leading to identical Friedmann equations, Eq.(1.30) and Eq.(1.31), as solutions to the Einstein's field equations.

Thus in General Relativity, and in particular in the ΛCDM framework, parity is conserved. To rigorously search for parity violations in non-standard models, we should explore possible detection methods, building on the statistical tools introduced in Ch.2. We have already discussed the 2PCF and the power spectrum, which relate to the probability of finding two galaxies at a comoving separation r. Higher-order statistics generalize this concept to larger numbers of configurations. For instance, considering three spatial volumes, $\mathrm{d}V_1$, $\mathrm{d}V_2$, and $\mathrm{d}V_3$, the probability of finding three galaxies centered within each of them is:

$$\mathrm{d}P = \bar{n}^3[1 + \xi_{12} + \xi_{23} + \xi_{31} + \zeta]\mathrm{d}V_1\mathrm{d}V_2\mathrm{d}V_3, \tag{3.9}$$

where ζ represents the three-point correlation function (3PCF) and $\xi_{12} := \xi(r_{12})$. Similarly, extending this approach to four volumes leads to the four-point correlation function η (4PCF) defined by [58]:

$$\begin{aligned} \mathrm{d}P = \bar{n}^4[&1 + \xi_{12} + \xi_{13} + \xi_{14} + \xi_{23} + \xi_{24} + \xi_{34} \\ &+ \xi_{12}\xi_{34} + \xi_{13}\xi_{24} + \xi_{14}\xi_{23} \\ &+ \zeta_{123} + \zeta_{124} + \zeta_{134} + \zeta_{234} + \eta]\mathrm{d}V_1\mathrm{d}V_2\mathrm{d}V_3\mathrm{d}V_4, \end{aligned} \tag{3.10}$$

where $\zeta_{123} := \zeta_(r_{12}, r_{23}, r_{31})$.

While the 2PCF considers pairs of galaxies, the 3PCF, considers triplets of galaxies, thus located at the vertices of triangles. Applying the parity transformation to such a configuration is effectively equivalent to a spatial rotation, as illustrated in Fig. 3.1. Therefore, detecting parity violation using the 2PCF or 3PCF is impossible, as these are equivalent to spatial rotations and remain symmetric due to the constraints of the Cosmological Principle. The lowest-order correlation function potentially sensitive to parity violation is the 4PCF, as illustrated in Fig.3.2. Similarly, this conclusion is valid in Fourier space. In this case studying parity-violating features in the galaxy distribution thus requires at least the trispectrum.

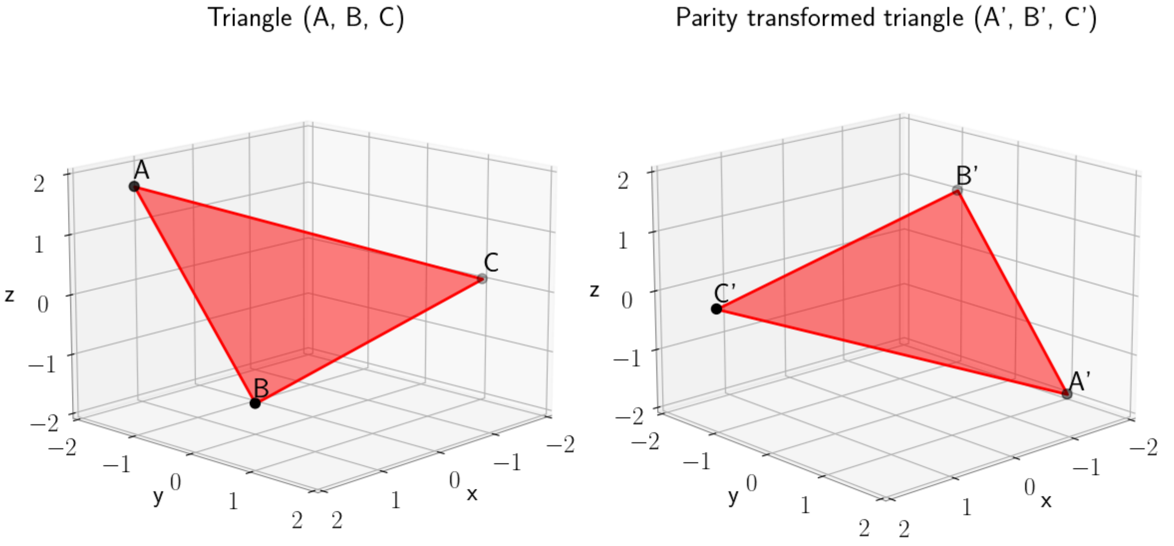


Figure 3.1: Applying a parity transformation to a two-dimensional triangle is equivalent to a spatial rotation by $\pi/2$ in three dimensions. This implies that the 3PCF can only violate parity in two-dimensional space.

A generic scalar field can be expressed under a parity transformation as

$$\mathbb{P} : \delta(\vec{x}) \to \delta(-\vec{x}), \tag{3.11}$$

and its Fourier modes will transform similarly:

$$\mathbb{P} : \delta(\vec{k}) \to \delta(-\vec{k}). \tag{3.12}$$

If $\delta(\mathbb{P}[\vec{k}]) = \delta(\vec{k})$, we say that the function is *parity-even*, while if $\delta(\mathbb{P}[\vec{k}]) = -\delta(\vec{k})$, the function is *parity-odd*. Moreover, if the scalar field is real-valued, as in the case of redshift, the Fourier modes will satisfy the condition

$$\delta(-\vec{k}) = \delta(\vec{k})^*, \tag{3.13}$$

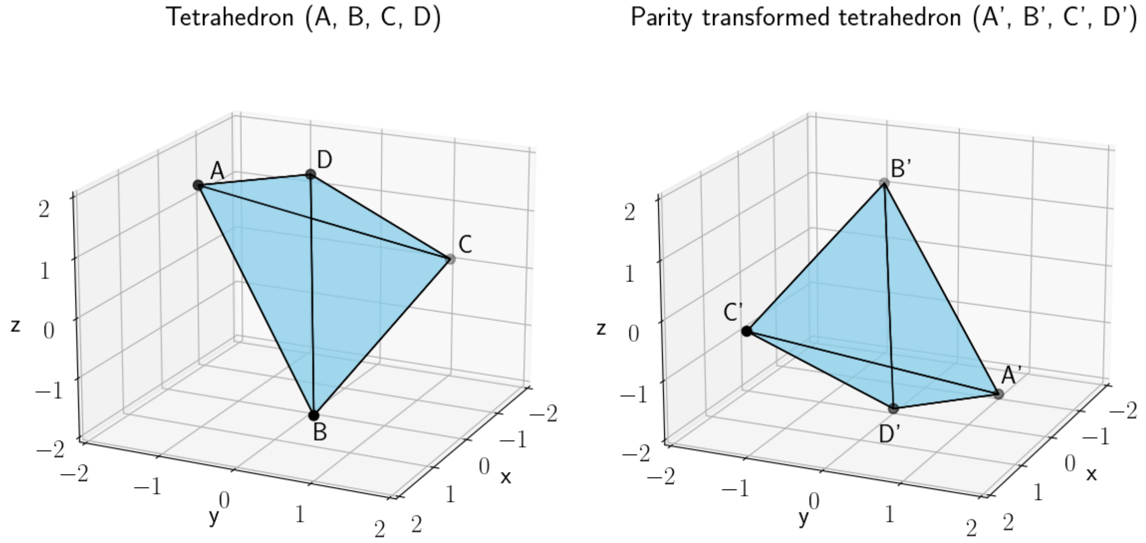


Figure 3.2: Tetrahedra are the lowest-order polygons capable of exhibiting parity-violating features in three dimensions. Studying parity violation requires the 4PCF or higher-order correlation functions.

meaning that the parity transformation replaces the modes of a field with the complex conjugate of the modes. This immediately leads to the conclusion that the power spectrum (as well as the 2PCF) cannot detect parity violation. From the definition in Eq.(2.80):

$$\left\langle \delta(\vec{k})\delta(\vec{k}')\right\rangle = \left\langle \delta(\vec{k})\delta(\vec{k})^*\right\rangle = \left\langle |\delta(\vec{k})|^2\right\rangle = (2\pi)^3\delta_D^{(3)}(\vec{k}+\vec{k}')P(k). \tag{3.14}$$

Since the 3-dimensional Dirac delta enforces translational invariance, it leads to a purely real power spectrum with only parity-even components. Therefore, the possibility of detecting parity violation hinges on whether we can have non-vanishing parity-odd components. For the bispectrum, which is the Fourier transform of the 3PCF, the situation is less immediate. In general terms, the sufficient condition for parity violation is the presence of *pseudo-scalars* quantities.

By definition, the bispectrum B is:

$$\left\langle \delta(\vec{k}_1)\delta(\vec{k}_2)\delta(\vec{k}_3)\right\rangle := (2\pi)^3\delta_D^{(3)}(\vec{k}_1+\vec{k}_2+\vec{k}_3)B(k_1,k_2,k_3), \tag{3.15}$$

and we can construct a pseudo-vector by taking the cross product of two Fourier modes. This is similar to the case of the angular momentum of a rotating object, which is a pseudo-vector obtained from the cross product of the position and momentum vectors of the object. For instance, $\vec{k}_1 \times \vec{k}_2$ is a pseudo-vector, and $\vec{k}_3 \cdot (\vec{k}_1 \times \vec{k}_2)$ is a pseudo-scalar (it changes sign under parity transformation). Due to the Dirac delta function, we have the constraint $\vec{k}_3 = -(\vec{k}_1 + \vec{k}_2)$, so our pseudo-vector must vanish. As a result, it is impossible to have parity-odd components in the bispectrum. This outcome is consistent with the fact that both the power spectrum and the bispectrum are parity-insensitive and real, which aligns with the geometrical construction shown in Fig.3.1.

Let us consider the trispectrum T, defined as

$$\begin{aligned}\left\langle \delta(\vec{k}_1)\delta(\vec{k}_2)\delta(\vec{k}_3)\delta(\vec{k}_4) \right\rangle := (2\pi)^3 \delta_D^{(3)}(\vec{k}_1 + \vec{k}_2 + \vec{k}_3 + \vec{k}_4) T(k_1, k_2, k_3, k_4, \\ |\vec{k}_1 + \vec{k}_2|, |\vec{k}_1 + \vec{k}_3|, \vec{k}_1 \cdot \vec{k}_2 \times \vec{k}_3).\end{aligned} \tag{3.16}$$

Differently from revious cases, here we have the possibility to construct irreducible non-vanishing pseudo-scalars through the triple product. For the trispectrum, we can have three possible linearly independent vectors to be used to construct pseudo-scalars in the form $\vec{k}_1 \cdot (\vec{k}_2 \times \vec{k}_3)$. Therefore, the parity-odd components of the trispectrum is proportional to the triple product, and the trispectrum can be decomposed into a purely real part (parity-even), which does not carry information about parity, and a purely imaginary part (parity-odd) containing the parity-violating features:

$$T = T^+ + iT^-. \tag{3.17}$$

Parity-odd components can be parametrized by introducing a six-dimensional function τ^- (reduced parity-odd trispectrum) that contains the shape of the parity-odd trispectrum:

$$\begin{aligned}T(k_1, k_2, k_3, k_4, |\vec{k}_1 + \vec{k}_2|, |\vec{k}_1 + \vec{k}_4|) =& i\vec{k}_1 \cdot |\vec{k}_2 \times \vec{k}_3| \\ & \tau^-(k_1, k_2, k_3, k_4, |\vec{k}_1 + \vec{k}_2|, |\vec{k}_1 + \vec{k}_4|).\end{aligned} \tag{3.18}$$

In Fourier space, we may visualize the trispectrum as a tetrahedron with four sides whose length is determined by the magnitude of the vectors $\vec{k}_1$, $\vec{k}_2$, $\vec{k}_3$, and $\vec{k}_4$. The other remaining sides are diagonals, so their length is given by the vector sum of the other sides, as in Fig.3.3. Thus, it is evident that $|\vec{k}_1 + \vec{k}_4| = |\vec{k}_2 + \vec{k}_3|$ and $|\vec{k}_1 + \vec{k}_3| = |\vec{k}_2 + \vec{k}_4|$.

3.2 Angular correlators

The trispectrum described above, which is the Fourier transform of the 4PCF, is a powerful tool for studying parity violation in the 3-dimensional LSS of the Universe. By analyzing the abundance of right- and left-handed configurations of galaxy quadruplets, it provides insights into potential parity-breaking symmetries [120, 38]. However, in this work we do not apply the statistical tools described above directly to the 3-dimensional

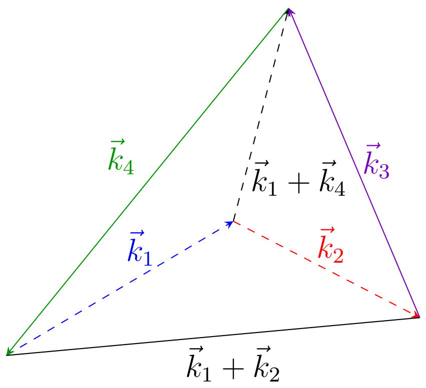


Figure 3.3: Fourier space tetrahedron representing a possible configuration for the trispectrum.

comoving space. Instead, our focus is on searching for parity violations within a 2-sphere framework, using the ARFs developed by C. H. Monteagudo in [69], which will be discussed in detail in the Ch.4.

To achieve this, we require a 2-dimensional version of angular correlations, represented by the projection of scalar fields onto the surface of a sphere.

3.2.1 Angular power spectrum and bispectrum

An arbitrary function defined on a 2-sphere, $\mathbb{S}^2$, can be expanded in a spherical harmonic series. Let us consider a generic zero-mean signal, labeled as a, that depends on the line of sight direction $\hat{n}$, which is specified by the azimuthal and polar angles (θ, ϕ), it can be expanded as follows:

$$a(\hat{n}) = \sum_{\ell=0}^{+\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{n}), \tag{3.19}$$

where $Y_{\ell m}(\hat{n}) = Y_{\ell m}(\theta, \phi)$ are the spherical harmonics, which are related to the Legendre polynomials by:

$$Y_{\ell m}(\theta, \phi) = \sqrt{\frac{2\ell+1}{2}\frac{(\ell-m)!}{(\ell+m)!}} P_{\ell m}(\cos\theta) \frac{e^{im\phi}}{\sqrt{2\pi}}, \tag{3.20}$$

a relationship that arises as a consequence of the *Sturm-Liouville theorem.*

The integer ℓ is known as the *angular quantum number* (or multipole) and describes the angular scale of the function on the 2-sphere. In particular:

- $\ell = 0$ is the monopole term and corresponds to an angular scale of $\theta = 2\pi$.
- $\ell = 1$ is the dipole term and corresponds to an angular scale of $\theta = \pi$.

- $\ell = 2$ is the quadrupole term and corresponds to an angular scale of $\theta = \pi/2$.
- $\ell > 3$ corresponds to progressively smaller angular scales, approximately $\theta \approx \pi/(6\ell)$.

The term m is called the *magnetic quantum number* (or azimuthal number), representing the longitudinal variation of the function along a specific axis. It is an integer within the range $-\ell \leq m \leq \ell$. The coefficients in the expansion, $a_{\ell m}$, are calculated by integrating the product of the signal $a(\hat{n})$ and the complex conjugate of the spherical harmonics over the 2-sphere:

$$a_{\ell m} = \int_{\mathbb{S}^2} \mathrm{d}\hat{n} \; a(\hat{n}) Y^*_{\ell m}. \tag{3.21}$$

The angular power spectrum is typically expressed as the average over statistical realizations of the signal as:

$$C^{\ell_1 \ell_2}_{m_1 m_2} := \left\langle a_{\ell_1 m_1} a^*_{\ell_2 m_2} \right\rangle = \delta_{\ell_1 \ell_2} \delta_{m_1 m_2} C_\ell, \tag{3.22}$$

where the Kronecker deltas impose statistical isotropy and homogeneity, ensuring no correlations between different angular scales of the field. As a result, the angular power spectrum depends only on the multipole ℓ, quantifying the signal's magnitude at a specific angular scale. In cosmology, this harmonic expansion is commonly applied to CMB temperature fluctuations to construct power spectra. The statistical isotropy and homogeneity of the angular power spectrum of the temperature field align with the Cosmological Principle. Since cosmic radiation is located at a specific redshift, it allows us to ignore the radial dependence of the temperature field.

Similarly, the angular bispectrum B is defined as:

$$B^{\ell_1 \ell_2 \ell_3}_{m_1 m_2 m_3} := \left\langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3} \right\rangle, \tag{3.23}$$

and, under isotropy and homogeneity conditions, it takes the form:

$$B^{\ell_1 \ell_2 \ell_3}_{m_1 m_2 m_3} = \mathcal{G}^{\ell_1 \ell_2 \ell_3}_{m_1 m_2 m_3} b_{\ell_1 \ell_2 \ell_3}. \tag{3.24}$$

Here, $b_{\ell_1 \ell_2 \ell_3}$ is the reduced bispectrum, while $\mathcal{G}$ is the Gaunt function, defined as the average over three spherical harmonics:

$$\begin{aligned} \mathcal{G}^{\ell_1 \ell_2 \ell_3}_{m_1 m_2 m_3} &:= \sqrt{\frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4\pi}} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \int_{\mathbb{S}^2} \mathrm{d}\hat{n} \, Y_{\ell_1 m_1}(\hat{n}) Y_{\ell_2 m_2}(\hat{n}) Y_{\ell_3 m_3}(\hat{n}) \end{aligned}, \tag{3.25}$$

where the 3×2 matrices are the *Wigner-3j* symbols, describing the combination of angular momenta. We distinguish between parity-even and parity-odd bispectrum components based on the parity of the sum of the three multipoles. The Gaunt function is non-zero only if $\ell_1 + \ell_2 + \ell_3$ is even, enforcing the fact that parity-odd bispectra do not exist in the absence of anisotropies [39].

3.2.2 Angular trispectrum

Finally, we introduce the tool used in this work for testing parity violation: the angular trispectrum. This statistics correlates four harmonic modes, thus containing redundant contributions (disconnected terms), given by lower-order correlations (such as between pairs of modes) combined inside the four-mode correlation. Therefore, the angular trispectrum is defined by taking the only connected part of the correlation of four harmonic modes:

$$T^{\ell_1\ell_2\ell_3\ell_4}_{m_1m_2m_3m_4} := \langle a_{\ell_1m_1}a_{\ell_2m_2}a_{\ell_3m_3}a_{\ell_4m_4}\rangle_c\,. \tag{3.26}$$

Unlike the bispectrum, which depends on three multipoles (one for each spherical harmonic), the trispectrum cannot be fully described by four multipoles. This is because, to achieve rotational invariance, we require an additional element called *effective angular momentum* (L), which represents the total angular momentum and is obtained by combining the other four harmonic modes. Specifically, L can be seen as the diagonal of the tetrahedron describing the angular trispectrum configuration in harmonic space. Thus, it must satisfy the following triangular relations:

$$|\ell_1-\ell_2| \le L \le \ell_1+\ell_2 \tag{3.27}$$

$$|\ell_3-\ell_4| \le L \le \ell_3+\ell_4, \tag{3.28}$$

which are induced by the Wigner-3j symbols. The isotropic and homogeneous angular trispectrum is

$$\langle a_{\ell_1m_1}a_{\ell_2m_2}a_{\ell_3m_3}a_{\ell_4m_4}\rangle_c = \sum_{L=0}^{+\infty}\sum_{M=-L}^{L}(-1)^M w^{L(-M)}_{\ell_1\ell_2m_1m_2}w^{LM}_{\ell_3\ell_4m_3m_4}t^{\ell_1\ell_2}_{\ell_3\ell_4}(L) + 23 \text{ perms}, \tag{3.29}$$

where the sum covers $4! = 24$ permutations of the four multipoles. Here, the azimuthal angular momentum $M = m_1 + m_2 = -m_3 - m_4$, and $t^{\ell_1\ell_2}_{\ell_3\ell_4}$ is the reduced trispectrum. The weighting function w generalizes the Gaunt function for the trispectrum and can be expressed as the average over three spin-weighted spherical harmonics ${}_sY_{\ell m}(\hat{n})$:

$$\begin{aligned} w^{LM}_{\ell_1\ell_2m_1m_2} &:= \sqrt{\frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4\pi}}\begin{pmatrix}\ell_1 & \ell_2 & L\\ m_1 & m_2 & M\end{pmatrix}\begin{pmatrix}\ell_1 & \ell_2 & L\\ -1 & -1 & 2\end{pmatrix} \\ &= \int_{\mathbb{S}^2}\mathrm{d}\hat{n}\,{}_{+1}Y_{\ell_1m_1}(\hat{n})\,{}_{+1}Y_{\ell_2m_2}(\hat{n})\,{}_{-2}Y_{LM}(\hat{n}). \end{aligned} \tag{3.30}$$

The reduced angular trispectrum, which embodies the full trispectrum's shape, behaves equivalently under parity transformation or complex conjugation:

$$\left[t^{\ell_1\ell_2}_{\ell_3\ell_4}(L)\right]^* = \mathbb{P}\left[t^{\ell_1\ell_2}_{\ell_3\ell_4}(L)\right] = (-1)^{\ell_1+\ell_2+\ell_3+\ell_4}t^{\ell_1\ell_2}_{\ell_3\ell_4}(L) \tag{3.31}$$

so the sum $\ell_1+\ell_2+\ell_3+\ell_4$ distinguishes between a purely real part (parity-odd) and a purely imaginary part (parity-even).

Thus, unlike the angular power spectrum and the angular bispectrum, the angular trispectrum is sensitive to parity transformation. This can be leveraged to test the ΛCDM model: since General Relativity and standard inflation introduce no parity-violating features, observing a non-zero signal from the parity-odd components would indicate parity symmetry violation. This will be tested by studying compatibility with a vanishing signal in the difference between parity-odd and parity-even components. If systematic uncertainties are properly modeled and accounted for, this signal would suggest a preference for left- or right-handed configurations, similar to the weak interaction.

Chapter 4

Angular redshift fluctuations

Modern cosmology is deeply focused on gathering diverse data to strengthen our understanding of the Universe and refine each individual data set's insights. Observations of CMB conducted by instruments such as *WMAP* [152, 86], *Planck* [9, 125], *SPT-3G* [23], *BICEP3* [7], *ACT* and *ACTPol* [10, 111] provide a detailed picture of the early Universe. Additionally, as CMB photons travel through the Universe, they interact with intervening matter, leading to secondary effects that offer insights into cosmic structures at lower redshifts (such as Sunyaev-Zel'dovich effect [25]). Meanwhile, to explore the later stages of the Universe, cosmology relies on the observation of luminous objects in the LSS. By studying features like the angular distribution of galaxies [49] and the abundance and properties of galaxy clusters and cosmic voids [14, 89], it has been possible to gather precise insights into the underlying density field. However, LSS surveys are significantly impacted by complexities such as how galaxies trace matter (bias), baryonic effects, and gravitational nonlinearities [176, 16, 154, 150, 140, 19, 139, 41, 36, 49], which add additional layers of observational systematics compared to CMB studies [32, 134]. Extensive efforts are underway to mitigate or model these uncertainties.

Ideally, findings from CMB and LSS data should be mutually consistent and complementary. However, discrepancies in key cosmological parameters, such as the one on the Hubble constant (see Sec.1.1.3) or the amplitude of linear matter fluctuations [3], are well known. This has increased the emphasis on refining error estimates, addressing systematic effects, and validating results through alternative cosmological probes [107]. This chapter focuses on density-weighted ARFs, an emerging observable developed to capture underlying density fluctuations across redshifts. This approach shows promise in refining cosmological constraints and serves as the foundation for the parity test conducted in this study.

4.1 Measuring angular redshift fluctuations

ARFs have emerged as a powerful probe for retrieving cosmological information from the LSS of the Universe. In particular, previous works have shown that ARFs are sensitive to the underlying galaxy velocity and density fields, while being relatively unaffected by systematic uncertainties and avoiding the need for assuming a fiducial cosmology in the distance-to-redshift relation [70]. Findings in [71] indicate that ARFs can constrain the quantity $f(z)\sigma_8(z)H(z)/H_0$, where $f := \mathrm{dlog}D_+(a)/\mathrm{dlog}a$ is the linear growth factor, with a level of agreement within 1.4σ with the fiducial ΛCDM cosmology from the *Planck satellite*.

Let us consider a spectroscopic galaxy survey, containing the redshifts z and the angular positions (RA, DEC) of galaxies in the sky. The celestial sphere is divided into N_{pix} regions, each of them pointed by a line of sight $\hat{n}(\theta, \phi)$. A certain number of galaxies will fall in each region, that from now on we identify as a pixel as commonly done when dealing to sky maps (see Sec.5.4). This number of galaxies will contribute to the value of the observable, depending on the generic window function W_j which weights each *j-th* galaxy.

The standard 2-dimensional clustering is computed from angular density fluctuations (ADFs), which describe the angular distribution of galaxies [71]:

$$1 + \delta_g(\hat{n}) = \frac{\sum_{j\in\hat{n}} W_j}{\left\langle \sum_{j\in\hat{n}} W_j \right\rangle_{\hat{n}}}, \tag{4.1}$$

where δ_g is the fluctuation of the angular number density of objects, and the angular average $\langle \dots \rangle_{\hat{n}}$ of the weights is computed over all the N_{pix} pixels included in the survey's footprint (it is an area average). The angular average of the number of galaxies under the window W, with the double sum running over the sky pixels and the galaxies falling in each pixel, is therefore:

$$\left\langle \sum_{j\in\hat{n}} W_j \right\rangle_{\hat{n}} := \frac{\sum_{i=1}^{N_{pix}} \sum_{j\in\hat{n}_i} W_j}{N_{pix}}, \tag{4.2}$$

where j runs over the galaxies in the i-th pixel.

The window function W can be any, as long as it is ends at the extremes. Howeven, for simplicity, we only refer to Gaussian window functions properly cut at a certain distance from their mean. In this work, we consider Gaussian shells

$$W_j = W_j(z_j; \sigma_z) := \exp\left\{ -\frac{(z_j - z_{cen})^2}{2\sigma_z^2} \right\}, \tag{4.3}$$

centered at an arbitrary reference redshift z_{cen} with width σ_z, which weights the galaxies at z_j. We note that, for tomographic aims, σ_z is typically smaller than 0.03 thus allowing us to assume constant bias within the Gaussian shell. For practical reasons, in the production of sky maps, we typically work in the interval $[-3\sigma_z, 3\sigma_z]$ (containing around 99.7% of the total area subtended by the Gaussian), thus using the W as a selection function.

While ADFs also convey cosmological information without relying on a distance-to-redshift relation, they are more prone to systematic uncertainties and nonlinear effects compared to ARFs. In ARFs, redshift is treated as a scalar field, allowing us to quantify the anisotropy around the average redshift (monopole) of all galaxies under the same Gaussian shell (4.3), regardless their angular location:

$$\bar{z} = \frac{\sum_j W_j z_j}{\sum_j W_j}, \tag{4.4}$$

which is typically close to the central redshift z_{cen}, and equal to it in case of flat distribution of galaxies. For tomographic analyses, z_{cen} will be very near to $\bar{z}$.

We consider two ARF estimators built upon Eq.(4.1):

$$\bar{z} + \delta_z(\hat{n}) = \frac{\sum_{j\in\hat{n}} W_j z_j}{\sum_{j\in\hat{n}} W_j}, \tag{4.5}$$

as proposed in [70], and

$$\delta_z(\hat{n}) = \frac{\sum_{j\in\hat{n}} W_j (z_j - \bar{z})}{\left\langle \sum_{j\in\hat{n}} W_j \right\rangle_{\hat{n}}}, \tag{4.6}$$

from [71]. Since, in Eq.(4.5), the denominator can be noisy or even vanish; thus, this work focuses on the latter implementation, as the ensamble average in the denominator helps addressing issues with sparse samples.

This clarifies what was mentioned at the beginning of Sec.3.2: despite working with LSS data, we do not conduct a 3-dimensional study, as ARF implementation depends only on the line of sight $\hat{n}$, with radial information embedded in the redshift itself. Consequently, we work on a 2-sphere, applying the "CMB technology" presented in the previous chapter.

As highlighted in [97], both ARF implementations yield the same results in linear cosmological perturbation theory. However, their sensitivity to systematics differs: Eq.(4.5) is robust against both multiplicative and additive systematics, while Eq.(4.6) is only resilient to additive systematics, as these do not vary significantly under the redshift shell. Additionally, it is important to note that the Gaussian window function used in the ARF implementation is chosen for simplicity; ARFs may be defined with any window function, provided it is restricted to a redshift interval.

4.1.1 Newtonian derivation

In the limit of weak gravitational fields and non-relativistic velocities, the observed redshift of a galaxy does not include the gravitational redshift $z_\phi(r,\hat{n})$, and all other relativistic contributions, and can thus be approximated as follows (c=1):

$$z_{obs} = z + (1+z)\vec{v}\cdot\hat{n} = z + z_v, \tag{4.7}$$

where $\vec{v}$ is the peculiar velocity of the galaxy, inducing a redshift or blueshift z_v.

Given the number density of galaxies $n(z,\hat{n})$ at redshift z from an observer pointing toward the $\hat{n}$ direction, the average number density of galaxies is defined as $\bar{n}_g(z) = \langle n_g(z,\hat{n})\rangle_{\hat{n}}$. Assuming a linear galaxy bias, $\delta_g^{3D} = b_g\delta_m^{3D}$, where δ_m^{3D} is the total matter density contrast defined in (2.77), we can write the ADF field as an integral along the line of sight [93][1]:

$$\delta_g(\hat{n}) = \frac{1}{N_g}\int dV_\Omega \bar{n}_g(z) b_g(z)\delta_m^{3D}(z,\hat{n})W(z_{obs}), \tag{4.8}$$

where

$$\mathrm{d}V_\Omega = \frac{\mathrm{d}V}{\mathrm{d}\Omega} = \frac{r^2(z)}{H(z)}\mathrm{d}z \tag{4.9}$$

describes the comoving volume of a thin spherical shell (with thickness dz) subtended by a solid angle dΩ. Here,

$$N_g = \int \mathrm{d}V_\Omega \bar{n}_g(z)W(z) \tag{4.10}$$

is the total number of tracers (galaxies) within the survey's footprint, selected by a Gaussian shell centered at z_{cen}.

The ARF field, which characterizes the spatial variation of the average redshift of galaxies, can be modeled in a similar way. It is expressed as an integral along the line of sight, restricted to the galaxies selected by the Gaussian shell:

$$\delta_z(\hat{n}) = \frac{1}{N_g}\int \mathrm{d}V_\Omega [z_{obs}(z,\hat{n}) - \bar{z}]\bar{n}_g(z)[1 + b_g(z)\delta_m^{3D}(z,\hat{n})]W[z_{obs}(z,\hat{n})]. \tag{4.11}$$

By expanding the Gaussian window and keeping only linear terms in density and velocity, [93] derived the ARF field as:

$$\begin{aligned}\delta_z(\hat{n}) = \frac{1}{N_g}\int \mathrm{d}V_\Omega \bar{n}_g(z)W(z)\Bigg\{&(z-\bar{z})b_g\delta_m^{3D}(z,\hat{n})\\ &+ z_v\left[1 + (z-\bar{z})\frac{\mathrm{d}\ln W}{\mathrm{d}z}\right]\Bigg\},\end{aligned} \tag{4.12}$$

---

[1] Hereafter, we drop the j subscript in indicating the window function

where the first term inside the curly brackets represents the dipole contribution, and the second term (quadrupole) accounts for peculiar velocity contributions.

Looking at the dipole term in Eq.(4.12), we see that the ARF field is sensitive to variations in the density field within the Gaussian shell, representing a redshift gradient. Similarly, the quadrupole term indicates that ARFs are sensitive to the radial variation of line-of-sight velocities.

4.2 Modelling angular redshift fluctuations with `ARFCAMB`

As mentioned in Ch.2, `CAMB` is a Fortran/Python code that solves the set of Einstein-Boltzmann equations Eqs.(2.21)-(2.26) and Eqs.(2.41)-(2.42), deriving the transfer function which is used to compute the time evolution of the linear matter power spectrum. C. H. Monteagudo provided me with a modified version of this code (dubbed as `ARFCAMB`), adapted to include ARFs as a new observable, as described in [97]. This includes also relativistic corrections for ARFs, which closely follow the ones of ADFs in [31].

Recovering the metric in the conformal-Newtonian gauge (Eq.2.13), given solely scalar perturbations, we can express it in terms of conformal time η:

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = a^2(\eta)[(1+2\psi)\mathrm{d}\eta - \delta_{ij}(1-2\phi)\mathrm{d}x^i\mathrm{d}x^j]. \tag{4.13}$$

We aim to evaluate the quantity $n(\hat{n}, z)(z-\bar{z})\mathrm{d}z\mathrm{d}\Omega$, which is the observed number of sources in the direction $\hat{n}$ over the solid angle $\mathrm{d}\Omega$ with redshift $(z-\bar{z})$ in a range $\mathrm{d}z$. In this chapter, we will use the $\cdot$ symbol to indicate the derivatives with respect to the conformal time.

Consider an observer measuring the redshift of an emitting light source. Hereafter, quantities with the subscript *obs* are evaluated at the observer's position while quantities with the subscript s refer to the source's position. A photon moving along a geodesic $x^\mu(\lambda)$, where λ is a parametrization for the geodesic, has 4-wavevector defined as:

$$k^\mu := \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} = \left(\frac{\mathrm{d}\eta}{\mathrm{d}\lambda}, \frac{\mathrm{d}\vec{x}}{\mathrm{d}\lambda}\right), \tag{4.14}$$

which can be decomposed into a space component using the direction of propagation $\vec{e} := \hat{n}$ and into a time component using the comoving frequency $\nu a^{-1} := k^\mu u_\mu$ (with $u_\mu = a(1+\psi)\delta_{\mu 0}$ the zero-shear 4-velocity field of a comoving observer) in the observer's rest frame. We can leverage the null condition of photons $\mathrm{d}s^2 = 0$ to write

$$\frac{\mathrm{d}\eta}{\mathrm{d}\lambda} = a^{-2}\nu(1-\psi). \tag{4.15}$$

for the time component, and we can obtain the space component as:

$$\frac{\mathrm{d}\vec{x}}{\mathrm{d}\eta} = \frac{\mathrm{d}\vec{x}}{\mathrm{d}\lambda}\frac{\mathrm{d}\lambda}{\mathrm{d}\eta} = (1+\phi+\psi)\vec{e}, \tag{4.16}$$

keeping only linear terms in ϕ, ψ. Eq.(4.15) can be rewritten using the geodesic equation to obtain the equation for the evolution of the comoving frequency in [31]:

$$\frac{\mathrm{d}\nu}{\mathrm{d}\eta} = -\nu\frac{\mathrm{d}\psi}{\mathrm{d}\eta} + \nu\left(\frac{\partial\phi}{\partial\eta} + \frac{\partial\psi}{\partial\eta}\right). \tag{4.17}$$

Integrating this equation can be done to compute the following ratio of frequencies:

$$\frac{\nu}{\nu_{obs}} = 1 + \psi_{obs} - \psi + \int_{\eta_{obs}}^{\eta} \mathrm{d}\eta' \left(\frac{\partial\phi}{\partial\eta} + \frac{\partial\psi}{\partial\eta}\right), \tag{4.18}$$

that contains the *Sachs-Wolfe* and the integrated Sachs-Wolfe (ISW) contributions, accounting for the variation of the potentials along the photon's trajectory [135].

The observed redshift of the source, z_s, is determined by the ratio of the observed frequency and the emitted frequency from Eq.(1.18):

$$1 + z_s = \frac{(k^\mu u_\mu)_s}{(k^\mu u_\mu)_{obs}}. \tag{4.19}$$

Writing the observer's 4-velocity as $u^\mu_{obs} = u^\mu + v^\mu_{obs}$, where we have explicited the contribution of the peculiar velocity v, the observed redshift of a source reads:

$$1 + z_s = \frac{a_{obs}}{a_s}\frac{\nu_s}{\nu_{obs}}(1 + \hat{n}\cdot[\vec{v} - \vec{v}_{obs}]). \tag{4.20}$$

Using Eq.(4.18), we can express the redshift along the line of sight $\hat{n}$ at conformal time η as:

$$1 + z(\eta) = \frac{a_{obs}}{a(\eta)}\left\{1 + \psi_{obs} - \psi + \hat{n}\cdot[\vec{v} - \vec{v}_{obs}] + \int_{\eta_{obs}}^{\eta} \mathrm{d}\eta' \left(\frac{\partial\phi}{\partial\eta} + \frac{\partial\psi}{\partial\eta}\right)\right\}. \tag{4.21}$$

Thus, the observed redshifts are mapped into perturbed radial coordinates of the sources so that, by setting $\eta = \eta_s + \delta\eta$ and $1 + z_s = a(\eta_s)^{-1}$, the perturbation of the conformal time assigned to the source is:

$$\mathcal{H}(\eta_s)\delta\eta := \Delta z(\eta_*) = \psi_{obs} - \psi + \hat{n}\cdot[\vec{v} - \vec{v}_{obs}] + \int_{\eta_{obs}}^{\eta_s} \mathrm{d}\eta' \left(\frac{\partial\phi}{\partial\eta} + \frac{\partial\psi}{\partial\eta}\right) + \mathcal{H}_0\delta\eta_0, \tag{4.22}$$

where $\mathcal{H}(\eta)$ is the conformal Hubble parameter. Following [31], one may express also the perturbed radial position of a photons at z_s as:

$$r(\hat{n}, z_s) = r_s + \delta r = \eta_{obs} - \eta_s - \delta\eta - \int_{\eta_{obs}}^{\eta_s} \mathrm{d}\eta' \left(\frac{\partial\phi}{\partial\eta} + \frac{\partial\psi}{\partial\eta}\right). \tag{4.23}$$

Starting from these results, [31] computed the angular power spectrum for ADFs by performing the Fourier transform of the curvature power spectrum $P(k)$ multiplied by the squared transfer function for ADFs, $T_\ell^{ADFs,W}(k)$:

$$C_\ell^{ADFs} = \frac{2}{\pi}\int \mathrm{d}k\, k^2 P(k) |T_\ell^{ADFs,W}(k)|. \tag{4.24}$$

For a survey with limit magnitude $m < m_*$, let us define the total number of sources observed over a redshift interval $\mathrm{d}z$ under a solid angle Ω as $N(\hat{n}, z, m < m_*)$.

Referring to [31] for a complete derivation, the transfer function for ADFs reads ($r' := \eta_{obs} - \eta'$):

$$\begin{aligned}
T_{N,\ell}^{ADFs,W}(k) = \int_0^{\eta_{obs}} \mathrm{d}\eta \Bigg[& W(\eta)\left(\delta_N j_\ell(kr) + \frac{kv}{\mathcal{H}} j_\ell''(kr)\right) + W_{\delta\eta}(\eta)[\psi j_\ell(kr) + v j_\ell'(kr) \\
& + \left(\frac{\partial\phi}{\partial\eta} + \frac{\partial\psi}{\partial\eta}\right) j_\ell(kr) \int_0^\eta \mathrm{d}\eta' W_{\delta\eta}(\eta') + (\psi+\phi) j_\ell(kr) \Bigg(\int_0^\eta \mathrm{d}\eta' (2-5b_s)\frac{W(\eta')}{r'} \\
& + \frac{\ell(\ell+1)}{2}\int_0^\eta \mathrm{d}\eta' \frac{r'-r}{rr'}(2-5b_s) W(\eta') \Bigg) \\
& + W(\eta) j_\ell(kr)\left(\frac{1}{\mathcal{H}}\frac{\partial\phi}{\partial\eta} + \psi + (5b_s - 2)\phi\right)\Bigg].
\end{aligned} \tag{4.25}$$

Let us explain the meaning of each term. Integrals are computed along the line of sight over the window function $W(\eta) = W(z)(1+z)\mathcal{H}$. $j_\ell''(kr)$ and $j_\ell'(kr)$ indicate, respectively, the second and the first derivative of the spherical Bessel function $j_\ell(kr)$. The magnification bias - a bias in the magnitude of an object given by gravitational lensing effects - is:

$$b_s(z, m_*) := \frac{\partial \log_{10} \bar{N}(z, m < m_*)}{\partial m_*}, \tag{4.26}$$

where $\bar{N}(z, m < m_*)$ is the background physical number density of sources with magnitude outside the survey's limit, and δ_N is its fractional perturbation associated.

In Eq.(4.25), the term

$$W_{\delta\eta}(\eta) := \left[\frac{2-5b_s}{\mathcal{H}r} + 5b_s - \frac{\partial \ln[a^3 \bar{N}(m < m_*)]}{\mathcal{H}\partial\eta} + \frac{\mathcal{H}'}{\mathcal{H}^2}\right]_\eta W(\eta) \tag{4.27}$$

accounts for the source evolution in redshift. The redshift z is the observed redshift and it contains all the information about the scalar potential and the magnification bias, encoding the effect of lensing on the luminous objects. The resolution for the transfer function has been implemented by [31, 30] in a code called `CAMB sources` (based on `CAMB`), that computes source count angular power spectra.

Eq.(4.25) is the starting point to compute the transfer function for ARFs. In this case, the anisotropies are sourced by both mass tracers and redshift fluctuations with respect to $\bar{z}$. For the sake of simplicity, the window function is assumed to be Gaussian in order to rewrite Eq.(4.25) using

$$\mathcal{W}(z, \bar{z}) := W(z)(z - \bar{z}), \tag{4.28}$$

with $\bar{z} = \int \mathrm{d}z W(z) z$. This means that the integrals of Eq.(4.25) must be recomputed. The integration over the derivatives of the spherical Bessel function can be eliminated by integrating by parts. Since `CAMB sources` is written in the *CDM gauge* (zero acceleration frame), the transfer function must be translated to this frame. The detailed pipeline for the computation of the transfer function for ARFs can be found in Appendix B of the reference paper [97], and it involves the usage of the `sympy`-based `symbolic` module of the code `CAMB sources`.

An example of the evolution of the angular power spectrum at different redshifts computed with `ARFCAMB` can be found in Fig.4.1.

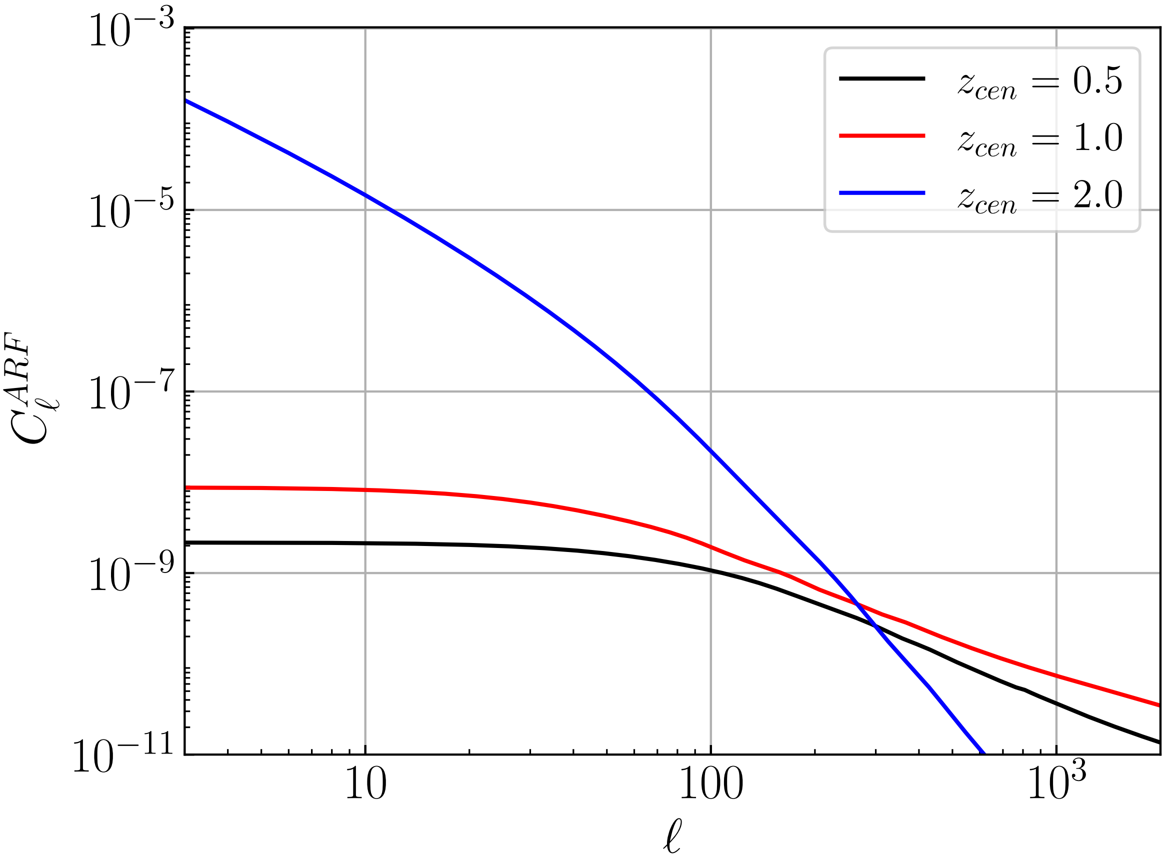


Figure 4.1: Linear angular auto-power spectrum for ARFs, computed using `ARFCAMB` at $\sigma_z = 0.01$ and for different reference redshifts z_{cen}. For comparison, all biases are set to unity. Notably, C_ℓ vanishes for multipoles $\ell < 2$.

Chapter 5

Codes and data sets

With the objectives of this work clearly defined and a focus on observing parity violation, we now present the data sets and computational tools used in the analysis. Each data set provided unique insights or testing grounds for parity violation, and each code facilitated different stages of data reprocessing and statistical analysis.

For real observational data, we used the Data Release 12 (DR12) of the Sloan Digital Sky Survey (SDSS) from the Baryon Oscillation Spectroscopic Survey (BOSS) catalogue, which offers an extensive galaxy redshift survey well-suited for parity analysis across the LSS. Alongside BOSS, we employed the MultiDark-`PATCHY` mocks [133, 80], which provide realistic simulated data designed to match the statistical properties of the BOSS catalogue, thereby enhancing the robustness of our analyses by offering a comparison data set with similar cosmic features.

We also incorporated two types of Quijote simulations [38, 170]: the Quijote-ODD simulations, which were specifically constructed with parity violation in their initial conditions, and standard Quijote simulations with ΛCDM cosmology that follow the conventional cosmological model without parity-breaking features. The Quijote-ODD simulations served as essential benchmarks, allowing us to test the sensitivity and accuracy of our detection methods on data with known parity-violating characteristics. In contrast, the standard ΛCDM Quijote simulations provided a baseline for comparison, enabling us to assess and interpret any asymmetries detected in the context of a standard cosmological model.

To process the data into a format suitable for analysis, we used a suite of computational tools. `CosmoBolognaLib` (hereafter `CBL`) [102] was particularly useful for the Quijote simulations, where it enabled the conversion of comoving halo coordinates into physical sky coordinates, necessary for constructing sky maps. This transformation bridged the gap between simulated catalogues and the formats required for our analysis, specifically when handling the spatial information in simulations.

For constructing pixelized maps of the sky, we used `HEALPix` [60, 61]. This software, a standard in cosmological data analysis, allows for high-resolution mapping of the celestial sphere by dividing it into equal-area pixels, thus providing an efficient structure for analyzing cosmic data. We built `HEALPix` maps tailored to ARFs through a Python/C++ implementation, allowing us to visualize and quantify the spatial distribution of LSS observables with high precision.

For calculating the angular power spectra and angular trispectra, we relied on `PolyBin` [122]. This tool is designed for unbiased estimation of higher-order statistics, which are fundamental in investigating parity properties across various scales. The spectra obtained with `PolyBin` yielded critical information on asymmetries within the spatial distribution of cosmic structures, contributing to a refined test for parity violation.

Together, these data sets and computational tools established a robust framework for building maps, generating mock data, and conducting detailed statistical analyses on parity-violating signals in the LSS.

5.1 Sloan Digital Sky Survey

SDSS is a large-scale photometric and spectroscopic redshift survey [173] conducted using the 2.5-meter wide-angle optical telescope at Apache Point Observatory (APO) [64], focusing on mapping the northern celestial hemisphere. This project has been foundational for numerous breakthroughs in observational astronomy and theoretical advancements, driven by its innovative instruments: a high-efficiency multi-array Charge-Coupled Device (CCD) camera for extensive sky imaging and a multi-fiber spectrograph capable of capturing spectra for several hundred objects simultaneously. Over the years, SDSS has evolved through various stages, with each iteration introducing new surveys and instruments. The stages, starting from SDSS-I (2000–2005) and continuing to SDSS-V (2020–present), reflect this development.

The third stage, SDSS-III [50, 11], ran from 2008 to 2014 and included four specialized spectroscopic surveys:

- MARVELS (Multi-object APO Radial Velocity Exoplanet Large-area Survey): This survey monitored the radial velocities of 11 000 bright stars with precision sufficient to detect gas giant planets.

- APOGEE (APO Galactic Evolution Experiment): This survey used high-resolution, high signal-to-noise infrared spectroscopy to study over 100 000 red giant stars throughout the Galactic bulge, bar, disk, and halo, penetrating the dust that obscures much of the Milky Way.

- SEGUE-2 (Sloan Extension for Galactic Understanding and Exploration): A continuation of SEGUE-1, this survey measured spectra for 240 000 stars, mapping the outer Milky Way to improve our understanding of the formation and growth of the stellar halo over time.

- BOSS: Designed to map the spatial distribution of luminous red galaxies (LRGs) and quasars, BOSS aimed to detect the characteristic scale imprinted by BAOs from the early Universe. Using a 1000-fiber spectrograph, it surveyed 1.5 million galaxies up to redshift 0.7 and 160 000 quasars at redshifts 2.2-3 and 1.5 million galaxies a redshift 0.7.

The BOSS DR12 survey [99], in particular, is used for examining parity violation in the cosmic distribution of galaxies of the CMASS and LOWZ catalogues.

5.1.1 BOSS Data Release 12

The survey includes observations on five colour bands (u, g, r, i, z) [59] of 1.5 million galaxies' spectra in 10 000 deg^2, using the doubled-armed spectrographs with wavelength coverage $3600 - 10000$ Å and resolving power of 1500 to 2600 [146]. Surveyed galaxies are divided into CMASS samples and LOWZ samples, each of them mapped in two separate sky regions: the Northern Galactic Cap (NGC) and Southern Galactic Cap (SGC).

The BOSS collaboration provides the galaxy catalogues, produced with an updated version of the code `mksample` [128], with specific colour-magnitude and colour-colour cuts to the photometric catalogue. The LOWZ sample comes from the original SDSS LRG catalogue specialized for low redshifts, while the CMASS catalogue contains higher redshift galaxies with almost constant stellar mass. In this work, following the pipeline of [99], we restricted the redshift domain of the two samples by applying the following cuts in order to simplify the tomographic analysis:

$$\begin{aligned} &\text{LOWZ: } 0.15 \leq z < 0.45 \\ &\text{CMASS: } 0.45 \leq z < 0.80. \end{aligned} \tag{5.1}$$

All the catalogues and masks of BOSS DR12 are pubblicly available at [1]. In Fig.5.1, the redshift distribution of galaxies is represented. The cut at $z = 0.45$ delimits the two catalogues.

LOWZ sample

The LOWZ sample contains LRGs with redshifts up to 0.45, and it is an extension of the SDSS-II LRG Cut I sample from [48]. The photometric selection criteria for this sample are as follows:

$$|c_\perp| < 0.2 \tag{5.2}$$

$$r_{cmod} < 13.5 + c_{\|}/0.3 \tag{5.3}$$

$$r_{psf} - r_{cmod} > 0.3 \tag{5.4}$$

$$16 < r_{cmod} < 19.6, \tag{5.5}$$

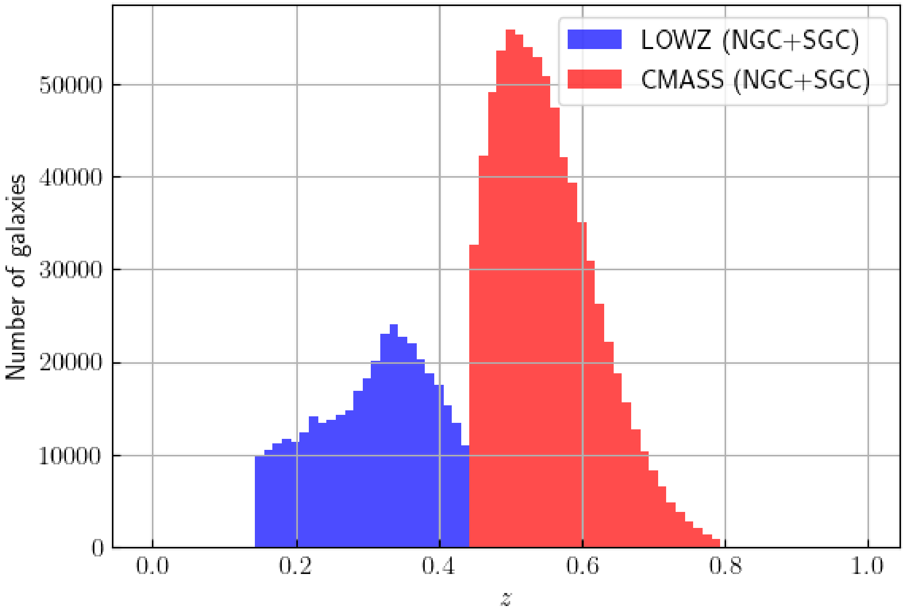


Figure 5.1: Redshift distribution of BOSS galaxies after applying the redshift cuts in Eq.(5.1), which prevent overlap between the CMASS and LOWZ catalogues. The cut survey contains a total of 1 158 301 objects.

where *cmod* stands for cmodel magnitude, representing the estimated total magnitude of a galaxy based on a composite model fit. The *psf* magnitude refers to the point spread function magnitude, which is optimized for the unresolved (point-like) component of an object. The parameters $c_{\parallel}$ and $c_{\perp}$ quantify a galaxy's position along the main colour locus and its deviation perpendicular to it, respectively, which are used to distinguish galaxies based on colour.

These criteria enable effective target selection by isolating objects near the expected colour locus, Eq.(5.2), while Eq.(5.3) selects bright, red objects. Eq.(5.4) introduces an r-band magnitude cut to help distinguish stars from LRGs, and Eq.(5.5) enforces a brightness limit in the r-band to ensure a high success rate in the galaxy detection.

It is important to note that the BOSS collaboration applied different selection criteria during the initial months of observation compared to those used later, which affected the redshift distribution in the LOWZE2 and LOWZE3 sky regions. Consequently, these two regions were excluded by the survey mask, as shown in the right panels of Fig.6.1, to maintain consistent redshift distributions across the entire sky footprint.

CMASS sample

The CMASS sample is targeted using the CMASS algorithm, which applies the following selection criteria, extending the Cut-II LRGs from SDSS-II [48]:

$$d_\perp > 0.55 \tag{5.6}$$

$$i_{mod} < \min\left[19.86 + 1.6(d_\perp - 0.8), 19.9\right] \tag{5.7}$$

$$17.5 < i_{cmod} < 19.9 \tag{5.8}$$

$$i_{fib2} < 21.5 \tag{5.9}$$

$$r_{mod} - i_{mod} > 0.2(21 - i_{mod}) \tag{5.10}$$

$$z_{psf} - z_{mod} > 0.46(19.8 - z_{mod}) \tag{5.11}$$

$$i_{psf} - i_{mod} > 0.2(21 - i_{mod}). \tag{5.12}$$

In these criteria, Eq.(5.6) excludes low-redshift objects, while Eq.(5.7) selects the brightest objects at each redshift to ensure approximately constant stellar mass limits, based on the evolving model from [101]. Eqs.(5.8)-(5.9) apply magnitude cuts to avoid misidentifying low-redshift objects and to ensure a high success rate. Problematic objects and outliers are further excluded by cuts in the i- and r-bands by Eq.(5.9). Finally, Eqs.(5.11)-(5.12) serve to separate stars from galaxies.

Galaxy weights

In galaxy surveys, galaxy weights are applied to correct for various biases that can affect clustering statistics derived from raw data. These weights help adjust for observational effects introduced by the survey's design and instrumental limitations, and they are provided by the BOSS collaboration as part of a composite weight parameter w_{tot} assigned to each galaxy [128, 134, 15]:

$$w_{tot} = w_{systot}(w_{cp} + w_{noz} - 1). \tag{5.13}$$

Each weight parameter accounts for a different observational effect:

- w_{systot}: This parameter adjusts for angular systematic effects based on the local conditions at each galaxy's position, including stellar density and seeing conditions (such as airmass and reddening).
- w_{cp}: Due to the finite size of fibers, which cannot overlap, objects separated by less than 62 arcseconds cannot both be observed on the same plate. The fiber collision weight w_{cp} compensates for this effect by assigning additional weight to galaxies that have collided with others within this threshold distance.

- w_{noz}: Redshift failures affect approximately 1.8% of CMASS galaxies and 0.4% of LOWZ galaxies. To mitigate this, w_{noz} up-weights the nearest neighbors of galaxies that experienced redshift measurement failures, ensuring they remain statistically representative.

The terms within the round brackets in Eq.(5.13) are constructed to conserve the total count of galactic targets. Although these observational systematics are designed for accuracy within the data processing pipeline, they have limited impact on sky map production with ARFs, as noted in Ch.4. In fact, observational systematics have a minimal effect on the ARFs power spectra, as highlighted in [71] and supported by findings in [70, 33].

5.1.2 MultiDark-`PATCHY` mocks

Mock catalogues, used alongside the real BOSS DR12 data, are essential for estimating uncertainties on the spectra and providing a benchmark for validating the data analysis pipeline. These mocks are generated using the BigMDPL simulation [81], one of the MultiDark N-body simulations created with the GADGET-2 code. GADGET-2 is a publicly available, massively parallel code that employs Tree-PM methods for efficient N-body simulations [153].

The BigMDPL simulation was produced assuming a flat Planck ΛCDM cosmology, with parameters summarized in Tab.5.1. It follows the evolution of 3840^3 dark matter particles, each with a mass of $2.356 \times 10^{10}\ h^{-1} M_{\odot}$, within a cosmological box with side length 2.5 h^{-1} Gpc, from an initial redshift of $z = 100$ to the present day.

Parameter	Symbol	Value
Hubble parameter	h	0.6777
Dark energy density parameter	Ω_Λ	0.692885
Matter density parameter	Ω_m	0.307115
Baryonic density parameter	Ω_b	0.048206
Normalization of the power spectrum	n_s	0.96
Amplitude of mass density fluctuation	σ_8	0.8228

Table 5.1: Cosmological parameters used for BigMDPL simulations, assuming a flat Planck ΛCDM cosmology.

The production of the mocks involves the use of the RockStar halo finder [20], which identifies spherical dark matter halos and subhalos using a friends-of-friends approach in 6-dimensional phase space. Additionally, RockStar constructs particle-merger trees [21] to calculate the peak circular velocity over a halo's history, which is useful for abundance matching.

Constructing light-cones that replicate the redshift-space monopole of the 2PCF of BOSS CMASS DR12 requires the `SUGAR` code. This code uses the halo abundance matching technique described in [133] to create galaxy catalogues from dark matter halo distributions.

The periodic mock boxes were created using the `PATCHY` code [80], which is based on the augmented Lagrangian perturbation theory and separates long- and short-range components of the dark matter particle displacement field. While the long-range components are computed with second-order Lagrangian perturbation theory [27, 29], the short-range components use spherical collapse approximation models [110]. This combined displacement field efficiently moves particles from their initial Lagrangian positions to final Eulerian positions. `PATCHY` also incorporates a local, nonlinear, scale-dependent, and stochastic biasing scheme to relate the discrete dark matter halo distribution with the underlying continuous dark matter particle field. In particular, this biasing scheme suppresses halo generation in low-density regions to model LRGs accurately [80].

Consequently, the resulting MultiDark-`PATCHY` mocks are designed to match observed clustering evolution across 1-point, 2-point, and 3-point clustering statistics and to account for dependencies on stellar mass. A publicly available set of 4096 CMASS mocks (NGC+SGC) and 4096 LOWZ mocks (NGC+SGC) can be accessed at [145], including detailed mock descriptions and weighting guidelines. Specifically, galaxies in the mocks are weighted as follows:

$$w_{tot} = w_{veto} w_{coll} w_{FKP}, \tag{5.14}$$

where w_{veto} is a binary flag that accepts or rejects a galaxy, and w_{coll} is the weight for fiber collisions. The term $w_{FKP} = (1 + n(z)P_0)^{-1}$ is the Feldman-Kaiser-Peacock weight [52], where $n(z)$ is the galaxy number density and $P_0 = 10^4 h^{-3}\mathrm{Mpc}^3$ represents the power spectrum at a reference scale around 100 h^{-1}Mpc. This weight is responsible for down-weighting overdense regions and up-weighting underdense regions to improve clustering statistics.

5.2 Quijote

The Quijote simulations consist of more than 82 000 full N-body simulations, created with two primary objectives in mind [170]. The first goal is to evaluate the information content on cosmological observables. For instance, one approach to quantifying the information contained in a set of cosmological parameters, given a particular statistic, is through the Fisher matrix formalism. This requires two key components: the covariance matrix of the statistic and the derivatives of the statistic with respect to the cosmological parameters. To enable this analysis, Quijote includes over 40.000 simulations specifically designed to numerically compute these components.

The second objective is to provide an extensive data set for training machine learning models. Quijote serves as an invaluable resource for machine learning applications. These simulations are generated using Latin hypercubes (which guarantee a uniform coverage of the parameter space) and Sobol sequences (which sample a multi-dimensional space evenly), and they cover a vast range of structures, including billions of halos, galaxies, and voids, as well as summary statistics such as power spectra, bispectra, correlation functions, marked power spectra and estimated probability density function. The sheer volume of data is crucial for training machine learning algorithms, where having a large and diverse data set significantly enhances the accuracy and robustness of the models.

The Quijote simulations are organized into different categories based on the cosmological models they implement. These can be broadly classified into three main *classes*:

- **Fiducial Simulations**: These simulations follow a fiducial cosmology that is consistent with the Planck data, with only the initial random seed varying.
- **Individual Parameter Variations**: These simulations alter the value of a single cosmological parameter with respect to the fiducial model, while keeping the initial random seed constant. They are designed for Fisher matrix analysis.
- **Multiple Parameter Variations**: These simulations simultaneously vary multiple cosmological parameters along with the initial random seed, intended for machine learning applications.

Furthermore, the Quijote simulations include a variety of *types*, each addressing different features of the Universe. These types include:

- **LCDM**: Standard simulations that vary cosmological parameters such as the Hubble constant, matter density, and cosmological constant.
- **Dark energy**: Simulations that investigate the impact of varying the expansion rate of the Universe through changes in the dark energy equation of state.
- **Massive neutrinos**: Simulations incorporating massive neutrinos among particles.
- **Separate Universe**: Simulations that involve an overall over- or under-density and modifications to the amplitude of the large-scale density fluctuations.
- **Primordial non-Gaussianities**: Simulations that explore different types of primordial non-Gaussianities.
- **Parity-violating**: Simulations that explore cosmological models with parity violation features.

- **Modified gravity**: Simulations that implement modified gravity models.

In this work, parity-violating simulations (Quijote-ODD [38]) are used alongside fiducial Quijote LCDM simulations (parameters in Tab.5.2) to test ARFs capability of detecting parity violation.

5.2.1 Quijote-ODD

Quijote-ODD is a type of parity-violating N-body simulation that propagates a parity violation from the initial conditions, generated at $z = 127$, forward in redshift. The pipeline for creating the initial conditions is detailed in [38] and involves generating a non-Gaussian primordial potential, ϕ^{NG}, by modulating a Gaussian one, $\phi^{(1)}$, according to the following transformation:

$$\phi^{NG}(\vec{x}) = \phi^{(1)}(\vec{x}) + p_{NL}\left[\epsilon^{ijk}(\partial_i|\partial|^\alpha\phi^{(1)})(\partial_j|\partial|^\beta\phi^{(1)})(\partial_k|\partial|^\gamma\phi^{(1)})\right](\vec{x}), \tag{5.15}$$

where ϵ^{ijk} is the 3-dimensional Levi-Civita symbol, an antisymmetric object often used to represent cross products. Here, $|\partial|^\alpha$ corresponds to a multiplication by k^α in Fourier space, with $\alpha \neq \beta \neq \gamma$. The parameter p_{NL} modulates the amplitude of the parity violation. Notably, in the right-hand side of Eq.(5.15), the parity violation is introduced by ϵ, as it changes sign under parity transformation (see Eq.(3.4)). In Fourier space, Eq.(5.15) can be rewritten as:

$$\begin{aligned}\phi^{NG}(\vec{k}) &= \phi^{(1)}(\vec{k}) \\ &+ ip_{NL}\int_{\vec{p}_1+\vec{p}_2+\vec{p}_3=\vec{k}}(2\pi)^3\delta_D(\vec{k}-\vec{p}_1-\vec{p}_2-\vec{p}_3)[\vec{p}_1\cdot\vec{p}_2\times\vec{p}_3]p_1^\alpha p_2^\beta p_3^\gamma\phi^{(1)}(\vec{p}_1)\phi^{(1)}(\vec{p}_2)\phi^{(1)}(\vec{p}_3),\end{aligned} \tag{5.16}$$

which confirms that the power spectra and bispectra cannot detect parity violations, as they lack terms involving the triple product and, consequently, pseudo-vectors.

Additionally, [38] rescaled the non-Gaussian, parity-violating initial conditions using

$$\phi^{NG}(\vec{k}) \to \sqrt{\frac{\left\langle\phi^{(1)}(\vec{k})\phi^{(1)}(\vec{k})\right\rangle}{\left\langle\phi^{NG}(\vec{k})\phi^{NG}(\vec{k})\right\rangle}}\phi^{NG}(\vec{k}), \tag{5.17}$$

to enforce a scale-invariant power spectrum and to avoid $\mathcal{O}(p_{NL}^2)$ corrections, which emerge at small scales in the power spectrum.

The initial conditions are generated by using a modified version of the `2LPTIC` code [141]. A total of 1000 initial conditions were generated: half with $p_{NL} = 10^6$ and the other half with $p_{NL} = -10^6$, ensuring averaged statistics as verified by [38]. The parity-violating initial conditions are propagated up to $z = 0$ using the transfer function from

`CAMB`, then rescaled back to $z = 127$ with a scale-independent growth factor. Finally, the density fields at $z = 127$ are corrected with second-order Lagrangian perturbation theory to compute the peculiar velocities and particle displacements.

Then, the particles evolve using the same pipeline as other QUIJOTE simulations, following the GADGET-3 code [153]. These simulations span a volume of 1000 $(h^{-1}\text{Mpc})^3$, hosting the evolution of 512^3 particles under the cosmological parameters detailed in Tab.5.2.

Simulation	$\mathbf{\Omega_m}$	$\mathbf{\Omega_\Lambda}$	$\mathbf{\Omega_b}$	σ_8	**h**	$\mathbf{n_s}$	$\mathbf{p_{NL}}$
ODD_P	0.3175	0.6825	0.049	0.834	0.6711	0.9624	$+10^6$
ODD_M	0.3175	0.6825	0.049	0.834	0.6711	0.9624	-10^6
LCDM	0.3175	0.6825	0.049	0.834	0.6711	0.9624	0

Table 5.2: Cosmological parameters used for QUIJOTE-ODD and fiducial QUIJOTE LCDM simulations. The simulations are dubbed ODD_P or ODD_M depending on the sign of the parity violation parameter p_{NL}.

Finally, ROCKSTAR was used to generate halo catalogues with a comprehensive set of properties, including comoving coordinates, virial masses, and peculiar velocities. These catalogues are stored on the Rusty cluster in New York and can be accessed via *Globus*, a platform designed for efficient transfer of large data volumes, following the instructions provided in [144]. For both the QUIJOTE-ODD and the fiducial QUIJOTE LCDM simulations, halo catalogues at redshift $z = 1$ were used to conduct parity tests.

5.3 CosmoBolognaLib

`CBL` is a comprehensive suite of C++ libraries for cosmological calculations [102], offering various functionalities including classes to handle catalogues, customize cosmologies, perform statistical analyses, and compute 2PCFs, 3PCFs, and power spectra using built-in wrappers for `CAMB` and `CLASS`, or accurate fitting formulae [47]. The backbone structure of the libraries relies on the object-oriented flexibility features of C++ that allow to organize efficiently functions and classes.

In particular, `CBL` provides a cbl::cosmology::Cosmology class designed to set various cosmological parameters, including the Hubble parameter h, the density parameters of cosmological species at $z = 0$ (Ω_m, Ω_b, Ω_ν, Ω_r and Ω_Λ), the primordial spectral index n_s, the initial scalar amplitude of the power spectrum A_s, the dark energy equation of state parameters w_o, w_a [34, 98], and the non-Gaussian amplitude f_{NL}. While each cosmological parameter can be set manually, `CBL` offers pre-built cosmologies from *WMAP5, WMAP7, WMAP9, Planck13, Planck15* and *Planck18* [85, 87, 73, 5, 6, 9]. This Cosmology class includes methods to estimate the number density and mass function of dark matter halos, as well as methods to estimate their effective linear bias by resolving the integral:

$$b_{\rm eff}(z) = \frac{\int_{M_{min}}^{M_{max}} n(M,z)b(M,z)\mathrm{d}M}{\int_{M_{min}}^{M_{max}} n(M,z)\mathrm{d}M}, \tag{5.18}$$

where n is the halo number density and b is the linear bias. This methodology has been widely employed to estimate the linear bias of dark matter halo catalogues from QUIJOTE.

The cbl::catalogue::Catalogue class is dedicated to handling samples of various astrophysical objects such as halos, galaxies, galaxy clusters, and voids. Each object has fields to record properties like positions (in both comoving and observed coordinates), masses, velocities, magnitudes, and other properties. The class also supports the creation of new customized catalogues, random catalogues, and subcatalogues, applying user-defined filters to mask objects selectively.

The cbl::measure::Measure class contains several subclasses for retrieving observables from data sets. In particular, these classes include methods to measure angular power spectra, 2PCF [91], and 3PCF [158] in both real and redshift spaces. The full inheritance diagram of the base class is shown in Fig.5.2.

In this work, a new class called cbl::measure::MapCalculator that inherits from the class Measure has been added to the project. This class computes full-sky maps using `HEALPix` according to the implementation of ARFs in Eq.(4.6). A detailed description of the class, along with the source code, is available in Appendix A.

5.3.1 Employed functionalities

In addition to the base classes and essential methods that every C++ program analyzing catalogues with `CBL` should implement, this work utilizes specific functions for different tasks. Firstly, `CBL` was employed to construct mock catalogues in physical coordinates (RA, DEC, z) from the halo catalogues in comoving coordinates produced by [38]. This process involves the function cbl::create_mocks(), which takes as input the positions and velocities, the box side length, the redshift of the halo catalogue snapshot, and other parameters to create a mock catalogue. The algorithm re-centers the simulation box by placing its center at a comoving distance that corresponds to the snapshot's redshift [104]. This approach allows for the creation of a light cone from a single snapshot at a specific redshift.

Additional functions were used to estimate the monopole of the 2PCF of the generated mocks in both real and redshift space, as well as to calculate the effective bias and model the 2PCF by Fourier transforming the power spectrum. These functions are part of the validation pipeline for the mocks, which will be detailed in the next chapter.

Finally, functions and classes within the cbl::statistics group were used to perform a *Markov Chain Monte Carlo* (MCMC) analysis, a widely adopted method in cosmology

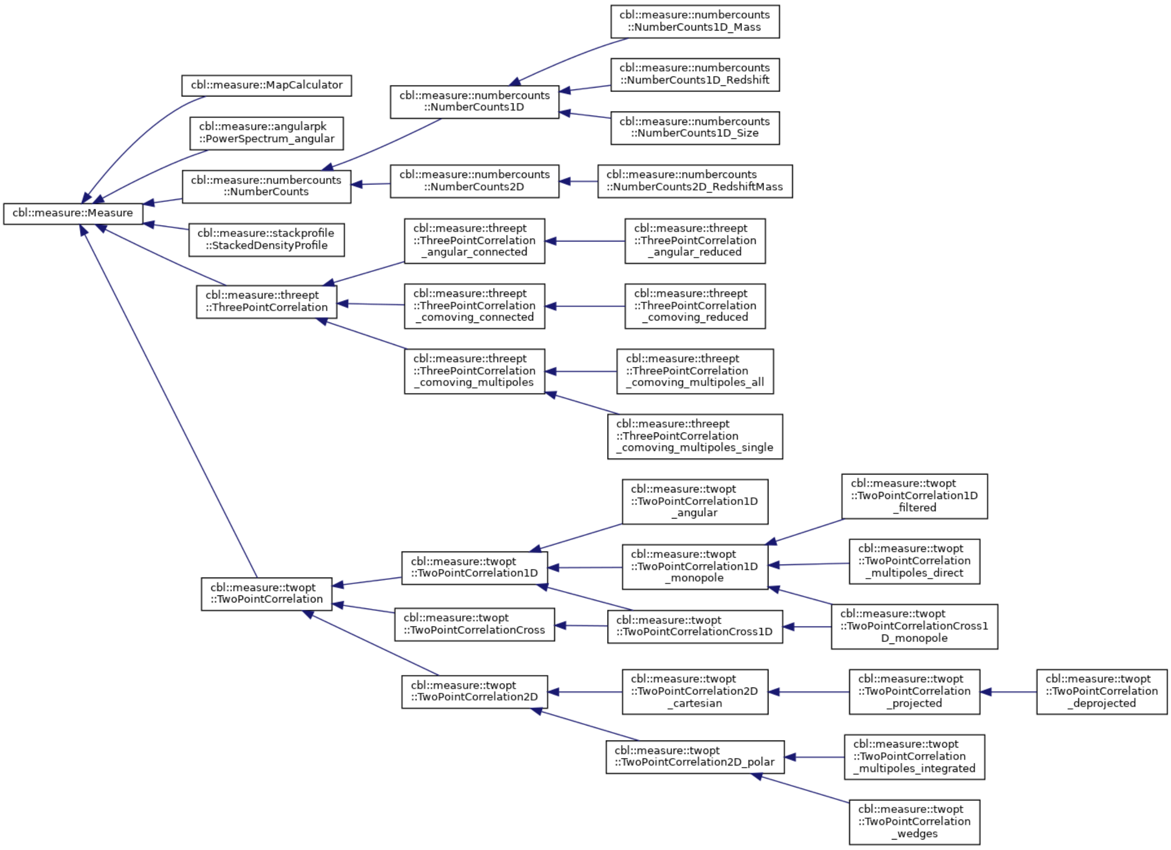


Figure 5.2: Inheritance diagram of cbl::measure::Measure. Among several subclasses, there is MapCalculator, a new addition to the project. Graph generated by DOXYGEN.

[75, 151]. MCMC is a sampling method for exploring parameter spaces to reconstruct posterior distributions. The algorithm produces statistically independent samples of the posterior arranged in a chain, where each sample depends only on the previous one, following the properties of Markov processes. After generating each sample λ, a conditional probability determines whether the next sample λ' is accepted or rejected. This probability satisfies the detailed balance condition:

$$P_{\lambda'} K_{\lambda|\lambda'} = P_{\lambda} K_{\lambda'|\lambda}, \tag{5.19}$$

which ensures that the chain is reversible, meaning the probability of moving forward or backward along the chain is equal. The conditional probability is chosen based on the *Metropolis-Hastings* algorithm [106, 68] and typically follows a Gaussian distribution. MCMC was employed to fit data from the parity violation test with a Gaussian function to detect potential signatures of parity violation. The complete validation pipeline will be illustrated in the next chapter.

5.4 `HEALPix`

`HEALPix` (Hierarchical Equal Area isoLatitude Pixelization) is a sophisticated framework used to map data on the sphere, originally designed to handle CMB data, and widely applied in astrophysics and cosmology. It implements an algorithm that divides a 2-sphere into pixels of equal area, using a pixelization scheme arranging pixels in a nested grid structure. In particular, the pixelation process consists by dividing the sphere's surface into curvilinear quadrilaterals in a hierarchical tessellation. At the lowest resolution, the sphere is partitioned into 12 base pixels. As the resolution increases, each pixel is recursively subdivided into four smaller ones making it possible to adjust pixel size and density as depicted in Fig.5.3, where the octahedral symmetry and the hierarchical subdivision of the grid are pointed out.

`HEALPix` is especially valuable for handling massive data sets since its structure allows data storage and retrieval to be computationally efficient. Indeed, pixels are distributed along iso-latitude lines allowing fast analyses in harmonic space. The complexity of `HEALPix` scales as $\mathcal{O}(N_{pix}^{1/2})$ while for other software that do not use iso-latitude pixelization, such as the Quadrilateralized Spherical Cube [161] (used for *COBE* data [149, 22, 56]) the complexity is $\mathcal{O}(N_{pix})$.

The software package is available in C, C++, Fortran90, IDL, Java, and Python, and offers a comprehensive set of tools for spherical data manipulation:

- `HEALPix` allows for filtering maps with arbitrary circular windows to mask specific regions and pixels of the sky.

- Includes an highly optimized library (`libsharp`) which is a collection of algorithms for efficient conversion between maps on the sphere and their spherical harmonic coefficients [129].

- `HEALPix` efficiently stores and organizes harmonic coefficients, while being also able to modify and rotate them which is useful when analyzing data from different reference frames. A typical example is the transformation of a map from one coordinate system to another (such as from equatorial to galactic coordinates).

- Enables pixelation down to sizes up to 0.4 milli-arcseconds, allowing for the creation of maps with up to 3.5×10^{18} pixels. Besides, it supports pixel queries in various geometric shapes, such as discs, triangles, polygons, and strips, and includes programs for identifying pixel neighbors.

- Support for multi order coverage maps used to represent different parts of the sky at varying levels of detail in a scalable way, particularly useful to describe localized crowded fields [55].

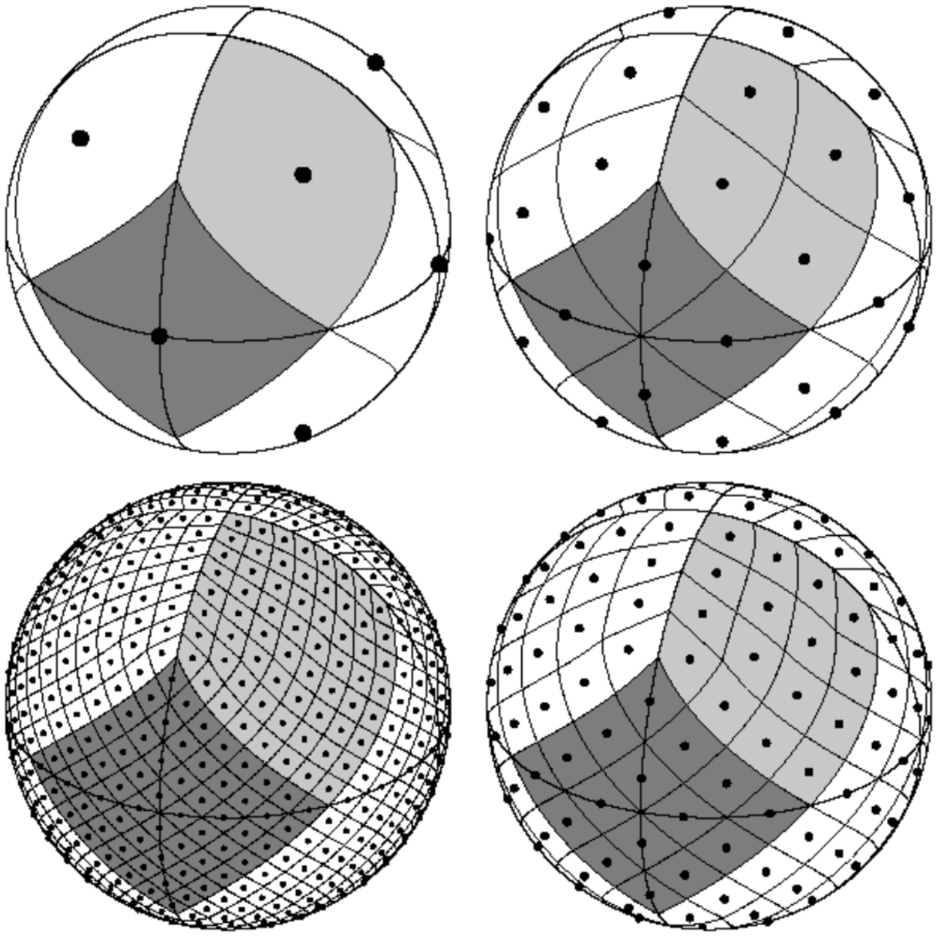


Figure 5.3: 2-sphere partition at progressively increase in resolution across four stages: starting with 12 pixels at the base level, followed by 48, 192, and 768 pixels. The black dots represent the center of the pixels occouring on 3, 7, 15 and 31 iso-latitude rings respectively. The resolution is refined in a clockwise manner from the top-left to bottom-left. Credits for the figure to [61].

- Visualization tools are available in multiple languages, and the software supports exporting maps to various formats, such as HDF5 files [13] and FITS files [174], which are commonly used to store scientific data. `HEALPix` provides also multiple routines to facilitate data handling in these formats.

5.4.1 Pixelization overview

The resolution of the `HEALPix` grid is controlled by the parameter N_{side}, which specifies the number of subdivisions along each side of the original base-resolution pixels to reach the desired finer resolution. Pixel centers are placed along $4N_{side} - 1$ rings of constant latitude, with the pixels on each ring spaced evenly in azimuth ϕ. Iso-latitude rings in the equatiorial region are divided into $4N_{side}$ pixels, while in the polar regions rings contain a varying number of pixels that increase by one between successive ring as the distance from the poles grows.

`HEALPix` offers two primary pixel indexing schemes:

- `RING`: Pixels are indexed sequentially from the north to the south pole, moving along each iso-latitude ring so that pixel indices follow a straightforward order across latitudes.
- `NESTED`: Pixels are indexed in a hierarchical tree structure rooted in the 12 base-resolution pixels. Each base pixel's hierarchy can be mapped to a $([0, 1] \times [0, 1])$ square, making nearest-neighbour searches more efficient and supporting fast Haar wavelet transforms [164] on `HEALPix` grids. In contrast, pixelization schemes like Gauss-Legendre Sky Pixelization [44] and icosahedron-based methods [160] lack hierarchical ordering, which is essential for handling large data sets efficiently by ensuring quick nearest-neighbor searches.

Pixel boundaries in `HEALPix` are not geodesics. Given (θ, ϕ) the coordinates on the sphere, in the equatorial zone pixel boundaries are given by $\cos\theta = a \pm b\phi$, while in the polar caps the forms are $\cos\theta = a + b/\phi^2$ or $\cos\theta = a + b/(\pi/2 - \phi)^2$. This geometric structure guarantees that all pixels have equal areas, enabling efficient computation of complex quantities, such as the Fourier transforms of individual pixels, which are commonly used to calculate power spectra. In contrast, other pixelization schemes, such as the Equidistant Cylindrical Projection [108], result in increasingly smaller pixel sizes as they approach the poles, leading to inefficient oversampling in those regions.

5.4.2 Main routines

One of the main `HEALPix` functions used for constructing sky maps is `ang2pix`. This function performs a conversion from the polar coordinates of a unit sphere to pixel indices, associating an integer between 0 and $12N_{\mathrm{side}}^2 - 1$ to each point on the 2-sphere.

This feature has proven particularly useful for indexing the positions of galaxies in the sky, as it tracks the full pixelization structure for the application of pixels weights and masks to exclude specific portions of the sky.

The other essential routine is `anafast`, a computationally efficient algorithm used to calculate both auto- and cross-angular power spectra C_ℓ of `HEALPix` maps, up to a user-defined maximum multipole $\ell_{\rm max}$ (with an upper bound of $3N_{\rm side} - 1$). This calculation is performed over the whole $\mathbb{S}^2$ sphere (or a subset if the map is cropped). The algorithm scales as $\mathcal{O}(N_{\rm pixel}^{1/2}\ell_{\rm max}^2)$, making it highly efficient for large data sets.

`anafast` offers several options to improve precision. It includes an iterative scheme that computes both forward and backward spherical harmonic transformations, ensuring high accuracy. Additionally, it incorporates pixel weights, which allow for the exclusion of specific pixels or adjust their importance based on the data's characteristics. In the context of this work, the pixel weights were designed to prioritize counts of galaxies near the mean of the Gaussian shells, as specified by the ARFs implementation in Ch.4.

Moreover, `anafast` includes a pixel window function option that adjusts the power distribution across pixels. This helps mitigate spectral leakage (the transfer of power between different frequency bands) and border effects (distortions in the power spectrum caused by pixel boundaries). These issues primarily arise at small angular scales ($\ell_{\rm max} > 2N_{\rm side}$), and therefore, this scale is used as the upper limit for the multipoles in the power spectrum analysis conducted in this work.

5.5 `PolyBin`

In Ch.3, we introduced the concepts of bispectrum, and trispectrum, fundamental tools for understanding the statistical properties of cosmological fields. These quantities are able to probe physical effects with unknown forms a priori, such as strong temperature bispectra from synchrotron emission and polarized bispectra from galactic dust emission [39], but they require significant computational resources. In this work we focus on the usage of the angular spectra described in Ch.3 alongside the implementation of ARFs described in Ch.4. However, there's an additional complication arising from observational effects: usually, it is not possible to measure a field in all points of the 2-sphere due to galactic pollution, too bright stars saturating the instruments, noisy regions and instrumental limitations. For this reason, data are provided with a mask W (aka window function) that works as a filter modulating the observed field. The mask is in general not isotropic, thereby complicating the interpretation of the angular correlators and leading to potentially dangerous effects when ignoring window effects. Therefore to compare measurements and models we need either to convolve the theory [162, 74] or to deconvolve the measurements in order to obtain something that can be readily compared to the theory.

`PolyBin` is a specialized code designed to provide estimators for various poly-spectra, including binned power spectra, bispectra, and trispectra [122]. The binned estimators implemented in `PolyBin` offer an advantage over modal decomposition as they separate dependence on individual models, enhancing the interpretability of the results. Efficient estimators for windowed polyspectra have been explored in various studies, such as [163, 100, 63] for the power spectrum; [137, 84, 137, 83] for the bispectrum; and [168, 127, 54, 105] for the trispectrum. However, `PolyBin` provides a robust treatment of observational masks by maximizing the likelihood of the observed field. Similar approaches have been used in the analysis of 3-dimensional field statistics [143, 39], though this technique was first introduced by [122] specifically for 2-sphere projections.

In this section, we present an overview of the polyspectra estimators implemented in `PolyBin`, following the methodology of the reference work [122]. These estimators play a significant role in this thesis, where they are used extensively for performing the parity tests.

5.5.1 Binning

PolyBin uses binned estimators for polyspectra by discretizing the correlators into ℓ-bins that, in the limit of narrow bins, are directly comparable to theory. To achieve this, [122] introduces a binary binning function $\Theta_\ell(b)$ that is defined to be unity if the multipole ℓ is contained in b and it vanish otherwise. The binning process provides a practical way to compute estimates of the power spectrum and bispectrum by grouping modes into discrete bins and summing their contributions. For the angular power spectrum Eq.(3.22), the binned estimator can be obtained by averaging over bins, yielding:

$$\langle a_{\ell_1 m_1} a_{\ell_2 m_2} \rangle \approx (-1)^{m_1} \delta^K_{\ell_1 \ell_2} \delta^K_{m_1(-m_2)} \sum_b \Theta_{\ell_1}(b) C(b), \tag{5.20}$$

where $C(b)$ is the binned power spectrum that we aim to estimate. The sum over bins accounts for the modes within each bin, and $\Theta_{\ell_1}(b)$ is an indicator function selecting which modes belong to each bin b.

Similarly, for the bispectrum Eq.(3.24), the binned version $b(\vec{b})$ can be expressed as:

$$\langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3} \rangle \approx \mathcal{G}^{\ell_1 \ell_2 \ell_3}_{m_1 m_2 m_3} \sum_{b_1 b_2 b_3} \frac{b(\vec{b})}{\Delta_3(\vec{b})} [\Theta_{\ell_1}(b_1)\Theta_{\ell_2}(b_2)\Theta_{\ell_3}(b_3) + 5 \text{ perms.}], \tag{5.21}$$

where $\vec{b} = \{b_1, b_2, b_3\}$ represents the triplet of bins for the bispectrum components, with the ordering condition $b_1 \leq b_2 \leq b_3$. Here, Δ_3 is a symmetry factor that adjusts for permutations within the set $\{\ell_1, \ell_2, \ell_3\}$, defined as:

$$\Delta_3(\mathbf{b}) := \begin{cases} 6 \; b_1 = b_2 = b_3 \\ 2 \; b_1 = b_2 \neq b_3 \text{ or } b_1 \neq b_2 = b_3 \\ 1 \text{ else,} \end{cases} \tag{5.22}$$

to ensure that each unique configuration is only counted once. This binning approach allows us to manage the complexity of bispectra estimation and avoid double-counting in the sums over permutations.

Finally, the binned trispectrum $t(\vec{b}, B)$ is defined analogously to Eq.(3.29):

$$\begin{aligned}\langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3} a_{\ell_4 m_4} \rangle_c \approx \sum_{LM} (-1)^M w^{L(-M)}_{\ell_1 \ell_2 m_1 m_2} w^{LM}_{\ell_3 \ell_4 m_3 m_4} \sum_{\vec{b}, B} \frac{t(\vec{b}, B)}{\Delta_4(\vec{b})} \Theta_L(B) \\ \times [\Theta_{\ell_1}(b_1)\Theta_{\ell_2}(b_2)\Theta_{\ell_3}(b_3)\Theta_{\ell_4}(b_4) + \text{ 7 perms.}] \\ + (2 \longleftrightarrow 3) + (2 \longleftrightarrow 4)\end{aligned} \tag{5.23}$$

Here, the sum runs over all permutations of $\{b_1, b_2, b_3, b_4\}$. $\vec{b} = \{b_1, b_2, b_3, b_4\}$ and B represents the bin corresponding to the effective angular momentum L. The sum includes only independent bins, so we require $b_1 \le b_2$, $b_3 \le b_4$, $b_1 \le b_3$, and $b_2 \ge b_4$ if $b_1 = b_3$, to avoid double counting. Additionally, the bins $\{b_1, b_2, B\}$ and $\{b_3, b_4, B\}$ must satisfy the triangle condition. To prevent double counting terms when summing over the permutations, we define a degeneracy factor $\Delta_4(\mathbf{b})$ to account for the occurrences of each unique configuration in Eq.(5.23):

$$\Delta_4(\mathbf{b}) := \begin{cases} 8 \; b_1 = b_2 = b_3 = b_4 \\ 6 \; b_1 = b_2 \text{ and } b_3 = b_4 \\ 2 \; b_1 = b_2 \text{ or } b_3 \neq b_4 \\ 2 \; b_1 = b_3 \text{ and } b_2 \neq b_4 \\ 1 \text{ else.} \end{cases} \tag{5.24}$$

5.5.2 Including the mask

As previously mentioned, observational limitations require applying a mask to exclude unwanted regions of the sky footprint. Moreover, the measured signal is affected by noise, $n(\hat{n})$, which adds to the observed signals, resulting in an observed field $\tilde{a}$ defined as:

$$\tilde{a}(\hat{n}) := W(\hat{n})a(\hat{n}) + n(\hat{n}). \tag{5.25}$$

We assume the simplest scenario, where the mask and the signal are uncorrelated, so that $\langle Wa \rangle = \langle W \rangle \langle a \rangle = 0$. If this assumption is violated, the complexity of the estimators will inevitably increase [94].

However, the presence of the mask and inhomogeneous noise may make the field statistics non-ideal, meaning we cannot assume the rotational symmetry of Eq.(3.22), Eq.(3.24), and Eq.(3.29) that led to the binned quantities previously described. Thus, we work in map space and define polyspectra in terms of observed fields. The non-ideal two-point harmonic correlator of the observed field can be written as:

$$\tilde{\mathsf{C}}^{ij} := \langle \tilde{a}(\hat{n}^i)\tilde{a}(\hat{n}^j) \rangle, \tag{5.26}$$

where the indices i and j indicate generic points on the sky. In this work, they specifically refer to `HEALPix` pixels. From Eq.(5.25), the 2-point correlator can be rewritten in terms of the unmasked quantities $\mathsf{C}^{ij} := \langle a(\hat{n}^i)a(\hat{n}^j)\rangle$ and $\mathsf{N}^{ij} := \langle a(\hat{n}^i)a(\hat{n}^j)\rangle$ as:

$$\tilde{\mathsf{C}}^{ij} = W(\hat{n}^i)C^{ij}W(\hat{n}^j) + \mathsf{N}^{ij} = \sum_{\ell m} B_\ell^2 C_\ell [W(\hat{n}^i)Y_{\ell m}(\hat{n}^i)][W(\hat{n}^j)Y^*_{\ell m}(\hat{n}^j)] + \mathsf{N}^{ij}. \quad (5.27)$$

This expression includes the isotropic beam B_ℓ and, if necessary, the pixel window function ω_ℓ through the substitution $B_\ell \to \omega_\ell B_\ell$, which ensures rotational invariance. The main advantage of the estimator given by Eq.(5.27) is that all terms in the expression can be estimated, and the estimation is simplified by assuming the same mask for both the signal and the noise (as is done in this work).

Analogous considerations lead to the definition of the non-ideal 3- and 4-point correlators:

$$\tilde{\mathsf{B}}^{ijk} := \left\langle \tilde{a}(\hat{n}^i)\tilde{a}(\hat{n}^j)\tilde{a}(\hat{n}^k)\right\rangle \ , \ \tilde{\mathsf{T}}^{ijkl} := \left\langle \tilde{a}(\hat{n}^i)\tilde{a}(\hat{n}^j)\tilde{a}(\hat{n}^k)\tilde{a}(\hat{n}^l)\right\rangle , \quad (5.28)$$

which can be expressed in terms of the ideal correlators B^{ijk}, T^{ijkl}.

5.5.3 Optimal estimators

To obtain optimal estimators from our binned polyspectra, we maximize the likelihood of the observed field, $\tilde{a}(\hat{n})$. Assuming weak non-Gaussianity, this likelihood can be expanded in terms of the non-ideal correlators using an *Edgeworth expansion*:

$$L[\tilde{a}] \propto \exp\left[-\frac{1}{2}h_i\tilde{\mathsf{C}}^{ij}h_j\right]\left\{1 + \frac{1}{3!}\tilde{\mathsf{B}}^{ijk}\mathcal{H}_{ijk} + \frac{1}{4!}\tilde{\mathsf{T}}^{ijkl}\mathcal{H}_{ijkl} + \dots\right\}, \quad (5.29)$$

where we assume an implicit summation over repeated indices. In this equation, the Wiener-filtered map is defined as

$$h(\hat{n}) := \left[\tilde{\mathsf{C}}^{-1}\tilde{a}\right](\hat{n}), \quad (5.30)$$

assuming $\tilde{\mathsf{C}}$ to be invertible. The *Hermite polynomials* are defined as:

$$\mathcal{H}_{ijk} := h_ih_jh_k - (h_i\tilde{\mathsf{C}}^{-1}_{jk} + 2 \text{ perms.}) \quad (5.31)$$

$$\mathcal{H}_{ijkl} := h_ih_jh_kh_l - (h_ih_j\tilde{\mathsf{C}}^{-1}_{kl} + 5 \text{ perms.}) + (\tilde{\mathsf{C}}^{-1}_{ij}\tilde{\mathsf{C}}^{-1}_{kl} + 2 \text{ perms.}). \quad (5.32)$$

In the Edgeworth expansion above, all cosmological information is contained in the map-space correlators, which are connected to the binned coefficient $C(b)$, $b(\vec{b})$, $t(\vec{b}, B)$ that we aim to estimate using Eqs.(3.22)-(3.29).

To see how this works, we introduce a generic binned quantity $\hat{x}(\vec{b})$ that arises within the N-point correlator $\tilde{\mathsf{X}}^{i_1\cdots i_N}$, with N$>$ 2. Maximizing the log-likelihood log$L[\tilde{a}](x)$ in the limit of small x leads to the optimal estimator for $x(\vec{b})$:

$$\hat{x}(\vec{b}) \propto \frac{1}{N!} \frac{\partial \tilde{\mathsf{X}}^{i_1 \dots i_N}}{\partial x(\vec{b})} \mathcal{H}_{i_1 \dots i_N}. \tag{5.33}$$

A normalization factor is necessary to remove biases from the estimator. It can be derived by requiring $\left\langle \hat{x}(\vec{b}) \right\rangle = x(\vec{b})$. The inverse of the normalization factor is referred to as the *Fisher matrix*, defined as:

$$\mathcal{F}_N(\vec{b}, \vec{b}') = \frac{1}{N!} \frac{\partial \tilde{\mathsf{X}}^{i_1 \dots i_N}}{\partial x(\vec{b})} \tilde{\mathsf{C}}^{-1}_{i_1 j_1} \dots \tilde{\mathsf{C}}^{-1}_{i_N j_N} \frac{\partial \tilde{\mathsf{X}}^{i_1 \dots i_N}}{\partial x(\vec{b}')}, \tag{5.34}$$

where all disconnected terms cancel when taking the expectation value of the Hermite tensor $\mathcal{H}_{i_1 \dots i_N}$.

This estimator is unbiased by construction, as ensured by the Fisher matrix, and it accounts for the response of the map to the underlying signal, as the partial derivatives are computed with respect to the unwindowed binned correlator $x(\vec{b})$. This estimation is optimal since we have maximized the likelihood, resulting in a covariance given by the inverse of the Fisher matrix. Additionally, this estimator is efficient, as both Eq.(5.33) and the Fisher matrix scale almost linearly with the total number of bins. A more general approach replaces the matrix $\tilde{\mathsf{C}}^{-1}$ with a weighting matrix S^{-1} in Eqs.(5.30)-(5.33) to ensure unbiased estimates at all times.

5.5.4 Power spectrum estimation

As in [122], the goal is to estimate the binned angular power spectrum using the formalism and concepts introduced thus far. We begin by taking the derivative of the log-likelihood with respect to $C(b)$:

$$\begin{aligned} \frac{\partial \log L[\tilde{a}]}{\partial C(b)} &= \frac{1}{2} \frac{\partial \tilde{C}_{ij}}{\partial C(b)} h_i h_j - \frac{1}{2} \mathrm{Tr} \left[\tilde{C}^{-1} \frac{\partial \tilde{C}}{\partial C(b)} \right], \\ \frac{\partial^2 \log L[\tilde{a}]}{\partial C(b) \partial C(b')} &= - \left[\frac{\partial \tilde{C}}{\partial C(b)} \tilde{C}^{-1} \frac{\partial \tilde{C}}{\partial C(b')} \right]^{ij} h_i h_j + \frac{1}{2} \mathrm{Tr} \left[\tilde{C}^{-1} \frac{\partial \tilde{C}}{\partial C(b)} \tilde{C}^{-1} \frac{\partial \tilde{C}}{\partial C(b')} \right]. \end{aligned} \tag{5.35}$$

The optimal estimator is then derived by expanding the likelihood up to second order around a fiducial power spectrum $\bar{C}(b)$ and maximizing it to solve for the true power spectrum $C(b)$:

$$\hat{C}(b) = \bar{C}(b) + \frac{1}{2} \sum_{b'} \mathcal{F}^{-1}_{2,\mathrm{opt}}(b, b') \frac{\partial \tilde{\mathsf{C}}^{ij}}{\partial C(b')} \left[h_i h_j - \tilde{\mathsf{C}}^{-1}_{ij} \right], \tag{5.36}$$

where we define the matrix $\mathcal{F}_{2,\mathrm{opt}}$ as:

$$\mathcal{F}_{2,\mathrm{opt}} := \frac{1}{2}\mathrm{Tr}\left[\tilde{C}^{-1}\frac{\partial \tilde{C}}{\partial C(b)}\tilde{C}^{-1}\frac{\partial \tilde{C}}{\partial C(b')}\right]. \tag{5.37}$$

The final estimator is obtained by substituting a generic weighting matrix S^{-1} for $\tilde{\mathsf{C}}^{-1}$ in both the Wiener-filtered map and in Eq.(5.36):

$$\begin{cases} \hat{C}(b) = \frac{1}{2}\sum_{b'}\mathcal{F}_2^{-1}(b,b')\left[\frac{\partial \tilde{\mathsf{C}}_{ij}}{\partial C(b')}h_i h_j - \mathrm{Tr}\left(\frac{\partial \tilde{\mathsf{C}}}{\partial C(b')}\mathsf{S}^{-1}\mathsf{N}\mathsf{S}^{-\mathrm{T}}\right)\right], \\ \mathcal{F}_2(b,b') = \frac{1}{2}\mathrm{Tr}\left[\mathsf{S}^{-\mathrm{T}}\frac{\partial \tilde{\mathsf{C}}}{\partial C(b)}\mathsf{S}^{-1}\frac{\partial \tilde{\mathsf{C}}}{\partial C(b')}\right]. \end{cases} \tag{5.38}$$

The trace term serves as a correction that accounts for the bias induced by the noise N in the estimator. This correction is valid under the assumption that the noise is independent of the true signal.

Ideal form

We now consider the power spectrum estimation in the idealized limit, where there is no masking, and the noise is assumed to be isotropic. By moving to harmonic space and applying Eq.(3.22), the derivative of the 2-point correlator becomes:

$$\frac{\partial \mathsf{C}^{ij}}{\partial C(b)} = \sum_{\ell m}\Theta_\ell(b)Y_{\ell m}(\hat{n}^i)Y^*_{\ell m}(\hat{n}^j). \tag{5.39}$$

Substituting this into Eq.(5.38) and rewriting the Wiener-filtered map as $h_{\ell m} = S_\ell^{-1}a_{\ell m}$ where S is assumed diagonal, we obtain the ideal estimator:

$$\begin{cases} \hat{C}_{\mathrm{ideal}}(b) = \frac{1}{2}\mathcal{F}^{-1}_{2,\mathrm{ideal}}(b)\sum_{\ell m}\Theta_\ell(b)\frac{|a_{\ell m}|^2}{S_\ell^2}, \\ \mathcal{F}_{2,\mathrm{ideal}}(b) = \frac{1}{2}\sum_\ell \Theta_\ell(b)\frac{2\ell+1}{S_\ell^2}. \end{cases} \tag{5.40}$$

Generic form

Similarly to the ideal case, but now incorporating the mask W, we can use Eq.(3.22) to express the derivative of the 2-point correlator as:

$$\frac{\partial \mathsf{C}^{ij}}{\partial C(b)} = \sum_{\ell m}B_\ell^2\Theta_\ell(b)[W(\hat{n}^i)Y_{\ell m}(\hat{n}^i)][W(\hat{n}^j)Y^*_{\ell m}(\hat{n}^j)]. \tag{5.41}$$

Using Eq.(5.38), we can write the numerator of the optimal power spectrum estimator as:

$$\hat{C}(b) \propto \frac{1}{2} \sum_{\ell m} B_\ell^2 \Theta_\ell(b) [Wh]_{\ell m} [Wh]^*_{\ell m}, \tag{5.42}$$

where $[Wh]_{\ell m}$ denotes the spherical harmonic expansion of the product $W(\hat{n})h(\hat{n}) := W(\hat{n})[\mathsf{S}^{-1}\tilde{a}](\hat{n})$.

The Fisher matrix in Eq.(5.38) is less straightforward to compute, as it involves $\mathcal{O}(N^2_{\text{bins}})$ operations. An efficient approach to estimate the Fisher matrix is to use the expectation value of a Gaussian random field (GRF) u , following methods in [147, 119]. This approach gives the Fisher matrix as:

$$\mathcal{F}_2(b, b') = \frac{1}{2} \left\langle \left(\frac{\partial \tilde{\mathsf{C}}}{\partial C(b)} \mathsf{S}^{-1} u \right)^{\mathrm{T}} \mathsf{S}^{-1} \left(\frac{\partial \tilde{\mathsf{C}}}{\partial C(b')} \mathsf{U}^{-1} u \right) \right\rangle_u \tag{5.43}$$

or, equivalently:

$$\mathcal{F}_2(b, b') = \frac{1}{2} \left\langle Q_2^{\mathrm{T}} [\mathsf{S}^{-1} u](b) \times W \mathsf{S}^{-1} W \times Q_2 [\mathsf{U}^{-1} u](b') \right\rangle_u . \tag{5.44}$$

Here, U is the invertible covariance matrix of u, and $Q_{2,\ell m}[x](b) = B_\ell^2 \Theta_\ell(b) [Wx]_{\ell m}$ is a filtered map whose harmonic space definition is derived from Eq.(5.38).

The optimal binned estimator for the angular power spectrum then reads:

$$\begin{cases} \hat{C}(b) = \dfrac{1}{2} \mathcal{F}_2 \sum_{\ell m} B_\ell^2 \Theta_\ell(b) [Wh]_{\ell m} [Wh]^*_{\ell m}, \\ \mathcal{F}_2(b, b') = \dfrac{1}{2} \left\langle Q_2^{\mathrm{T}} [\mathsf{S}^{-1} u](b) \times W \mathsf{S}^{-1} W \times Q_2 [\mathsf{U}^{-1} u](b') \right\rangle_u . \end{cases} \tag{5.45}$$

To perform power spectrum estimation, it is essential to provide the window function W, the weighting matrix S^{-1} (which reduces the weight of noisy regions), and the Fisher GRF covariance U. While the mask W is typically included with survey data, the choice of S^{-1} is crucial. Ideally, S^{-1} should be the inverse of the data covariance $\tilde{\mathsf{C}}^{-1}$; however, since $\tilde{\mathsf{C}}$ is generally not invertible, approximations are often necessary. A practical solution is to use a diagonal approximation in harmonic space, which preserves unbiasedness with only minor optimality losses. For the Fisher GRF matrix, a choice that ensures fast convergence is setting $\mathsf{U}^{-1} = \mathsf{S}^{-1}$, assuming both are in diagonal form.

5.5.5 Trispectrum estimation

We now turn to the estimation of the trispectrum, as the bispectrum was not used in the parity test for the reasons discussed in Ch.3. For those interested in bispectrum estimation, we refer to the detailed explanation in [122], which complements the overview of polyspectra estimators presented in this thesis.

In statistical and cosmological analyses, the nomenclature of zero-field, one-field, two-field, etc., terms denotes the number of fields or data realizations contributing to a particular statistical quantity. This classification helps distinguish between various contributions to correlators. For instance, the zero-field term represents the theoretical expectation value independent of any specific data realization, relying solely on the assumed statistical properties of the underlying distribution. The one-field term, instead, includes the contribution from a single realization of the data field adjusting the expected value based on the observed data. Higher-order terms, describe contributions from correlations across multiple fields and are crucial for capturing complex correlations. In the trispectrum estimation, we will make widely use of this nomenclature as the trispectrum will be subdivided into four-, two-, zero-field terms.

As discussed for the power spectrum, the trispectrum estimation is performed by maximizing the likelihood for a 4-point correlator accordingly to Eq.(5.33) and Eq.(5.34):

$$\begin{aligned}\hat{t}(\vec{b},B) = \frac{1}{4!}\sum_{\vec{b}'}\mathcal{F}_4^{-1}(\vec{b},B;\vec{b}',B')\frac{\partial\tilde{\mathsf{T}}_{ijkl}}{\partial t(\vec{b},B')}&[h_ih_jh_kh_l - (h_ih_j\langle h_kh_l\rangle + 5\text{ perms.})\\ &+ (\langle h_ih_j\rangle\langle h_kh_l\rangle + 2\text{ perms.})],\end{aligned} \tag{5.46}$$

$$\mathcal{F}_4(\vec{b},B;\vec{b}',B') = \frac{1}{4!}\frac{\partial\tilde{\mathsf{T}}_{ijkl}}{\partial t(\vec{b},B)}\mathsf{S}^{-1}_{im}\mathsf{S}^{-1}_{jn}\mathsf{S}^{-1}_{ko}\mathsf{S}^{-1}_{lp}\frac{\partial\tilde{T}_{mnop}}{\partial t(\vec{b}',B')},$$

which is fully optimal in the limit of $\mathsf{S}^{-1}\to\mathsf{C}^{-1}$, as in prior cases. Unlike the power spectrum (and similarly the bispectrum), the trispectrum contains distinct zero-, two-, and four-field terms, which introduce additional complexity to the estimation process. For clarity, we denote parity-even trispectra as $t_+(\vec{b},L)$ and to parity-odd trispectra as $t_-(\vec{b},L)$.

Ideal form

Let us begin with the four-field term. From Eq.(5.46), we can write the numerator of the ideal trispectrum in harmonic space. This term involves a sum over multiple configurations, incorporating both permutations and expectation values between different Wiener-filtered maps:

$$\begin{aligned}\hat{t}_\pm(\vec{b},B) \propto \frac{1}{4!}\sum_{\vec{b}'}\frac{\partial\tilde{\mathsf{T}}^{\ell_1\ell_2\ell_3\ell_4}_{(-m_1)(-m_2)(-m_3)(-m_4)}}{\partial t_\pm(\vec{b},B')}&[h_{\ell_1m_1}h_{\ell_2m_2}h_{\ell_2m_3}h_{\ell_4m_4}\\ &- (h_{\ell_1m_1}h_{\ell_2m_2}\langle h_{\ell_3m_3}h_{\ell_4m_4}\rangle + 5\text{ perms.}) + (\langle h_ih_j\rangle\langle h_kh_l\rangle + 2\text{ perms.})].\end{aligned} \tag{5.47}$$

This can be reformulated using the binned equation, Eq.(3.29). In this case, the four-field trispectrum takes a more compact form:

$$\begin{aligned}\hat{t}^{(4)}_{\pm,\mathrm{ideal}}(\vec{b},B) \propto \frac{1}{\Delta_4(\vec{b})}&\sum_{\ell_i m_i}(-1)^{\ell_{1234}}\sum_{LM}(-1)^M w^{L(-M)}_{\ell_1\ell_2 m_1 m_2} w^{LM}_{\ell_3\ell_4 m_3 m_4}\Theta_L(B)\\ \times &\left[\frac{1\pm(-1)^{\ell_{1234}}}{2}\right]\Theta_{\ell_1}(b_1)\ldots\Theta_{\ell_4}(b_4)\ h_{\ell_1 m_1}\ldots h_{\ell_4 m_4},\end{aligned} \tag{5.48}$$

where, since all 24 permutations are identical, and due to the symmetry of the four h fields, we restrict the permutations to be either even or odd, for $\ell_{1234} := \ell_1+\ell_2+\ell_3+\ell_4$. Expanding the sum over the four multipoles and introducing the definitions

$$\begin{aligned}A^{\mathrm{ideal}}_{b_1 b_2}(L,M) &= \sum_{\ell_1\ell_2 m_1 m_2} w^{LM}_{\ell_1\ell_2 m_1 m_2}\Theta_{\ell_1}(b_1)\Theta_{\ell_2}(b_2)h_{\ell_1 m_1}h_{\ell_2 m_2},\\ \bar{A}^{\mathrm{ideal}}_{b_1 b_2}(L,M) &= \sum_{\ell_1\ell_2 m_1 m_2} (-1)^{\ell_1+\ell_1+L} w^{LM}_{\ell_1\ell_2 m_1 m_2}\Theta_{\ell_1}(b_1)\Theta_{\ell_2}(b_2)h_{\ell_1 m_1}h_{\ell_2 m_2},\end{aligned} \tag{5.49}$$

which are symmetric under exchange of bins $b_1 \leftrightarrow b_2$, we can express the ideal four-field numerator in Eq.(5.48) as:

$$\begin{aligned}\hat{t}^{(4)}_{\pm,\mathrm{ideal}}(\vec{b},B) \propto \pm\frac{1}{2\Delta_4(\vec{b})}&\sum_{LM}(-1)^M\Theta_L(B)[A^{\mathrm{ideal}}_{b_1 b_2}(L,-M)A^{\mathrm{ideal}}_{b_3 b_4}(L,M)\\ &\pm\bar{A}^{\mathrm{ideal}}_{b_1 b_2}(L,-M)\bar{A}^{\mathrm{ideal}}_{b_3 b_4}(L,M)].\end{aligned} \tag{5.50}$$

As highlighted by [122], this reformulation of the estimator is significantly more computationally efficient, as the calculation, given a certain maximum scale $\ell_{\max}$, scales as $\mathcal{O}(\ell^6_{\max})$ rather than the $\mathcal{O}(\ell^{10}_{\max})$ scaling of Eq.(5.48).

We now turn to the two-field term, which can readily be obtained by noting that, assuming uniform weights S_ℓ, the expectation value between two Wiener-filtered maps can be written in terms of Kronecker deltas:

$$\langle h_{\ell m}h_{\ell' m'}\rangle = (-1)^m\delta^K_{\ell\ell'}\delta^K_{m(-m)}\frac{C_\ell}{S^2_\ell}. \tag{5.51}$$

Thus, the two-field estimator is given by:

$$\begin{aligned}\hat{t}^{(2)}_{\pm,\mathrm{ideal}}(\vec{b},B) \propto -\frac{1}{\Delta_4(\vec{b})}&\sum_{\ell_i m_i}(-1)^{\ell_{1234}}\sum_{LM}(-1)^M w^{L(-M)}_{\ell_1\ell_2 m_1 m_2} w^{LM}_{\ell_3\ell_4 m_3 m_4}\Theta_L(B)\\ \times&\left[\frac{1\pm(-1)^{\ell_{1234}}}{2}\right]\Theta_{\ell_1}(b_1)\ldots\Theta_{\ell_4}(b_4)\\ \times&\left[h_{\ell_1 m_1}h_{\ell_2 m_2}(-1)^{m_3}\delta^K_{\ell_3\ell_4}\delta^K_{m_3(-m_4)}\frac{C_{\ell_3}}{S^2_{\ell_3}}+5\text{ perms.}\right].\end{aligned} \tag{5.52}$$

Leveraging the properties of the Wigner-3j symbols, this equation can be further simplified by constraining ℓ_{1234}. Specifically, ℓ_{1234} must be even, which implies that any parity-odd two-field term vanishes. In this case, the parity-even two-field numerator takes the form:

$$\begin{aligned}\hat{t}^{(2)}_{+,\mathrm{ideal}}(\vec{b},B) \propto -\frac{1}{\Delta_4(\vec{b})}(\delta^K_{b_1b_4}\delta^K_{b_2b_3}+\delta^K_{b_1b_3}\delta^K_{b_2b_4})\sum_{\ell_1\ell_2L}\frac{(2\ell_1+1)(2L+1)}{4\pi}\\ \begin{pmatrix}\ell_1 & \ell_2 & L\\ -1 & -1 & 2\end{pmatrix}^2 \times (-1)^{\ell_1+\ell_2+L}\Theta_L(B)(\Theta_{\ell_1}(b_1)\Theta_{\ell_2}(b_2)\\ +\Theta_{\ell_2}(b_1)\Theta_{\ell_1}(b_2))\frac{C_{\ell_1}}{S^2_{\ell_1}}\sum_{m_2}|h_{\ell_2m_2}|^2.\end{aligned} \tag{5.53}$$

Finally, the zero-field term can be obtained similarly, noting that due to the $1\pm(-1)^{\ell_{1234}}$ term, the odd spectrum must vanish. For the same reason, the parity-even trispectrum has only two non-vanishing permutations, leading to:

$$\begin{aligned}\hat{t}^{(2)}_{+,\mathrm{ideal}}(\vec{b},B) \propto -\frac{1}{\Delta_4(\vec{b})}(\delta^K_{b_1b_4}\delta^K_{b_2b_3}+\delta^K_{b_1b_3}\delta^K_{b_2b_4})\sum_{\ell_1\ell_2L}\frac{(2\ell_1+1)(2L+1)}{4\pi}\\ \begin{pmatrix}\ell_1 & \ell_2 & L\\ -1 & -1 & 2\end{pmatrix}^2 \times (-1)^{\ell_1+\ell_2+L}\Theta_{\ell_1}(b_1)\Theta_{\ell_2}(b_2)\frac{C_{\ell_1}}{S^2_{\ell_1}}\frac{C_{\ell_2}}{S^2_{\ell_2}}.\end{aligned} \tag{5.54}$$

To conclude, by combining the results from Eq.(5.48), Eq.(5.53), and Eq.(5.54), we can express the ideal trispectrum estimator as follows:

$$\begin{aligned}\hat{t}_{+,\mathrm{ideal}}(\vec{b},B) &= \sum_{\vec{b}'B'}\mathcal{F}^{\mathrm{ideal},-1}_{4+}(\vec{b},B;\vec{b}',B')\left[\hat{t}^{(4)}_{+,\mathrm{ideal}}(\vec{b}',B')+\hat{t}^{(2)}_{+,\mathrm{ideal}}(\vec{b}',B')+\hat{t}^{(0)}_{+,\mathrm{ideal}}(\vec{b}',B')\right],\\ \hat{t}_{-,\mathrm{ideal}}(\vec{b},L) &= \sum_{\vec{b}'B'}\mathcal{F}^{\mathrm{ideal},-1}_{4-}(\vec{b},B;\vec{b}',B')\hat{t}^{(4)}_{-,\mathrm{ideal}}(\vec{b}',B'),\end{aligned} \tag{5.55}$$

where we refer to [122] for the detailed derivation of the Fisher matrices used for the parity-odd and parity-even components of the trispectrum.

Generic form

Let us study the case of binned trispectrum estimation in the presence of a masked field. In this case, the parity-odd components in the zero-, two-, and four-field terms generally do not vanish, as they do in the ideal form, due to the effect of the window function.

We begin with the four-field term. As with the ideal estimator, we can insert the binned definition, Eq.(3.29), into the optimal estimator, Eq.(5.46), to obtain:

$$
\begin{aligned}
\hat{t}^{(4)}_{\pm}(\vec{b},B) \propto \pm \frac{1}{\Delta_4(\vec{b})} \sum_{\ell_i m_i} \left[\frac{1 \pm (-1)^{\ell_{1234}}}{2}\right] \sum_{LM} (-1)^M w^{L(-M)}_{\ell_1\ell_2 m_1 m_2} w^{LM}_{\ell_3\ell_4 m_3 m_4} B_{\ell_1} B_{\ell_2} B_{\ell_3} B_{\ell_4} \\
\times \Theta_L(B)\Theta_{\ell_1}(b_1)\dots\Theta_{\ell_4}(b_4)[Wh]_{\ell_1 m_1}\dots[Wh]_{\ell_4 m_4},
\end{aligned}
\tag{5.56}
$$

noting that this differs from the ideal case, Eq.(5.48), by the replacement $Wh := W\mathsf{S}^{-1}\tilde{a}$. By introducing some mask-dependent definitions:

$$
\begin{aligned}
\bar{A}_{b_1b_2}[x,y](L,M) &= (-1)^L \int \mathrm{d}\hat{n}\, {}_{-2}Y_{LM}(\hat{n}) \bar{H}^{+}_{b_1}[x](\hat{n}) \bar{H}^{+}_{b_2}[y](\hat{n}), \\
A_{b_1b_2}[x,y](L,M) &= \int \mathrm{d}\hat{n}\, {}_{-2}Y_{LM}(\hat{n}) H^{+}_{b_1}[x](\hat{n}) H^{+}_{b_2}[y](\hat{n}), \\
H^{+}_{b}[x](\hat{n}) &= \sum_{\ell m} [Wx]_{\ell m} B_\ell \Theta_\ell(b)\, {}_{\pm 1}Y_{\ell m}(\hat{n}), \\
\bar{H}^{+}_{b}[x](\hat{n}) &= \sum_{\ell m} (-1)^\ell [Wx]_{\ell m} B_\ell \Theta_\ell(b)\, {}_{\pm 1}Y_{\ell m}(\hat{n}),
\end{aligned}
\tag{5.57}
$$

we can express the four-field terms in the same form as the idealized one:

$$
\begin{aligned}
\hat{t}^{(4)}_{\pm}(\vec{b},B) \propto \pm \frac{1}{2\Delta_4(\vec{b})} \sum_{LM} (-1)^M \Theta_L(B) [A_{b_1b_2}(L,-M) A_{b_3b_4}(L,M) \\
\pm \bar{A}_{b_1b_2}(L,-M) \bar{A}_{b_3b_4}(L,M)].
\end{aligned}
\tag{5.58}
$$

As previously noted, the parity-odd two-field term does not, in general, vanish due to the effect of the mask. In particular, it may occur that even ℓ_{1234} in the true map does not necessarily correspond to even ℓ_{1234} in the windowed map. Generally, the two-field term can be obtained from the four-field term by taking the expectation value of two distinct Wiener-filtered maps, $h_i h_j \to \langle h_i h_j \rangle$, which can be estimated from a set of simulations $\{\alpha\}$ with covariance $\tilde{\mathsf{C}}_\alpha$. Therefore, the two-field term becomes:

$$
\begin{aligned}
\hat{t}^{(2)}_{\pm}(\vec{b},B) \propto \mp \frac{1}{2\Delta_4(\vec{b})} \sum_{LM} (-1)^M \Theta_L(B) \{ A_{b_1b_2}[h,h](L,-M) \langle A_{b_1b_2}[\mathsf{S}^{-1}_\alpha, \mathsf{S}^{-1}_\alpha](L,M) \rangle_\alpha \\
\pm \bar{A}_{b_1b_2}[h,h](L,-M) \langle \bar{A}_{b_3b_4}[\mathsf{S}^{-1}_\alpha, \mathsf{S}^{-1}_\alpha](L,M) \rangle_\alpha \} + \text{ 5 perms.},
\end{aligned}
\tag{5.59}
$$

where the permutations refer to the α realizations of the mocks.

The zero-field term is computed similarly, by using two sets of independent Monte Carlo simulations $\{\alpha_1\}$, $\{\alpha_2\}$ with the same covariance:

$$\begin{aligned}\hat{t}^{(0)}_{\pm}(\vec{b}, B) \propto \pm \frac{1}{4\Delta_4(\vec{b})} \sum_{LM} (-1)^M\, \Theta_L(B) \{ &\langle A_{b_1 b_2} \left[\mathsf{S}^{-1}\alpha_1, \mathsf{S}^{-1}\alpha_2\right] (L, -M)\, A_{b_3 b_4} \\ &\times \left[\mathsf{S}^{-1}\alpha_1, \mathsf{S}^{-1}\alpha_2\right] (L, M) \rangle_{\alpha_1, \alpha_2} - \langle \bar{A}_{b_1 b_2} \left[\mathsf{S}^{-1}\alpha_1, \mathsf{S}^{-1}\alpha_2\right] (L, -M)\, \bar{A}_{b_3 b_4} \\ &\times \left[\mathsf{S}^{-1}\alpha_1, \mathsf{S}^{-1}\alpha_2\right] (L, M) \rangle_{\alpha_1, \alpha_2} + 5\,\text{perms.} \}.\end{aligned} \tag{5.60}$$

For further details on the Fisher matrix estimation for masked-field trispectra using Gaussian random field (GRF) simulations, we refer the reader to [122].

Denoting the two possible parity states with $\lambda, \lambda' \in \{\pm 1\}$, we can summarize the estimation of the masked binned angular trispectrum:

$$\begin{aligned}\hat{t}_{\lambda}(\vec{b}, B) = \sum_{\vec{b}' B' \lambda'} \mathcal{F}^{-1}_{4\lambda\lambda'}(\vec{b}, B; \vec{b}', B') \{ &\tau_{\lambda'}[h, h, h, h](\vec{b}', B') - 6 \langle \tau_{\lambda'}[h, h, \mathsf{S}^{-1}\alpha, \mathsf{S}^{-1}\alpha](\vec{b}', B') \rangle_{\alpha} \\ &+ 3 \langle \tau_{\lambda'}[\mathsf{S}^{-1}\alpha_1, \mathsf{S}^{-1}\alpha_1, \mathsf{S}^{-1}\alpha_2, \mathsf{S}^{-1}\alpha_2](\vec{b}', B') \rangle_{\alpha 1, \alpha 2} \},\end{aligned} \tag{5.61}$$

where the unnormalized trispectrum estimator (without Fisher matrix) can be written explicitly as:

$$\begin{aligned}\tau_{\pm}[\alpha, \beta, \gamma, \delta](\vec{b}, B) = \pm \frac{1}{48\, \Delta_4(\vec{b})} \sum_{LM} (-1)^M\, \Theta_L(B) \{ &A_{b_1 b_2}[\alpha, \beta](L, -M) A_{b_3 b_4}[\gamma, \delta](L, M) \\ &\pm \bar{A}_{b_1 b_2}[\alpha, \beta](L, -M) \bar{A}_{b_3 b_4}[\gamma, \delta](L, M) \} + 23\,\text{permutations}.\end{aligned} \tag{5.62}$$

The permutations run over the positions $\alpha, \beta, \gamma, \delta$, where α represents random fields that satisfy the condition $\langle \alpha^{\mathrm{T}} \alpha \rangle = \tilde{\mathsf{C}}^{-1}$. This expression allows the removal of the odd-even coupling by setting $\lambda = \lambda'$.

Chapter 6

Testing for parity violation

In this chapter, we present results from our parity violation tests conducted on LSS data using ARFs. As described in Ch.4, ARFs offer a robust method that minimizes systematic biases ensuring a reasonable signal-to-noise ratio (SNR). While the SNR is typically lower for LSS compared to CMB analyses, the reduced bias provided by ARFs enhances the reliability of our findings.

6.1 Test on BOSS DR12

The first step in this analysis involved constructing `HEALPix` masks for the BOSS DR12 catalogue. The BOSS collaboration provides masks that define the survey's geometry, for both the LOWZ and CMASS samples, in `mangle` format [155]. These masks were converted into a `HEALPix`-readable format (FITS files) using a Python script provided by C. H. Monteagudo, which we modified to produce masks with resolutions up to $N_{pix} = 256$. Additionally, BOSS provides veto masks to account for effects such as seeing and extinction, the invisibility of objects behind the plate centerpost, and the exclusion of bright objects and high-priority quasars (within 62 arcseconds). The results of the application of the the survey's geometry masks, both before and after applying the veto masks, are shown in Fig.6.1. These are consistent with previous studies such as [99].

Subsequently, we produced sky maps for ARFs. This was achieved using another Python code originally written by C. H. Monteagudo, that was suitably modified for handling BOSS catalogues (the C++ rewriting of the code can be found in App.A). The code processes galaxy catalogues, masks, and weights (as described in Sec.5.1.1) to produce masked sky maps for ADFs, ARFs, and shot noise.

For this analysis, we adopted $N_{side} = 64$, corresponding to a map resolution of $N_{pix} = 49152$. This resolution was selected to match the average density of the BOSS DR12 catalogue while ensuring that nonlinear effects on ARFs (dominant at multipoles ℓ larger than 100) do not compromise the linear modelling described in Ch.4.

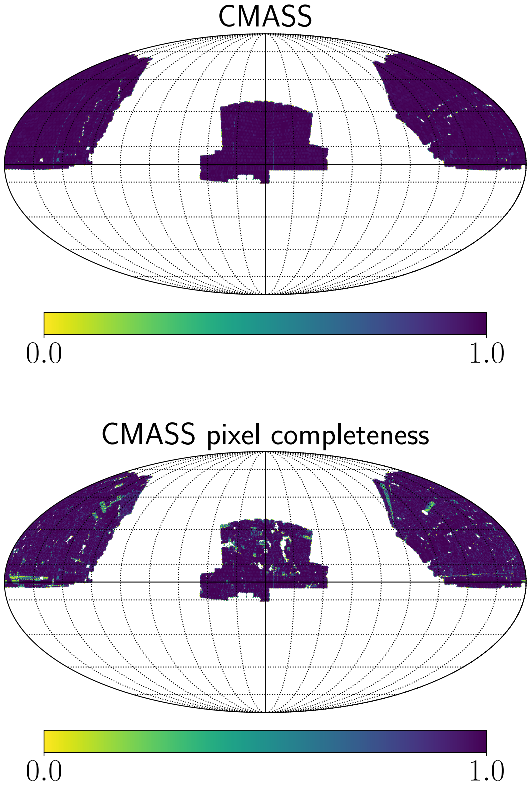


Figure 6.1: Survey geometry masks for BOSS DR12 generated at $N_{side} = 256$, shown in Mollweide projection. Top panels: masks for the CMASS (left) and LOWZ (right) samples before applying vetoes. Bottom panels: the same masks after applying vetoes. Note the gaps in the LOWZE2 and LOWZE3 regions.

We focused on two reference redshifts: $z_{cen} = 0.26$ and $z_{cen} = 0.5$, with a Gaussian window of standard deviation $\sigma_z = 0.01$. While $z_{cen} = 0.5$ corresponds approximately to the maximum in the BOSS galaxy redshift distribution and belongs to the CMASS sample, $z_{cen} = 0.26$ is part of the LOWZ sample, as shown in Fig.5.1. These two redshifts require different masks, as they belong to distinct galaxy samples.

From the MultiDark-`PATCHY` mocks, we generated 370 maps for $z_{cen} = 0.5$ and 100 maps for $z_{cen} = 0.26$. Examples of these maps for real galaxy catalogues from BOSS DR12 are displayed in Fig.6.2, highlighting the differing survey geometries.

Another essential step for modelling the trispectrum measures is the computation of the theoretical C_ℓ using `ARFCAMB`. This requires constructing a smooth redshift distribution of galaxies, $\mathrm{d}N/\mathrm{d}z$, for the BOSS catalogues. This distribution was generated by interpolating the histogram of galaxy counts from all 370 MultiDark-`PATCHY` mocks using quadratic interpolation. To ensure additional smoothness, a *Savitzky-Golay* filter

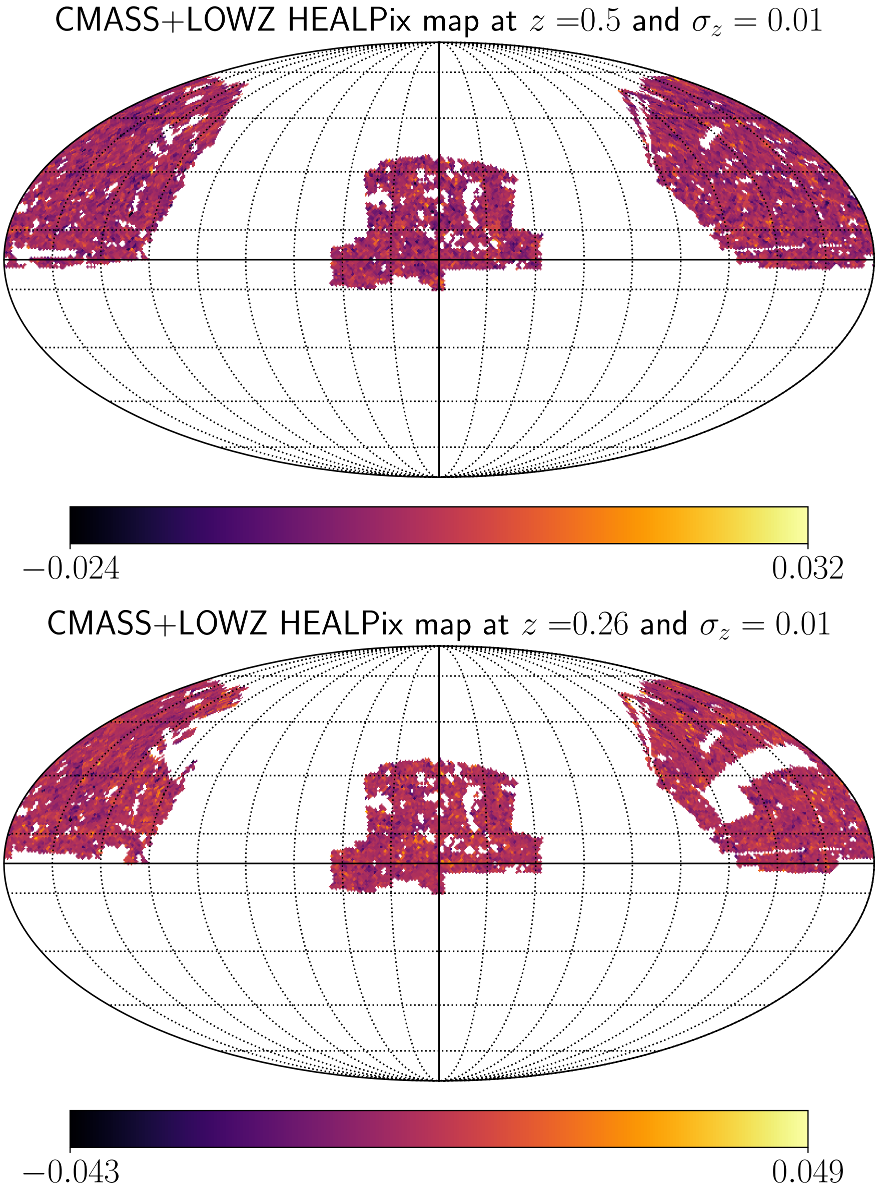


Figure 6.2: Mollwide projection of ARF maps computed at $N_{side} = 64$ for real galaxy catalogues of joint CMASS+LOWZ samples from BOSS DR12, combining the NGC and SGC. The colorbar represents the value of the redshift fluctuation δ_z with respect to the average redshift $\bar{z}$ of the objects in the sky map. Top: Gaussian shell with $z_{cen} = 0.5$ and $\sigma_z = 0.01$. Bottom: Gaussian shell with $z_{cen} = 0.26$.

was applied [138]. The resulting smooth distribution of BOSS DR12 galaxies in the redshift range $z = 0.15$ to $z = 0.75$, based on the mocks, is displayed in Fig.6.3.

Once this distribution is obtained, it is used to construct the Gaussian window function, as `ARFCAMB` requires the effective redshift distribution of the sources. This step accounts for the BOSS survey's selection function, which detects a specific number of LRGs per unit redshift. By imposing a Gaussian window on the BOSS redshift distribution, the effective distribution is given by $(\mathrm{d}N/\mathrm{d}z)\ \times W(z_{cen},\ \sigma_z)$. Although the slope of $\mathrm{d}N/\mathrm{d}z$ is typically irrelevant within the narrow Gaussian window used in tomographic analyses, it still introduces a slight dependency in the theoretical C_ℓ. The linear bias was set to $b = 2$ across all redshifts, consistent with findings in the literature for this catalogue [71]. The output of `ARFCAMB` provides the theoretical C_ℓ for auto-correlation (ARFxARF).

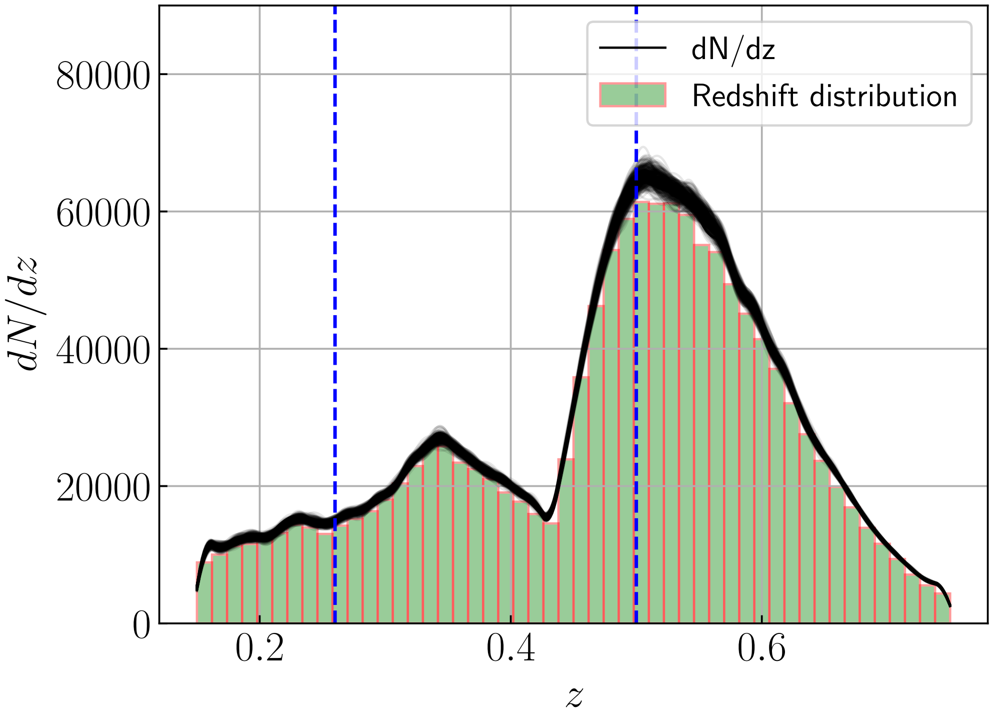


Figure 6.3: Smoothed redshift distribution of galaxies from 370 MultiDark-`PATCHY` mock catalogues. The redshift range extends from $z = 0.15$ to $z = 0.75$, covering both the LOWZ and CMASS samples of BOSS DR12, as shown in Fig.5.1. The green histogram represents the underlying redshift distribution of one mock. The vertical blue lines indicate the two reference redshifts, $z_{cen} = 0.26$ and $z_{cen} = 0.5$, employed in this work.

The final preliminary phase for the parity test is to measure the power spectra from the maps. The code employed for ARFs maps provides also shot noise estimation by filtering purely random mocks with the survey masks. Power spectra were computed using the `anafast` routine from `HEALPix` (see Sec.5.4) for both the ARFs maps and the shot noise maps. These power spectra were normalized with the sky fraction (approximately 21% for the CMASS mask and 23% for the LOWZ mask) to avoid underestimations. Finally, the theoretical C_ℓ curve was corrected for shot noise to obtain the fiducial observed power spectrum.

With this groundwork laid, the `PolyBin` base class was initialized to estimate the power spectra. Knowing N_{side} and having a fiducial shape for the power spectrum, the class was configured as excluding the pixel window and cross-correlations, focusing solely on the estimator for the auto-correlated angular power spectrum. The auto-correlation (ARFxARF) was renamed as (TxT) to align with `PolyBin`'s dictionary, originally designed for CMB temperature and polarization analyses [1].

The next step was initializing the `PolyBin` class for the power spectrum estimation. For this, a weighting choice S^{-1} was required, as discussed in the previous chapter. A quasi-optimal weight was obtained a method from the base class, which employs a diagonal weighting scheme based on the diagonal approximation of the true inverse covariance $\tilde{\mathsf{C}}^{-1}$ [2]. In this analysis, bins were restricted to the range $[0,\ 2N_{side}]$ with a step size of 4ℓ.

The ideal estimator was computed by invoking the corresponding routine and passing to it the *data* array, which contained the ARFs maps [3]. For the optimal estimator, the Fisher matrix was required. This matrix was produced averaging over $N\_it = 50$ realizations[4] and passing the result as input to the method for the unwindowed estimation [5]. The results of the power spectrum analysis are shown in Fig.6.4. The average estimation from `PolyBin` is consistent with the theoretical predictions from `ARFCAMB` across all bins. This confirms that the pipeline is functioning correctly, and all the necessary codes have been accurately implemented.

---

[1] `base = pb.PolyBin(ns, Sl_fiducial, include_pixel_window=False, pol=False, backend="healpix")`

[2] `pspec = pb.PSpec(base, mask, base.applyAinv, l_bins, fields=['TT'])`

[3] `Cl_opt = pspec.Cl_unwindowed(data, fish=fish)`

[4] `for i in range(N_it): fish += pspec.compute_fisher_contribution(i, verb=(i==0))/N_it`

[5] `Cl_ideal = pspec.Cl_ideal(data)`

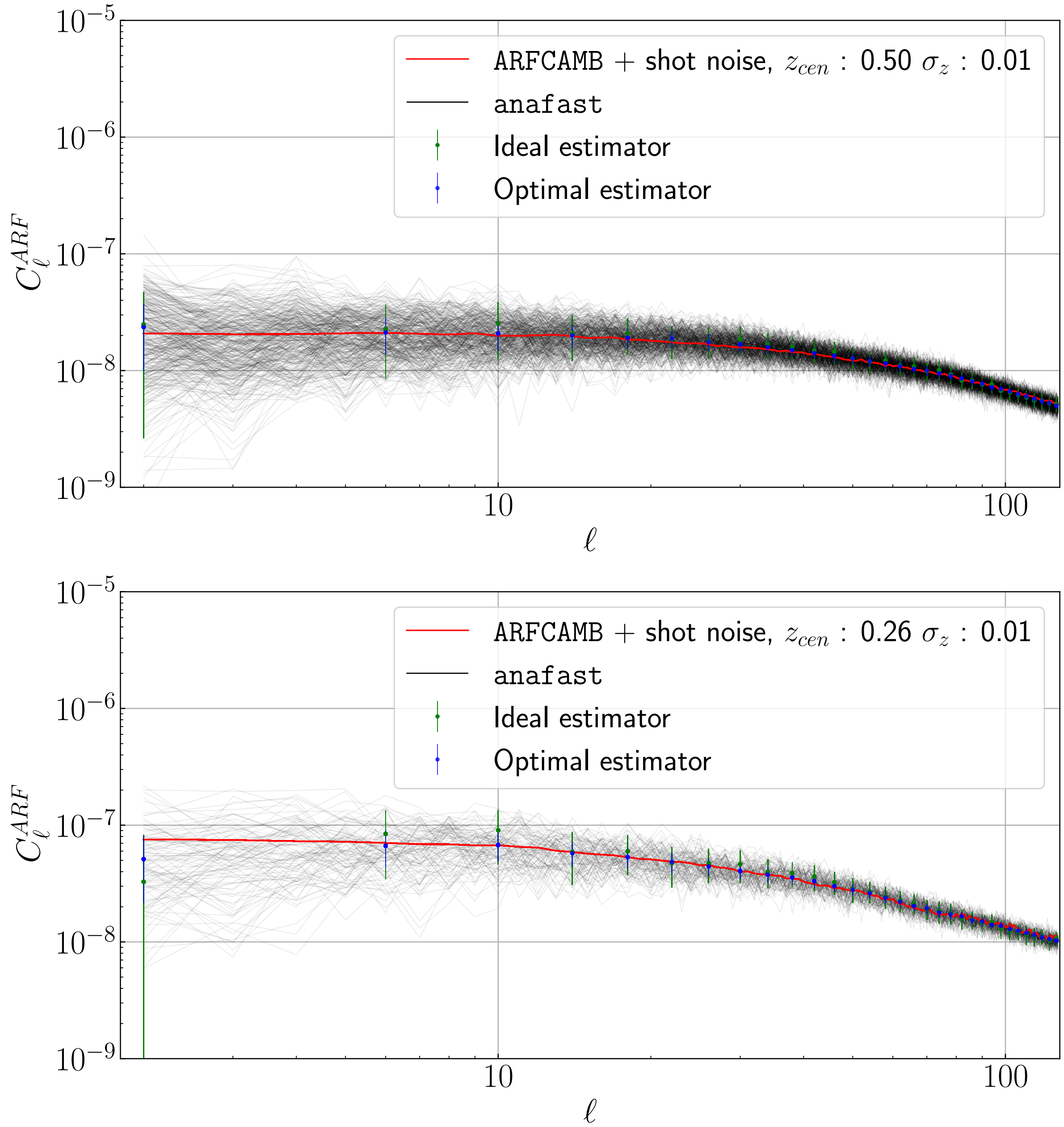


Figure 6.4: Angular power spectrum obtained from 370 MultiDark-`PATCHY` mock maps under a Gaussian shell with $z_{cen} = 0.5$ and $\sigma_z = 0.01$ (top panel). The gray curves represent the power spectrum C_ℓ estimated from the maps using `HEALPix`, while the blue and green dots show the mean values of the optimal and ideal estimations from `PolyBin` (with associated error bars). The red curve represents the average of the `ARFCAMB` output for each mock, including the shot noise contribution. Same for the bottom panel, but on 100 MultiDark-`PATCHY` mock maps, with Gaussian shells having $z_{cen} = 0.26$ and $\sigma_z = 0.01$.

6.1.1 Trispectrum analysis

With our pipeline validated, we now address the central aspect of the parity test by analyzing the trispectra. The goal is to measure, on average, a statistically significant difference between the odd and even components of trispectrum as indicator of parity violation in the BOSS DR12 galaxy catalogue. From this point onward, we focus exclusively on the optimal estimator.

We start with the estimation of the reduced trispectrum using `PolyBin`. To begin, a binning scheme for the estimators is defined. Since the trispectrum is a high-dimensional object with $\mathcal{O}(N_{bins}^6)$ elements, we adopt the following linear binning choice: $\ell_{min} = 10$, bin width $\Delta_\ell = 6$, $N_{bins} = 6$, and we drop the largest bin to avoid edge effects.

This binning choice is justified because at too large ℓ the angular trispectrum becomes almost flat [122]. This would reduce it to behave like a two-dimensional object, where a parity transformation becomes equivalent to a spatial rotation, making the parity test unreliable. Moreover, computations at large multipoles are prohibitive, as the complexity of the Fisher matrix computation scales as $\mathcal{O}(\ell_{max}^6)$.

This binning scheme keeps memory usage low (few gigabytes) and computational times manageable (around 3000 seconds wall-time)[6]. However, this also restricted our analysis to few hundreds of mocks rather than the full set of 2048 CMASS+LOWZ galaxy mocks due to the long execution times.

In this work, we present all the allowed trispectra within the bins $\{b_1, b_2, b_3, b_4, B\}$, which satisfy the triangle conditions at the bin centers and follow the ordering scheme described in Sec.5.5.1, condensed into a single dimension for visualization. The characteristic ℓ values in the bin gradually increase in size from the left to the right, following the bin index.

The initialization of the trispectrum class relies on `PolyBin`'s base class, on the previously defined mask, on the same weighting function as the power spectrum, on the specified binning scheme, on the (TxTxTxT) field referring to (ARFxARFxARFxARF) auto-correlation, and on a keyword to compute both parity-even and parity-odd components of the (reduced) angular trispectrum [7].

Finally, the trispectrum is estimated using 32 CPUs for fast computation. The normalization factor of Eq.(5.61) is estimated by averaging over $N\_it = 50$ Fisher matrix realizations. Parity-conserving terms (disconnected terms) are subtracted during the computation, focusing the estimation solely on parity-violating contributions[8].

We are now ready to investigate for parity violations. To this end, we model the SNR from the difference between the odd and even components of the optimal trispectrum estimators, using the set of mocks to estimate the noise associated to each bin, shown in

---

[6] Executions were performed on the supercomputer *Hyperion* [157] and the cluster *Atlas* [156]. Note that execution time is significantly influenced by the level of busyness of the nodes.

[7] `tspec = pb.TSpec(base, mask_used, base.applyAinv, l_bins, fields=['TTTT'], parity='both')`

[8] `t_unwindowed = tspec.Tl_unwindowed(data, verb=True, fish=fish, include_disconnected_term=True)`

Fig.6.5. The SNR, averaged over 370 mocks for $z_{cen} = 0.5$ and $z_{cen} = 0.26$, is shown as a function of the bin index in Fig.6.6.

As expected in absence of parity violation, the mean SNR values are distributed around a zero signal for each bin index. To reduce the impact of the scattering of the SNR across the ℓ multipoles, we average the SNR values over the bins for each mock. In this way, each mock produces a single value with an associated uncertainty. For a parity-symmetric process, these values are expected to be distributed around zero. To test this hypothesis, we construct a histogram of the mean SNR values from all mocks, along with their associated uncertainties on the frequency, which were estimated using a Monte Carlo method. Specifically, the uncertainties were estimated by generating random numbers that are Gaussian-distributed, with means and standard deviations corresponding to the scale-average SNR and its associated uncertainty of each mock.

To assess whether deviations from parity symmetry are statistically significant and to quantify them, we fit the histogram using a Gaussian model through an MCMC-based fitting process. We indicate as $\langle \mathrm{SNR} \rangle$ the best-fit mean value and with σ_{SNR} its standard deviation.

For $z_{cen} = 0.5$, the fitting process produced convergent chains and a successful fit, as depicted by the black curve in Fig.6.9. The contour plot in Fig.6.10 reveals $\langle \mathrm{SNR} \rangle = 0.003^{+0.004}_{-0.004}$ and a standard deviation $\sigma_{\mathrm{SNR}} = 0.043^{+0.004}_{-0.003}$. The mean value being compatible with zero aligns with expectations, as the mocks are parity-symmetric by construction. Using real data, the resulting SNR plot in Fig.6.7 has a scale-averaged value of 0.035 ± 0.89, which falls within $1-\sigma$ from the best-fit mean value. This outcome supports the conclusion that no parity violations are evident in the real data.

For $z_{cen} = 0.26$ we have larger noise than in the previous case as seen in Fig.6.5. This is mainly related to the less number density of objects entering in the ARFs statistics at this redshift, and partially due to lower number of mocks employed. This fact reflects negatively in the uncertainties on the frequencies of the SNR histogram in Fig.6.9, making them considerably large. As a consequence, the MCMC fit did not produce very constraining results due to the large dispersion, as shown by the large uncertainties in Fig.6.11. The fit illustrated in 6.9 returned $\langle \mathrm{SNR} \rangle = 0.01^{+0.10}_{-0.11}$ and $\sigma_{\mathrm{SNR}} = 0.3^{+0.6}_{-0.2}$. From the plot of the SNR of real data in Fig.6.8, we obtained a scale-averaged SNR of -0.013 ± 0.92 which is compatible with the best-fit mean value.

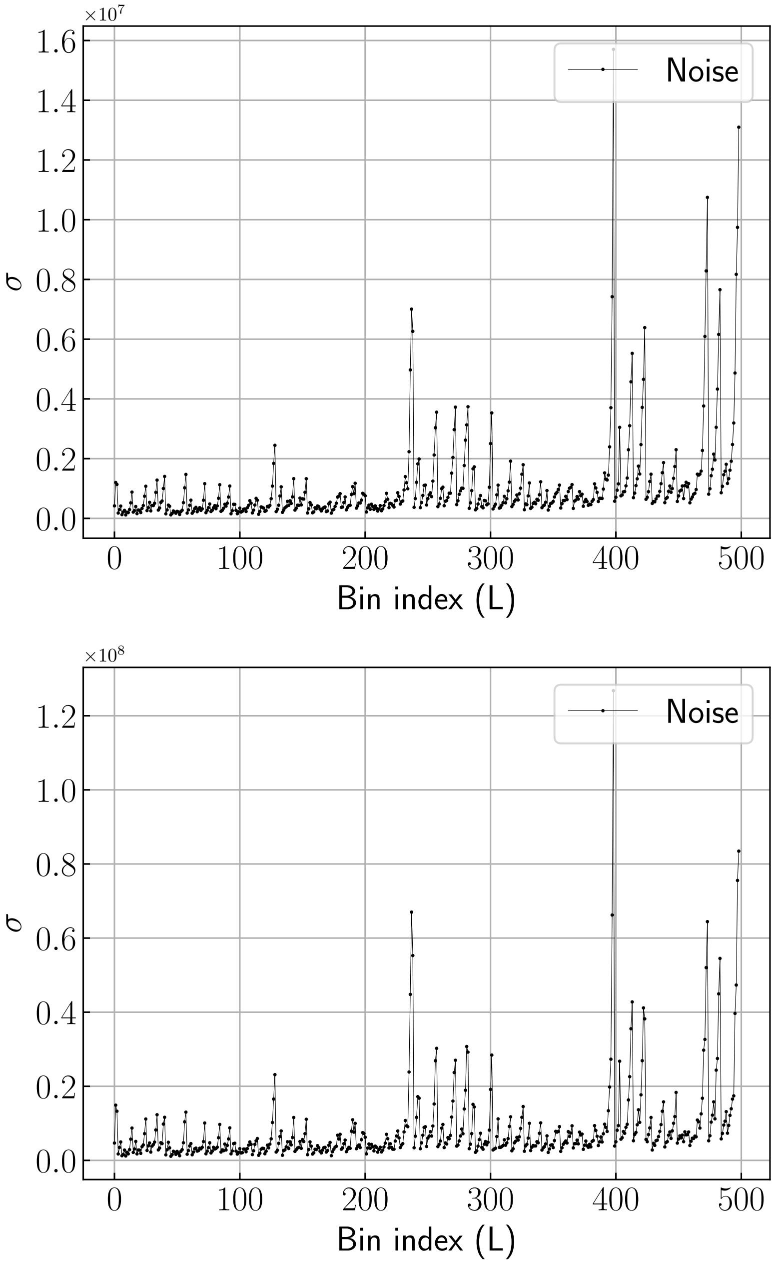


Figure 6.5: Noise associated to each bin of the optimal (reduced) angular trispectrum estimation. The top panel shows the noise estimation for $z_{cen} = 0.5$, while the bottom panel corresponds to $z_{cen} = 0.26$. Following the choice of [120], the optimal estimators are scaled by a factor $\ell_1\ell_2\ell_3\ell_4$. Additionally, we applied a further rescaling by a factor of $10^{24}\sigma_z^{-4}$ to enhance readability. The bin index is a combination of bins corresponding to a valid configuration under the triangular condition, and it indicates gradually increasing scales starting from $\ell = 13$ up to $\ell = 37$ (for the adopted binning scheme).

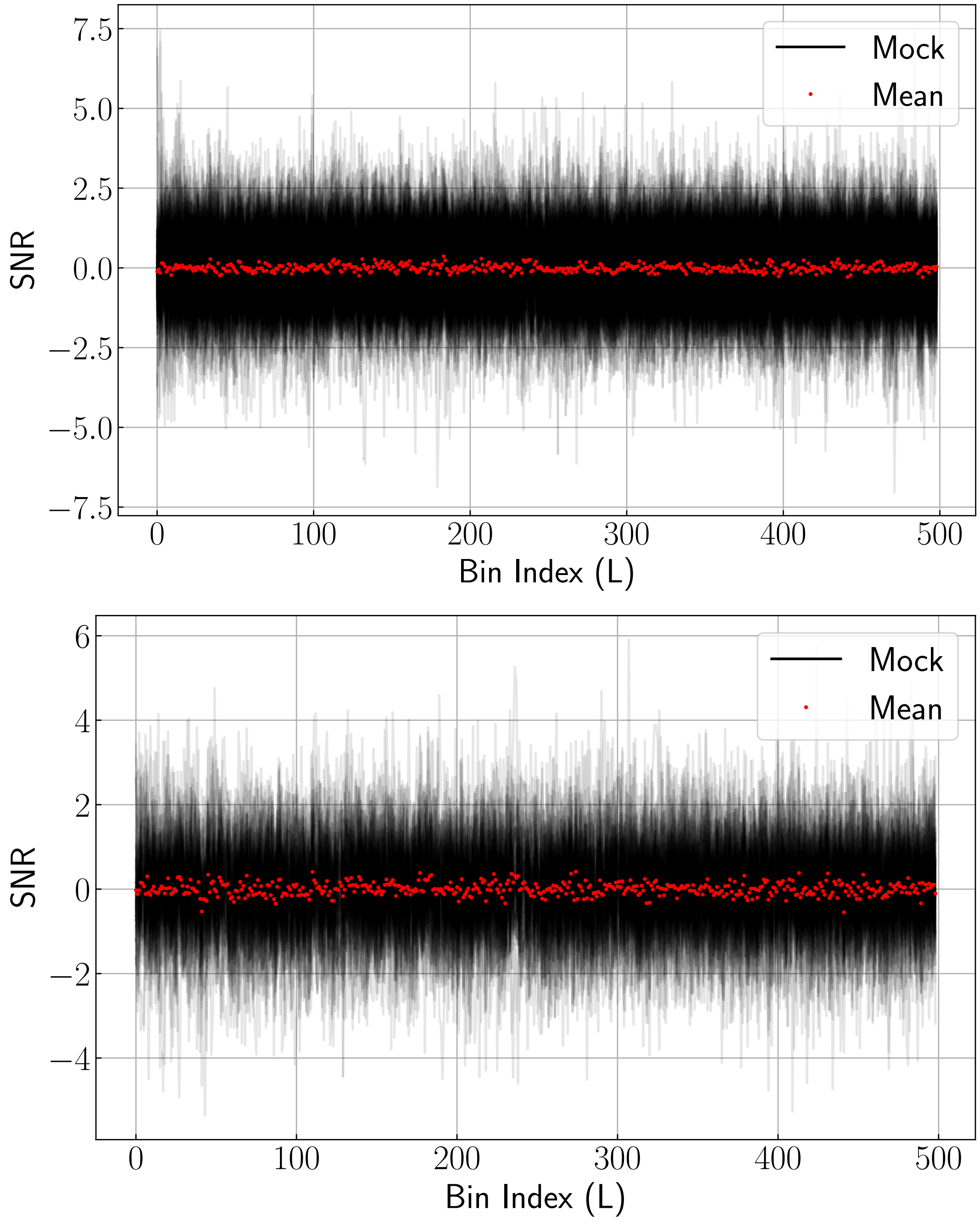


Figure 6.6: Mean SNR (black dots) varying through bins of MultiDark-`PATCHY` mocks. The top panel shows the results obtained from 370 mocks (each one represented by a grey curve) imposing Gaussian shells with $z_{cen} = 0.5$ and $\sigma_z = 0.01$. The bottom panel illustrates the same study on 100 mocks at $z_{cen} = 0.26$ and $\sigma_z = 0.01$.

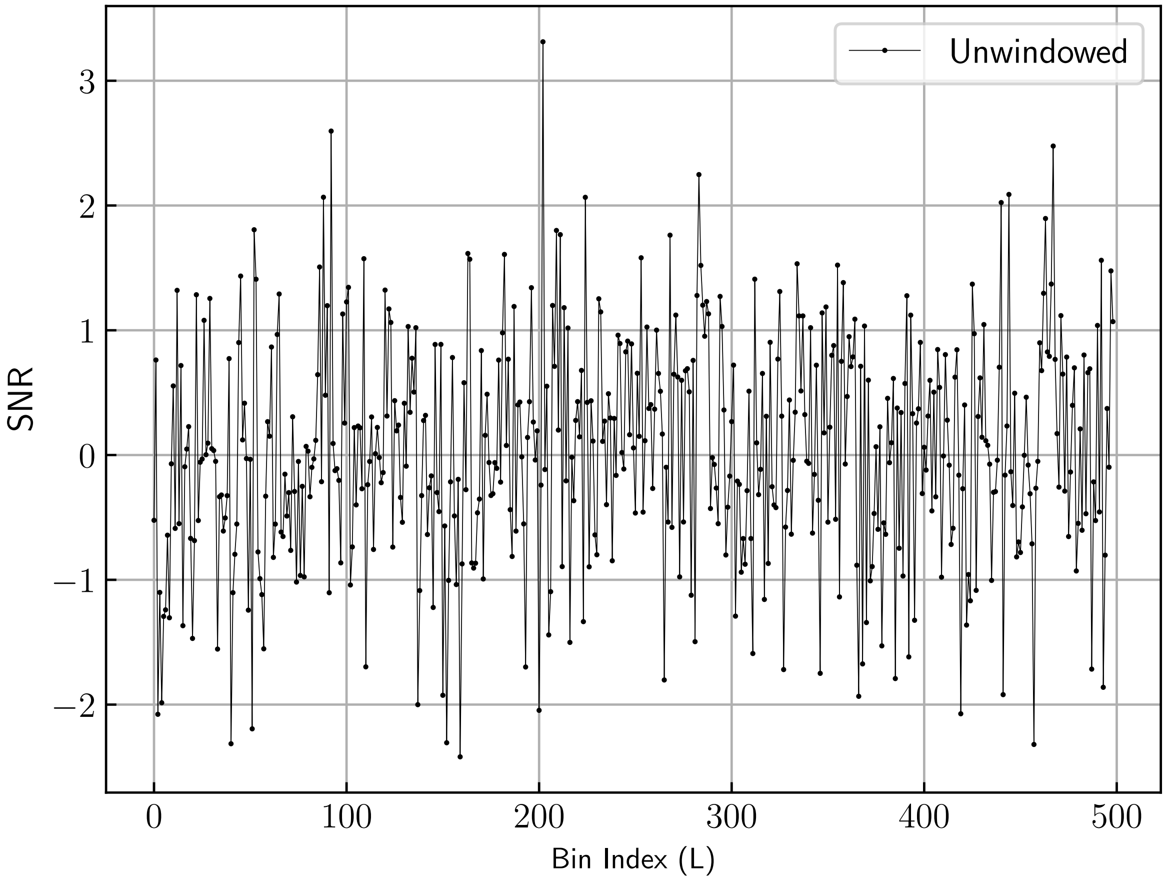


Figure 6.7: Variability of the SNR of the real galaxy catalogue of BOSS DR12, imposing Gaussian shells with $z_{cen} = 0.5$ and $\sigma_z = 0.01$.

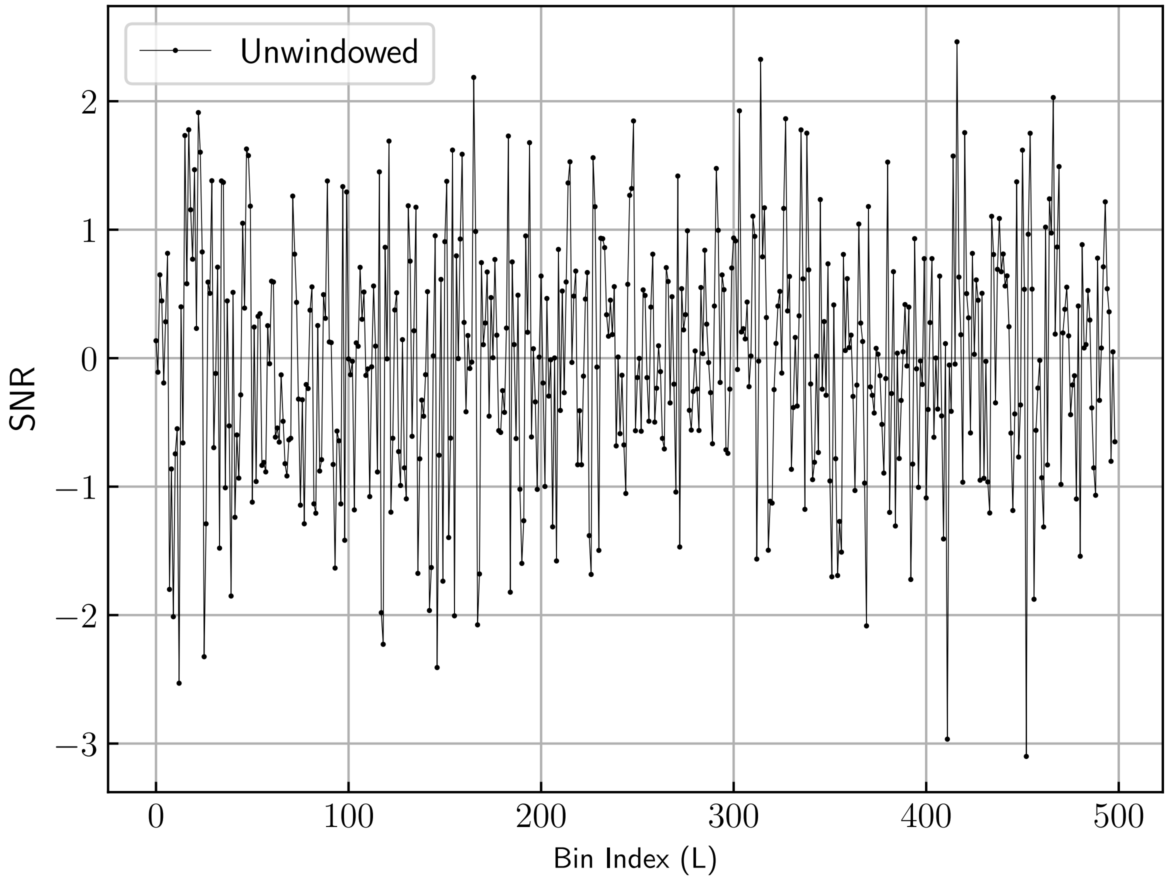


Figure 6.8: As in Fig.6.7 but with $z_{cen} = 0.26$ and $\sigma_z = 0.01$.

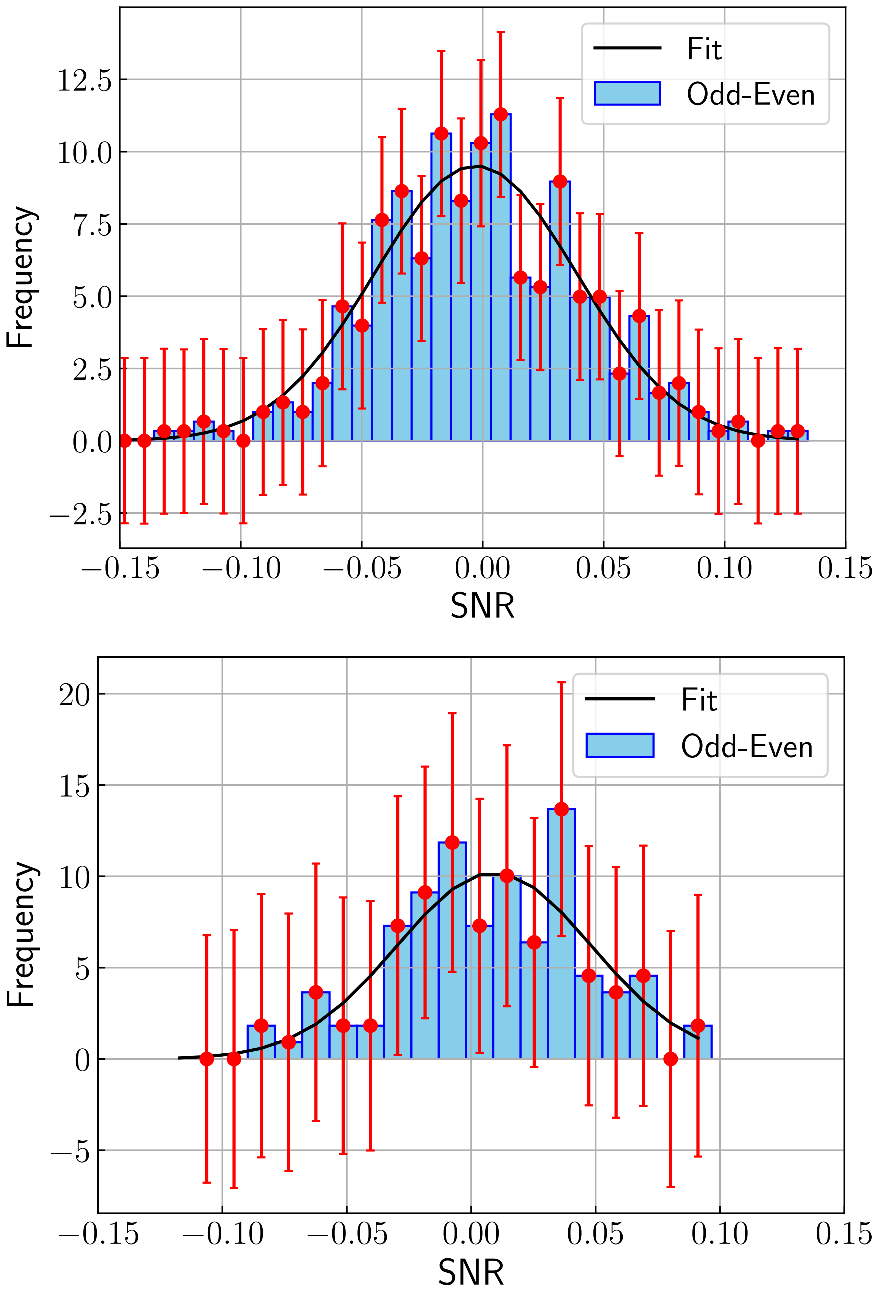


Figure 6.9: Gaussian fit (black curve) of the scale-averaged SNR distribution for the difference between the odd and even components of the MultiDark-`PATCHY` mocks trispectra. The top panel shows the distribution at $z_{cen} = 0.50$, while the bottom panel corresponds to $z_{cen} = 0.26$. Red error bars represent the uncertainties on the frequencies.

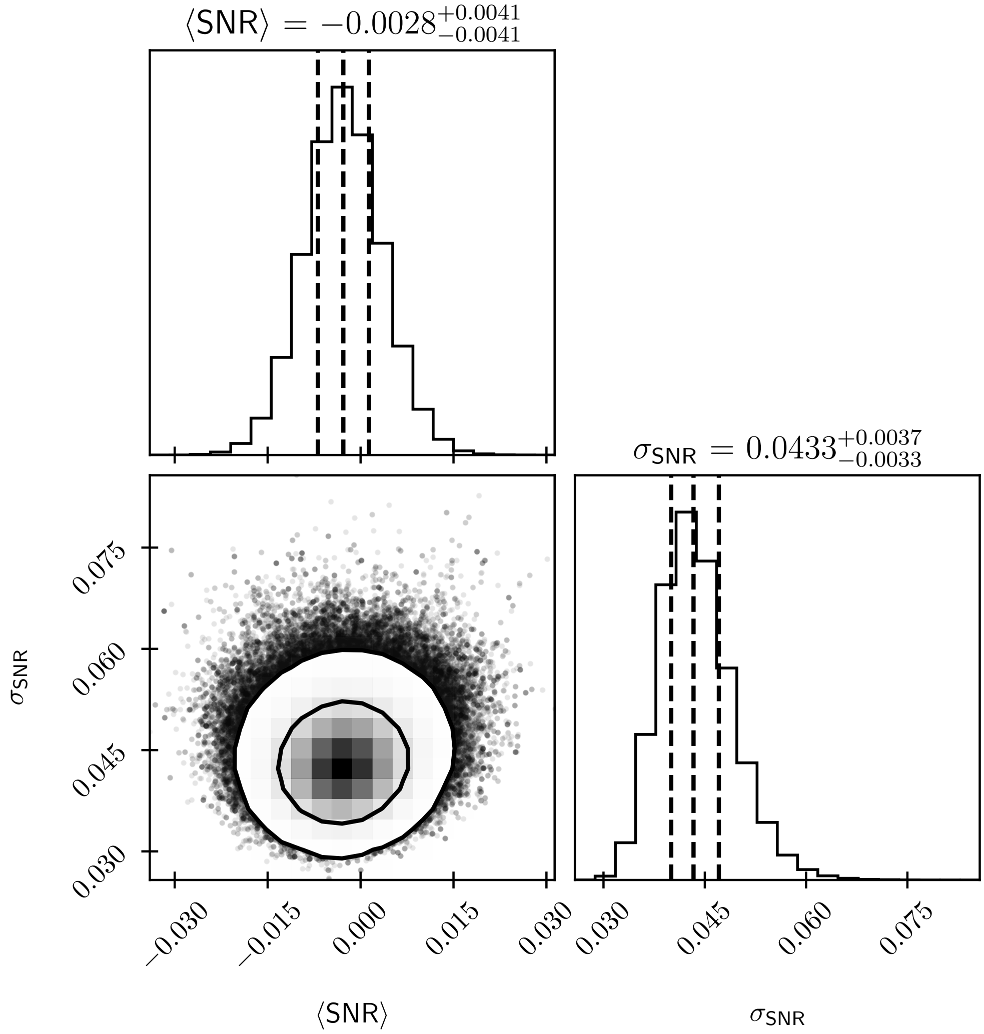


Figure 6.10: Contour plot showing the mean and standard deviation of the Gaussian MCMC fit for the parity test at $z_{cen} = 0.5$ and $\sigma_z = 0.01$. The contours represent the 68% and 95% confidence levels. Additionally, the marginalized posterior distributions are plotted for each parameter, including vertical dashed lines for the median and the quartiles.

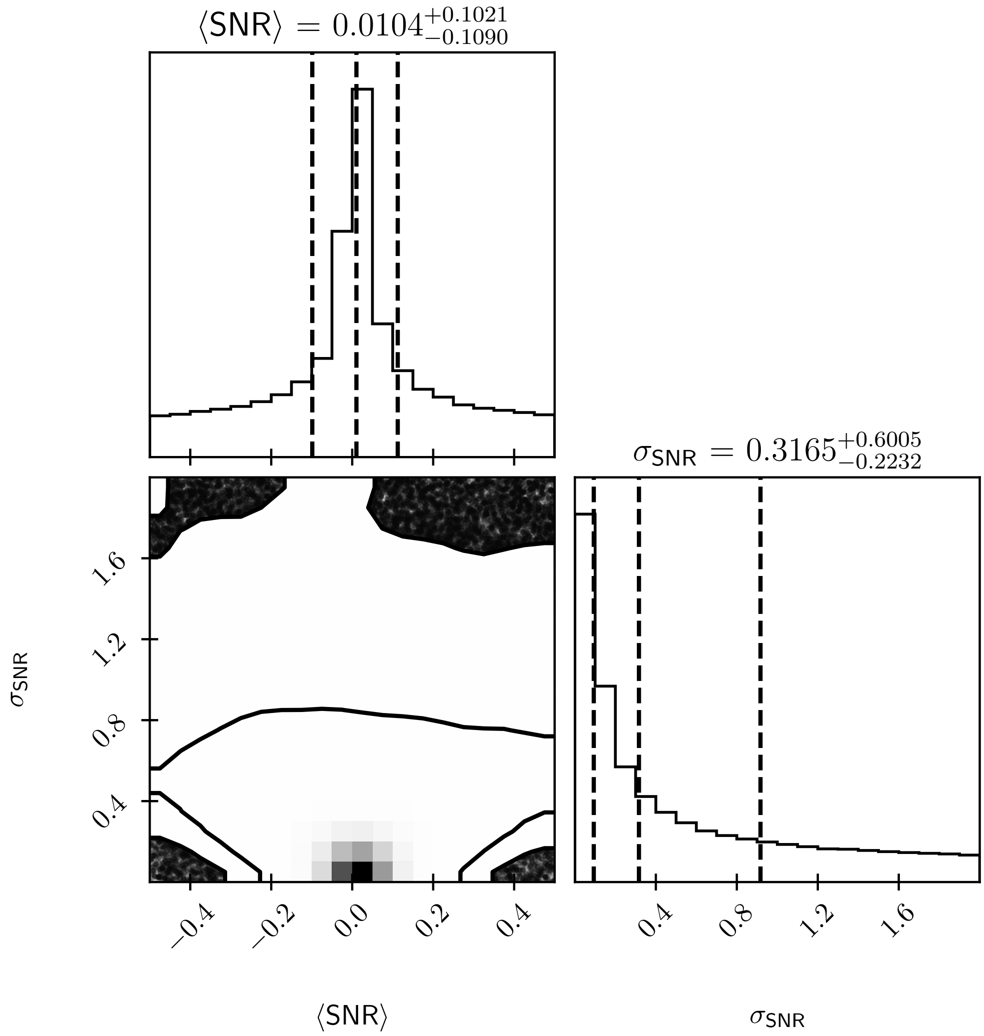


Figure 6.11: As in Fig6.10 but with $z_{cen} = 0.26$ and $\sigma_z = 0.01$.

6.2 Tests on Quijote

So far, in this work no parity-violating features have been detected in the analyzed data sets. To further validate our methods, it is worthwhile to investigate whether parity violation can be detected in simulations where it is inherently present by construction. In this section, we perform a parity test by searching for signatures of non-zero signals in the difference between the odd components of the trispectrum obtained from the QUIJOTE-ODD simulations, which include parity violation, and those obtained from the fiducial LCDM QUIJOTE simulations, which are parity-symmetric.

6.2.1 Construction of mock catalogues

As mentioned in the previous chapter, the QUIJOTE simulations provide halo catalogues in comoving coordinates. To work with ARFs, we constructed light-cones to transform the simulation coordinates into observed coordinates (RA, Dec, z). These mocks were generated by constructing a light cone from a single snapshot's redshift (see Sec.5.3). A validation pipeline was subsequently applied to ensure the accuracy of the generated mocks, focusing in particular on comparing the halo 2PCF with theoretical predictions.

The first observation was that the halo catalogues were incomplete due to the resolution limitations of the simulation, having fewer low-mass halos than expected. Specifically, the effective halo bias, calculated by averaging the linear bias of halos across a given mass range, is given by:

$$b_{\text{eff-mass}}(z) = \frac{1}{N_{\text{halo}}} \sum_{i=1}^{N_{\text{halo}}} b(M_i, z_i), \tag{6.1}$$

where N_{halo} is the number of halos in the catalogue and b is the linear bias. This effective bias was found to be inconsistent with the bias derived averaging on the theoretical mass function over the full halo mass range, defined in Eq.(5.18). In this analysis, we adopted the linear bias model of [166] and the mass function of [165].

To overcome this issue, as a standard practice we applied a mass cut, removing low-mass halos below the resolution limit. After this selection, the two bias estimates are consistent at $b_{\text{eff}} \sim 4.8$ for all the mocks of QUIJOTE_ODD and fiducial LCDM QUIJOTE. Figures 6.13a and 6.13b illustrate the effects of this mass cut and its impact on the total number of objects in a halo catalogue of ODD_M.

Finally, we measured the 2PCF from the mocks and compared it to the Fourier transform of the `CAMB` power spectrum. The 2PCF estimation was performed using the Landy-Szalay estimator [91], an improved version of Eq.(2.68):

$$\xi(r) = \frac{DD(r) - 2DR(r) + RR(r)}{RR(r)}, \tag{6.2}$$

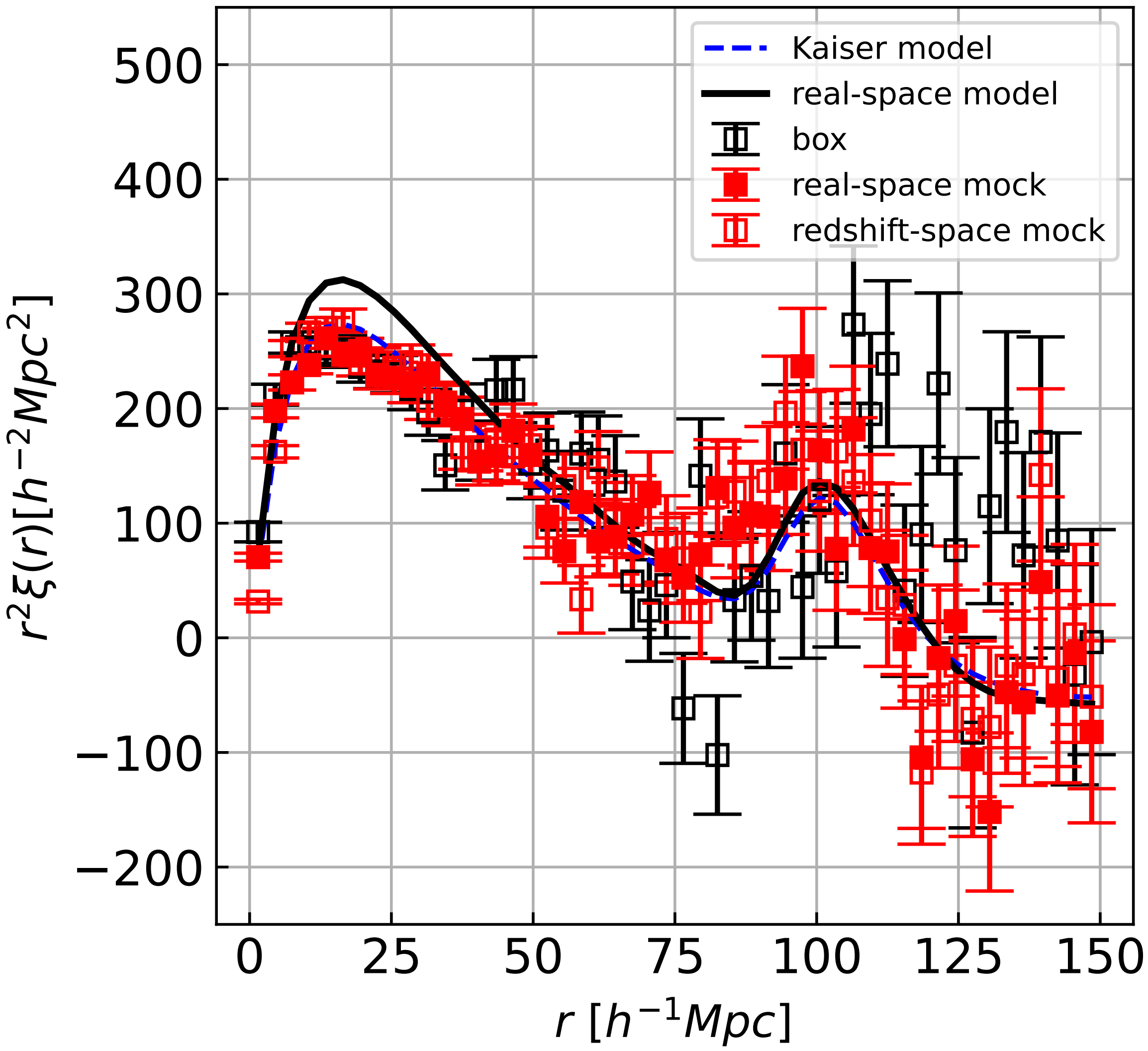
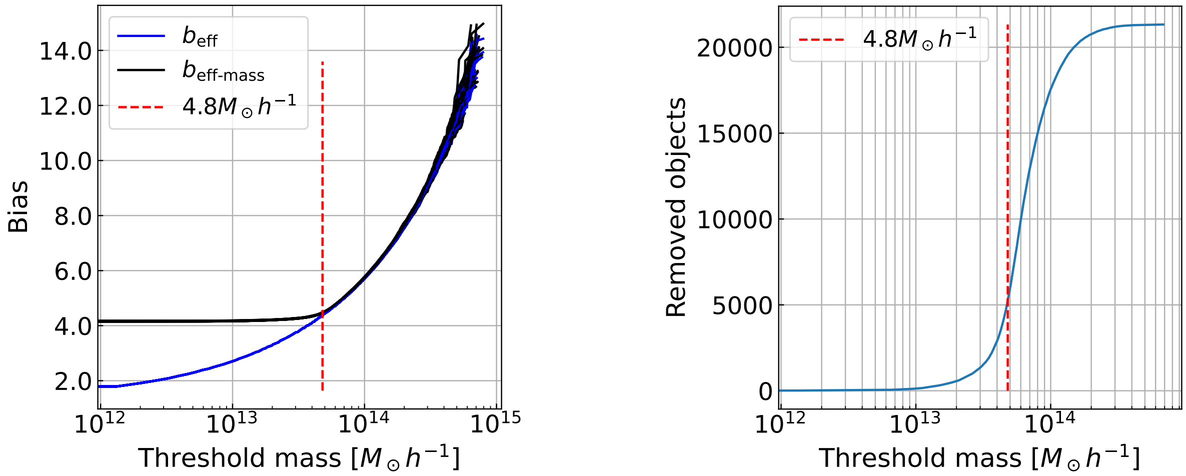


Figure 6.12: The 2PCF measured from a fiducial LCDM Quijote simulation. The solid black curve and the blue dashed curve represent the 2PCF in real-space and redshift-space, respectively, as predicted by `CAMB`. The squares, shown with their associated Poissonian errors, correspond to 2PCF measurements using the Landy-Szalay estimator. Specifically, the black-filled squares indicate the 2PCF measured from the original box, while the red squares represent the 2PCF from the constructed mock, with filled squares for real-space and empty squares for redshift-space measurements.



(a) The two effective bias estimates as a function of the threshold mass. b_{eff} and $b_{\text{eff-mass}}$ refer, respectively, to the definitions in Eq.(5.18) and Eq.(6.1).

(b) The number of removed objects for different mass cuts. The mass cut application reduced the number of halos by 5000.

Figure 6.13: Effects of the completeness cut performed on the ODD_M mocks. The red dashed vertical line sets the mass threshold for the cut at $M \leq 4.8 \times 10^{13}\ M_\odot$.

where the $DR(r)$ term refers to the normalized number of data-random pairs as a function of the comoving separation r, which helps to reduce the estimation bias at large scales. Similarly, $DD(r)$ represents the normalized number of data-data pairs, and $RR(r)$ the number of random-random pairs. The redshift-space 2PCF model, $\xi(s)$, was estimated with the so-called Kaiser model [103]:

$$\xi(s) = 1 + \left[\frac{2}{3}\frac{f}{b} + \frac{1}{5}\left(\frac{f}{b}\right)^2\right]\xi(r). \tag{6.3}$$

Here, $\xi(r)$ is the real-space 2PCF introduced in Sec.2.3.1 while $\xi(s)$ is the redshift-space 2PCF, which depend on the observed separation s and accounts for the distortion effects given by the peculiar velocities of the underlying density field. b is the linear bias term and f is the linear growth factor of cosmic structures, already defined at the beginning of Sec.4.1.

The measurements are statistically consistent with the theoretical expectations, both in real and redshift space, confirming the correctness of the mock generation process. An example of the 2PCF is presented in Fig.6.12, where we used a halo catalogue from the fiducial LCDM Quijote simulation, after applying a small redshift cut ($\Delta z = 0.01$) at the borders of the mock to preserve its cubic shape. Specifically, we used `CAMB` to compute the 2PCF in both real- and redshift-space, and applied the Landy-Szalay estimator in

Eq.(6.2) to measure the 2PCF for the mock in both comoving and observed coordinates.

6.2.2 Power spectrum and trispectrum

We now retrace the entire BOSS pipeline described in Sec.6.1, this time without applying any weighting to the pixels. To avoid underestimation of the power spectrum signal and overestimation of the shot noise, we introduced a mask to filter out empty sky regions. We worked with two choices of reference redshift for the ARFs' Gaussian shell, $z_{cen} = 0.8$ and $z_{cen} = 1.25$, while keeping $\sigma_z = 0.01$ for tomographic purposes. As shown in Fig.6.14, the halo catalogues occupy only a small fraction of the full survey footprint, resulting in a sky fraction of approximately 2%. This limited coverage negatively impacted the power spectrum estimation, as the shot noise was comparable to the `ARFCAMB`-predicted C_ℓ. Consequently, the observed signal exhibited a quasi-flat shape, as illustrated in Fig.6.15 referring to ODD_M catalogues. Notice how the ideal estimator provides a biased power spectrum, due to the drastic mask employed which introduces significant effects not accounted for by this estimator.

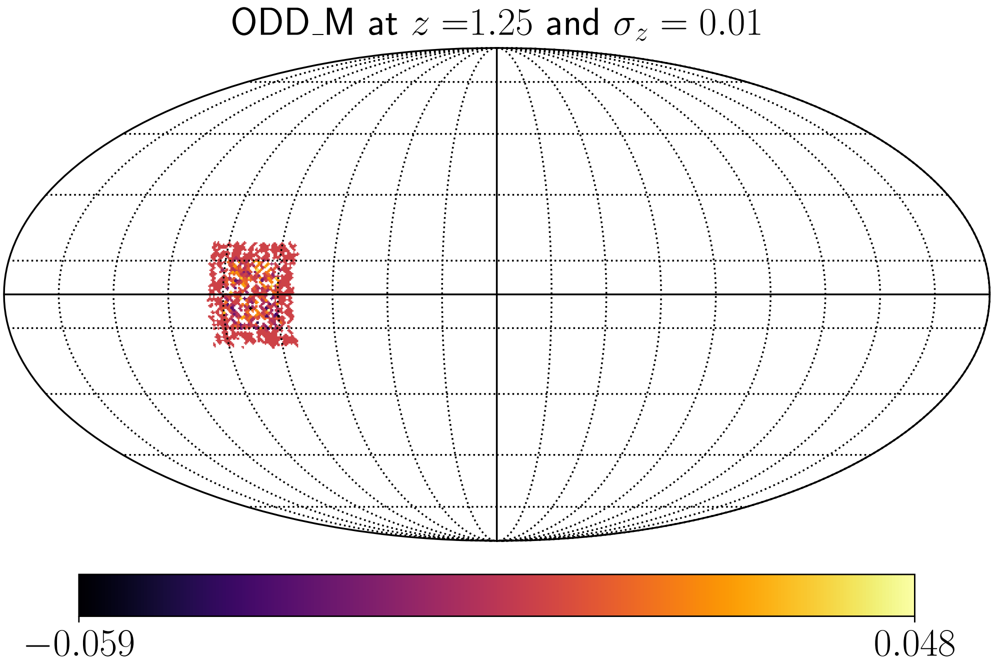


Figure 6.14: Mollwide projection of ARF maps computed at $N_{side} = 64$ for a Quijote-ODD_M mock. We imposed a Gaussian shell with $z_{cen} = 1.25$ and $\sigma_z = 0.01$, and a mask to obtain a occupied sky fraction of around 2%.

In this work, we analyzed a total of 1050 mocks (350 for each family of Quijote simulation halo catalogues) at $z_{cen} = 0.80$ and 1500 mocks (500 for each family of Qui-

JOTE simulation halo catalogues, thus utilizing all the available mocks) at $z_{cen} = 1.25$, always keeping $\sigma_z = 0.01$ for tomographic purposes. The binning scheme adopted for the trispectrum comprised $N_{bins} = 3$ bins, with $\ell_{min} = 12$ and a bin width of $\Delta_\ell = 2$. The last bin was excluded to mitigate edge effects. To estimate the Fisher matrix, we generated $N_{it} = 50$ GRF realizations, executed using 32 CPUs. This configuration enabled fast execution times (around 500 wall-clock seconds per mock), making it feasible to analyze a large number of mocks. To reduce the impact of systematic uncertainties, the fit was performed by subtracting the odd-trispectra of the fiducial ΛCDM catalogues from those of the parity-violating ODD_M and ODD_P catalogues. The average SNR as a function of the bin index is shown in Fig.6.16 for both parity-violating simulations at $z_{cen} = 0.8$ and $z_{cen} = 1.25$.

From the scale-averaged SNR of each mock, we constructed histograms, estimating uncertainties on the frequencies using Monte Carlo sampling, as described in Sec.6.1. The resulting distributions were then fitted using an MCMC procedure.

For $z_{cen} = 1.25$, the histograms and their corresponding fitting curves are shown in the top panel of Fig.6.21 for all QUIJOTE mock sets. The contour plots of the fitted parameters, presented in Figs.6.17-6.18, reveal that for ODD_M, the mean signal is $\langle \mathrm{SNR} \rangle = 0.0011^{+0.0008}_{-0.0008}$, consistent with zero within one $\sigma_{\mathrm{SNR}} = 0.0227^{+0.0008}_{-0.0007}$. Similarly, for ODD_P, the mean value $\langle \mathrm{SNR} \rangle = 0.0004^{+0.0008}_{-0.0008}$ is compatible with zero, and $\sigma_{\mathrm{SNR}} = 0.0229^{+0.0008}_{-0.0007}$.

At $z_{cen} = 0.80$, the same analysis yielded best-fit Gaussian parameters of $\langle \mathrm{SNR} \rangle = -0.0030^{+0.0010}_{-0.0010}$, $\sigma_{\mathrm{SNR}} = 0.0236^{+0.0010}_{-0.0009}$ for the ODD_M mocks, and $\langle \mathrm{SNR} \rangle = -0.0031^{+0.0011}_{-0.0010}$, $\sigma_{\mathrm{SNR}} = 0.0240^{+0.0010}_{-0.0010}$ for the ODD_P mocks, as depicted in the contour plots of Figs.6.19-6.20 and in the fits at the bottom of Fig.6.21.

These findings lead to the conclusion that the analyzed distributions exhibit no significant evidence of parity-violating features in the analyzed mock catalogues, as all results remain consistent with a parity-symmetric scenario within the statistical uncertainties. This indicates that the sensitivity of the ARF-based pipeline employed is not sufficient to detect parity-violating features, even when such features are introduced by construction in the catalogues. This result is given by the fact that the employed mocks do not have enough statistics, meaning that the halo density and the sky fraction with the employed mask are not sufficient to generate an appreciable signal.

Future works could focus on conducting a more detailed analysis using a larger number of mock catalogues for both BOSS DR12 and QUIJOTE simulations. For the BOSS DR12 catalogues, this would help reduce uncertainties in the frequency estimates, leading to more stringent and robust results. In the case of QUIJOTE simulations, an increased number of realizations could highlight potential parity-violating effects, and could provide more awareness of systematics. Moreover, improving the choice of masks and pixel weightings in QUIJOTE simulations could mitigate noise contributions and yield more reliable results. Exploring alternative binning schemes and different window function configurations for ARFs might also enhance the precision and accuracy of the measurements, providing further insight into possible parity asymmetries.

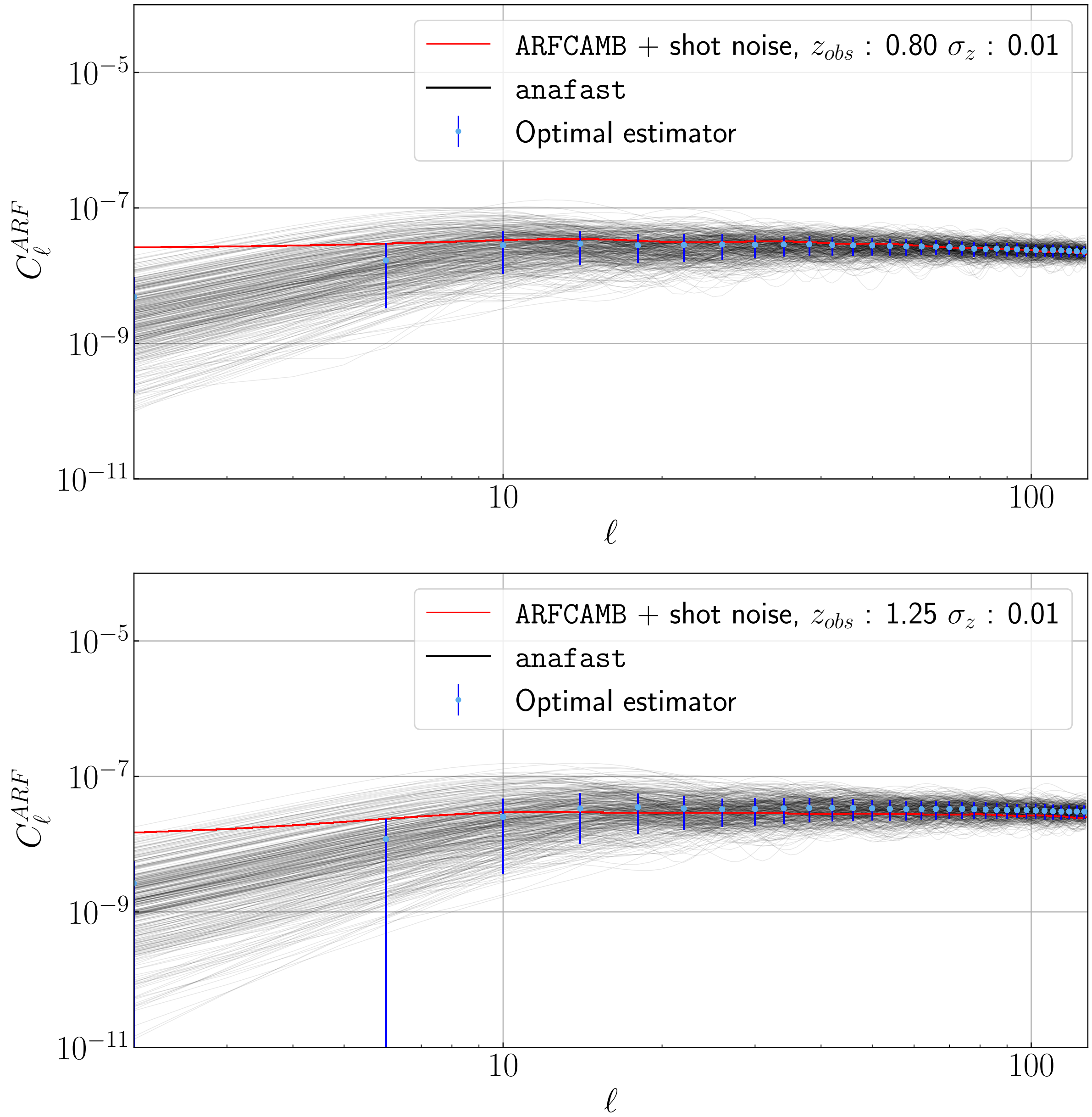


Figure 6.15: Auto-angular power spectrum obtained from the maps of 500 mock catalogues of ODD_M simulations under a Gaussian shell with $\sigma_z = 0.01$, at $z_{cen} = 0.8$ (top panel) and $z_{cen} = 1.25$ (bottom panel). The gray curves show the measured C_ℓ from the maps, while the blue and green dots represent the mean values of the optimal and ideal estimations from `PolyBin`, with associated error bars. The red curve corresponds to the average `ARFCAMB` output for each mock, including the shot noise contribution.

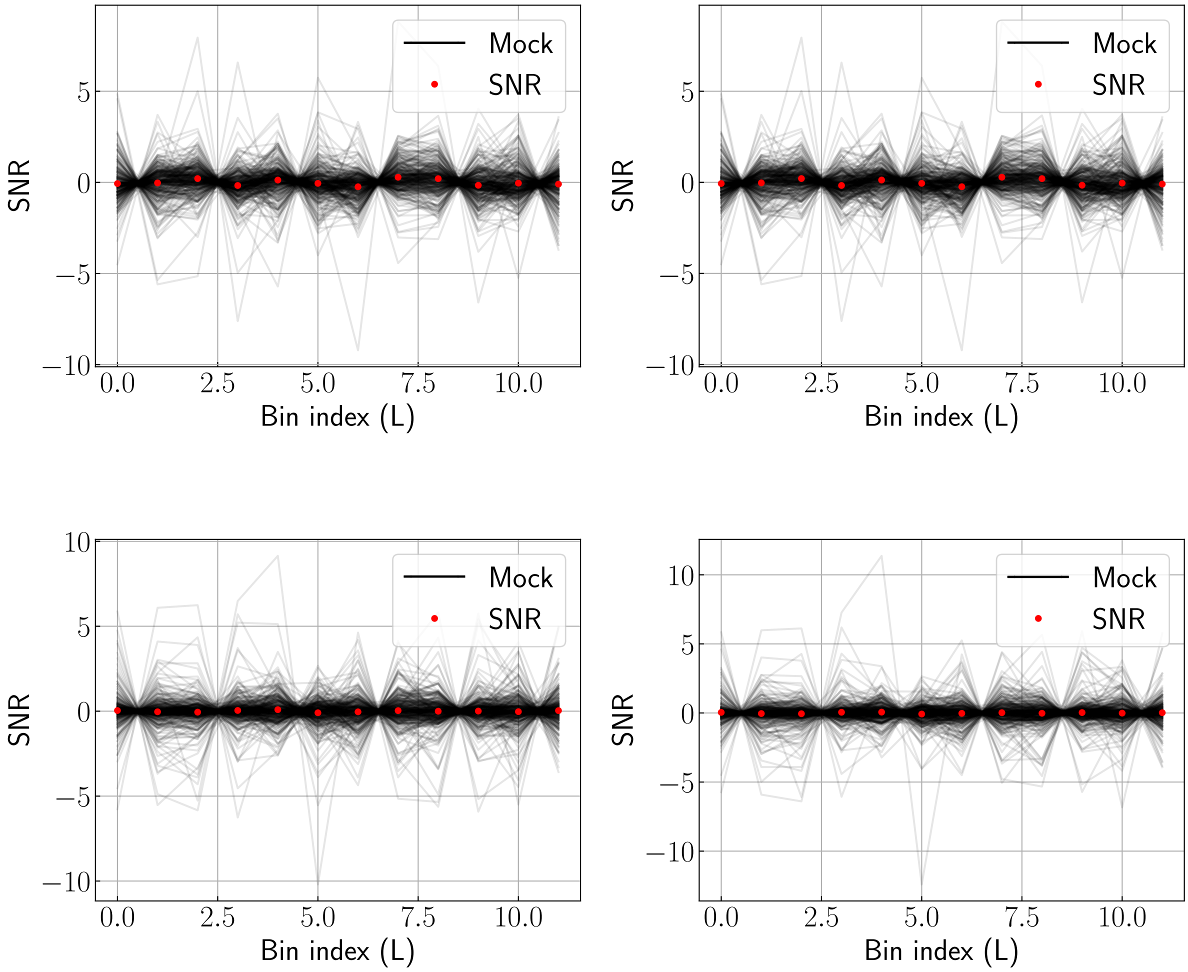


Figure 6.16: Mean SNR (red dots) varying through bins of Quijote-ODD mocks. The top panel shows the results obtained from ODD_M (left) and ODD_P (right) mocks, each one represented by a grey curve, imposing Gaussian shells with $z_{cen} = 1.25$ and $\sigma_z = 0.01$. The bottom panel illustrates the same study at $z_{cen} = 0.80$ and $\sigma_z = 0.01$, with the ODD_M SNR on the left and the ODD_P SNR on the right.

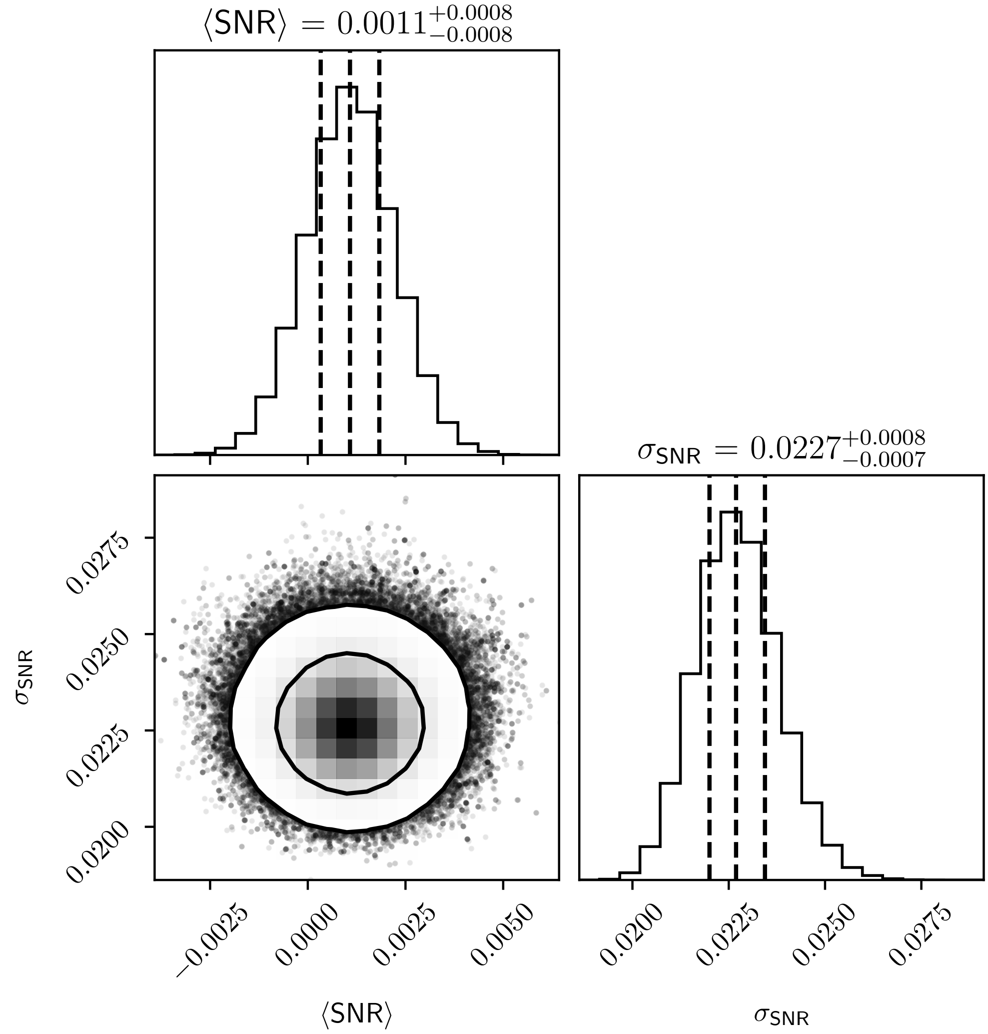


Figure 6.17: Contour plot showing the mean and standard deviation of the Gaussian MCMC fit for the parity test on ODD_M mocks at $z_{cen} = 1.25$. The contours represent the 68% and 95% confidence levels. Additionally, the marginalized posterior distributions are plotted for each parameter, including vertical dashed lines for the median and quartiles.

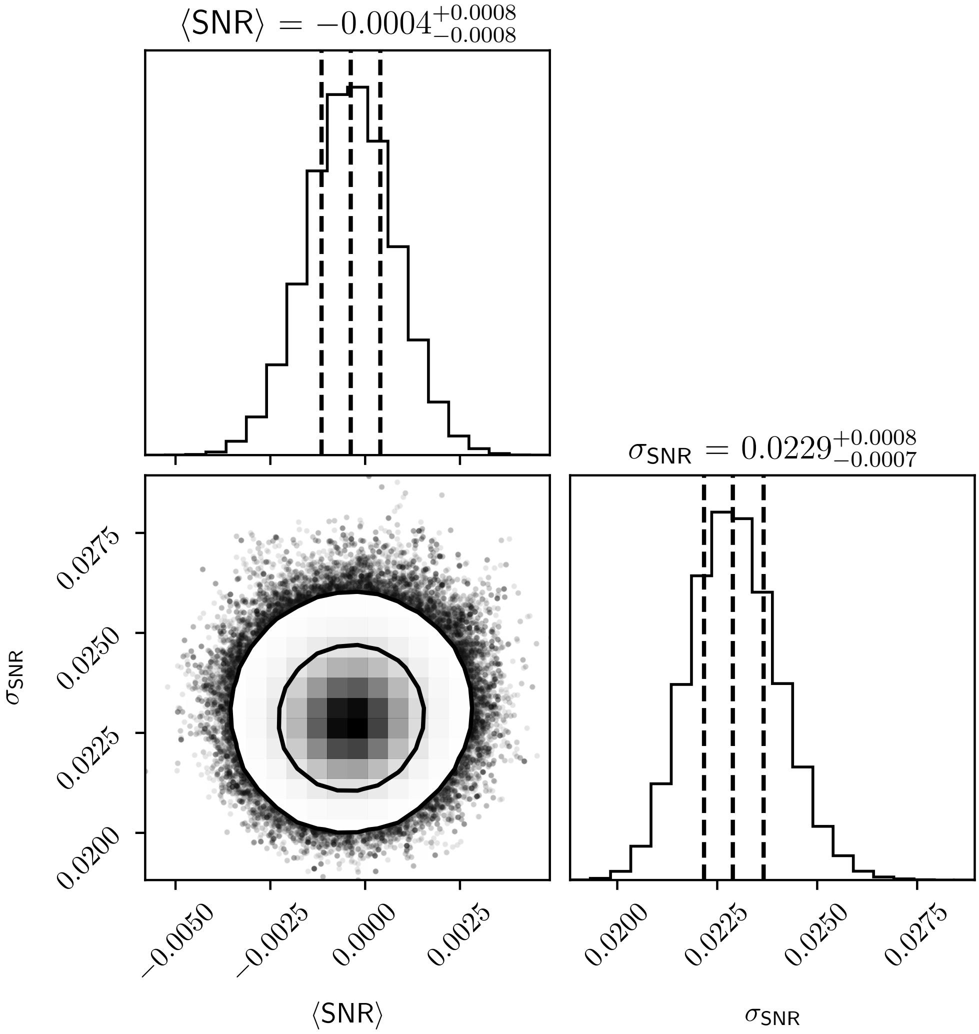


Figure 6.18: As in Fig.6.17 but for ODD_P mocks at $z_{cen} = 1.25$.

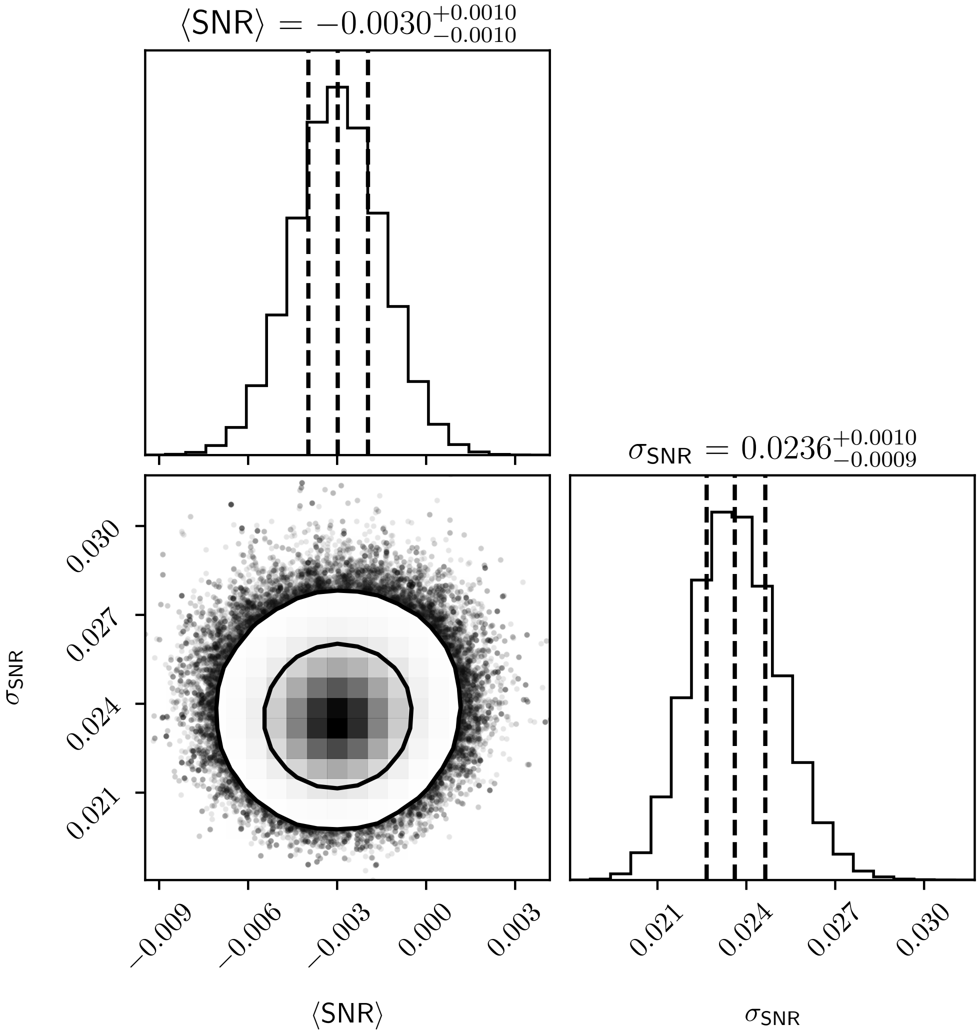


Figure 6.19: As in Fig.6.17 but for ODD_M mocks at $z_{cen} = 0.8$.

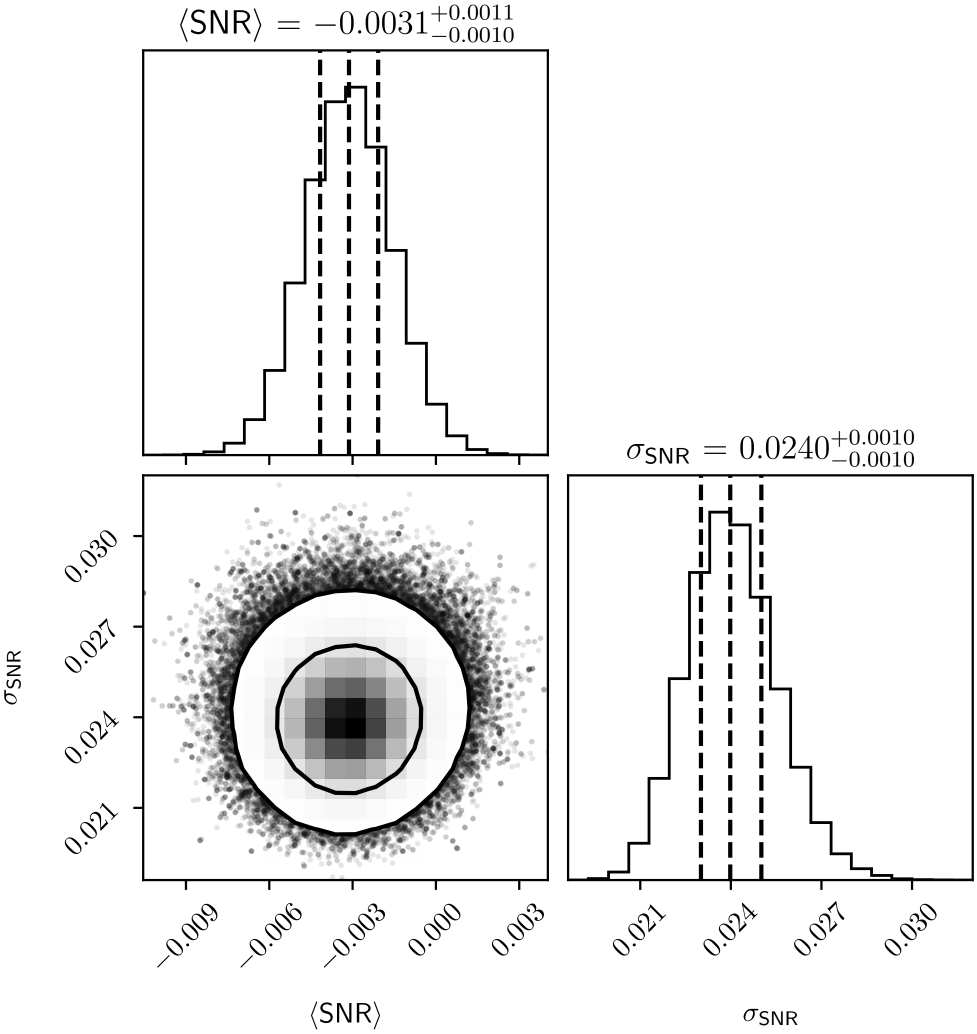


Figure 6.20: As in Fig.6.17 but for ODD_P mocks at $z_{cen} = 0.8$.

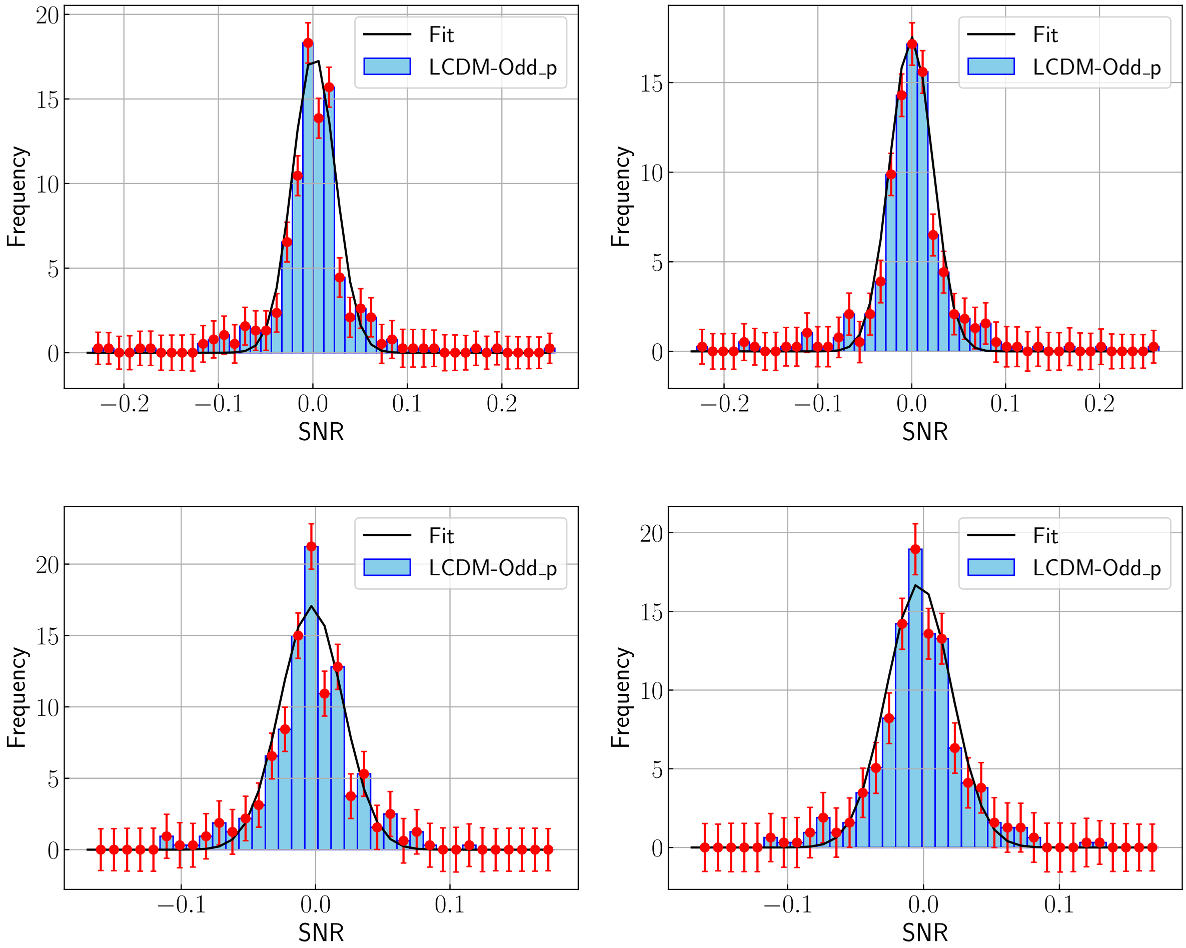


Figure 6.21: Gaussian fit (black curve) of the scale-averaged SNR distribution for the difference between the odd components of the fiducial LCDM QUIJOTE mocks' trispectra and the QUIJOTE-ODD mocks' trispectra. The top panel displays the fits at $z_{cen} = 1.25$ for ODD_M (left) and ODD_P (right), while the bottom panel shows the fits at $z_{cen} = 0.8$ for ODD_M (left) and ODD_P (right). Red error bars indicate the uncertainties on the frequencies.

Conclusions

In this work, we first provided an introduction to cosmology, starting from spacetime modelling and arriving at a description of the emergence of the LSS, offering an overview of the main statistical tools used to study matter clustering. We discussed the concept of parity violation, how it can be investigated, and detailed the ARFs statistics used as a probe for testing parity symmetry. Then, we provided a detailed description of the BOSS DR12 catalogues, the Quijote simulations, and all the codes employed in our analysis, with particular attention to `PolyBin` and the role of trispectra. Finally, we presented the main result of our analysis and discussed about possible future improvements of the method.

One notable contribution of this thesis is the development of the MapCalculator class, which was added to the `CBL` project, expanding this extensive library for cosmological calculations and data manipulation. While the code is functional and well-structured, there is room for improvement. Future work could focus on enhancing its computational efficiency, adding new methods to improve the user experience, and allowing for greater customization in the production of maps. Additionally, this class lays the groundwork for integrating `HEALPix` into `CBL`, opening up possibilities for creating generic sky maps useful in cosmology, beyond the specific implementation of ARFs and ADFs.

Regarding the parity violation tests, no evidence of parity asymmetry was found in the BOSS DR12 catalogues analyzed. While at first glance, this result may seem inconsistent with previous findings using the 4PCF on the same data set [120, 28, 76] (though [120] updated their results to a lower detection significance [123]) it aligns with other null results reported in [90, 72].

In particular, [90] introduced a new statistic correcting biases arising from the mismatch between the parity-even 8PCF of observational data and mocks, concluding that there is no evidence of parity violation in BOSS. Additionally, [114] highlighted how relativistic corrections to Newtonian redshift-space distortions [79] could lead to parity-violating effects, emphasizing the importance of addressing such systematics. This suggests also that parity asymmetries could arise even without primordial violations.

From the perspective of the CMB, recent techniques investigating cosmic birefringence [51, 42] have identified evidence of parity-violating features, while the analysis in [121] reported null results.

We also tested the ARF statistics' sensitivity on QUIJOTE-ODD simulations [38], designed to include parity violations from the initial conditions. Even in this case, no evidence of parity violation was detected. This suggests that, with the employed pipeline, the signal from ARFs is not sufficiently strong to detect parity-violating features.

Determining whether the Universe exhibits parity violation is a non-trivial challenge. While current evidence is insufficient to claim a definitive parity violation, significant work remains to validate existing results, address biases due to systematic uncertainties, and refine statistical tools capable of detecting parity asymmetries.

The take-home message is clear: while the Universe's parity symmetry remains an open question, progresses in addressing systematic effects and refining observational techniques will be key to advancing our understanding in this domain.

Bibliography

[1] SDSS Science Archive Server (SAS). *BOSS DR12*. URL: `https://data.sdss.org/sas/dr12/boss/`.

[2] B P Abbott et al. "LIGO: the Laser Interferometer Gravitational-Wave Observatory". In: *Reports on Progress in Physics* 72.7 (June 2009), p. 076901. ISSN: 1361-6633. DOI: `10.1088/0034-4885/72/7/076901`. URL: `http://dx.doi.org/10.1088/0034-4885/72/7/076901`.

[3] Guillermo F. Abellán, Riccardo Murgia, and Vivian Poulin. "Linear cosmological constraints on two-body decaying dark matter scenarios and the S_8 tension". In: *Physical Review D* 104.12 (Dec. 2021). ISSN: 2470-0029. DOI: `10.1103/physrevd.104.123533`. URL: `http://dx.doi.org/10.1103/PhysRevD.104.123533`.

[4] T. Accadia et al. "Virgo: a laser interferometer to detect gravitational waves". In: *Journal of Instrumentation* 7.3 (Mar. 2012), p. 3012. DOI: `10.1088/1748-0221/7/03/P03012`.

[5] P. A. R. Ade et al. "Planck2013 results. XVI. Cosmological parameters". In: *Astronomy & Astrophysics* 571 (Oct. 2014), A16. ISSN: 1432-0746. DOI: `10.1051/0004-6361/201321591`. URL: `http://dx.doi.org/10.1051/0004-6361/201321591`.

[6] P. A. R. Ade et al. "Planck2015 results: XIII. Cosmological parameters". In: *Astronomy & Astrophysics* 594 (Sept. 2016), A13. ISSN: 1432-0746. DOI: `10.1051/0004-6361/201525830`. URL: `http://dx.doi.org/10.1051/0004-6361/201525830`.

[7] P. A. R. Ade et al. "Bicep/Keck XV: The Bicep3 Cosmic Microwave Background Polarimeter and the First Three-year Data Set". In: *The Astrophysical Journal* 927.1 (Mar. 2022), p. 77. DOI: `10.3847/1538-4357/ac4886`. URL: `https://dx.doi.org/10.3847/1538-4357/ac4886`.

[8] Peter Adshead et al. "Gauge-preheating and the end of axion inflation". In: *Journal of Cosmology and Astroparticle Physics* 2015.12 (Dec. 2015), pp. 034–034. ISSN: 1475-7516. DOI: `10.1088/1475-7516/2015/12/034`. URL: `http://dx.doi.org/10.1088/1475-7516/2015/12/034`.

[9] N. Aghanim et al. “Planck2018 results: VI. Cosmological parameters”. In: *Astronomy & Astrophysics* 641 (Sept. 2020), A6. ISSN: 1432-0746. DOI: `10.1051/0004-6361/201833910`. URL: `http://dx.doi.org/10.1051/0004-6361/201833910`.

[10] Simone Aiola et al. “The Atacama Cosmology Telescope: DR4 maps and cosmological parameters”. In: *Journal of Cosmology and Astroparticle Physics* 2020.12 (Dec. 2020), pp. 047–047. ISSN: 1475-7516. DOI: `10.1088/1475-7516/2020/12/047`. URL: `http://dx.doi.org/10.1088/1475-7516/2020/12/047`.

[11] Shadab Alam et al. “The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III”. In: *Astrophysical Journal Supplement* 219.1, 12 (July 2015), p. 12. DOI: `10.1088/0067-0049/219/1/12`. arXiv: `1501.00963 [astro-ph.IM]`.

[12] Stephon Alexander and Nicolás Yunes. “Chern–Simons modified general relativity”. In: *Physics Reports* 480.1–2 (Aug. 2009), pp. 1–55. ISSN: 0370-1573. DOI: `10.1016/j.physrep.2009.07.002`. URL: `http://dx.doi.org/10.1016/j.physrep.2009.07.002`.

[13] A. Alexov et al. “Status of LOFAR Data in HDF5 Format”. In: *Astronomical Data Analysis Software and Systems XXI.* Ed. by P. Ballester, D. Egret, and N. P. F. Lorente. Vol. 461. Astronomical Society of the Pacific Conference Series. Sept. 2012, p. 283.

[14] Steven W. Allen, August E. Evrard, and Adam B. Mantz. “Cosmological Parameters from Observations of Galaxy Clusters”. In: *Annual Review of Astronomy and Astrophysics* 49.Volume 49, 2011 (2011), pp. 409–470. ISSN: 1545-4282. DOI: `https://doi.org/10.1146/annurev-astro-081710-102514`. URL: `https://www.annualreviews.org/content/journals/10.1146/annurev-astro-081710-102514`.

[15] Lauren Anderson et al. “The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Releases 10 and 11 Galaxy samples”. In: *Monthly Notices of the Royal Astronomical Society* 441.1 (Apr. 2014), pp. 24–62. ISSN: 0035-8711. DOI: `10.1093/mnras/stu523`. URL: `http://dx.doi.org/10.1093/mnras/stu523`.

[16] Giovanni Aricò et al. “Modelling the large-scale mass density field of the universe as a function of cosmology and baryonic physics”. In: *Monthly Notices of the Royal Astronomical Society* 495.4 (May 2020), pp. 4800–4819. ISSN: 1365-2966. DOI: `10.1093/mnras/staa1478`. URL: `http://dx.doi.org/10.1093/mnras/staa1478`.

[17] Nima Arkani-Hamed and Juan Maldacena. *Cosmological Collider Physics*. 2015. arXiv: `1503.08043 [hep-th]`. URL: `https://arxiv.org/abs/1503.08043`.

[18] James M. Bardeen. "Gauge-invariant cosmological perturbations". In: *Phys. Rev. D* 22 (8 Oct. 1980), pp. 1882–1905. DOI: `10.1103/PhysRevD.22.1882`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.22.1882`.

[19] Peter Behroozi et al. "UniverseMachine: The correlation between galaxy growth and dark matter halo assembly from z = 0-10". In: *Monthly Notices of the Royal Astronomical Society* 488.3 (May 2019), pp. 3143–3194. ISSN: 0035-8711. DOI: `10.1093/mnras/stz1182`. eprint: `https://academic.oup.com/mnras/article-pdf/488/3/3143/29016136/stz1182.pdf`. URL: `https://doi.org/10.1093/mnras/stz1182`.

[20] Peter S. Behroozi, Risa H. Wechsler, and Hao-Yi Wu. "The RockStar phase-space temporal halo finder and the velocity offsets of cluster cores". In: *The Astrophysical Journal* 762.2 (Dec. 2012), p. 109. ISSN: 1538-4357. DOI: `10.1088/0004-637x/762/2/109`. URL: `http://dx.doi.org/10.1088/0004-637X/762/2/109`.

[21] Peter S. Behroozi et al. "Gravitationally Consistent Halo Catalogs and Merger Trees for Precision Cosmology". In: *Astrophysical Journal* 763.1, 18 (Jan. 2013), p. 18. DOI: `10.1088/0004-637X/763/1/18`. arXiv: `1110.4370 [astro-ph.CO]`.

[22] C. L. Bennett et al. "Four-Year COBE DMR Cosmic Microwave Background Observations: Maps and Basic Results". In: *The Astrophysical Journal* 464.1 (June 1996), pp. L1–L4. ISSN: 0004-637X. DOI: `10.1086/310075`. URL: `http://dx.doi.org/10.1086/310075`.

[23] B. A. Benson et al. "SPT-3G: a next-generation cosmic microwave background polarization experiment on the South Pole telescope". In: *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII.* Ed. by Wayne S. Holland and Jonas Zmuidzinas. Vol. 9153. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. July 2014, 91531P, 91531P. DOI: `10.1117/12.2057305`. arXiv: `1407.2973 [astro-ph.IM]`.

[24] M. Betoule et al. "Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples". In: *Astronomy & Astrophysics* 568, A22 (Aug. 2014), A22. DOI: `10.1051/0004-6361/201423413`. arXiv: `1401.4064 [astro-ph.CO]`.

[25] M Birkinshaw. "The Sunyaev–Zel'dovich effect". In: *Physics Reports* 310.2–3 (Mar. 1999), pp. 97–195. ISSN: 0370-1573. DOI: `10.1016/s0370-1573(98)00080-5`. URL: `http://dx.doi.org/10.1016/S0370-1573(98)00080-5`.

[26] Lorenzo Bordin and Giovanni Cabass. "Graviton non-Gaussianities and parity violation in the EFT of inflation". In: *Journal of Cosmology and Astroparticle Physics* 2020.07 (July 2020), pp. 014–014. ISSN: 1475-7516. DOI: `10.1088/1475-7516/2020/07/014`. URL: `http://dx.doi.org/10.1088/1475-7516/2020/07/014`.

[27] Thomas Buchert. “Lagrangian theory of gravitational instability of Friedman–Lemaître cosmologies – a generic third-order model for non-linear clustering”. In: *Monthly Notices of the Royal Astronomical Society* 267.4 (Apr. 1994), pp. 811–820. ISSN: 0035-8711. DOI: `10.1093/mnras/267.4.811`. eprint: `https://academic.oup.com/mnras/article-pdf/267/4/811/4056911/mnras267-0811.pdf`. URL: `https://doi.org/10.1093/mnras/267.4.811`.

[28] Robert N. Cahn, Zachary Slepian, and Jiamin Hou. *A Test for Cosmological Parity Violation Using the 3D Distribution of Galaxies.* 2021. arXiv: `2110.12004 [astro-ph.CO]`. URL: `https://arxiv.org/abs/2110.12004`.

[29] Paolo Catelan. “Lagrangian dynamics in non-flat universes and non-linear gravitational evolution”. In: *Monthly Notices of the Royal Astronomical Society* 276.1 (Sept. 1995), pp. 115–124. ISSN: 0035-8711. DOI: `10.1093/mnras/276.1.115`. eprint: `https://academic.oup.com/mnras/article-pdf/276/1/115/2969897/mnras276-0115.pdf`. URL: `https://doi.org/10.1093/mnras/276.1.115`.

[30] Anthony Challinor and Antony Lewis. *CAMB Sources: Number Counts, Lensing & Dark-age 21cm Power Spectra.* Astrophysics Source Code Library, record ascl:1105.013. May 2011.

[31] Anthony Challinor and Antony Lewis. “Linear power spectrum of observed source number counts”. In: *Physical Review D* 84.4 (Aug. 2011). ISSN: 1550-2368. DOI: `10.1103/physrevd.84.043516`. URL: `http://dx.doi.org/10.1103/PhysRevD.84.043516`.

[32] Jonás Chaves-Montero, Raúl E Angulo, and Carlos Hernández-Monteagudo. “The effect of photometric redshift uncertainties on galaxy clustering and baryonic acoustic oscillations”. In: *Monthly Notices of the Royal Astronomical Society* 477.3 (Apr. 2018), pp. 3892–3909. ISSN: 0035-8711. DOI: `10.1093/mnras/sty924`. eprint: `https://academic.oup.com/mnras/article-pdf/477/3/3892/24812692/sty924.pdf`. URL: `https://doi.org/10.1093/mnras/sty924`.

[33] Jonás Chaves-Montero et al. “Measuring the evolution of intergalactic gas from z = 0 to 5 using the kinematic Sunyaev–Zel’dovich effect”. In: *Monthly Notices of the Royal Astronomical Society* 503.2 (Mar. 2020), pp. 1798–1814. ISSN: 1365-2966. DOI: `10.1093/mnras/staa3782`. URL: `http://dx.doi.org/10.1093/mnras/staa3782`.

[34] Michel Chevallier and David Polarski. “Accelerating Universes with Scaling Dark Matter”. In: *International Journal of Modern Physics D* 10.2 (Jan. 2001), pp. 213–223. DOI: `10.1142/S0218271801000822`. arXiv: `gr-qc/0009008 [gr-qc]`.

[35] DESI Collaboration et al. *DESI 2024 VI: Cosmological Constraints from the Measurements of Baryon Acoustic Oscillations.* 2024. arXiv: `2404.03002 [astro-ph.CO]`. URL: `https://arxiv.org/abs/2404.03002`.

[36] S Contreras, R E Angulo, and M Zennaro. “A flexible subhalo abundance matching model for galaxy clustering in redshift space”. In: *Monthly Notices of the Royal Astronomical Society* 508.1 (Sept. 2021), pp. 175–189. ISSN: 0035-8711. DOI: `10.1093/mnras/stab2560`. eprint: `https://academic.oup.com/mnras/article-pdf/508/1/175/40430519/stab2560.pdf`. URL: `https://doi.org/10.1093/mnras/stab2560`.

[37] Jessica L. Cook and Lorenzo Sorbo. “Particle production during inflation and gravitational waves detectable by ground-based interferometers”. In: *Physical Review D* 85.2 (Jan. 2012). ISSN: 1550-2368. DOI: `10.1103/physrevd.85.023534`. URL: `http://dx.doi.org/10.1103/PhysRevD.85.023534`.

[38] William R. Coulton, Oliver H. E. Philcox, and Francisco Villaescusa-Navarro. *Signatures of a Parity-Violating Universe*. 2023. arXiv: `2306.11782 [astro-ph.CO]`.

[39] William R. Coulton and David N. Spergel. “The bispectrum of polarized galactic foregrounds”. In: *Journal of Cosmology and Astroparticle Physics* 2019.10 (Oct. 2019), pp. 056–056. ISSN: 1475-7516. DOI: `10.1088/1475-7516/2019/10/056`. URL: `http://dx.doi.org/10.1088/1475-7516/2019/10/056`.

[40] Cyril Creque-Sarbinowski et al. *Parity-Violating Trispectrum from Chern-Simons Gravity*. 2023. arXiv: `2303.04815 [astro-ph.CO]`.

[41] Darren J. Croton et al. “The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies”. In: *Monthly Notices of the Royal Astronomical Society* 365.1 (Jan. 2006), pp. 11–28. ISSN: 0035-8711. DOI: `10.1111/j.1365-2966.2005.09675.x`. eprint: `https://academic.oup.com/mnras/article-pdf/365/1/11/5530258/365-1-11.pdf`. URL: `https://doi.org/10.1111/j.1365-2966.2005.09675.x`.

[42] P. Diego-Palazuelos et al. “Cosmic Birefringence from the Planck Data Release 4”. In: *Physical Review Letters* 128.9 (Mar. 2022). ISSN: 1079-7114. DOI: `10.1103/physrevlett.128.091302`. URL: `http://dx.doi.org/10.1103/PhysRevLett.128.091302`.

[43] Scott Dodelson and Fabian Schmidt. *Modern Cosmology*. Academic Press, 2020. DOI: `10.1016/C2017-0-01943-2`.

[44] A. G. Doroshkevich et al. “Gauss-Legendre Sky Pixelization (GLESP) for CMB maps”. In: *International Journal of Modern Physics D* 14.02 (Feb. 2005), pp. 275–290. ISSN: 1793-6594. DOI: `10.1142/s0218271805006183`. URL: `http://dx.doi.org/10.1142/S0218271805006183`.

[45] Marco Drewes. “The Phenomenology of Right Handed Neutrinos”. In: *International Journal of Modern Physics E* 22.08 (Aug. 2013), p. 1330019. ISSN: 1793-6608. DOI: `10.1142/s0218301313300191`.

[46] Cora Dvorkin et al. *Dark Matter Physics from the CMB-S4 Experiment.* 2022. arXiv: `2203.07064 [hep-ph]`. URL: `https://arxiv.org/abs/2203.07064`.

[47] Daniel J. Eisenstein and Wayne Hu. "Power Spectra for Cold Dark Matter and Its Variants". In: *The Astrophysical Journal* 511.1 (Jan. 1999), pp. 5–15. ISSN: 1538-4357. DOI: `10.1086/306640`. URL: `http://dx.doi.org/10.1086/306640`.

[48] Daniel J. Eisenstein et al. "Spectroscopic Target Selection for the Sloan Digital Sky Survey: The Luminous Red Galaxy Sample". In: *The Astronomical Journal* 122.5 (Nov. 2001), pp. 2267–2280. ISSN: 0004-6256. DOI: `10.1086/323717`. URL: `http://dx.doi.org/10.1086/323717`.

[49] Daniel J. Eisenstein et al. "Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies". In: *Astrophysical Journal* 633.2 (Nov. 2005), pp. 560–574. DOI: `10.1086/466512`. arXiv: `astro-ph/0501171 [astro-ph]`.

[50] Daniel J. Eisenstein et al. "SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way, and Extra-solar Planetary Systems". In: *The Astronomical Journal* 142.3 (Aug. 2011), p. 72. DOI: `10.1088/0004-6256/142/3/72`. URL: `https://dx.doi.org/10.1088/0004-6256/142/3/72`.

[51] Johannes R. Eskilt and Eiichiro Komatsu. "Improved constraints on cosmic birefringence from the WMAP and Planck cosmic microwave background polarization data". In: *Physical Review D* 106.6 (Sept. 2022). ISSN: 2470-0029. DOI: `10.1103/physrevd.106.063503`. URL: `http://dx.doi.org/10.1103/PhysRevD.106.063503`.

[52] Hume A. Feldman, Nick Kaiser, and John A. Peacock. "Power-spectrum analysis of three-dimensional redshift surveys". In: *The Astrophysical Journal* 426 (May 1994), p. 23. ISSN: 1538-4357. DOI: `10.1086/174036`. URL: `http://dx.doi.org/10.1086/174036`.

[53] Bo Feng et al. "Searching for CPT Violation with Cosmic Microwave Background Data from WMAP and BOOMERANG". In: *Physical Review Letters* 96.22 (June 2006). ISSN: 1079-7114. DOI: `10.1103/physrevlett.96.221302`. URL: `http://dx.doi.org/10.1103/PhysRevLett.96.221302`.

[54] J.R. Fergusson. "Efficient optimal non-Gaussian CMB estimators with polarization". In: *Physical Review D* 90.4 (Aug. 2014). ISSN: 1550-2368. DOI: `10.1103/physrevd.90.043533`. URL: `http://dx.doi.org/10.1103/PhysRevD.90.043533`.

[55] Pierre Fernique et al. *MOC - HEALPix Multi-Order Coverage map Version 1.0.* IVOA Recommendation 02 June 2014. June 2014. DOI: `10.5479/ADS/bib/2014ivoa.spec.0602F`. arXiv: `1505.02937 [astro-ph.IM]`.

[56] D. J. Fixsen et al. “The Cosmic Microwave Background Spectrum from the FullCOBEFIRAS Data Set”. In: *The Astrophysical Journal* 473.2 (Dec. 1996), pp. 576–587. ISSN: 1538-4357. DOI: `10.1086/178173`. URL: `http://dx.doi.org/10.1086/178173`.

[57] C.S. Frenk and S.D.M. White. “Dark matter and cosmic structure”. In: *Annalen der Physik* 524.9–10 (Sept. 2012), pp. 507–534. ISSN: 1521-3889. DOI: `10.1002/andp.201200212`. URL: `http://dx.doi.org/10.1002/andp.201200212`.

[58] J. N. Fry and P. J. E. Peebles. “Statistical analysis of catalogs of extragalactic objects. IX. The four-point galaxy correlation function.” In: *Astrophysical Journal* 221 (Apr. 1978), pp. 19–33. DOI: `10.1086/156001`.

[59] M. Fukugita et al. “The Sloan Digital Sky Survey Photometric System”. In: *Astronomical Journal* 111 (Apr. 1996), p. 1748. DOI: `10.1086/117915`.

[60] K. M. Gorski et al. “HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere”. In: *The Astrophysical Journal* 622.2 (Apr. 2005), pp. 759–771. ISSN: 1538-4357. DOI: `10.1086/427976`. URL: `http://dx.doi.org/10.1086/427976`.

[61] Krzysztof M. Gorski et al. *The HEALPix Primer*. 1999. arXiv: `astro-ph/9905275 [astro-ph]`. URL: `https://arxiv.org/abs/astro-ph/9905275`.

[62] Dario Grasso and Hector R. Rubinstein. “Magnetic fields in the early Universe”. In: *Physics Reports* 348.3 (July 2001), pp. 163–266. ISSN: 0370-1573. DOI: `10.1016/s0370-1573(00)00110-1`. URL: `http://dx.doi.org/10.1016/S0370-1573(00)00110-1`.

[63] H. F. Gruetjen et al. “Using inpainting to construct accurate cut-sky CMB estimators”. In: *Phys. Rev. D* 95 (4 Feb. 2017), p. 043532. DOI: `10.1103/PhysRevD.95.043532`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.95.043532`.

[64] James E. Gunn et al. “The 2.5 m Telescope of the Sloan Digital Sky Survey”. In: *Astronomical Journal* 131.4 (Apr. 2006), pp. 2332–2359. DOI: `10.1086/500975`. arXiv: `astro-ph/0602326 [astro-ph]`.

[65] Alan H. Guth. *Inflation*. 2004. arXiv: `astro-ph/0404546 [astro-ph]`. URL: `https://arxiv.org/abs/astro-ph/0404546`.

[66] Balakrishna S. Haridasu et al. “Strong evidence for an accelerating Universe”. In: *Astronomy & Astrophysics* 600 (Mar. 2017), p. L1. ISSN: 1432-0746. DOI: `10.1051/0004-6361/201730469`. URL: `http://dx.doi.org/10.1051/0004-6361/201730469`.

[67] E. R. Harrison. “Fluctuations at the Threshold of Classical Cosmology”. In: *Phys. Rev. D* 1 (10 May 1970), pp. 2726–2730. DOI: `10.1103/PhysRevD.1.2726`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.1.2726`.

[68] W. K. Hastings. "Monte Carlo sampling methods using Markov chains and their applications". In: *Biometrika* 57.1 (Apr. 1970), pp. 97–109. ISSN: 0006-3444. DOI: `10.1093/biomet/57.1.97`. eprint: `https://academic.oup.com/biomet/article-pdf/57/1/97/23940249/57-1-97.pdf`. URL: `https://doi.org/10.1093/biomet/57.1.97`.

[69] C. Hernández-Monteagudo. "Angular Redshift Fluctuations. A New Cosmological Observable". In: *Highlights on Spanish Astrophysics X*. Ed. by B. Montesinos et al. Mar. 2019, pp. 134–139.

[70] Carlos Hernández–Monteagudo, Jonás Chaves-Montero, and Raúl E Angulo. "Density weighted angular redshift fluctuations: a new cosmological observable". In: *Monthly Notices of the Royal Astronomical Society: Letters* 503.1 (Mar. 2020), pp. L56–L61. ISSN: 1745-3933. DOI: `10.1093/mnrasl/slaa172`. URL: `http://dx.doi.org/10.1093/mnrasl/slaa172`.

[71] Carlos Hernández–Monteagudo et al. "Tomographic constraints on gravity from angular redshift fluctuations in the late Universe". In: *Monthly Notices of the Royal Astronomical Society: Letters* 503.1 (Mar. 2020), pp. L62–L66. ISSN: 1745-3933. DOI: `10.1093/mnrasl/slab021`. URL: `http://dx.doi.org/10.1093/mnrasl/slab021`.

[72] Samuel Hewson, Will J. Handley, and Christopher G. Lester. *On the spatial distribution of the Large-Scale structure: An Unsupervised search for Parity Violation*. 2024. arXiv: `2410.16030 [astro-ph.CO]`. URL: `https://arxiv.org/abs/2410.16030`.

[73] G. Hinshaw et al. "Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results". In: *Astrophysical Journal Supplement* 208.2, 19 (Oct. 2013), p. 19. DOI: `10.1088/0067-0049/208/2/19`. arXiv: `1212.5226 [astro-ph.CO]`.

[74] Eric Hivon et al. "MASTER of the Cosmic Microwave Background Anisotropy Power Spectrum: A Fast Method for Statistical Analysis of Large and Complex Cosmic Microwave Background Data Sets". In: *The Astrophysical Journal* 567.1 (Mar. 2002), p. 2. DOI: `10.1086/338126`. URL: `https://dx.doi.org/10.1086/338126`.

[75] David W. Hogg and Daniel Foreman-Mackey. "Data Analysis Recipes: Using Markov Chain Monte Carlo*". In: *The Astrophysical Journal Supplement Series* 236.1 (May 2018), p. 11. DOI: `10.3847/1538-4365/aab76e`. URL: `https://dx.doi.org/10.3847/1538-4365/aab76e`.

[76] Jiamin Hou, Zachary Slepian, and Robert N Cahn. “Measurement of parity-odd modes in the large-scale 4-point correlation function of Sloan Digital Sky Survey Baryon Oscillation Spectroscopic Survey twelfth data release CMASS and LOWZ galaxies”. In: *Monthly Notices of the Royal Astronomical Society* 522.4 (May 2023), pp. 5701–5739. ISSN: 0035-8711. DOI: `10.1093/mnras/stad1062`. eprint: `https://academic.oup.com/mnras/article-pdf/522/4/5701/50582849/stad1062.pdf`. URL: `https://doi.org/10.1093/mnras/stad1062`.

[77] Donghui Jeong and Marc Kamionkowski. “Clustering Fossils from the Early Universe”. In: *Physical Review Letters* 108.25 (June 2012). ISSN: 1079-7114. DOI: `10.1103/physrevlett.108.251301`. URL: `http://dx.doi.org/10.1103/PhysRevLett.108.251301`.

[78] Raul Jimenez and Abraham Loeb. “Constraining Cosmological Parameters Based on Relative Galaxy Ages”. In: *Astrophysical Journal* 573.1 (July 2002), pp. 37–42. DOI: `10.1086/340549`. arXiv: `astro-ph/0106145 [astro-ph]`.

[79] Nick Kaiser. “Clustering in real space and in redshift space”. In: *Monthly Notices of the Royal Astronomical Society* 227.1 (July 1987), pp. 1–21. ISSN: 0035-8711. DOI: `10.1093/mnras/227.1.1`. eprint: `https://academic.oup.com/mnras/article-pdf/227/1/1/18522208/mnras227-0001.pdf`. URL: `https://doi.org/10.1093/mnras/227.1.1`.

[80] F. -S. Kitaura, G. Yepes, and F. Prada. “Modelling baryon acoustic oscillations with perturbation theory and stochastic halo biasing.” In: *Monthly Notices of the Royal Astronomical Society* 439 (Mar. 2014), pp. L21–L25. DOI: `10.1093/mnrasl/slt172`. arXiv: `1307.3285 [astro-ph.CO]`.

[81] Anatoly Klypin et al. “MultiDark simulations: the story of dark matter halo concentrations and density profiles”. In: *Monthly Notices of the Royal Astronomical Society* 457.4 (Apr. 2016), pp. 4340–4359. DOI: `10.1093/mnras/stw248`. arXiv: `1411.4001 [astro-ph.CO]`.

[82] Hideo Kodama and Misao Sasaki. “Cosmological Perturbation Theory”. In: *Progress of Theoretical Physics Supplement* 78 (Jan. 1984), pp. 1–166. ISSN: 0375-9687. DOI: `10.1143/PTPS.78.1`. eprint: `https://academic.oup.com/ptps/article-pdf/doi/10.1143/PTPS.78.1/5321391/78-1.pdf`. URL: `https://doi.org/10.1143/PTPS.78.1`.

[83] E. Komatsu, D. N. Spergel, and B. D. Wandelt. “Measuring Primordial Non-Gaussianity in the Cosmic Microwave Background”. In: *The Astrophysical Journal* 634.1 (Nov. 2005), pp. 14–19. ISSN: 1538-4357. DOI: `10.1086/491724`. URL: `http://dx.doi.org/10.1086/491724`.

[84] E. Komatsu et al. “Measurement of the Cosmic Microwave Background Bispectrum on theCOBEDMR Sky Maps”. In: *The Astrophysical Journal* 566.1 (Feb. 2002), pp. 19–29. ISSN: 1538-4357. DOI: `10.1086/337963`. URL: `http://dx.doi.org/10.1086/337963`.

[85] E. Komatsu et al. “Five-Year Wilkinson Microwave Anisotropy Probe Observations: Cosmological Interpretation”. In: *Astrophysical Journal Supplement* 180.2 (Feb. 2009), pp. 330–376. DOI: `10.1088/0067-0049/180/2/330`. arXiv: `0803.0547 [astro-ph]`.

[86] E. Komatsu et al. “Five-year Wilkinson Microwave Anisotropy Probe observations: cosmological interpretation”. In: *The Astrophysical Journal Supplement Series* 180.2 (Feb. 2009), pp. 330–376. ISSN: 1538-4365. DOI: `10.1088/0067-0049/180/2/330`. URL: `http://dx.doi.org/10.1088/0067-0049/180/2/330`.

[87] E. Komatsu et al. “Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation”. In: *Astrophysical Journal Supplement* 192.2, 18 (Feb. 2011), p. 18. DOI: `10.1088/0067-0049/192/2/18`. arXiv: `1001.4538 [astro-ph.CO]`.

[88] Eiichiro Komatsu. “New physics from the polarized light of the cosmic microwave background”. In: *Nature Reviews Physics* 4.7 (July 2022), pp. 452–469. DOI: `10.1038/s42254-022-00452-4`. arXiv: `2202.13919 [astro-ph.CO]`.

[89] Elisabeth Krause et al. “The weight of emptiness: the gravitational lensing signal of stacked voids”. In: *The Astrophysical Journal Letters* 762.2 (Dec. 2012), p. L20. DOI: `10.1088/2041-8205/762/2/L20`. URL: `https://dx.doi.org/10.1088/2041-8205/762/2/L20`.

[90] Alex Krolewski et al. *No evidence for parity violation in BOSS*. 2024. arXiv: `2407.03397 [astro-ph.CO]`. URL: `https://arxiv.org/abs/2407.03397`.

[91] Stephen D. Landy and Alexander S. Szalay. “Bias and Variance of Angular Correlation Functions”. In: *Astrophysical Journal* 412 (July 1993), p. 64. DOI: `10.1086/172900`.

[92] T. D. Lee and C. N. Yang. “Question of Parity Conservation in Weak Interactions”. In: *Phys. Rev.* 104 (1 Oct. 1956), pp. 254–258. DOI: `10.1103/PhysRev.104.254`. URL: `https://link.aps.org/doi/10.1103/PhysRev.104.254`.

[93] L. Legrand et al. “High-resolution tomography for galaxy spectroscopic surveys with angular redshift fluctuations”. In: *Astronomy & Astrophysics* 646 (Feb. 2021), A109. ISSN: 1432-0746. DOI: `10.1051/0004-6361/202039049`. URL: `http://dx.doi.org/10.1051/0004-6361/202039049`.

[94] Margherita Lembo et al. "CMB lensing reconstruction biases from masking extragalactic sources". In: *Physical Review D* 106.2 (July 2022). ISSN: 2470-0029. DOI: `10.1103/physrevd.106.023525`. URL: `http://dx.doi.org/10.1103/PhysRevD.106.023525`.

[95] Julien Lesgourgues. *The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview.* 2011. arXiv: `1104.2932 [astro-ph.IM]`. URL: `https://arxiv.org/abs/1104.2932`.

[96] Antony Lewis, Anthony Challinor, and Anthony Lasenby. "Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models". In: *The Astrophysical Journal* 538.2 (Aug. 2000), pp. 473–476. ISSN: 1538-4357. DOI: `10.1086/309179`. URL: `http://dx.doi.org/10.1086/309179`.

[97] Adal Lima-Hernández, Carlos Hernández-Monteagudo, and Jonás Chaves-Montero. "Relativistic angular redshift fluctuations embedded in large scale varying gravitational potentials". In: *Journal of Cosmology and Astroparticle Physics* 2022.09 (Sept. 2022), p. 038. ISSN: 1475-7516. DOI: `10.1088/1475-7516/2022/09/038`. URL: `http://dx.doi.org/10.1088/1475-7516/2022/09/038`.

[98] Eric V. Linder. "Exploring the Expansion History of the Universe". In: *Physical Review Letters* 90.9 (Mar. 2003). ISSN: 1079-7114. DOI: `10.1103/physrevlett.90.091301`. URL: `http://dx.doi.org/10.1103/PhysRevLett.90.091301`.

[99] Arthur Loureiro et al. "Cosmological measurements from angular power spectra analysis of BOSS DR12 tomography". In: *Monthly Notices of the Royal Astronomical Society* 485.1 (Jan. 2019), pp. 326–355. ISSN: 1365-2966. DOI: `10.1093/mnras/stz191`. URL: `http://dx.doi.org/10.1093/mnras/stz191`.

[100] Mathew S. Madhavacheril et al. "CMB lensing power spectrum estimation without instrument noise bias". In: *Journal of Cosmology and Astroparticle Physics* 2021.05 (May 2021), p. 028. DOI: `10.1088/1475-7516/2021/05/028`. URL: `https://dx.doi.org/10.1088/1475-7516/2021/05/028`.

[101] Claudia Maraston et al. "Modelling the colour evolution of luminous red galaxies – improvements with empirical stellar spectra". In: *Monthly Notices of the Royal Astronomical Society: Letters* 394.1 (Mar. 2009), pp. L107–L111. ISSN: 1745-3925. DOI: `10.1111/j.1745-3933.2009.00621.x`. eprint: `https://academic.oup.com/mnrasl/article-pdf/394/1/L107/54679761/mnrasl\_394\_1\_l107.pdf`. URL: `https://doi.org/10.1111/j.1745-3933.2009.00621.x`.

[102] F. Marulli, A. Veropalumbo, and M. Moresco. "CosmoBolognaLib: C++ libraries for cosmological calculations". In: *Astronomy and Computing* 14 (Jan. 2016), pp. 35–42. ISSN: 2213-1337. DOI: `10.1016/j.ascom.2016.01.005`. URL: `http://dx.doi.org/10.1016/j.ascom.2016.01.005`.

[103] F. Marulli et al. "The XXL Survey: XVI. The clustering of X-ray selected galaxy clusters atz 0.3". In: *Astronomy & Astrophysics* 620 (Nov. 2018), A1. ISSN: 1432-0746. DOI: `10.1051/0004-6361/201833238`. URL: `http://dx.doi.org/10.1051/0004-6361/201833238`.

[104] Federico Marulli et al. "Cosmology with clustering anisotropies: disentangling dynamic and geometric distortions in galaxy redshift surveys: Disentangling dynamics and geometry". In: *Monthly Notices of the Royal Astronomical Society* 426.3 (Oct. 2012), pp. 2566–2580. ISSN: 0035-8711. DOI: `10.1111/j.1365-2966.2012.21875.x`. URL: `http://dx.doi.org/10.1111/j.1365-2966.2012.21875.x`.

[105] Kareem Marzouk, Antony Lewis, and Julien Carron. "Constraints on τ_{NL} from Planck temperature and polarization". In: *Journal of Cosmology and Astroparticle Physics* 2022.08 (Aug. 2022), p. 015. ISSN: 1475-7516. DOI: `10.1088/1475-7516/2022/08/015`. URL: `http://dx.doi.org/10.1088/1475-7516/2022/08/015`.

[106] Nicholas Metropolis et al. "Equation of State Calculations by Fast Computing Machines". In: *The Journal of Chemical Physics* 21.6 (June 1953), pp. 1087–1092. ISSN: 0021-9606. DOI: `10.1063/1.1699114`. eprint: `https://pubs.aip.org/aip/jcp/article-pdf/21/6/1087/18802390/1087\_1\_online.pdf`. URL: `https://doi.org/10.1063/1.1699114`.

[107] Michele Moresco et al. "Unveiling the Universe with emerging cosmological probes". In: *Living Reviews in Relativity* 25.1 (2022), p. 6. ISSN: 1433-8351. DOI: `10.1007/s41114-022-00040-z`. URL: `https://doi.org/10.1007/s41114-022-00040-z`.

[108] P. F. Muciaccia, P. Natoli, and N. Vittorio. "Fast Spherical Harmonic Analysis: A Quick Algorithm for Generating and/or InvertingFull-Sky, High-Resolution Cosmic Microwave Background Anisotropy Maps". In: *The Astrophysical Journal* 488.2 (Oct. 1997), pp. L63–L66. ISSN: 0004-637X. DOI: `10.1086/310921`. URL: `http://dx.doi.org/10.1086/310921`.

[109] Seshadri Nadathur. "Seeing patterns in noise: gigaparsec-scale 'structures' that do not violate homogeneity". In: *Monthly Notices of the Royal Astronomical Society* 434.1 (July 2013), pp. 398–406. ISSN: 1365-2966. DOI: `10.1093/mnras/stt1028`. URL: `http://dx.doi.org/10.1093/mnras/stt1028`.

[110] Mark C. Neyrinck. "Quantifying distortions of the Lagrangian dark-matter mesh in cosmology". In: *Monthly Notices of the Royal Astronomical Society* 428.1 (Oct. 2012), pp. 141–153. ISSN: 0035-8711. DOI: `10.1093/mnras/sts027`. eprint: `https://academic.oup.com/mnras/article-pdf/428/1/141/3551359/141.pdf`. URL: `https://doi.org/10.1093/mnras/sts027`.

[111] M. D. Niemack et al. “ACTPol: a polarization-sensitive receiver for the Atacama Cosmology Telescope”. In: *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy V.* Ed. by Wayne S. Holland and Jonas Zmuidzinas. Vol. 7741. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. July 2010, 77411S, 77411S. DOI: `10.1117/12.857464`. arXiv: `1006.5049 [astro-ph.IM]`.

[112] Maria Okounkova et al. “Constraining gravitational wave amplitude birefringence and Chern-Simons gravity with GWTC-2”. In: *Physical Review D* 106.4 (Aug. 2022). ISSN: 2470-0029. DOI: `10.1103/physrevd.106.044067`. URL: `http://dx.doi.org/10.1103/PhysRevD.106.044067`.

[113] Bob Osano. *Evolution of Cosmological Total Energy Density and Transient Periods in Cosmology.* 2024. arXiv: `2002.08875 [gr-qc]`. URL: `https://arxiv.org/abs/2002.08875`.

[114] Pritha Paul, Chris Clarkson, and Roy Maartens. *The Odd-Parity Part of the Observed Galaxy Trispectrum.* 2024. arXiv: `2411.10897 [astro-ph.CO]`. URL: `https://arxiv.org/abs/2411.10897`.

[115] P. J. E. Peebles. *Principles of Physical Cosmology.* Princeton University Press, Sept. 2020. ISBN: 978-0-691-20981-4.

[116] P. J. E. Peebles and M. G. Hauser. “Statistical Analysis of Catalogs of Extragalactic Objects. III. The Shane-Wirtanen and Zwicky Catalogs”. In: *Astrophysical Journal Supplement* 28 (Nov. 1974), p. 19. DOI: `10.1086/190308`.

[117] P. J. E. Peebles and J. T. Yu. “Primeval Adiabatic Perturbation in an Expanding Universe”. In: *Astrophysical Journal* 162 (Dec. 1970), p. 815. DOI: `10.1086/150713`.

[118] S. Perlmutter et al. “Measurements of Ω and Λ from 42 High-Redshift Supernovae”. In: *Astrophysical Journal* 517.2 (June 1999), pp. 565–586. DOI: `10.1086/307221`. arXiv: `astro-ph/9812133 [astro-ph]`.

[119] Oliver H. E. Philcox. “Cosmology without window functions. II. Cubic estimators for the galaxy bispectrum”. In: *Phys. Rev. D* 104 (12 Dec. 2021), p. 123529. DOI: `10.1103/PhysRevD.104.123529`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.104.123529`.

[120] Oliver H. E. Philcox. “Probing parity violation with the four-point correlation function of BOSS galaxies”. In: *Physical Review D* 106.6 (Sept. 2022). ISSN: 2470-0029. DOI: `10.1103/physrevd.106.063501`. URL: `http://dx.doi.org/10.1103/PhysRevD.106.063501`.

[121] Oliver H. E. Philcox. *Do the CMB Temperature Fluctuations Conserve Parity?* 2023. arXiv: `2303.12106 [astro-ph.CO]`. URL: `https://arxiv.org/abs/2303.12106`.

[122] Oliver H. E. Philcox. “Optimal estimation of the binned mask-free power spectrum, bispectrum, and trispectrum on the full sky: Scalar edition”. In: *Physical Review D* 107.12 (June 2023). ISSN: 2470-0029. DOI: `10.1103/physrevd.107.123516`. URL: `http://dx.doi.org/10.1103/PhysRevD.107.123516`.

[123] Oliver H. E. Philcox and Julia Ereza. *Could Sample Variance be Responsible for the Parity-Violating Signal Seen in the BOSS Galaxy Survey?* 2024. arXiv: `2401.09523 [astro-ph.CO]`. URL: `https://arxiv.org/abs/2401.09523`.

[124] Oliver H. E. Philcox and Maresuke Shiraishi. *Testing Parity Symmetry with the Polarized Cosmic Microwave Background.* 2024. arXiv: `2308.03831 [astro-ph.CO]`. URL: `https://arxiv.org/abs/2308.03831`.

[125] Planck Collaboration et al. “Planck 2013 results. I. Overview of products and scientific results”. In: *Astronomy & Astrophysics* 571, A1 (Nov. 2014), A1. DOI: `10.1051/0004-6361/201321529`. arXiv: `1303.5062 [astro-ph.CO]`.

[126] Levon Pogosian and Mark Wyman. “B-modes from cosmic strings”. In: *Physical Review D* 77.8 (Apr. 2008). ISSN: 1550-2368. DOI: `10.1103/physrevd.77.083509`. URL: `http://dx.doi.org/10.1103/PhysRevD.77.083509`.

[127] D. M. Regan, E. P. S. Shellard, and J. R. Fergusson. “General CMB and primordial trispectrum estimation”. In: *Phys. Rev. D* 82 (2 July 2010), p. 023520. DOI: `10.1103/PhysRevD.82.023520`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.82.023520`.

[128] Beth Reid et al. “SDSS-III Baryon Oscillation Spectroscopic Survey Data Release 12: galaxy target selection and large-scale structure catalogues”. In: *Monthly Notices of the Royal Astronomical Society* 455.2 (Nov. 2015), pp. 1553–1573. ISSN: 0035-8711. DOI: `10.1093/mnras/stv2382`. eprint: `https://academic.oup.com/mnras/article-pdf/455/2/1553/18511627/stv2382.pdf`. URL: `https://doi.org/10.1093/mnras/stv2382`.

[129] M. Reinecke and D. S. Seljebotn. “Libsharp – spherical harmonic transforms revisited”. In: *Astronomy & Astrophysics* 554 (June 2013), A112. ISSN: 1432-0746. DOI: `10.1051/0004-6361/201321494`. URL: `http://dx.doi.org/10.1051/0004-6361/201321494`.

[130] Adam G. Riess et al. “Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant”. In: *The Astronomical Journal* 116.3 (Sept. 1998), pp. 1009–1038. ISSN: 0004-6256. DOI: `10.1086/300499`. URL: `http://dx.doi.org/10.1086/300499`.

[131] Adam G. Riess et al. “Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond ΛCDM”. In: *The Astrophysical Journal* 876.1 (May 2019), p. 85. ISSN: 1538-4357. DOI: `10.3847/1538-4357/ab1422`. URL: `http://dx.doi.org/10.3847/1538-4357/ab1422`.

[132] Adam G. Riess et al. *A Comprehensive Measurement of the Local Value of the Hubble Constant with 1 km/s/Mpc Uncertainty from the Hubble Space Telescope and the SH0ES Team*. 2022. DOI: `https://doi.org/10.3847/2041-8213/ac5c5b`. arXiv: `2112.04510 [astro-ph.CO]`. URL: `https://arxiv.org/abs/2112.04510`.

[133] Sergio A. Rodrıguez-Torres et al. “The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: modelling the clustering and halo occupation distribution of BOSS CMASS galaxies in the Final Data Release”. In: *Monthly Notices of the Royal Astronomical Society* 460.2 (Aug. 2016), pp. 1173–1187. DOI: `10.1093/mnras/stw1014`. arXiv: `1509.06404 [astro-ph.CO]`.

[134] Ashley J. Ross et al. “The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: observational systematics and baryon acoustic oscillations in the correlation function”. In: *Monthly Notices of the Royal Astronomical Society* 464.1 (Sept. 2016), pp. 1168–1191. ISSN: 0035-8711. DOI: `10.1093/mnras/stw2372`. eprint: `https://academic.oup.com/mnras/article-pdf/464/1/1168/18517623/stw2372.pdf`. URL: `https://doi.org/10.1093/mnras/stw2372`.

[135] R. K. Sachs and A. M. Wolfe. “Perturbations of a Cosmological Model and Angular Variations of the Microwave Background”. In: *Astrophysical Journal* 147 (Jan. 1967), p. 73. DOI: `10.1086/148982`.

[136] A. D. Sakharov. “Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe”. In: *Pisma Zh. Eksp. Teor. Fiz.* 5 (1967), pp. 32–35. DOI: `10.1070/PU1991v034n05ABEH002497`.

[137] M. G. Santos et al. “Multiple methods for estimating the bispectrum of the cosmic microwave background with application to the MAXIMA data”. In: *Monthly Notices of the Royal Astronomical Society* 341.2 (May 2003), pp. 623–643. ISSN: 0035-8711. DOI: `10.1046/j.1365-8711.2003.06438.x`. eprint: `https://academic.oup.com/mnras/article-pdf/341/2/623/3866993/341-2-623.pdf`. URL: `https://doi.org/10.1046/j.1365-8711.2003.06438.x`.

[138] A. Savitzky and M. J. E. Golay. “Smoothing and differentiation of data by simplified least squares procedures”. In: *Analytical Chemistry* 36 (Jan. 1964), pp. 1627–1639. DOI: `10.1021/ac60214a047`.

[139] Joop Schaye et al. "The EAGLE project: simulating the evolution and assembly of galaxies and their environments". In: *Monthly Notices of the Royal Astronomical Society* 446.1 (Nov. 2014), pp. 521–554. ISSN: 0035-8711. DOI: `10.1093/mnras/stu2058`. eprint: `https://academic.oup.com/mnras/article-pdf/446/1/521/4139718/stu2058.pdf`. URL: `https://doi.org/10.1093/mnras/stu2058`.

[140] Aurel Schneider and Romain Teyssier. "A new method to quantify the effects of baryons on the matter power spectrum". In: *Journal of Cosmology and Astroparticle Physics* 2015.12 (Dec. 2015), p. 049. DOI: `10.1088/1475-7516/2015/12/049`. URL: `https://dx.doi.org/10.1088/1475-7516/2015/12/049`.

[141] Román Scoccimarro et al. "Large-scale bias and efficient generation of initial conditions for nonlocal primordial non-Gaussianity". In: *Physical Review D* 85.8 (Apr. 2012). ISSN: 1550-2368. DOI: `10.1103/physrevd.85.083002`. URL: `http://dx.doi.org/10.1103/PhysRevD.85.083002`.

[142] M. Shiraishi and S. Yokoyama. "Violation of the Rotational Invariance in the CMB Bispectrum". In: *Progress of Theoretical Physics* 126.5 (Nov. 2011), pp. 923–935. ISSN: 1347-4081. DOI: `10.1143/ptp.126.923`. URL: `http://dx.doi.org/10.1143/PTP.126.923`.

[143] Maresuke Shiraishi, Michele Liguori, and James R. Fergusson. "General parity-odd CMB bispectrum estimation". In: *Journal of Cosmology and Astroparticle Physics* 2014.05 (May 2014), pp. 008–008. ISSN: 1475-7516. DOI: `10.1088/1475-7516/2014/05/008`. URL: `http://dx.doi.org/10.1088/1475-7516/2014/05/008`.

[144] Quijote Simulations. *Data Access*. 2024. URL: `https://quijote-simulations.readthedocs.io/en/latest/access.html`.

[145] Skies and Universes. *BOSS-LRG DR12 MultiDark-Patchy mocks*. URL: `https://skiesanduniverses.org/Products/MockCatalogues/SDSS/BOSSLRGDR12MDP/`.

[146] Stephen A. Smee et al. "The Multi-object, Fiber-fed Spectrographs for the Sloan Digital Sky Survey and the Baryon Oscillation Spectroscopic Survey". In: *Astronomical Journal* 146.2, 32 (Aug. 2013), p. 32. DOI: `10.1088/0004-6256/146/2/32`. arXiv: `1208.2233 [astro-ph.IM]`.

[147] Kendrick M. Smith and Matias Zaldarriaga. "Algorithms for bispectra: forecasting, optimal analysis and simulation". In: *Monthly Notices of the Royal Astronomical Society* 417.1 (Oct. 2011), pp. 2–19. ISSN: 0035-8711. DOI: `10.1111/j.1365-2966.2010.18175.x`. eprint: `https://academic.oup.com/mnras/article-pdf/417/1/2/3015641/mnras0417-0002.pdf`. URL: `https://doi.org/10.1111/j.1365-2966.2010.18175.x`.

[148] Lee Smolin. *Lessons from Einstein's 1915 discovery of general relativity*. 2015. arXiv: `1512.07551 [physics.hist-ph]`.

[149] George F. Smoot. "COBE observations and results". In: *Conference on 3K cosmology.* ASCE, 1999, pp. 1–10. DOI: 10.1063/1.59326. URL: http://dx.doi.org/10.1063/1.59326.

[150] Rachel S. Somerville et al. "A semi-analytic model for the co-evolution of galaxies, black holes and active galactic nuclei". In: *Monthly Notices of the Royal Astronomical Society* 391.2 (Nov. 2008), pp. 481–506. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2008.13805.x. eprint: https://academic.oup.com/mnras/article-pdf/391/2/481/5764146/mnras0391-0481.pdf. URL: https://doi.org/10.1111/j.1365-2966.2008.13805.x.

[151] Joshua S. Speagle. "A Conceptual Introduction to Markov Chain Monte Carlo Methods". In: *arXiv: Other Statistics* (2019). URL: https://api.semanticscholar.org/CorpusID:203591670.

[152] D. N. Spergel et al. "First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters". In: *Astrophysical Journal Supplement* 148.1 (Sept. 2003), pp. 175–194. DOI: 10.1086/377226. arXiv: astro-ph/0302209 [astro-ph].

[153] Volker Springel. "The cosmological simulation code gadget-2". In: *Monthly Notices of the Royal Astronomical Society* 364.4 (Dec. 2005), pp. 1105–1134. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2005.09655.x. eprint: https://academic.oup.com/mnras/article-pdf/364/4/1105/18657201/364-4-1105.pdf. URL: https://doi.org/10.1111/j.1365-2966.2005.09655.x.

[154] Volker Springel et al. "First results from the IllustrisTNG simulations: matter and galaxy clustering". In: *Monthly Notices of the Royal Astronomical Society* 475.1 (Dec. 2017), pp. 676–698. ISSN: 0035-8711. DOI: 10.1093/mnras/stx3304. eprint: https://academic.oup.com/mnras/article-pdf/475/1/676/23534347/stx3304.pdf. URL: https://doi.org/10.1093/mnras/stx3304.

[155] M. E. C. Swanson et al. "Methods for rapidly processing angular masks of next-generation galaxy surveys". In: *Monthly Notices of the Royal Astronomical Society* 387.4 (July 2008), pp. 1391–1402. ISSN: 1365-2966. DOI: 10.1111/j.1365-2966.2008.13296.x. URL: http://dx.doi.org/10.1111/j.1365-2966.2008.13296.x.

[156] Atlas EDR system. *DIPC Technical Documentation.* URL: https://scc.dipc.org/docs/systems/atlas-edr/.

[157] Hyperion system. *DIPC Technical Documentation.* URL: https://scc.dipc.org/docs/systems/hyperion/overview/.

[158] István Szapudi and Alexander S. Szalay. "A New Class of Estimators for the N-Point Correlations". In: *Astrophysical Journal* 494.1 (Feb. 1998), pp. L41–L44. DOI: 10.1086/311146.

[159] Ryuichi Takahashi et al. “Revising the halofiti model for the nonlinear matter power spectrum”. In: *The Astrophysical Journal* 761.2 (Dec. 2012), p. 152. ISSN: 1538-4357. DOI: `10.1088/0004-637x/761/2/152`. URL: `http://dx.doi.org/10.1088/0004-637X/761/2/152`.

[160] Max Tegmark. “An Icosahedron-Based Method for Pixelizing the Celestial Sphere”. In: *Astrophysical Journal* 470 (Oct. 1996), p. L81. DOI: `10.1086/310310`. arXiv: `astro-ph/9610094 [astro-ph]`.

[161] Max Tegmark. “An Icosahedron-based Method for Pixelizing the Celestial Sphere”. In: *The Astrophysical Journal* 470.2 (Oct. 1996), pp. L81–L84. ISSN: 0004-637X. DOI: `10.1086/310310`. URL: `http://dx.doi.org/10.1086/310310`.

[162] Max Tegmark. “How to measure CMB power spectra without losing information”. In: *Physical Review D* 55.10 (May 1997), pp. 5895–5907. ISSN: 1089-4918. DOI: `10.1103/physrevd.55.5895`. URL: `http://dx.doi.org/10.1103/PhysRevD.55.5895`.

[163] Max Tegmark, Andy N. Taylor, and Alan F. Heavens. “Karhunen-Loeve Eigenvalue Problems in Cosmology: How Should We Tackle Large Data Sets?” In: *The Astrophysical Journal* 480.1 (May 1997), pp. 22–35. ISSN: 1538-4357. DOI: `10.1086/303939`. URL: `http://dx.doi.org/10.1086/303939`.

[164] Sergios Theodoridis and Konstantinos Koutroumbas. “Chapter 6 - Feature Generation I: Data Transformation and Dimensionality Reduction”. In: *Pattern Recognition (Fourth Edition)*. Ed. by Sergios Theodoridis and Konstantinos Koutroumbas. Fourth Edition. Boston: Academic Press, 2009, pp. 323–409. ISBN: 978-1-59749-272-0. DOI: `https://doi.org/10.1016/B978-1-59749-272-0.50008-6`. URL: `https://www.sciencedirect.com/science/article/pii/B9781597492720500086`.

[165] Jeremy Tinker et al. “Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality”. In: *The Astrophysical Journal* 688.2 (Dec. 2008), pp. 709–728. ISSN: 1538-4357. DOI: `10.1086/591439`. URL: `http://dx.doi.org/10.1086/591439`.

[166] Jeremy L. Tinker et al. “The Large Scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests”. In: *The Astrophysical Journal* 724.2 (Nov. 2010), pp. 878–886. ISSN: 1538-4357. DOI: `10.1088/0004-637x/724/2/878`. URL: `http://dx.doi.org/10.1088/0004-637X/724/2/878`.

[167] Tommaso Treu. “Strong Lensing by Galaxies”. In: *Annual Review of Astronomy and Astrophysics* 48 (Sept. 2010), pp. 87–125. DOI: `10.1146/annurev-astro-081309-130924`. arXiv: `1003.5567 [astro-ph.CO]`.

[168] Antonino Troja et al. *The Needlet CMB Trispesctrum*. 2014. arXiv: `1407.0624 [astro-ph.CO]`. URL: `https://arxiv.org/abs/1407.0624`.

[169] T.S. van Albada et al. "Distribution of dark matter in the spiral galaxy NGC 3198." In: *Astrophysical Journal* 295 (Aug. 1985), pp. 305–313. DOI: `10.1086/163375`.

[170] Francisco Villaescusa-Navarro et al. "The Quijote Simulations". In: *The Astrophysical Journal Supplement Series* 250.1 (Aug. 2020), p. 2. ISSN: 1538-4365. DOI: `10.3847/1538-4365/ab9d82`. URL: `http://dx.doi.org/10.3847/1538-4365/ab9d82`.

[171] C. S. Wu et al. "Experimental Test of Parity Conservation in Beta Decay". In: *Phys. Rev.* 105 (4 Feb. 1957), pp. 1413–1415. DOI: `10.1103/PhysRev.105.1413`. URL: `https://link.aps.org/doi/10.1103/PhysRev.105.1413`.

[172] Jun-Qing Xia. "Cosmological CPT violation and CMB polarization measurements". In: *Journal of Cosmology and Astroparticle Physics* 2012.01 (Jan. 2012), pp. 046–046. ISSN: 1475-7516. DOI: `10.1088/1475-7516/2012/01/046`. URL: `http://dx.doi.org/10.1088/1475-7516/2012/01/046`.

[173] Donald G. York et al. "The Sloan Digital Sky Survey: Technical Summary". In: *Astronomical Journal* 120.3 (Sept. 2000), pp. 1579–1587. DOI: `10.1086/301513`. arXiv: `astro-ph/0006396 [astro-ph]`.

[174] N. Zárate, R. Seaman, and D. Tody. "FITS Foreign File Encapsulation". In: *Astronomical Data Analysis Software and Systems XVI*. Ed. by R. A. Shaw, F. Hill, and D. J. Bell. Vol. 376. Astronomical Society of the Pacific Conference Series. Oct. 2007, p. 351.

[175] Ya B Zeldovich. "Hypothesis, unifying the structure and the entropy of the Universe." In: *Mon. Notic. Roy. Astron. Soc. 160: No. 1, 1P-3P(1972).* (Jan. 1972). DOI: `10.1093/mnras/160.1.1P`. URL: `https://www.osti.gov/biblio/4640375`.

[176] Zheng Zheng et al. "Theoretical Models of the Halo Occupation Distribution: Separating Central and Satellite Galaxies". In: *The Astrophysical Journal* 633.2 (Nov. 2005), p. 791. DOI: `10.1086/466510`. URL: `https://dx.doi.org/10.1086/466510`.

[177] F. Zwicky. "Republication of: The redshift of extragalactic nebulae". In: *General Relativity and Gravitation* 41.1 (Jan. 2009), pp. 207–224. DOI: `10.1007/s10714-008-0707-4`.

Appendix A

MapCalculator

Here is the class that was implemented in the `CBL`, used to produce sky maps for ADFs, shot noise and ARFs following the definition in Eq.(4.6). The MapCalculator class, whose hierarchy is depicted in Fig.5.2, comprehends also some methods to wrap as `CBL` functions some commonly used `HEALPix` routines that returns information on the pixelization process. Finally, multiple constructors have been written to increase the elasticity and facilitate the `CBL`'s user programming pipelines.

MapCalculator.h

```
/*********************************************************************
 *  Copyright (C) 2010 by Federico Marulli                           *
 *  federico.marulli3@unibo.it                                       *
 *                                                                   *
 *  This program is free software; you can redistribute it and/or    *
 *  modify it under the terms of the GNU General Public License as   *
 *  published by the Free Software Foundation; either version 2 of   *
 *  the License, or (at your option) any later version.              *
 *                                                                   *
 *  This program is distributed in the hope that it will be useful,  *
 *  but WITHOUT ANY WARRANTY; without even the implied warranty of   *
 *  MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.  See the    *
 *  GNU General Public License for more details.                     *
 *                                                                   *
```

```
 *  You should have received a copy of the GNU General Public       *
 *  License along with this program; if not, write to the Free      *
 *  Software Foundation, Inc.,                                      *
 *  59 Temple Place - Suite 330, Boston, MA  02111-1307, USA.       *
 ********************************************************************/

/**
 *  @file Headers/MapCalculator.h
 *
 *  @brief The class MapCalculator
 *
 *  This file defines the interface of the class MapCalculator, used to
 *  build the sky map for angular density fluctuations (ADF), angular
 *  redshift fluctuations (ARF) and shot noise random realisations (SN)
 *
 *  @author Matteo Santini
 *
 *  @author matteo.santini7@studio.unibo.it
 */

#ifndef __MAP__
#define __MAP__


#include "Measure.h"
#include "Catalogue.h"
#include "Func.h"
#include <healpix_base.h>


// ========================================================================


namespace cbl {

  namespace measure {

    /**
     *  @brief Class for <B> HEALPix map generation </B>
     *
     *  The \e measure::MapCalculator namespace contains all the functions
     *  and classes for building HEALPix maps for ADF and ARF
     */
```

```
class MapCalculator : public Measure {

private:

  /**
  *  @param m_ns: the resolution nside of Healpix (must be a power of 2)
  */
  int m_ns;

  /**
  *  @param m_scheme: scheme for the HEALPix map. RING (default) or NESTED
  */
  Healpix_Ordering_Scheme m_scheme = RING; //RING is default

  /**
  *  @param m_nsg: number of sigmas for source selection (speeds up the computation)
  */
  int m_nsg;

  /**
  *  @param m_zobs: central redshifts of the redshift shells
  */
  std::vector<double> m_zobs;

  /**
  *  @param m_sgzar: width of the gaussian shells
  */
  std::vector<double> m_sgzar;

  /**
  *  @param m_seed: seed for random number generation
  */
  int m_seed = 123;

protected:
  /**
  *  @param m_data: points to the catalogue of data
  */
  std::shared_ptr<catalogue::Catalogue> m_data;

public:
```

```

    /**
     *  @name Constructors/destructors
     */
    ///@{

    /**
     *  @brief default constructor
     */
    MapCalculator () = default;

    /**
     *  @brief constructor with specified parameters.
     *
     *  @param ns_value resolution nside of Healpix (must be a power of
↪  2).
     *
     *  @param nsg_value number of sigmas for source selection.
     *
     *  @param zobs central redshifts of the redshift shells.
     *
     *  @param sgzar width of the Gaussian shells.
     *
     *  @param data catalogue of data.
     */
    MapCalculator (int ns_value, int nsg_value, std::vector<double>& zobs,
↪  std::vector<double>& sgzar, const catalogue::Catalogue data)
    : m_ns(ns_value), m_nsg(nsg_value), m_zobs(zobs), m_sgzar(sgzar) ,
↪  m_data(std::make_shared<catalogue::Catalogue>
↪  (catalogue::Catalogue(std::move(data)))) {};

    /**
     *  @brief constructor with specified parameters.
     *
     *  @param ns_value resolution nside of Healpix (must be a power of
↪  2).
     *
     *  @param scheme scheme for the HEALPix map (RING or NESTED).
     *
     *  @param nsg_value number of sigmas for source selection.
     *
     *  @param zobs central redshifts of the redshift shells.
     *
```

```
     *  @param sgzar width of the Gaussian shells.
     *
     *  @param data catalogue of data.
     */
    MapCalculator (int ns_value, Healpix_Ordering_Scheme scheme, int nsg_value, std::vector<double>& zobs, std::vector<double>& sgzar, const catalogue::Catalogue data)
    : m_ns(ns_value), m_scheme(scheme), m_nsg(nsg_value), m_zobs(zobs), m_sgzar(sgzar) , m_data(std::make_shared<catalogue::Catalogue>(catalogue::Catalogue(std::move(data)))) {};

    /**
     *  @brief constructor with specified parameters.
     *
     *  @param ns_value resolution nside of Healpix (must be a power of 2).
     *
     *  @param scheme scheme for the HEALPix map (RING or NESTED).
     *
     *  @param nsg_value number of sigmas for source selection.
     *
     *  @param zobs central redshifts of the redshift shells.
     *
     *  @param sgzar width of the Gaussian shells.
     *
     *  @param seed seed for random number generation.
     *
     *  @param data catalogue of data.
     */
    MapCalculator (int ns_value, Healpix_Ordering_Scheme scheme, int nsg_value, std::vector<double>& zobs, std::vector<double>& sgzar, const int seed, const catalogue::Catalogue data)
    : m_ns(ns_value), m_scheme(scheme), m_nsg(nsg_value), m_zobs(zobs), m_sgzar(sgzar) , m_seed(seed), m_data(std::make_shared<catalogue::Catalogue>(catalogue::Catalogue(std::move(data)))) {};

    //@}

    /**
     *  @name Functions to get the private members of the class
     */
    ///@{
```

```

    /**
     *  @brief get the value of the private variable m_ns
     *  @return the value of m_ns
     */
    int getNside () const {return m_ns;};

    /**
     *  @brief get the value of the private variable m_scheme
     *  @return the value of m_scheme
     */
    Healpix_Ordering_Scheme getScheme () {return m_scheme;};

    /**
     *  @brief get the value of the private variable m_nsg
     *  @return the value of m_nsg
     */
    int getNsigma () const {return m_nsg;};

    /**
     *  @brief get the values of the private variable m_zobs
     *  @return the value of m_zobs
     */
    std::vector<double> getZobs () const {return m_zobs;};

    /**
     *  @brief get the values of the private variable m_sgzar
     *  @return the value of m_sgzar
     */
    std::vector<double> getWidths () const {return m_sgzar;};

    /**
     *  @brief get the value of the private variable m_seed
     *  @return the value of m_seed
     */
    int getSeed () const {return m_seed;};

    /**
     *  @brief get the values of the private variable m_data
     *  @return pointer to the member m_data of the class Catalogue
     */
    std::shared_ptr<catalogue::Catalogue> getCatalogue () const {return
↪  m_data;};
```

```

      ///@}

      /**
       *  @name Functions to set the private members of the class
       */
      ///@{

      /**
       *  @brief set the value of the private variable m_ns
       *  @param ns resolution nside of Healpix (must be a power of 2).
       */
      void setNside (const int ns) {m_ns = ns;};

      /**
       *  @brief set the value of the private variable m_scheme
       *  @param scheme scheme for the HEALPix map (RING or NESTED).
       */
      void setScheme (Healpix_Ordering_Scheme scheme) {m_scheme = scheme;};

      /**
       *  @brief set the value of the private variable m_nsg
       *  @param nsg number of sigmas for source selection.
       */
      void setNsigma (const int nsg) {m_nsg=nsg;};

      /**
       *  @brief set the values of the private variable m_zobs
       *  @param zobs central redshifts of the redshift shells.
       */
      void setZobs (const std::vector<double> zobs) {m_zobs=zobs;};

      /**
       *  @brief set the values of the private variable m_sgzar
       *  @param sgzar width of the Gaussian shells.
       */
      void setWidths (const std::vector<double> sgzar) {m_sgzar=sgzar;};

      /**
      *  @brief set the value of the private variable m_seed
      *  @param seed seed for random number generation.
      */
      void setSeed (const int seed) {m_seed=seed;};
```

```

    /**
     *  @brief set the values of the private variable m_data
     *  @param data catalogue of data.
     */
    void setCatalogue (const catalogue::Catalogue data)
     {m_data = std::make_shared<catalogue::Catalogue>
↪  (catalogue::Catalogue(std::move(data)));};

    //@}

    /**
     *  @name Functions to retrieve HEALPix functions
     */
    ///@{

    /**
     *  @brief performs the conversion from angles to pixels
     * to build the HEALPix map. The angles are treated in radians units
     * and they refer to spherical polar coordinates theta phi with domain
     * [0, pi[ and [0, 2pi[
     *  @return integer array of pixels
     */
    std::vector<int> HealpixMapping () const;

    //@}

    /**
     *  @name Functions to generate HEALPix maps
     */
    ///@{

    /**
     *  @brief builds the map for the angular density fluctuations, given
↪  the input mask
     *  @param mskin: Healpix sky mask
     *  @return 3D tensor containing the values associated to the map's
↪  pixels
     */
    std::vector<std::vector<std::vector<double>>> ADF (std::vector<double>
↪  const& mskin) const;

    /**
```

```
   *  @brief builds the map for the angular density fluctuations and the angular redshift fluctuations,
   *  given the input mask
   *  @param mskin: Healpix sky mask
   *  @return standard tuple containing the two maps
   */
  std::tuple<
  std::vector<std::vector<std::vector<double>>> ,
  std::vector<std::vector<std::vector<double>>>
  > ADF_ARF (std::vector<double> const& mskin) const;

  /**
   *  @brief builds the map for the redshifted angular density fluctuations, given the input mask
   *  @param mskin: Healpix sky mask
   *  @param zH: Hubble drift redshifts
   *  @return 3D tensor containing the values associated to the map's pixels
   */
  std::vector<std::vector<std::vector<double>>> ADFr (std::vector<double> const& mskin, std::vector<double> const& zH) const;

  /**
   *  @brief builds the map for the redshifted angular density fluctuations and the redshifted
   * angular redshift fluctuations, given the input mask
   *  @param mskin: Healpix sky mask
   *  @param zH: Hubble drift redshifts
   *  @return standard tuple containing the two maps
   */
  std::tuple<
  std::vector<std::vector<std::vector<double>>> ,
  std::vector<std::vector<std::vector<double>>>
  > ADFr_ARFr (std::vector<double> const& mskin, std::vector<double> const& zH) const;

  /**
   *  @brief builds the map for the shot noise, given the input mask
   *  @param mskin: Healpix sky mask
   *  @param nran: number of mocks computing purely random ADF/ARF mocks
   *  @return 5D tensor containing 2 * nran maps
   */
```

```
   std::vector<std::vector<std::vector<std::vector<std::vector<double>>>>>
↪  SN (std::vector<double> const& mskin, const int nran) const;

   /**
   *  @brief prints a map on an output file
   *  @param map: your map as a 3D tensor
   *  @param filename: the name of the output file
   */
   void print_map (const std::vector<std::vector<std::vector<double>>>&
↪  map, const std::string& filename) const;

   ///@}

  };

 }

}

#endif
```

MapCalculator.cpp

```
/********************************************************************
 *  Copyright (C) 2015 by Federico Marulli                          *
 *  federico.marulli3@unibo.it                                      *
 *                                                                  *
 *  This program is free software; you can redistribute it and/or   *
 *  modify it under the terms of the GNU General Public License as  *
 *  published by the Free Software Foundation; either version 2 of  *
 *  the License, or (at your option) any later version.             *
 *                                                                  *
 *  This program is distributed in the hope that it will be useful, *
 *  but WITHOUT ANY WARRANTY; without even the implied warranty of  *
 *  MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.  See the   *
 *  GNU General Public License for more details.                    *
 *                                                                  *
```

```
 *  You should have received a copy of the GNU General Public       *
 *  License along with this program; if not, write to the Free      *
 *  Software Foundation, Inc.,                                      *
 *  59 Temple Place - Suite 330, Boston, MA  02111-1307, USA.       *
 ********************************************************************/

/**
 *  @file MapCalculator.cpp
 *
 *  @brief Methods of the class MapCalculator
 *
 *  This file contains the implementation of the methods of the class
 *  MapCalculator, used to produce sky maps with HEALPix
 *
 *  @author Matteo Santini
 *
 *  @author matteo.santini7@studio.unibo.it
 */

#include "MapCalculator.h"
#include "RandomNumbers.h"

std::vector<int> cbl::measure::MapCalculator::HealpixMapping () const
{
  Healpix_Base healpixBase(m_ns, m_scheme, SET_NSIDE);  // RING ordering
↪  scheme as default
  std::vector<int> ipx1;
  auto size = m_data->nObjects();
  //need to convert from RA,DEC into polar spherical coordinates
  for (size_t i=0; i<size; ++i) {
    //the angles must be between[0, pi[ and [0, 2pi[
    pointing ang((cbl::par::pi/2)-m_data->dec(i), m_data->ra(i));
    ipx1.push_back(healpixBase.ang2pix(ang));
  }
  return ipx1;
}

std::vector<std::vector<std::vector<double>>>
↪  cbl::measure::MapCalculator::ADF (std::vector<double> const& mskin) const
{

  std::vector<bool> cnzero(mskin.size()); // For the mask application

```

```
  std::transform(mskin.begin(), mskin.end(), cnzero.begin(), [](double x)
  {
    return x>0.0;
  });
  // Now cnzero is a boolean array (0 and 1), values depending on the
↪   matching of the condition in the lambda function

  const unsigned int nzobs = m_zobs.size();
  const unsigned int nsgz = m_sgzar.size();
  const unsigned int size = m_data->nObjects();

  auto ipx1 = HealpixMapping();

  std::vector<std::vector<std::vector<double>>> map_ADF (nzobs,
↪   std::vector<std::vector<double>>(nsgz, std::vector<double>(12 * m_ns *
↪   m_ns, 0.0)));
  std::vector<std::vector<std::vector<double>>> nocounts (nzobs,
↪   std::vector<std::vector<double>>(nsgz, std::vector<double>(12 * m_ns *
↪   m_ns, 0.0)));
  std::vector<std::vector<double>> nang (nzobs, std::vector<double>(nsgz,
↪   0.0));

  for (size_t iz=0; iz<nzobs; ++iz) {
    for (size_t isgz=0; isgz<nsgz; ++isgz) {
      std::vector<int> ipxA;
      std::vector<double> W1a;
      std::vector<int> true_inda;
      std::vector<bool> csel1a(size, 0);
      for (size_t i=0; i<ipx1.size(); ++i) {
        csel1a[i] = ((std::abs(m_data->redshift(i) - m_zobs[iz]) < (m_nsg *
↪   m_sgzar[isgz])) && (mskin[ipx1[i]] != 0));
        if (csel1a[i]==true) {
          true_inda.push_back(i);
        }
      }

      std::vector<double> red;
      for (size_t ind=0; ind<size; ++ind)
        red.push_back(m_data->redshift(ind));
      auto zin_sel = cbl::select<double>(red, true_inda);
      for (size_t i=0; i<true_inda.size(); ++i)
        W1a.push_back(std::exp(-0.5*(std::pow((zin_sel[i]-m_zobs[iz]),
↪   2)/std::pow(m_sgzar[isgz], 2))));
```

```
      auto ipx_sel = select<int>(ipx1, true_inda);
      for (size_t i=0; i<true_inda.size(); ++i)
        ipxA.push_back(ipx_sel[i]);
      for (size_t ig=0; ig<true_inda.size(); ++ig)
        map_ADF[iz][isgz][ipxA[ig]] += W1a[ig];
      for (size_t i=0; i<mskin.size(); ++i) {
        if (cnzero[i]==true) {
          map_ADF[iz][isgz][i] = map_ADF[iz][isgz][i]/mskin[i];
          nocounts[iz][isgz][i] = map_ADF[iz][isgz][i];
          nang[iz][isgz] += map_ADF[iz][isgz][i];
        }
      }

      auto iter_n = std::next(cnzero.begin(), 12*m_ns*m_ns);
      nang[iz][isgz] = nang[iz][isgz]/std::accumulate(cnzero.begin(), iter_n,
↪   0.0);
      double meanADF = 0;
      for (size_t i=0; i<mskin.size(); ++i) {
        if (cnzero[i]==true) {
          map_ADF[iz][isgz][i] = map_ADF[iz][isgz][i]/nang[iz][isgz];
          meanADF += map_ADF[iz][isgz][i];
        }
      }
      meanADF = meanADF/std::accumulate(cnzero.begin(), iter_n, 0.0);
      for (size_t i=0; i<mskin.size(); ++i) {
        if (cnzero[i]==true)
          map_ADF[iz][isgz][i] = map_ADF[iz][isgz][i] - meanADF;
      }
    }
  }
  return map_ADF;
}

std::tuple<std::vector<std::vector<std::vector<double>>>
↪   ,std::vector<std::vector<std::vector<double>>>>
↪   cbl::measure::MapCalculator::ADF_ARF (std::vector<double> const& mskin)
↪   const
{

  std::vector<bool> cnzero(mskin.size()); // For the mask application

  std::transform(mskin.begin(), mskin.end(), cnzero.begin(), [](double x)
  {
```

```
    return x>0.0;
  });
  // Now cnzero is a boolean array (0 and 1), values depending on the
↪   matching of the condition in the lambda function

  const unsigned int nzobs = m_zobs.size();
  const unsigned int nsgz = m_sgzar.size();
  const unsigned int size = m_data->nObjects();

  auto ipx1 = HealpixMapping();

  std::vector<std::vector<std::vector<double>>> map_ADF (nzobs,
↪   std::vector<std::vector<double>>(nsgz, std::vector<double>(12 * m_ns *
↪   m_ns, 0.0)));
  std::vector<std::vector<std::vector<double>>> nocounts (nzobs,
↪   std::vector<std::vector<double>>(nsgz, std::vector<double>(12 * m_ns *
↪   m_ns, 0.0)));
  std::vector<std::vector<std::vector<double>>> map_ARF (nzobs,
↪   std::vector<std::vector<double>>(nsgz, std::vector<double>(12 * m_ns *
↪   m_ns, 0.0)));
  std::vector<std::vector<double>> nang (nzobs, std::vector<double>(nsgz,
↪   0.0));

  for (size_t iz=0; iz<nzobs; ++iz) {
    for (size_t isgz=0; isgz<nsgz; ++isgz) {
      std::vector<int> ipxA;
      std::vector<double> W1a;
      std::vector<int> true_inda;
      std::vector<bool> csel1a(size, 0);
      std::vector<double> zeff;
      std::vector<double> zav1a_vec;
      std::vector<bool> cnzero2(12*m_ns*m_ns, false);
      for (size_t i=0; i<ipx1.size(); ++i) {
        csel1a[i] =
↪   ((std::abs(m_data->redshift(i)-m_zobs[iz])<(m_nsg*m_sgzar[isgz])) &&
↪   (mskin[ipx1[i]]!=0));
        if (csel1a[i]==true)
          true_inda.push_back(i);
      }

      std::vector<double> red;
      for (size_t ind=0; ind<size; ++ind)
        red.push_back(m_data->redshift(ind));
```

```
      auto zin_sel = cbl::select<double>(red, true_inda);
      for (size_t i=0; i<true_inda.size(); ++i)
        W1a.push_back(std::exp(-0.5*(std::pow((zin_sel[i]-m_zobs[iz]),
↪   2)/std::pow(m_sgzar[isgz], 2))));
      auto ipx_sel = select<int>(ipx1, true_inda);
      for (size_t i=0; i<true_inda.size(); ++i) {
        ipxA.push_back(ipx_sel[i]);
        zav1a_vec.push_back((W1a[i]*zin_sel[i]));
        zeff.push_back(zin_sel[i]);
      }

      double zav1a = std::accumulate(zav1a_vec.begin(), zav1a_vec.end(),
↪   0.0)/std::accumulate(W1a.begin(), W1a.end(), 0.0);
      for (size_t ig=0; ig<true_inda.size(); ++ig) {
        map_ADF[iz][isgz][ipxA[ig]] += W1a[ig];
        map_ARF[iz][isgz][ipxA[ig]] += W1a[ig]*(zeff[ig]-zav1a);
      }
      for (size_t i=0; i<mskin.size(); ++i)
        cnzero2[i] = (map_ADF[iz][isgz][i]!=0); // assigning bools
      for (size_t i=0; i<mskin.size(); ++i) {
        if (cnzero[i]==true) {
          map_ADF[iz][isgz][i] = map_ADF[iz][isgz][i]/mskin[i];
          nocounts[iz][isgz][i] = map_ADF[iz][isgz][i];
          nang[iz][isgz] += map_ADF[iz][isgz][i];
        }
        if (cnzero2[i]==true)
          map_ARF[iz][isgz][i] = map_ARF[iz][isgz][i]/mskin[i];
      }

      auto iter_n = std::next(cnzero.begin(), 12*m_ns*m_ns);
      nang[iz][isgz] = nang[iz][isgz]/std::accumulate(cnzero.begin(), iter_n,
↪   0.0);
      double meanADF = 0;
      double meanARF = 0;
      for (size_t i=0; i<mskin.size(); ++i) {
        if (cnzero[i]==true) {
          map_ADF[iz][isgz][i] = map_ADF[iz][isgz][i]/nang[iz][isgz];
          meanADF += map_ADF[iz][isgz][i];
        }
        if (cnzero2[i]==true) {
          map_ARF[iz][isgz][i] = map_ARF[iz][isgz][i]/nang[iz][isgz];
          meanARF += map_ARF[iz][isgz][i];
        }
```

```
      }
      meanADF = meanADF/std::accumulate(cnzero.begin(), iter_n, 0.0);
      auto iter_n2 = std::next(cnzero2.begin(), 12*m_ns*m_ns);
      meanARF = meanARF/std::accumulate(cnzero2.begin(), iter_n2, 0.0);
      for (size_t i=0; i<mskin.size(); ++i) {
        if (cnzero[i]==true)
          map_ADF[iz][isgz][i] = map_ADF[iz][isgz][i]-meanADF;
        if (cnzero2[i]==true)
          map_ARF[iz][isgz][i]=map_ARF[iz][isgz][i]-meanARF;
      }
    }
  }
  return std::make_tuple(map_ADF, map_ARF);
}

std::vector<std::vector<std::vector<double>>>
↪   cbl::measure::MapCalculator::ADFr (std::vector<double> const& mskin,
↪   std::vector<double> const& zH) const
{
  std::vector<bool> cnzero(mskin.size()); // For the mask application

  std::transform(mskin.begin(), mskin.end(), cnzero.begin(), [](double x)
  {
    return x>0.0;
  });
  // Now cnzero is a boolean array (0 and 1), values depending on the
↪   matching of the condition in the lambda function

  const unsigned int nzobs = m_zobs.size();
  const unsigned int nsgz = m_sgzar.size();
  const unsigned int size = m_data->nObjects();

  auto ipx1 = HealpixMapping();

  std::vector<std::vector<std::vector<double>>> map_ADFr (nzobs,
↪   std::vector<std::vector<double>>(nsgz, std::vector<double>(12 * m_ns *
↪   m_ns, 0.0)));
  std::vector<std::vector<double>> nangr (nzobs, std::vector<double>(nsgz,
↪   0.0));

  for (size_t iz=0; iz<nzobs; ++iz) {
    for (size_t isgz=0; isgz<nsgz; ++isgz) {
      std::vector<int> ipxA;
```

```
    std::vector<double> W1b;
    std::vector<int> true_indb;
    std::vector<bool> csel1b(size, 0);
    std::vector<double> zeff;
    std::vector<bool> cnzero2(12*m_ns*m_ns, false);
    for (size_t i=0; i<ipx1.size(); ++i) {
      csel1b[i] = (std::abs(zH[i]-m_zobs[iz])<m_nsg*m_sgzar[isgz]) &&
↪  (mskin[ipx1[i]]!=0);
      if (csel1b[i]==true)
        true_indb.push_back(i);
    }

    std::vector<double> red;
    for (size_t ind=0; ind<size; ++ind)
      red.push_back(m_data->redshift(ind));
    auto zH_sel = cbl::select<double>(red, true_indb);
    for (size_t i=0; i<true_indb.size(); ++i)
      W1b.push_back(std::exp(-0.5*(std::pow((zH_sel[i]-m_zobs[iz]),
↪  2)/std::pow(m_sgzar[isgz], 2))));
    auto ipx_sel = select<int>(ipx1, true_indb);
    for (size_t i=0; i<true_indb.size(); ++i)
      ipxA.push_back(ipx_sel[i]);
    for (size_t ig=0; ig<true_indb.size(); ++ig)
      map_ADFr[iz][isgz][ipxA[ig]] += W1b[ig];
    for (size_t i=0; i<mskin.size(); ++i)
      cnzero2[i] = (map_ADFr[iz][isgz][i]!=0); // assigning bools
    for (size_t i=0; i<mskin.size(); ++i) {
      if (cnzero[i]==true) {
        map_ADFr[iz][isgz][i] = map_ADFr[iz][isgz][i]/mskin[i];
        nangr[iz][isgz] += map_ADFr[iz][isgz][i];
      }
    }

    auto iter_n = std::next(cnzero.begin(), 12*m_ns*m_ns);
    nangr[iz][isgz] = nangr[iz][isgz]/std::accumulate(cnzero.begin(),
↪  iter_n, 0.0);
    double meanADFr = 0;
    for (size_t i=0; i<mskin.size(); ++i) {
      if (cnzero[i]==true) {
        map_ADFr[iz][isgz][i] = map_ADFr[iz][isgz][i]/nangr[iz][isgz];
        meanADFr += map_ADFr[iz][isgz][i];
      }
    }
```

```
      meanADFr=meanADFr/std::accumulate(cnzero.begin(), iter_n, 0.0);
      for (size_t i=0; i<mskin.size(); ++i) {
        if (cnzero[i]==true)
          map_ADFr[iz][isgz][i] = map_ADFr[iz][isgz][i] - meanADFr;
      }
    }
  }
  return map_ADFr;
}

std::tuple<std::vector<std::vector<std::vector<double>>>
↪  ,std::vector<std::vector<std::vector<double>>>>
↪  cbl::measure::MapCalculator::ADFr_ARFr (std::vector<double> const& mskin,
↪  std::vector<double> const& zH) const
{
  std::vector<bool> cnzero(mskin.size()); // For the mask application

  std::transform(mskin.begin(), mskin.end(), cnzero.begin(), [](double x)
  {
    return x>0.0;
  });
  // Now cnzero is a boolean array (0 and 1), values depending on the
↪  matching of the condition in the lambda function

  const unsigned int nzobs = m_zobs.size();
  const unsigned int nsgz = m_sgzar.size();
  const unsigned int size = m_data->nObjects();

  auto ipx1 = HealpixMapping();

  std::vector<std::vector<std::vector<double>>> map_ADFr (nzobs,
↪  std::vector<std::vector<double>>(nsgz, std::vector<double>(12 * m_ns *
↪  m_ns, 0.0)));
  std::vector<std::vector<std::vector<double>>> map_ARFr (nzobs,
↪  std::vector<std::vector<double>>(nsgz, std::vector<double>(12 * m_ns *
↪  m_ns, 0.0)));
  std::vector<std::vector<double>> nangr (nzobs, std::vector<double>(nsgz,
↪  0.0));

  for (size_t iz=0; iz<nzobs; ++iz) {
    for (size_t isgz=0; isgz<nsgz; ++isgz) {
      std::vector<int> ipxA;
      std::vector<double> W1b;
```

```
      std::vector<int> true_indb;
      std::vector<bool> csel1b(size, 0);
      std::vector<double> zeff;
      std::vector<double> zav1b_vec;
      std::vector<bool> cnzero2(12*m_ns*m_ns, false);
      for (size_t i=0; i<ipx1.size(); ++i) {
        csel1b[i] = (std::abs(zH[i]-m_zobs[iz])<m_nsg*m_sgzar[isgz]) &&
↪   (mskin[ipx1[i]]!=0);
        if (csel1b[i]==true)
          true_indb.push_back(i);
      }

      std::vector<double> red;
      for (size_t ind=0; ind<size; ++ind)
        red.push_back(m_data->redshift(ind));
      auto zH_sel = cbl::select<double>(red, true_indb);
      for (size_t i=0; i<true_indb.size(); ++i)
        W1b.push_back(std::exp(-0.5*(std::pow((zH_sel[i]-m_zobs[iz]),
↪   2)/std::pow(m_sgzar[isgz], 2))));
      auto ipx_sel = select<int>(ipx1, true_indb);
      for (size_t i=0; i<true_indb.size(); ++i) {
        zav1b_vec.push_back(W1b[i]*zH_sel[i]);
        ipxA.push_back(ipx_sel[i]);
        zeff.push_back(zH_sel[i]);
      }

      double zav1b = std::accumulate(zav1b_vec.begin(), zav1b_vec.end(),
↪   0.0)/std::accumulate(W1b.begin(), W1b.end(), 0.0);
      for (size_t ig=0; ig<true_indb.size(); ++ig) {
        map_ADFr[iz][isgz][ipxA[ig]] += W1b[ig];
        map_ARFr[iz][isgz][ipxA[ig]] += W1b[ig]*(zeff[ig]-zav1b);
      }
      for (size_t i=0; i<mskin.size(); ++i)
        cnzero2[i] = (map_ADFr[iz][isgz][i]!=0); // assigning bools
      for (size_t i=0; i<mskin.size(); ++i) {
        if (cnzero[i]==true) {
          map_ADFr[iz][isgz][i] = map_ADFr[iz][isgz][i]/mskin[i];
          nangr[iz][isgz] += map_ADFr[iz][isgz][i];
        }
        if (cnzero2[i]==true)
          map_ARFr[iz][isgz][i] = map_ARFr[iz][isgz][i]/mskin[i];
      }

```

```
      auto iter_n = std::next(cnzero.begin(), 12*m_ns*m_ns);
      nangr[iz][isgz] = nangr[iz][isgz]/std::accumulate(cnzero.begin(),
 ↪  iter_n, 0.0);
      double meanADFr = 0;
      double meanARFr = 0;
      for (size_t i=0; i<mskin.size(); ++i) {
        if (cnzero[i]==true) {
          map_ADFr[iz][isgz][i] = map_ADFr[iz][isgz][i]/nangr[iz][isgz];
          meanADFr += map_ADFr[iz][isgz][i];
        }
        if (cnzero2[i]==true) {
          map_ARFr[iz][isgz][i] = map_ARFr[iz][isgz][i]/nangr[iz][isgz];
          meanARFr += map_ARFr[iz][isgz][i];
        }
      }
      meanADFr = meanADFr/std::accumulate(cnzero.begin(), iter_n, 0.0);
      auto iter_n2 = std::next(cnzero2.begin(), 12*m_ns*m_ns);
      meanARFr = meanARFr/std::accumulate(cnzero2.begin(), iter_n2, 0.0);
      for (size_t i=0; i<mskin.size(); ++i) {
        if (cnzero[i]==true)
          map_ADFr[iz][isgz][i] = map_ADFr[iz][isgz][i]-meanADFr;
        if (cnzero2[i]==true)
          map_ARFr[iz][isgz][i] = map_ARFr[iz][isgz][i]-meanARFr;
      }
    }
  }
  return std::make_tuple(map_ADFr, map_ARFr);
}

std::vector<std::vector<std::vector<std::vector<std::vector<double>>>>>
 ↪  cbl::measure::MapCalculator::SN (std::vector<double> const& mskin, const
 ↪  int nran) const
{
  std::vector<bool> cnzero(mskin.size()); // For the mask application

  std::transform(mskin.begin(), mskin.end(), cnzero.begin(), [](double x)
  {
    return x>0.0;
  });
  // Now cnzero is a boolean array (0 and 1), values depending on the
 ↪  matching of the condition in the lambda function

  const unsigned int nzobs = m_zobs.size();
```

```
  const unsigned int nsgz = m_sgzar.size();
  const unsigned int size = m_data->nObjects();
  cbl::random::UniformRandomNumbers rnd(0., 1., m_seed);

  auto ipx1 = HealpixMapping();

  auto map_SN =
↪  std::vector<std::vector<std::vector<std::vector<std::vector<double>>>>>
↪  (nzobs, std::vector<std::vector<std::vector<std::vector<double>>>>(nsgz,
↪  std::vector<std::vector<std::vector<double>>>(nran,
↪  std::vector<std::vector<double>>(2, std::vector<double>(12 * m_ns * m_ns,
↪  0.0)))));

  for (size_t iz=0; iz<nzobs; ++iz) {
    for (size_t isgz=0; isgz<nsgz; ++isgz) {
      std::vector<int> ipxA;
      std::vector<double> W1a;
      std::vector<int> true_inda;
      std::vector<bool> csel1a(size, 0);
      std::vector<double> zeff;
      std::vector<double> zav1a_vec;
      std::vector<bool> cnzero2(12*m_ns*m_ns, false);
      for (size_t i=0; i<ipx1.size(); ++i) {
        csel1a[i] =
↪  ((std::abs(m_data->redshift(i)-m_zobs[iz])<(m_nsg*m_sgzar[isgz])) &&
↪  (mskin[ipx1[i]]!=0));
        if (csel1a[i]==true)
          true_inda.push_back(i);
      }

      std::vector<double> red;
      for (size_t ind=0; ind<size; ++ind)
        red.push_back(m_data->redshift(ind));
      auto zin_sel = cbl::select<double>(red, true_inda);
      for (size_t i=0; i<true_inda.size(); ++i)
        W1a.push_back(std::exp(-0.5*(std::pow((zin_sel[i]-m_zobs[iz]),
↪  2)/std::pow(m_sgzar[isgz], 2))));
      auto ipx_sel = select<int>(ipx1, true_inda);
      for (size_t i=0; i<true_inda.size(); ++i) {
        ipxA.push_back(ipx_sel[i]);
        zav1a_vec.push_back((W1a[i]*zin_sel[i]));
        zeff.push_back(zin_sel[i]);
      }
```

```

    double zav1a = std::accumulate(zav1a_vec.begin(), zav1a_vec.end(),
↪ 0.0)/std::accumulate(W1a.begin(), W1a.end(), 0.0);
    double cumsum = 0;
    std::vector<int> true_cnzero;
    for (size_t i=0; i<cnzero.size(); ++i) {
      if (cnzero[i]==true)
        true_cnzero.push_back(i);
    }
    std::vector<double> cum = cbl::select<double>(mskin, true_cnzero);
    for (size_t i=0; i<cum.size(); ++i) {
      if (i>0)
        cumsum = cum[i-1];
      cum[i]= cum[i]+cumsum;
    }
    std::vector<double> Cprobar;
    for (size_t i=0; i<cum.size(); ++i)
      Cprobar.push_back(cum[i]/std::accumulate(mskin.begin(), mskin.end(),
↪ 0.0));
    std::vector<double> jeff = cbl::arange(0, static_cast<unsigned
↪ int>(std::accumulate(cnzero.begin(), cnzero.end(), 0)), 1);
    assert(Cprobar.size()==jeff.size());
    std::vector<int> jindex;
    auto all_jindex = cbl::arange(12*m_ns*m_ns);
    for (size_t i=0; i<cnzero.size(); ++i) {
      if (cnzero[i]==true)
        jindex.push_back(all_jindex[i]);
    }
    std::vector<std::vector<double>> probg (nran,
↪ std::vector<double>(true_inda.size(), 0.0));

    // Filling with randoms
    for (int r=0; r<nran; ++r) {
      for (size_t c=0; c<true_inda.size(); ++c)
        probg[r][c] = rnd();
    }
    std::vector<std::vector<int>> jgar0 (nran,
↪ std::vector<int>(true_inda.size(), 0.0));
    for (int r=0; r<nran; ++r) {
      for (size_t c=0; c<true_inda.size(); ++c) {
        //sets vanishing outrange values
        if (probg[r][c]>*std::max_element(Cprobar.begin(), Cprobar.end())
↪ || probg[r][c]<*std::min_element(Cprobar.begin(), Cprobar.end()))
```

```
        jgar0[r][c] = 0;
      else jgar0[r][c] = static_cast<int>(cbl::interpolated(probg[r][c], Cprobar, jeff, "Linear"));
    }
  }
  std::vector<std::vector<double>> jgar (nran, std::vector<double>(true_inda.size(), 0.0));
  for (int r=0; r<nran; ++r) {
    for (size_t c=0; c<true_inda.size(); ++c)
      jgar[r][c] = jindex[jgar0[r][c]];
  }
  for (int jran=0; jran<nran; ++jran) {
    for (size_t ig=0; ig<true_inda.size(); ++ig) {
      int jg = jgar[jran][ig];
      map_SN[iz][isgz][jran][0][jg] += W1a[ig];
      map_SN[iz][isgz][jran][1][jg] += W1a[ig]*(zeff[ig]-zav1a);
    }
    for (size_t i=0; i<mskin.size(); ++i)
      cnzero2[i] = map_SN[iz][isgz][jran][0][i]!=0; // assigning bools
    std::vector<double> nangT_vec;
    std::vector<double> nangT2_vec;
    for (size_t i=0; i<mskin.size(); ++i) {
      if (cnzero[i]==true) {
        map_SN[iz][isgz][jran][0][i] = map_SN[iz][isgz][jran][0][i]/mskin[i];
        nangT_vec.push_back(map_SN[iz][isgz][jran][0][i]);
      }
    }
    for (size_t i=0; i<mskin.size(); ++i) {
      if (cnzero2[i]==true)
        map_SN[iz][isgz][jran][1][i] = map_SN[iz][isgz][jran][1][i]/mskin[i];
    }
    double nangT = std::accumulate(nangT_vec.begin(), nangT_vec.end(), 0.0)/nangT_vec.size();
    double meanSN0 = 0;
    double meanSN1 = 0;
    for (size_t i=0; i<mskin.size(); ++i) {
      if (cnzero[i]==true) {
        map_SN[iz][isgz][jran][0][i] = map_SN[iz][isgz][jran][0][i]/nangT;
        meanSN0 += map_SN[iz][isgz][jran][0][i];
      }
```

```
          if (cnzero2[i]==true) {
            map_SN[iz][isgz][jran][1][i] =
↪ map_SN[iz][isgz][jran][1][i]/nangT;
            meanSN1 += map_SN[iz][isgz][jran][1][i];
          }
        }
        meanSN0 = meanSN0/std::accumulate(cnzero.begin(), cnzero.end(), 0.0);
        meanSN1 = meanSN1/std::accumulate(cnzero2.begin(), cnzero2.end(),
↪ 0.0);
        for (size_t i=0; i<mskin.size(); ++i) {
          if (cnzero2[i]==true)
            map_SN[iz][isgz][jran][1][i] =
↪ map_SN[iz][isgz][jran][1][i]-meanSN1;
          if (cnzero[i]==true)
            map_SN[iz][isgz][jran][0][i] =
↪ map_SN[iz][isgz][jran][0][i]-meanSN0;
        }
      }
    }
  }
  return map_SN;
}

void cbl::measure::MapCalculator::print_map (const
↪ std::vector<std::vector<std::vector<double>>>& map, const std::string&
↪ filename) const {

  std::ofstream outputFile(filename);

  if (!outputFile.is_open()) {
    ErrorCBL("An error occured in opening the file "+filename, "print_map",
↪ "MapCalculator.cpp");
    return;
  }
  // Printing map_ADF as 3D array
  for (size_t ra=0; ra<map.size(); ++ra) {
    for (size_t dec=0; dec<map[ra].size(); ++dec) {
      for (size_t z=0; z<map[ra][dec].size(); ++z) {
        outputFile << map[ra][dec][z] << " ";
      }
      outputFile << std::endl;
    }
    outputFile << std::endl;
```

```
  }
  // Closing file
  outputFile.close();
  coutCBL << "Data printed on " << filename << std::endl;
}

```