UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia “Galileo Galilei”

Master Degree in Astrophysics and Cosmology

Final Dissertation

A study of marked power spectra including the matter tidal field

Thesis supervisor
Prof. Michele Liguori

Thesis co-supervisors
Dr. Marco Marinucci

Candidate
Matilde Schianchi

Academic Year 2025/2026

Abstract

Upcoming galaxy surveys such as Euclid will map the Large-Scale Structure (LSS) of the Universe with unprecedented precision and the main challenge is to extract as much cosmological information as possible from these new datasets.
The galaxy power spectrum is the standard statistics used in LSS analyses, but non linear gravitational evolution transfers cosmological information into higher-order correlators, which are expensive to measure and model. Marked statistics offer a compelling alternative: by non linearly weighting the density field before computing its power spectrum, marked statistics can recover part of the higher-order information while retaining the simplicity of a two-point correlation function.
Existing implementations, however, rely solely on marks constructed from the local density field, which does not capture the anisotropic geometry of the environment. This information is instead encoded in the matter tidal tensor.
For this reason, this thesis extends the marked power spectrum formalism to a mark constructed from the matter tidal field, recovering the tidal marked power spectrum at one-loop order for both dark matter and biased tracers in real space within the Effective Field Theory of the Large-Scale Structure, complemented by a dedicated numerical implementation.
Our Fisher forecast analysis shows that combining the standard power spectrum with the proposed tidal marked power spectrum substantially tightens cosmological constraints on the physical cold dark matter density parameter ω_{cdm}, on the dimensionless Hubble parameter h and on the scalar spectral index n_s, which are the most robust results of this study, by factors ~ 1.7, ~ 1.7 and ~ 2 respectively. The quadratic and tidal bias parameters b_2 and $b_{\mathcal{G}_2}$, generally well constrained with the inclusion of the galaxy bispectrum, are also measured with substantially higher precision, improving by up to a factor ~ 3.2 and ~ 2.7 respectively. Finally, the scalar amplitude of primordial fluctuations A_s and the linear bias b_1 show a remarkable improvement, driven by the breaking of the $A_s - b_1$ degeneracy present in real space.
These results confirm the effectiveness of the tidal mark in recovering higher-order cosmological information, providing a new tool towards the broader goal of maximizing the scientific return of next-generation galaxy surveys.

Acknowledgments

I am deeply grateful to my supervisor, Prof. Michele Liguori, and my co-supervisor, Dr. Marco Marinucci, for guiding me through this journey, for the time they invested and for always being available to discuss every aspect of this work.

A special thank you goes to my family and friends for always encouraging and supporting me. Their support made even the most difficult moments lighter.

I would also like to thank my colleagues, who shared with me the challenges, the successes, and the moments that made this experience unique. Their friendship created memories that I will always cherish.

Contents

Introduction

Over the last decades, a combination of observations has led to the emergence and confirmation of the ΛCDM cosmological model [1]. In this model, the Universe has evolved from a homogeneous state after the Big Bang, to a hierarchical assembly of galaxies, clusters and superclusters at our epoch. The energy density of the resulting Universe is dominated by two mysterious components: 69% of its energy density is in the form of Dark Energy, which is causing the universe expansion to accelerate, another 26% of the energy is in the form of dark matter, which interacts exclusively via gravitational interactions. Only 5% of the Universe is made up of the ordinary baryonic matter composed of protons, neutrons and electrons.

The existence of Dark Energy is in conflict with our knowledge of fundamental physics and a key unresolved question is whether it behaves like the cosmological constant Λ introduced by Einstein [2]. Also the nature of dark matter is unknown, even if several candidates exist in particle physics. One possibility to explain these puzzles is that the theory of General Relativity needs to be revised on cosmological scales.

The next generation of galaxy surveys, such as Euclid [3], DESI [4] and Vera C. Rubin Observatory's LSST [5], are designed to answer these and other fundamental questions in cosmology.

In particular, Euclid is designed to answer key questions such as: is the Dark Energy a cosmological constant as introduced by Einstein, or is it a scalar field that evolves dynamically with the expansion of the universe? Alternatively, is the Dark Energy a manifestation of a breakdown of General Relativity on the largest scales? What are the nature and properties of dark matter? What are the initial conditions in the Early Universe, which seeded the formation of cosmic structure? [3].

The central goal of modern cosmology is therefore to extract as much cosmological information as possible from these new datasets which are expected to be bigger and to allow a higher precision in constraining cosmological parameters. For example, compared to BOSS [6], which measured ~ 1.2 million luminous red galaxies over $10,000$ deg^2 in the range $0.1 < z < 0.7$, Euclid will measure ~ 35 million spectroscopic redshifts over $14,000$ deg^2 at higher redshift $0.9 < z < 1.8$. We expect this to translate into tighter constraints on the equation of state for dark energy, neutrino masses and deviations from General Relativity [3].

Today the Large-Scale Structure analysis relies on the galaxy power spectrum, the Fourier transform of the galaxy density two-point correlation function, to extract information from galaxy surveys datasets. At early times, when deviations of the primordial fluctuations from Gaussianity are minimal the power spectrum is a near complete statistic. At later times, non linear gravitational evolution generates significant deviations from Gaussianity and information is redistributed into higher-order statistics, notably the bispectrum and the trispectrum.

The practical implementation of the bispectrum and higher-order correlators, however, becomes increasingly complex, especially in the mildly non linear regime, where the signal is expected to be richer. This complexity arises both from the computational cost of modelling a larger configuration space, the challenges in estimating its covariance and the complexity in the analytical calculations [7].

Marked statistics [8, 9, 10, 11, 12] offer a compelling alternative. Rather than measuring higher-order correlators directly, one weights the density field with a suitably chosen non linear function, the mark, before computing the power spectrum. Because a non linear transformation reshapes the statistics of the field, part of the information that would reside in the higher-order correlators is transferred back to the power spectrum [13].

Existing implementations of this approach, however, rely only on marks constructed from the local density field. But the structure of the cosmic web, voids, sheets, filaments and knots, is determined by gravitational collapse, a strongly anisotropic process whose geometry is encoded in the tidal tensor $T_{ij} = \partial_i \partial_j \phi$, where ϕ is the gravitational field. Since the formation and evolution of galaxies and halos depends on both density and tidal environment [14, 15, 16], extending the marked statistics framework to tidal field dependent marks constitute a natural extension of the formalism.

This thesis develops such an extension. Working within the Effective Field Theory of the Large-Scale Structure (EFTofLSS) [17, 18, 19] to correctly account for mildly non linear scales, we build a mark from the contraction of the tidal tensor with itself smoothed over a distance R and derive the corresponding one-loop marked power spectrum, both for dark matter and biased tracers in real space. In parallel to the analytical part, we develop a numerical implementation to run a Fisher forecast to quantify how much cosmological constraining power this tidal field observable adds on top of the standard power spectrum alone analysis. We present the results in terms of the improvement σ_P / σ_{P+M} for different values of the smoothing radius $R = \{10, 15, 20\}$Mpc$/h$ and maximum wavevector $k_{\max} = \{0.2,\ 0.25,\ 0.3\} h/$Mpc to push our analysis towards the mildly non linear regime.

This thesis is organized as follows:

- Chapter 1 introduces the background cosmology. It reviews the Cosmological Principle and the FLRW metric, the dynamics of the homogeneous background through the Friedmann equations and the thermal history of the Universe. It then discusses the horizon and flatness problems and shows how inflation solves them, presenting the inflationary paradigm and the generation of primordial quantum fluctuations from the inflaton field. The chapter closes with a first description of the cosmic web, the large-scale pattern of voids, filaments and clusters that emerges from structure formation.

- Chapter 2 develops the theoretical framework used to describe Large-Scale Structure. Starting from the Vlasov equation, it derives the Eulerian equations of motion for a pressureless fluid and constructs the perturbative solution for the density contrast order by order. It then gives the statistical description of the density field: the physical origin of its fluctuations from inflation, and their characterization through correlation functions and power spectra, including the Wick theorem for Gaussian fields. The chapter discusses the limitations of standard perturbation theory at small scales and introduces the Effective Field Theory of Large-Scale Structure (EFTofLSS) as a consistent framework to extend the validity of the perturbative expansion. It closes with the galaxy bias expansion, which relates the galaxy density field to the underlying matter field through a set of bias parameters.

- Chapter 3 turns to the observational side of the framework. It introduces the power spectrum and bispectrum of the density field as cosmological estimators, illustrating how they are measured from galaxy survey data and used to constrain cosmological parameters. It discusses how the computational and analytical cost of a bispectrum analysis motivates the modeling of the marked power spectrum as a statistics that compresses part of the higher-order information in a two point correlation function. The chapter also introduces the Redshift Space Distortions (RSD), one of the main observational complications affecting real spectroscopic surveys. Finally, it presents the Fisher forecasting formalism, a statistical tool used to quantify the constraining power of a given statistic.

- Chapter 4 contains the original contribution of the thesis. It first reviews the marked power spectrum formalism for a mark built from the local filtered density field: the definition of the mark, its Eulerian perturbation theory expansion, the resulting one-loop marked power spectrum for matter and biased tracers, and the associated counterterms and shot noise contributions. The chapter then introduces the original extension to a mark based on the contraction of the matter tidal tensor, $T_{ij}T^{ij}$, related to the density contrast through the gravitational potential via the Poisson equation. The mark is defined, expanded perturbatively following the same procedure used for the density based mark, and the corresponding one-loop marked power spectrum is derived for both dark matter and biased tracers. The chapter also describes the numerical implementation used to evaluate these expressions, and discusses alternative marks that could be built from the tidal tensor.

- Chapter 5 presents the original results from a Fisher forecast analysis. Using the volume and redshift of a BOSS-like survey, it compares the constraints on cosmological, bias and EFT parameters obtained from the standard power spectrum alone with those obtained by combining it with the tidal marked power spectrum, for different choices of $k_{\max}$ and smoothing scale R.

- Conclusions: presents the conclusions of this work. We discuss the Fisher forecast results, acknowledge the limitations of the current framework, and propose future research directions.

- Appendix A: contains the analytical derivation of the 1-loop contributions of power spectrum and tree-level bispectrum.

- Appendix B: contains the analytical derivation of the 1-loop contributions of the tidal marked power spectrum.

- Appendix C: presents the complete triangle plot for all the cosmological, bias and EFT parameters considered in this thesis for different values of $k_{\max} = \{0.2,\, 0.25,\, 0.3\}h/\mathrm{Mpc}$.

Chapter 1

Standard cosmology

1.1 The Copernican and the Cosmological principle

One of the fundamental assumptions of modern cosmology is the Copernican principle. It states that humans are not privileged observers of the universe, that is: observations from the Earth are representative of observations from the average position in the universe. This principle can be combined with the fact that we observe isotropy to yield the Cosmological principle:

Every comoving observer, at large enough scales, sees the Universe, at a fixed time in their reference frame, as homogeneous and isotropic.

Isotropy means that the characteristics of the universe are the same in every direction and homogeneity means that the characteristics of the universe are the same at every point in space.
By comoving observer we mean an observer that moves coherently with the cosmic rest frame, that is the rest frame of the cosmic fluid. But how do we define the cosmic rest frame? Observations of the Cosmic Microwave Background (CMB) show that it is an almost perfect blackbody with a temperature of $T_{\mathrm{CMB}} = 2.72\,\mathrm{K}$ [20]. However, this radiation is not uniformly distributed in the sky: it presents a dipole anisotropy of $\Delta T \sim 3.4\,\mathrm{mK}$, caused by the Doppler effect arising from the motion of the Solar System with respect to the CMB[20]. From this dipole we can measure the peculiar velocity of the Solar System relative to the CMB to be roughly $370\,\mathrm{km/s}$. And given that the Local Group, which includes the Milky Way, also moves relative to the CMB at roughly $620\,\mathrm{km/s}$ [20], the cosmic rest frame can be operationally defined as the reference frame in which the CMB dipole anisotropy vanishes. The proper time of a comoving observer is what we refer to with “fixed time”, and it is called cosmic time.
A crucial feature of the cosmological principle is that it is expected to hold only on large scales since at small scales we see structures, such as galaxies, so we do not observe homogeneity.

1.2 Spacetime

The best description for gravity we have so far is the General theory of Relativity[21]. In General Relativity, spacetime is modelled as a $3 + 1$-dimensional smooth pseudo-Riemannian manifold. Such a spacetime can have up to 10 global continuous symmetries: time translation, three Lorentz boosts, three spatial translations and three

spatial rotations.
In cosmology we lose the time translation symmetry, since our Universe is expanding[22, 23], and the three Lorentz boosts symmetry.
The most general form of a metric that satisfies these conditions is the Friedmann-Lemaitre-Robertson-Walker (FLRW) line element. In the comoving frame it reads:

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + a^2(t)\Big(\frac{\mathrm{d}r^2}{1-kr^2} + r^2\mathrm{d}\Omega^2\Big)\,, \tag{1.1}$$

where $\mathrm{d}\Omega^2 = \mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2$.
The parameter $a(t)$ is called the scale factor and is related to the expansion of the universe and it relates the comoving coordinate to the physical coordinate via

$$\lambda_{phys} = a(t)\lambda_{com}\,. \tag{1.2}$$

The parameter k is a constant describing the spatial curvature, which can always be normalized to ±1 or 0. A universe with $k=1$ is a closed universe, $k=0$ is a flat universe, and $k=-1$ is an open universe.
The expansion rate of the universe is quantified by the Hubble parameter, defined as

$$H(t) \equiv \frac{\dot{a}(t)}{a(t)}\,, \tag{1.3}$$

where the dot denotes differentiation with respect to the cosmic time t.

1.3 Friedmann equations

The dynamical evolution of the universe is determined by Einstein's field equations[24]:

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\mathcal{R} = 8\pi G T_{\mu\nu}\,. \tag{1.4}$$

Here $G_{\mu\nu}$ is the Einstein tensor and it is a function of the metric $g_{\mu\nu} = \mathrm{diag}(-1,+1,+1,+1)$, meaning it describes the geometry of space-time. The energy-momentum tensor $T_{\mu\nu}$ describes the matter and energy content of the universe. $R_{\mu\nu}$ is the Ricci tensor, that depends on the metric and its derivatives, while $\mathcal{R}$ is the Ricci scalar, which is the contraction of the Ricci tensor. The Ricci tensor is the contraction of the Riemann tensor:

$$R_{\mu\nu} = g^{\rho\sigma}R_{\mu\rho\nu\sigma}\,, \tag{1.5}$$

$$R = g^{\mu\nu}R_{\mu\nu}\,, \tag{1.6}$$

and the Riemann tensor is defined as

$$R^{\rho}{}_{\mu\sigma\nu} = \frac{\partial\Gamma^{\rho}_{\mu\nu}}{\partial x^{\sigma}} - \frac{\partial\Gamma^{\rho}_{\mu\sigma}}{\partial x^{\nu}} + \Gamma^{\eta}_{\mu\nu}\Gamma^{\rho}_{\sigma\eta} - \Gamma^{\eta}_{\mu\sigma}\Gamma^{\rho}_{\nu\eta}. \tag{1.7}$$

$\Gamma^{\eta}_{\mu\nu}$ are the Christoffel symbols

$$\Gamma^{\eta}_{\mu\nu} = \frac{1}{2}g^{\eta\lambda}\big(\partial_{\mu}g_{\lambda\nu} + \partial_{\nu}g_{\mu\lambda} - \partial_{\lambda}g_{\mu\nu}\big)\,. \tag{1.8}$$

On scales larger than approximately $100\,h^{-1}$ Mpc, the matter distribution of the Universe can be considered statistically homogeneous, allowing it to be modeled as a perfect fluid [25]. Such a fluid is completely described by the energy density ρ and the pressure p. The energy-momentum tensor for a perfect fluid takes the following form:

$$T_{\mu\nu} = (\rho + p)u_\mu u_\nu + p g_{\mu\nu}\,, \tag{1.9}$$

with u_μ the fluid element 4-velocity. In the comoving frame, adopting the mostly-plus signature for the metric, the four-velocity is $u^\mu = (1, 0, 0, 0)$. Lowering the index with the metric, $u_\mu = g_{\mu\nu}u^\nu = (-1, 0, 0, 0)$, so we obtain:

$$T_{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix}\,. \tag{1.10}$$

Starting from Eq. 1.4 we can derive two independent equations

$$H^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}\,, \tag{1.11}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P)\,. \tag{1.12}$$

These are called first 1.11 and second 1.12 Friedmann equations[26]. Taking the covariant derivative of Einstein's field equations and using the contracted Bianchi identity, we get $\nabla_\mu G^{\mu\nu} = 0$. This, together with the metric compatibility condition, $\nabla_\mu g^{\mu\nu} = 0$, implies that

$$\nabla_\mu T^{\mu\nu} = 0\,, \tag{1.13}$$

that represent the conservation of energy and momentum: the covariant derivative accounts for the curvature of spacetime and ensures that this conservation law is formulated in a coordinate-independent way.

For a homogeneous and isotropic universe filled with a perfect fluid, this relation reduces to the continuity equation:

$$\dot{\rho} = -3H(\rho + P)\,, \tag{1.14}$$

also called third Friedmann equation.

To summarize, the system involves three unknown quantities, $a(t)$, $\rho(t)$, and $p(t)$. However, only two of the three equations are independent, so we need an equation of state relating the pressure to the energy density to close the system:

$$p = w\rho\,. \tag{1.15}$$

If we plug 1.15 in 1.14 we find:

$$\rho \propto a^{-3(1+w)}\,, \tag{1.16}$$

where the parameter w takes different values depending on which constituent of the cosmic fluid we are considering: $w = 0$ corresponds to pressureless dust (non relativistic matter), $w = 1/3$ corresponds to radiation and $w = -1$ corresponds to dark energy.

Now we know how the energy density evolves with the scale factor:

$$\rho(a) \propto a^{-3} \qquad \text{matter}\,, \tag{1.17}$$

$$\rho(a) \propto a^{-4} \qquad \text{radiation}\,, \tag{1.18}$$

$$\rho(a) = \text{constant} \qquad \text{dark energy}\,. \tag{1.19}$$

Dark energy is defined to be any kind of fluid which is uniformly distributed in space and has a negative pressure: this negative pressure has the effect of a tension, pulling the universe apart.
Now, if we assume a flat universe, which is consistent with observations [27], we can find the time evolution of the scale factor, of the energy density and of the Hubble parameter:

$$a(t) = a_0 \big[1 + 3/2(1+w)H_0(t-t_0)\big]^{\frac{2}{3(1+w)}} , \tag{1.20}$$
$$\rho(t) = \rho_0 \big[1 + 3/2(1+w)H_0(t-t_0)\big]^{-2} , \tag{1.21}$$
$$H(t) = h_0 \big[1 + 3/2(1+w)H_0(t-t_0)\big]^{-1} . \tag{1.22}$$

The subscript zero stands for "evaluated today". In particular we have $a_0 = 1$ and $H_0 = 67.4 \pm 0.5 \ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ [27].
If we look at 1.20, we see that there is a time when $a(t) = 0$. This time, called t_{BB} and defined by

$$1 + 3/2(1+w)H_0(t_{BB} - t_0) = 0 , \tag{1.23}$$

corresponds to the Big Bang singularity. Indeed, as $t \to t_{BB}$, the scale factor approaches zero while the energy density and the Hubble parameter diverge,

$$a(t) \to 0, \qquad \rho(t) \to \infty, \qquad H(t) \to \infty. \tag{1.24}$$

This indicates that the classical cosmological solution cannot be extended to times earlier than t_{BB}.

1.3.1 Cosmological constant

At the time Einstein formulated the general theory of relativity, the prevailing view was that of a static Universe[28]. However this hypothesis is not compatible with the Friedmann equations, unless $\ddot{a} = 0$, that implies

$$\rho = -3p . \tag{1.25}$$

Since such a condition did not correspond to any known form of matter, Einstein modified his field equations by introducing the cosmological constant in order to obtain a static cosmological solution:

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu} , \tag{1.26}$$

with Λ a constant scalar called cosmological constant. One can now obtain a static solution with an appropriate choice of Λ.
After the discovery of the expansion of the Universe[22, 23], the need for a cosmological constant to support a static solution faded. However, now Λ is the simplest explanation for the observed accelerated expansion of our Universe [29].
To properly account for this cosmological constant we define an effective stress-energy tensor

$$\tilde{T}_{\mu\nu} = T_{\mu\nu} + \frac{\Lambda}{8\pi G} g_{\mu\nu} , \tag{1.27}$$

so we perform the following corrections

$$\rho \to \rho + \frac{\Lambda}{8\pi G}\,, \tag{1.28}$$

$$p \to p - \frac{\Lambda}{8\pi G}\,, \tag{1.29}$$

A cosmological model featuring Λ is the de Sitter universe. In this model, the universe is dominated by a positive cosmological constant, with negligible contributions from matter and radiation. Under these conditions

$$p = -\rho = -\frac{\Lambda}{8\pi G}\,. \tag{1.30}$$

Substituting this into the Friedmann equations we find

$$H^2 = \frac{\Lambda}{3}\,, \tag{1.31}$$

indicating a constant Hubble parameter. The solution for the scale factor takes the form:

$$a(t) = a_0 \exp\left(\sqrt{\frac{\Lambda}{3}}(t - t_0)\right)\,. \tag{1.32}$$

This result illustrates the key dynamical role of the cosmological constant. Since Λ corresponds to an equation-of-state $w = -1$, its energy density remains constant as the Universe expands. As a consequence, once the cosmological constant dominates the energy budget, the expansion becomes exponential, as in the de Sitter solution. This provides the basic mechanism by which a positive cosmological constant can drive the accelerated expansion observed in the late Universe.

1.4 Components of the cosmic fluid

The Friedmann equations show that the expansion of the Universe is determined by its energy content. It is therefore useful to examine the different components of the cosmic fluid and their contribution to the total energy density.
We define the critical energy density as the energy density required for the Universe to be spatially flat, $k = 0$. Imposing this condition in the first Friedmann equation 1.11 gives

$$\rho_{crit} = \frac{3H^2}{8\pi G}\,. \tag{1.33}$$

The critical density provides a reference value against which the actual energy density of the Universe can be compared. In fact, the first Friedmann equation can be rewritten as

$$\frac{k}{a^2} = \frac{8\pi G}{3}(\rho - \rho_c)\,, \tag{1.34}$$

showing that $\rho = \rho_c$ corresponds to a spatially flat Universe, while deviations from the critical density are associated with non-zero spatial curvature.
We then introduce the density parameter

$$\Omega_i = \frac{\rho_i}{\rho_{crit}}\,, \tag{1.35}$$

that is dimensionless and i stands for different components of the universe (radiation, matter, curvature, and dark energy).
We can now rewrite the Friedmann equation 1.11 as

$$H^2(a) = H_0^2 \left[\Omega_{r,0} \left(\frac{a_0}{a} \right)^4 + \Omega_{m,0} \left(\frac{a_0}{a} \right)^3 + \Omega_{k,0} \left(\frac{a_0}{a} \right)^2 + \Omega_{\Lambda,0} \right] . \tag{1.36}$$

By convention, the present-day value of the scale factor is set to unity, so the first Friedmann equation reduces to

$$\frac{H^2(a)}{H_0^2} = \Omega_{r,0} a^{-4} + \Omega_{m,0} a^{-3} + \Omega_{k,0} a^{-2} + \Omega_{\Lambda,0} . \tag{1.37}$$

Evaluating this final expression at the present time, and dropping the subscript 0 for simplicity, we obtain

$$\Omega_r + \Omega_m + \Omega_k + \Omega_\Lambda = 1 . \tag{1.38}$$

Planck satellite constraints at 68% confidence level the values for the reduced Hubble parameter h and for physical density parameters ω_i [27]:

$$h = \frac{H_0}{100\,\mathrm{km/s\,Mpc}} = 0.674 \pm 0.005 , \tag{1.39}$$

$$\omega_b = \Omega_b h^2 = 0.0224 \pm 0.0001 , \tag{1.40}$$

$$\omega_{cdm} = \Omega_{cdm} h^2 = 0.1200 \pm 0.0012 , \tag{1.41}$$

$$\omega_k = \Omega_k h^2 = 0.0007 \pm 0.0019 , \tag{1.42}$$

$$\Omega_\Lambda = 0.6847 \pm 0.0073 , \tag{1.43}$$

where the matter component is conventionally divided into two contributions: baryonic matter, composed of the ordinary matter described by the Standard Model and responsible for the formation of stars, planets, and galaxies, and cold dark matter (CDM), a non-luminous component that interacts predominantly through gravity and plays a fundamental role in the formation of Large-Scale Structure.
Dark matter is commonly classified as cold or hot according to the thermal velocities of its constituent particles at the time of decoupling. CDM consists of particles that were already non-relativistic when they decoupled from the thermal bath. In contrast, hot dark matter (HDM) is composed of particles that decoupled when they were ultrarelativistic. This distinction is crucial because cold dark matter clusters gravitationally, forming the seeds of Large-Scale Structure in the universe, while hot dark matter, due to its high velocities, cannot efficiently form such structures [30]. Consequently, the standard ΛCDM model, the same model used by Planck satellite, assumes that the dominant dark matter component is cold. This assumption is strongly supported by observations, as it provides an excellent description of the large-scale distribution of galaxies and of the anisotropies of the Cosmic Microwave Background.

1.5 Thermal history of the Universe

The Friedmann equations describe the evolution of the scale factor once the matter content of the Universe is specified through its equation of state. However, a complete description of the Universe requires more than its dynamical evolution: it also demands

an understanding of its thermal history.
As the Universe expands, its temperature decreases, giving rise to a sequence of fundamental physical processes such as particle decoupling, primordial nucleosynthesis, and recombination. Together, these processes shaped the composition of the Universe and left observable imprints, like the light-element abundances and the Cosmic Microwave Background, providing a direct connection between the physics of the early Universe and present cosmological observations.
Let's consider the phase space distribution function $f(\mathbf{x}, \mathbf{p}, t, \mu)$, where μ is the chemical potential. If we assume homogeneity, isotropy and use the relation $E = \sqrt{p^2 + m^2}$, the distribution function reduces to a function of only energy, temperature and chemical potential: $f(E, T, \mu)$.
From the phase space distribution we can define the number density $n(T, \mu)$, the energy density $\rho(T, \mu)$ and the pressure $p(T, \mu)$:

$$n(T, \mu) = \frac{g}{(2\pi)^3} \int \mathrm{d}^3 q\, f(E, T, \mu)\,, \tag{1.44}$$

$$\rho(T, \mu) = \frac{g}{(2\pi)^3} \int \mathrm{d}^3 q\, E(q) f(E, T, \mu)\,, \tag{1.45}$$

$$p(T, \mu) = \frac{g}{(2\pi)^3} \int \mathrm{d}^3 q\, \frac{q^2}{3E(q)} f(E, T, \mu)\,. \tag{1.46}$$

where $E(q)$ is the energy per particle and g is the number of helicity states, that is related to the number of internal degrees of freedom for a particle species.
The phase space distribution function is in general an arbitrary function that solves the Boltzmann equation, but, if we consider thermal equilibrium, it takes the form

$$f(E, T, \mu) = \frac{1}{e^{\frac{E-\mu}{T}} \pm 1}\,, \tag{1.47}$$

this is the Bose-Einstein distribution if we consider bosons (-1) or the Fermi Dirac distribution if we consider fermions $(+1)$. If we consider photons instead it reduces to the Planck distribution

$$f(E, T, \mu) = \frac{1}{e^{\frac{E}{T}} - 1}\,, \tag{1.48}$$

since the chemical potential for photons is 0.
Solving the integrals both in the non-relativistic and relativistic limits we find the results in table 1.5.
The thermal history of the early Universe is closely tied to its expansion. As the Universe expands, the temperature of the cosmic plasma decreases, while evolving according to the laws of thermodynamics. Assuming that the expansion is adiabatic, the entropy contained in a comoving volume is conserved, implying that, as long as the number of relativistic degrees of freedom remains constant, the temperature scales with the scale factor as [1]

$$T \sim \frac{1}{a} = 1 + z\,, \tag{1.49}$$

where we used the definition of the cosmic redshift:

$$1 + z = \frac{a_0}{a}\,, \tag{1.50}$$

[1] More generally, entropy conservation implies $g_{*S} T^3 a^3 = \text{const.}$, where g_{*S} denotes the effective number of relativistic degrees of freedom contributing to the entropy density of the thermal plasma. Since g_{*S} is constant away from particle annihilation epochs, one recovers the scaling $T \propto a^{-1}$.

Ultrarelativistic ($T \gg m$)	Non-relativistic ($T \ll m$)
$n = g\,\dfrac{\zeta(3)}{\pi^2}\,T^3 \begin{cases} 1\,, & \text{bosons} \\ \frac{3}{4}\,, & \text{fermions} \end{cases}$	$n = g\left(\dfrac{mT}{2\pi}\right)^{3/2} e^{-m/T}$
$\rho = g\,\dfrac{\pi^2}{30}\,T^4 \begin{cases} 1\,, & \text{bosons} \\ \frac{7}{8}\,, & \text{fermions} \end{cases}$	$\rho = mn$
$P = \dfrac{1}{3}\rho$	$P = nT$

Table 1.1: Thermodynamic quantities for particles in thermal equilibrium in the ultra-relativistic and non-relativistic limits.

and the fact that the present-day scale factor is conventionally normalized to a 1. From these considerations follows that:

$$T(z) = T_0(1+z)\,. \tag{1.51}$$

Therefore, looking back to higher redshifts corresponds to probing an increasingly hotter and denser Universe. This simple relation underlies the Hot Big Bang model and provides the framework for understanding the sequence of physical processes that shaped the early Universe. The main stages of this thermal history are briefly summarized below.

Baryogenesis

During the radiation-dominated era, at extremely high temperatures, the Universe is believed to have developed a slight asymmetry between matter and antimatter. Although the mechanism responsible for this asymmetry remains unknown, it resulted in a tiny excess of baryons over antibaryons. As the Universe continued to cool, matter and antimatter annihilated almost completely, leaving behind the small amount of baryonic matter that constitutes the visible matter observed today.

Neutrino decoupling

When the temperature dropped to approximately $T \sim 1\,\mathrm{MeV}$, neutrinos decoupled and started propagating freely through the expanding Universe.

Big Bang Nucleosynthesis

At $T \sim 100\,\mathrm{keV}$, nuclear reactions became efficient so protons and neutrons fused to form light nuclei such as hydrogen, helium, and lithium, leaving an imprint on the chemical composition of the universe.

Radiation-matter equality

Approximately 5×10^4 years after the Big Bang, the energy density of matter became equal to that of radiation. From this epoch onward, the expansion of the Universe was matter dominated, allowing density perturbations to grow efficiently under the action of gravity and paving the way for the subsequent formation of cosmic structures.

Recombination

Approximately 3.8×10^5 years after the Big Bang, the temperature fell below about $0.3\,\mathrm{eV}$, allowing free electrons and protons to combine into neutral hydrogen atoms.

Cosmic Microwave Background

Following recombination, photons decoupled from matter and began to propagate almost freely throughout the Universe, defining the last scattering surface. These relic photons constitute the Cosmic Microwave Background (CMB), that provides a direct snapshot of the Universe at the epoch of last scattering. Tiny anisotropies in the CMB encode information about the primordial density fluctuations that later evolved into cosmic structures.

Large-Scale Structure formation

The small density perturbations present in the early Universe grew through gravitational instability after matter became dynamically dominant. Over billions of years, these perturbations evolved into galaxies, galaxy clusters, and the cosmic web observed today, tracing the large-scale distribution of matter in the Universe.

Dark energy domination

At relatively recent times, the energy density of dark energy exceeded that of matter, marking the onset of the present dark-energy-dominated era. The resulting accelerated expansion of the Universe slows down the growth of Large-Scale Structure, although it does not halt it completely. Today, dark energy accounts for about 70% of the total energy density of the Universe and governs its large-scale dynamics.

1.6 Cosmological problems

Despite its remarkable success, the standard Hot Big Bang model faces several conceptual problems when extrapolated to the earliest stages of cosmic evolution. In particular, the horizon and flatness problems point to the need for an explanation of the observed large-scale homogeneity and near-flatness of the Universe.
In this section, we briefly review the two problems of the standard Hot Big Bang model. In the following section, we will discuss how these issues can be addressed by the inflationary paradigm.

1.6.1 Horizon problem

Before addressing the horizon problem, we first review the relevant notions of distance in an expanding Universe. In particular, the concepts of comoving distance and particle

horizon will be introduced, as they provide the appropriate framework for discussing the causal structure of the early Universe.
The light cone of an event is defined by $\mathrm{d}s^2 = 0$ where $\mathrm{d}s^2$ is the space-time interval between two infinitesimally separated events. The light cone separates the spacetime into regions that can be causally connected to a given event from those that cannot. Events lying on the light cone or inside can be connected by signals traveling at or below the speed of light, while events outside the light cone can not be causally connected.
Starting from FLRW metric and imposing isotropy we can set $\mathrm{d}\Omega^2 = 0$, so that evaluating $\mathrm{d}s^2 = 0$ we get

$$\frac{\mathrm{d}t}{a(t)} = \frac{\mathrm{d}r}{\sqrt{1-kr^2}} \,. \tag{1.52}$$

This expression can be integrated to find the comoving particle horizon

$$l_H(t) \equiv \int_0^t \frac{\mathrm{d}t'}{a(t')} \,. \tag{1.53}$$

The corresponding physical distance, the particle horizon,

$$d_H(t) = a(t) \int_0^t \frac{\mathrm{d}t'}{a(t')} \,, \tag{1.54}$$

defines the radius of a sphere centered in an observed O that contains all the spatial points which are causally connected with to observer.
Recalling the expression for $a(t)$ 1.20, we see that d_H exists finite only if $\frac{2}{3(1+w)} < 1$, that is $w > -1/3$. In this case, the expression for the particle horizon takes the form

$$d_H(t) = \frac{3(1+w)}{1+3w} t \,. \tag{1.55}$$

The condition $w > -1/3$ also implies, looking at the second Friedmann equation 1.12, that we have a finite particle horizon if and only if we have a decelerating expansion of the universe $\ddot{a} < 0$.
Now, we define the Hubble radius as the physical distance traveled by a photon in a Hubble time τ_H, the typical expansion time of the Universe:

$$R_c = \tau_H = \frac{1}{H(t)} \,, \tag{1.56}$$

this defines the maximum distance from which light or signals could have traveled to us since the beginning of the Universe.
We recover the corresponding comoving Hubble radius as usual using 1.2:

$$r_H(t) = \frac{R_H}{a(t)} = \frac{1}{\dot{a}} \,. \tag{1.57}$$

With these concepts in place, we are now equipped to examine the horizon problem and understand why it poses a challenge to the standard Hot Big Bang model.
Consider a region of the Universe characterized by a comoving length scale λ. A region can be considered causally connected when its characteristic scale is smaller than the comoving Hubble radius, r_H. Conversely, regions with $\lambda > r_H$ lie beyond this characteristic scale and cannot efficiently exchange information. As the Universe expands,

the Hubble radius evolves, so different regions of the Universe can become causally connected at different stages of cosmic evolution.
This picture leads to a fundamental puzzle when confronted with observations of the CMB. The CMB is remarkably homogeneous and isotropic on large scales, with temperature fluctuations of only $\Delta T/T \sim 10^{-5}$. In particular, regions sky separated by distances larger than the Hubble radius at recombination appear to have nearly identical temperatures, despite having been apparently causally disconnected. How could such distant regions have reached such a high degree of homogeneity if they had no opportunity to exchange information? This apparent contradiction constitutes the horizon problem.

1.6.2 Flatness problem

Now, let us consider the flatness problem. As we can see from 1.42, our Universe is compatible with flatness. Using the definitions introduced in section 1.4 we can rewrite the first Friedmann equation as

$$\Omega(t) - 1 = \frac{k}{a^2H^2} = kr_H^2(t)\,. \tag{1.58}$$

This equation tells us that if the universe is perfectly flat, then $\Omega = 1$. However, for any deviation of Ω from unity, the difference increases proportionally to $r_H(t)$. During matter or radiation dominated, the comoving Hubble radius grows over time. Consequently, when we trace backward in time, $\Omega(t)$ becomes progressively closer to 1.
This means that the value of Ω at the beginning of the Universe was closer to 1 than it is today ($\omega_k = 0.0007 \pm 0.0019$ at 68% confidence level [27]). Taking as a reference the Planck time $t_{Pl} \sim 10^{-44}\,$s, a straightforward calculation leads to

$$|\Omega(t_{Pl}) - 1| \sim |\Omega_0 - 1| \cdot 10^{-60}\,. \tag{1.59}$$

This extreme sensitivity to the initial value of Ω is known as the flatness problem and represents a fine-tuning problem.

1.7 Inflation

1.7.1 Solution to the cosmological problems

The inflationary paradigm provides a solution to these problems.
Inflation is a period of accelerated expansion in the Early Universe, that means $\ddot{a}(t) > 0$. If we recall the definition of comoving Hubble radius we can easily see how this condition reflects on r_H:

$$\dot{r}_H = -\frac{\ddot{a}}{a^2H^2} < 0\,. \tag{1.60}$$

This means that inflation is a period in which the comoving Hubble radius decreases over time (depicted in figure 1.7.1): such a behaviour ensures that regions of the universe which have only recently come into causal contact were actually causally connected during the very Early Universe. This early connection solves the horizon problem by allowing particles in those regions to interact and exchange information,

leading to the homogeneous and isotropic conditions we observe today.
Thus, to solve the horizon problem we require

$$r_H(t_0) \leq r_H(t_i)\,, \tag{1.61}$$

that looking at 1.12 is equivalent to having a negative pressure $P < -\rho/3$. The case of the cosmological constant ($w = -1$) is compatible with this condition: during inflation we have an approximately constant Hubble factor and an exponentially growing in time scale factor $a(t) \propto e^{Ht}$.
The condition 1.61 can be translated in a constraint on the minimum duration of inflation:

$$N_{min} \in [50, 70]\,, \tag{1.62}$$

where we have defined the number of e-folds

$$N_{min} = \int_{t_i}^{t_f} H\mathrm{d}t = \log\left(\frac{a(t_f)}{a(t_i)}\right)\,, \tag{1.63}$$

t_i and t_f are the time when inflation correspondingly begins and ends. The exact prediction for the minimum number of e-folds depends on the model and on the post-inflationary thermal history.
Note that condition 1.62 is equivalent to requiring an expansion of the Universe during the inflationary period of a factor $a(t_f)/a(t_i) \sim 10^{26} - 10^{30}$.
Inflation also solves the flatness problem: by allowing r_H to decrease we also allow $|\Omega(t) - 1|$ to decrease:

$$|\Omega(t) - 1| \propto r_H^2 \propto e^{-2Ht}\,. \tag{1.64}$$

Regardless of the initial deviation, inflation exponentially suppresses the difference from 1.
The flatness problem can be solved if we require the density parameter at the onset of inflation to be further from 1 compared to today:

$$\frac{\Omega_i^{-1} - 1}{\Omega_0^{-1} - 1} \geq 1\,. \tag{1.65}$$

It can be shown that the number of e-folds required to solve the flatness problem is of the same order as that required to solve the horizon problem. This result is not surprising, given the dependence of the density parameter on the Hubble horizon.

1.7.2 Inflationary paradigm

In the previous chapter we defined inflation as a period of accelerated expansion in the very Early Universe. We have also seen that the condition $\ddot{a}(t) > 0$ is equivalent to requiring an equation of state $w < -1/3$.
The simplest way to achieve such an accelerated expansion is through a cosmological constant[2] ($w = -1$); we have already seen that this leads to $a \propto e^{Ht}$ a constant Hubble parameter $H = \sqrt{\Lambda/3}$.
In principle, accelerated expansion driven by a cosmological constant would persist indefinitely, preventing the Universe from transitioning to the radiation- and matter-dominated epochs. For inflation to be physically viable, the accelerated expansion must be driven by a mechanism that temporarily mimics the behaviour of a cosmological constant, while allowing inflation to end after a finite period.

[2] In principle there are other possibilities: the so called power law inflation ($-1 < w < -1/3 \rightarrow a(t) \propto t^{\frac{2}{3(1+w)}}$, or the pole inflation ($w < -1$). In this work we focus on the case $w = -1$.

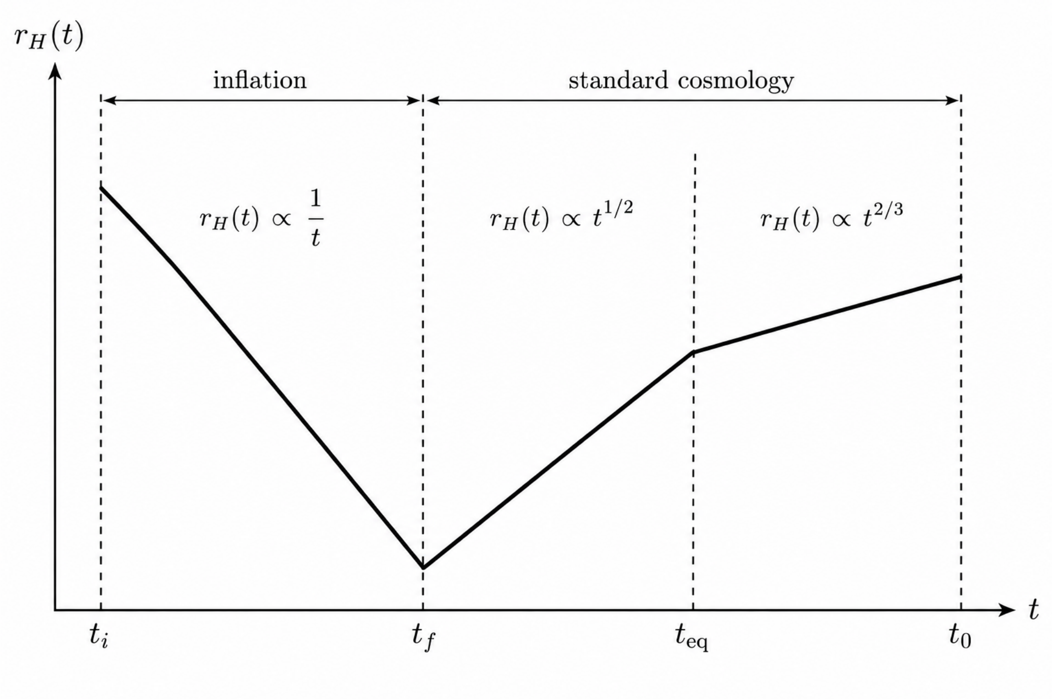

Figure 1.1: Schematic evolution of the comoving Hubble radius r_H during inflation and the subsequent standard cosmological evolution. During inflation $r_H \propto t^{-1}$, while during radiation and matter domination $r_H \propto t^{1/2}$ and $r_H \propto t^{2/3}$ respectively.

The inflaton

A natural realization is provided by a minimally coupled scalar field ϕ, called the inflaton, consistent with the homogeneity and isotropy of the background, evolving in a potential $V(\phi)$.
The dynamics of a minimally coupled scalar field is described by the Lagrangian:

$$\mathcal{L} = \frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi)\,, \tag{1.66}$$

from which the corresponding energy-momentum tensor can be derived:

$$T^\phi_{\mu\nu} = \phi_{,\mu}\, \phi_{,\nu} + g_{\mu\nu} \left[-\frac{1}{2} g^{\alpha\beta} \phi_{,\alpha}\, \phi_{,\beta} - V(\phi) \right]\,, \tag{1.67}$$

where $\phi_{,\mu} \equiv \partial_\mu \phi$.
The equation of motion for ϕ is the Klein-Gordon equation:

$$\Box \phi(\mathbf{x}, t) = \frac{\partial V(\phi)}{\partial \phi(\mathbf{x}, t)}\,, \tag{1.68}$$

where, in curved space, the d'Alembert operator is

$$\Box \phi = \frac{1}{\sqrt{-g}} \left(g^{\mu\nu} \sqrt{-g} \phi_{,\mu} \right)_{,\nu}\,. \tag{1.69}$$

The resulting equation of motion is:

$$\ddot{\phi}(\mathbf{x}, t) + 3H\dot{\phi}(\mathbf{x}, t) - \frac{\nabla^2 \phi(\mathbf{x}, t)}{a^2(t)} = -\frac{\partial V(\phi)}{\partial \phi}\,. \tag{1.70}$$

The term $3H\dot{\phi}(\mathbf{x},t)$ represents the expansion of the universe and acts as a friction term.
The assumption of homogeneity and isotropy applies to the background Universe and allows us to separate the scalar field into a homogeneous background component and small fluctuations around it. We can therefore write:

$$\phi(\mathbf{x},t) = \phi_0(t) + \delta\phi(\mathbf{x},t)\,, \tag{1.71}$$

where $\phi_0(t)$ describes the homogeneous background, while $\delta\phi(\mathbf{x},t)$ represents the inhomogeneous fluctuations. This decomposition is useful because the background field determines the dynamics of the expanding Universe, while the fluctuations describe the small departures from homogeneity and isotropy that can later seed the primordial density perturbations.
The approximation 1.71 holds if the perturbation is significantly smaller than the background value, which can be expressed as:

$$\langle\delta\phi^2\rangle \ll \phi_0^2\,. \tag{1.72}$$

The smallness of the primordial perturbations is consistent with the observed high degree of homogeneity and isotropy of the CMB.

Background dynamics

The background dynamics describes the accelerated expansion of the Universe.
The stress tenor takes the form 1.9, and if we compare it with 1.67 we find:

$$T^0_0 = -\rho_\phi(t) = -\left(\frac{1}{2}\dot{\phi}^2 + V(\phi)\right)\,, \tag{1.73}$$

$$T^i_j = p_\phi(t)\delta^i_j = \frac{1}{2}\dot{\phi}^2 - V(\phi)\,. \tag{1.74}$$

The class of models of inflation we are interested in is called single field models of slow roll inflation. These models are defined by requiring that $1/2\dot{\phi}^2 \ll V(\phi)$, this condition, often referred to as 1^{st} slow roll condition, ensures that the scalar field is evolving slowly.
We also have a 2^{nd} slow roll condition that is motivated by an analogy with viscous motion. For a particle moving in a viscous medium, the equation of motion contains an acceleration term and a friction term. In the regime of strong damping, the acceleration becomes negligible compared to the friction force. Analogously, for the inflaton, the term $3H\dot{\phi}$ plays the role of an effective friction term induced by the expansion of the Universe. We therefore require the acceleration term to be negligible compared to this friction term,

$$\ddot{\phi}_0 \ll 3H\dot{\phi}_0\,. \tag{1.75}$$

If we use these slow roll conditions, the first Friedmann equation 1.11 becomes:

$$H^2 \simeq \frac{8\pi G}{3}V(\phi)_0\,, \tag{1.76}$$

since from the 1^{st} slow roll condition we have $\rho_\phi \sim V(\phi) \sim -p_\phi$.
The equation of motion for the background ϕ_0 takes the form:

$$\dot{\phi}_0(t) \sim -\frac{V'(\phi)_0}{3H} \tag{1.77}$$

where, starting from 1.70, we used the slow roll conditions and the fact that ϕ_0 depends only on time.

Slow roll parameters

Concerning the duration of inflation we define the two slow roll parameters:

$$\epsilon = -\frac{\dot{H}}{H^2}\,, \tag{1.78}$$

$$\eta = -\frac{\ddot{\phi}}{H\dot{\phi}}\,. \tag{1.79}$$

For inflation to occur we need $\ddot{a} > 0$:

$$\frac{\ddot{a}}{a} = H^2 + \dot{H} = H^2(1-\epsilon)\,, \tag{1.80}$$

meaning that as long as $\epsilon < 1$ we have inflation.
While ϵ gives us the condition to have inflation, η is directly related to how much time it lasts. We require $|\eta| \ll 1$; if, on contrary, we take $\eta = 0$, ϵ would be constant and inflation would never end.
With some straightforward calculation we can relate the condition on ϵ to a condition on the potential $V(\phi)$:

$$\epsilon = \frac{1}{16\pi G}\left(\frac{V'(\phi)}{V(\phi)}\right)^2 \ll 1 \quad \longrightarrow \quad V'(\phi) \ll V(\phi)\,, \tag{1.81}$$

that implies that the slope of the potential is very shallow, ensuring that the inflaton field rolls slowly down the potential $V(\phi)$ maintaining a nearly constant value of the potential energy.

Reheating

To recover the well established prediction of the Big Bang model, after inflation the radiation dominated universe should start.
At the end of inflation, the energy density of the Universe is still dominated by the inflaton field, while the expansion has diluted any previously existing radiation. Reheating is the transitional period in which the universe is reheated and which comes just before the radiation era.
When ϵ or $|\eta|$ approach 1 the potential that drives inflation starts to deviate from flatness: as depicted in 1.2 the inflaton starts rolling down towards the minimum of the potential. Once the inflaton reaches the minimum, the potential induces oscillations in the scalar field with a frequency $\omega^2 \gg H^2$: during these oscillations, the scalar field decays into light, relativistic particles, leading to a radiation dominated era.

Quantum fluctuations of the inflaton field

Having described the evolution of the homogeneous background during inflation, we now turn our attention to the small perturbations around it. Although these perturbations are initially tiny, they play a fundamental role in the subsequent evolution of the Universe: amplified during inflation and later evolved by gravity, they provide the

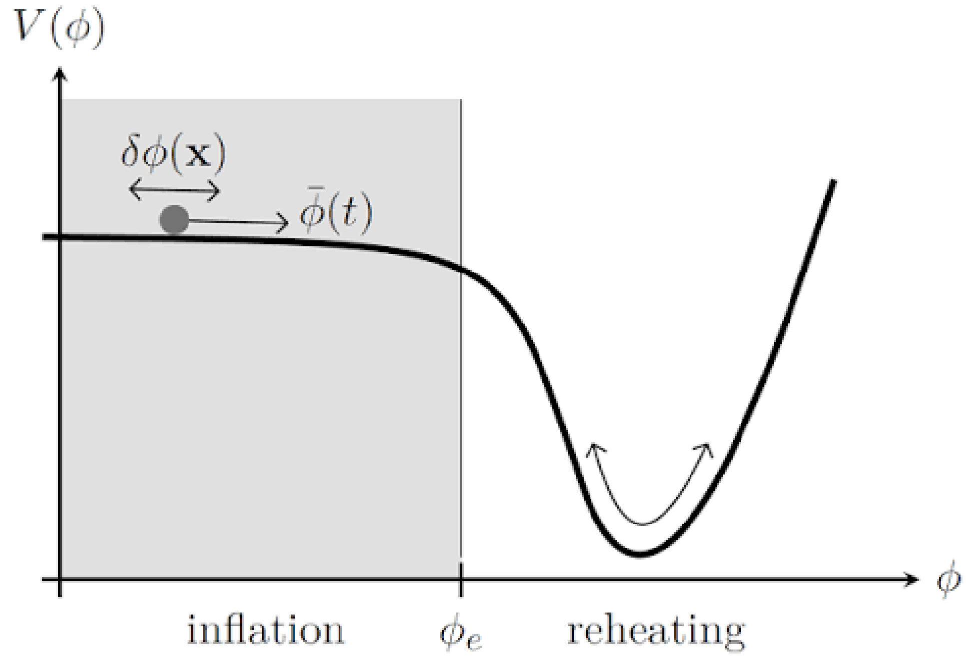

Figure 1.2: A schematic representation of the inflationary potential $V(\phi)$ showing the inflaton's slow roll and the subsequent reheating phase. What in the picture is called $\bar{\phi}(t)$ is our homogeneous background $\phi_0(t)$.

seeds from which the Large-Scale Structure of the Universe emerges.
At the classical level, a perfectly homogeneous inflaton field would remain homogeneous and would not provide a mechanism for generating the primordial inhomogeneities observed in the Universe. The origin of these perturbations is instead quantum mechanical. Even when the inflaton is in its vacuum state, quantum mechanics implies unavoidable fluctuations around its expectation value. We can therefore say that $\phi_0(t) \equiv \langle 0|\phi(\mathbf{x},t)|0\rangle$ [3] and $\delta\phi(\mathbf{x},t)$ represents the quantum fluctuations around the background.
Starting from 1.70 and plugging in the equation of motion for the background we find that the equation of motion for the perturbations is:

$$\delta\ddot{\phi}(\mathbf{x},t) + 3H\delta\dot{\phi}(\mathbf{x},t) - \frac{\nabla^2\delta\phi(\mathbf{x},t)}{a^2(t)} = -\frac{\partial^2 V(\phi)}{\partial\phi^2}\delta\phi(\mathbf{x},t)\,, \tag{1.82}$$

where ∇^2 is the spatial Laplacian with respect to the comoving coordinates.
Taking the equation of motion for the background and derive it, we get:

$$\ddot{\psi} + 3H\dot{\psi} = -\frac{\partial^2 V}{\partial\phi^2}\psi\,, \tag{1.83}$$

with $\psi = \dot{\phi}_0$.
If we now compare 1.82 and 1.83 we notice that they differ only by the spatial Laplacian term. But on large scale this term is negligible: it is evident in Fourier space where the Laplacian is proportional to k^2 (recalling that at large scales $k \ll aH$). Therefore, the

[3] The spatial dependence vanishes because of the homogeneity of the vacuum state

perturbations evolve in the same way as the time derivative of the background field. To analyze the behaviour of these perturbations we go to Fourier space, where the equation of motions 1.82 takes the form:

$$\delta\ddot{\phi}_{\mathbf{k}} + 3H\delta\dot{\phi}_{\mathbf{k}} + \frac{\mathbf{k}^2}{a^2}\delta\phi_{\mathbf{k}} = -V''\delta\phi_{\mathbf{k}}\,. \tag{1.84}$$

The behaviour of $\delta\phi_{\mathbf{k}}$ varies widely among different scales:

- Sub-horizon regime ($k \gg aH$): it is possible to show that in this regime the equation of motion reduces to the one fo an harmonic oscillator with a time dependent frequency. The solution is an oscillating function with amplitude $1/a\sqrt{2k}$,

- Super-horizon regime ($k \ll aH$): in this regime the Laplace operator becomes negligible. The solution shows that perturbations become constant over time.

The quantitative characterization of these fluctuations in terms of their power spectrum will be addressed in Chapter 2, where the tools of cosmological perturbation theory are introduced.

1.8 Large-Scale Structure formation

Quantum fluctuations of the inflaton field provide the initial seeds for the formation of cosmic structures. During inflation, the expansion of the Universe stretches these fluctuations to scales larger than the Hubble horizon. On these super-horizon scales, the corresponding perturbations become effectively frozen, preserving the information about their primordial amplitude. These fluctuations can be described in terms of the curvature perturbation ζ, which characterizes perturbations in the geometry of spacetime. After inflation, as the relevant scales re-enter the Hubble horizon, these primordial perturbations are converted into density fluctuations, which subsequently evolve under gravitational instability and eventually give rise to the Large-Scale Structure of the Universe. Their imprints are also observed as temperature anisotropies in the Cosmic Microwave Background. Schematically, the connection between the different stages can be summarized as

$$\delta\phi \;\rightarrow\; \zeta \;\rightarrow\; \frac{\delta\rho}{\rho} \;\rightarrow\; \frac{\delta T}{T}\,. \tag{1.85}$$

A complete description of this process requires the treatment of cosmological perturbations within the framework of General Relativity and is beyond the scope of this work; for a detailed discussion please refer to [31].

1.8.1 The cosmic web

As these density perturbations evolve under gravity, matter becomes increasingly inhomogeneously distributed on large scales. While the initial perturbations are well described by linear theory, their subsequent evolution becomes increasingly non-linear as structures collapse, giving rise to the complex network known as the cosmic web.

The cosmic web can be described in terms of four distinct environments of the matter density field: voids, walls, filaments, and nodes. Voids are large underdense regions, while walls are sheet-like structures surrounding them. Filaments are elongated structures connecting the densest regions of the web, known as nodes, where galaxy groups and clusters are preferentially found. For further details on the cosmic web please refer to [32].
This web-like pattern primarily reflects the underlying dark matter distribution, which provides the gravitational framework for the formation and evolution of cosmic structures. Since dark matter cannot be observed directly, galaxies are used as tracers of its underlying distribution. However, galaxies do not trace the dark matter field perfectly: their clustering is biased with respect to the underlying matter distribution.

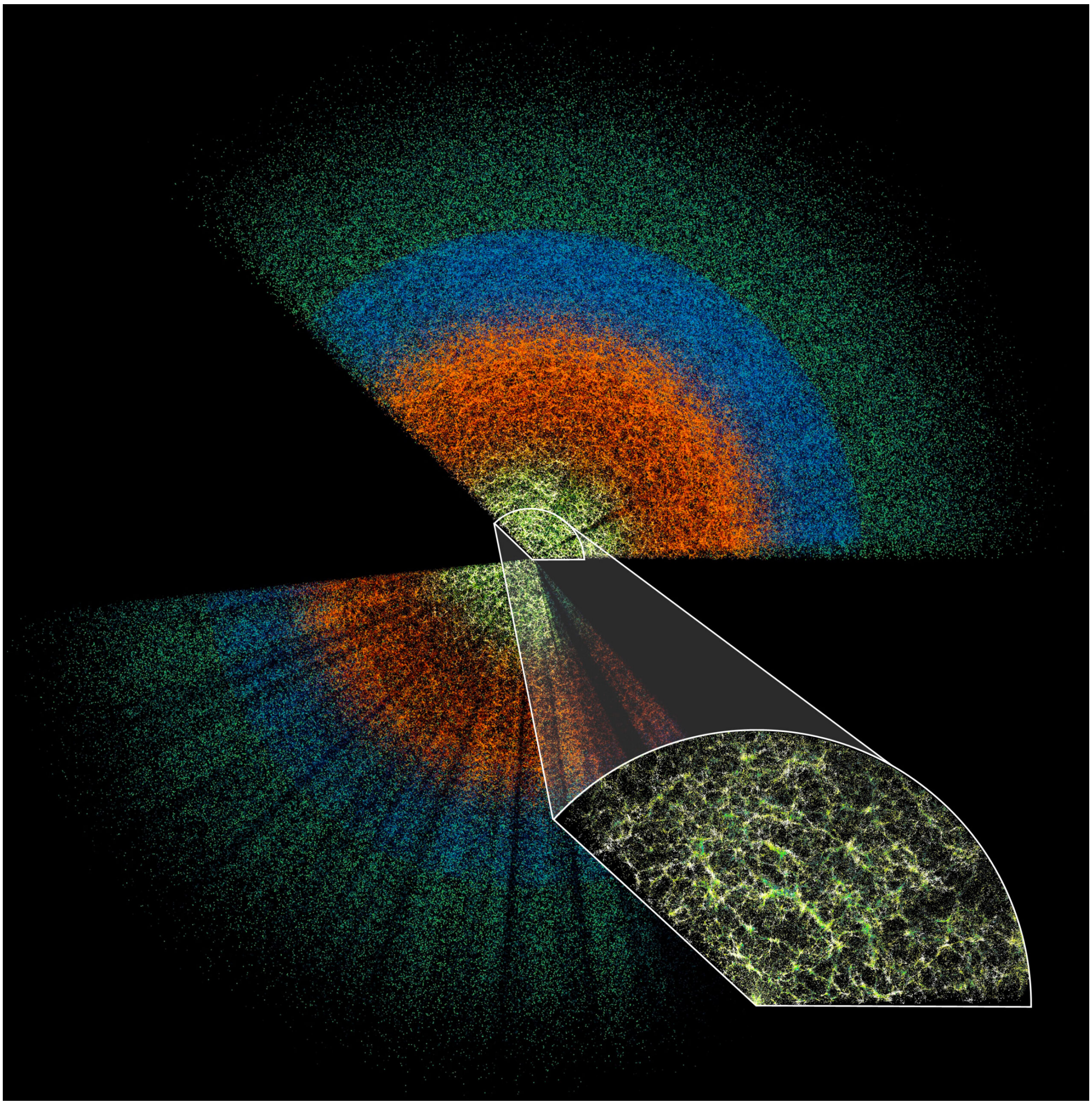

Figure 1.3: A slice of the three-dimensional distribution of galaxies mapped by the Dark Energy Spectroscopic Instrument (DESI). The map illustrates the Large-Scale Structure of the Universe, characterized by filaments, walls, nodes and voids. Source: DESI Collaboration/DOE/KPNO/NOIRLab/NSF/AURA/C. Lamman.

Chapter 2

The large scale structure theoretical framework

2.1 Gravitational instability

The Large-Scale Structure seen in galaxy surveys is believed to be the result of gravitational amplification of small primordial fluctuations due to gravitational interaction of collisionless cold dark matter (CDM) particles in an expanding universe. Although the nature of dark matter has not yet been identified, all candidates for CDM particles are hypothesized to be non-relativistic and extremely light compared to the mass scale of typical galaxies, with expected number densities of at least 10^{50} particles per Mpc^3. In this framework, the vast number of particles ensures that discreteness effects, such as two-body relaxation, are negligible, enabling the treatment of dark matter as a collisionless fluid. The behavior of such a fluid is described by the Vlasov equation, which, together with the Poisson equation for the gravitational potential, forms the foundations of gravitational instability theory.
Since CDM particles are non-relativistic, at scales much smaller than the Hubble radius the equations of motion are those of Newtonian gravity. In this regime, the equations of motion for CDM particles are adapted to account for the expanding universe through a redefinition of variables for position, momentum, and gravitational potential.
Throughout this chapter we will discuss how perturbation theory (PT) can be used to understand the physics of gravitational instability, starting from the fundamental equations that govern CDM dynamics.

2.1.1 The Vlasov equation

This chapter closely follows [33].
Consider a set of particles of mass m that interact in an expanding universe only via gravity. For a particle of velocity $\mathbf{v}$ at position $\mathbf{r}$, the equation of motion reads

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = Gm \sum_i \frac{\mathbf{r}_i - \mathbf{r}}{|\mathbf{r}_i - \mathbf{r}|^3}\,, \tag{2.1}$$

where the summation is made over all other particles at positions $\mathbf{r}_i$.
In the limit of a large number of particles, this equation can be rewritten in terms of

a smooth gravitational potential

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\frac{\partial \phi}{\partial \mathbf{r}}\,, \tag{2.2}$$

here ϕ is the Newtonian potential induced by the local mass density $\rho(\mathbf{r})$

$$\phi(\mathbf{r}) = G\int \mathrm{d}^3\mathbf{r}'\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,. \tag{2.3}$$

To incorporate cosmic expansion we use the conformal time τ, related to cosmic time by $\mathrm{d}t = a(\tau)\mathrm{d}\tau$. The scale factor $a(\tau)$ obeys the Friedmann equations for an homogeneous and isotropic background, that can be written as:

$$\frac{\partial \mathcal{H}}{\partial \tau} = -\frac{\Omega_m(\tau)}{2}\mathcal{H}^2(\tau) + \frac{\Lambda}{3}a^2(\tau) \equiv \Big(\Omega_\Lambda(\tau) - \frac{\Omega_m(\tau)}{2}\Big)\mathcal{H}^2\,, \tag{2.4}$$

$$\big(\Omega_{tot} - 1\big)\mathcal{H}^2 = k\,, \tag{2.5}$$

where $\mathcal{H} = \mathrm{d}\mathbf{ln}a/\mathrm{d}\tau = Ha$ is the conformal expansion rate, H is the Hubble constant, Ω_m is the ratio of matter density to critical density, Λ is the cosmological constant and $k = -1, 0, 1$ for $\Omega_{tot} < 1$, $\Omega_{tot} = 1$ and $\Omega_{tot} > 1$ respectively ($\Omega_{tot} = \Omega_m + \Omega_\Lambda$).
When studying gravitational instability a key quantity is the density contrast $\delta(\mathbf{x})$

$$\rho(\mathbf{x},\tau) = \bar{\rho}(\tau)[1 + \delta(\mathbf{x},\tau)]\,, \tag{2.6}$$

where $\bar{\rho}(\tau)$ is the homogeneous background density.
We then define the peculiar velocity $\mathbf{u}(\mathbf{x},\tau)$ bu splitting the total velocity $\mathbf{v}(\mathbf{x},\tau)$

$$\mathbf{v}(\mathbf{x},\tau) = \mathcal{H}\mathbf{x} + \mathbf{u}(\mathbf{x},\tau)\,. \tag{2.7}$$

We can define also the cosmological gravitational potential Φ by splitting the Newtonian gravitational potential

$$\phi(\mathbf{x},\tau) = -\frac{1}{2}\frac{\partial \mathcal{H}}{\partial \tau}x^2 + \Phi(\mathbf{x},\tau)\,, \tag{2.8}$$

so that the latter is sourced only by fluctuations in the matter distribution.
Finally, in comoving coordinates, the Poisson equation reads

$$\nabla^2\Phi(\mathbf{x},\tau) = 4\pi G\bar{\rho}a^2\delta(\mathbf{x},\tau) = \frac{3}{2}\Omega_m(\tau)\mathcal{H}^2(\tau)\delta(\mathbf{x},\tau)\,, \tag{2.9}$$

since $\nabla_r^2 = \nabla_x^2/a^2$. In the following we will only use comoving coordinates so all spatial derivatives refer to the coordinate $\mathbf{x}$.
The equation of motion 2.2 then reads

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}\tau} = -am\nabla\Phi(\mathbf{x})\,, \tag{2.10}$$

where $\mathbf{p} = am\mathbf{u}$ is the canonical momentum in comoving coordinates.
Now, if we consider the particle number density in phase space $f(\mathbf{x},\mathbf{p},\tau)$, the conservation of phase-space density is expressed by the Vlasov equation:

$$\begin{aligned}\frac{\mathrm{d}f}{\mathrm{d}\tau} &= \frac{\partial f}{\partial \tau} + \frac{\partial f}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial \tau} + \frac{\partial f}{\partial \mathbf{p}}\frac{\partial \mathbf{p}}{\partial \tau} \\ &= \frac{\partial f}{\partial \tau} + \frac{\mathbf{p}}{ma}\cdot\nabla f - am\nabla\Phi\cdot\frac{\partial f}{\partial \mathbf{p}} = 0\,.\end{aligned} \tag{2.11}$$

Needless to say, this equation is very difficult to solve, being a non-linear partial differential equation involving multiple variables. The non-linearity arises because Φ depends on $\rho(\mathbf{x},\tau)$, which is obtained by integrating f over all momenta.

2.1.2 Eulerian Dynamics

In practice, we are usually interested in solving the evolution of the spatial distribution rather than the full phase-space dynamics. This is obtained by taking momentum moments of the distribution function. The zeroth order moment simply relates the phase space density to the local mass density field

$$\int \mathrm{d}^3\mathbf{p} f(\mathbf{x},\mathbf{p},\tau) \equiv \rho(\mathbf{x},\tau)\,, \tag{2.12}$$

the first order momentum define the peculiar velocity flow $\mathbf{u}(\mathbf{x},\tau)$

$$\int \mathrm{d}^3\mathbf{p} \frac{\mathbf{p}}{ma} f(\mathbf{x},\mathbf{p},\tau) \equiv \rho(\mathbf{x},\tau)\mathbf{u}(\mathbf{x},\tau)\,, \tag{2.13}$$

and the second order moment

$$\int \mathrm{d}^3\mathbf{p} \frac{p_i p_j}{m^2a^2} f(\mathbf{x},\mathbf{p},\tau) \equiv \rho(\mathbf{x},\tau)\mathbf{u}_i(\mathbf{x},\tau)\mathbf{u}_j(\mathbf{x},\tau) + \sigma_{ij}(\mathbf{x},\tau)\,, \tag{2.14}$$

defines the stress tensor $\sigma_{ij}(\mathbf{x},\tau)$ that describes deviations of individual particle velocity from bulk velocity $\mathbf{u}$ and contributes to the anisotropic stress of the system.
The equation for these fields follow from taking moments of the Vlasov equation. From the zeroth moment,

$$\int \mathrm{d}^3\mathbf{p} \Big(\frac{\partial f}{\partial \tau} + \frac{\mathbf{p}}{ma}\cdot\nabla f - am\nabla\Phi\cdot\frac{\partial f}{\partial \mathbf{p}} \Big)\,, \tag{2.15}$$

and substituting ρ in terms of δ

$$\frac{\partial \delta(\mathbf{x},\tau)}{\partial \tau} + \nabla\cdot\big\{\big[1+\delta(\mathbf{x},\tau)\big]\mathbf{u}(\mathbf{x},\tau)\big\} = 0\,, \tag{2.16}$$

we obtain the conservation of mass, or continuity equation.
Taking the first moment of the Vlasov equation and subtracting $\mathbf{u}(\mathbf{x},\tau)$ times the continuity equation we obtain the Euler equation

$$\frac{\partial \mathbf{u}(\mathbf{x},\tau)}{\partial \tau} + \mathcal{H}(\mathbf{x},\tau)\mathbf{u}(\mathbf{x},\tau) + \mathbf{u}(\mathbf{x},\tau)\cdot\nabla\mathbf{u}(\mathbf{x},\tau) = -\nabla\Phi(\mathbf{x},\tau) - \frac{1}{\rho}\nabla_j(\rho\sigma_{ij})\,, \tag{2.17}$$

which describes momentum conservation.
The first moment of the distribution function, represented by the peculiar velocity $\mathbf{u}(\mathbf{x},\tau)$, is dynamically linked via the Euler equation to the stress tensor σ_{ij}, which quantifies deviations of particle motions from a single coherent flow. To close the hierarchy we need to postulate an equation of state of the cosmological fluid that would represent an ansatz for the stress tensor. The equation of state relies on the assumption that cosmological structure formation is driven by matter with negligible velocity dispersion; from this follows that the first term in 2.14 will be the dominant contribution. Therefore is a good approximation to set $\sigma_{ij} \simeq 0$, at least in the first stages of gravitational instability, when structures did not have time to collapse and virialize. As time goes on, this approximation will break down at progressively larger scales.
We now begin by discussing the solutions of the Poisson equation, the continuity equation, and the Euler equation in the case of a vanishing stress tensor, corresponding to a perfect fluid.

2.1.3 Eulerian perturbation theory

Linear PT

At large scales, where we expect the Universe to become smooth, the fluctuation fields $\delta(\mathbf{x},\tau)$, $\mathbf{u}(\mathbf{x},\tau)$ and $\Phi(\mathbf{x},\tau)$ can be assumed to be small compared to the homogeneous contribution. Therefore, we can linearize Eqs 2.16, 2.17 to obtain

$$\begin{aligned}&\frac{\partial\delta(\mathbf{x},\tau)}{\partial\tau}+\nabla\cdot\mathbf{u}(\mathbf{x},\tau)=0\\&\frac{\partial\mathbf{u}(\mathbf{x},\tau)}{\partial\tau}+\mathcal{H}(\mathbf{x},\tau)\mathbf{u}(\mathbf{x},\tau)=-\nabla\Phi(\mathbf{x},\tau)\,,\end{aligned} \tag{2.18}$$

The velocity field, as any vector field, can be completely described by its divergence $\theta(\mathbf{x},\tau)\equiv\nabla\cdot\mathbf{u}(\mathbf{x},\tau)$ and its vorticity $\mathbf{w}(\mathbf{x},\tau)\equiv\nabla\times\mathbf{u}(\mathbf{x},\tau)$, whose equations of motion follow from 2.18

$$\frac{\partial\theta(\mathbf{x},\tau)}{\partial\tau}+\mathcal{H}(\tau)\theta(\mathbf{x},\tau)+\frac{3}{2}\Omega_m(\tau)\mathcal{H}^2(\tau)\delta(\mathbf{x},\tau)=0 \tag{2.19}$$

$$\frac{\partial\mathbf{w}(\mathbf{x},\tau)}{\partial\tau}+\mathcal{H}(\tau)\mathbf{w}(\mathbf{x},\tau)=0\,, \tag{2.20}$$

where we also used the Poisson equation for Φ. The vorticity evolution readily follows from 2.20: $\mathbf{w}(\tau)\propto a^{-1}$, meaning that in the linear regime any initial vorticity decays due to the expansion of the Universe.
The continuity equation can be instead rewritten as:

$$\frac{\partial\delta(\mathbf{x},\tau)}{\partial\tau}+\theta(\mathbf{x},\tau)=0\,. \tag{2.21}$$

Now, taking the time derivative of Eq. 2.19 and replacing in 2.21 we get

$$\frac{\mathrm{d}^2D_1(\tau)}{\mathrm{d}\tau^2}+\mathcal{H}(\tau)\frac{\mathrm{d}D_1(\tau)}{\mathrm{d}\tau}=\frac{3}{2}\Omega_m(\tau)\mathcal{H}^2(\tau)D_1(\tau)\,, \tag{2.22}$$

were we wrote $\delta(\mathbf{x},\tau)=D_1(\tau)\delta(\mathbf{x},0)$, with $D_1(\tau)$ the linear growth factor.
This equation, together with the Friedmann equations 2.4, 2.5, determines the growth of density perturbations in the linear regime as a function of cosmology.
Eq. 2.22 is a second-order differential equation and thus have two independent solutions, let's denote with $D_1^+(\tau)$ the fastest growing mode and $D_1^-(\tau)$ the slowest one. The evolution of the density is then

$$\delta(\mathbf{x},\tau)=D_1^+(\tau)A(\mathbf{x})+D_1^-(\tau)B(\mathbf{x})\,, \tag{2.23}$$

where $A(\mathbf{x})$ and $B(\mathbf{x})$ are two arbitrary functions describing the initial density field configuration.

Non-linear PT

We will now consider the evolution of density and velocity fields beyond the linear approximation. To proceed, we adopt a self-consistent approximation by describing the velocity field through its divergence, while justifying the neglect of vorticity. The evolution of vorticity, from 2.17, is

$$\frac{\partial\mathbf{w}(\mathbf{x},\tau)}{\partial\tau}+\mathcal{H}(\mathbf{x},\tau)\mathbf{w}(\mathbf{x},\tau)-\nabla\times\big[\mathbf{u}(\mathbf{x},\tau)\times\mathbf{w}(\mathbf{x},\tau)\big]=\nabla\Big(\frac{1}{\rho}\nabla\cdot\vec{\sigma}\Big)\,, \tag{2.24}$$

where we have temporarily restored the stress tensor. We see that if $\sigma_{ij} \simeq 0$, if the primordial vorticity vanishes, it remains 0 at all times. On the other hand, if the initial vorticity is non-zero, we saw that in the linear regime vorticity decays due to the expansion of the Universe. However, non-linear interactions, such as those represented by the third term in the equation above, can amplify vorticity. For simplicity, we assume vanishing primordial vorticity and neglect σ_{ij}, ensuring that vorticity remains zero during evolution.
This framework forms the basis of perturbation theory (PT), where density contrast and velocity fields are expanded as perturbative series around their linear solutions. Assuming the variance of linear fluctuations is small and the velocity field is irrotational, we write:

$$\delta(\mathbf{x},\tau) = \sum_{n=1}^{\infty} \delta^{(n)}(\mathbf{x},\tau)\,, \tag{2.25}$$

$$\theta(\mathbf{x},\tau) = \sum_{n=1}^{\infty} \theta^{(n)}(\mathbf{x},\tau)\,. \tag{2.26}$$

Here, $\delta^{(1)}$ and $\theta^{(1)}$ are linear terms proportional to the initial density contrast field, $\delta^{(2)}$ and $\theta^{(2)}$ are quadratic in the initial density field, and so on.

Fourier Representation

When we consider large scales, fluctuations are small and linear perturbation theory provides an adequate description of cosmological fields.
In this regime, different Fourier modes evolve independently conserving the primordial statistics. Therefore, it is natural to transform 2.16 and 2.17

$$\frac{\partial\tilde{\delta}(\mathbf{k},\tau)}{\partial\tau} + \tilde{\theta}(\mathbf{k},\tau) = -\int \frac{\mathrm{d}^3\mathbf{k}_1\mathrm{d}^3\mathbf{k}_2}{(2\pi)^6}\delta_D(\mathbf{k}-\mathbf{k}_{12})\alpha(\mathbf{k}_1,\mathbf{k}_2)\tilde{\theta}(\mathbf{k}_1,\tau)\tilde{\delta}(\mathbf{k}_2,\tau)\,, \tag{2.27}$$

$$\begin{aligned}\frac{\partial\tilde{\theta}(\mathbf{k},\tau)}{\partial\tau} &+ \mathcal{H}(\tau)\tilde{\theta}(\mathbf{k},\tau) + \frac{3}{2}\Omega_m\mathcal{H}^2(\tau)\tilde{\delta}(\mathbf{k},\tau) = \\ &= -\int \frac{\mathrm{d}^3\mathbf{k}_1\mathrm{d}^3\mathbf{k}_2}{(2\pi)^6}\delta_D(\mathbf{k}-\mathbf{k}_{12})\beta(\mathbf{k}_1,\mathbf{k}_2)\tilde{\theta}(\mathbf{k}_1,\tau)\tilde{\theta}(\mathbf{k}_2,\tau)\,,\end{aligned} \tag{2.28}$$

where the functions

$$\alpha(\mathbf{k}_1,\mathbf{k}_2) = \frac{\mathbf{k}_{12}\cdot\mathbf{k}_1}{k_1^2}\,, \tag{2.29}$$

$$\beta(\mathbf{k}_1,\mathbf{k}_2) = \frac{k_{12}^2(\mathbf{k}_1\cdot\mathbf{k}_2)}{2k_1^2k_2^2}\,, \tag{2.30}$$

encode the non linearity of the evolution and come from the non-linear term in the continuity equation 2.16 and Euler equation 2.17 respectively.

General solution for Einstein-de Sitter Universe

Let's consider an Einstein-de Sitter Universe, for which we have $\Omega_m = 1$ and $\Omega_\Lambda = 0$. In this case, equation 2.4 implies $a(\tau) \propto \tau^2$ and $\mathcal{H}(\tau) = 2/\tau$. This allows us to simplify the equations of motion for perturbations by factoring out $\mathcal{H}$ from the velocity field, rendering Eqs. 2.27 and 2.28 homogeneous in either τ or equivalently in $a(\tau)$. In this framework, the continuity and Euler equations in Fourier space can be solved using a perturbative expansion of the following form

$$\tilde{\delta}(\mathbf{k},\tau) = \sum_{n=1}^{\infty} a^n(\tau)\delta_n(\mathbf{k})\,, \tag{2.31}$$

$$\tilde{\theta}(\mathbf{k},\tau) = -\mathcal{H}\sum_{n=1}^{\infty} a^n(\tau)\theta_n(\mathbf{k})\,, \tag{2.32}$$

where only the fastest-growing mode of perturbations is considered. At small a the series are dominated by their first term, and since $\theta_1(\mathbf{k}) = \delta_1(\mathbf{k})$ from the continuity equation, in follows that $\delta_1(\mathbf{k})$ completely characterize the linear fluctuations.
Equations 2.27 and 2.28, determine $\delta_n(\mathbf{k})$ and $\theta_n(\mathbf{k})$ in terms of the linear fluctuations:

$$\delta_n(\mathbf{k}) = \int \mathrm{d}^3\mathbf{q}_1...\mathrm{d}^3\mathbf{q}_n\delta_d(\mathbf{k}-\mathbf{q}_{1...n})F_n(\mathbf{q}_1,...,\mathbf{q}_n)\delta_q(\mathbf{q}_1)...\delta_1(\mathbf{q}_n)\,, \tag{2.33}$$

$$\theta_n(\mathbf{k}) = \int \mathrm{d}^3\mathbf{q}_1...\mathrm{d}^3\mathbf{q}_n\delta_d(\mathbf{k}-\mathbf{q}_{1...n})G_n(\mathbf{q}_1,...,\mathbf{q}_n)\delta_q(\mathbf{q}_1)...\delta_1(\mathbf{q}_n)\,, \tag{2.34}$$

where F_n and G_n are the perturbative kernels capturing the non-linear mode coupling due to gravitational interactions. They are constructed from the fundamental mode coupling functions $\alpha(\mathbf{k}_1,\mathbf{k}_2)$ and $\beta(\mathbf{k}_1,\mathbf{k}_2)$ according to the recurrence relation

$$\begin{aligned}F_n(\mathbf{q}_1,\ldots,\mathbf{q}_n) = \sum_{m=1}^{n-1}\frac{G_m(\mathbf{q}_1,\ldots,\mathbf{q}_m)}{(2n+3)(n-1)}\Big[&(2n+1)\alpha(\mathbf{k}_1,\mathbf{k}_2)F_{n-m}(\mathbf{q}_{m+1},\ldots,\mathbf{q}_n)\\ &+2\beta(\mathbf{k}_1,\mathbf{k}_2)G_{n-m}(\mathbf{q}_{m+1},\ldots,\mathbf{q}_n)\Big]\,,\end{aligned} \tag{2.35}$$

$$\begin{aligned}G_n(\mathbf{q}_1,\ldots,\mathbf{q}_n) = \sum_{m=1}^{n-1}\frac{G_m(\mathbf{q}_1,\ldots,\mathbf{q}_m)}{(2n+3)(n-1)}\Big[&3\alpha(\mathbf{k}_1,\mathbf{k}_2)F_{n-m}(\mathbf{q}_{m+1},\ldots,\mathbf{q}_n)\\ &+2n\beta(\mathbf{k}_1,\mathbf{k}_2)G_{n-m}(\mathbf{q}_{m+1},\ldots,\mathbf{q}_n)\Big]\,,\end{aligned} \tag{2.36}$$

where $\mathbf{k}_1 \equiv \mathbf{q}_1 + ... + \mathbf{q}_m$, $\mathbf{k}_2 \equiv \mathbf{q}_{m+1} + ... + \mathbf{q}_n$, $\mathbf{k} \equiv \mathbf{k}_1 + \mathbf{k}_2$ and $F_1 = G_1 = 1$.
For $n = 2$, after symmetrizing the expressions with respect to $\mathbf{q}_1$ and $\mathbf{q}_2$ we have

$$F_2(\mathbf{q}_1,\mathbf{q}_2) = \frac{5}{7} + \frac{1}{2}\frac{\mathbf{q}_1\cdot\mathbf{q}_2}{q_1q_2}\left(\frac{q_1}{q_2}+\frac{q_2}{q_1}\right) + \frac{2}{7}\frac{(\mathbf{q}_1\cdot\mathbf{q}_2)^2}{q_1^2q_2^2}\,, \tag{2.37}$$

$$G_2(\mathbf{q}_1,\mathbf{q}_2) = \frac{3}{7} + \frac{1}{2}\frac{\mathbf{q}_1\cdot\mathbf{q}_2}{q_1q_2}\left(\frac{q_1}{q_2}+\frac{q_2}{q_1}\right) + \frac{4}{7}\frac{(\mathbf{q}_1\cdot\mathbf{q}_2)^2}{q_1^2q_2^2}\,. \tag{2.38}$$

Symmetrized kernels are used in perturbation theory to ensure solutions respect the symmetries of the universe, including isotropy, homogeneity, and invariance under wave-vector permutations. Symmetrization ensures physical consistency by enforcing momentum conservation among interacting wave-vectors in Fourier space, ensuring that the solutions remain consistent with the fundamental equations governing the evolution of the density field. Furthermore, it eliminates redundant contributions in higher-order calculations, simplifying both analytical and computational efforts.

2.2 Statistical description of matter density fluctuations

Tests of cosmological theories which characterize the primordial fluctuations are not deterministic in nature but rather statistical, for the following reasons. First, we do not have direct observational access to primordial fluctuations, which would provide definite initial conditions for the deterministic evolution equations. In addition, the time-scale for cosmological evolution is so much larger that that over which we can make observations that is not possible to follow the evolution of single systems. Therefore, testing the evolution of structure must be done statistically.
The observable universe is modeled as a stochastic realization of a statistical ensemble of possibilities. The goal is to make statistical predictions which in turn depend on the statistical properties of primordial perturbations.

2.2.1 Physical origin of Fluctuations from Inflation

Among the models that have emerged to explain the Large-Scale Structure of the universe the most widely considered are those based on the inflationary paradigm. In models of inflation the stochastic properties of the fields originate from quantum fluctuations of a scalar field, the inflaton. It is beyond the scope of this thesis to describe inflationary models in any detail but it is worth noting that, at least for the simplest single-field models within the slow-roll approximation, all fluctuation originate from scalar adiabatic perturbations. In adiabatic perturbations, the total density perturbation is non-zero, giving rise to perturbations in the spatial curvature; by contrast, isocurvature fluctuations is a set of individual perturbations such that the total density perturbation vanishes.
During inflation, the quantum fluctuations of the inflaton field are described within the framework of quantum field theory in an expanding background. As the perturbation modes exit the Hubble horizon, their quantum evolution effectively freezes and the field behaves as a classical stochastic field. In this regime, statistical ensemble averages can be identified with the vacuum expectation values of the corresponding quantum operators

$$\langle \ldots \rangle \equiv \langle 0| \ldots |0\rangle \ . \tag{2.39}$$

After re-entering the horizon these perturbations seed fluctuations in the gravitational potential, transferring their statistical properties to the cosmological fields.

2.2.2 Correlation functions and Power Spectra

The density contrast field $\delta(\mathbf{x})$, has an ensemble average of zero by definition, but its statistical properties can be studied through correlations.
To that purpose we assume that it is statistically homogeneous and isotropic. A random field is called statistically homogeneous if all the joint multi-point probability distribution functions $p(\delta_1, \delta_2, ...)$ or its moment, remain the same under translation of the coordinates $\mathbf{x}_1, \mathbf{x}_2, ...$ in space (here $\delta_i \equiv \delta(\mathbf{x}_i)$), meaning that probabilities only depend on the relative positions. A stochastic field is called statistically isotropic if $p(\delta_1, \delta_2, ...)$ is invariant under spatial rotations.

Two-Point Correlation Function

The two-point correlation function is defined as the joint ensemble average of the density at two different locations

$$\xi(r) = \langle \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r})\rangle \ . \tag{2.40}$$

Note that, as stated in the previous paragraph, the correlation function only depend on the norm of $\mathbf{r}$ due to statistical homogeneity and isotropy.
We can compute correlators in Fourier space. The Fourier transform of the density contrast $\delta(\mathbf{x})$ is

$$\delta(\mathbf{x}) = \int \mathrm{d}^3\mathbf{k}\ \delta(\mathbf{k})\,\mathbf{exp}(i\mathbf{k}\cdot\mathbf{x})\,, \tag{2.41}$$

where $\delta(\mathbf{k})$ are complex variables satisfying $\delta(\mathbf{k}) = \delta^*(-\mathbf{k})$.
The Fourier space correlator

$$\begin{aligned}
\langle \delta(\mathbf{k})\delta(\mathbf{k}')\rangle =& \int \frac{\mathrm{d}^3\mathbf{x}\,\mathrm{d}^3\mathbf{r}}{(2\pi)^6}\,\langle \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r})\rangle\,\mathbf{exp}[-i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x} - i\mathbf{k}'\cdot\mathbf{r}] \\
=& \int \frac{\mathrm{d}^3\mathbf{x}\,\mathrm{d}^3\mathbf{r}}{(2\pi)^6}\xi(r)\mathbf{exp}[-i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x} - i\mathbf{k}'\cdot\mathbf{r}] \\
=& \delta_D(\mathbf{k}+\mathbf{k}')\int \frac{\mathrm{d}^3\mathbf{r}}{(2\pi)^3}\xi(r)\mathbf{exp}(i\mathbf{k}\cdot\mathbf{r}) \\
=& \delta_D(\mathbf{k}+\mathbf{k}')P(k)\,,
\end{aligned} \tag{2.42}$$

where $P(k)$ is by definition the density power spectrum.
The inverse relation between two-point correlation function and power spectrum is then

$$\xi(r) = \int \mathrm{d}^3\mathbf{k}\,P(k)\,\mathbf{exp}(i\mathbf{k}\cdot\mathbf{r})\,. \tag{2.43}$$

We can also evaluate the variance of the field

$$\begin{aligned}
\langle \delta(\mathbf{k})^2\rangle = \xi(0) = \int \mathrm{d}k\,P(k) =& \frac{1}{2\pi^2}\int_0^\infty \mathrm{d}k\,k^2\,P(k) \\
=& \frac{1}{2\pi^2}\int_0^\infty \mathrm{d}k\,\frac{k^3}{k}\,P(k) = \int_0^\infty \frac{\mathrm{d}k}{k}\Delta(k)\,,
\end{aligned} \tag{2.44}$$

where $\Delta(k) = \frac{k^3}{2\pi^2}P(k)$ is the dimensionless power spectrum.

The Wick Theorem for Gaussian Fields

The power spectrum $P(k)$ is a fundamental quantity for characterizing almost all homogeneous random fields. Its utility becomes evident for Gaussian random fields, where the joint distribution of density values is Gaussian. For such fields, the Wick theorem applies: the ensemble average of any product of variables can be expressed in terms of two-point correlators. This is a foundational concept in classical and quantum field theories. Explicitly, for the Fourier modes, the Wick theorem states:

$$\langle \delta(\mathbf{k}_1)\ldots\delta(\mathbf{k}_{2p+1})\rangle = 0 \tag{2.45}$$

$$\langle \delta(\mathbf{k}_1)\ldots\delta(\mathbf{k}_{2p})\rangle = \sum_{\text{pair associations}}\ \prod_{p\,\text{pairs (i,j)}} \langle \delta(\mathbf{k}_i)\delta(\mathbf{k}_j)\rangle \ . \tag{2.46}$$

Thus, the statistical properties of the random variables $\delta(\mathbf{k})$ are fully specified by the shape and normalization of $P(k)$.

Higher Order Correlators

In general it is possible to define higher-order correlation functions. They are defined as the connected part of the joint ensemble average of the density constrast. The N-point correlation function is formally written as

$$\begin{aligned}\xi_N(\mathbf{x}_1,\ldots,\mathbf{x}_N) &= \langle\delta(\mathbf{x}_1),\ldots,\delta(\mathbf{x}_n)\rangle_c \\ &= \langle\delta(\mathbf{x}_1),\ldots,\delta(\mathbf{x}_n)\rangle + \\ &- \sum_{\mathcal{S}\in\mathcal{P}(\{\mathbf{x}_\infty,\ldots,\mathbf{x}_\mathrm{N}\})}\prod_{s_i\in\mathcal{S}}\xi_{\#s_i}(\mathbf{x}_{s_i(1)},\ldots,\mathbf{x}_{s_i(\#s_i)})\,,\end{aligned} \tag{2.47}$$

where the sum is over all proper partitions of $\{\mathbf{x}_1,\ldots,\mathbf{x}_N\}$, excluding the set itself. When the average of $\delta(\mathbf{x})$ is defined as zero, only partitions that contain no singlets contributes. For Gaussian fields, all connected correlators beyond the two-point function vanish, making ξ_2 the sole non-zero connected part.
In Fourier space, due to homogeneity, connected N-point correlators are proportional to $\delta_D(\mathbf{k}_1+\cdots+\mathbf{k}_N)$. Then we can define $P_N(\mathbf{k}_1,\ldots,\mathbf{k}_N)$

$$\langle\delta(\mathbf{k}_1)\ldots\delta(\mathbf{k}_N)\rangle_c = \delta_D(\mathbf{k}_1+\cdots+\mathbf{k}_N)P_N(\mathbf{k}_1,\ldots,\mathbf{k}_N)\,. \tag{2.48}$$

For $N=3$, this corresponds to the bispectrum $B(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)$.

Linear and One-loop Power Spectrum

In linear theory, the two-point correlation of the density contrast in Fourier space is given by:

$$\langle\delta^{(1)}(\mathbf{k})\delta^{(1)}(\mathbf{k}')\rangle = (2\pi)^3\delta_D(\mathbf{k}+\mathbf{k}')P_L(k,\tau)\,, \tag{2.49}$$

where $P_L(k,\tau)=D_+^2(\tau)P_L(k)$.
To estimate the typical density contrast variance at a given point $\mathbf{x}$, we compute:

$$\langle\delta^2(\mathbf{x})\rangle = \int\frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3}P_L(q)\,. \tag{2.50}$$

Using the non-linear perturbation theory results we can compute the power spectrum at order $[\delta^{(1)}]^3$. Higher-order corrections to tree-level PT, namely one-loop PT, describes the first effects of mode-mode coupling in the evolution of the power spectrum.
One-loop corrections to power spectrum have been extensively studied in literature [34][35][36]; we review here the results.
We can write the power spectrum up to one-loop corrections as

$$P(k,\tau) = P^{(0)}(k,\tau)+P^{(1)}(k,\tau)\,, \tag{2.51}$$

where the superscript (n) denotes the n-loop contribution, the tree-level (0-loop) contribution is just the linear power spectrum $P_L(k,\tau)=D_+^2(\tau)P_L(k)$ and the one-loop contribution consists of two terms

$$P^{(1)}(k,\tau) = P_{22}(k,\tau)+P_{13}(k,\tau)\,, \tag{2.52}$$

with

$$P_{22}(k,\tau) \equiv 2\int\mathrm{d}^3\mathbf{q}[F_2(\mathbf{k}-\mathbf{q},\mathbf{q})]^2P_L(|\mathbf{k}-\mathbf{q}|,\tau)P_L(q,\tau)\,, \tag{2.53}$$

$$P_{13}(k,\tau) \equiv 6\int\mathrm{d}^3\mathbf{q}\,F_3(\mathbf{k},\mathbf{q},-\mathbf{q})P_L(k,\tau)P_L(q,\tau)\,. \tag{2.54}$$

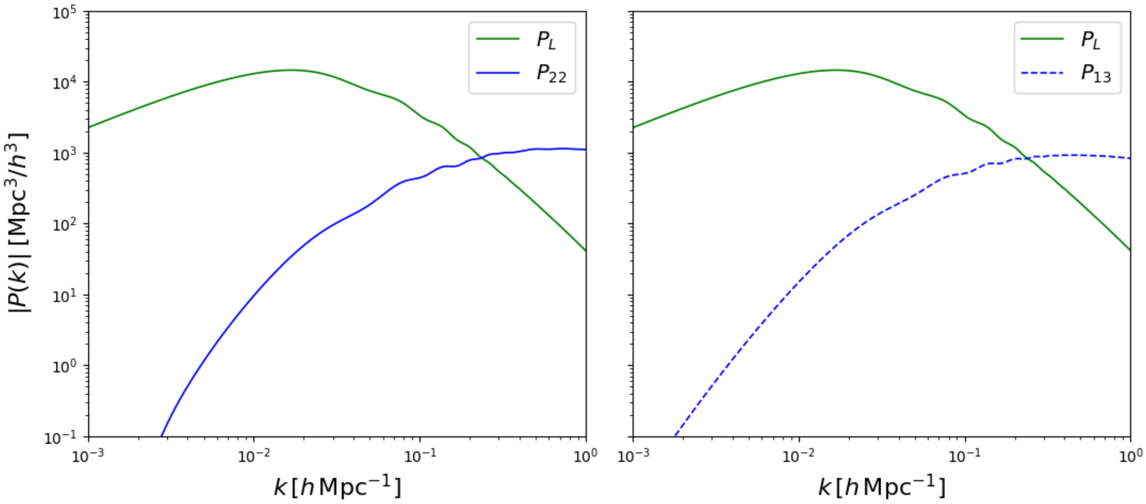

Figure 2.1: Left panel: Plot of the linear power spectrum and the P_{22} one-loop contribution at $z = 0$. Right panel: Similar to the left panel, but showing the P_{13} one-loop contribution instead. The dashed line indicates a negative contribution.

Collecting everything we obtain

$$\begin{aligned} P_{1-loop}(k,\tau) =& P_L(k,\tau)+ \\ & + 2\int \mathrm{d}^3\mathbf{q}[F_2(\mathbf{k}-\mathbf{q},\mathbf{q})]^2 P_L(|\mathbf{k}-\mathbf{q}|,\tau)P_L(q,\tau)+ \\ & + 6\int \mathrm{d}^3\mathbf{q}\, F_3(\mathbf{k},\mathbf{q},-\mathbf{q})P_L(k,\tau)P_L(q,\tau)\,. \end{aligned} \tag{2.55}$$

The computation of P_{22} and P_{13} can be found in Appendix A.
In Fig. 2.1, we present a plot of P_{22} and P_{13} compared to P_L. The one-loop corrections to the integral increase significantly with increasing wavenumber. Notably, $P_{22}(k)$ is positive, while $P_{13}(k)$ is negative. As a result, the two corrections nearly cancel each other, leading to a relatively small overall correction. However, starting at $k \simeq 0.1h/\mathrm{Mpc}$, the corrections to the linear power spectrum become non-negligible, indicating that the linear regime is no longer valid.

Linear bispectrum

The three-point correlator, the bispectrum, of the density contrast in Fourier space is defined by 2.48:

$$\langle \delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3)\rangle_c = \delta_D(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)B(\mathbf{k}_1,\mathbf{k}_2,\tau)\,. \tag{2.56}$$

The tree-level expression for the bispectrum is:

$$B(k_1,k_2,\tau) \equiv 2P_L(k_1,\tau)P_L(k_2,\tau)F_2(\mathbf{k}_1,\mathbf{k}_2) + \text{permutations}\,. \tag{2.57}$$

The derivation of this expression is in Appendix A.
Since the Dirac delta function imposes the condition $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0$, the three wavevectors form a closed triangle in Fourier space. Therefore, the bispectrum is not only characterized by the magnitudes of the three wavenumbers, but also by the shape of the corresponding triangle. In particular, three commonly considered configurations are the equilateral, squeezed, and flattened configurations, shown in 2.2.

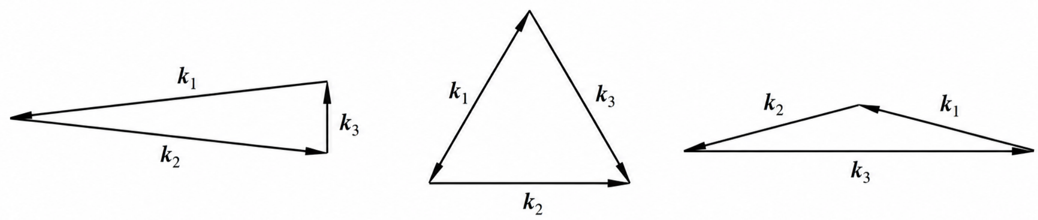

Figure 2.2: Left panel: the squeezed configuration is characterized by one wavenumber much smaller than the other two $k_3 \ll k_1 \simeq k_2$. Central panel: in the equilateral configuration the three wavenumbers have approximately equal magnitudes $k_1 \simeq k_2 \simeq k_3$. Right panel: in the flattened configuration the three wavevectors are nearly collinear $k_3 \simeq k_1 + k_2$. Figure is from [37].

2.3 Limitations of Perturbation Theory

In the analysis of the one-loop power spectrum integrals, two major sources of divergence arise: the ultraviolet and infrared singularities. This can be seen by considering the spectra in the limit of small and large internal momentum q:

$$P^{UV}_{1-loop} \sim k^2 P_L(k) \int_{q \gg k} \mathrm{d}q \, q^2 \frac{P_L(q)}{q^2} \,, \tag{2.58}$$

$$P^{IR}_{1-loop} \sim P_L(k) \int_{q \ll k} \mathrm{d}q \, q^2 P_L(q) \,. \tag{2.59}$$

UV singularities originate from the behavior of the one-loop integrals at small scales, where nonlinear physics becomes significant. While this integral converge for the ΛCDM power spectrum, it fails to capture the physics of small-scale virialization, where halos form.

The physical IR effects arise from the coherent displacements induced by large-scale flows. These flows move structures coherently over scales comparable to the BAO wavelength, leading to the suppression and shift of the Baryonic Acoustic Oscillations (BAO) peak. BAO refers to the periodic fluctuations in the density of visible baryonic matter in the universe, which were imprinted during the early universe when sound waves propagated through the hot plasma before recombination. Since these coherent displacements are not captured accurately by standard perturbation theory, they require resummation techniques for an accurate treatment.

However these divergencies are not the only problems intrinsic to standard perturbation theory, there are hidden subtleties in the treatment of the fluid equations. In particular, when solving 2.16 and 2.17 perturbatively, we have implicitly assumed that density field and velocity divergence are small. This is needed for a perturbative solution to exist, since if, for example, the second order solution is larger than the first

order piece, the series does not converge. If we average over sufficiently large scales, we would expect the overdensity to be small, however such a filtering is not built into the underlying SPT equations. A consequence is that we necessarily obtain contributions from regimes where $|\delta|$ is large, which breaks the perturbative hierarchy.
The Effective Field Theory of Large-Scale Structure addresses the UV and the $|\delta| < 1$ issues of standard perturbation theory that we just presented.

2.3.1 Effective Field Theory of the Large-Scale Structure

In this paragraph we follow the review [38], the formalism is described in depth in [17] and [18].
The starting point is the notion of coarse graining, that consists in the replacement of the standard fluid variables with those smoothed on some scale Λ^{-1}:

$$\begin{aligned} \delta(\mathbf{x},\tau) &\to \delta_\Lambda(\mathbf{x},\tau)\,, \\ \mathbf{v}(\mathbf{x},\tau) &\to \mathbf{v}_\Lambda(\mathbf{x},\tau)\,, \\ \phi(\mathbf{x},\tau) &\to \phi_\Lambda(\mathbf{x},\tau)\,, \end{aligned} \tag{2.60}$$

where the smoothing is defined by $X_\Lambda(\mathbf{k},\tau) = W_\Lambda(k)X(\mathbf{k},\tau)$ in Fourier space, with W_Λ being some window excising scales with $k > \Lambda$.
By working with smoothed fields, we are explicitly restricting our theory models to scales $k < \Lambda$. If Λ is below the non-linear scale k_{NL}, the scale at which perturbation theory breaks down, we avoid explicit contributions from short-scales modes. Their effect on the long-wavelength dynamics is then captured by adding effective operators, as will become clear in the following.
The value of k_{NL} is conventionally determined by requiring that

$$\Delta(k_{NL}) = \frac{k_{NL}^3}{2\pi^2}P(k_{NL}) = 1\,, \tag{2.61}$$

since the dimensionless power spectrum $\Delta(k)$ quantifies the amplitude of primordial fluctuations as a function of scale.
We now want to recover the underlying equations for the smoothed density field, to do that we start from the fluid equations given in 2.16, 2.17 and2.9 and apply the smoothing operator W_Λ. For full generality we retain the stress-tensor τ in the Euler equation:

$$\dot{\delta}_\Lambda(\mathbf{x},\tau) + \nabla\cdot[(1+\delta_\Lambda(\mathbf{x},\tau))\mathbf{v}_\Lambda(\mathbf{x},\tau)] = 0\,, \tag{2.62}$$

$$\begin{aligned} &\dot{\mathbf{v}}_\Lambda(\mathbf{x},\tau) + [\mathbf{v}_\Lambda(\mathbf{x},\tau)\cdot\nabla]\mathbf{v}_\Lambda(\mathbf{x},\tau) = \\ &\qquad = -\mathcal{H}(\tau)\mathbf{v}_\Lambda(\mathbf{x},\tau) - \nabla\phi_\Lambda(\mathbf{x},\tau) - \frac{1}{\rho_\Lambda(\mathbf{x},\tau)}\left[\nabla\tau_\Lambda(\mathbf{x},\tau) + \nabla\tau_\Lambda^{\mathrm{UV}}(\mathbf{x},\tau)\right]\,, \end{aligned} \tag{2.63}$$

$$\nabla^2\phi_\Lambda(\mathbf{x},\tau) = \frac{3}{2}\mathcal{H}^2(\tau)\Omega_m(\tau)\delta_\Lambda(\mathbf{x},\tau)\,, \tag{2.64}$$

where the variables are now the long wavelength density contrast and velocity fields. Indeed the operation of coarse graining corresponds to "integrating out" short-wavelengths modes.
These equations are identical to the starting fluid equations except for the appearance of a new term $\tau_\Lambda^{\mathrm{UV}}$ in the Euler equation. This term arises because multiplication and smoothing do not commute, and takes the form

$$\nabla\tau_\Lambda^{\mathrm{UV}} = \Big[[\rho\nabla\phi]_\Lambda - \rho_\Lambda\nabla\phi_\Lambda\Big] + \Big[[(\rho\mathbf{v}\cdot\nabla)\mathbf{v}]_\Lambda - (\rho_\Lambda\mathbf{v}_\Lambda\cdot\nabla)\mathbf{v}_\Lambda\Big]\,. \tag{2.65}$$

The meaning of this stress tensor is that small scales mode appearing in the fluid equations non-linearly modify the dynamics of large scale modes.
Since the exact dynamics of the short-wavelength modes are unknown, the EFTofLSS adopts an agnostic approach by parameterizing the UV stress tensor in terms of the long-wavelength fields available in the theory. Following symmetry principles, specifically isotropy, homogeneity, and the equivalence principle, the stress tensor is expanded into all possible operators compatible with these symmetries at a given order.
At the lowest order in smoothed variables:

$$\tau^{UV}_{ij,\Lambda} = \delta_{ij} \left(p_\Lambda + \bar{\rho}\tilde{c}^2_{s,\Lambda}\delta_\Lambda - \bar{\rho}\frac{c^2_{v,b,\Lambda}}{\mathcal{H}}\partial^k v_{k,\Lambda} \right) + \dots \,, \tag{2.66}$$

where p_Λ is an isotropic pressure, $\tilde{c}_{s,\Lambda}$ is the sound-speed and $c^2_{s,b,\Lambda}$ is the bulk viscosity. Physically, the speed of sound captures the back-reaction of small-scale non-linear processes, such as dark matter halo formation, on the large-scale dynamics.
To see how the stress tensor modifies the dynamics, we consider its contribution to the evolution of the velocity divergence $\theta_\Lambda(k, \tau)$. The impact of the UV stress tensor is mediated by its divergence in the Euler equation, specifically through the operator:

$$\tau_{\theta,\Lambda}(\mathbf{x}, \tau) \equiv \partial_i \left[\frac{1}{\rho_\Lambda(\mathbf{x}, \tau)} \partial_j \tau^{UV}_{ij,\Lambda}(\mathbf{x}, \tau) \right] \approx c^2_{s,\Lambda}(\tau)\nabla^2\delta_\Lambda(\mathbf{x}, \tau) + \dots \,, \tag{2.67}$$

where we have expanded to the lowest order in the fields and identified the effective speed of sound $c^2_{s,\Lambda}$ as the combination of the physical sound speed and viscosities[1]. In Fourier space, the Laplacian operator ∇^2 becomes $-k^2$, allowing us to solve the fluid equations perturbatively. The resulting EFT prediction for the density field at third order can be written as:

$$\delta_{EFT}(\mathbf{k}, \tau) = \delta^{(1-3)}_{SPT}(\mathbf{k}, \tau) - c^2_s(\tau)k^2 D(\tau)\delta^{(1)}(\mathbf{k}) + \dots \,, \tag{2.68}$$

where $\delta^{(1-3)}_{SPT}(\mathbf{k}, \tau)$ represents the standard perturbative solution up to third order.
The one-loop matter power spectrum receives the correction:

$$P_{EFT}(k, \tau) = P_{SPT,\text{1-loop}}(k, \tau) - 2c^2_{s,\Lambda}(\tau)k^2 P_{\text{lin}}(k) \,. \tag{2.69}$$

This final expression illustrates the skeleton of the EFTofLSS: the c^2_s term acts as a scale-dependent pressure that corrects the ideal fluid behavior of SPT, absorbing the sensitivity to the UV cut-off Λ and providing a theoretically consistent description of the mildly non-linear regime.

Renormalization

From this discussion we may conclude that both the truncated loop integrals and the microscopic parameters depend on the value adopted for Λ. This seems a problem since Λ is an arbitrary separation scale introduced by the EFT rather than a physical parameter. We would like our observables, such as the power spectrum, to not depend on the choice of the smoothing scale.
In fact, the EFT predictions do not depend on Λ due to renormalization: although the individual loop integrals (e.g. P^{13}) and our counterterm (c^2_s) vary as a function of Λ,

[1] If you consider the complete form for 2.66, both the bulk and the shear viscosity are present

their sum does not.
Mathematically:

$$D^2(\tau)P_{\Lambda}^{(13)}(k) - c_{s,\Lambda}^2(\tau)k^2 P_L(k) = D^2(\tau)P_{\Lambda'}^{(13)}(k) - c_{s,\Lambda'}^2(\tau)k^2 P_L(k)\,, \tag{2.70}$$

for all $\Lambda, \Lambda' \gg k$ and at all times τ.
Physically, as we vary the value of the cutoff scale, modes near the threshold are simply redistributed between the explicit perturbative calculation (the loops) and the effective parameters of the stress tensor (the counterterms). Consequently, the value of c_s^2 shifts to exactly compensate for any variation in the loop integrals, ensuring that the final theoretical predictions are robust and independent of the specific choice of the smoothing scale.

Numerical validation

We report here the numerical validation performed by [38].
Comparing the standard perturbation theory predictions for the matter power spectrum with those obtained from N-body simulations assuming a ΛCDM cosmology, and evolving the Universe down to z = 0 we can assess the performance of SPT with respect to the one of the effective theory. This is shown in 2.3, where the linear theory prediction for the power spectrum and both the one-loop corrections (SPT and EFT) are plotted.
At low-k, we generally find good agreement between simulations and data: these correspond to large-scales, on which the Universe is close to a perfect fluid and traces the initial conditions with high accuracy. On smaller scales, linear theory starts to break down, due to the impact of gravitational evolution on Large-Scale Structure. In principle, this deviation between simulations and theory should be captured by the addition of one-loop SPT contributions, at least on relatively large scales. However, this is not what we see in the figure: adding one-loop SPT corrections does not appear to improve the fit to data; similarly, the addition of two-loop terms does not solve the problem, with the theory now underestimating the spectrum. From this simple test, we conclude that to yield accurate predictions we need the EFTofLSS.

2.3.2 Infrared resummation

While the EFT framework addresses the sensitivity of perturbative predictions to short-scale physics, we now turn to the resummation of the infrared contributions.
The Eulerian formulation of the EFTofLSS causes the power spectra to be overly wiggly. The model correctly predicts the shape of the spectrum but it fails to fully capture the oscillations. In the initial conditions the BAO feature is a sharp oscillation induced by preferred clustering of matter on small distances. Due to large-wavelengths bulk flows galaxies shift with respect to their initial position, thus any sharp phenomenon gets blurred out. Such a feature manifests in correlators as bumps or wiggles whose amplitude decay in time. The Eulerian formalism does not capture this dumping.
It is customary to decompose the primordial power spectrum into two pieces, denoted as "non-wiggly" and "wiggly", and, working in a Lagrangian framework, to perform a partial Taylor expansion for the wiggly part, keeping the non wiggly part exponentiated. For further details or the complete derivation, please refer to [19].

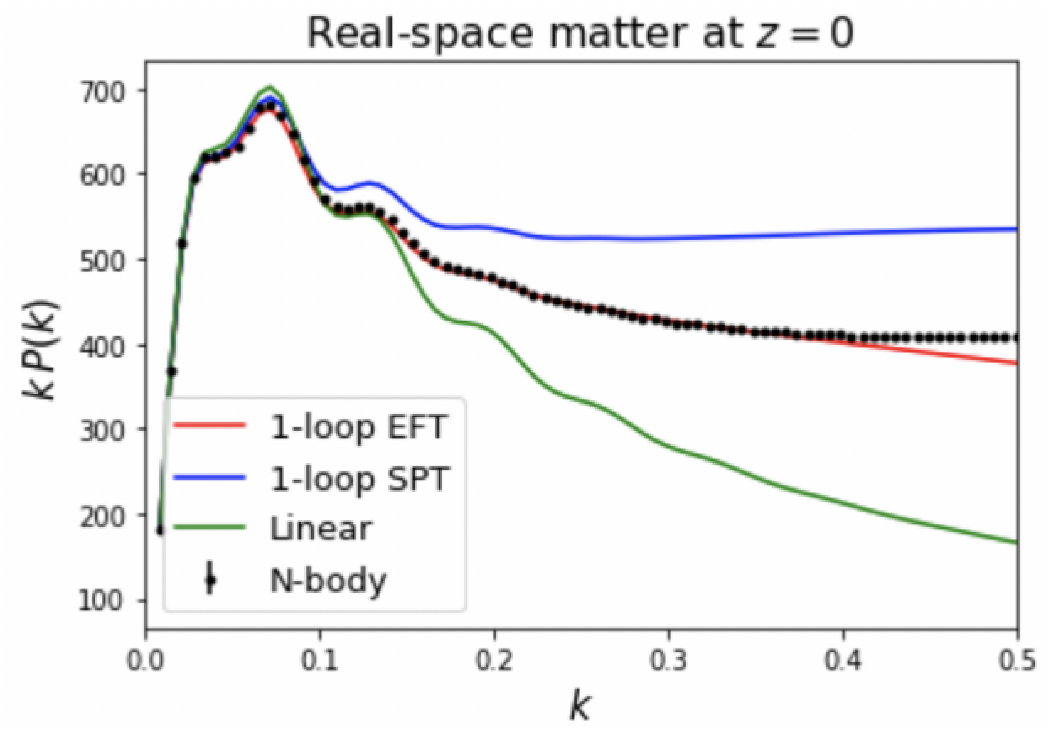

Figure 2.3: Comparison of matter power spectra obtained from simulations (black points), linear theory (green), one-loop SPT (blue) and one-loop EFTofLSS (red). The numerical results are obtained from the Quijote simulations at redshift zero [39], whilst theory spectra are computed with Class-PT. All quantities are given in h Mpc^{-1} units. Figure is from [38].

The IR resummed version of the 1-loop power spectrum is:

$$P_{\mathrm{L}}^{\mathrm{IR-res}}(k) = P_L^{\mathrm{nw}}(k) + e^{-k^2\Sigma^2} P_L^{\mathrm{w}} \,, \tag{2.71}$$

$$P_{\mathrm{NL}}^{\mathrm{IR-res}}(k) = P_L^{\mathrm{nw}}(k) + P_{\mathrm{1-loop}}^{\mathrm{nw}}(k) + e^{-k^2\Sigma^2}\left[(1+k^2\Sigma^2)P_L^{\mathrm{w}} + P_{\mathrm{1-loop}}^{\mathrm{w}}\right] \,, \tag{2.72}$$

where Σ is the velocity dispersion. The factor $e^{-\Sigma^2 k^2}$ acts as the IR damping factor, suppressing contributions from small scales.
The resulting power spectrum is shown in 2.4, and, as expected, finds a much improved fit to the data, without the spurious wiggle confirming that this resummation method improves the description of the matter power spectrum at large scales where IR modes play a significant role.

2.4 Galaxy Bias

The distribution of galaxies, quasars and clusters of galaxies is one of the foundations of our knowledge about the history of the Universe. These tracers can be observed out to cosmological distances, and can thus be used to survey significant fractions of the observable Universe. If we understand how the distribution of tracers is related to the underlying distribution of matter, we can access a wealth of information on the properties of dark matter, dark energy and gravity, as well as the nature of the process that produced the initial seeds of structure. Such a relation between luminous tracers

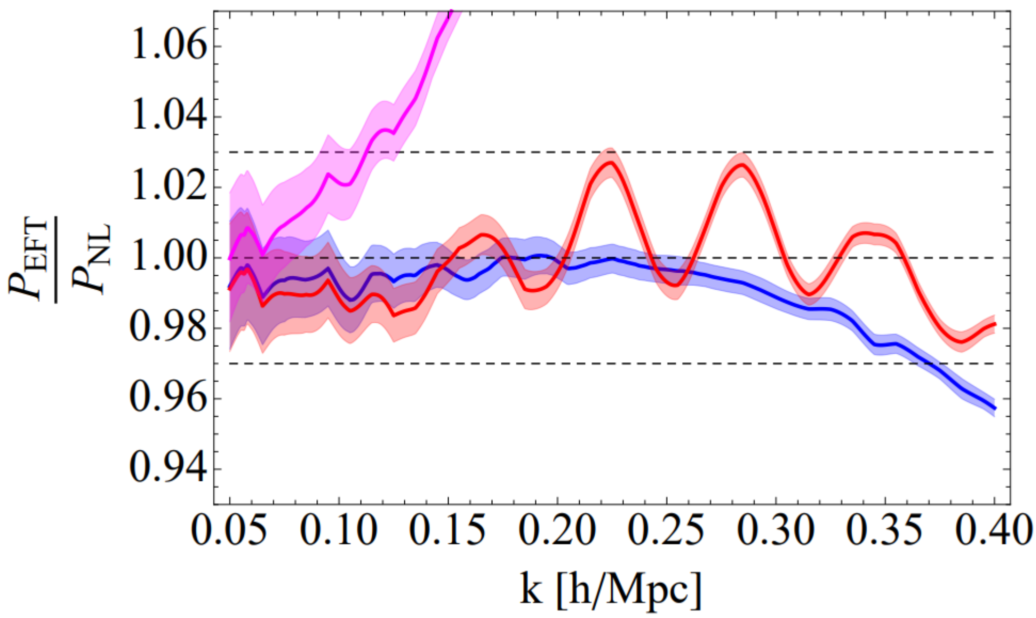

Figure 2.4: The impact of IR resummation on the matter power spectrum. We plot the ratio of perturbative predictions for the matter power spectrum to that obtained from simulations for three models: two-loop SPT (pink), the EFTofLSS without IR resummation (red) and the EFTofLSS including IR resummation (blue). IR resummation provides a much improved fit to data. Figure is from [19].

and matter is known as bias.

We focus on large scales where the cosmic density fields are quasi-linear. On these scales, the clustering of galaxies can be described by a perturbative expansion, and the complicated physics of galaxy formation is absorbed by a finite set of coefficients of the expansion, called bias parameters. This result relies on the fact that on large scales, structure formation is completely determined by the action of gravity.

The bias parameter can be understood as coefficients of operators $\mathcal{O}(\mathbf{x},\tau)$ in an expansion of the galaxy number density perturbation:

$$\delta_g(\mathbf{x},\tau) = \sum_{\mathcal{O}} b_{\mathcal{O}}(\tau)\mathcal{O}(\mathbf{x},\tau)\,. \tag{2.73}$$

Thus, the general perturbative bias expansion consists of an enumeration of all operators that are relevant at a given order in perturbation theory. To describe the one-loop power spectrum for biased tracers, we need to consider all relevant operators up to third order in the bias expansion:

$$\begin{aligned}\delta_g(\mathbf{x}) =& b_1\delta(\mathbf{x}) + \varepsilon + \frac{b_2}{2}\delta^2(\mathbf{x}) + b_{G_2}\mathcal{G}_2(\mathbf{x}) + \frac{b_3}{6}\delta^3(\mathbf{x}) + b_{G_3}\mathcal{G}_3(\mathbf{x}) + \\ &+ b_{G_2\delta}\mathcal{G}_2(\mathbf{x})\delta(\mathbf{x}) + b_{\Gamma_3}\Gamma_3(\mathbf{x})\,,\end{aligned} \tag{2.74}$$

where ε represents the stochastic contribution, which is uncorrelated with the large-scale density field. This term represents the fact that whether a galaxy forms at a given location depends on the initial conditions on very small scales, whose random phases are not determined by the large-scale perturbations included in the bias expansion.

The operators in 2.74 are defined by:

$$\mathcal{G}_2(\mathbf{x}) \equiv (\partial_i\partial_j\phi(\mathbf{x}))^2 - (\partial^2\phi(\mathbf{x}))^2\,, \tag{2.75}$$

$$\mathcal{G}_3(\mathbf{x}) \equiv -\partial_i\partial_j\phi\partial_j\partial_k\phi(\mathbf{x})\partial_k\partial_i\phi(\mathbf{x}) - \frac{1}{2}\partial^2\phi + \frac{3}{2}(\partial_i\partial_j\phi)^2\partial^2\phi\,, \tag{2.76}$$

$$\Gamma_3(\phi_g,\phi_v)(\mathbf{x}) \equiv \mathcal{G}_2(\phi_g)(\mathbf{x}) - \mathcal{G}_2(\phi_v)(\mathbf{x})\,, \tag{2.77}$$

with ϕ_g and ϕ_v being the gravitational and velocity potential perturbations (in linear theory they coincide, but they differ at higher-order perturbations).
It is worth noting that contributions from operators like δ^3, $\delta\mathcal{G}_2$ and $\mathcal{G}_3$ disappear after renormalization. Renormalization is the process of absorbing divergent or redundant contributions into redefined parameters, ensuring that only physically meaningful quantities remain in the final expressions. This means that higher-order bias terms that are not independent after loop corrections can be systematically removed or absorbed into lower-order bias parameters.
For further details on the form of the one-loop auto-power spectrum of the bias tracers using the bias 2.74 please refer to [40].

Chapter 3

Large scale structure data analysis

Up to this point, the discussion has focused on the theoretical framework. We now turn to the observational side of this framework, discussing how these theoretical predictions are compared with real data and introducing a complementary tool that allows error bars estimation even in the absence of data: the Fisher forecast formalism.
Galaxy surveys are the primary observational tool for probing the Large-Scale Structure of the Universe: they measure the three-dimensional positions of galaxies by combining their angular position on the sky with their redshift. From this discrete set of positions, one can construct estimators of the statistics discussed in the previous chapters, the power spectrum and the bispectrum, and use them to test the theoretical predictions against real data.
However, measuring and modelling the bispectrum is significantly more demanding than the power spectrum, both computationally and theoretically, and this cost grows further as the volume and precision of upcoming surveys increase. For this reason, in recent years alternative strategies have been developed to extract additional information from galaxy survey data without the full cost of a bispectrum analysis.
Finally among the several observational complications that must be accounted for when comparing theoretical predictions with real survey data this chapter presents one of the most significant, redshift-space distortions, even though it is not implemented in the analysis presented in this thesis, which constitutes a first, real-space marked power spectrum study.

3.1 Estimators of Large-Scale Structure

The power spectrum and the bispectrum can be used to constrain cosmological parameters by comparing the theoretical prediction, computed as discussed in 2, with the same statistic measured from real galaxy survey data. The parameters of the theoretical model, cosmological, bias and EFT, are then adjusted until the model matches the data as closely as possible, typically through a likelihood analysis.
To measure the power spectrum from a galaxy catalogue, one first builds the density field by comparing the observed galaxy positions to a synthetic random catalogue with the same survey geometry and selection effects, which removes artefacts coming from the shape of the survey itself. The power spectrum is then obtained from the Fourier transform of this field, using standard Fast Fourier Transform (FFT) algorithms. For further details on how to construct the power spectrum starting from a galaxy survey dataset please refer to [41].

The bispectrum is measured in an analogous way, but as a function of three wavevectors. Since it depends on the shape of the triangle formed by the three wavevectors, it contains different information to that encoded in the power spectrum, in particular, it probes the nonlinear evolution of structure and the galaxy bias. Moreover, its dependence on triangle configurations makes the bispectrum a particularly powerful probe of primordial non-Gaussianity, as different shapes can be sensitive to different types of primordial signals. This is why including the bispectrum, together with the power spectrum, typically improves parameter constraints. This is clearly depicted in figure 3.1 that shows the statistical errors on the combination of cosmological parameters $f\sigma_8$, where f is the growth factor while σ_8 is the root mean square amplitude of matter density fluctuations smoothed on a scale of $8\,\mathrm{Mpc}/h$.

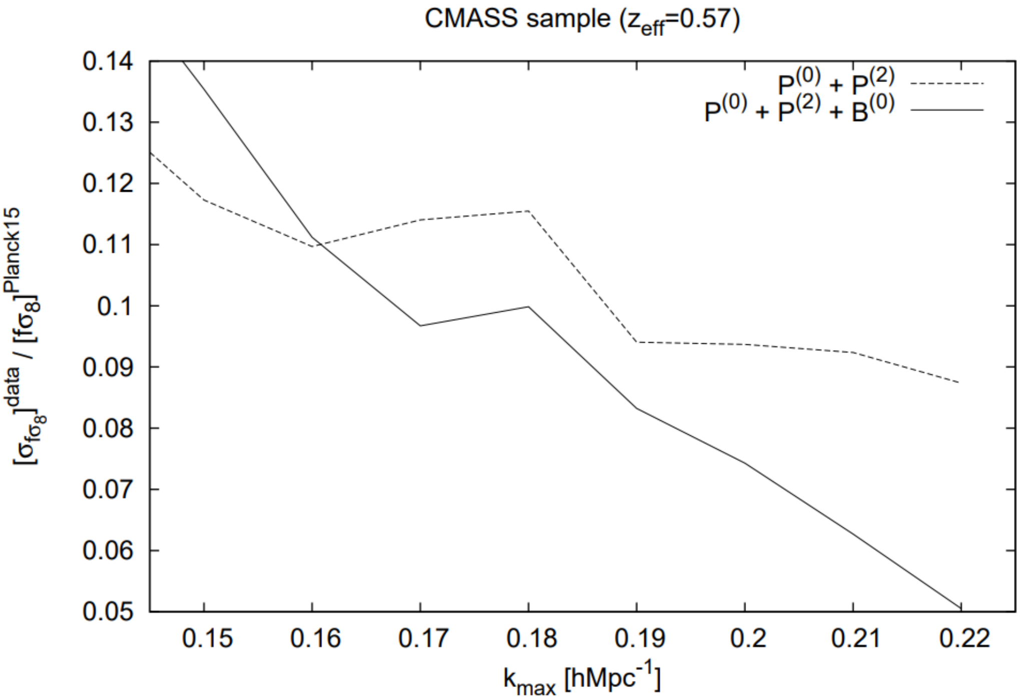

Figure 3.1: Statistical errors on the $f\sigma_8$ parameters combination as a function of the truncation scale k_{max}, extracted from the MCMC-chains of the CMASS-DR12 data. On dashed lines the prediction where the power spectrum monopole and quadrupole are used, and in solid lines when the bispectrum monopole is added to the analysis. At large scale we see that the effect of adding the bispectrum worsen the errors on $f\sigma_8$ (likely due to noise), but for k_{max} big enough the errors on $f\sigma_8$ are reduced by the effect of the bispectrum signal. Figure is from [42].

An example of the measurement of these two statistics from real galaxy survey data is shown in figure 3.2. The figure shows in the top sub-panels the monopole and quadrupole power spectrum (top) and the monopole bispectrum for different triangular shapes (bottom) measured from the CMASS sample of the BOSS DR12 galaxy survey at an effective redshift of $z_{\mathrm{eff}} = 0.57$. The middle sub-panels show the ratio between the power spectrum (top) and bispectrum (bottom) multipoles measurements and the best-fitting models. In the bottom sub-panels the difference between the data and the model relative to the statistical error of the data are presented. The black dashed lines represent a 2σ deviation (95.4%) confidence level. By varying the parameters of the theoretical model, one can identify the parameter combination that provides the best agreement with the measured statistics. This best-fit combination determines the corresponding cosmological constraints; for the BOSS DR12 example shown in Fig. 3.2, the resulting parameter values are reported in Table 3 of [42].

3.1.1 Bispectrum challenges

As we have also mentioned before, the bispectrum depends on three wavenumbers, which results in a significantly higher computational cost compared to the power spectrum: the number of independent Fourier configurations grow rapidly with the binning resolution, making both the evaluation and storage of the bispectrum more demanding. The theoretical modelling is also more challenging, since one-loop and higher-order perturbative calculations involve a larger number of terms and loop integrals. Approaches based on FFTLog can simplify this problem by decomposing the linear power spectrum into a sum of power laws, allowing loop integrals to be evaluated more efficiently, but the calculation remains considerably more involved than for the power spectrum [43, 44].

Another important challenge is the estimation of the covariance matrix. Since the bispectrum is computed for a large number of different triangular configurations, its data vector and corresponding covariance matrix can become very large. Moreover, different triangle configurations can be correlated due to non-Gaussian contributions, making an accurate covariance estimate computationally demanding and often requiring a large number of simulations [45].

3.2 Compressing higher-order information

We have established that the bispectrum carries additional cosmological information with respect to the power spectrum but extracting this information comes at a significant cost. This raises a natural question: is it possible to retain part of this additional information without measuring the bispectrum directly?

Marked correlation functions, first formalized in [8], answer this question in a simple way. Instead of computing the power spectrum of the density field itself, one first weights the field with a non linear function, called "mark", and computes the power spectrum of this weighted field.

This formalism was originally introduced in the context of modified gravity, where the effect of gravity can depend strongly on the local environment. In particular, screening mechanisms can suppress deviations from General Relativity in dense regions, while leaving them significant in low-density environments. This means that structure formation can proceed differently in different environments. A standard two-point statistic, however, treats the density field without explicitly distinguishing between different environments.

Martin White [8], with his marked power spectrum formalism, introduced a mark based on the smoothed local density field, so that each tracer is assigned a weight according to the density of the environment in which it resides. The resulting statistic gives different importance to overdense and underdense regions, making it particularly sensitive to the environmental dependence of structure formation. The effectiveness of this approach follows from the fact that the mark is a non-linear transformation of the density field. If the marked density field is schematically written as $m(\mathbf{x})[1 + \delta(\mathbf{x})]$,

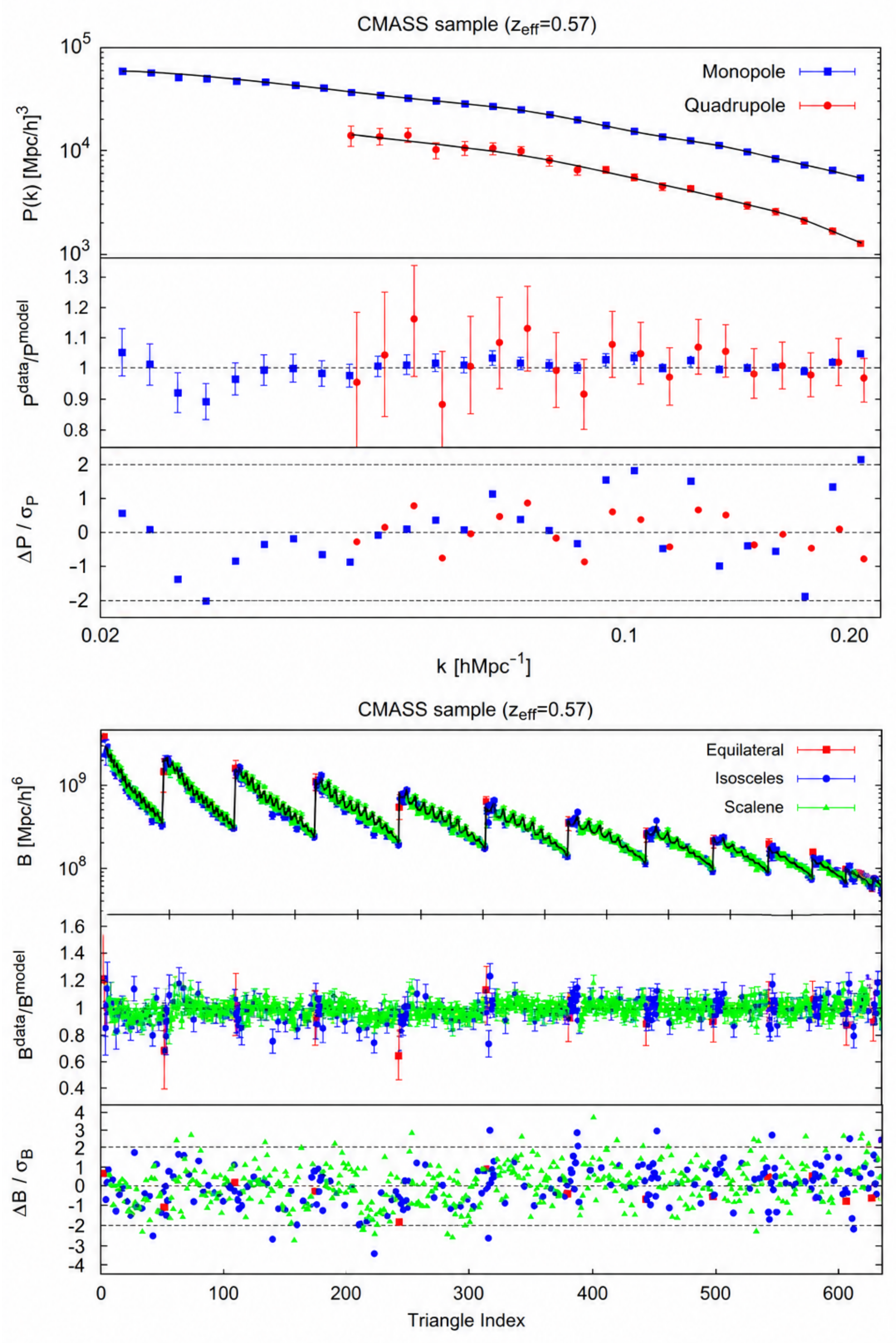

Figure 3.2: Plots of the measured power spectrum (top) and bispectrum (bottom) from the CMASS sample of the BOSS DR12 galaxy survey at an effective redshift of $z_{\rm eff} = 0.57$. The black solid lines in the top sub-panels correspond to the best-fitting model, calculated from a full fit to the power spectrum and bispectrum moments. For more details see the main text. Figure is from [42].

with $m(x)$ depending on the locally smoothed density, its two-point function contains products of higher powers of δ. Expanding the mark in powers of the density therefore generates terms such as $\langle\delta^2\delta\rangle$ and $\langle\delta^2\delta^2\rangle$, which are typical of the bispectrum and higher-order correlation functions. In this sense, the marked power spectrum does not simply reweight the information already contained in the ordinary power spectrum: the non-linear transformation mixes different orders of the density field, allowing higher-order, non-Gaussian information to enter a statistic that is still measured as a two-point function.
Thus, the marked power spectrum can be thought of as a way of compressing part of the bispectrum and higher-order correlators signal into a two point function that is not expensive to measure.
A detailed description of the formalism is provided in Chapter 4.

3.3 Galaxy surveys and observational effects

Galaxy surveys do not directly measure the three-dimensional position of a galaxy: they measure its angular position in the sky and its redshift, which is then converted into a radial distance assuming a fiducial cosmology.
In an expanding universe, the redshift has two contributions. The first contribution is the cosmological redshift, that is directly related to the comoving distance of the galaxy through the Hubble law. The second contribution is due to the peculiar velocity of the galaxy, induced by the gravitational pull of the surrounding Large-Scale Structure. Since these two contributions cannot be disentangled observationally, the reconstructed position of each galaxy along the line of sight is displaced with respect to its true position. The presence of these peculiar velocities thus introduces anisotropic distortions in the observed clustering signal, known as redshift space distortions (RSD), that are typical of the redshift space, as the name suggests.
Redshift space distortions are not a subdominant effect: they leave a characteristic, scale-dependent signature in the power spectrum and bispectrum. They are one of the primary tools used to measure the growth rate of structure f, since peculiar velocities are sourced by the same gravitational potential that drives structure formation.
When talking about RSD, two regimes are usually distinguished. On linear scales, coherent infall towards overdensities produces an apparent compression of structures along the line of sight, known as the Kaiser effect [47]. While on non linear scales, we observe an apparent elongation along the line of sight, commonly referred to as the Fingers of God effect [48]. Figure 3.3 illustrates both effects. It shows the redshift-space two point correlation function $\xi(\sigma, \pi)$ measured from the 2dFGRS survey [46], as a function of the galaxies pair separation perpendicular σ and parallel π to the line of sight. In real space, the correlation function would be isotropic, and the contours of constant ξ would simply be circles, since the clustering signal does not depend on direction. The observed contours are instead clearly anisotropic: at small separations, the elongation along the π axis is produced by the Fingers of God effect, while at larger separations the flattening along the σ axis reflects the Kaiser compression.
The analysis presented in this thesis is developed entirely in real space, and the marked power spectrum derived in Chapter 4 does not include the modelling of RSD. This is a necessary simplification for a first, consistent derivation, but it also represents a genuine limitation: real spectroscopic data are observed in redshift space, and no observable can be directly compared to real data without accounting for this effect.

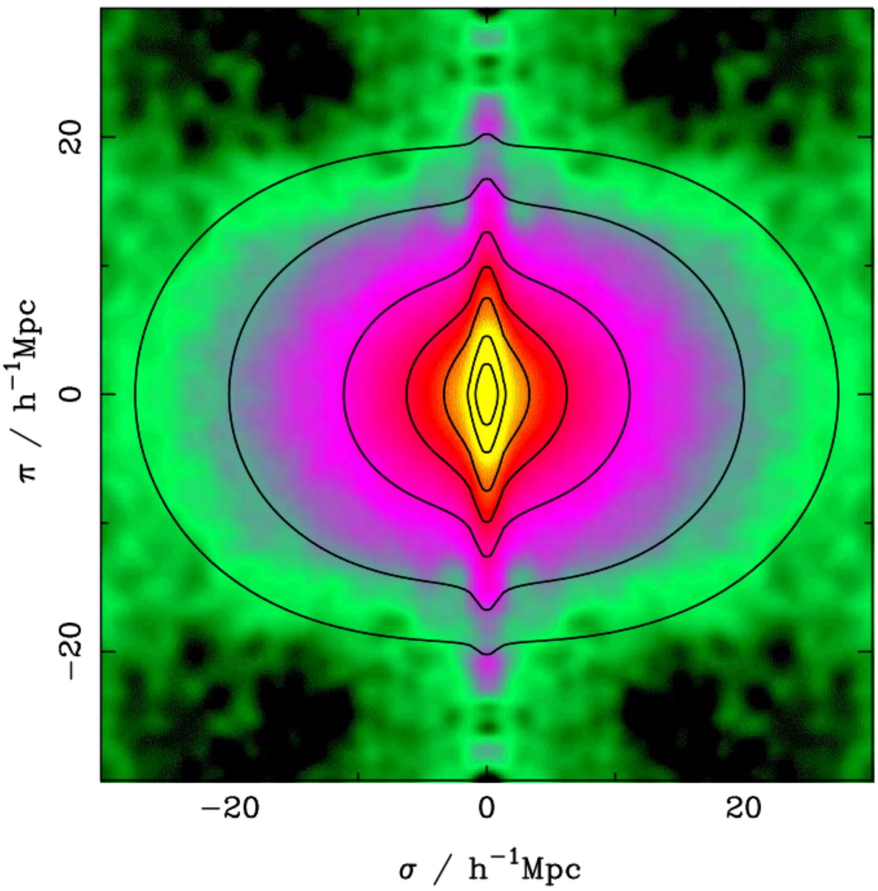

Figure 3.3: The redshift-space correlation function $\xi(\sigma, \pi)$ measured from the 2dF Galaxy Redshift Survey, as a function of the pair separation perpendicular (σ) and parallel (π) to the line of sight. If measured in real space, the contours of constant ξ would be circular, reflecting the isotropy of the clustering signal. The observed anisotropy is instead a direct signature of redshift-space distortions: the elongation along the line of sight at small separations is due to the Fingers-of-God effect, while the flattening at larger separations reflects the Kaiser compression from coherent infall. Figure is from [46].

3.4 Fisher forecast formalism

We now introduce the Fisher forecasting formalism: a tool to estimate the expected constraints on cosmological parameters.

As discussed in the previous chapter 2, the LSS of the Universe is described statistically, since the observed matter distribution represents only one realization of an underlying stochastic process. Because the primordial density fluctuations and their non-linear evolution cannot be studied deterministically, cosmological observables are characterized in terms of their statistical properties, namely their expectation values and covariances. This framework naturally motivates the use of parameter estimation techniques, which provide a quantitative assessment of the cosmological information encoded in Large-Scale Structure data.

Among these techniques, the Fisher matrix formalism occupies a central role, since it allows to forecast, prior to the availability of real data, the precision with which a given observable can constrain the parameters of a cosmological model.

3.4.1 Parameter estimation

In cosmology, our primary task is to analyze collected data and interpret it within the framework of a theoretical model. This model is assumed to be valid and typically includes a set of parameters $\boldsymbol{\theta}$, which we aim to determine. The ultimate objective is to estimate these parameters along with their associated uncertainties, or ideally, to obtain the full probability distribution of $\boldsymbol{\theta}$ given the observed data $\mathbf{x}$. This distribution is referred to as the posterior probability $p(\boldsymbol{\theta}|\mathbf{x})$, it is the probability that the parameters take certain values, after doing the experiment.
From this posterior distribution, one can compute both the expectation values of the parameters and their corresponding uncertainties.
Often, what may be easily calculable is not this, rather the opposite, $p(\mathbf{x}|\boldsymbol{\theta})$. This is referred to as forward modelling: if we know what the parameters are, we can compute the expected distribution of the data.
Following the example in [49], let's consider a model which is a gaussian with mean μ and variance σ^2. The model has two parameters, $\boldsymbol{\theta} = (\mu, \sigma)$, and the probability of a single variable x given the parameters is

$$p(x|\boldsymbol{\theta}) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \tag{3.1}$$

we can relate this to $p(\boldsymbol{\theta}|\mathbf{x})$ using Bayes' Theorem:

$$p(\boldsymbol{\theta}|\mathbf{x}) = \frac{p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{x})}. \tag{3.2}$$

Here, $p(\boldsymbol{\theta}|\mathbf{x})$ is the posterior probability for the parameters, $p(\mathbf{x}|\boldsymbol{\theta})$ is called the likelihood, $p(\mathbf{x})$ is the evidence and $p(\boldsymbol{\theta})$ is called the prior, and expresses what we know about the parameters prior to the experiment being done.
The evidence simply acts to normalize the probabilities

$$p(\mathbf{x}) = \int \mathrm{d}\boldsymbol{\theta}\, p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta}). \tag{3.3}$$

In the absence of any prior information, the prior is often assumed to be a constant, i.e. all parameter values are equally likely, this is referred to as flat prior. Under this assumption, setting $p(\boldsymbol{\theta}) = \text{const.}$, the posterior probability simplifies to

$$p(\boldsymbol{\theta}|\mathbf{x}) \propto L(\mathbf{x};\boldsymbol{\theta}), \tag{3.4}$$

where $L(\mathbf{x};\boldsymbol{\theta})$ is the likelihood.

3.4.2 Errors and marginalization

Let us assume we have a posterior probability distribution which is single-peaked. A common estimator of the parameters, denoted with a hat $\hat{\boldsymbol{\theta}}$, is the mean,

$$\hat{\boldsymbol{\theta}} = \int \mathrm{d}\boldsymbol{\theta}\, \boldsymbol{\theta}\, p(\boldsymbol{\theta}|\mathbf{x}). \tag{3.5}$$

An estimator is said to be unbiased if its expectation value coincides with the true value $\boldsymbol{\theta}_0$,

$$\langle\hat{\boldsymbol{\theta}}\rangle = \boldsymbol{\theta}_0\,. \tag{3.6}$$

Assuming for now a flat prior, so that the posterior is proportional to the likelihood, a Taylor expansion of the log-likelihood close to its peak shows that, locally, it is well described by a multivariate gaussian in parameter space,

$$\ln L(\mathbf{x};\boldsymbol{\theta}) = \ln L(\mathbf{x};\boldsymbol{\theta}_0) + \frac{1}{2}(\theta_\alpha - \theta_{0\alpha})\frac{\partial^2 \ln L}{\partial\theta_\alpha\partial\theta_\beta}(\theta_\beta - \theta_{0\beta}) + \dots\,, \tag{3.7}$$

or equivalently

$$L(\mathbf{x};\boldsymbol{\theta}) = L(\mathbf{x};\boldsymbol{\theta}_0)\exp\Big[-\frac{1}{2}(\theta_\alpha - \theta_{0\alpha})H_{\alpha\beta}(\theta_\beta - \theta_{0\beta})\Big]\,, \tag{3.8}$$

where we have defined the Hessian matrix

$$H_{\alpha\beta} \equiv -\frac{\partial^2 \ln L}{\partial\theta_\alpha\partial\theta_\beta}\,. \tag{3.9}$$

Comparing Eq. 3.8 with the general expression for a multivariate gaussian distribution, it is clear that $H_{\alpha\beta}$ plays the role of the precision matrix, i.e. the inverse of the covariance matrix of the parameters,

$$H_{\alpha\beta} = (C^{-1})_{\alpha\beta}\,, \qquad C_{\alpha\beta} \equiv \langle(\theta_\alpha - \theta_{0\alpha})(\theta_\beta - \theta_{0\beta})\rangle\,. \tag{3.10}$$

This identification allows the Hessian to be directly related to the uncertainties on the parameters: the diagonal elements of $H^{-1} \equiv C$ give the variances of the individual estimates, while the off-diagonal elements encode their correlations. Note that this is a statement about the estimates of the parameters, not about the parameters themselves, which may be physically independent but still have correlated estimates if they affect the data in a similar way.

Conditioning

If all the parameters except one, say θ_α, are held fixed, the error on θ_α is given by the curvature of the likelihood (or posterior, if the prior is not flat) along the θ_α direction,

$$\sigma_{\text{conditional},\alpha} = \frac{1}{\sqrt{H_{\alpha\alpha}}}\,. \tag{3.11}$$

This is called the conditional error: it is the minimum error bar attainable on θ_α, achieved only if all the other parameters are known exactly. Because this is rarely the case, the conditional error is rarely the relevant quantity, and should almost never be quoted as the final result of a parameter estimation.

Marginalizing

The correct treatment consists instead in marginalizing the posterior over all the other parameters. The marginal distribution of θ_1, for instance, is obtained by integrating over $\theta_2, \dots, \theta_n$,

$$p(\theta_1) = \int \mathrm{d}\theta_2 \dots \mathrm{d}\theta_n\, p(\boldsymbol{\theta})\,, \tag{3.12}$$

a procedure referred to as *marginalization*.
Assuming a multivariate gaussian likelihood, as in Eq. 3.8, one can show that the marginal error on the parameter θ_α is

$$\sigma_\alpha = \sqrt{(H^{-1})_{\alpha\alpha}}\,, \tag{3.13}$$

i.e. one inverts the Hessian matrix, and then takes the square root of the diagonal elements.

3.4.3 Fisher matrix

In forecasting applications, where no data are yet available, one works instead with the expectation value of the Hessian, known as the Fisher matrix,

$$F_{\alpha\beta} \equiv \langle H_{\alpha\beta} \rangle = \Big\langle -\frac{\partial^2 \ln L}{\partial\theta_\alpha \partial\theta_\beta} \Big\rangle\,. \tag{3.14}$$

The expected marginal error on θ_α is then

$$\sigma_\alpha = \sqrt{(F^{-1})_{\alpha\alpha}}\,, \tag{3.15}$$

which is always at least as large as the corresponding conditional error, $\sigma_{\text{conditional},\alpha} = 1/\sqrt{F_{\alpha\alpha}}$.

The Cramér-Rao bound

The power of the Fisher matrix derives from the Cramér-Rao inequality, which states that:

$$\Delta\theta_i \geq \sqrt{(\mathbf{F}^{-1})_{ii}}\,. \tag{3.16}$$

This relation places a firm lower limit on the error bars that one can attain. Eq, 3.16 can be proven using the Schwarz and assuming we have an unbiased estimator, as shown in [50].

3.4.4 Triangle plot

To visualize the results of a Fisher forecast, it is common to present the constraints in a triangle plot, like the one in figure 3.4. In this representation, the diagonal panel show the marginalized one-dimensional probability distribution for each parameter, while the off-diagonal panels show, for every pair of parameters, an ellipse. These ellipses arise from the assumption, already used to define the Fisher matrix, that the likelihood of the observable is well approximated by a multivariate Gaussian distribution: under this assumption, the parameters themselves are also approximately Gaussian distributed around their true values, with covariance matrix given by the inverse of the Fisher matrix. An ellipse is a contour of constant probability of this Gaussian distribution, restricted to the two considered parameters, and is fully determined by the corresponding 2×2 sub-matrix of F^{-1}.
These plots are read as follows: the size of the ellipse reflects how tightly the two parameters are constrained, while the orientation and elongation reveal the degree of correlation, or degeneracy, between them. A narrow, tilted ellipse indicates that the two parameters are strongly degenerate, meaning that a change in one parameter

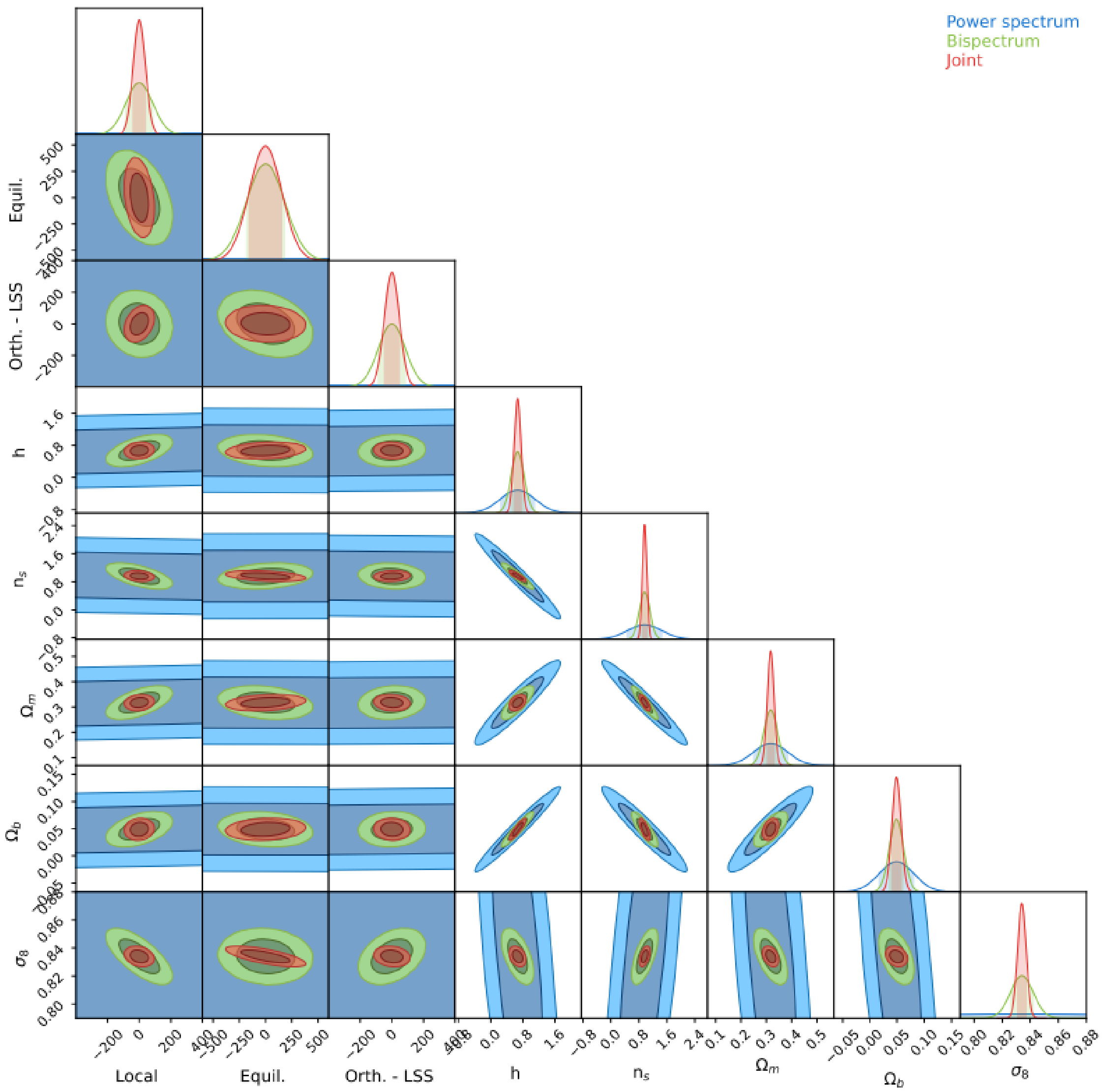

Figure 3.4: Joint constraints on the standard cosmological parameters and three shapes of primordial non-Gaussianity. Figure is from [51]

has the same effect on the data as a change in the other parameter. A nearly circular ellipse instead indicates that the two parameters are constrained independently. Comparing the ellipses obtained from different observables make it possible to visualize how and if the inclusion of additional information reduces the size of the ellipses or change their orientation by breaking parameter degeneracies. From the figure, it is clear that the bispectrum constraints are consistently tighter than those from the power spectrum alone, confirming that the bispectrum carries additional cosmological information. Combining the two observables tightens the constraints even more.

Chapter 4

The Marked Power Spectrum

Galaxy surveys constitute one of the primary probes of modern cosmology, providing increasingly precise measurements of the Large-Scale Structure (LSS) of the Universe and offering unique constraints on the nature of dark matter, dark energy and gravity. In recent years, galaxy surveys have made remarkable progress and forthcoming LSS observations will give us access to a significant amount of new cosmological information from non linear scales in the galaxy density field. The main challenge will be to extract as much cosmological information as possible from the resulting datasets.

In the previous sections, we established that the primary method of extracting cosmological information is through Fourier transform of the two-point correlation function, namely the power spectrum $P(k)$. At early times, deviations from Gaussianity in the Universe are minimal, making the power spectrum a sufficient statistic that captures nearly all cosmological information. However, at late times, non-linear effects become prominent, leading to significant deviations from Gaussianity and redistributing information into higher-order statistics, notably the bispectrum and the trispectrum.

To address this problem, the most direct approach consists in measuring higher order correlators of the field, starting with the bispectrum, the Fourier transform of the three-point correlation function [52, 53, 54, 55, 56, 57]. The primary limit of this approach is that extending it to higher-order correlators presents both numerical and computational challenges, especially when accounting for realistic data analysis where, for example, we need to evaluate covariances and include window functions.

For this reason, recently few alternatives have been developed: the first one is to design new summary statistics of the density field that are cheap to compute, analytically tractable and yet capable of compressing information from higher order correlators, the second one is to extract information directly at field-level, avoiding any statistical compression [58, 59, 60, 61].

A partial list of these new summary statistics includes the following methods: Wavelet Scattering Transform [62, 63, 64, 65, 66], k-nearest neighbors [67, 68], skew-spectra [69, 70], clipping [71] and Minkowski functionals [72, 73].

In this thesis we focus on another statistic that has recently attracted considerable interest in literature: the marked power spectrum [74, 11, 9]. The idea behind the marked statistics is to apply a non-linear transformation to the density field to obtain a weighted version of the density field where the weight (or “mark”) can encode galaxy properties, halo merger history, or information on the local environmental density, as in this work. In [13], the authors show that applying a non-linear transformation to the cosmic density field can partially restore its Gaussianity. Therefore, part of the information that is encoded in the higher order correlators is transferred back to the power spectrum, significantly improving parameter constraints from two-point statistics.

Summarizing, this additional constraining power emerges from two key factors. First, as demonstrated in [12], by correlating the density field with its environment the marked power spectrum effectively incorporates information from higher-order correlators. Secondly, the mark can be chosen to up-weight low-density regions [8]. This is crucial because these regions are the ones from which the signal is most accessible since they are the least affected by non linear gravitational evolution.

It is worth noticing that the claim of [12] is that the information is encoded in the higher-order correlation functions rather than in the environment, as stated in previous works [11, 75].

In this chapter we follow [12] and present the one-loop perturbative description of the marked power spectrum $M(k)$ for general tracers.

4.1 Definitions

The standard statistic used in LSS analysis is the power spectrum of the galaxy number density perturbation:

$$n_g(x) = \bar{n}_g(1 + \delta_g(x)) \tag{4.1}$$

where $\bar{n}_g$ is the mean galaxy density and $\delta_g(x)$ is the overdensity field.

In the considered marked formalism the galaxy number density is weighted by a space-dependent mark:

$$\rho_M(\mathbf{x}) = m(\mathbf{x})n_g(\mathbf{x}) = m(\mathbf{x})\bar{n}\,[1 + \delta_g(\mathbf{x})]\ ,$$

The mark function usually adopted in literature is [9, 8]

$$m(\mathbf{x}) = \left(\frac{1 + \delta_s}{1 + \delta_s + \delta_{g,R}(\mathbf{x})}\right)^p\ , \tag{4.2}$$

where $\delta_{g,R}(\mathbf{x})$ is the smoothed galaxy overdensity field. It is the convolution of the smoothing kernel W_R with the density contrast field $\delta(\mathbf{x})$, given by:

$$\delta_R(\mathbf{x}) = \int \mathrm{d}^3\mathbf{x}'\, W_R(\mathbf{x} - \mathbf{x}')\delta(\mathbf{x}')\,, \tag{4.3}$$

where the window function is assumed to be normalized

$$\int \mathrm{d}^3\mathbf{x}'\, W_R(\mathbf{x} - \mathbf{x}') = 1\,. \tag{4.4}$$

We illustrate the effect of such a smoothing in Figure 4.1 for varying values of R.

In 4.2 R, δ_s and p are parameter of the model. In particular, p can be used to enhance underdense ($p > 0$) or overdense ($p < 0$) regions.

This is clear in the limit for $\delta_s \to 0$, that leads to $m(x) \to [\bar{\rho}/\rho_R(\mathbf{x})]^p$. In this case it is evident that for $p > 0$ we are increasing the weight of points in low-density regions. It's worth noting that enhancing the weight in low-density regions is conceptually similar to studying the properties of voids, which are often used as probes for modified gravity [77], but without the need to identify voids.

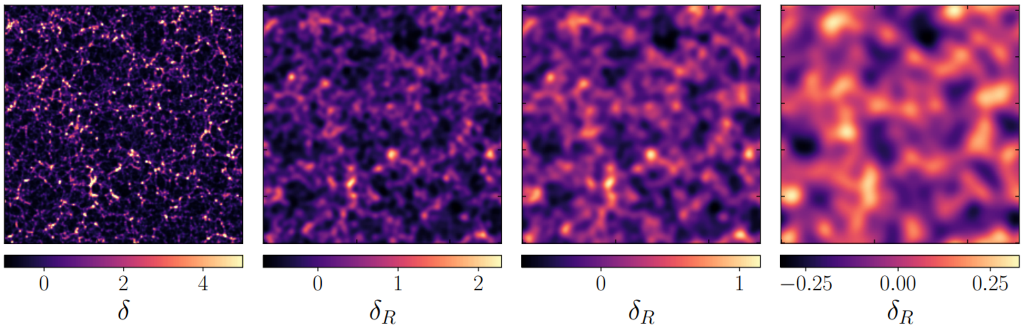

Figure 4.1: Density field for a slice of one realization of the Quijote simulations at $z = 0$ (left plot). We also show the smoothed density field at the same redshift for the radii $R = 10\text{Mpc}/h$, $15\text{Mpc}/h$, $30\text{Mpc}/h$. Figure is from [76].

By defining the mean mark $\bar{m}$ as $\langle \rho_M(\mathbf{x})\rangle = \langle m(\mathbf{x}) n_g(\mathbf{x})\rangle \equiv \bar{m}\bar{n}$, we can define the marked overdensity field δ_M:

$$\delta_M(\mathbf{x}) \equiv \frac{\rho_M(\mathbf{x}) - \langle \rho_M\rangle}{\langle \rho_M\rangle} = \frac{m(\mathbf{x})}{\bar{m}} \left[1 + \delta_g(\mathbf{x})\right] - 1 \,. \tag{4.5}$$

In figure 4.2, we show the effect of the mark on the matter density field. From the figure is further evident that a positive value of p has the effect of down-weighting high density regions, while it up-weights low density regions.

4.2 Eulerian perturbation theory

We proceed by expanding the marked overdensity in powers of the non linear fields δ_g and $\delta_{g,R}$, which will allow power spectra to be computed perturbatively. As specified also in [9] to obtain a consistent theory at one-loop order, we need to expand up to third order.
We start by approximating the mark by its Taylor series in $\delta_{g,R}(\mathbf{x})$

$$\begin{aligned} m(\mathbf{x}) &= \left(\frac{1+\delta_s}{1+\delta_s+\delta_{g,R}(\mathbf{x})}\right)^p = \left(1 + \frac{\delta_{g,R}(\mathbf{x})}{1+\delta_s}\right)^{-p} \\ &= 1 - \frac{p}{1+\delta_s}\delta_{g,R}(\mathbf{x}) + \frac{1}{2}\frac{p(p+1)}{(1+\delta_s)^2}\delta^2_{g,R}(\mathbf{x}) - \frac{1}{6}\frac{p(p+1)(p+2)}{(1+\delta_s)^3}\delta^3_{g,R}(\mathbf{x}) + \dots \\ &= \sum_{n=0}^{\infty}(-1)^n C_n \delta^n_{g,R}(\mathbf{x}) \,, \end{aligned} \tag{4.6}$$

where

$$C_n = \frac{p(p+1)(p+2)\dots(p+n-1)}{n!}\frac{1}{(1+\delta_s)^n} \,. \tag{4.7}$$

Before proceeding, we need to fix the condition for the series in 4.6 to converge: taking a large enough R, the time evolution of the smoothed field is $\delta_{g,R} \simeq b_1(z)D(z)\delta_0$, with b_1 and D being the linear bias and the growth factor, resulting in the condition $|\delta_{g,R}(\mathbf{x})/(1+\delta_s)| < 1$. Noting that the fluctuation of the smoothed field is $\sigma^2_{RR}(z) =$

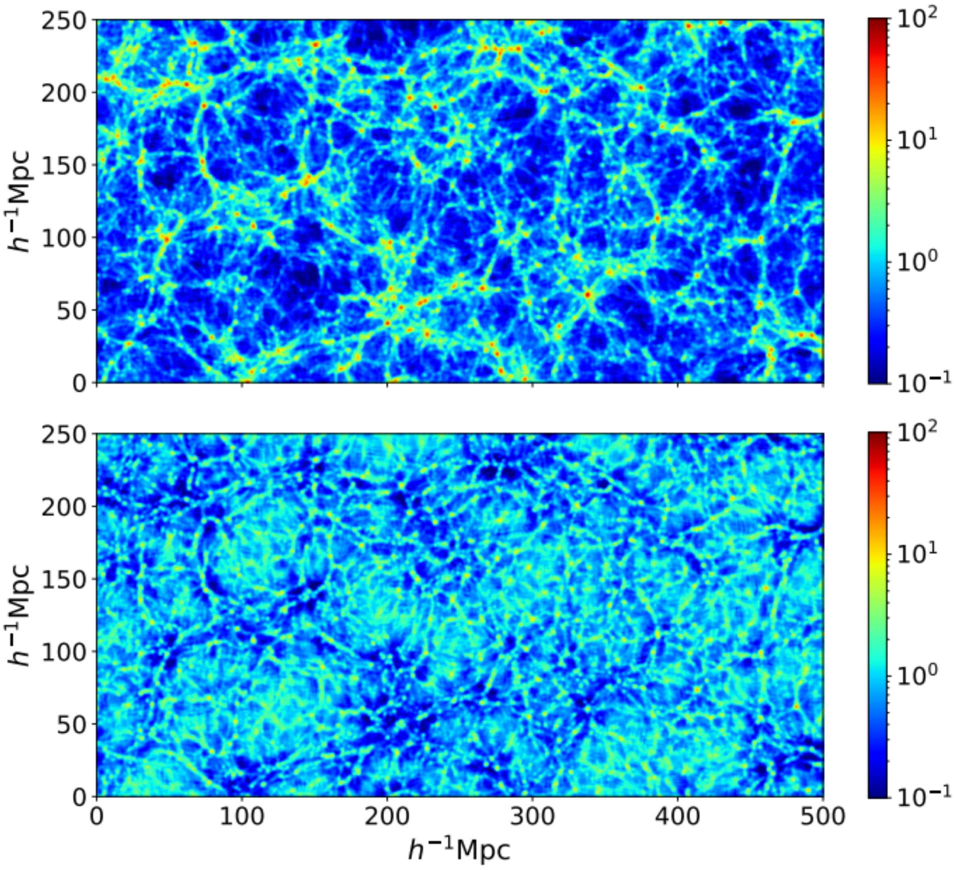

Figure 4.2: Projected matter density field (top panel) and marked density field (bottom panel) with parameters $R = 10h^{-1}$Mpc, $p = 2$ and $\delta_s = 0.25$. Figure is from [11].

$\langle \delta^2_{g,R}(\mathbf{x}) \rangle$, we expect convergence if

$$\sigma_{RR}(0) < \frac{1 + \delta_S}{b_1(z) D(z)} \,. \tag{4.8}$$

We proceed by expanding perturbatively the fields

$$\delta_g(\mathbf{x}) = \sum_{n=0}^{\infty} \delta_g^{(n)}(\mathbf{x}) \qquad\qquad \delta_{g,R}(\mathbf{x}) = \sum_{n=0}^{\infty} \delta_{g,R}^{(n)}(\mathbf{x}) \,, \tag{4.9}$$

where, in Fourier space, we have $\delta^{(n)}_{g,R}(\mathbf{k}) = W_R(k) \delta^{(n)}_g(\mathbf{k})$.
Inserting 4.9 in 4.5, we obtain

$$\begin{aligned} \bar{m}(\delta_M + 1) &= [1 + \delta_g(\mathbf{x})] \, m(\mathbf{x}) \\ &= \left[1 + \sum_{i=1}^{\infty} \delta_g^{(i)}(\mathbf{x})\right] \sum_{n=0}^{\infty} (-1)^n C_n \Big(\sum_{j=1}^{\infty} \delta_g^{(j)} \Big) \\ &= 1 + \left(\delta_M^{(1)} + \delta_M^{(2)} + \delta_M^{(3)} + \dots \right) \,, \end{aligned} \tag{4.10}$$

from which

$$
\begin{aligned}
\delta_M^{(1)}(\mathbf{x}) =& C_0\delta_g^{(1)}(\mathbf{x}) - C_1\delta_{g,R}^{(1)}(\mathbf{x})\,,\\
\delta_M^{(2)}(\mathbf{x}) =& C_0\delta_g^{(2)}(\mathbf{x}) - C_1\delta_{g,R}^{(1)}(\mathbf{x})\delta_g^{(1)}(\mathbf{x}) - C_1\delta_{g,R}^{(2)}(\mathbf{x}) + C_2\delta_{g,R}^{(1)}(\mathbf{x})\delta_{g,R}^{(1)}(\mathbf{x})\,,\\
\delta_M^{(3)}(\mathbf{x}) =& C_0\delta_g^{(3)}(\mathbf{x}) - C_1\delta_{g,R}^{(1)}(\mathbf{x})\delta_g^{(2)}(\mathbf{x}) - C_1\delta_{g,R}^{(2)}(\mathbf{x})\delta_g^{(1)}(\mathbf{x}) - C_1\delta_{g,R}^{(3)}(\mathbf{x})\\
&+ C_2\delta_{g,R}^{(1)}(\mathbf{x})\delta_{g,R}^{(1)}(\mathbf{x})\delta_g^{(1)}(\mathbf{x}) + 2C_2\delta_g^{(1)}(\mathbf{x})\delta_{g,R}^{(2)}(\mathbf{x}) +\\
&- C_3\delta_{g,R}^{(1)}(\mathbf{x})\delta_{g,R}^{(1)}(\mathbf{x})\delta_{g,R}^{(1)}(\mathbf{x})\,.
\end{aligned}
\tag{4.11}
$$

Next, we expand the various field involved into power of the linear matter perturbation field $\delta^{(1)}(\mathbf{x})$. For biased tracers in real space:

$$
\begin{aligned}
\delta_g^{(n)}(\mathbf{k},z) = \int \frac{\mathrm{d}^3\mathbf{q}_n}{(2\pi)^3}\cdots\int \frac{\mathrm{d}^3\mathbf{q}_n}{(2\pi)^3}(2\pi)^3\delta^D(\mathbf{k}-\mathbf{q}_{1\ldots n})\\
\times K_n(\mathbf{q}_1,\ldots,\mathbf{q}_n;z)\delta^{(1)}(\mathbf{q}_1)\ldots\delta^{(1)}(\mathbf{q}_n)\,,
\end{aligned}
\tag{4.12}
$$

where K_n are the real space kernels for a generic biased tracer [33]. Analogously, the kernels for the marked field are defined as

$$
\begin{aligned}
\delta_M^{(n)}(\mathbf{k},z) = \int \frac{\mathrm{d}^3\mathbf{q}_n}{(2\pi)^3}\cdots\int \frac{\mathrm{d}^3\mathbf{q}_n}{(2\pi)^3}(2\pi)^3\delta^D(\mathbf{k}-\mathbf{q}_{1\ldots n})\\
\times H_n(\mathbf{q}_1,\ldots,\mathbf{q}_n;z)\delta^{(1)}(\mathbf{q}_1)\ldots\delta^{(1)}(\mathbf{q}_n)\,.
\end{aligned}
\tag{4.13}
$$

with

$$
\begin{aligned}
H_1(\mathbf{k}) =& C_{\delta_M}(k)K_1(\mathbf{k})\,,\\
H_2(\mathbf{k}_1,\mathbf{k}_2) =& C_{\delta_M}(k)K_2(\mathbf{k}_1,\mathbf{k}_2) + C_{\delta_M^2}(k_1,k_2)K_1(\mathbf{k}_1)K_2(\mathbf{k}_2)\,,\\
H_3(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) =& C_{\delta_M}(k)K_3(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) + 2C_{\delta_M^2}(k_1,k_{23})K_1(\mathbf{k}_1)K_2(\mathbf{k}_2,\mathbf{k}_3) +\\
& C_{\delta_M{}^3}(k_1,k_2,k_3)K_1(\mathbf{k}_1)K_1(\mathbf{k}_2)K_1(\mathbf{k}_3)\,,
\end{aligned}
$$

where the H_3 still needs to be symmetrized. We define the functions

$$
\begin{aligned}
C_{\delta_M}(k) =& 1 - C_1W_R(k)\,,\\
C_{\delta_M^2}(k_1,k_2) =& C_2W_R(k_1)W_R(k_2) - \frac{C_1}{2}\left[W_R(k_1)+W_R(k_2)\right]\,,\\
C_{\delta_M^3}(k_1,k_2,k_3) =& -C_3W_R(k_1)W_R(k_2)W_R(k_3) +\\
& \frac{C_2}{3}\left[W_R(k_2)W_R(k_3)+W_R(k_1)W_R(k_3)+W_R(k_1)W_R(k_2)\right]\,.
\end{aligned}
$$

4.2.1 Final expressions for the marked power spectra

Finally from this expansion it is possible to compute the power spectrum of the marked density field up to third perturbative order

$$
\bar{m}^2M(k) = \bar{m}^2\left|\delta_M(\mathbf{k})\right|^2 = M_{11}(k) + M_{22}(k) + 2M_{13}(k)\,,
$$

with

$$
M_{11}(k) = \delta_M^{(1)}(\mathbf{k})\delta_M^{(1)}(\mathbf{k}') = H_1(\mathbf{k})\delta^{(1)}(\mathbf{k})H_1(\mathbf{k}')\delta^{(1)}(\mathbf{k}') = H_1^2(\mathbf{k})P_L(k)\,, \tag{4.14}
$$

$$
M_{22}(k) = \delta_M^{(2)}(\mathbf{k})\delta_M^{(2)}(\mathbf{k}') = 2\int\frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3}\left|H_2(\mathbf{k}-\mathbf{q},\mathbf{q})\right|^2P_L(q)P_L(|\mathbf{k}-\mathbf{q}|)\,, \tag{4.15}
$$

$$
M_{13}(k) = \delta_M^{(1)}(\mathbf{k})\delta_M^{(3)}(\mathbf{k}') = 3H_1(\mathbf{k})P_L(k)\int\frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3}H_3(\mathbf{k},\mathbf{q},-\mathbf{q})P_L(q)\,, \tag{4.16}
$$

where P_L is the linear power spectrum. The explicit calculation for $M_{13}(\mathbf{k})$ is reported in Appendix B.

4.3 Bias expansion

The kernels K_n need to be evaluated using a given prescription for the bias expansion. We follow the bias expansion described in [17] up to third order:

$$\delta_g(\mathbf{x}) = b_1\delta(\mathbf{x}) + \frac{b_2}{2}\delta^2(\mathbf{x}) + b_{G_2}\mathcal{G}_2(\mathbf{x}) + b_{\Gamma_3}\Gamma_3(\mathbf{x})\,, \tag{4.17}$$

The only operators that will enter the one-loop power spectrum are the four renormalized operators δ, $[\delta^2]$, $[\mathcal{G}_2]$ and $[\Gamma_3]$, related to the usual field via

$$[\delta^2](\mathbf{x}) = \delta^2 - \sigma^2\,, \qquad\qquad [\mathcal{G}_2](\mathbf{x}) = \mathcal{G}_2(\mathbf{x}) \tag{4.18}$$

with σ^2 being the variance of the unsmoothed field. By subtracting this term, we ensure that $\langle\delta_g\rangle = 0$. This renormalization step removes the contribution from large-scale fluctuations and guarantees that only the short-scale fluctuations contribute to the power spectrum, keeping the theory physically consistent. For simplicity, we set b_{Γ_3} to zero.
The other terms in 2.74, the operators δ^3, $\delta\mathcal{G}_2$ and $\mathcal{G}_3$, at one-loop order do not introduce new k-dependence and can therefore be absorbed into the bias parameters of lower-order terms.
The inclusion of renormalized operators leads to the addition of a zero-lag contribution. The contribution due to renormalized operators for galaxy bias, in Fourier space, manifests as

$$\delta_g(\mathbf{k}) \to \delta_g(\mathbf{k}) - \frac{b_2}{2}\sigma^2(2\pi)^3\delta_D(\mathbf{k}), \tag{4.19}$$

and, for the marked field at second and third perturbative orders,

$$\delta_M^{(2)}(\mathbf{k}) \to \delta_M^{(2)}(\mathbf{k}) - \frac{b_2}{2}\sigma^2 C_{\delta_M}(k)(2\pi)^3\delta_D(\mathbf{k}), \tag{4.20}$$

$$\delta_M^{(3)}(\mathbf{k}) \to \delta_M^{(3)}(\mathbf{k}) - b_2\sigma^2 C_{\delta_M^2}(k,0)\delta^{(1)}(\mathbf{k}). \tag{4.21}$$

The inclusion of the third-order contribution is crucial for the small-scale safety of the one-loop marked power spectrum.

4.4 Counterterms, shot noise and IR resummation

These results must be corrected in order to achieve better agreement with non linear scales. Following the formalism in 2.3.1 we add a counterterm of the form

$$M_{\rm ct}(k) = -2c_s^2k^2C^2_{\delta_M}(k)P_L(k)\,, \tag{4.22}$$

with c_s^2 being redshift dependent. Ref. [9] shows that this c_s^2 counterterm is the only one needed to capture the UV divergences of the one-loop $M(k)$ theory in real space. All other terms are manifestly convergent for hard loop momenta $p \gg k$ due to the presence of smoothing windows that depend on the physical scale R used in the

definition of the mark.
In the case of biased tracers, additional UV divergences arise for M_{13} and M_{22}: the M_{13} contribution vanishes when including the proper renormalized biased operator as in 4.18, while the M_{22} contains a divergent term that can be completely absorbed when considering the shot noise term. The stochastic shot noise term for the marked power spectrum takes the form:

$$M_{\rm stoch}(k) = C^2_{\delta_M}(k) P_{\rm shot} \,. \tag{4.23}$$

The last aspect to consider is the effect of infrared resummation. Following the approximation presented in [9, 10] we replace P_{NL} with its IR-resummed form in terms involving $C^2_{\delta_M}(k)P_{NL}$, while for other terms that involve the linear power spectrum $P_L(k)$, we replace it with the IR-resummed one.
The IR-resummed power spectrum for matter has the form 2.71. Similar relations hold for biased tracers.
Therefore the final expression for the EFT marked power spectrum for biased tracers in real space is the following:

$$M(k) = M_{11}(k) + M_{22}(k) + 2M_{13}(k) + M_{\rm ct}(k) + M_{\rm stoch}(k) \,.$$

4.5 Matter tidal field

Existing studies of the marked power spectrum have mainly focused on marks constructed from the local density field [11, 9, 8], but the local density provides only a partial description of the environment in which galaxies form and evolve. The structure of the cosmic web is determined by gravitational collapse, which is strongly anisotropic: its geometry is encoded in the tidal tensor, whose eigenvalues determine the local morphology, distinguishing voids, sheets, filaments and knots.
Since halo formation and galaxy evolution are expected to depend on both density and tidal environment[14, 15, 16], tidal field observables are a physically motivated extension of density-based marked statistics.
For this reason, in the following we will extend the marked statistics framework presented in the previous chapter to a broader class of marks. In particular, we will evaluate the marked power spectrum at one-loop order for a mark that depends on the contraction of the matter tidal field: $T_{ij}(x)T^{ij}(x) = \partial_i\partial_j\phi(x)\partial^i\partial^j\phi(x)$, where the gravitational potential $\phi(x)$ is related to the density contrast $\delta(x)$ via the Poisson equation. After the development of the theoretical framework we also perform a numerical implementation to evaluate the spectra, which will be used to perform a Fisher forecast analysis.

4.5.1 Definitions

In analogy with the previous section, we define the smoothed operator:

$$T_{R,ij}(x) = \int \mathrm{d}^3x' \, W_R(x - x') T_{ij}(x') \,, \tag{4.24}$$

this convolution, in Fourier space, becomes $T_{Rij} = W_R T_{ij}$.
We define the mark:

$$m(x) = \left(1 + \frac{T_{Rij}(x)T^{ij}_R(x) - \lambda^2}{1 + \delta_{s^2}}\right)^{-p} \,, \tag{4.25}$$

where δ_{s^2} and p are parameters of the model. The parameter λ^2 is introduced since the operator $T_{ij}T^{ij}$ does not oscillate around zero, unlike δ.
Also in this case we can increase the weight of the regions characterized by a lower value of the tidal tensor if we choose $p > 0$.

4.5.2 Eulerian perturbation theory

We perform Eulerian perturbation theory:

$$\begin{aligned} m(x) \simeq & 1 - \frac{p}{1+\delta_{s^2}} \left(T_{Rij} T_R^{ij} - \lambda^2 \right) + \frac{1}{2} \frac{p(p+1)}{(1+\delta_{s^2})^2} \left(T_{Rij} T_R^{ij} - \lambda^2 \right)^2 \\ & - \frac{1}{3!} \frac{p(p+1)(p+2)}{(1+\delta_{s^2})^3} \left(T_{Rij} T_R^{ij} - \lambda^2 \right)^3 + \dots \\ = & \sum_{n=0}^{\infty} (-1)^n D_n \left(T_{Rij} T_R^{ij} - \lambda^2 \right)^n , \end{aligned} \tag{4.26}$$

with

$$D_n = \frac{p(p+1)(p+2)\dots(p+n-1)}{n!\,(1+\delta_{s^2})^n} . \tag{4.27}$$

Also here we need to fix the condition for the series in 4.26 to converge. Since it is a binomial series, we expect convergence if:

$$\left| \frac{T_{Rij} T_R^{ij} - \lambda^2}{1+\delta_{s^2}} \right| < 1 , \tag{4.28}$$

defining $\sigma^2_{T^2}(z) = \left\langle (T_{Rij} T_R^{ij} - \lambda^2)^2 \right\rangle$, and using the linear evolution relation $T_{ij}(z) = D(z) T_{ij}^{(0)}$ this results is requiring

$$\sigma^2_{T^2}(0) < \frac{1+\delta_{s^2}}{D^2(z)} . \tag{4.29}$$

The expansions of δ_g and ϕ:

$$\delta_g(\mathbf{x}) = \sum_{n=0}^{\infty} \delta_g^{(n)}(\mathbf{x}) \qquad\qquad \phi(\mathbf{x}) = \sum_{n=0}^{\infty} \phi^{(n)}(\mathbf{x}) . \tag{4.30}$$

We expand also out tidal operator:

$$\begin{aligned} T_{Rij}(x) {T_R}^{ij}(x) = & \int \mathrm{d}^3 x_1 \, W_R(x-x_1) \partial_i \partial_j \left[\phi^{(1)}(x_1) + \phi^{(2)}(x_1) + \dots \right] \\ & \int \mathrm{d}^3 x_2 \, W_R(x-x_2) \partial_i \partial_j \left[\phi^{(1)}(x_2) + \phi^{(2)}(x_2) + \dots \right] \\ = & \int \mathrm{d}^3 x_1 \, W_R(x-x_1) \partial_i \partial_j \phi^{(1)}(x_1) \int \mathrm{d}^3 x_2 \, W_R(x-x_2) \partial_i \partial_j \phi^{(1)}(x_2) + \\ & + \int \mathrm{d}^3 x_1 \, W_R(x-x_1) \partial_i \partial_j \phi^{(1)}(x_1) \int \mathrm{d}^3 x_2 \, W_R(x-x_2) \partial_i \partial_j \phi^{(2)}(x_2) + \\ & + \int \mathrm{d}^3 x_1 \, W_R(x-x_1) \partial_i \partial_j \phi^{(2)}(x_1) \int \mathrm{d}^3 x_2 \, W_R(x-x_2) \partial_i \partial_j \phi^{(1)}(x_2) \\ = & [t_R^2]^{(2)} + [t_R^2]^{(3)} . \end{aligned} \tag{4.31}$$

Plugging everything into 4.5:

$$
\begin{aligned}
\bar{m}(1+\delta_M) =& m(x)\left[1+\delta_g(x)\right] \\
=&\Big[D_0 - D_1\left(T_{Rij}T_R^{ij} - \lambda^2\right) + D_2\left(T_{ij}T^{ij} - \lambda^2\right)^2 + \\
& D_3\left(T_{Rij}T_R^{ij} - \lambda^2\right)^3\Big]\Big[1+\delta_g^{(1)}+\delta_g^{(2)}+\delta_g^{(3)}+\dots\Big] \\
=& 1+\left[\delta_m^{(1)}+\delta_m^{(2)}+\delta_m^{(3)}+\dots\right] ,
\end{aligned} \tag{4.32}
$$

with

$$
\delta_M^{(1)} = D_0\delta_g^{(1)} , \tag{4.33}
$$

$$
\delta_M^{(2)} = D_0\delta_g^{(2)} - D_1[t_R^2]^{(2)} - 2D_2\lambda^2[t_R^2]^{(2)} , \tag{4.34}
$$

$$
\begin{aligned}
\delta_M^{(3)} =& D_0\delta_g^{(3)} - D_1[t_R^2]^{(2)}\delta_g^{(1)} - D_1[t_R^2]^{(3)} - 2D_2\lambda^2[t_R^2]^{(3)} + \\
& - 2D_2\lambda^2[t_R^2]^{(2)}\delta_g^{(1)} - 3D_3\lambda^4[t_R^2]^{(2)}\delta_g^{(1)} - 3D_3\lambda^4[t_R^2]^{(3)} .
\end{aligned} \tag{4.35}
$$

To evaluate the marked kernels we go to Fourier space, where:

$$
\partial_i\partial_j\phi^{(n)}(\mathbf{k}) = \frac{k_i k_j}{\mathbf{k}^2}\delta^{(1)}(\mathbf{k}) , \tag{4.36}
$$

$$
\begin{aligned}
[t_R^2]^{(2)}(\mathbf{k}) =& \int \frac{\mathrm{d}^3 q_1}{(2\pi)^3}\frac{\mathrm{d}^3 q_2}{(2\pi)^3}\frac{(\mathbf{q}_1\cdot\mathbf{q}_2)^2}{\mathbf{q}_1^2\mathbf{q}_2^2} W_R(q_1)W_R(q_2)\delta^{(1)}(\mathbf{q}_1)\delta^{(1)}(\mathbf{q}_2)\times \\
& \times (2\pi)^3\delta_D(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2) ,
\end{aligned} \tag{4.37}
$$

$$
\begin{aligned}
[t_R^2]^{(3)}(\mathbf{k}) =& 2\int \frac{\mathrm{d}^3 q_1}{(2\pi)^3}\frac{\mathrm{d}^3 q_2}{(2\pi)^3}\frac{\mathrm{d}^3 q_3}{(2\pi)^3}\frac{q_{1i}q_{1j}}{\mathbf{q}_1^2}\frac{(q_{2i}+q_{3i})(q_{2j}+q_{3j})}{(\mathbf{q}_2+\mathbf{q}_3)^2}F_2(\mathbf{q}_2,\mathbf{q}_3)\times \\
& \times W_R(q_1)W_R(q_2)\delta^{(1)}(\mathbf{q}_1)\delta^{(1)}(\mathbf{q}_2)\delta^{(1)}(\mathbf{q}_3)(2\pi)^3\times \\
& \times \delta_D(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2-\mathbf{q}_3) .
\end{aligned} \tag{4.38}
$$

Finally, using the definition of kernel 4.12 we recover:

$$
H_1(\mathbf{q}_1) = D_0 K_1(\mathbf{q}_1) , \tag{4.39}
$$

$$
\begin{aligned}
H_2(\mathbf{q}_1,\mathbf{q}_2) =& D_0K_2(\mathbf{q}_1,\mathbf{q}_2) - D_1 W_R(q_1)W_R(q_2)\frac{(\mathbf{q}_1\cdot\mathbf{q}_2)^2}{\mathbf{q}_1^2\mathbf{q}_2^2} + \\
& - 2D_2\lambda^2 W_R(q_1)W_R(q_2)\frac{(\mathbf{q}_1\cdot\mathbf{q}_2)^2}{\mathbf{q}_1^2\mathbf{q}_2^2} \\
=& D_0K_2(\mathbf{q}_1,\mathbf{q}_2) - (D_1+2D_2\lambda^2)W_R(q_1)W_R(q_2)\mathcal{G}(\mathbf{q}_1,\mathbf{q}_2) ,
\end{aligned} \tag{4.40}
$$

$$
\begin{aligned}
H_3(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) =& D_0K_3(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) - \Big(\frac{D_1}{3}+\frac{2}{3}D_2\lambda^2+D_3\lambda^4\Big)\times \\
& \times\Big[K_1(\mathbf{q}_3)W_R(q_1)W_R(q_2)\mathcal{G}(\mathbf{q}_1,\mathbf{q}_2) + K_1(\mathbf{q}_2)W_R(q_1)W_R(q_3)\mathcal{G}(\mathbf{q}_1,\mathbf{q}_3) + \\
& + K_1(\mathbf{q}_1)W_R(q_2)W_R(q_3)\mathcal{G}(\mathbf{q}_2,\mathbf{q}_3)\Big] - \Big(\frac{2}{3}D_1+\frac{4}{3}D_2\lambda^2+2D_3\lambda^4\Big)\times \\
& \times\Big[W_R(q_1)W_R(|\mathbf{q}_2+\mathbf{q}_3|)\mathcal{G}(\mathbf{q}_1,\mathbf{q}_2+\mathbf{q}_3)F_2(\mathbf{q}_2,\mathbf{q}_3) + W_R(q_2)\times \\
& \times W_R(|\mathbf{q}_1+\mathbf{q}_3|)\mathcal{G}(\mathbf{q}_2,\mathbf{q}_1+\mathbf{q}_3)F_2(\mathbf{q}_1,\mathbf{q}_3) + W_R(q_3)W_R(|\mathbf{q}_1+\mathbf{q}_2|)\times \\
& \times \mathcal{G}(\mathbf{q}_3,\mathbf{q}_1+\mathbf{q}_2)F_2(\mathbf{q}_1,\mathbf{q}_2)\Big]
\end{aligned} \tag{4.41}
$$

where we have defined $\mathcal{G}(\mathbf{q}_1,\mathbf{q}_2) = (\mathbf{q}_1\cdot\mathbf{q}_2)^2/(\mathbf{q}_1^2\mathbf{q}_2^2)$.

Final expressions for the marked power spectra

To evaluate the marked power spectra we use the results 4.14, 4.15 and 4.16:

$$M_{11}(k) = D_0^2 P_{11,g}(k)\,, \tag{4.42}$$

$$\begin{aligned} M_{22}(k) = D_0^2 P_{22,g}(k) + 2\int \frac{d^3q}{(2\pi)^3}\Big[&-2D_0\big(D_1+2D_2\lambda^2\big)K_2(\mathbf{q},\mathbf{k}-\mathbf{q})W_R(q)\times \\ &\times W_R(|\mathbf{k}-\mathbf{q}|)\mathcal{G}(\mathbf{q},\mathbf{k}-\mathbf{q}) + \big(D_1+2D_2\lambda^2\big)^2 W_R^2(q)W_R^2(|\mathbf{k}-\mathbf{q}|)\times \\ &\times \mathcal{G}^2(\mathbf{q},\mathbf{k}-\mathbf{q})\Big] P_L(k)P_L(|\mathbf{k}-\mathbf{q}|)\,, \end{aligned} \tag{4.43}$$

$$\begin{aligned} M_{13}(k) = D_0^2 P_{13,g}(k) - 3D_0P_L(k)K_1(\mathbf{k})\int \frac{d^3q}{(2\pi)^3}\Big\{ &K_1(\mathbf{q})\Big[\frac{D_1}{3} + 2\frac{D_2}{3}\lambda^2 + D_3\lambda^4\Big] \\ &\times \Big[2W_R(k)W_R(q)\mathcal{G}(\mathbf{k},\mathbf{q}) + W_R^2(q)\Big] - \frac{2}{3}\big(D_1+2D_2\lambda^2+3D_3\lambda^4\big)\times \\ &\times \Big[W_R(q)W_R(|-\mathbf{k}-\mathbf{q}|)\mathcal{G}(\mathbf{q},\mathbf{k}+\mathbf{q})F_2(-\mathbf{k},-\mathbf{q})+ \\ &+ W_R(q)W_R(|\mathbf{q}-\mathbf{k}|)\mathcal{G}(\mathbf{q},\mathbf{q}-\mathbf{k})F_2(-\mathbf{k},\mathbf{q})\Big]\Big\} P_L(q)\,, \end{aligned} \tag{4.44}$$

where to simplify the expression we use $K_1(-\mathbf{k}) = K_1(\mathbf{k})$, $W_R(0) = 1$, $W_R(-k) = W_R(k)$, $\mathcal{G}(-\mathbf{k},\mathbf{q}) = \mathcal{G}(-\mathbf{k},-\mathbf{q}) = \mathcal{G}(\mathbf{k},\mathbf{q})$, $\mathcal{G}(\mathbf{q},-\mathbf{q}) = 1$ and $F_2(\mathbf{q},-\mathbf{q}) = 0$.
The complete expression for the marked power spectrum is:

$$M(k) = M_{11}(k) + M_{22}(k) + 2M_{13}(k) + M_{ct}(k) + M_{stoch}(k)\,, \tag{4.45}$$

where $M_{ct}(k)$ and $M_{stoch}(k)$ have expressions similar to 4.22 and 4.23:

$$M_{ct} = D_0^2 b_1^2 P_{ct}(k) = -2D_0^2 b_1^2 c_s^2 k^2 P_{11}(k)\,, \tag{4.46}$$

$$M_{sthoch}(k) = D_0^2 b_1^2 P_{shot}\,. \tag{4.47}$$

To evaluate the kernels K_n we used the same prescription we used in 4.4. Same holds for IR-resummation, we follow the approximation of [9, 10].
It is worth noticing that these calculations should be understood in the context of an idealized, noiseless reconstruction of the matter tidal field from galaxy survey data. In realistic applications, the reconstruction is affected by sources of uncertainty, such as galaxy bias, redshift-space distortions, shot noise and survey geometry.
When we analyze a galaxy survey we don't observe neither the dark matter density field nor the dark matter tidal field. In a realistic analysis the reconstruction pipeline begins with the angular position and redshift of each galaxy, from which we compute the comoving position and we correct it for the selection function of the survey. Since the galaxy distribution is discrete, the continuous galaxy density field is obtained either by assigning galaxies to a three-dimensional grid [78] or by expanding the discrete distribution in a basis of continuous functions [79]. We then smooth the resulting field to suppress the shot noise that arises from the discrete sampling. The field is then corrected for redshift-space distortions and, because this field traces the galaxy distribution rather than the underlying dark matter distribution, we apply the galaxy bias to convert it into an estimate of the dark matter density field. Finally we solve the Poisson equation to find the gravitational potential, from which we evaluate the tidal tensor via $T_{ij} = \partial_i\partial_j\phi$.

For further details on the reconstruction procedure, please refer to [80, 78, 79].
Lastly, we evaluate λ^2:

$$\begin{aligned}
\lambda^2 &= \langle T_{Rij}(\mathbf{x}) T_R^{ij}(\mathbf{x})\rangle \\
&= \left\langle \int d^3x W_R(x-x_1)\partial_i\partial_j\phi(x_1) \int d^3x W_R(x-x_2)\partial_i\partial_j\phi(x_2) \right\rangle \\
&= \left\langle \int \frac{d^3k_1}{(2\pi)^3} e^{ik_1x} W_R(k_1)\frac{k_{1i}k_{1j}}{k_1^2}\delta(k_1) \int \frac{d^3k_2}{(2\pi)^3} e^{ik_2x} W_R(k_2)\frac{k_{2i}k_{2j}}{k_2^2}\delta(k_2) \right\rangle \\
&= \int \frac{d^3k_1}{(2\pi)^3}\frac{d^3k_2}{(2\pi)^3} e^{ix(k_1+k_2)} W_R(k_1)W_R(k_2)\frac{k_{1i}k_{1j}}{k_1^2}\frac{k_{2i}k_{2j}}{k_2^2}\langle\delta(k_1)\delta(k_2)\rangle \\
&= \int \frac{d^3k_1}{(2\pi)^3}\frac{d^3k_2}{(2\pi)^3} e^{ix(k_1+k_2)} W_R(k_1)W_R(k_2)\frac{k_{1i}k_{1j}}{k_1^2}\frac{k_{2i}k_{2j}}{k_2^2}(2\pi)^3\delta_D(k_1+k_2)P_L(k_1) \\
&= \int \frac{d^3k_1}{(2\pi)^3} W_R^2(k_1)\left(\frac{k_{1i}k_{1j}}{k_1^2}\right)^2 P_L(k_1) \\
&= \int \frac{d^3k}{(2\pi)^3} W_R^2(k)P_L(k)\,.
\end{aligned} \tag{4.48}$$

4.5.3 Numerical implementation

Galaxies power spectrum

The galaxies power spectrum

$$P_{1-loop,g}(k) = P_{11,g}(k) + P_{22,g}(k) + P_{13,g}(k)\,, \tag{4.49}$$

can be reconstructed using CLASS-PT [1] [81], which is a modification of the CLASS code [82, 83, 84]. The Cosmic Linear Anisotropy Solving System (CLASS) is a Boltzmann code with the purpose of simulating the evolution of linear perturbations in the universe and to compute CMB and LSS observables. CLASS-PT, instead, computes the non-linear power spectra of dark matter and biased tracers in one-loop cosmological perturbation theory, for both Gaussian and non-Gaussian initial conditions. This code uses the Logarithmic Fast Fourier Transform (FFTLog) to evaluate the one-loop integrals and compute the spectra. These spectra also incorporate all ingredients required for a direct application to data, for example, non-linear bias, redshift-space distortions and infra-red resummation.
CLASS-PT provides all the necessary pieces and we only need to multiply them by the corresponding bias parameters combination.
For the galaxy biases we adopt a semi-analytic model of galaxy formation: for the linear bias we adopt the model [85] [2]

$$b_1(z) = 0.9 + 0.4z\,. \tag{4.50}$$

For the higher order biases we adopt the following fitting formulae [86], obtained from a combination of N-body simulations and halo occupation modeling:

$$b_2(z) = -0.704 - 0.208z + 0.183z^2 - 0.00771z^3\,. \tag{4.51}$$

[1] https://github.com/Michalychforever/CLASS-PT

[2] Although this model was originally developed for Euclid-like Hα galaxies, we also apply it to our analysis because we focus on the relative differences between statistics, rather than their absolute values.

Lastly, for $b_{\mathcal{G}_2}$ we use instead the co-evolution model [87]:

$$b_{\mathcal{G}_2} = -\frac{2}{7}(b_1 - 1)\,. \tag{4.52}$$

Following [12] we fix the redshift to $z = 0.6$, this results in:

$$b_1 = 1.14\,, \tag{4.53}$$
$$b_2 = -0.38\,, \tag{4.54}$$
$$b_{\mathcal{G}_2} = -0.041\,. \tag{4.55}$$

From this we can straightforwardly implement $D_0^2 P_{1-loop,g}(k)$, to which we sum also the counterterm and the shot noise contribution.
For these last two contributions we follow [12] and we choose:

$$c_s^2 = 13.21[\mathrm{Mpc}/h]^2\,, \tag{4.56}$$
$$P_{\mathrm{shot}} = 3333\,[\mathrm{Mpc}/h]^3\,, \tag{4.57}$$

The prescription for c_s is the one from [88]:

$$c_s = -25D^2(z)\,, \tag{4.58}$$

with $D(z)$ being the usual growth factor for linear perturbations. For the stochastic term we simply adopt the Poissonian sampling prediction:

$$P_{\mathrm{shot}} = \bar{n}_g^{-1}\,. \tag{4.59}$$

One-loop term M_{13}

Starting from the formula 4.44, we split it into three terms:

- $M_{13}^A = D_0^2 P_{13_g}(k)\,,$

- $$M_{13}^B = -3D_0 P_L(k) K_1(\mathbf{k}) \int \frac{d^3q}{(2\pi)^3} \Big\{ K_1(\mathbf{q}) \Big[\frac{D_1}{3} + 2\frac{D_2}{3}\lambda^2 + D_3\lambda^4\Big] \times \\ \times \Big[2W_R(k)W_R(q)\mathcal{G}(\mathbf{k},\mathbf{q}) + W_R^2(q)\Big]\Big\} P_L(q)\,,$$

- $$M_{13}^C = 2D_0 P_L(k) K_1(\mathbf{k}) \int \frac{d^3q}{(2\pi)^3} \Big\{ \big(D_1 + 2D_2\lambda^2 + 3D_3\lambda^4\big) \times \\ \times \Big[W_R(q)W_R(|-\mathbf{k}-\mathbf{q}|)\mathcal{G}(\mathbf{q},\mathbf{k}+\mathbf{q})F_2(-\mathbf{k},-\mathbf{q}) + \\ + W_R(q)W_R(|\mathbf{q}-\mathbf{k}|)\mathcal{G}(\mathbf{q},\mathbf{q}-\mathbf{k})F_2(-\mathbf{k},\mathbf{q})\Big]\Big\} P_L(q)\,.$$

The first term is already accounted for by 4.49.
The second term can be simplified by solving the angular integral:

$$M_{13}^B(k) = -3D_0 b_1^2 P_L(k) \Big[\frac{D_1}{3} + 2\frac{D_2}{3}\lambda^2 + D_3\lambda^4\Big] \times \\ \times \int \frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3} \Big[2W_R(k)W_R(q)\mathcal{G}(\mathbf{k},\mathbf{q}) + W_R^2(q)\Big] P_L(q)\,, \tag{4.60}$$

since $K_1(\mathbf{q}) = K_1(\mathbf{k}) = b_1$.
We immediately recognize that the second term of the integral is just a rescaling of 4.48.
The first term can be evaluated as follows:

$$\begin{aligned}
&\int \frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3}\Big[2W_R(k)W_R(q)\mathcal{G}(\mathbf{k},\mathbf{q})\Big]P_L(q) = \\
&= \frac{2}{(2\pi)^3}W_R(k)\int_0^\infty \mathrm{d}q\, q^2 W_R(q)P_L(q)\int_0^{2\pi}\mathrm{d}\phi\int_0^\pi \mathrm{d}\theta\,\sin\theta\mathcal{G}(\mathbf{k},\mathbf{q}) \\
&= \frac{2}{(2\pi)^3}W_R(k)\int_0^\infty \mathrm{d}q\, q^2 W_R(q)P_L(q)\,(2\pi)\int_0^\pi \mathrm{d}\theta\,\sin(\theta)\cos^2\theta \\
&= \frac{2}{(2\pi)^2}W_R(k)\int_0^\infty \mathrm{d}q\, q^2 W_R(q)P_L(q)\int_{-1}^1 \mathrm{d}\mu\,\mu^2 \\
&= \frac{4}{3(2\pi)^2}W_R(k)\int_0^\infty \mathrm{d}q\, q^2 W_R(q)P_L(q)\,,
\end{aligned} \tag{4.61}$$

we now evaluate the one-dimensional integral using Simpson's rule.
Looking at M_{13}^C, we note that it is not possible to solve the angular integral due to the complex μ-dependence encoded in the window functions $W_R(|\mathbf{k}+\mathbf{q}|)$ and $W_R(|\mathbf{q}-\mathbf{k}|)$, so we perform the three-dimensional integration using Simpson's rule.

One-loop term M_{22}

As for $M_{13}(k)$, the first term of 4.43 is already accounted for by 4.49.
For the other terms we follow the same approach we use for M_{13}^C, since a complex μ-dependence doesn't allow to simplify the integral.

Results

Figure 4.3 compares the galaxy power spectrum and the marked power spectrum at redshift $z = 0.6$, computed with a Gaussian smoothing radius of $R = 10\mathrm{Mpc}/h$ for different values of the parameter $p = \{2, 1, -1\}$.
The marked power spectrum shows a clear dependence on the amplitude of the tidal field. As we already mentioned, for larger value of p the mark up-weights regions with a tidal field amplitude that is lower than the average. From this image we clearly see that the marked power spectrum $M(k)$ as enhanced amplitude with respect to $P(k)$ alone for these low-tidal regions, in particular, the effect is increasingly stronger for increasing p. Conversely, for negative values of p the mark is up-weighting regions with a value of the tidal field amplitude higher than the average and this results in a suppression of the marked power spectrum with respect to the power spectrum.
Figure 4.4 shows the individual contributions to the marked power spectrum $M(k)$, negative contributions are identified by dashed lines. The mark parameter are fixed to $R = 10\,\mathrm{Mpc}/h$, $p = 2$ and $\delta_{s^2} = 0.25$. In contrast with the power spectrum shown in figure 2.1 we notice that the 1-loop corrections M_{13} and M_{22} show a change of sign in the considered k interval. In particular, M_{22}, positive for the power spectrum, becomes negative towards the upper end of the interval, while M_{13} becomes positive at intermediate k. This behavior reflects the additional contributions introduced by the mark, which modify the kernels entering the one-loop corrections and therefore allow

the marked 1-loop contributions to have a different sign structure from their unmarked counterparts.

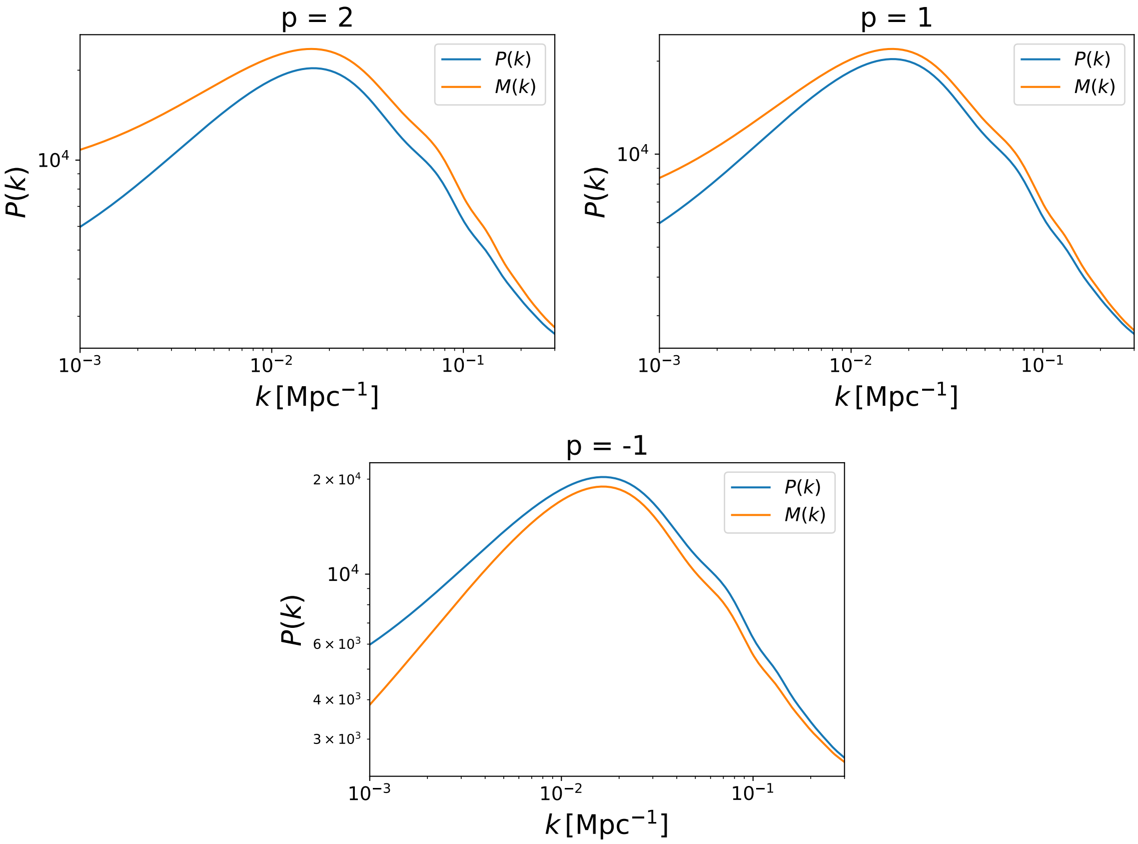

Figure 4.3: Plot of the power spectrum and the marked power spectrum for biased tracers at $z = 0.6$ for different values of the mark parameter $p = \{2, 1, -1\}$. Mark parameters: $R = 10\,\mathrm{Mpc}/h$ and $\delta_{s^2} = 0.25$.

4.5.4 Other marks

In this work, we adopt $T_{ij}T^{ij}$ as the tidal field mark, representing the most straightforward scalar quantity that can be constructed from the tidal tensor. This choice is particularly suitable as a first application of tidal information within the marked statistics framework. However, this contraction does not retain the full information encoded in the tidal tensor. In particular, it is insensitive to the individual eigenvalues and therefore does not directly distinguish between different cosmic-web environments, such as voids, sheets, filaments, and clusters.

The eigenvalues of the tidal tensor provide a more complete description of its local structure. In this section, we briefly discuss alternative scalar quantities that can be constructed from the eigenvalues.

The tidal tensor is a symmetric tensor and can therefore be characterized by its three eigenvalues, $\lambda_1, \lambda_2, \lambda_3$. To find the eigenvalues we use the condition:

$$\det\big(T - \lambda I\big) = 0\,, \tag{4.62}$$

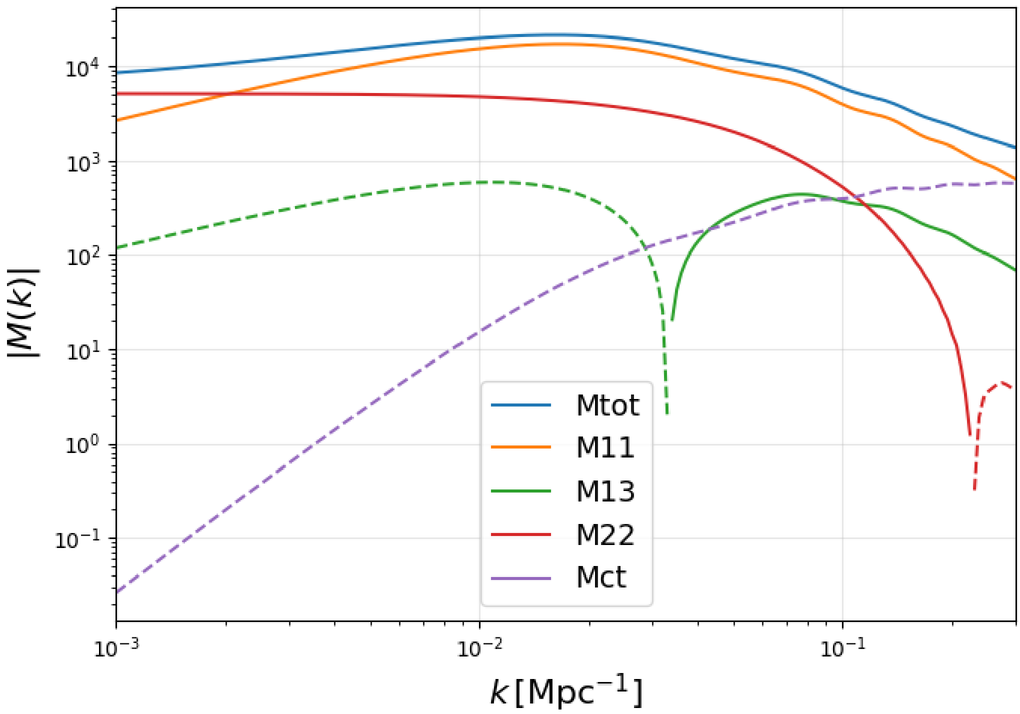

Figure 4.4: Plot of the individual components of the marked power spectrum for biased tracers at $z = 0.6$. The dashed line indicates a negative contribution. Mark parameters: $R = 10\,\mathrm{Mpc}/h$, $\delta_{s^2} = 0.25$ and $p = 2$.

where I is the identity matrix. This condition can be rewritten as:

$$\lambda^3 - I_1\lambda^2 + I_2\lambda - I_3 = 0\,, \tag{4.63}$$

where I_1, I_2, I_3 are scalars defined as:

$$I_1 = \mathrm{Tr}(T) = \phi_{xx} + \phi_{yy} + \phi_{zz} = \partial^2\phi = \delta\,, \tag{4.64}$$

$$\begin{aligned} I_2 &= \begin{vmatrix} \phi_{xx} & \phi_{xy} \\ \phi_{xy} & \phi_{yy} \end{vmatrix} + \begin{vmatrix} \phi_{yy} & \phi_{yz} \\ \phi_{yz} & \phi_{zz} \end{vmatrix} + \begin{vmatrix} \phi_{xx} & \phi_{xz} \\ \phi_{xz} & \phi_{zz} \end{vmatrix} \\ &= \phi_{xx}\phi_{yy} + \phi_{yy}\phi_{zz} + \phi_{zz}\phi_{xx} - \phi_{xy}^2 - \phi_{yz}^2 - \phi_{xz}^2\,, \end{aligned} \tag{4.65}$$

$$\begin{aligned} I_3 = \det(T) &= \phi_{xx}(\phi_{yy}\phi_{zz} - \phi_{yz}^2) + \\ &\quad - \phi_{xy}(\phi_{xy}\phi_{zz} - \phi_{yz}\phi_{xz}) + \phi_{xz}(\phi_{xy}\phi_{yz} - \phi_{yy}\phi_{xz})\,. \end{aligned} \tag{4.66}$$

These can be rewritten as:

$$I_1 = \partial^2\phi\,, \tag{4.67}$$

$$I_2 = \left(\partial^2\phi\right)^2 - (\partial_i\partial_j\phi)(\partial_i\partial_j\phi)\,, \tag{4.68}$$

$$I_3 = \frac{1}{6}\varepsilon_{ijk}\varepsilon_{lmn}\phi_{il}\phi_{jm}\phi_{kn}\,. \tag{4.69}$$

One can rewrite these scalars also in terms of the eigenvalues of the tidal tensor using Vieta's formula:

$$I_1 = \lambda_1 + \lambda_2 + \lambda_3\,, \tag{4.70}$$

$$I_2 = \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3\,, \tag{4.71}$$

$$I_3 = \lambda_1\lambda_2\lambda_3\,. \tag{4.72}$$

The three solutions to 4.63, $\{\lambda_1, \lambda_2, \lambda_3\}$ define the environment:

- *voids*: the region of space where T_{ij} has no eigenvalues bigger than a threshold λ_{th};
- *sheets*: the set of points with one eigenvalue bigger than λ_{th} and two eigenvalues smaller;
- *filaments*: the sites with two eigenvalues bigger than λ_{th} and one smaller eigenvalue;
- *clusters*: the zones with three eigenvalues bigger than λ_{th}.

Typical values used in the literature for the threshold λ lie in the range $\lambda_{th} \simeq 0.1 - 0.4$, depending on the smoothing scale [89]. For simplicity, in what follows we assume $\lambda_{th} = 0$.
Ordering the eigenvalues as $\lambda_1 > \lambda_2 > \lambda_3$, we know that the invariants I_1, I_2, I_3 themselves identify specific environments. First, I_3 depends only on how many negative eigenvalues there are, so a consequence is that if $I_3 > 0$ then that region belongs to a cluster or a sheet, if $I_3 < 0$ belongs to a filament or a void.

Clusters: $I_1 > 0,\ I_2 > 0,\ I_3 > 0$
Voids: $I_1 < 0,\ I_2 > 0,\ I_3 < 0$
Sheets: $I_3 > 0$ and NOT clusters
Filaments: $I_3 < 0$ and NOT voids

This classification may be very useful to construct mark functions that have the form:

$$m(\mathbf{x}) = f(I_1, I_2, I_3)\,, \tag{4.73}$$

An interesting function is the one proposed in [90], the so called tidal torque

$$t^2 = \frac{1}{2}\Big[(\lambda_1 - \lambda_2)^2 + (\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_3)^2\Big]\,, \tag{4.74}$$

which have a very intuitive physical interpretation. Indeed, using the Vieta's formulas we obtain

$$t^2 = I_1^2 - 3I_2\,. \tag{4.75}$$

This is strictly connected to the anisotropic part of the tidal field. We can see this by decomposing the tidal tensor into isotropic and traceless part

$$T_{ij} = \frac{I_1}{3}\delta_{ij} + T_{ij}^{\mathrm{TL}}\,, \tag{4.76}$$

where T_{ij}^{TL} is the traceless part, for which the eigenvalues are, by construction

$$\tilde{\lambda}_i = \lambda_i - \frac{I_1}{3}\,, \quad \text{with} \quad \tilde{\lambda}_1 + \tilde{\lambda}_2 + \tilde{\lambda}_3 = 0\,. \tag{4.77}$$

If we compute the trace of the squared traceless part we obtain

$$\mathrm{Tr}\Big[T_{ij}^{\mathrm{TL}} T_{ij}^{\mathrm{TL}}\Big] = \tilde{\lambda}_1^2 + \tilde{\lambda}_2^2 + \tilde{\lambda}_3^2 = \frac{2}{3}t^2\,, \tag{4.78}$$

so

$$t^2 = \frac{3}{2}\mathrm{Tr}\left[T_{ij}^{\mathrm{TL}}T_{ij}^{\mathrm{TL}}\right] . \tag{4.79}$$

This clearly shows that t measures the strength of the anisotropic traceless part of the tidal field: if $t^2 = 0$ then the tidal field is purely isotropic (all eigenvalues are equal $\lambda_1 = \lambda_2 = \lambda_3 = I_1/3$), if $t^2 \gg 0$ then the tidal field is strongly anisotropic.

Chapter 5

Forecast results

Fisher forecast is particularly useful in the study of the marked power spectrum. As we have already mentioned many times before, the idea behind the marked power spectrum is to build a two-point statistics that is able to compress information that is usually encoded in the bispectrum.
To check whether or not our statistic encode this additional information with respect to the standard power spectrum we perform a Fisher forecast analysis.
In chapter 3 we introduced the formalism, and now we apply it to our observables using as references the volume and effective redshift of the BOSS survey [6]. We compare the constraints obtained from the standard power spectrum alone with those obtained by combining $P(k)$ with the proposed marked power spectrum $M(k)$, for a set of cosmological, bias and EFT parameters, exploring how the resulting improvement depends on the scale, i.e. maximum wavenumber $k_{\max}$, and on the smoothing scale R used to construct the mark.

5.1 Analysis specifics

By assuming that the posterior distribution of the parameters can be approximated as a multivariate Gaussian distribution around the maximum of the likelihood, we compute the Fisher matrix as:

$$F_{ij} = \sum_{k,l} \frac{\partial D_k}{\partial \theta_i} C^{-1}_{kl} \frac{\partial D_l}{\partial \theta_j} \,, \tag{5.1}$$

where the vector $\mathbf{D} = \{P(k), M(k)\}$ is our data vector, $\boldsymbol{\theta}$ is the vector that contains all our cosmological and EFT parameters and C is the covariance matrix.
We now have all the ingredients to assess the constraining power of the marked power spectrum and compare it with the standard power spectrum analysis. Our marked power spectrum 4.45 contains the following parameters:

$$\{b_1, b_2, b_{\mathcal{G}_2}, c_s^2, P_{\rm shot}\} \,, \tag{5.2}$$

whose fiducial values are the ones presented in 4.5.3 and summarized in 5.1.
We also want to constrain the following cosmological parameters:

$$\{A_s, h, n_s, \omega_{\rm cdm}\} \,, \tag{5.3}$$

which are, respectively, the amplitude of scalar perturbations, the dimensionless Hubble parameter, the scalar spectral index and physical cold dark matter density. The fiducial cosmology was adopted from the default parameters of CLASS-PT:

$$\{2.089 \times 10^{-9},\, 0.6736,\, 0.9649,\, 0.12\}\,. \tag{5.4}$$

b_1	b_2	$b_{\mathcal{G}_2}$	$c_s^2\,[\mathrm{Mpc}/h]^2$	$P_{\mathrm{shot}}\,[\mathrm{Mpc}/h]^3$
1.14	−0.38	−0.041	13.21	3333

Table 5.1: Fiducial values at redshift $z = 0.6$ for the bias, counterterm and shot noise parameters.

For the covariance we followed [12] and adopt diagonal Gaussian covariances:

$$C_P(k) = \frac{2(2\pi)^2}{VV_s} P^2(k)\,, \tag{5.5}$$

$$C_M(k) = \frac{2(2\pi)^2}{VV_s} M^2(k)\,, \tag{5.6}$$

$$C_{PM}(k) = \frac{2(2\pi)^2}{VV_s} P(k)M(k)\,. \tag{5.7}$$

Here, V is the volume of the survey, which determines the total number of independent Fourier modes available for the measurement. V_s represents the volume of the considered shell in Fourier space, $V_s = 4\pi k^2 \delta k$ for a spherical k-bin of width δk. Our forecast is performed using the redshift and survey volume of the BOSS CMASS2 sample:

$$z = 0.61\,, \tag{5.8}$$

$$V_{BOSS} = 3.38(\mathrm{Gpc}/h)^3\,. \tag{5.9}$$

In this work we fix the mark parameters $\{p = 2, \delta_{s^2} = 0.25\}$ and we study how the constraining power of our statistics changes varying $R = \{10, 15, 20\}\,\mathrm{Mpc}/h$ and $k_{\max} = \{0.20, 0.25, 0.30\}\,h/\mathrm{Mpc}$.

5.2 Results

We display the results of our Fisher forecast in terms of improvements:

$$I = \frac{\sigma_P}{\sigma_{P+M}}\,, \tag{5.10}$$

in tables 5.2, 5.3, 5.4.
These tables show a substantial improvement in the marginalized constraints on all the considered cosmological, bias and EFT parameters.

Primordial scalar amplitude and linear bias

A particularly interesting result is the one concerning A_s and b_1. We can see from the tables and also from the triangle plot 5.1 the huge improvement on the marginal

	Improvement ($R = 10\text{Mpc}/h$)	Improvement ($R = 15\text{Mpc}/h$)	Improvement ($R = 20\text{Mpc}/h$)
A_s	13.067	11.384	9.850
h	1.507	1.453	1.408
n_s	1.793	1.745	1.676
ω_{cdm}	1.583	1.542	1.505
b_1	26.272	15.159	10.194
b_2	2.740	2.205	1.696
$b_{\mathcal{G}_2}$	2.664	2.393	2.027
P_{shot}	2.276	1.986	1.657
c_s^2	1.203	1.227	1.214

Table 5.2: Improvement factors on the marginalized parameter uncertainties obtained by including the tidal-field marked power spectrum, for smoothing scales $R = 10$, 15, and $20\,h^{-1}$Mpc at $k_{\max} = 0.3\,h/\text{Mpc}$.

	Improvement ($R = 10\text{Mpc}/h$)	Improvement ($R = 15\text{Mpc}/h$)	Improvement ($R = 20\text{Mpc}/h$)
A_s	20.447	17.774	14.911
h	1.666	1.628	1.576
n_s	1.854	1.823	1.784
ω_{cdm}	1.717	1.691	1.648
b_1	40.181	23.025	14.759
b_2	3.119	2.567	1.900
$b_{\mathcal{G}_2}$	2.146	2.099	2.017
P_{shot}	2.562	2.245	1.777
c_s^2	1.244	1.285	1.302

Table 5.3: Same as table 5.2 but for $k_{\max} = 0.25\,h/\text{Mpc}$.

errors for this parameters. This improvement represent a degeneracy breaking that is expected on physical grounds. At tree level, the galaxy power spectrum depends on A_s and b_1 only through the combination:

$$P_g^{tree}(k) \propto b_1^2 P_{11}(k) \propto b_1^2 A_s\,, \tag{5.11}$$

so that rescaling properly both A_s and b_1 does not affect $P_g^{tree}(k)$: this means that the two parameters are exactly degenerate.
This degeneracy can be verified also analytically by noting that $\partial P/\partial A_s$ and $\partial P/\partial b_1$ are parallel vectors in k-space at tree level.
In particular, we have:

$$A_s\,\partial P/\partial A_s = \tfrac{1}{2}\,b_1\,\partial P/\partial b_1 \quad \forall\, k\,. \tag{5.12}$$

One-loop corrections break this exact degeneracy only mildly, since their dependence on A_s and b_1 does not follow the same simple rescaling; as a result, the correlation coefficient ρ between A_s and b_1 extracted from $P(k)$ alone remains close to unity.
Because the marginalized error on a parameter scales as $(1-\rho^2)^{-1/2}$ in the presence of a

	Improvement ($R = 10\mathrm{Mpc}/h$)	Improvement ($R = 15\mathrm{Mpc}/h$)	Improvement ($R = 20\mathrm{Mpc}/h$)
A_s	39.394	33.206	26.324
h	1.603	1.546	1.424
n_s	1.544	1.508	1.488
ω_{cdm}	1.631	1.582	1.481
b_1	76.938	47.042	28.919
b_2	3.196	2.759	1.981
$b_{\mathcal{G}_2}$	2.176	1.179	1.158
P_{shot}	2.071	1.939	1.546
c_s^2	1.876	1.927	1.925

Table 5.4: Same as table 5.2 but for $k_{\max} = 0.20\,h/\mathrm{Mpc}$.

strong pairwise correlation, even a modest reduction in ρ can produce a large improvement in the marginalized constraints. The marked power spectrum provides exactly this kind of reduction: since it encodes parameter dependencies that are not simply proportional to the combination $b_1^2 A_s$, it lowers the A_s-b_1 correlation and substantially tightens the constraints on both parameters.

This behaviour can be visualized by comparing, for each k-bin, the rescaled derivatives $A_s\,\partial X/\partial A_s$ and $b_1\,\partial X/\partial b_1$, for $X = \{P(k), M(k)\}$, as shown in 5.2. For the power spectrum, all points lie on a single straight line through the origin with slope 1/2, providing a direct geometric confirmation of the parallel-vector condition discussed above. For the marked power spectrum, the points clearly depart from this linear relation.

This degeneracy breaking is also confirmed by the triangle plot 5.1 where we can see that considering only $P(k)$ or only $M(k)$ the tilt of the ellipses is very close to 45°, while once we consider the combination $P(k) + M(k)$ the tilt is significantly closer to 0°.

The trend of the improvements on these two parameters with $k_{\max}$ is consistent with this picture: lowering $k_{\max}$ suppresses the loop contributions that partially break the degeneracy, causing $\sigma_P(A_s)$ and $\sigma_P(b_1)$ to grow much faster than their $P + M$ counterparts, which explains the increase of the improvement of these two parameters at lower $k_{\max}$.

While this qualitative behaviour is expected, the size of the effect is sufficiently large to merit a dedicated check of whether it genuinely reflect new, independent information carried by the power spectrum or is instead mainly driven by the tighter constraint on b_1 provided by $M(k)$, which helps break the A_s-b_1 degeneracy present in $P(k)$.

To disentangle these two possibilities, we repeat the forecast imposing an external Gaussian prior on b_1 with a 10% uncertainty, chosen as a conservative benchmark for an external constraint on the linear galaxy bias. If the resulting improvement on $\sigma(A_s)$ is comparable to the one obtained by adding $M(k)$, the bulk of the gain is mediated by its ability to constrain b_1, rather than genuinely new information on A_s. The results are displayed in 5.5.

Even after imposing a 10% external prior on b_1, the addition of $M(k)$ yields a substantial further improvement in the constraint on A_s. These results indicate that the constraining power of the marked spectrum is not solely mediated by its ability to constrain b_1. The improvement decreases as $k_{\max}$ is reduced, as expected from the loss

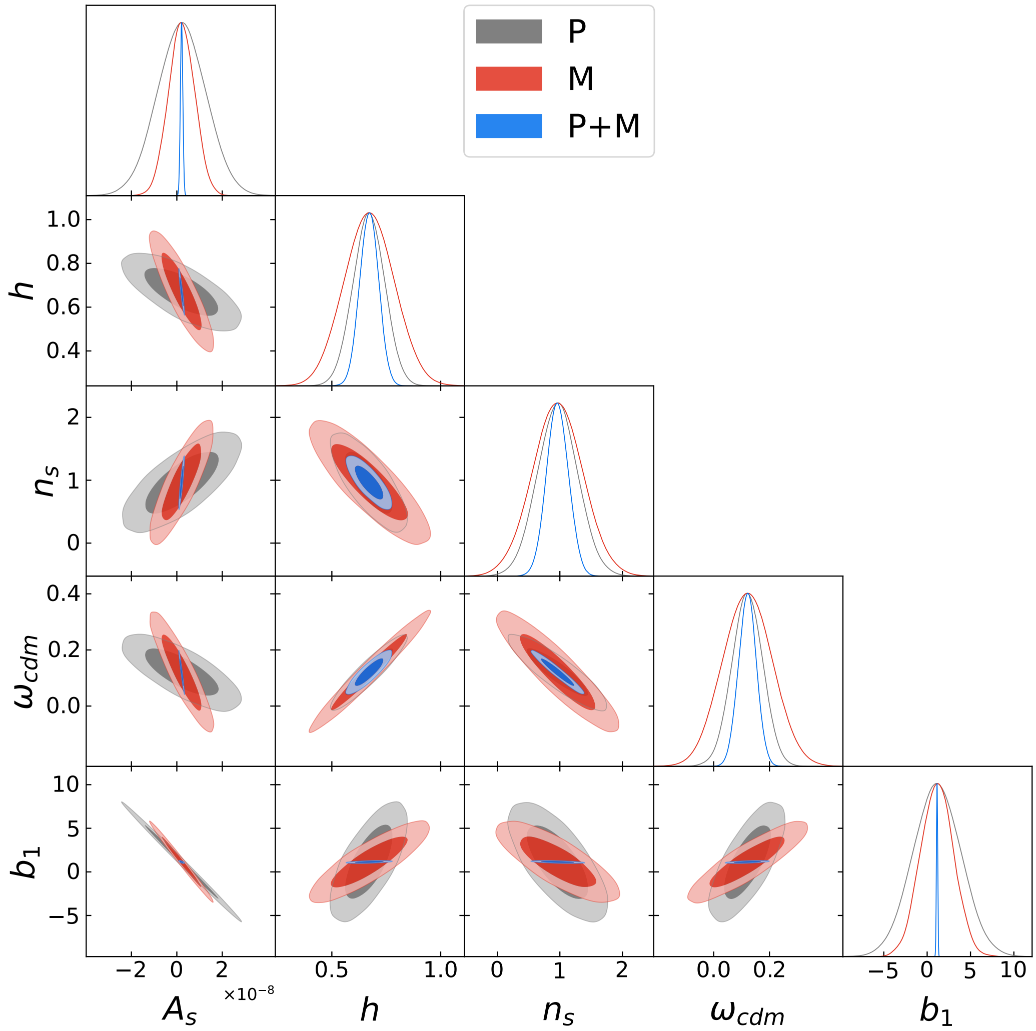

Figure 5.1: Triangle plot for the cosmological parameters, $k_{\max} = 0.25\,h\text{Mpc}^{-1}$ and $R = 10\,\text{Mpc}/h$. From this image we clearly see the strong degeneracy breaking between A_s and b_1.

of information at smaller scales.
It is worth noticing that these results are obtained in real space, where, as we already mentioned, A_s and b_1 are degenerate at tree level. In redshift space this degeneracy is expected to be already partially broken by $P(k)$ alone: decomposing the power spectrum into multipoles, the monopole and the quadrupole probe different combinations of the two parameters. The other cosmological parameters are expected to remain more robust when moving to redshift space. For this reason, we regard the results concerning ω_{cdm}, h and n_s as the most reliable results of this forecast. The graphical results for these parameters are displayed in figure 5.3.

Higher-order bias parameters

The improvement obtained on the quadratic and tidal bias parameters b_2 and $b_{\mathcal{G}_2}$ also deserves a dedicated comment. At tree level, the bispectrum depends explicitly on these two parameters and several studies have shown that combining the bispectrum with the power spectrum substantially tightens the constraints on these parameters

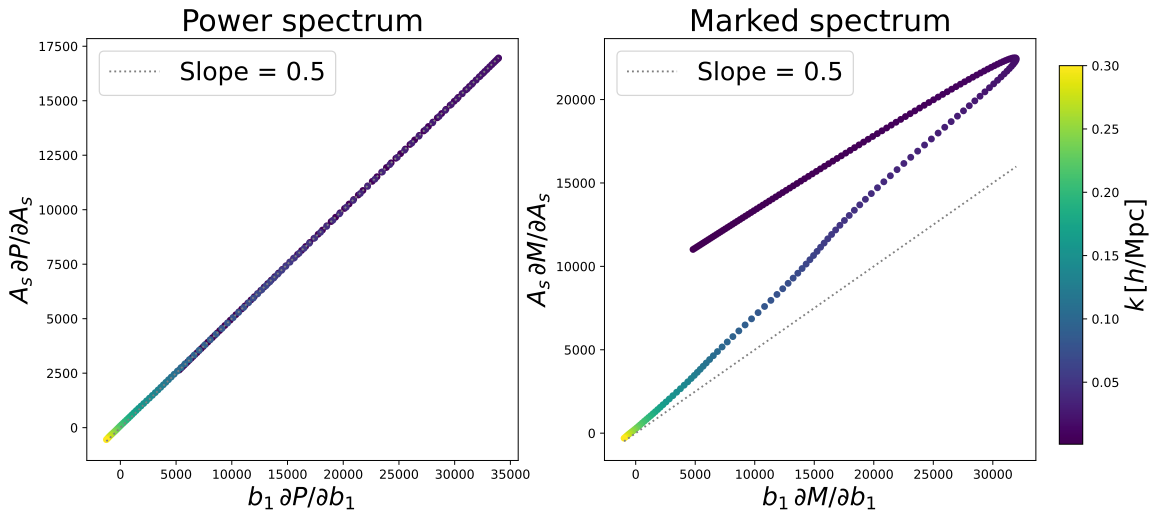

Figure 5.2: Scatter plot of $A_s\,\partial X/\partial A_s$ against $b_1\,\partial X/\partial b_1$, with one point per k-bin up to $k_{\max} = 0.30\,h/\mathrm{Mpc}$ (colour-coded by k), for the power spectrum $X = P$ (left) and the marked power spectrum $X = M$ (right). The points for $P(k)$ lie close to a single straight line through the origin with slope $1/2$: as expected the two rescaled derivatives satisfy $A_s\,\partial P/\partial A_s = \frac{1}{2}\,b_1\,\partial P/\partial b_1$ at every k, reflecting the near-exact A_s-b_1 degeneracy. The points for $M(k)$, by contrast, clearly depart from this linear relation.

$k_{\max}$	$R = 10\mathrm{Mpc}/h$	$R = 15\mathrm{Mpc}/h$	$R = 20\mathrm{Mpc}/h$
0.30	2.086	1.886	1.743
0.25	2.078	1.883	1.726
0.20	1.682	1.510	1.386

Table 5.5: Improvement factors on $\sigma(A_s)$ obtained by adding the marked power spectrum, after imposing a Gaussian prior with 10% uncertainty on b_1. Results are shown for the different values of $k_{\max}$ and smoothing scale R.

compared to $P(k)$ alone [91], [92], [93]. Since the marked power spectrum is built to capture part of the information encoded in the bispectrum, we expect a significant improvement on b_2 and $b_{\mathcal{G}_2}$. This expectation is indeed confirmed by our results: as we can see in 5.2, 5.3, 5.4, we find an improvement of a factor of a few on both the parameters across all values of $k_{\max}$ and smoothing scale R considered. The graphical counterpart of these results is in figure 5.1.

EFT parameter

Another interesting result concerns the EFT parameter c_s^2. As we can see from the tables 5.3 and 5.2, the improvements on this parameter increase with increasing $k_{\max}$. This can be understood intuitively: this parameter encodes small scales effects, therefore as we push the analysis to smaller scales, i.e. bigger k, we expect to constrain it better.
From the same tables, c_s^2 also appears to be unaffected by varying the smoothing scale R. This is because this parameter is already well constrained by $P(k)$ alone: indeed we can see that the improvement is small, meaning that the effect of $M(k)$, and consequently of R, is minimal.

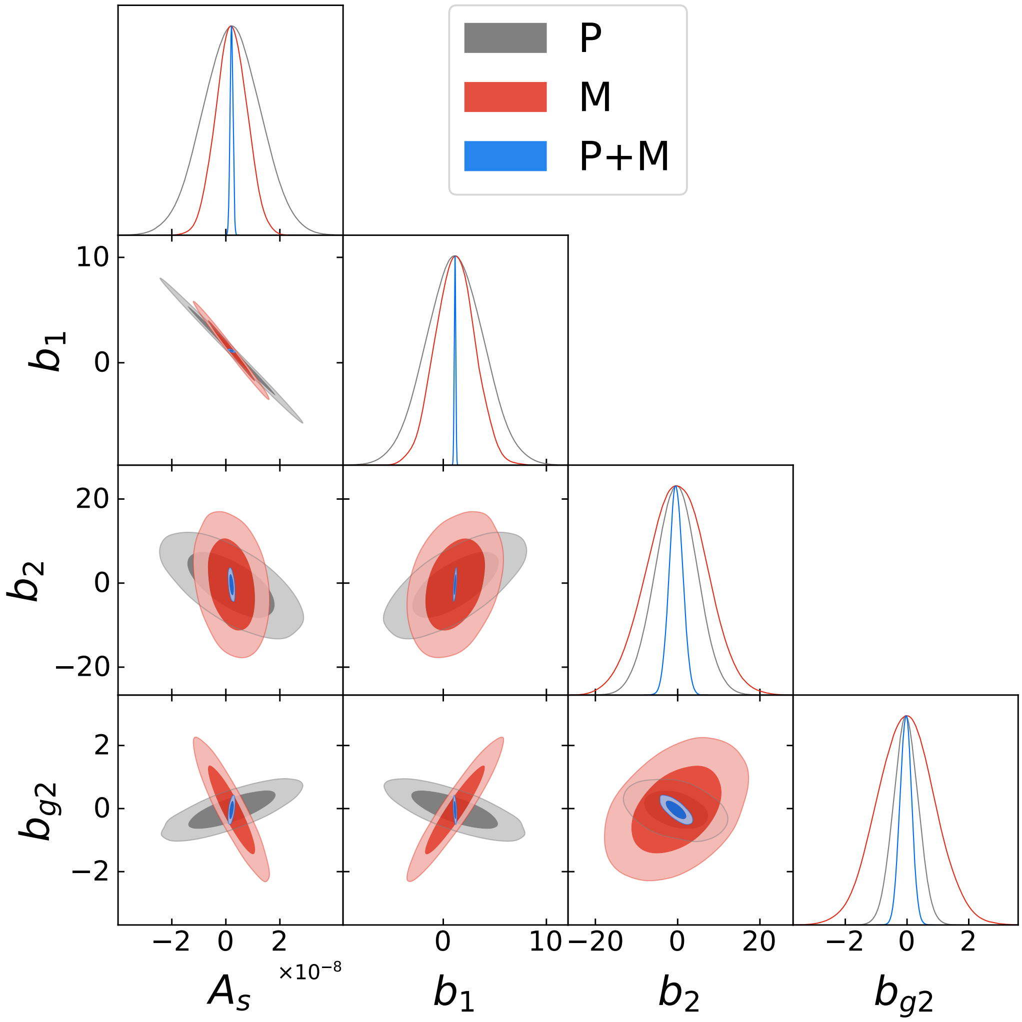

Figure 5.3: Triangle plot for the bias and cosmological parameters. Forecast specifics: $k_{\max} = 0.25\, h\mathrm{Mpc}^{-1}$ and $R = 10\,\mathrm{Mpc}/h$.

Overall results

In figure 5.4 we show the triangle plot for $P + M$ at the three different $k_{\max} = \{0.20,\, 0.25,\, 0.30\}h/\mathrm{Mpc}$ considered in our analysis. The plot shows the constraints for the parameters $\{A_s,\, h,\, n_s,\, \omega_{cdm},\, b_1\}$, with fixed mark parameters $R = 10\mathrm{Mpc}/h$, $p = 2$ and $\delta_{s^2} = 0.25$. From the image we see that contours shrink as $k_{\max}$ increases, indicating that the total information content grows with the number of modes included, as expected. However, from the tables 5.2,5.3,5.4 we see that the improvements decrease with increasing $k_{\max}$: this is because σ_P also shrinks, and does so more rapidly in proportion. This is the same behavior already discussed for A_s and b_1, which is now confirmed across the full parameter set. In appendix C we show the same triangle plot for the complete set of parameters: cosmological, bias and EFT.
It is worth commenting also on the R-dependence of our results: for almost all parameters we notice an increase in the improvements as R decreases. This is expected because a smaller smoothing radius retains more small-scale information. However, as shown in [76] a smaller smoothing radius R also makes the theoretical description for the marked field less perturbative. This trade-off motivated our choice of setting

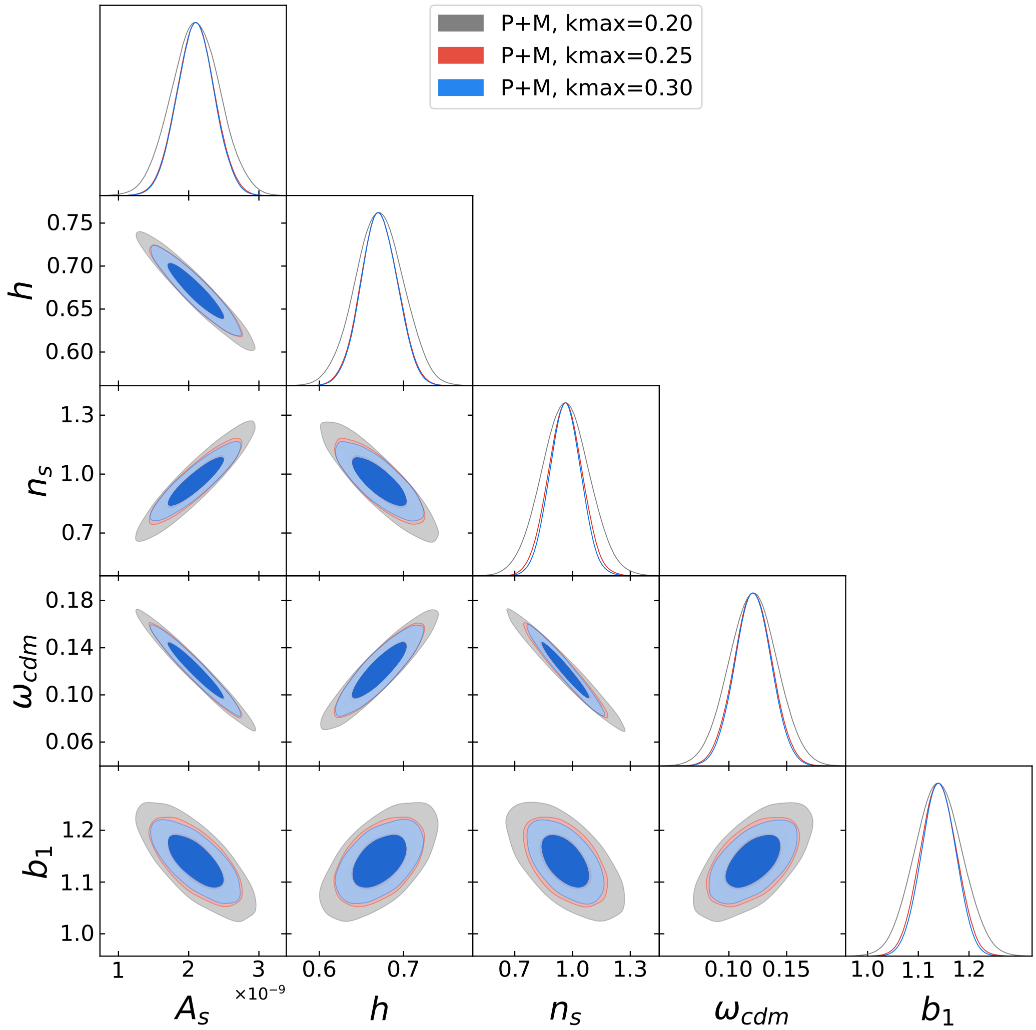

Figure 5.4: Triangle plot of the constraint from $P + M$ on the cosmological parameters considered in the analysis at different k_{max}. Mark parameters: $R = 10\,\text{Mpc}/h$, $p = 2$ and $\delta_{s^2} = 0.25$.

$R = 10\,\text{Mpc}/h$ as the lower limit for the smoothing radius in our analysis. Indeed, the paper shows that for $R = 10\,\text{Mpc}/h$ the conditions required for the perturbative expansion to remain under control are already rather restrictive.

We also want to analyze the R-dependence of the constraining power of $P + M$ alone, rather than the improvement relative to P. Fig. 5.5 shows the 95% confidence level marginalized uncertainty on each parameter as a function of $k_{\text{max}} = \{0.1,\, 0.15\,, 0.2,\, 0.25,\, 0.3\}h/\text{Mpc}$, for the three smoothing radii $R = \{10, 15, 20\}\,\text{Mpc}/h$. As expected, the uncertainties decrease monotonically with k_{max} for all values of R, reflecting the growing number of Fourier modes included in the analysis. At fixed k_{max}, however, we observe that a smaller smoothing radius consistently yields tighter constraints, confirming that decreasing R includes more small-scale information in the marked field and therefore increases its constraining power.

This trend in the constraining power of $P+M$ is fully consistent with the R-dependence of the improvements discussed above. Since σ_P does not depend on R, the improvement σ_P/σ_{P+M} increases as σ_{P+M} decreases with R.

In Appendix C, for completeness we show the triangle plots for all parameters considered in the Fisher forecast, cosmological, bias and EFT.

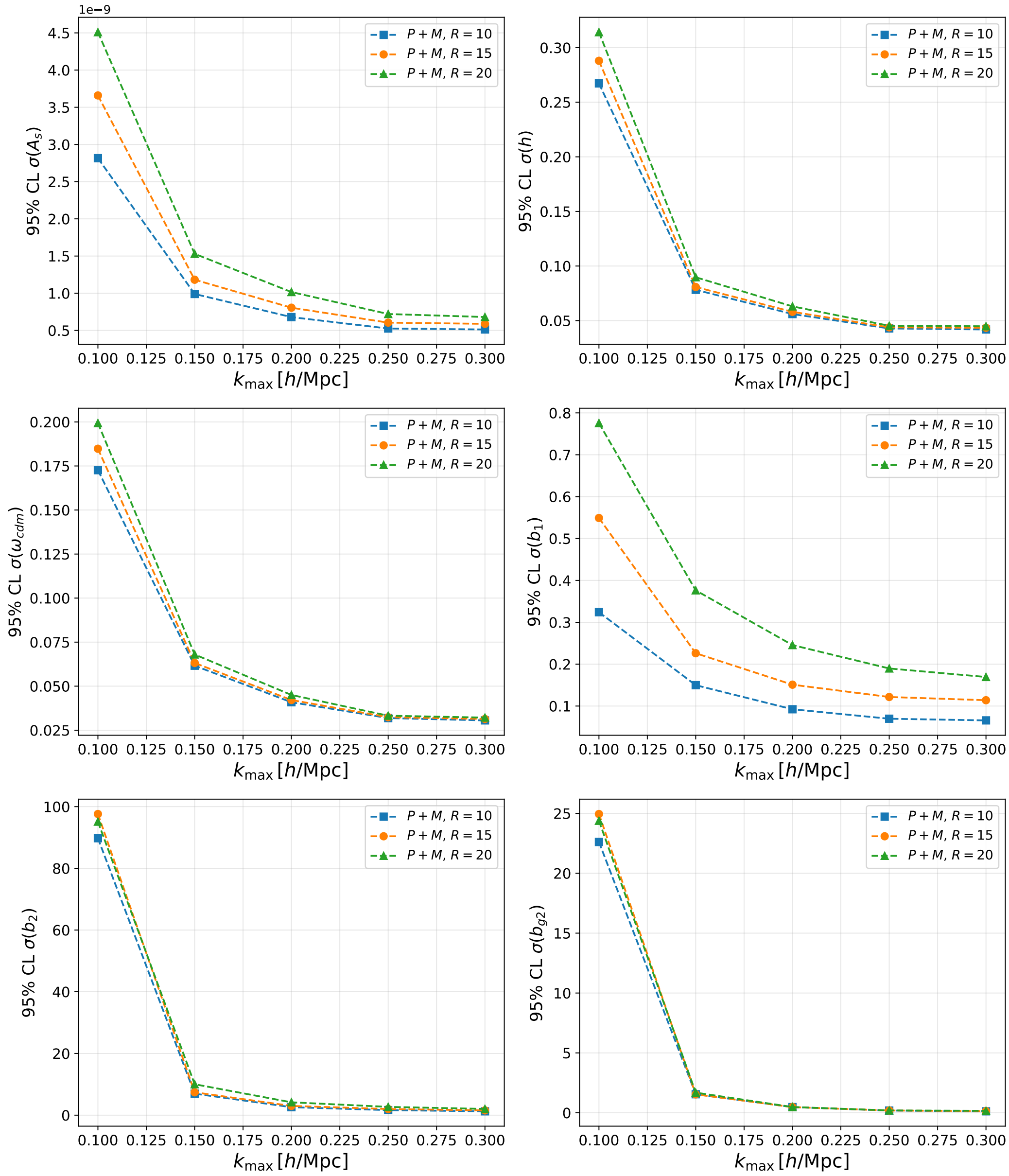

Figure 5.5: 95% CL marginalized uncertainty on A_s, h, ω_{cdm}, b_1, b_2, and b_{g2} as a function of $k_{\max}$, for $P + M$ with smoothing radii $R = 10, 15, 20\,\mathrm{Mpc}/h$.

As shown in the tables 5.4, 5.3 and 5.2, all the considered parameters, cosmological, bias and EFT, present substantial improvements when we consider the combined observable $P(k) + M(k)$. These results confirm that the tidal marked power spectrum carries genuinely new cosmological information with respect to the standard power spectrum alone. The magnitude of this gain depends on the specific choice of $k_{\max}$ and smoothing radius R, as discussed above: larger $k_{\max}$ and smaller R both enhance the amount of small-scale, non-Gaussian information accessible through the marking procedure, within the limits imposed by the perturbative validity of the model.

Conclusions

The advent of a new generation of large-scale galaxy surveys is fundamentally reshaping the landscape of observational cosmology. Missions such as Euclid [94], alongside ground-based programs like DESI [4] and the Vera C. Rubin Observatory's LSST [5], are delivering galaxy catalogues of unprecedented volume and precision. This leap in observational precision calls for new theoretical models and new statistical tools able to extract this cosmological information from the galaxy density field, while remaining robust against the systematic and theoretical uncertainties that accompany such high-precision measurements.
Traditionally, cosmological inference from Large-Scale Structure has relied primarily on the galaxy power spectrum and bispectrum. The power spectrum captures the full information content of the density field under the assumption of Gaussianity, while the bispectrum, the Fourier transform of the three-point correlation function, accesses additional information generated by nonlinear gravitational evolution, bias, and redshift-space distortions. However, both statistics face significant limitations in the mildly nonlinear regime, where perturbative approaches start to lose accuracy and a hierarchical extension to higher-order correlators becomes computationally impractical and statistically inefficient.
In this context, the marked power spectrum [8, 9, 10, 11, 12, 10, 74] has recently gained interest as a promising alternative. By reweighting the density field with a suitably chosen mark before computing its two-point correlator, the marked power spectrum is able to effectively repackage higher-order information into a two-point function, offering enhanced sensitivity to features such as low-density environments or the imprint of massive neutrinos and modified gravity, without incurring the full computational and statistical cost of estimating higher-order correlators directly. This marked statistics therefore provides a computationally efficient way to extracting information beyond the two-point correlator, complementing standard power spectrum and bispectrum analyses in the effort to fully exploit the statistical power of next-generation cosmological datasets.
This thesis contributes to this direction by extending the analytical modeling of marked statistics exploring a new class of marks. In particular, we developed a theoretically consistent model for the marked power spectrum at one-loop order in real space for a mark that depends on the contraction of the matter tidal field: $T_{ij}(x)T^{ij}(x) = \partial_i\partial_j\phi(x)\partial^i\partial^j\phi(x)$, where the gravitational potential $\phi(x)$ is related to the density contrast $\delta(x)$ via the Poisson equation.
We choose the matter tidal field because the local density, on which most of the existing studies relies, provides only a partial description of the environment in which galaxies form and evolve. The structure of the cosmic web is determined by gravitational collapse, which is strongly anisotropic and its geometry is encoded in the tidal tensor, whose eigenvalues determine the local morphology, distinguishing voids, sheets,

filaments and knots. Unlike the density field, which measures the amplitude of local fluctuations, the tidal tensor contains information about the geometry of the surrounding matter distribution and since halo formation and galaxy evolution are expected to depend on both density and tidal environment [14, 15, 16], tidal field observables are a physically motivated extension of density-based marked formalism.
Building on the theoretical framework laid out in [9, 12], we derived the one-loop expression for the marked power spectrum using the mark built from the matter tidal tensor, extending the perturbative treatment previously developed for density-based marks. This required computing the relevant standard perturbation theory kernels associated with the new mark up to third order in the density contrast, and deriving the corresponding loop integrals entering the marked power spectrum. In order to obtain a theoretically consistent and realistic description, the calculations are carried out within the Effective Field Theory of the Large-Scale Structure (EFTofLSS) [17, 18, 19]. We also included the galaxy bias expansion, so that the resulting model applies not only to the dark matter field alone, but to a general biased tracer. Careful attention was given to ensuring that the resulting expressions reduce consistently to the standard matter power spectrum in the appropriate limits, providing an internal validation of the calculation.
Once the analytical model was established, we developed a numerical code to evaluate the various components of the marked power spectrum. The building blocks of our code, as for example the cosmology or the linear matter power spectrum, are provided by CLASS-PT.
Finally, using this numerical implementation we perform a Fisher forecast analysis to assess the constraining power of our new tidal field statistics on a set of cosmological, bias and EFT parameters. In particular, we compare the constraining power of the combined marked and standard power spectrum with respect to the standard power spectrum alone. To do so we define the improvement of a given parameter as the ratio between the marginalized error we find for that parameter from the $P(k)$ Fisher forecast and the one we obtain from the $P(k) + M(k)$ forecast: σ_P/σ_{P+M}. We decided to present these improvement for different values of $k_{\max}$ and smoothing scale R, while we fixed the mark parameters to $p = 2$, to upweight regions with lower amplitude of the tidal field, and $\delta_{s^2} = 0.25$.
The results of this analysis show that the tidal marked power spectrum carries substantial additional cosmological information with respect to the standard power spectrum alone. The bigger improvement concerns the primordial scalar amplitude A_s and the linear bias b_1: at tree level these two parameters enter the power spectrum only through the combination $A_s b_1^2$, making them nearly perfectly degenerate. The tidal marked power spectrum does not depend on these parameters with the same combination, and its inclusion breaks this degeneracy, improving the marginalized error on A_s by up to a factor of ~ 40, depending on the maximum wavenumber $k_{\max}$ and the smoothing scale R. To verify that this large improvement was not simply a consequence of a tighter constrain on b_1, we repeated the forecast imposing a 10% Gaussian prior on b_1: the improvement on A_s remained substantial, confirming that the tidal field provides genuinely new cosmological information. The remaining cosmological parameters, ω_{cdm}, h and n_s, also show consistent, though more modest, improvements across all considered configurations.
The quadratic and tidal bias parameters, b_2 and $b_{\mathcal{G}_2}$, also benefit strongly, with improvements of up to a factor ~ 3.2 and ~ 2.7 respectively. This is consistent with

the fact that these higher-order bias parameters are typically well constrained by the bispectrum, and this is precisely the kind of information the marked power spectrum is designed to recover.
Looking ahead, several developments will be necessary to bring our tidal field marked statistics closer to application on real survey data.
Future works should extend the present formalism to redshift space, incorporating redshift space distortions, which are anisotropic effects due gravitational pull that generates peculiar velocities in galaxies' motion. This effects are unavoidable in any realistic analysis of spectroscopic galaxy surveys [46, 47, 48].
Future developments should also account for primordial non Gaussianity (PNG), especially the one of the non local type [12], which we expect to manifest when we measure the bispectrum. Primordial non Gaussianity is crucial because is one of the few direct probes of the physics of the early Universe and allows us, for example, to distinguish between different inflationary models.
These extensions are essential prerequisites for applying the marked statistics we developed to upcoming datasets from new generation galaxy surveys such as Euclid.
A further natural direction for future work concerns the choice of the mark itself. In [76], the authors develop a framework to optimize the choice of the mark so as to maximize the constraining power on cosmological parameter. They show that an optimized density mark can substantially reduce parameter uncertainties compared to ad hoc choices. It would therefore be valuable to extend this optimization procedure to the matter tidal field mark studied in this thesis. Such an optimization could further enhance the constraining power of our statistics, and represents a promising step toward establishing marked statistics as a standard tool for the analysis of next-generation Large-Scale Structure surveys.
In summary, through analytical modelling, numerical implementation, and Fisher forecast analysis, this thesis has shown that the tidal marked power spectrum is a genuinely informative and computationally inexpensive complement to standard Large-Scale Structure statistics. Its strength is the fact that it is a two-point statistic but behaves, in many respects, like a much more expensive one. It inherits part of the sensitivity to non-linear structure and galaxy bias that is normally the domain of the bispectrum, while remaining as cheap to compute as the ordinary power spectrum.
As the next generation of galaxy surveys begins to deliver datasets of a size and precision without precedent, the marked power spectrum, and its extension to the tidal field, is emerging as computationally efficient way to extract cosmological information. This combination of sensitivity and efficiency provided by this statistic is precisely what large surveys such as Euclid will require.

Appendix A

Power Spectrum and bispectrum

In this section we present the computation of the one-loop power spectrum and the tree-level bispectrum. In the following calculations, we consider equal-time correlators, and therefore, the time dependence has been omitted.

A.1 Power spectrum

We consider the expressions for the two point correlator at one-loop order 2.52.
Here we compute the two non trivial terms $P_{22}(k)$ and $P_{13}(k)$.

$$
\begin{aligned}
P_{22}(k) =& \langle \delta^{(2)}(\mathbf{k}_1)\delta^{(2)}(\mathbf{k}_2)\rangle \\
=& \Big\langle \int \frac{\mathrm{d}^3\mathbf{q}_1}{(2\pi)^3} \int \frac{\mathrm{d}^3\mathbf{q}_2}{(2\pi)^3} (2\pi)^3 \delta_D(\mathbf{k}_1 - \mathbf{q}_1 - \mathbf{q}_2) F_2(\mathbf{q}_1, \mathbf{q}_2) \delta^{(1)}(\mathbf{q}_1)\delta^{(1)}(\mathbf{q}_2) \times \\
& \times \int \frac{\mathrm{d}^3\mathbf{p}_1}{(2\pi)^3} \int \frac{\mathrm{d}^3\mathbf{p}_2}{(2\pi)^3} (2\pi)^3 \delta_D(\mathbf{k}_2 - \mathbf{p}_1 - \mathbf{p}_2) F_2(\mathbf{p}_1, \mathbf{p}_2) \delta^{(1)}(\mathbf{p}_1)\delta^{(1)}(\mathbf{p}_2) \Big\rangle \\
=& \int \frac{\mathrm{d}^3\mathbf{q}_1}{(2\pi)^3} \int \frac{\mathrm{d}^3\mathbf{p}_1}{(2\pi)^3} F_2(\mathbf{q}_1, \mathbf{k}_1 - \mathbf{q}_1) F_2(\mathbf{p}_1, \mathbf{k}_2 - \mathbf{p}_1) \times \\
& \times \langle \delta^{(1)}(\mathbf{q}_1)\delta^{(1)}(\mathbf{k}_1 - \mathbf{q}_1)\delta^{(1)}(\mathbf{p}_1)\delta^{(1)}(\mathbf{k}_2 - \mathbf{p}_1)\rangle \ ,
\end{aligned}
$$

and now we apply the Wick theorem 2.2.2

$$
\begin{aligned}
\langle \delta^{(1)}(\mathbf{q}_1)&\delta^{(1)}(\mathbf{k}_1 - \mathbf{q}_1)\delta^{(1)}(\mathbf{p}_1)\delta^{(1)}(\mathbf{k}_2 - \mathbf{p}_1)\rangle = \\
&= \langle \delta^{(1)}(\mathbf{q}_1)\delta^{(1)}(\mathbf{k}_1 - \mathbf{q}_1)\rangle \langle \delta^{(1)}(\mathbf{p}_1)\delta^{(1)}(\mathbf{k}_2 - \mathbf{p}_1)\rangle + \\
&\quad + \langle \delta^{(1)}(\mathbf{q}_1)\delta^{(1)}(\mathbf{p}_1)\rangle \langle \delta^{(1)}(\mathbf{k}_1 - \mathbf{q}_1)\delta^{(1)}(\mathbf{k}_2 - \mathbf{p}_1)\rangle \\
&\quad + \langle \delta^{(1)}(\mathbf{q}_1)\delta^{(1)}(\mathbf{k}_2 - \mathbf{p}_1)\rangle \langle \delta^{(1)}(\mathbf{p}_1)\delta^{(1)}(\mathbf{k}_1 - \mathbf{q}_1)\rangle \\
&= (2\pi)^6 \big[\delta_D(\mathbf{k}_1)\delta_D(\mathbf{k}_2) P_L(|\mathbf{k}_1 - \mathbf{q}_1|) P_L(|\mathbf{k}_2 - \mathbf{p}_1|) + \\
&\quad + \delta_D(\mathbf{q}_1 + \mathbf{p}_1)\delta_D(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{q}_1 - \mathbf{p}_1) P_L(q_1) P_L(|\mathbf{k}_1 - \mathbf{q}_1|) + \\
&\quad + \delta_D(\mathbf{q}_1 + \mathbf{k}_2 - \mathbf{p}_1)\delta_D(\mathbf{k}_1 + \mathbf{p}_1 - \mathbf{q}_1) P_L(q_1) P_L(|\mathbf{k}_1 - \mathbf{q}_1|) \big] \ ,
\end{aligned}
$$

where the first term only contributes at $\mathbf{k}_1 = \mathbf{k}_2 = 0$ so we eliminate it.
We can rewrite everything as

$$\begin{aligned}
P_{22}(k) = &\int \frac{\mathrm{d}^3\mathbf{q}_1}{(2\pi)^3} F_2(\mathbf{q}_1, \mathbf{k}_1 - \mathbf{q}_1) F_2(\mathbf{q}_1 - \mathbf{k}_1, -\mathbf{q}_1) \times \\
&\times (2\pi)^3 \big[\delta_D(\mathbf{k}_1 + \mathbf{k}_2) P_L(q_1) P_L(|\mathbf{k}_1 - \mathbf{q}_1|) + \\
&+ \delta_D(\mathbf{k}_1 + \mathbf{k}_2) P_L(q_1) P_L(|\mathbf{k}_1 - \mathbf{q}_1|)\big] \\
=&(2\pi)^3 \delta_D(\mathbf{k}_1 + \mathbf{k}_2) \int \frac{\mathrm{d}^3\mathbf{q}_1}{(2\pi)^3} [F_2(\mathbf{q}, \mathbf{k}_1 - \mathbf{q})]^2 P_L(q_1) P_L(|\mathbf{k}_1 - \mathbf{q}|) \,,
\end{aligned} \tag{A.1}$$

where we have used the property of the kernel $F_2(\mathbf{k}, \mathbf{q}) = F_2(-\mathbf{q}, -\mathbf{k})$.
The calculation for $P_{13}(k)$ is completely analogous

$$\begin{aligned}
P_{13}(k) =&\langle \delta^{(3)}(\mathbf{k}_1) \delta^{(1)}(\mathbf{k}_2) \rangle = \\
=&2 \int \frac{\mathrm{d}^3\mathbf{q}_1}{(2\pi)^3} \int \frac{\mathrm{d}^3\mathbf{q}_2}{(2\pi)^3} \int \frac{\mathrm{d}^3\mathbf{q}_3}{(2\pi)^3} (2\pi)^3 \delta_D(\mathbf{k}_1 - \mathbf{q}_1 - \mathbf{q}_2 - \mathbf{q}_3) \\
&\times F_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) \langle \delta(\mathbf{q}_1) \delta(\mathbf{q}_2) \delta(\mathbf{q}_3) \delta(\mathbf{k}_2) \rangle \\
=&2 \int \frac{\mathrm{d}^3\mathbf{q}_1}{(2\pi)^3} \int \frac{\mathrm{d}^3\mathbf{q}_2}{(2\pi)^3} F_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{k}_1 - \mathbf{q}_1 - \mathbf{q}_2) \\
&\times \langle \delta(\mathbf{q}_1) \delta(\mathbf{q}_2) \delta(\mathbf{k}_1 - \mathbf{q}_1 - \mathbf{q}_2) \delta(\mathbf{k}_2) \rangle \\
=&2 \int \frac{\mathrm{d}^3\mathbf{q}_1}{(2\pi)^3} \int \frac{\mathrm{d}^3\mathbf{q}_2}{(2\pi)^3} F_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{k}_1 - \mathbf{q}_1 - \mathbf{q}_2) \\
&\times (2\pi)^6 \Big[\delta_D(\mathbf{q}_1 + \mathbf{q}_2) P_L(q_1) \delta_D(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{q}_1 - \mathbf{q}_2) P_L(|\mathbf{k}_1 - \mathbf{q}_1 - \mathbf{q}_2|) \\
&+ \delta_D(\mathbf{k}_1 - \mathbf{q}_2) P_L(q_1) \delta_D(\mathbf{k}_2 + \mathbf{q}_2) P_L(k_2) \\
&+ \delta_D(\mathbf{q}_1 + \mathbf{k}_2) P_L(q_1) \delta_D(\mathbf{k}_1 - \mathbf{q}_1) P_L(|\mathbf{k}_1 - \mathbf{q}_1 - \mathbf{q}_2|)\Big] \\
=&2\,(2\pi)^3 \delta_D(\mathbf{k}_1 + \mathbf{k}_2) \int \frac{\mathrm{d}^3\mathbf{q}_1}{(2\pi)^3} F_3(\mathbf{q}_1, -\mathbf{q}_1, \mathbf{k}_1) \Big[P_L(q_1) P_L(k_1) + \\
&+ P_L(q_1) P_L(k_1) + P_L(q_1) P_L(k_1)\Big] \\
=&6\,(2\pi)^3 \delta_D(\mathbf{k}_1 + \mathbf{k}_2) \int \frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3} F_3(\mathbf{q}, -\mathbf{q}, \mathbf{k}_1) P_L(q) P_L(k_1) \,.
\end{aligned} \tag{A.2}$$

A.2 Bispectrum

Now we show how to recover 2.57.
At tree level, the bispectrum is obtained by:

$$\begin{aligned}
\langle \delta(\mathbf{k}_1) \delta(\mathbf{k}_2) \delta(\mathbf{k}_3) \rangle = &\left\langle \delta^{(2)}(\mathbf{k}_1) \delta^{(1)}(\mathbf{k}_2) \delta^{(1)}(\mathbf{k}_3) \right\rangle \\
&+ \left\langle \delta^{(1)}(\mathbf{k}_1) \delta^{(2)}(\mathbf{k}_2) \delta^{(1)}(\mathbf{k}_3) \right\rangle \\
&+ \left\langle \delta^{(1)}(\mathbf{k}_1) \delta^{(1)}(\mathbf{k}_2) \delta^{(2)}(\mathbf{k}_3) \right\rangle \,.
\end{aligned} \tag{A.3}$$

Using the second-order solution,

$$\delta^{(2)}(\mathbf{k}) = \int \frac{d^3q}{(2\pi)^3} \frac{d^3p}{(2\pi)^3} (2\pi)^3 \delta_D(\mathbf{k} - \mathbf{q} - \mathbf{p}) F_2(\mathbf{q}, \mathbf{p}) \delta^{(1)}(\mathbf{q}) \delta^{(1)}(\mathbf{p}) \,, \tag{A.4}$$

the first contribution becomes:

$$\begin{aligned}
&\left\langle \delta^{(2)}(\mathbf{k}_1)\delta^{(1)}(\mathbf{k}_2)\delta^{(1)}(\mathbf{k}_3)\right\rangle \\
&\quad = \int \frac{d^3q}{(2\pi)^3}\frac{d^3p}{(2\pi)^3}(2\pi)^3\delta_D(\mathbf{k}_1-\mathbf{q}-\mathbf{p})F_2(\mathbf{q},\mathbf{p})\times \qquad \text{(A.5)}\\
&\qquad\qquad \times \left\langle \delta^{(1)}(\mathbf{q})\delta^{(1)}(\mathbf{p})\delta^{(1)}(\mathbf{k}_2)\delta^{(1)}(\mathbf{k}_3)\right\rangle \,. \qquad \text{(A.6)}
\end{aligned}$$

Applying Wick's theorem,

$$\begin{aligned}
&\left\langle \delta^{(1)}(\mathbf{q})\delta^{(1)}(\mathbf{p})\delta^{(1)}(\mathbf{k}_2)\delta^{(1)}(\mathbf{k}_3)\right\rangle \\
&\quad = \left\langle \delta^{(1)}(\mathbf{q})\delta^{(1)}(\mathbf{p})\right\rangle \left\langle \delta^{(1)}(\mathbf{k}_2)\delta^{(1)}(\mathbf{k}_3)\right\rangle \\
&\qquad + \left\langle \delta^{(1)}(\mathbf{q})\delta^{(1)}(\mathbf{k}_2)\right\rangle \left\langle \delta^{(1)}(\mathbf{p})\delta^{(1)}(\mathbf{k}_3)\right\rangle \\
&\qquad + \left\langle \delta^{(1)}(\mathbf{q})\delta^{(1)}(\mathbf{k}_3)\right\rangle \left\langle \delta^{(1)}(\mathbf{p})\delta^{(1)}(\mathbf{k}_2)\right\rangle \,. \qquad \text{(A.7)}
\end{aligned}$$

Using the definition of the linear power spectrum,

$$\left\langle \delta^{(1)}(\mathbf{k})\delta^{(1)}(\mathbf{k}')\right\rangle = (2\pi)^3\delta_D(\mathbf{k}+\mathbf{k}')P_L(k)\,, \qquad \text{(A.8)}$$

the first Wick contraction gives a disconnected contribution proportional to $\delta_D(\mathbf{k}_1)$ and can be neglected for non-zero external momenta. The remaining two contractions give

$$\begin{aligned}
&\left\langle \delta^{(2)}(\mathbf{k}_1)\delta^{(1)}(\mathbf{k}_2)\delta^{(1)}(\mathbf{k}_3)\right\rangle \\
&\quad = (2\pi)^3\delta_D(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)\, 2F_2(\mathbf{k}_2,\mathbf{k}_3)P_L(k_2)P_L(k_3)\,. \qquad \text{(A.9)}
\end{aligned}$$

Including the two cyclic permutations, one obtains

$$\begin{aligned}
&\langle \delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3)\rangle \\
&\quad = (2\pi)^3\delta_D(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)\Big[2F_2(\mathbf{k}_1,\mathbf{k}_2)P_L(k_1)P_L(k_2) \\
&\qquad\qquad + 2F_2(\mathbf{k}_2,\mathbf{k}_3)P_L(k_2)P_L(k_3) \\
&\qquad\qquad + 2F_2(\mathbf{k}_1,\mathbf{k}_3)P_L(k_1)P_L(k_3)\Big]\,. \qquad \text{(A.10)}
\end{aligned}$$

Therefore, the tree-level bispectrum is

$$B(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) = 2F_2(\mathbf{k}_1,\mathbf{k}_2)P_L(k_1)P_L(k_2) + \text{cyc.} \qquad \text{(A.11)}$$

Appendix B

Derivation of the Marked Power Spectrum

In this section, we will recover the result 4.16.
Starting from the expression for the marked power spectrum and inserting the expressions for the two delta

$$\begin{aligned}M_{13}(k) =&\delta^{(1)}_M(\mathbf{k})\delta^{(3)}_M(-\mathbf{k})\\ =&H_1(\mathbf{k})\delta^{(1)}(\mathbf{k})\mathcal{I}_{\mathbf{k},\mathbf{q_1},\mathbf{q_2},\mathbf{q_3}}H_3(\mathbf{q_1},\mathbf{q_2},\mathbf{q_3})\delta^{(1)}(\mathbf{q_1})\delta^{(1)}(\mathbf{q_2})\delta^{(1)}(\mathbf{q_3})\,,\end{aligned} \tag{B.1}$$

where

$$\mathcal{I}_{\mathbf{k},\mathbf{q_1},\mathbf{q_2},\mathbf{q_3}} \equiv \int\frac{d^3\mathbf{q_1}}{(2\pi)^3}\int\frac{d^3\mathbf{q_2}}{(2\pi)^3}\int\frac{d^3\mathbf{q_3}}{(2\pi)^3}(2\pi)^3\delta_D(\mathbf{k}-\mathbf{q_1}-\mathbf{q_2}-\mathbf{q_3})\,, \tag{B.2}$$

Putting everything together and applying Wick theorem 2.2.2, we obtain three contributions

$$M_{13}(k) =H_1(\mathbf{k})\int\frac{d^3\mathbf{q_1}}{(2\pi)^3}\int\frac{d^3\mathbf{q_2}}{(2\pi)^3}\int\frac{d^3\mathbf{q_3}}{(2\pi)^3}H_3(\mathbf{q}_1,\mathbf{q}_2\mathbf{q}_3)(A+B+C) \tag{B.3}$$

$$(2\pi)^3\delta_D(-\mathbf{k}-\mathbf{q_1}-\mathbf{q_2}-\mathbf{q_3})\,, \tag{B.4}$$

where A, B and C are the contractions:

$$\begin{aligned}A &= \langle\delta^{(1)}(\mathbf{k})\delta^{(1)}(\mathbf{q_1})\rangle\,\langle\delta^{(1)}(\mathbf{q_2})\delta^{(1)}(\mathbf{q_3})\rangle\\ &= (2\pi)^6\delta_D(\mathbf{k}+\mathbf{q_1})\delta_D(\mathbf{q_2}+\mathbf{q_3})P_L(k)P_L(q_2)\,, && \text{(B.5)}\\ B &= \langle\delta^{(1)}(\mathbf{k})\delta^{(1)}(\mathbf{q_2})\rangle\,\langle\delta^{(1)}(\mathbf{q_1})\delta^{(1)}(\mathbf{q_3})\rangle\\ &= (2\pi)^6\delta_D(\mathbf{k}+\mathbf{q_2})\delta_D(\mathbf{q_1}+\mathbf{q_3})P_L(k)P_L(q_1)\,, && \text{(B.6)}\\ C &= \langle\delta^{(1)}(\mathbf{k})\delta^{(1)}(\mathbf{q_3})\rangle\,\langle\delta^{(1)}(\mathbf{q_1})\delta^{(1)}(\mathbf{q_2})\rangle\\ &= (2\pi)^6\delta_D(\mathbf{k}+\mathbf{q_3})\delta_D(\mathbf{q_1}+\mathbf{q_2})P_L(k)P_L(q_1)\,. && \text{(B.7)}\\ & && \text{(B.8)}\end{aligned}$$

Using the Dirac deltas to integrate over two of the q_i we obtain:

$$\begin{aligned}M_{13}(k) =&H_1(\mathbf{k})\int d^3\mathbf{q_2}H_3(-\mathbf{k},\mathbf{q_2},-\mathbf{q_2})P_L(k)P_L(q_2)\delta_D(-\mathbf{k}+\mathbf{k}-\mathbf{q_2}+\mathbf{q_2})\\ &+H_1(\mathbf{k})\int d^3\mathbf{q_1}H_3(\mathbf{q_1},-\mathbf{k},-\mathbf{q_1})P_L(k)P_L(q_1)\delta_D(-\mathbf{k}-\mathbf{q_1}+\mathbf{k}+\mathbf{q_1})\\ &+H_1(\mathbf{k})\int d^3\mathbf{q_1}H_3(\mathbf{q_1},-\mathbf{q_1},-\mathbf{k})P_L(k)P_L(q_1)\delta_D(-\mathbf{k}-\mathbf{q_1}+\mathbf{q_1}+\mathbf{k})\,.\end{aligned} \tag{B.9}$$

These three contributions are equivalent, so we can rewrite everything as:

$$\begin{aligned}M_{13}(k) =&3H_1(\mathbf{k})P_L(k)\int\frac{d^3\mathbf{q}}{(2\pi)^3}H_3(-\mathbf{k},\mathbf{q},-\mathbf{q})P_L(q)(2\pi)^3\delta_D(0)\\ =&3H_1(\mathbf{k})P_L(k)\int\frac{d^3\mathbf{q}}{(2\pi)^3}H_3(-\mathbf{k},\mathbf{q},-\mathbf{q})P_L(q)\,.\end{aligned} \tag{B.10}$$

The calculation for $M_{22}(\mathbf{k})$ is analogous.

Appendix C

Full Fisher forecast

Here we present the triangle plot for all the parameters varied in the Fisher forecast for $k_{\max} = \{0.20, 0.25, 0.30\} h/\text{Mpc}$.

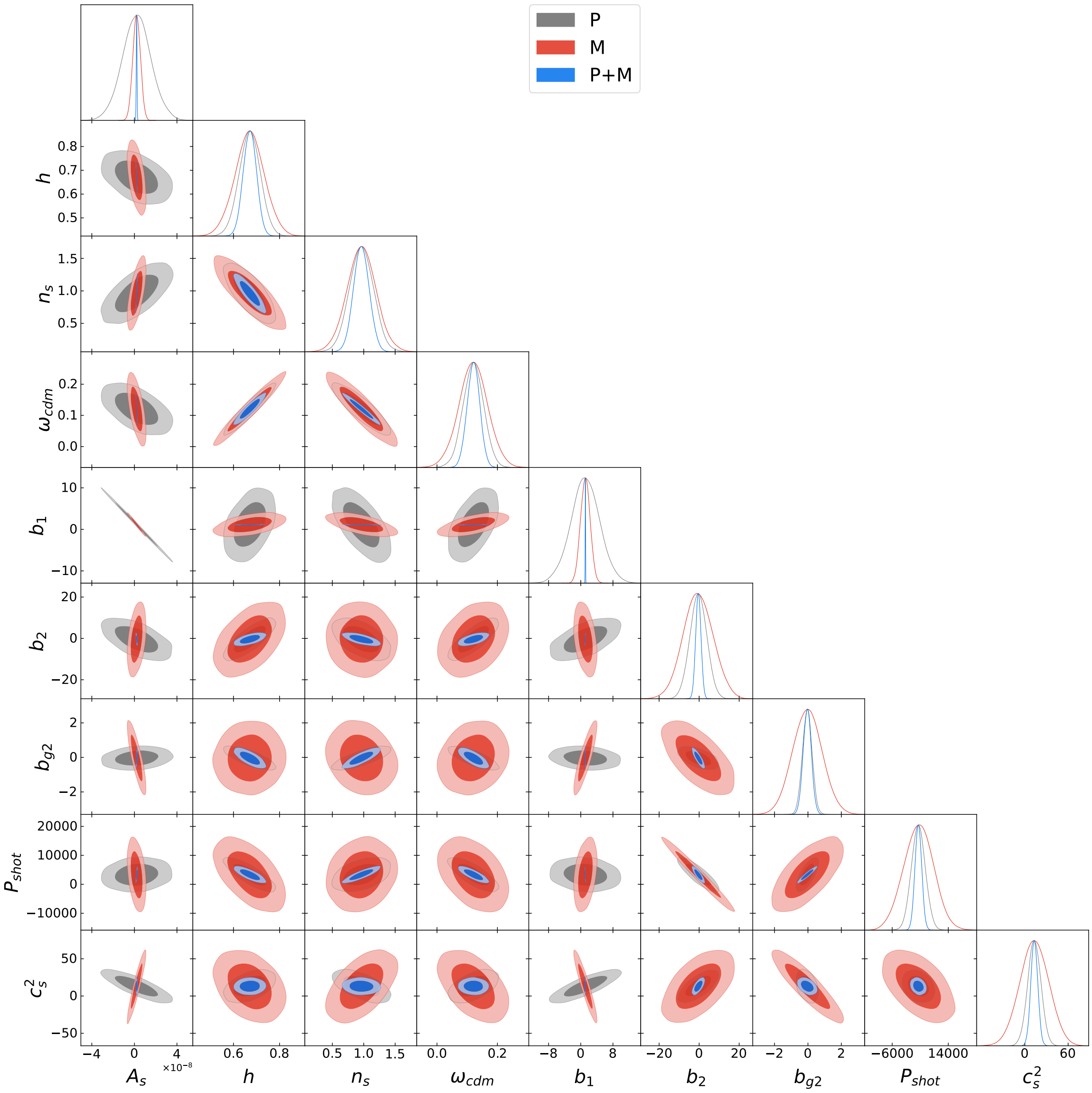

Figure C.1: Triangle plot: Fisher forecast constraints on the cosmological, EFT and bias parameters. Here we compare the constraining power of P only, M only and $P + M$ with $k_{\max} = 0.20\,h\text{Mpc}^{-1}$ and $R = 10\,\text{Mpc}/h$.

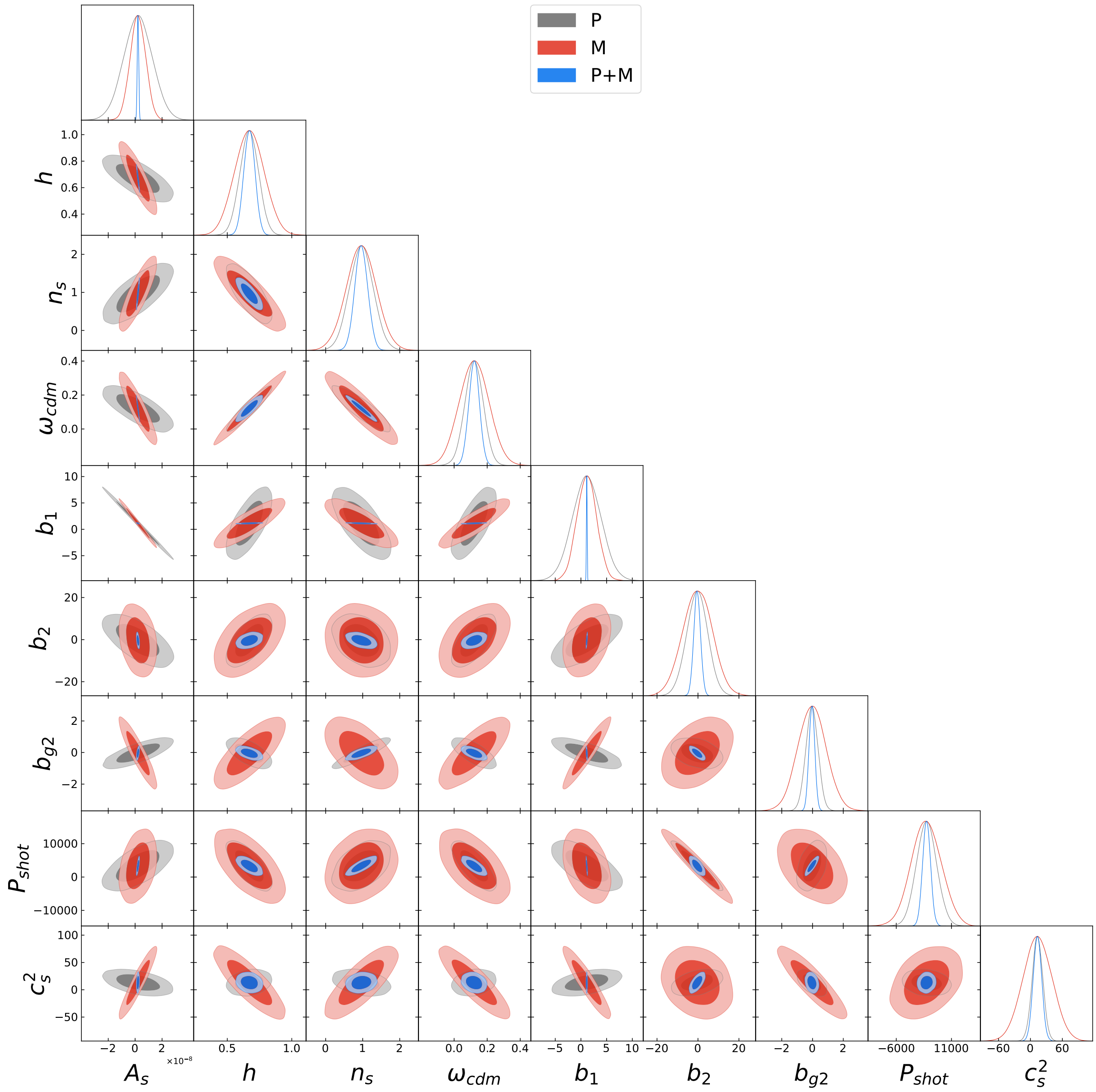

Figure C.2: Triangle plot: Fisher forecast constraints on the cosmological, EFT and bias parameters. Here we compare the constraining power of P only, M only and $P + M$ with $k_{\rm max} = 0.25\,h{\rm Mpc}^{-1}$ and $R = 10\,{\rm Mpc}/h$.

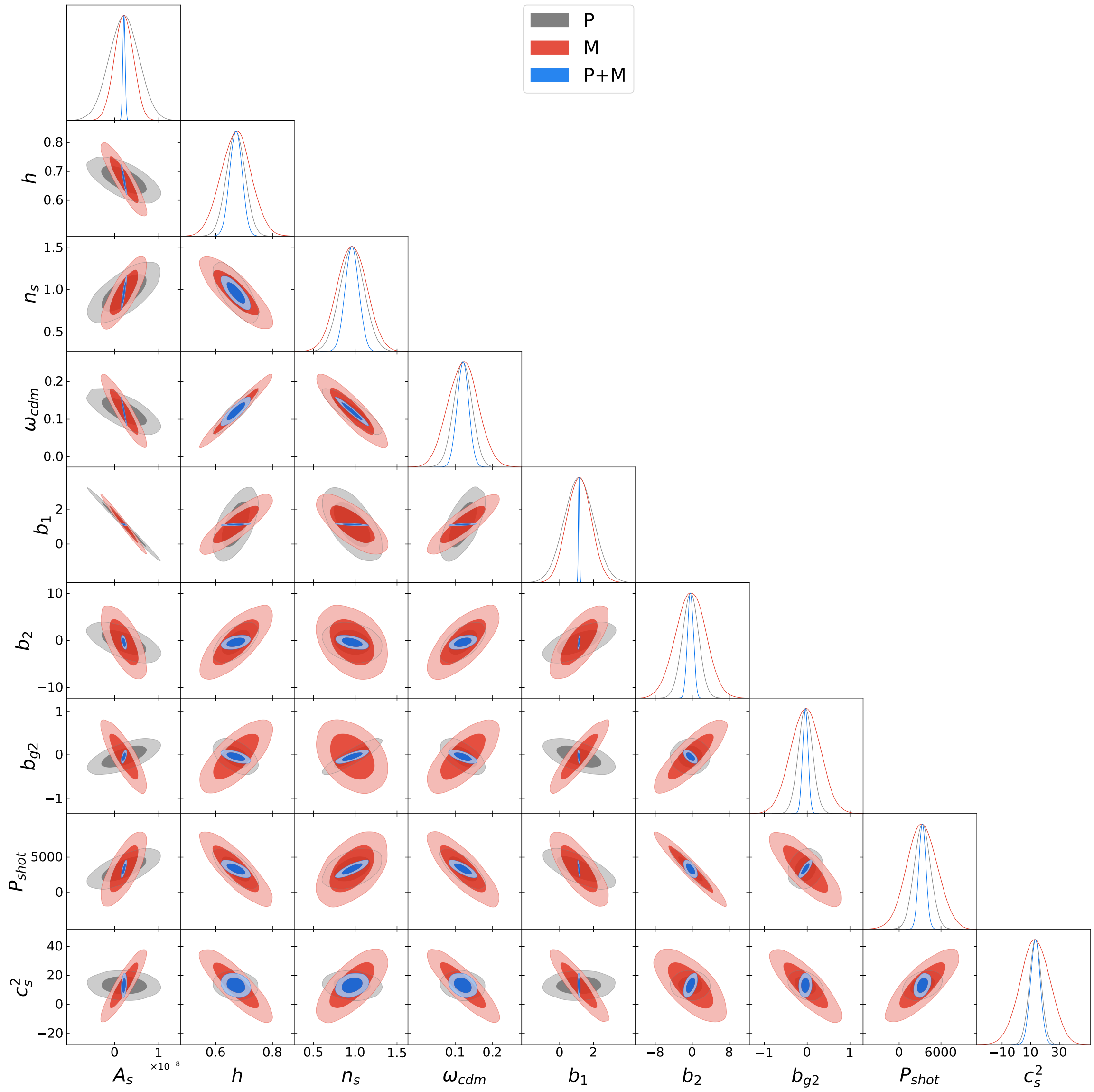

Figure C.3: Triangle plot: Fisher forecast constraints on the cosmological, EFT and bias parameters. Here we compare the constraining power of P only, M only and $P + M$ with $k_{\text{max}} = 0.30\, h\text{Mpc}^{-1}$ and $R = 10\,\text{Mpc}/h$.

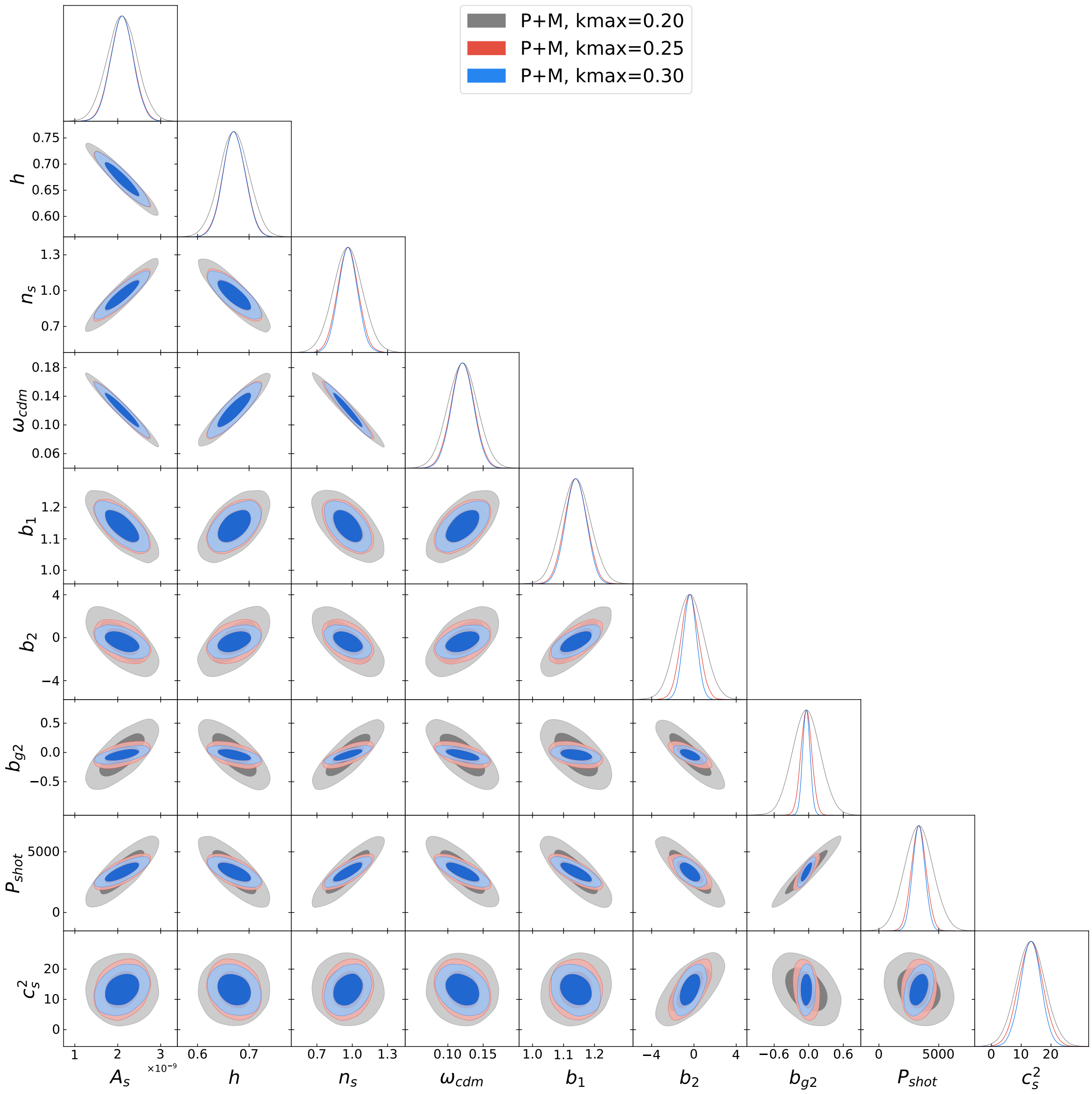

Figure C.4: Triangle plot of the constraint from $P+M$ on all the parameters considered in the analysis at different $k_{\max}$. Mark parameters: $R = 10\,\mathrm{Mpc}/h$, $p = 2$ and $\delta_{s^2} = 0.25$.

Bibliography

[1] G. Efstathiou, W. J. Sutherland, and S. J. Maddox. “The cosmological constant and cold dark matter”. In: *Nature* 348 (1990), pp. 705–707. DOI: 10.1038/348705a0.

[2] A. Einstein. “Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie”. In: *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften* (1917), pp. 142–152.

[3] R. Laureijs. *Euclid Assessment Study Report for the ESA Cosmic Visions*. 2009. arXiv: 0912.0914 [astro-ph.CO]. URL: https://arxiv.org/abs/0912.0914.

[4] DESI Collaboration et al. *Data Release 1 of the Dark Energy Spectroscopic Instrument*. 2026. arXiv: 2503.14745 [astro-ph.CO]. URL: https://arxiv.org/abs/2503.14745.

[5] Igor Andreoni et al. *Rubin ToO 2024: Envisioning the Vera C. Rubin Observatory LSST Target of Opportunity program*. 2024. arXiv: 2411.04793 [astro-ph.IM]. URL: https://arxiv.org/abs/2411.04793.

[6] Kyle S. Dawson et al. “THE BARYON OSCILLATION SPECTROSCOPIC SURVEY OF SDSS-III”. In: *The Astronomical Journal* 145.1 (Dec. 2012), p. 10. ISSN: 1538-3881. DOI: 10.1088/0004-6256/145/1/10. URL: http://dx.doi.org/10.1088/0004-6256/145/1/10.

[7] Azadeh Moradinezhad Dizgah et al. “Capturing non-Gaussianity of the large-scale structure with weighted skew-spectra”. In: *Journal of Cosmology and Astroparticle Physics* 2020.04 (2020), p. 011. DOI: 10.1088/1475-7516/2020/04/011. arXiv: 1911.05763 [astro-ph.CO].

[8] Martin White. “A marked correlation function for constraining modified gravity models”. In: *Journal of Cosmology and Astroparticle Physics* 2016.11 (Nov. 2016), pp. 057–057. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2016/11/057. URL: http://dx.doi.org/10.1088/1475-7516/2016/11/057.

[9] Oliver H. E. Philcox, Elena Massara, and David N. Spergel. “What does the marked power spectrum measure? Insights from perturbation theory”. In: *Physical Review D* 102.4 (Aug. 2020). ISSN: 2470-0029. DOI: 10.1103/physrevd.102.043516. URL: http://dx.doi.org/10.1103/PhysRevD.102.043516.

[10] Oliver H.E. Philcox, Alejandro Aviles, and Elena Massara. “Modeling the marked spectrum of matter and biased tracers in real- and redshift-space”. In: *Journal of Cosmology and Astroparticle Physics* 2021.03 (Mar. 2021), p. 038. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2021/03/038. URL: http://dx.doi.org/10.1088/1475-7516/2021/03/038.

[11] Elena Massara et al. “Using the Marked Power Spectrum to Detect the Signature of Neutrinos in Large-Scale Structure”. In: *Physical Review Letters* 126.1 (Jan. 2021). ISSN: 1079-7114. DOI: 10.1103/physrevlett.126.011301. URL: http://dx.doi.org/10.1103/PhysRevLett.126.011301.

[12] Marco Marinucci et al. *The constraining power of the Marked Power Spectrum: an analytical study.* 2024. arXiv: 2411.14377 [astro-ph.CO]. URL: https://arxiv.org/abs/2411.14377.

[13] Mark C. Neyrinck, István Szapudi, and Alexander S. Szalay. “REJUVENATING THE MATTER POWER SPECTRUM: RESTORING INFORMATION WITH A LOGARITHMIC DENSITY MAPPING”. In: *The Astrophysical Journal* 698.2 (May 2009), pp. L90–L93. ISSN: 1538-4357. DOI: 10.1088/0004-637x/698/2/l90. URL: http://dx.doi.org/10.1088/0004-637X/698/2/L90.

[14] Oliver Hahn et al. “Tidal effects and the environment dependence of halo assembly”. In: *Monthly Notices of the Royal Astronomical Society* 398.4 (Oct. 2009), pp. 1742–1756. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2009.15271.x. eprint: https://academic.oup.com/mnras/article-pdf/398/4/1742/3046988/mnras0398-1742.pdf. URL: https://doi.org/10.1111/j.1365-2966.2009.15271.x.

[15] Aseem Paranjape, Oliver Hahn, and Ravi K Sheth. “Halo assembly bias and the tidal anisotropy of the local halo environment”. In: *Monthly Notices of the Royal Astronomical Society* 476.3 (May 2018), pp. 3631–3647. ISSN: 0035-8711. DOI: 10.1093/mnras/sty496. eprint: https://academic.oup.com/mnras/article-pdf/476/3/3631/24503388/sty496.pdf. URL: https://doi.org/10.1093/mnras/sty496.

[16] Heling Yan, Zuhui Fan, and Simon D. M. White. *On the Tidal Dependence of Galaxy Properties.* 2012. arXiv: 1203.1225 [astro-ph.GA]. URL: https://arxiv.org/abs/1203.1225.

[17] Daniel Baumann et al. “Cosmological non-linearities as an effective fluid”. In: *Journal of Cosmology and Astroparticle Physics* 2012.07 (July 2012), pp. 051–051. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2012/07/051. URL: http://dx.doi.org/10.1088/1475-7516/2012/07/051.

[18] John Joseph M. Carrasco, Mark P. Hertzberg, and Leonardo Senatore. “The effective field theory of cosmological large scale structures”. In: *Journal of High Energy Physics* 2012.9 (Sept. 2012). ISSN: 1029-8479. DOI: 10.1007/jhep09(2012)082. URL: http://dx.doi.org/10.1007/JHEP09(2012)082.

[19] Leonardo Senatore and Matias Zaldarriaga. “The IR-resummed Effective Field Theory of Large Scale Structures”. In: *Journal of Cosmology and Astroparticle Physics* 2015.02 (Feb. 2015), pp. 013–013. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2015/02/013. URL: http://dx.doi.org/10.1088/1475-7516/2015/02/013.

[20] N. Aghanim et al. “Planck 2018 results: I. Overview and the cosmological legacy of Planck”. In: *Astronomy amp; Astrophysics* 641 (Sept. 2020), A1. ISSN: 1432-0746. DOI: 10.1051/0004-6361/201833880. URL: http://dx.doi.org/10.1051/0004-6361/201833880.

[21] Albert Einstein. "Die Grundlage der allgemeinen Relativitätstheorie". In: *Annalen der Physik* 354.7 (1916), pp. 769–822. DOI: 10.1002/andp.19163540702.

[22] Georges Lemaitre. "Un univers homogène de masse constante et de rayon croissant, rendant compte de la vitesse radiale des nébuleuses extra-galactiques". In: *Annales de la Société Scientifique de Bruxelles* 47 (1927), pp. 49–59.

[23] Edwin Hubble. "A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae". In: *Proceedings of the National Academy of Sciences* 15.3 (1929), pp. 168–173. DOI: 10.1073/pnas.15.3.168.

[24] Albert Einstein. "Die Feldgleichungen der Gravitation". In: *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften* (1915), pp. 844–847.

[25] Savvas Nesseris and Manuel Trashorras. "Can the homogeneity scale be used as a standard ruler?" In: *Physical Review D* 99.6 (Mar. 2019). ISSN: 2470-0029. DOI: 10.1103/physrevd.99.063539. URL: http://dx.doi.org/10.1103/PhysRevD.99.063539.

[26] Alexander Friedmann. ""Uber die Kr"ummung des Raumes". In: *Zeitschrift f"ur Physik* 10 (1922), pp. 377–386. DOI: 10.1007/BF01332580.

[27] N. Aghanim et al. "Planck 2018 results: VI. Cosmological parameters". In: *Astronomy and Astrophysics* 641 (Sept. 2020), A6. ISSN: 1432-0746. DOI: 10.1051/0004-6361/201833910. URL: http://dx.doi.org/10.1051/0004-6361/201833910.

[28] Cormac O'Raifeartaigh et al. "Einstein's 1917 static model of the universe: a centennial review". In: *The European Physical Journal H* 42.3 (July 2017), pp. 431–474. ISSN: 2102-6467. DOI: 10.1140/epjh/e2017-80002-5. URL: http://dx.doi.org/10.1140/epjh/e2017-80002-5.

[29] 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.

[30] Joel R. Primack, David Seckel, and Bernard Sadoulet. "Detection of Cosmic Dark Matter". In: *Annual Review of Nuclear and Particle Science* 38 (1988), pp. 751–807. DOI: 10.1146/annurev.ns.38.120188.003535.

[31] David H. Lyth and Andrew R. Liddle. *The Primordial Density Perturbation: Cosmology, Inflation and the Origin of Structure*. Cambridge, UK: Cambridge University Press, 2009.

[32] James Sunseri, Adrian E. Bayer, and Jia Liu. "Power of the cosmic web". In: *Physical Review D* 112.6 (Sept. 2025). ISSN: 2470-0029. DOI: 10.1103/grx3-hj7w. URL: http://dx.doi.org/10.1103/grx3-hj7w.

[33] F. Bernardeau et al. "Large-scale structure of the Universe and cosmological perturbation theory". In: *Physics Reports* 367.1-3 (Sept. 2002), pp. 1–248. ISSN: 0370-1573. DOI: 10.1016/s0370-1573(02)00135-7. URL: http://dx.doi.org/10.1016/S0370-1573(02)00135-7.

[34] Peter Coles. "Second-order evolution of Cold Dark Matter perturbations". In: *Monthly Notices of the Royal Astronomical Society* 243.2 (Dec. 1990), pp. 171–176. ISSN: 0035-8711. DOI: 10.1093/mnras/243.2.171. eprint: https://academic.oup.com/mnras/article-pdf/243/2/171/4051535/mnras243-0171.pdf. URL: https://doi.org/10.1093/mnras/243.2.171.

[35] Bhuvnesh Jain and Edmund Bertschinger. "Second-order power spectrum and nonlinear evolution at high redshift". In: *The Astrophysical Journal* 431 (Aug. 1994), p. 495. ISSN: 1538-4357. DOI: 10.1086/174502. URL: http://dx.doi.org/10.1086/174502.

[36] Roman Juszkiewicz. "On the evolution of cosmological adiabatic perturbations in the weakly non-linear regime". In: *Monthly Notices of the Royal Astronomical Society* 197.4 (Dec. 1981), pp. 931–940. ISSN: 0035-8711. DOI: 10.1093/mnras/197.4.931. eprint: https://academic.oup.com/mnras/article-pdf/197/4/931/3095469/mnras197-0931.pdf. URL: https://doi.org/10.1093/mnras/197.4.931.

[37] Michele Liguori et al. "Primordial Non-Gaussianity and Bispectrum Measurements in the Cosmic Microwave Background and Large-Scale Structure". In: *Advances in Astronomy* 2010.1 (Jan. 2010). Ed. by Dragan Huterer. ISSN: 1687-7977. DOI: 10.1155/2010/980523. URL: http://dx.doi.org/10.1155/2010/980523.

[38] Oliver H. E. Philcox. *An Introduction to the EFTofLSS*. https://oliverphilcox.github.io/assets/pdf/eft_intro.pdf. Unpublished lecture notes. 2023.

[39] 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.

[40] Anton Chudaykin et al. "Nonlinear perturbation theory extension of the Boltzmann code CLASS". In: *Physical Review D* 102.6 (Sept. 2020). ISSN: 2470-0029. DOI: 10.1103/physrevd.102.063533. URL: http://dx.doi.org/10.1103/PhysRevD.102.063533.

[41] 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.

[42] Héctor Gil-Marín et al. "The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: RSD measurement from the power spectrum and bispectrum of the DR12 BOSS galaxies". In: *Monthly Notices of the Royal Astronomical Society* 465.2 (Oct. 2016), pp. 1757–1788. ISSN: 1365-2966. DOI: 10.1093/mnras/stw2679. URL: http://dx.doi.org/10.1093/mnras/stw2679.

[43] Marko Simonović et al. "Cosmological perturbation theory using the FFTLog: formalism and connection to QFT loop integrals". In: *Journal of Cosmology and Astroparticle Physics* 2018.04 (Apr. 2018), pp. 030–030. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2018/04/030. URL: http://dx.doi.org/10.1088/1475-7516/2018/04/030.

[44] Oliver H.E. Philcox et al. “Cosmology with the redshift-space galaxy bispectrum monopole at one-loop order”. In: *Physical Review D* 106.4 (Aug. 2022). ISSN: 2470-0029. DOI: 10.1103/physrevd.106.043530. URL: http://dx.doi.org/10.1103/PhysRevD.106.043530.

[45] Joyce Byun et al. “Towards optimal cosmological parameter recovery from compressed bispectrum statistics”. In: *Monthly Notices of the Royal Astronomical Society* 471.2 (July 2017), pp. 1581–1618. ISSN: 1365-2966. DOI: 10.1093/mnras/stx1681. URL: http://dx.doi.org/10.1093/mnras/stx1681.

[46] John A. Peacock et al. “A measurement of the cosmological mass density from clustering in the 2dF Galaxy Redshift Survey”. In: *Nature* 410.6825 (Mar. 2001), pp. 169–173. ISSN: 1476-4687. DOI: 10.1038/35065528. URL: http://dx.doi.org/10.1038/35065528.

[47] 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.

[48] J. C. Jackson. “A Critique of Rees’s Theory of Primordial Gravitational Radiation”. In: *Monthly Notices of the Royal Astronomical Society* 156.1 (Feb. 1972), 1P–5P. ISSN: 1365-2966. DOI: 10.1093/mnras/156.1.1p. URL: http://dx.doi.org/10.1093/mnras/156.1.1P.

[49] Alan Heavens. *Statistical techniques in cosmology*. 2010. arXiv: 0906.0664 [astro-ph.CO]. URL: https://arxiv.org/abs/0906.0664.

[50] Andrew J. S. Hamilton. *Power Spectrum Estimation I. Basics*. 2005. arXiv: astro-ph/0503603 [astro-ph]. URL: https://arxiv.org/abs/astro-ph/0503603.

[51] William R Coulton et al. “Quijote-PNG: Simulations of Primordial Non-Gaussianity and the Information Content of the Matter Field Power Spectrum and Bispectrum”. In: *The Astrophysical Journal* 943.1 (Jan. 2023), p. 64. ISSN: 1538-4357. DOI: 10.3847/1538-4357/aca8a7. URL: http://dx.doi.org/10.3847/1538-4357/aca8a7.

[52] Héctor Gil-Marín et al. “The power spectrum and bispectrum of SDSS DR11 BOSS galaxies – I. Bias and gravity”. In: *Monthly Notices of the Royal Astronomical Society* 451.1 (July 2015), pp. 539–580. ISSN: 0035-8711. DOI: 10.1093/mnras/stv961. eprint: https://academic.oup.com/mnras/article-pdf/451/1/539/4177830/stv961.pdf. URL: https://doi.org/10.1093/mnras/stv961.

[53] Davide Gualdi, Héctor Gil-Marín, and Licia Verde. “Joint analysis of anisotropic power spectrum, bispectrum and trispectrum: application to N-body simulations”. In: *Journal of Cosmology and Astroparticle Physics* 2021.07 (July 2021), p. 008. DOI: 10.1088/1475-7516/2021/07/008. URL: https://doi.org/10.1088/1475-7516/2021/07/008.

[54] Giovanni Cabass et al. “Constraints on Single-Field Inflation from the BOSS Galaxy Survey”. In: *Phys. Rev. Lett.* 129 (2 July 2022), p. 021301. DOI: 10.1103/PhysRevLett.129.021301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.129.021301.

[55] Giovanni Cabass et al. “Constraints on multifield inflation from the BOSS galaxy survey”. In: *Phys. Rev. D* 106 (4 Aug. 2022), p. 043506. DOI: 10.1103/PhysRevD.106.043506. URL: https://link.aps.org/doi/10.1103/PhysRevD.106.043506.

[56] ChangHoon Hahn et al. *SIMBIG: The First Cosmological Constraints from Non-Gaussian and Non-Linear Galaxy Clustering.* 2023. arXiv: 2310.15246 [astro-ph.CO]. URL: https://arxiv.org/abs/2310.15246.

[57] ChangHoon Hahn et al. *SIMBIG: The First Cosmological Constraints from the Non-Linear Galaxy Bispectrum.* 2023. arXiv: 2310.15243 [astro-ph.CO]. URL: https://arxiv.org/abs/2310.15243.

[58] Massimo Pietroni and Fabian Schmidt. *On the Relation Between Field-Level Posteriors, Correlators, and their Likelihoods.* 2026. arXiv: 2604.25385 [astro-ph.CO]. URL: https://arxiv.org/abs/2604.25385.

[59] Eleni Tsaprazi et al. “Field-level inference of galaxy intrinsic alignment from the SDSS-III BOSS survey”. In: *Journal of Cosmology and Astroparticle Physics* 2022.08 (Aug. 2022), p. 003. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2022/08/003. URL: http://dx.doi.org/10.1088/1475-7516/2022/08/003.

[60] Hugo Simon-Onfroy, François Lanusse, and Arnaud Mattia. “Benchmarking field-level cosmological inference from galaxy redshift surveys”. In: (Apr. 2025). DOI: 10.48550/arXiv.2504.20130.

[61] Florent Leclercq. *Field-level inference in cosmology.* 2025. arXiv: 2509.13435 [astro-ph.CO]. URL: https://arxiv.org/abs/2509.13435.

[62] Bruno Régaldo-Saint Blancard et al. “Galaxy clustering analysis with SimBIG and the wavelet scattering transform”. In: *Physical Review D* 109.8 (Apr. 2024). ISSN: 2470-0029. DOI: 10.1103/physrevd.109.083535. URL: http://dx.doi.org/10.1103/PhysRevD.109.083535.

[63] Georgios Valogiannis, Francisco Villaescusa-Navarro, and Marco Baldi. “Towards unveiling the large-scale nature of gravity with the wavelet scattering transform”. In: *Journal of Cosmology and Astroparticle Physics* 2024.11 (Sept. 2024), p. 061. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2024/11/061. URL: http://dx.doi.org/10.1088/1475-7516/2024/11/061.

[64] Michael Eickenberg et al. *Wavelet Moments for Cosmological Parameter Estimation.* 2022. arXiv: 2204.07646 [astro-ph.CO]. URL: https://arxiv.org/abs/2204.07646.

[65] Georgios Valogiannis, Sihan Yuan, and Cora Dvorkin. “Precise cosmological constraints from BOSS galaxy clustering with a simulation-based emulator of the wavelet scattering transform”. In: *Phys. Rev. D* 109 (10 May 2024), p. 103503. DOI: 10.1103/PhysRevD.109.103503. URL: https://link.aps.org/doi/10.1103/PhysRevD.109.103503.

[66] Georgios Valogiannis and Cora Dvorkin. “Towards an optimal estimation of cosmological parameters with the wavelet scattering transform”. In: *Phys. Rev. D* 105 (10 May 2022), p. 103534. DOI: 10.1103/PhysRevD.105.103534. URL: https://link.aps.org/doi/10.1103/PhysRevD.105.103534.

[67] Arka Banerjee and Tom Abel. “Tracer-field cross-correlations with k-nearest neighbour distributions”. In: *Monthly Notices of the Royal Astronomical Society* 519.4 (Mar. 2023), pp. 4856–4868. ISSN: 0035-8711. DOI: 10.1093/mnras/stac3813. eprint: https://academic.oup.com/mnras/article-pdf/519/4/4856/48745149/stac3813.pdf. URL: https://doi.org/10.1093/mnras/stac3813.

[68] William R Coulton, Tom Abel, and Arka Banerjee. “Small-scale signatures of primordial non-Gaussianity in k-nearest neighbour cumulative distribution functions”. In: *Monthly Notices of the Royal Astronomical Society* 534.3 (Nov. 2024), pp. 1621–1633. ISSN: 0035-8711. DOI: 10.1093/mnras/stae2108. eprint: https://academic.oup.com/mnras/article-pdf/534/3/1621/59594366/stae2108.pdf. URL: https://doi.org/10.1093/mnras/stae2108.

[69] Marcel Schmittfull and Azadeh Moradinezhad Dizgah. “Galaxy skew-spectra in redshift-space”. In: *Journal of Cosmology and Astroparticle Physics* 2021.03 (Mar. 2021), p. 020. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2021/03/020. URL: http://dx.doi.org/10.1088/1475-7516/2021/03/020.

[70] Jiamin Hou et al. “Cosmological constraints from the redshift-space galaxy skew spectra”. In: *Phys. Rev. D* 109 (10 May 2024), p. 103528. DOI: 10.1103/PhysRevD.109.103528. URL: https://link.aps.org/doi/10.1103/PhysRevD.109.103528.

[71] Fergus Simpson, Alan F. Heavens, and Catherine Heymans. “Clipping the cosmos. II. Cosmological information from nonlinear scales”. In: *Phys. Rev. D* 88 (8 Oct. 2013), p. 083510. DOI: 10.1103/PhysRevD.88.083510. URL: https://link.aps.org/doi/10.1103/PhysRevD.88.083510.

[72] Martha Lippich and Ariel G Sánchez. “medusa: Minkowski functionals estimated from Delaunay tessellations of the three-dimensional large-scale structure”. In: *Monthly Notices of the Royal Astronomical Society* 508.3 (Dec. 2021), pp. 3771–3784. ISSN: 0035-8711. DOI: 10.1093/mnras/stab2820. eprint: https://academic.oup.com/mnras/article-pdf/508/3/3771/40767773/stab2820.pdf. URL: https://doi.org/10.1093/mnras/stab2820.

[73] Wei Liu, Aoxiang Jiang, and Wenjuan Fang. “Probing massive neutrinos with the Minkowski functionals of the galaxy distribution”. In: *Journal of Cosmology and Astroparticle Physics* 2023.09 (Sept. 2023), p. 037. DOI: 10.1088/1475-7516/2023/09/037. URL: https://doi.org/10.1088/1475-7516/2023/09/037.

[74] Dietrich Stoyan. “On Correlations of Marked Point Processes”. In: *Mathematische Nachrichten* 116.1 (1984), pp. 197–207. DOI: https://doi.org/10.1002/mana.19841160115. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mana.19841160115. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/mana.19841160115.

[75] Tony Bonnaire et al. “Cosmology with cosmic web environments: I. Real-space power spectra”. In: *Astronomy and Astrophysics* 661 (May 2022), A146. ISSN: 1432-0746. DOI: 10.1051/0004-6361/202142852. URL: http://dx.doi.org/10.1051/0004-6361/202142852.

[76] Marco Marinucci et al. *Optimal marked statistics from the Effective Field Theory of Large Scale Structure*. 2026. arXiv: 2606.14332 [astro-ph.CO]. URL: https://arxiv.org/abs/2606.14332.

[77] Joseph Clampitt, Yan-Chuan Cai, and Baojiu Li. "Voids in modified gravity: excursion set predictions". In: *Monthly Notices of the Royal Astronomical Society* 431.1 (Mar. 2013), pp. 749–766. ISSN: 1365-2966. DOI: 10.1093/mnras/stt219. URL: http://dx.doi.org/10.1093/mnras/stt219.

[78] Huiyuan Wang et al. "Reconstructing the cosmic velocity and tidal fields with galaxy groups selected from the Sloan Digital Sky Survey: Cosmic velocity and tidal fields". In: *Monthly Notices of the Royal Astronomical Society* 420.2 (Dec. 2011), pp. 1809–1824. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2011.20174.x. URL: http://dx.doi.org/10.1111/j.1365-2966.2011.20174.x.

[79] P. Erdogdu et al. "Reconstructed density and velocity fields from the 2MASS Redshift Survey". In: *Monthly Notices of the Royal Astronomical Society* 373.1 (Nov. 2006), pp. 45–64. ISSN: 1365-2966. DOI: 10.1111/j.1365-2966.2006.11049.x. URL: http://dx.doi.org/10.1111/j.1365-2966.2006.11049.x.

[80] Huiyuan Wang et al. "Reconstructing the cosmic density field with the distribution of dark matter haloes". In: *Monthly Notices of the Royal Astronomical Society* 394.1 (Mar. 2009), pp. 398–414. ISSN: 1365-2966. DOI: 10.1111/j.1365-2966.2008.14301.x. URL: http://dx.doi.org/10.1111/j.1365-2966.2008.14301.x.

[81] Anton Chudaykin et al. "Nonlinear perturbation theory extension of the Boltzmann code CLASS". In: *Physical Review D* 102.6 (Sept. 2020). ISSN: 2470-0029. DOI: 10.1103/physrevd.102.063533. URL: http://dx.doi.org/10.1103/PhysRevD.102.063533.

[82] 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.

[83] Diego Blas, Julien Lesgourgues, and Thomas Tram. "The Cosmic Linear Anisotropy Solving System (CLASS). Part II: Approximation schemes". In: *Journal of Cosmology and Astroparticle Physics* 2011.07 (July 2011), pp. 034–034. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2011/07/034. URL: http://dx.doi.org/10.1088/1475-7516/2011/07/034.

[84] Julien Lesgourgues. *The Cosmic Linear Anisotropy Solving System (CLASS) III: Comparision with CAMB for LambdaCDM*. 2011. arXiv: 1104.2934 [astro-ph.CO]. URL: https://arxiv.org/abs/1104.2934.

[85] Anton Chudaykin and Mikhail M. Ivanov. "Measuring neutrino masses with large-scale structure: Euclid forecast with controlled theoretical error". In: *Journal of Cosmology and Astroparticle Physics* 2019.11 (Nov. 2019), p. 034. DOI: 10.1088/1475-7516/2019/11/034. URL: https://doi.org/10.1088/1475-7516/2019/11/034.

[86] Vincent Desjacques, Donghui Jeong, and Fabian Schmidt. "Large-scale galaxy bias". In: *Physics Reports* 733 (Feb. 2018), pp. 1–193. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2017.12.002. URL: http://dx.doi.org/10.1016/j.physrep.2017.12.002.

[87] Muntazir Mehdi Abidi and Tobias Baldauf. "Cubic halo bias in Eulerian and Lagrangian space". In: *Journal of Cosmology and Astroparticle Physics* 2018.07 (July 2018), pp. 029–029. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2018/07/029. URL: http://dx.doi.org/10.1088/1475-7516/2018/07/029.

[88] Anton Chudaykin and Mikhail M. Ivanov. "Measuring neutrino masses with large-scale structure: Euclid forecast with controlled theoretical error". In: *Journal of Cosmology and Astroparticle Physics* 2019.11 (Nov. 2019), pp. 034–034. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2019/11/034. URL: http://dx.doi.org/10.1088/1475-7516/2019/11/034.

[89] Mikel Martin Barandiaran et al. *Weighted Webs: Morphology-Informed Marked Fields*. July 2026. DOI: 10.48550/arXiv.2607.26021. arXiv: 2607.26021.

[90] M. Kärcher, J. Bel, and S. de la Torre. "Towards an optimal marked correlation function analysis for the detection of modified gravity". In: *Astronomy and; Astrophysics* 694 (Feb. 2025), A253. ISSN: 1432-0746. DOI: 10.1051/0004-6361/202450977. URL: http://dx.doi.org/10.1051/0004-6361/202450977.

[91] Oliver H.E. Philcox and Mikhail M. Ivanov. "BOSS DR12 full-shape cosmology: constraints from the large-scale galaxy power spectrum and bispectrum monopole". In: *Physical Review D* 105.4 (Feb. 2022). ISSN: 2470-0029. DOI: 10.1103/physrevd.105.043517. URL: http://dx.doi.org/10.1103/PhysRevD.105.043517.

[92] Mikhail M. Ivanov et al. "Precision analysis of the redshift-space galaxy bispectrum". In: *Physical Review D* 105.6 (Mar. 2022). ISSN: 2470-0029. DOI: 10.1103/physrevd.105.063512. URL: http://dx.doi.org/10.1103/PhysRevD.105.063512.

[93] Dionysios Karagiannis et al. "Constraining primordial non-Gaussianity with bispectrum and power spectrum from upcoming optical and radio surveys". In: *Monthly Notices of the Royal Astronomical Society* 478.1 (July 2018), pp. 1341–1376. ISSN: 0035-8711. DOI: 10.1093/mnras/sty1029. eprint: https://academic.oup.com/mnras/article-pdf/478/1/1341/25027360/sty1029.pdf. URL: https://doi.org/10.1093/mnras/sty1029.

[94] Y. and Mellier et al. "Euclid: I. Overview of the Euclid mission". In: *Astronomy and Astrophysics* 697 (Apr. 2025), A1. ISSN: 1432-0746. DOI: 10.1051/0004-6361/202450810. URL: http://dx.doi.org/10.1051/0004-6361/202450810.