SISSA

Scuola
Internazionale
Superiore di
Studi Avanzati

Physics Area – PhD course in

Astroparticle Physics

Toward a Complete Cosmological Analysis of Galaxy Clustering Measurements from Spectroscopic Redshift Surveys

Candidate: Advisor:

Andrea Oddo   Dr. Emiliano Sefusatti

Academic Year 2019-20

Abstract

The forthcoming generation of galaxy redshift surveys will sample the large-scale structure of the Universe over unprecedented volumes with high-density tracers. This advancement will make robust measurements of three-point clustering statistics not only possible, but necessary in order to exploit the surveys full potential to constrain cosmological models.

My Ph.D. project is conceived specifically for this improvement. Its main goal is the development of a software pipeline for the analysis of the joined galaxy power spectrum and bispectrum. In a first stage, my collaborators and I investigate how several methodological choices can influence inferences based on the bispectrum about galaxy bias and shot noise. We consider dark-matter halos, extracted from N-body simulations, of at least $\sim 10^{13} M_\odot$. While these are not representative of a realistic distribution of objects that is observed by redshift surveys, it is possible to extract a large number of synthetic catalogs of this type of objects from N-body simulations, and this still allows for comparison of perturbative models. We estimate the covariance matrix of the measurement errors by using 10,000 mock catalogues generated with the PINOCCHIO code, and then we fit a series of theoretical models based on tree-level perturbation theory to the numerical data. We study how the model constraints are influenced by the binning strategy for the bispectrum configurations and by the form of the likelihood function. We also use Bayesian model-selection techniques to single out the optimal theoretical description of our data. We find that a three-parameter bias model at tree-level combined with Poissonian shot noise is necessary to model the halo bispectrum up to scales of $k_\mathrm{max} \lesssim 0.09 \, h \, \mathrm{Mpc}^{-1}$, although fitting formulae that relate the bias parameters can be helpful to reduce the freedom of the model without compromising accuracy. Our data clearly disfavour local Eulerian and local Lagrangian bias models and do not require corrections to Poissonian shot noise.

We then approach our final goal of a simultaneous analysis of the power spectrum and bispectrum in real space. We fit measurements of power spectrum and bispectrum of dark-matter halos from the same set of N-body simulations, with a full covariance, including cross correlations between power spectrum and bispectrum, estimated by the same 10,000 mock catalogues. The theoretical models employed are perturbative predictions at tree-level for the bispectrum, and at one-loop level for the power spectrum, both based on the Effective Field Theory of the Large Scale Structure, including infrared resummation, that represent the state of the art in the analysis of galaxy clustering measurements. We focus on the constraints of bias and shot-noise parameters as a function of k_max, we study extensions of the parameter space and possible reductions through either phenomenological of physically-motivated bias relations; we explore the impact of different covariance approximations and binning effects on the theoretical predictions. We find that a joint fit of power spectrum and bispectrum with 4 bias parameters, one EFT counterterm and two shot-noise parameters gives a good description of our data up to $k_\mathrm{max,P} \simeq 0.21 \, h \, \mathrm{Mpc}^{-1}$ and $k_\mathrm{max,B} \simeq 0.09 \, h \, \mathrm{Mpc}^{-1}$. In this particular setup, we perform a simultaneous fit of power spectrum and bispectrum including cosmological parameters, and consistently evaluating the theoretical predictions at each sampled point in parameter space. We recover the correct value of the cosmological parameters used to run the N-body simulations.

We envision that the addition of the galaxy bispectrum to the galaxy power spectrum, being able to break degeneracies between the model parameters, will allow much tighter constraints on cosmological parameters in future analyses of actual data.

Contents

Chapter 1

Introduction

The standard ΛCDM cosmological model has been proven successful in describing the early-time evolution of our Universe ([1, 2], and companion papers), its late-time accelerated expansion [3, 4, 5], and the distribution of matter and galaxies at large scales ([6, 7, 8, 9]). Its main success is related to the fact that it is capable of describing a large variety of cosmological observations in terms of a relatively small set of parameters. However, the ΛCDM model fails to describe the nature of two of the main components of the Universe: dark matter and dark energy. The existence of a form of cold dark matter – a non-relativistic type of matter which is not made of known species of the standard model of particle physics – has been inferred by observations of rotation curves of galaxies [10], of the motion of objects in galaxy clusters [11], and from measurements of the power spectrum of temperature anisotropies of the Cosmic Microwave Background (CMB) radiation [12, 2]. In addition, the late-time accelerated expansion of the Universe – first probed by supernovae observations two decades ago [3, 4] – cannot be explained by the standard cosmological model without introducing a negative-pressure component, dark energy, that, again, cannot be trivially related to specific species of the standard model of particle physics. According to the most precise measurements coming from CMB experiments [2], roughly 95% of the Universe is composed by either cold dark matter or dark energy, and these represent two of the main directions of research in modern cosmology.

While opening the way to the era of precision cosmology, CMB experiments mainly probe the early-time history of the Universe; they focus on the epochs of recombination and decoupling of baryonic matter from the early-age radiation of the cosmos, when density fluctuations were extremely small compared to the homogeneous component, and their dynamics could be well described by linear Perturbation Theory (PT). In particular, CMB experiments measure the temperature anisotropies and the polarization of the cosmic microwave background radiation, and through summary statistics – with these being mainly the angular temperature power spectrum, the polarization power spectra, and all non-vanishing cross-spectra – cosmological information can be inferred.

Complementing the study of the early stages of the cosmos, the Large-Scale Structure (LSS), that is the distribution of dark matter, galaxies, and galaxy clusters at large scales, is the main observational probe of the late-time Universe. As opposed to the CMB, the matter density at late times and at small scales is characterised by large fluctuations: very dense objects, such as the dark-matter halos surrounding clusters of galaxies, are separated by large voids. This highly nonlinear evolution is mainly the result of gravitational instability, while we will describe later additional non-linear effects responsible for the observed galaxy distribution, the main focus of this thesis.

Detailed studies of the LSS of the Universe are today possible thanks to several observational probes. One of these is weak gravitational lensing, that measures the shape deformation of background galaxies whose light, travelling toward us, is deflected and distorted by the dark matter distribution. Such measurements can be used to infer matter perturbations along the line-of-sight, making weak lensing one of the most direct probes of the dark matter density. Another important probe is represented by large galaxy redshift surveys, that measure the redshifts and angular positions of millions of galaxies. As the galaxy distribution traces the underlying matter distribution, the clustering of galaxies is strictly related to the clustering of matter. For this reason it provides information on the late-time growth of matter perturbations, an important dynamical test for models of dark energy, as well as direct information on the accelerated expansion by means of the so-called Baryonic Acoustic Oscillations.

In the study of galaxy clustering of the LSS, the main observables are the two-point correlation function (2PCF) and the power spectrum. The 2PCF measures the excess probability of observing a pair of galaxies at a given distance in space. It shows a distinct peak at physical separations of order 150 Mpc. This feature is a remnant of the density oscillations experienced by the tightly coupled baryon-photon fluid before the decoupling of photons, the so-called Baryon Acoustic Oscillations, or BAOs [13]. Since the physics of the BAO, and specifically the distance travelled by the acoustic waves in the baryon density, is well known from CMB observations, the position of the peak can be used as a standard ruler, that is a known physical distance that, detected in a galaxy population at a given redshift, allows to relate this to the physical distance. Modelling and measuring the 2PCF, however, is non-trivial, since it relies on procedures that can be technically complicated and numerically expensive. Its Fourier-space counterpart, the power spectrum, has the advantage of separating perturbations at different scales and, as we will see, more directly relate to the theoretical modelling. For this reason, the power spectrum is commonly employed to study also extensions of our cosmological model, including warm dark matter species [14], neutrino masses [15, 16], and non-Gaussian initial conditions imprinted into the primordial LSS [17].

The Baryon Oscillation Spectroscopic Survey (*BOSS*) [18] is one of the most recent galaxy redshift surveys, with the analysis of its clustering data being presented in a large number of papers (*e.g.* [6, 7, 8, 9]). It observed the position of approximately 1.2 million galaxies over a redshift range $z = 0.15$ to 0.7. The constraints on cosmological parameters obtained from its BAO measurements are now routinely complementing CMB observations as the standard constraints on the ΛCDM cosmological models. The next batch of galaxy redshift surveys, aiming at uncovering the origin of the accelerated expansion of the Universe, and expected to collect data in the next few years, is designed to map the distribution of tracers of the LSS over comoving volumes of much larger size and at larger redshift, typically in the range $z = 1$ to 2, *e.g.* [19, 20, 21].

In order to make full use of such observational efforts, however, it is necessary to go beyond the traditional studies based solely on the power spectrum and 2PCF. The 3-point correlation function (3PCF), measuring the excess probability of observing three galaxies in a given triangular configuration, is the first step forward, alongside its Fourier-space counterpart, the galaxy bispectrum.

The bispectrum is the lowest order correlation function in Fourier space quantifying the non-Gaussianity of the galaxy distribution and, as such, has received considerable attention in the last years, with many works aiming at a proper assessment of the additional cosmological information it can provide. In this context, efforts have been made in different directions: simple signal-to-

noise measurements and Fisher forecasts analyses have quantified the information content of the redshift-space galaxy bispectrum in the standard ΛCDM model [22, 23, 24], as well as in models with massive neutrinos [25, 26]; due to its direct relation to any source of non-Gaussianity of the galaxy distribution, forecasts have focused also on the possible constraints that it can provide on different types of primordial non-Gaussianities [27, 28, 29]; moreover, the impact of the bispectrum in constraining other extensions to the standard cosmological model has been explored, such as in the context of dynamical dark energy [30] and modified gravity [31]. Other Fisher information studies on the redshift-space power spectrum and bispectrum show that the inclusion of the bispectrum helps significantly in breaking degeneracies between parameters, and that for Euclid-like surveys this could lead to a 10% measurement of the linear growth rate of density perturbations [32].

The galaxy bispectrum has been measured most recently in the BOSS galaxy survey and analysed in [33, 34, 35] with a first detection of BAOs presented in [36] (see also [37] for an even earlier detection in the 3PCF).

An inherent difficulty in the analysis of the galaxy bispectrum derives from the fact that the signal is distributed over a large number of triangle configurations, typically compressed into ~ 1000 bins for the range of scales covering the linear and quasi-linear regimes. Making unbiased inferences thus requires the robust estimation of a large covariance matrix for the bispectrum measurements, and of its inverse (the precision matrix). These are often obtained numerically, using a large set of mock catalogs or N-body simulations [38, 39, 40, 23, 33, 35]. In this case, it is necessary to account for statistical and systematic errors in the precision matrix when writing a likelihood function for the model parameters [35]. Other viable alternatives that can be employed in order to avoid estimating large covariance matrices are represented by different approaches involving the compression of the information content of the bispectrum into a small number of data points [41, 42].

Recent analyses of the galaxy power spectrum employ perturbative predictions at one-loop level, that is the first correction to the linear theory prediction. In particular, the modeling of the power spectrum based on the Effective Field Theory of the Large Scale Structure (EFTofLSS), including galaxy biasing, redshift-space distortions, and infrared resummation, [43, 44, 45, 46] has been thoroughly tested in [47], where a set of large N-body simulations has been used in a blinded challenge, in order to compare parameters constraints coming from the fit of power spectrum multipoles in redshift space, and to test the performance of perturbation theory with very small statistical uncertainties. The same perturbative model for the one-loop power spectrum, along with some alternative ones, has been used as well to re-analyse the data from BOSS [48, 16, 49]. These works represent the state of the art in the analysis of the galaxy power spectrum.

At the same time, there has been a considerable effort to develop similar perturbative predictions for the galaxy bispectrum as well. One-loop corrections to the leading order, *i.e.* tree-level, prediction of the matter and galaxy bispectrum, have been explored extensively over the last twenty years [50, 51, 52, 53, 54, 55, 56, 57, 58, 59] as well as full models for the redshift-space bispectrum [60]. In the context of Lagrangian perturbation theory, the modelling of the matter bispectrum has been studied as well up to one-loop level [61, 62]. However, the inclusion of such predictions in a full pipeline for the analysis of the redshift-space bispectrum is yet to come, one main difficulty being the numerical expensive evaluation (see e.g. [63] for an effort in solving this problem).

In addition, relatively little work has been done on understanding how several methodological decisions impact the inference of model parameters based on measurements of the galaxy bispectrum.

5

Even assessing the goodness of fit of bispectrum models is impaired by the limited accuracy with which measurement errors and covariances are known. This unsatisfactory situation provides the first motivation for this study.

In this dissertation, I investigate the consequences of several assumptions that are routinely made or overlooked in the construction of a likelihood function for fitting a model to measurements of the galaxy bispectrum, both alone and in conjunction with the galaxy power spectrum. Most importantly, we will quantify, when possible, the systematic errors induced by such assumptions and related approximations, on the constraints on the model parameters. My collaborators and I use measurements of the power spectrum and the bispectrum of dark-matter halos extracted from nearly 300 cosmological N-body simulations (the Minerva set, first presented in [64]) covering a total volume of approximately $1000 \, h^{-3} \, \mathrm{Gpc}^3$. While such a large volume will never be observed, using a large number of simulations lowers significantly the sample variance, and therefore any analysis of this set-up really represents a fundamental test of both the theoretical and technical framework. Covariance matrices are estimated using an even larger set of 10,000 mock halo catalogs generated with the 3LPT PINOCCHIO code [65, 66, 67]. This allows to base our results on an unprecedented robust estimation of bispectrum errors.

We will focus our analysis on relatively large scales, and we therefore consider the simple tree-level perturbative prediction as the benchmark model for the halo bispectrum, while for the power spectrum we use a 1-loop expression, which is currently the main standard for analyses based on perturbation theory (e.g. [48, 16]). As we will see, the numerically inexpensive evaluation of the bispectrum model will allow a number of tests of the full pipeline, otherwise impossible with more sophisticated but demanding loop corrections. We will nevertheless show how this large-scale bispectrum data can add valuable information to the standard power spectrum analysis.

When fitting only the bispectrum, we compare the results obtained from adopting different likelihood functions that account for errors in the precision matrix [68, 69]. Furthermore, we study how the best-fit parameters are affected by the binning strategy adopted for the bispectrum measurements, and on how we account for such binning in the theoretical model.

We apply model-selection techniques to identify the optimal number of model parameters that are needed to describe the numerical data. In this case, we focus on simple extensions to our reference model, given by possible corrections to the shot-noise contributions, and on the possibility of reducing the parameter space by adopting fitting functions or theoretically-motivated relations among the bias parameters [70, 71].

We then investigate the joint analysis of the galaxy power spectrum and bispectrum, with the final target being the constraining of cosmological parameters. In particular, this is done in order to see how the addition of the bispectrum impacts the posteriors of both bias and cosmological parameters, compared to the case in which the power spectrum is fitted alone.

We also explore possible simple extensions of the models used in the joint analysis: we consider including contributions coming from higher-derivative bias or k-dependent stochasticity, but at the same time we assess whether the model, given the statistical uncertainties of the measurements, needs to include these parameters in order to correctly describe the data. However, even when performing a simultaneous fit of power spectrum and bispectrum, degeneracies between parameters still remain; for this reason, when trying to extract cosmological information from the joint analysis of power spectrum and bispectrum, it might be useful to consider phenomenological or theory-driven relations between

the parameters in order to reduce the dimensionality of the sampled parameter space [71, 59, 72].

Moreover, given a maximum Fourier mode, the bispectrum is described by a large number of triangle configurations; even if the goodness-of-fit statistics signal a failure of the model to correctly reproduce the data at a specific k_{max}, there might still be some triangle configurations at larger values of k that are still consistent with the model. We investigate the possibility of identifying and including these triangles into our analysis, in order to use them to get better constraints on the model parameters.

This dissertation is organised as follows. Chapter 2 will start with a basic introduction of cosmology, based on Einstein's theory of General Relativity, and with some definitions related to the concepts of random fields and correlation functions. A presentation of perturbation theory will follow, in order to attempt to provide a summary of the state of the art predictions I mentioned above. However, one should keep in mind that an introduction that follows the development of this theoretical framework, even considering exclusively the last few years, goes well beyond the main scope of this dissertation. It would be an impossible task to try and justify these models under every single aspect, since the description of these models often lacks of some adequate physical motivation. I then describe estimators commonly used to measure the correlation functions in Fourier space from surveys and simulations, I study the impact that these estimators have on a consistent theoretical modelling, and discuss the state of the art of applications of these models to actual data.

In chapter 3, I present the data sets used in the rest of the text: these are based on the set of almost 300 N-body simulations called Minerva, and on the set of 10,000 mock halo catalogs generated with the Lagrangian PT-based Pinocchio code. The mock catalogs are extremely useful to estimate a covariance matrix for our observables, that can be used in a full likelihood analysis to constrain nuisance and cosmological parameters. Moreover, these data sets provide us with independent measurements of some of the parameters we fit: this becomes crucial in a likelihood analysis of power spectrum and bispectrum in which cosmological parameters are fixed to the real values, since the recovery of unbiased values of the fitted parameters has to represent one of the criteria to assess the quality of the fit to the data that the model we consider can provide.

In chapter 4, I describe the analysis of the galaxy bispectrum alone, that has been published in [73]. I introduce all the tree-level models considered, the adopted likelihood functions, and the statistical tools employed to assess the goodness of the fit to the data and to select models based on their performances; I then present the results of the analyses, where the impact of different choices in the determination of the theoretical model and the covariance matrix is explored.

In chapter 5, I describe the current effort we are putting forward of a joint analysis of power spectrum and bispectrum: we study the impact that the addition of the bispectrum has on both the bias parameters and the cosmological parameters. This requires a fast and accurate code to compute loop corrections for the power spectrum and to correctly average power spectrum and bispectrum models onto bins defined over the same discrete Fourier grid used to perform the measurements.

In chapter 6, I discuss more projects to which I gave my contribution, related to the analysis of matter bispectrum measurements compared to different PT-based one-loop models, to the comparison between the "modal estimator" and the standard bispectrum estimator aiming at the compression of the information content of the bispectrum, and on the ongoing effort of modelling redshift-space correlation functions finalized to the joint analysis of power spectrum and bispectrum in redshift space.

Finally, in chapter 7 I present my conclusions and some future outlooks to be considered toward a full likelihood analysis of power spectrum and bispectrum measurements from galaxy redshift surveys.

Chapter 2

The Large Scale Structure of the Universe

In this chapter, I discuss the foundations upon which the fields of Cosmology and of the LSS of the Universe are based. First, I will briefly introduce Einstein's theory of General Relativity as the physical law describing gravitational interactions in the Universe and I will draw a link from General Relativity to Cosmology. I will then introduce random fields and correlation functions as basic tools in the description of inhomogeneous cosmology. A brief introduction to linear and non-linear perturbation theory of the matter overdensity field will follow. After that, I will discuss the connection between matter field and galaxy number density in terms of galaxy biasing, redshift space distortions, and the correct modelling of the BAO feature in the non-linear power spectrum. I will then describe estimators in Fourier space, and will conclude with a short review of the state of the art.

2.1 Fundamentals of Cosmology

Modern cosmology has its foundation on Einstein's theory of General Relativity (GR) [74], according to which given a metric $g_{\mu\nu}$ defined on a manifold $\mathcal{M}$ and the stress-energy-momentum tensor $T_{\mu\nu}$ describing the distribution of matter and energy, these two quantities can be related by Einstein's field equations:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu}, \tag{2.1}$$

where $R_{\mu\nu}$ is the so-called Ricci tensor, defined from the metric $g_{\mu\nu}$, $R = R_\mu^\mu$ is the Ricci scalar, G is Newton's gravitational constant and c is the speed of light in vacuum.

The physical interpretation of this equation is that the matter-energy distribution of the Universe induces a curvature of spacetime, and the motion of test particles is described by curved geodesic trajectories determined by the metric tensor.

GR has a large number of applications in the description of phenomena at astrophysical and cosmological scales, ranging from black holes, to gravitational waves, to the evolution of the Universe itself. A number of experiments and observations have proven GR to be valid on all regimes tested: classical tests include the observed deviation of light by massive bodies (e.g. [75]) and the explanation for the observed rate of the perihelion precession of Mercury [76], while most recent experiments involve the observation of gravitational waves emission from compact binary systems (e.g. [77]) and

the imaging of the shadow of the supermassive black hole at the center of the galaxy M87 ([78] and companion papers).

The main focus of this text, however, is going to be on the evolution and the Large Scale Structure of the Universe. In order to study the evolution of the Universe, we will work under the assumption of the so-called **cosmological principle**, according to which the Universe is homogeneous, meaning that it has the same average properties in every point in space, and isotropic, meaning that it has no preferred direction. This is very well consistent with observations of the Large Scale Structure of the Universe when observed over very large scales. Working under this assumption implies that the large-scale Universe has constant curvature, and the most general spacetime metric satisfying this properties is the Friedman-Robertson-Walker (FRW) metric [79, 80]. In spherical coordinates (t, r, θ, φ), this metric reads:

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = c^2 dt^2 - a^2(t) \left[\frac{dr^2}{1 - \kappa r^2} + r^2 \left(d\theta^2 + \sin^2 \theta \, d\varphi^2 \right) \right]; \tag{2.2}$$

here κ is a constant whose sign is related directly to the constant curvature of the Universe, with $\text{sign}(\kappa) = 0$ representing a flat Universe, while 1 and -1 describe an elliptic and a hyperbolic geometry respectively. The factor $a(t)$ is the scale factor, describing the expansion or contraction of the spatial part of the metric, and is defined in such a way that in present time $a(t) = 1$; r is also called the comoving radial distance, while t is defined as the cosmic time.

The physical (or proper) distance d can be defined by considering the motion along the radial coordinate r ($d\theta = d\varphi = 0$) of light ($ds^2 = 0$) emitted at time t_e:

$$d(t) = a(t) \int_{t_e}^{t} \frac{c \, dt'}{a(t')} \equiv a(t) \chi(t); \tag{2.3}$$

the quantity $\chi(t)$ implicitly defined here is the comoving distance. Because of the time-dependent scale factor $a(t)$, this means that the physical wavelength of light will increase due to the expansion of the Universe occurred between the time of emission t_e and the time of observation t_o. Radiation will thus be shifted toward the low-energy end of the electromagnetic spectrum by a factor z, called redshift, and defined by

$$1 + z = \frac{a(t_o)}{a(t_e)}. \tag{2.4}$$

In our convention where $a(t) = 1$ today, the expression for the redshift can be rewritten as $1 + z = a(t_e)^{-1}$, and therefore redshift can be used both as a measure of the time of emission and as a measure of distance from a source to us.

A "conformal time" τ may also be defined by means of the relation $dt^2 = a^2 d\tau^2$, in such a way that the metric at equation 2.2 can be written more conveniently as a conformally transformed metric,

$$ds^2 = a^2(\tau) \left[c^2 d\tau^2 - \frac{dr^2}{1 - \kappa r^2} - r^2 d\Omega^2 \right], \tag{2.5}$$

where the notation for the solid angle differential $d\Omega^2 = d\theta^2 + \sin^2 \theta \, d\varphi^2$ has been introduced. In the following, units with $c = 1$ will be assumed.

The explicit form of the scale factor $a(t)$ can be derived using equation 2.1 and from the conser-

vation of the stress-energy-momentum tensor (which in GR is a consequence of the Bianchi identity)

$$\nabla_\mu T^{\mu\nu} = 0. \tag{2.6}$$

In order to find a specific solution for the scale factor, it is therefore necessary to assume the energy density to be dominated by a single perfect fluid, a fluid characterised completely in terms of its density ρ and pressure p. For a perfect fluid, the stress-energy-momentum tensor in the rest frame of the fluid reads

$$T^\mu_\nu = \mathrm{diag}(\rho, -p, -p, -p). \tag{2.7}$$

In order to relate pressure and density, an equation of state can be introduced, $p = w\rho$, and in this specific case, if w is constant in time, the system of differential equations defined by the stress-energy-momentum conservation (equation 2.6) assuming the FRW metric (equation 2.2), can be solved. For a single species, the solution is

$$\rho(t) = \rho_0 a^{-3(1+w)}(t), \tag{2.8}$$

where ρ_0 is the present density of the species considered. In particular, the equation of state parameter can be related to the most common species observed in the Universe: $w = 0$ corresponds to a pressure-less fluid (e.g. dark matter, and approximately baryonic matter), $w = \frac{1}{3}$ to relativistic fluids (e.g. radiation), and $w = -1$ to a constant-density component. Specifically, this last case corresponds to a stress-energy-tensor

$$T_{\mu\nu} = \frac{\Lambda}{8\pi G}, \tag{2.9}$$

with constant Λ. This is directly related to the cosmological constant Einstein introduced in his field equations [81], and is currently associated to dark energy.

The dependence of the density to the scale factor, together with the fact that all the components mentioned above are of the same order of magnitude at present time, implies that the very-early Universe underwent a radiation-dominated epoch (RD); this epoch was then followed by a matter-dominated epoch (MD), and is currently in a dark energy-dominated epoch (DED).

Starting from equations 2.1, 2.2, 2.7, and 2.6, it is possible to derive the Friedmann equations [82, 83]:

$$H^2 = \frac{8\pi G}{3}\rho - \frac{\kappa}{a^2} \tag{2.10a}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p), \tag{2.10b}$$

where $H = \frac{\dot{a}}{a}$ is the Hubble parameter; for historical reasons, sometimes the Hubble parameter is written in terms of the adimensional Hubble parameter h, with $H_0 = 100\,h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. A critical density can also be defined,

$$\rho_{\mathrm{crit}} = \frac{3H^2}{8\pi G}, \tag{2.11}$$

corresponding to the total energy density in the case of a flat Universe. Using the critical density, we can define for each species j in the Universe a quantity representing the relative contribution of the species to the total matter-energy content of the Universe

$$\Omega_j(t) = \frac{\rho_j(t)}{\rho_{\mathrm{crit}}(t)}. \tag{2.12}$$

Assuming that the Universe is flat, CMB experiments and observations of the galaxy distribution suggest that our Universe today is characterized by a relative abundance of relativistic species (including radiation and relativistic neutrinos) $\Omega_{\gamma,0} \sim 10^{-4}$ and a relative abundance of baryonic matter $\Omega_{b,0} \sim 0.05$; this implies that 95% of the energy content of the Universe is in dark components, with cold dark matter having a relative abundance $\Omega_{c,0} \sim 0.25$ and dark energy with $\Omega_{\Lambda,0} \sim 0.7$. These contributions to the total energy-density of the Universe define the ΛCDM cosmological model. Notably, our Universe is currently undergoing a phase dominated by a mysterious dark-energy component, as shown by the observed accelerated expansion. Also notice that while it is required for a species to have $w < -1/3$ in order to produce an accelerated expansion of the Universe, the standard ΛCDM model assumes that the dark energy component is a species of constant energy density, and therefore that $w = -1$; possible extensions of the ΛCDM model assume as well a free, *i.e.* unknown, equation of state parameter for the dark energy component that in principle could also depend on time.

When a set of relative abundances of species in the Universe is assumed, the scale factor as a function of time for a generic flat ΛCDM Universe can be computed from

$$H_0 \, \mathrm{d}t = \frac{\mathrm{d}a}{\left[\Omega_{\gamma,0}/a^2 + (\Omega_{b,0} + \Omega_{c,0})/a + \Omega_{\Lambda,0} a^2 \right]^{1/2}}. \tag{2.13}$$

Clearly, observational constraints on the expansion of the Universe provide information on the energy content: Taylor-expanding this expression close to the present epoch, one can relate directly the distance of objects and their redshift with parameters that depend on the present abundance of dark matter and dark energy. Therefore, measuring at the same time distances and redshifts of relatively close objects can help in constraining cosmological parameters: for example, the second derivative of the scale factor is related to the deceleration parameter, that in a ΛCDM Universe can be written as

$$q_0 = \Omega_{\gamma,0} + \frac{1}{2}(\Omega_{b,0} + \Omega_{c,0}) - \Omega_{\Lambda,0}; \tag{2.14}$$

using Supernovae Ia as standard candles, it has been possible to constrain a negative q_0: finding a negative value of the deceleration parameter implies therefore an accelerated expansion of the Universe, and a non-zero dark energy component. This cosmographic measurement is what introduced dark energy in the standard cosmological model [3, 4].

However, other possible probes exist apart from cosmography, that can provide even tighter constraints on cosmological parameters, as proven for example by CMB experiments, that study the microwave background emission of photons from the very high-redshift Universe, during the stage of decoupling of matter and radiation ($z \sim 1100$). In particular, one main focus of CMB experiments is the measurement of temperature fluctuations δT in the cosmic microwave background radiation

$$\frac{\delta T}{\bar{T}}(\hat{\mathbf{n}}) = \frac{T(\hat{\mathbf{n}}) - \bar{T}}{\bar{T}}, \tag{2.15}$$

where $\hat{\mathbf{n}}$ is the line-of-sight direction and $\bar{T}$ is the average temperature of the CMB. Temperature fluctuations arise because of the interaction of CMB photons with the gravitational potentials produced by the matter distribution, mainly at early time, and therefore these temperature fluctuations can be related to the perturbations in the dark matter field. Since the temperature fluctuations are of order 10^{-5} with respect to the average temperature of the CMB, this suggests that the matter density field

at very early times was very close to being homogeneous and isotropic, and a linear perturbation theory scheme can be employed to model these fluctuations. It is common to study the fluctuations using the angular power spectrum as a summary statistics

$$C_\ell = \frac{1}{2\ell + 1} \sum_m |a_{\ell m}|^2, \tag{2.16}$$

where the coefficients $a_{\ell m}$ are the coefficients of the expansion of the temperature fluctuations in spherical harmonics,

$$\frac{\delta T}{T}(\hat{\mathbf{n}}) = \sum_{\ell, m} a_{\ell m} Y_\ell^m(\hat{\mathbf{n}}). \tag{2.17}$$

While the temperature power spectrum contains the bulk of the information of the CMB, adding also the polarization spectra can help in constraining cosmological parameters and extensions of the standard ΛCDM model, for example the tensor-to-scalar ratio of primordial fluctuations [2].

The Large Scale Structure of the Universe as a cosmological probe will be the main focus of this dissertation. Future galaxy redshift surveys, such as *Euclid*, will probe the Universe using high-density tracers of the large-scale structures, over volumes of considerable size [84]. The large number of galaxies that will be observed requires that the distribution of objects is treated statistically, using summary statistics as it is done, for example, in CMB experiments. Galaxies are assumed to form in high-density dark matter regions, and therefore statistical properties of the galaxy distribution can and do contain information on the underlying dark matter distribution. Moreover, dark energy has an effect on the growth rate of dark matter fluctuations. This makes large-scale structure a proficient probe of the Universe both at low and high redshift.

Thus, it is necessary, to be able to model both the statistical and the physical properties of the dark matter density field, of the galaxy density field, and also to establish a link between the properties of these two density fields. The rest of this chapter will be dedicated to this specific purpose.

2.2 Statistics of classical random fields

Most of the observables employed in the field of cosmology are functionals of the perturbations of the matter density field,

$$\delta(\mathbf{x}, t) = \frac{\rho_R(\mathbf{x}, t)}{\bar{\rho}(t)} - 1; \tag{2.18}$$

the quantity $\bar{\rho}(t)$ is the spatial average of the matter density, while $\rho_R(\mathbf{x}, t)$ here indicates that the matter density field $\rho(\mathbf{x}, t)$ has been smoothed on some large scale R,

$$\rho_R(\mathbf{x}, t) \equiv \int \mathrm{d}^3 y \, W_R(|\mathbf{x} - \mathbf{y}|) \rho(\mathbf{y}, t), \tag{2.19}$$

where $W_R(x)$ is a window function of characteristic scale R, whose purpose is just to smooth the matter density field. Of course one also works with the galaxy number density $n_g(\mathbf{x})$, which however is always related to the perturbations in the matter density field itself, and therefore also the galaxy number density field is just a functional of the matter density field. This field is a classical field, and while it has a definite value for every single point in space and time $(\mathbf{x}, t)$, we are not able to predict quantities at every position in space and at any point in time; instead, we are only able to predict average quantities. For this reason, it is necessary to work with summary statistics, that depend on

the density field, treated however as a random field.

Random fields are a generalization of random variables. Given a random variable ϕ following a probability density function (PDF) $\mathcal{P}(\phi)$, we can characterize the n-th order moment of ϕ as

$$\langle \phi^n \rangle = \int \mathrm{d}\phi \, \mathcal{P}(\phi) \phi^n; \tag{2.20}$$

for example, the 1st order moment $\langle \phi \rangle$ is the mean value. In general, any n-th order moment can be computed; for standard probability density functions, all or some of these moments can be computed explicitly. For example, given a normally distributed random variable with PDF

$$\mathcal{P}(\phi) = \frac{1}{\sqrt{2\pi\sigma_\phi^2}} \exp\left[-\frac{(\phi - \mu_\phi)^2}{2\sigma_\phi^2} \right] \tag{2.21}$$

we will have $\langle \phi \rangle = \mu_\phi$, $\langle \phi^2 \rangle = \sigma_\phi^2 + \mu_\phi^2$, and $\langle \phi^n \rangle = 0$ for any other value of $n > 2$.

The concept of n-th order moment can be generalized, in the case of random fields, into the broader notion of a correlation function: given a random field $\phi(\mathbf{x})$, the general n-point correlation function can be defined as

$$\langle \phi_1(\mathbf{x}_1) \ldots \phi_n(\mathbf{x}_n) \rangle \equiv \int \mathrm{d}\phi_1 \ldots \mathrm{d}\phi_n \, \mathcal{P}[\phi_1, \ldots, \phi_n] \phi_1 \ldots \phi_n, \tag{2.22}$$

where in the right-hand side the position dependence has been suppressed to have a more compact notation, $\phi_i = \phi_i(\mathbf{x}_i)$. Working under the assumptions of statistical homogeneity and isotropy – that in cosmology is a consequence of the cosmological principle – all n-point correlation functions depend only on the moduli of the distances between each and every pair of positions $\mathbf{x}_i$ and $\mathbf{x}_j$:

$$\langle \phi_1(\mathbf{x}_1) \ldots \phi_n(\mathbf{x}_n) \rangle = f(\{|\mathbf{x}_i - \mathbf{x}_j|, \, \forall i = 1, \ldots n, \, \forall j = i + 1, \ldots, n\}). \tag{2.23}$$

Moreover, a generic n-point correlation function can be written as the sum of a *connected* n-point correlation function plus the products of all possible combinations of lower-order correlation functions evaluated with all the cyclic permutations of the fields $\phi_i(\mathbf{x}_i)$:

$$\begin{aligned}
\langle \phi_1(\mathbf{x}_1) \ldots \phi_n(\mathbf{x}_n) \rangle = {} & \langle \phi_1(\mathbf{x}_1) \ldots \phi_n(\mathbf{x}_n) \rangle_c \\
& + \langle \phi_1(\mathbf{x}_1) \rangle \langle \phi_2(\mathbf{x}_2) \ldots \phi_n(\mathbf{x}_n) \rangle + \mathrm{cyc.} + \\
& + \langle \phi_1(\mathbf{x}_1) \phi_2(\mathbf{x}_2) \rangle \langle \phi_3(\mathbf{x}_3) \ldots \phi_n(\mathbf{x}_n) \rangle + \mathrm{cyc.} + \ldots \\
& + \langle \phi_1(\mathbf{x}_1) \rangle \ldots \langle \phi_n(\mathbf{x}_n) \rangle.
\end{aligned} \tag{2.24}$$

This is particularly useful in the context of cosmology, where one works with two main density fields – the already mentioned matter density field $\rho(\mathbf{x}, t)$ (or the related matter density contrast $\delta(\mathbf{x}, t)$), and the galaxy number density – and where n-point correlation functions are expressed as products of lower-order correlation functions.

We can only predict statistical properties of fields, which are meant to be considered as averages over the ensemble of all possible universes. However, our Universe represents only one particular realization of this ensemble, and therefore a direct comparison cannot be properly performed, because we would need many realizations of our Universe to be able to perform this comparison. For this reason,

we work under the assumption of the hergodic hypothesis, according to which ensamble averages are equal to spatial averages, provided that the sampled volume V is large enough,

$$\int \mathrm{d}\phi\, \mathcal{P}(\phi)\phi \equiv \langle \phi(\mathbf{x}) \rangle = \lim_{V \to \infty} \frac{1}{V} \int_V \mathrm{d}^3\mathbf{x}\, \phi(\mathbf{x}). \qquad (2.25)$$

While the dark matter density field dominates the matter content of the Universe, what we observe is the galaxy density field. At a given time t, it is possible to write the galaxy number density $n_g(\mathbf{x})$ in terms of the mean number density over all of space plus a perturbation $\delta_g(\mathbf{x})$,

$$n_g(\mathbf{x}) \equiv \bar{n}_g + \delta n_g(\mathbf{x}) = \bar{n}_g \left[1 + \delta_g(\mathbf{x}) \right]; \qquad (2.26)$$

under the hergodic hypothesis, the ensemble average of the number density field is equal to the spatial average, $\langle n_g(\mathbf{x}) \rangle = \bar{n}_g$, therefore it is straightforward to see that this implies $\langle \delta_g(\mathbf{x}) \rangle = 0$: thus, the first non-zero correlation function of the galaxy number density $n_g(\mathbf{x})$ is actually the two-point correlation function (2PCF); this can be written as

$$\langle n_g(\mathbf{x}_1) n_g(\mathbf{x}_2) \rangle = \bar{n}_g^2 \left[1 + \xi_g(|\mathbf{x}_1 - \mathbf{x}_2|) \right]; \qquad (2.27)$$

here, the function $\xi_g(|\mathbf{x}_1 - \mathbf{x}_2|)$ has been introduced, which is the connected part of the 2PCF of the galaxy overdensity $\delta_g(\mathbf{x})$

$$\xi_g(|\mathbf{x}_1 - \mathbf{x}_2|) = \langle \delta_g(\mathbf{x}_1) \delta_g(\mathbf{x}_2) \rangle_c; \qquad (2.28)$$

in a similar way, it is also possible to define a 3PCF of the galaxy overdensity, that due to the fact that $\langle \delta_g(\mathbf{x}) \rangle = 0$, is just equal to the connected part of the full 3PCF

$$\zeta_g(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \langle \delta_g(\mathbf{x}_1) \delta_g(\mathbf{x}_2) \delta_g(\mathbf{x}_3) \rangle_c. \qquad (2.29)$$

The 4PCF, however, will contain also 2-point contributions, as well as the connected part of the 4PCF

$$\langle \delta_g(\mathbf{x}_1) \delta_g(\mathbf{x}_2) \delta_g(\mathbf{x}_3) \delta_g(\mathbf{x}_4) \rangle = \langle \delta_g(\mathbf{x}_1) \delta_g(\mathbf{x}_2) \delta_g(\mathbf{x}_3) \delta_g(\mathbf{x}_4) \rangle_c +$$
$$+ \langle \delta_g(\mathbf{x}_1) \delta_g(\mathbf{x}_2) \rangle \langle \delta_g(\mathbf{x}_3) \delta_g(\mathbf{x}_4) \rangle + 2\ \mathrm{cyc}. \qquad (2.30)$$

Since we work with matter density perturbations defined through a smoothed density field, on large enough scales these perturbations are going to be small. Then, a natural way of separating fluctuations on different scales is working in Fourier space, where the density field can be written as a linear combination of components of characteristic wavenumbers $\mathbf{k}$. The Fourier transform definitions that will be adopted in the rest of the text are defined as follows:

$$\phi_{\mathbf{k}} = \int \frac{\mathrm{d}^3\mathbf{x}}{(2\pi)^3}\, \mathrm{e}^{-i\mathbf{k}\cdot\mathbf{x}}\, \phi(\mathbf{x})$$
$$\phi(\mathbf{x}) = \int \mathrm{d}^3\mathbf{k}\, \mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}}\, \phi_{\mathbf{k}}; \qquad (2.31)$$

I will also adopt this definition for the integral representation of the Dirac delta function

$$\delta_D(\mathbf{x}) = \frac{1}{(2\pi)^3} \int \mathrm{d}^3\mathbf{k}\, \mathrm{e}^{-i\mathbf{k}\cdot\mathbf{x}}. \qquad (2.32)$$

In cosmology, one usually deals with real random fields, since both $\rho(\mathbf{x}, t)$ and $n_g(\mathbf{x}, t)$ are clearly real-valued fields; this means that for this fields we can use the reality condition, according to which given a random field ϕ, if $\phi(\mathbf{x}) \in \mathbb{R}$, then $\phi_{\mathbf{k}}^* = \phi_{-\mathbf{k}}$, and the field is random both in configuration space and in Fourier space.

Consider now the matter overdensity field $\delta(\mathbf{x})$ and its Fourier-transformed $\delta_{\mathbf{k}}$; the Fourier-transformed field has zero mean, $\langle \delta_{\mathbf{k}} \rangle = 0$; moreover, the 2PCF of $\delta_{\mathbf{k}}$ can be written as:

$$\langle \delta_{\mathbf{k}_1} \delta_{\mathbf{k}_2} \rangle = \delta_D(\mathbf{k}_1 + \mathbf{k}_2) P(k_1) \tag{2.33}$$

with $P(k_1) = \mathcal{FT}\left[\xi_\delta(|\mathbf{x}|)\right]$; the quantity $P(k)$ is the so-called **power spectrum** of the matter density fluctuations. In a similar way, it is possible to define a **bispectrum** of the matter density field,

$$\langle \delta_{\mathbf{k}_1} \delta_{\mathbf{k}_2} \delta_{\mathbf{k}_3} \rangle = \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3), \tag{2.34}$$

that is the Fourier counterpart of the 3PCF of the matter overdensity field in configuration space, $\delta(\mathbf{x})$.

A specific case is represented by Gaussian random fields, defined as fields that have vanishing connected n-point correlation functions with $n > 2$. In cosmology, simple inflationary models predict that the initial overdensity is a Gaussian random field, and results from CMB experiments are consistent with an initially Gaussian density field [85]. The fact that all n-point connected correlation functions of Gaussian fields vanish for $n > 2$, implies that the statistical properties of a Gaussian random field are completely characterized by its mean and its 2PCF (or power spectrum). As an example, for the matter overdensity field, its PDF can be written as

$$\mathcal{P}[\delta(\mathbf{x})] = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\frac{\delta^2(\mathbf{x})}{\sigma^2}\right], \tag{2.35}$$

and the multivariate PDF can therefore be written as

$$\mathcal{P}[\delta(\mathbf{x}_1), \ldots, \delta(\mathbf{x}_n)] = \frac{1}{\sqrt{(2\pi)^n \det(\mathbb{C})}} \exp\left[-\frac{1}{2}\delta_i C_{ij}^{-1} \delta_j\right], \tag{2.36}$$

where $\delta_i \equiv \delta(\mathbf{x}_i)$ for brevity, and where the matrix $\mathbb{C}$ of components C_{ij} is the 2PCF,

$$C_{ij} \equiv \langle \delta(\mathbf{x}_i)\delta(\mathbf{x}_j) \rangle; \tag{2.37}$$

in Fourier space, this becomes

$$\mathcal{P}[\delta_{\mathbf{k}}] = \frac{1}{\sqrt{2\pi P(k)}} \exp\left[-\frac{1}{2}\frac{|\delta_{\mathbf{k}}|^2}{P(k)}\right]. \tag{2.38}$$

For random Gaussian fields, therefore, all higher-order statistics – bispectrum, trispectrum, and so on – are identically zero. However, this is not true at all times, since non-linear gravitational evolution can introduce non-Gaussianities in initially Gaussian fields. This means that, in presence of gravitational instability, it is possible to measure higher-order correlation functions that are different from zero. For this exact reason, higher-order correlation functions are more effective in studying the nature of the process causing the gravitational collapse. Moreover, they can give further information on the

three-dimensional distribution of the large-scale structure of the Universe. Other effects contribute in introducing non-linearities in the density fields, although in the galaxy density fields, such as non-linear galaxy biasing and redshift-space distortions. Therefore, in the observed galaxy density field it is expected to observe also higher-order statistics that are different from zero, since all these other non-linear effects introduce as well non-Gaussianities in the fields.

2.3 Equations of motion and linear perturbation theory

As we have seen, starting from overdensity fields it is straightforward to define correlation functions of these fields; we have also seen that if we can smooth the density field over a large-scale mode $k \sim 1/R$, the matter overdensity field will actually be small. It seems natural, then, to look for theoretical predictions for the correlation functions following a perturbative scheme in the overdensity field δ. In order to do so, however, it is crucial to be able to describe the evolution of the overdensity field both in space and time. Once this is done, a perturbation theory of the matter overdensity field can be fully developed. This perturbative approach will go under the name of Standard Perturbation Theory, and while I will discuss the main point of this theory, a more detailed review of this topic can be found in [86].

To start, a few assumptions are in order. We will work assuming the Newtonian limit as a good approximation to the full general-relativistic treatment. This is valid as long as we consider scales much smaller than the Hubble horizon, and as long as peculiar velocities are much smaller than the speed of light. The total velocity of a galaxy is in fact given by the sum of two contributions,

$$v(t) \equiv \frac{\mathrm{d}}{\mathrm{d}t}d(t) = H(t)d(t) + u(t),\tag{2.39}$$

where the first term $H(t)d(t)$ represents the Hubble flow, the speed of an object due to the expansion of the Universe, while the second term $u(t) \equiv a(t)\dot{\chi}(t)$ is the aforementioned peculiar velocity of the object, that depends on the motion of large-scale matter perturbations (bulk flow) as well as on a velocity dispersion component for objects within collapsed dark-matter halos.

In this introduction, for simplicity, we work under the approximation of an Einstein-deSitter Universe, meaning a Universe that is flat ($\kappa = 0$) and matter dominated ($\Omega_m = 1$). This approximation in particular is very good from $z \sim 500$, down to $z \sim 1$, that is during matter domination, and therefore it is particularly relevant at large redshift. Moreover, we consider the single-stream approximation, according to which we assign to the velocity density field a unique value at each point in space and time. We expect this to be a valid description at large scales, where perturbations move coherently, while we expect it to break down at small scales, due to the collapse and virialization of matter over-densities into dark-matter halos, where the so-called shell-crossing takes place. The final assumption that we will make is that the entirety of matter is made of cold dark matter particles.

The equations of motion that we will consider are the continuity equation, Euler equation, expressing respectively conservation of matter and momentum, and Poisson's equation, relating matter density perturbations in the gravitational potential,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot [\rho \mathbf{v}] = 0 \tag{2.40a}$$

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} = -\frac{\nabla p}{\rho} - \nabla \Phi \tag{2.40b}$$

$$\nabla^2 \Phi = 4\pi G \rho. \tag{2.40c}$$

Since we are interested in developing a perturbation theory, it is useful to rewrite these equations in terms of the perturbations of the density field (meaning the matter overdensity), of the velocity field (meaning the peculiar velocity), and of the gravitational potential,

$$\rho(\mathbf{x}, t) = \bar{\rho}(t) \left[1 + \delta(\mathbf{x}, t) \right] \tag{2.41}$$

$$\mathbf{v}(\mathbf{x}, t) = H(t)\mathbf{d} + \mathbf{u}(\mathbf{x}, t) \tag{2.42}$$

$$\Phi(\mathbf{x}, t) = \frac{2\pi G}{3} \bar{\rho} d^2 + \phi(\mathbf{x}, t). \tag{2.43}$$

Working with conformal time instead of proper time, and defining the velocity divergence field $\theta(\mathbf{x}, t) = \nabla \cdot \mathbf{u}$, the system of equations can be linearized as

$$\begin{cases} \dfrac{\partial \delta}{\partial \tau} + \theta = 0 \\ \dfrac{\partial \theta}{\partial \tau} + \mathcal{H}\theta + \dfrac{3}{2}\mathcal{H}^2 \Omega_m \delta = 0 \end{cases}, \tag{2.44}$$

where $\mathcal{H}(\tau) = \frac{1}{a}\frac{\mathrm{d}a}{\mathrm{d}\tau} = a(\tau)H(\tau)$. Here δ and θ are random fields, function of the coordinates $\mathbf{x}$ of configuration space; in this case, it is convenient to work instead in Fourier space, with the Fourier-transformed fields $\delta_{\mathbf{k}}$ and $\theta_{\mathbf{k}}$; in particular, this allows to rewrite the linearized system of equations only in the variable $\delta_{\mathbf{k}}$:

$$\frac{\partial^2 \delta_{\mathbf{k}}}{\partial \tau^2} + \mathcal{H}\frac{\partial \delta_{\mathbf{k}}}{\partial \tau} - \frac{3}{2}\mathcal{H}^2 \Omega_m \delta_{\mathbf{k}} = 0. \tag{2.45}$$

It is possible to write an *ansatz* for the general solution of these equations, as the linear combination of a growing mode and a decaying mode, with growth and decay factors $D_\pm(\tau)$:

$$\delta_{\mathbf{k}} = D_+(\tau)A_{\mathbf{k}} + D_-(\tau)B_{\mathbf{k}} \tag{2.46}$$

$$\theta_{\mathbf{k}} = D_+'(\tau)A_{\mathbf{k}} + D_-'(\tau)B_{\mathbf{k}}, \tag{2.47}$$

with a prime symbol indicating derivation with respect to conformal time. In an Einstein-deSitter Universe one finds that the growth and decay factors are $D_+(a) \propto a$ and $D_-(a) \propto a^{-3/2} \propto H(a)$ respectively. Thus, the decaying mode goes rapidly to zero, while the growing mode quickly becomes the dominant contribution of the full solution. For this reason, in the following, the decaying mode will be neglected. Recall that the form of the growth factor that we have found is valid only in an Einstein-deSitter Universe; in the more general case of a flat ΛCDM cosmology, the general structure of the solution is the same, with the decaying mode again going very rapidly to zero; however the growth factor has the integral solution

$$D_+(a) = \frac{5}{2}H_0^2 \Omega_{m,0} \mathcal{H}(a) \int_0^a \frac{\mathrm{d}\tilde{a}}{[\tilde{a}H(\tilde{a})]^3}. \tag{2.48}$$

An important related quantity is the growth rate of structure f, defined as

$$f \equiv \frac{\partial \ln D}{\partial \ln a};$$

(2.49)

for a flat ΛCDM Universe, it is possible to parametrize f in terms of the relative abundance of total matter, $\Omega_m = \Omega_b + \Omega_c$, and in that case one finds $f \simeq \Omega_m^\gamma(a)$ with $\gamma = 6/11 \simeq 0.545\ldots$; for an Einstein-deSitter Universe, the exponent changes slightly, $\gamma = 4/7 \simeq 0.57$; in general, this exponent γ is even different in modified gravity theories, therefore measuring f can be seen either as directly measuring the matter content of the Universe (under the assumption that the Universe can be really described by a flat ΛCDM model), or as probing the validity of General Relativity (if an independent measure of $\Omega_m(a)$ is possible). In fact, measuring f at different ages of the Universe provides a dynamical test of the accelerated expansion of the Universe, to contrast the geometrical one provided by the BAOs.

Neglecting the decaying mode of the linear solution of the matter overdensity field, it is then possible to compute the power spectrum at the linear level as the two-point correlation function of $\delta_{\mathbf{k}}(\tau)$ at a given time τ, and this can be explicitly related to the power spectrum $P_{\mathrm{in}}(k)$ of the initial perturbations $\delta_{\mathbf{k}}^{\mathrm{in}}$:

$$\delta_D(\mathbf{k}_{12})P(k_1,\tau) = \langle \delta_{\mathbf{k}_1}(\tau)\delta_{\mathbf{k}_2}(\tau)\rangle = D^2(\tau)\langle \delta_{\mathbf{k}_1}^{\mathrm{in}}\delta_{\mathbf{k}_2}^{\mathrm{in}}\rangle = \delta_D(\mathbf{k}_{12})D^2(\tau)P_{\mathrm{in}}(k_1).$$

(2.50)

Therefore, at the linear level, the power spectrum at a given time is simply equal to the initial linear power spectrum rescaled by the square of the growth factor, $P_{\mathrm{L}}(k,\tau) = D^2(\tau)P_{\mathrm{in}}(k)$. However, under the assumption that the initial density field is Gaussian, all higher-order statistics, including the bispectrum, are identically zero at the linear level.

In the large-scale structure of the Universe, the cosmological information is mainly contained into the linear power spectrum, since a unique set of cosmological parameters correspond to a unique linear power spectrum. In the standard ΛCDM model, it can be written as the product of a primordial power spectrum $P_0(k)$ and the square of a function $T(k)$:

$$P_{\mathrm{L}}(k) = P_0(k)T^2(k);$$

(2.51)

the primordial power spectrum of scalar perturbations is usually assumed to be a power-law in k with tilt given by n_s and amplitude given by A_s

$$P_0(k) = A_s k^{n_s},$$

(2.52)

and the particular case with $n_s = 1$ corresponds to a scale-invariant power spectrum; this is not exactly true in our Universe, where n_s has been measured by a number of experiments: the Planck Collaboration has constrained the scalar tilt to be $n_s = 0.9652 \pm 0.0042$ [2], and thus to be completely inconsistent with 1. The function $T(k)$ is the so-called transfer function, that describes the suppression of the small scales fluctuations during the radiation dominated era, relative to the evolution of the large scale modes that enter the horizon during matter domination. In the very large scale limit, $T(k)$ approaches 1, and therefore for $k \to 0$ the linear power spectrum of matter has the same k-dependence as the primordial power spectrum. The transfer function depends on the cosmological parameters in a non-trivial way. In order to compute it accurately, Boltzmann solvers like `CAMB` [87] or `CLASS` [88] are
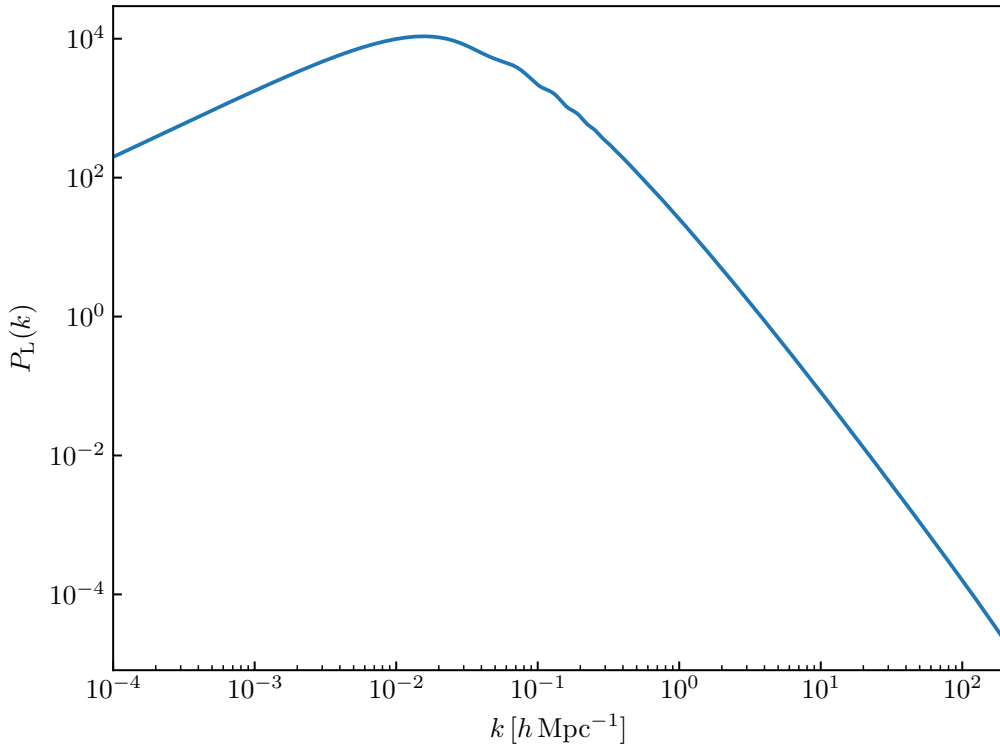
Figure 2.1: An example of a linear power spectrum computed with the use of the Boltzmann solver code `CAMB`.

commonly used. Under the assumption that the Universe can be described by the flat ΛCDM model, a set of six independent cosmological parameters can be used to completely characterize the Universe through a uniquely determined linear power spectrum of density perturbations.

2.4 Nonlinear perturbation theory

Now that we have a linear solution of the equations of motion of the matter overdensity field, we can try and push the perturbation theory scheme further. In particular, we assume that a fully non-linear solution to the equations of motion exists and that the non-linear matter overdensity field and the velocity-divergence field can be written as

$$\delta_{\mathbf{k}} = \delta_{\mathbf{k}}^{(1)} + \delta_{\mathbf{k}}^{(2)} + \delta_{\mathbf{k}}^{(3)} + \cdots = \sum_n \delta_{\mathbf{k}}^{(n)} \tag{2.53}$$

$$\theta_{\mathbf{k}} = \theta_{\mathbf{k}}^{(1)} + \theta_{\mathbf{k}}^{(2)} + \theta_{\mathbf{k}}^{(3)} + \cdots = \sum_n \theta_{\mathbf{k}}^{(n)}, \tag{2.54}$$

where the linear solution is itself an exact solution of the linearized equations of motion, and thus $\delta_{\mathbf{k}}^{(1)} = -\mathcal{H}\theta_{\mathbf{k}}^{(1)} = \delta_{\mathbf{k}}^{\mathrm{L}}$. In order for the perturbative expansion to be well-behaved, we will have to require that each perturbative order introduces increasingly "small" corrections to the previous order, otherwise the full non-linear solution would not converge to an actual solution of the equations of motion.

If we introduce the doublet notation for the fields,

$$\Psi_{\mathbf{k}} = \begin{pmatrix} \delta_{\mathbf{k}} \\ \Theta_{\mathbf{k}} \end{pmatrix}, \qquad \Psi_{\mathbf{k}} = \sum_n \Psi_{\mathbf{k}}^{(n)} \tag{2.55}$$

20

where $\Theta_{\mathbf{k}} = -\mathcal{H}\theta_{\mathbf{k}}$, the non-linear version of the equations of motion can be rewritten as

$$\partial_\eta \Psi_{\mathbf{k}}^a + \Omega^a{}_b \Psi_{\mathbf{k}}^b = \int \mathrm{d}^3\mathbf{k}_1 \, \mathrm{d}^3\mathbf{k}_2 \, \gamma^a{}_{bc}(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2) \Psi_{\mathbf{k}_1}^b \Psi_{\mathbf{k}_2}^c, \tag{2.56}$$

where the objects $\Omega^a{}_b$ and $\gamma^a{}_{bc}$ have been introduced, with

$$\Omega^a{}_b = \begin{pmatrix} 0 & -1 \\ -\frac{3}{2} & \frac{1}{2} \end{pmatrix}, \tag{2.57}$$

$$\gamma^1{}_{12}(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2) = \alpha(\mathbf{k}_1, \mathbf{k}_2)\delta_D(\mathbf{k} - \mathbf{k}_{12}) = \frac{\mathbf{k}_{12} \cdot \mathbf{k}_2}{k_2^2}\delta_D(\mathbf{k} - \mathbf{k}_{12}) \tag{2.58}$$

$$\gamma^2{}_{22}(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2) = \beta(\mathbf{k}_1, \mathbf{k}_2)\delta_D(\mathbf{k} - \mathbf{k}_{12}) = \frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)k_{12}^2}{2k_1^2 k_2^2}\delta_D(\mathbf{k} - \mathbf{k}_{12}) \tag{2.59}$$

$$\gamma^a{}_{bc}(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2) = 0 \qquad \text{otherwise.} \tag{2.60}$$

Using this notation, the linear solution to the equations of motion can be written simply as

$$\Psi_{\mathbf{k}}^{(1)} = \delta_{\mathbf{k}}^{\mathrm{L}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}; \tag{2.61}$$

with all of this, it is now possible to derive the higher-order solutions, meaning the successive higher-order contributions of the perturbation theory, that in general assume the form

$$\Psi_{\mathbf{k}}^{(n)} = \int \mathrm{d}^3\mathbf{k}_1 \ldots \mathrm{d}^3\mathbf{k}_n \, \delta_D(\mathbf{k} - \mathbf{k}_{1\ldots n}) \begin{pmatrix} F_n(\mathbf{k}_1, \ldots, \mathbf{k}_n) \\ G_n(\mathbf{k}_1, \ldots, \mathbf{k}_n) \end{pmatrix} \delta_{\mathbf{k}_1}^{\mathrm{L}} \ldots \delta_{\mathbf{k}_n}^{\mathrm{L}}. \tag{2.62}$$

Here, the functions F_n and G_n are fully symmetrized kernels of the $(\mathbf{k}_1, \ldots, \mathbf{k}_n)$ generated by appropriate combinations of the $\alpha(\mathbf{k}_i, \mathbf{k}_j)$ and $\beta(\mathbf{k}_i, \mathbf{k}_j)$, and these kernels derive from the equations of motion themselves. With this, it is possible to give an expression for the higher-order perturbative terms of the fully non-linear fields, as a function of the initial linear fields, where the non-linearities are introduced by the non-linear gravitational evolution of matter given by the equations of motion. Also worth noticing is the fact that each n-th perturbative term is of order $\mathcal{O}(\delta_{\mathrm{L}}^n)$, thus it is required that δ_{L} is small in order for the solution to converge.

In particular, the second-order kernels $F_2(\mathbf{k}_1, \mathbf{k}_2)$ and $G_2(\mathbf{k}_1, \mathbf{k}_2)$ in an Einstein-deSitter cosmology can be written explicitly as

$$F_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{1}{2}(\hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2)\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) + \frac{2}{7}(\hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2)^2 \tag{2.63}$$

$$G_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{3}{7} + \frac{1}{2}(\hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2)\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) + \frac{4}{7}(\hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2)^2; \tag{2.64}$$

in the more general ΛCDM model, the second order kernels have similar k-dependencies, however the numerical coefficients change in a non-trivial way, and acquire a time dependence. Higher-order kernels can be written iteratively starting from the second-order ones [86, 89].

Standard Perturbation Theory assumes that the fields doublet Ψ can be written as the sum of the perturbative contributions in equation 2.62 at all orders. This perturbative scheme allows also to
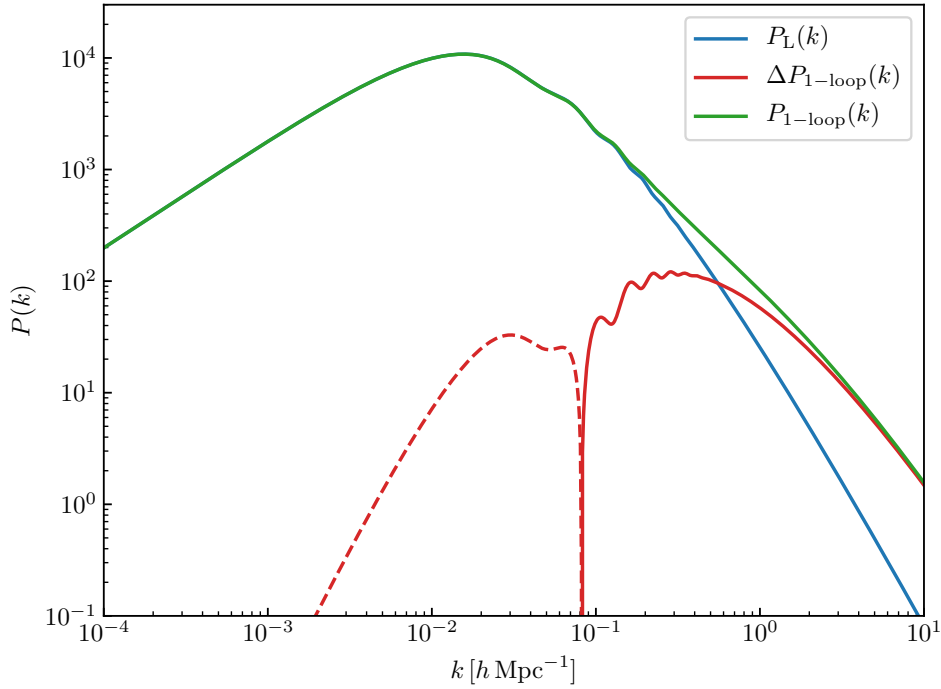
Figure 2.2: Comparison of the linear power spectrum (solid blue line) with the non-linear power spectrum at one-loop level (solid green line) and with the loop correction (in red, the dashed line indicating a negative contribution).

write non-linear, perturbative corrections to the power spectrum, as well as predictions for higher-order statistics arising from the non-Gaussian corrections due to the non-linear evolution of the fields.

For example, with this formalism, one can write the next-to-leading order contributions of the matter power spectrum:

$$\langle \delta_{\mathbf{k_1}} \delta_{\mathbf{k_2}} \rangle = \langle (\delta_{\mathbf{k_1}}^{(1)} + \delta_{\mathbf{k_1}}^{(2)} + \delta_{\mathbf{k_1}}^{(3)} + \dots)(\delta_{\mathbf{k_2}}^{(1)} + \delta_{\mathbf{k_2}}^{(2)} + \delta_{\mathbf{k_2}}^{(3)} + \dots) \rangle = \tag{2.65}$$

$$= \langle \delta_{\mathbf{k_1}}^{(1)} \delta_{\mathbf{k_2}}^{(1)} \rangle + \langle \delta_{\mathbf{k_1}}^{(2)} \delta_{\mathbf{k_2}}^{(2)} \rangle + \langle \delta_{\mathbf{k_1}}^{(1)} \delta_{\mathbf{k_2}}^{(3)} \rangle + \langle \delta_{\mathbf{k_1}}^{(3)} \delta_{\mathbf{k_2}}^{(1)} \rangle + \dots = \tag{2.66}$$

$$= \delta_D(\mathbf{k}_{12}) \left(P_{\mathrm{L}}(k_1) + P_{22}(k_1) + P_{13}(k_1) + \dots \right), \tag{2.67}$$

where the terms with odd powers of the linear density field $\delta_{\mathbf{k}}^{\mathrm{L}}$ are zero because of Gaussian initial conditions, and where the terms

$$P_{22}(k) = 2 \int \mathrm{d}^3 \mathbf{q} \, F_2(\mathbf{q}, \mathbf{k} - \mathbf{q})^2 P_{\mathrm{L}}(q) P_{\mathrm{L}}(|\mathbf{k} - \mathbf{q}|) \tag{2.68}$$

$$P_{13}(k) = 6 P_{\mathrm{L}}(k) \int \mathrm{d}^3 \mathbf{q} \, F_3(\mathbf{k}, \mathbf{q}, -\mathbf{q}) P_{\mathrm{L}}(q) \tag{2.69}$$

constitute the so-called one-loop corrections to the matter power spectrum in standard perturbation theory, $\Delta P_{1-\mathrm{loop}}(k) = P_{22}(k) + P_{13}(k)$. A comparison between the linear power spectrum, the one-loop correction and the non-linear power spectrum at one-loop level can be seen in figure 2.2.

In general, one can define any N-loop correction of the power spectrum in SPT, each of order $\mathcal{O}(\delta_{\mathrm{L}}^{2N+2})$, but under the assumption that SPT is well-defined and converges to the real non-linear power spectrum[1], they are all perturbatively negligible at large scales. For example, one could also

---

[1]Which is not necessarily true at small scales.

define two-loop corrections to the power spectrum,

$$\Delta P_{2\text{-loop}}(k) = P_{15}(k) + P_{24}(k) + P_{33}(k), \tag{2.70}$$

but this will be order $\mathcal{O}(\delta_{\mathrm{L}}^6)$, and therefore negligible with respect to the linear and one-loop terms at large enough scales.

Now that we can write the fully non-linear solution of the density field, we can really see the impact of gravitational instability on the Gaussianity of the fields. Earlier we said that Gaussian fields have all n-point statistics with $n > 2$ equal to zero. However, using the second-order solution of the density field, the leading order contribution (tree-level) to the matter bispectrum can be computed:

$$\delta_D(\mathbf{k}_{123}) B_{\mathrm{TL}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \langle \delta_{\mathbf{k}_1} \delta_{\mathbf{k}_2} \delta_{\mathbf{k}_3} \rangle =$$
$$= \langle \delta_{\mathbf{k}_1}^{(1)} \delta_{\mathbf{k}_2}^{(1)} \delta_{\mathbf{k}_3}^{(1)} \rangle + \langle \delta_{\mathbf{k}_1}^{(2)} \delta_{\mathbf{k}_2}^{(1)} \delta_{\mathbf{k}_3}^{(1)} \rangle + \langle \delta_{\mathbf{k}_1}^{(1)} \delta_{\mathbf{k}_2}^{(2)} \delta_{\mathbf{k}_3}^{(1)} \rangle + \langle \delta_{\mathbf{k}_1}^{(1)} \delta_{\mathbf{k}_2}^{(1)} \delta_{\mathbf{k}_3}^{(2)} \rangle; \tag{2.71}$$

the first term is zero because the connected part vanishes for Gaussian initial conditions, and because $\langle \delta_{\mathbf{k}} \rangle = 0$; the other terms can be grouped into

$$B_{\mathrm{TL}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 2 F_2(\mathbf{k}_1, \mathbf{k}_2) P_{\mathrm{L}}(k_1) P_{\mathrm{L}}(k_2) + 2 \text{ cyc.}, \tag{2.72}$$

which is evidently non-zero. This represents the matter bispectrum induced by gravity at tree-level, meaning at leading order in perturbation theory, and shows explicitly how gravitational instabilities introduce non-Gaussianities in initially Gaussian fields. Similarly to the the power spectrum, loop corrections for the bispectrum can be computed as well. For example, the one-loop contributions of the matter bispectrum can be written schematically as

$$\Delta B_{1\text{-loop}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = B_{114}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + B_{222}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + B_{123}^{(I)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + B_{123}^{(II)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3), \tag{2.73}$$

where each of these terms is given by the following equations:

$$B_{114}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 12 P_{\mathrm{L}}(k_1) P_{\mathrm{L}}(k_2) \int \mathrm{d}^3 \mathbf{q}\, F_4(\mathbf{q}, -\mathbf{q}, -\mathbf{k}_1, -\mathbf{k}_2) P_{\mathrm{L}}(q) + \text{cyc.} \tag{2.74}$$

$$B_{222}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 8 \int \mathrm{d}^3 \mathbf{q}\, F_2(\mathbf{q} - \mathbf{k}_2, -\mathbf{k}_3 - \mathbf{q}) F_2(\mathbf{k}_2 - \mathbf{q}, \mathbf{q}) F_2(\mathbf{k}_3 + \mathbf{q}, -\mathbf{q}) \times$$
$$\times P_{\mathrm{L}}(|\mathbf{q} - \mathbf{k}_2|) P_{\mathrm{L}}(|\mathbf{k}_3 + \mathbf{q}|) P_{\mathrm{L}}(q) \tag{2.75}$$

$$B_{123}^{(I)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 6 F_2(\mathbf{k}_1, \mathbf{k}_3) P_{\mathrm{L}}(k_1) P_{\mathrm{L}}(k_3) \int \mathrm{d}^3 \mathbf{q}\, F_3(\mathbf{k}_3, \mathbf{q}, -\mathbf{q}) P_{\mathrm{L}}(q) + \text{cyc.} \tag{2.76}$$

$$B_{123}^{(II)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 6 P_{\mathrm{L}}(k_1) \int \mathrm{d}^3 \mathbf{q}\, F_2(\mathbf{k}_2 + \mathbf{q}, \mathbf{q}) F_3(-\mathbf{k}_2 - \mathbf{q}, \mathbf{q}, -\mathbf{k}_1) P_{\mathrm{L}}(q) P_{\mathrm{L}}(|\mathbf{k}_2 + \mathbf{q}|) + \text{cyc.} \tag{2.77}$$

2.4.1 Effective field theory of Large Scale Structure (EFTofLSS)

Up until now, we have derived perturbative solutions to the equations of motion that assume that the matter fluid is exactly a pressureless perfect fluid. In reality, however, this is not the case: dark matter can undergo shell-crossing at small-scales, and therefore the single-stream approximation would be violated; moreover, one should consider the full Vlasov equation, that in the equations of motion introduces an extra anisotropic stress tensor, which is usually neglected; finally, even the pressure of the baryon fluid cannot be neglected at small scales. All of this introduces a series of additional terms

in the equations of motion (specifically in Euler equation). The Effective Field Theory of the Large Scale Structure (EFTofLSS) accounts for such additional terms and includes them in a consistent perturbative treatment by considering their dependence on long wavelength modes while averaging on short-scale perturbations [43]. This introduces additional terms with respect to the standard perturbation theory approach, that can be seen similar to higher-derivative operators (see below) multiplied by coefficients that are not known *a priori*.

Considering a third-order expansion of the density field, the prediction for the non-linear matter power spectrum at one-loop level requires an additional term,

$$P_{1-\text{loop}}^{\text{EFT}}(k) = P_{\text{L}}(k) + \Delta P_{1-\text{loop}}(k) + P_{\text{ct}}(k), \tag{2.78}$$

where $P_{\text{ct}}(k)$ is the so-called "counterterm", that at this order in perturbation theory takes the form

$$P_{\text{ct}}(k) = -2c_s^2 k^2 P_{\text{L}}(k). \tag{2.79}$$

The parameter c_s^2 can be regarded as an effective sound speed. Since this parameter, expected to be of order $\mathcal{O}(1)$ in units of $h^{-2}\,\text{Mpc}^2$, is not known *a priori* and cannot either be determined from cosmological parameters, it has to be fitted to simulations or considered as a free parameter in the analysis of observative data-sets.

Figure 2.3 shows the matter power spectrum from the Minerva set of N-body simulations (see Section 3.1) normalized by the linear power spectrum, compared to the one-loop power spectrum in SPT and the one-loop power spectrum in EFTofLSS (predictions shown in the plot account for infrared resummation, see section 2.7). The error bars represent the standard error on the mean of the full set of simulations, while the shaded area represents the standard deviation. The counterterm amplitude has been fitted assuming a $k_{\text{max}} = 0.15\,h\,\text{Mpc}^{-1}$, and in this case is equal to $c_s^2 \simeq 0.65\,h^{-2}\,\text{Mpc}^2$. It is apparent how the SPT approach introduces differences at the percent level even at large scales, while the EFT counterterm fitted up to $k_{\text{max}} = 0.15\,h\,\text{Mpc}^{-1}$ reduces the differences between the datapoints and the EFT one-loop power spectrum even at scales smaller than the ones of the fit.

A similar approach can be applied also to other correlation functions; the bispectrum acquires counterterms as well, all modulated by $\mathcal{O}(1)$ coefficients (when expressed in units of $h^{-2}\,\text{Mpc}^2$) that have to be fitted to simulations.

2.5 Non-linearities from tracers: galaxy bias

The topic of galaxy bias plays a critical role in establishing the link between the correlation functions of the matter field, that contains the cosmological information we can extract, and the correlation functions of the galaxy field, that is what we actually measure in a galaxy survey. In this section, I will try to describe the main points of its modelling, considering specifically those aspects that directly affect my project; this however will not represent an exhaustive and thorough description of the topic, since this would go beyond the main scope of this dissertation. I redirect the reader interested in a comprehensive review of the topic to [90].

The physics behind galaxy formation is very complicated, however we can easily imagine that galaxy formation occurs, in a very general sense, in regions where the matter distribution is denser: following gravitational instability, dark matter forms halos, and gas made of baryonic matter will
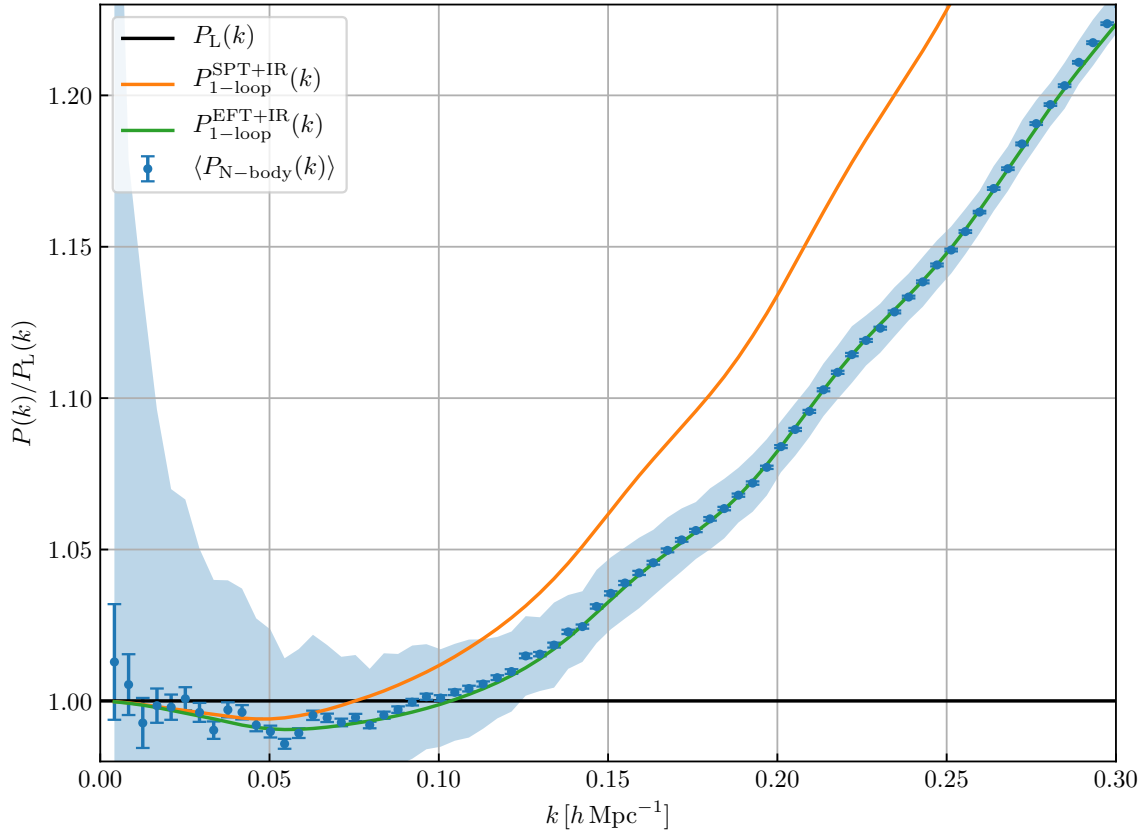
Figure 2.3: Comparison between the measurements of the matter power spectrum of the Minerva set of 298 N-body simulations and linear theory (black solid line), 1-loop SPT (orange solid line) and 1-loop EFT (the EFT counterterm amplitude has been fitted from the measurements with a $k_{\max} = 0.15h\,\mathrm{Mpc}^{-1}$). In all cases, the spectra have been IR-resummed (see below). The dots with error-bars represent the average of the N-body spectra along with their standard errors, while the shaded area represents the standard deviation of the measurements.

collapse as well, and even more so since it can emit radiation, become hotter, and thus form stars and galaxies. The latter are clearly effects that occur at smaller scales than the ones considered in the LSS, however this justifies why galaxies form in dense dark matter regions. This allows us to write the galaxy density field as a functional of the matter density field, and under the assumption that matter fluctuations are small – and this is true whenever the density field is smoothed over a large-scale mode – we can Taylor-expand the galaxy density field into a power series of the matter density field

$$\delta_g(\mathbf{x}) = \sum_n \frac{b_n}{n!}\delta^n(\mathbf{x}). \tag{2.80}$$

The coefficients b_n are called local bias parameters, and depend on the galaxy population that is being considered.

At the largest scales, we therefore can expect a direct proportionality of the matter density field and the galaxy field,

$$\delta_g(\mathbf{x}) \simeq b_1\delta(\mathbf{x}), \tag{2.81}$$

and this relation would then propagate to the correlation functions, with the galaxy power spectrum at the linear level being directly proportional to the matter, linear power spectrum, with the

proportionality constant given by the square of the linear bias b_1

$$P_{\mathrm{L}}^g(k) = b_1^2 P_{\mathrm{L}}(k). \tag{2.82}$$

The presence of non-linear bias, however, breaks this direct proportionality, and therefore introduces further non-linearities in the correlation functions; this is a new cause of non-linearities, that is independent from those sourced by the evolution of the matter density fields.

Notice that we can see the local bias expansion also as a series of operators of the density field $\delta(\mathbf{x})$; therefore, the linear bias can be seen as associated to the operator $\delta(\mathbf{x})$, that in Fourier space is $\delta_{\mathbf{k}}$, while the quadratic local bias would be related to the operator $\delta^2(\mathbf{x})$, that in Fourier space is a convolution of the matter density field,

$$\mathcal{FT}[\delta^2(\mathbf{x})]_{\mathbf{k}} \sim \int \mathrm{d}^3\mathbf{q}\, \delta_{\mathbf{q}}\delta_{\mathbf{k}-\mathbf{q}}. \tag{2.83}$$

The analysis of numerical simulations in recent years made evident, however, that the assumption of a purely local bias expansion does not provide an accurate and consistent description of the halo power spectrum and bispectrum [91, 70, 92]. This lead to the introduction of non-local corrections, that is operators of the density field, as well as of second derivatives of the gravitational potential, as for instance tidal terms.

Using symmetry arguments, it is possible to write down all possible local and non-local bias operators in a general bias expansion. Up to third order in the linear density field, the most general bias expansion including local and non-local operators is

$$\begin{aligned} \delta_g = & b_1\delta + \frac{b_2}{2}\delta^2 + \frac{b_3}{6}\delta^3 + \\ & + b_{\mathcal{G}_2}\mathcal{G}_2 + b_{\mathcal{G}_3}\mathcal{G}_3 + b_{\mathcal{G}_2\delta}\mathcal{G}_2\delta + b_{\Gamma_3}\Gamma_3; \end{aligned} \tag{2.84}$$

at this order in perturbation theory, the non-local operators that appear in the bias expansion are $\{\mathcal{G}_2, \mathcal{G}_3, \mathcal{G}_2\delta, \Gamma_3\}$, where the new operators are defined by

$$\mathcal{G}_2 \equiv \left[(\partial_i\partial_j\hat{\Phi})^2 - (\nabla^2\hat{\Phi})^2\right] \tag{2.85}$$

$$\mathcal{G}_3 \equiv -\frac{1}{2}\left[2(\partial_i\partial_j\hat{\Phi})^3 + \delta^3 - 3(\partial_i\partial_j\hat{\Phi})^2\delta\right] \tag{2.86}$$

$$\Gamma_3 \equiv (\partial_i\partial_j\hat{\Phi})^2 - (\nabla^2\hat{\Phi})^2 - \left[(\partial_i\partial_j\Phi_v)^2 - (\nabla^2\Phi_v)^2\right], \tag{2.87}$$

and where $\hat{\Phi}$ and Φ_v are defined by

$$\hat{\Phi} \equiv \nabla^{-2}\delta \tag{2.88}$$

$$\Phi_v \equiv -\frac{\nabla^{-2}\theta}{f\mathcal{H}}. \tag{2.89}$$

This is not the full story, however. Since halo and galaxy formation occur in finite regions of space, extra terms arise that depend on the Laplacian of local operators. These new operators introduce a series of higher-derivative bias coefficients; the higher-derivative operator relative to the linear bias operator, for example, takes the form $b_{\nabla^2\delta}\nabla^2\delta$, and in Fourier space this becomes $-b_{\nabla^2\delta}k^2\delta$. Finally, small-scale perturbations influence galaxy formation, which is instead a stochastic process. To account

for this, stochastic operators can be added to the bias expansion, that give rise to stochastic bias. These operators are responsible, for instance, for the shot-noise contributions of the galaxy correlation functions. At third order in the linear fields, the stochastic part of the galaxy density field can be written as

$$\delta_g^{\text{stoch}} = \epsilon + \epsilon_\delta \delta + \epsilon_{\delta^2} \delta^2 + \epsilon_{\mathcal{G}_2} \mathcal{G}_2. \tag{2.90}$$

Therefore, at third-order in the linear field, the full bias expansion for the galaxy density field including local, non-local, higher-derivative, and stochastic operators is given by

$$\begin{aligned}
\delta_g = b_1 \delta &+ \frac{b_2}{2} \delta^2 + \frac{b_3}{6} \delta^3 + \\
&+ b_{\mathcal{G}_2} \mathcal{G}_2 + b_{\mathcal{G}_3} \mathcal{G}_3 + b_{\mathcal{G}_2 \delta} \mathcal{G}_2 \delta + b_{\Gamma_3} \Gamma_3 + \\
&+ \epsilon + \epsilon_\delta \delta + \epsilon_{\delta^2} \delta^2 + \epsilon_{\mathcal{G}_2} \mathcal{G}_2 + b_{\nabla^2 \delta} \nabla^2 \delta.
\end{aligned} \tag{2.91}$$

Keep in mind that the same bias expansion can be written in terms of different bias operators and coefficients, but the final theoretical prediction for correlation functions should be the same regardless of the expansion.

Given this bias expansion, the model for the one-loop galaxy power spectrum in real space in EFT can be written down, under the working assumption that the $k^2 \delta_{\mathbf{k}}$ term is already second-order:

$$\begin{aligned}
P_g^{1\text{-loop}}(k) = b_1^2 \left[P_{\text{L}}(k) + \Delta P_{1\text{-loop}}(k) \right] &- 2\tilde{c}_0 k^2 P_{\text{L}}(k) + b_1 b_2 P_{b_1 b_2}(k) + b_1 b_{\mathcal{G}_2} P_{b_1 b_{\mathcal{G}_2}}(k) + \\
&+ b_1 b_{\Gamma_3} P_{b_1 b_{\Gamma_3}}(k) + b_2^2 P_{b_2 b_2}(k) + b_2 b_{\mathcal{G}_2} P_{b_2 b_{\mathcal{G}_2}}(k) + b_{\mathcal{G}_2}^2 P_{b_{\mathcal{G}_2} b_{\mathcal{G}_2}}(k) + P_{\text{shot}}(k);
\end{aligned} \tag{2.92}$$

here $\tilde{c}_0 = b_1^2 c_s^2 + b_1 b_{\nabla^2 \delta}$ is a combination of c_s^2 and $b_{\nabla^2 \delta}$, due to the fact that the EFT counterterm and the higher-derivative term in the power spectrum have the same k-dependence, and are therefore completely degenerate; $P_{\text{shot}}(k)$ is the shot-noise contribution to the galaxy power spectrum, and it arises due to the auto-correlation of any galaxy with itself, as well as from stochastic operators in the bias expansion; at large scales, this term can be modeled as

$$P_{\text{shot}}(k) = (1 + \alpha_P + \epsilon_{k^2} k^2 + \dots) P_{\text{PSN}}, \tag{2.93}$$

where P_{PSN} is the constant Poisson prediction of the shot-noise, $P_{\text{PSN}} = \bar{n}^{-1} (2\pi)^{-3}$, α_P is a constant parametrizing deviations from the Poisson prediction at large scales, and with further large-scales k^2 corrections arising from stochastic terms. In general, more terms can be added to the shot-noise term due to stochastic bias of the different operators; however, it is usually observed that at smaller scales the shot-noise does approach the Poisson prediction, and therefore this parametrization should be intended as a description of the large-scale contribution, including some scale-dependence.

In equation 2.92, a number of k-dependent terms were introduced, that are loop-correction terms arising from the non-linearities induced by the biased tracer. These terms include "mode-coupling" type integrals, where two linear power spectra are convolved together through a kernel, and "propagator" type terms, that have the structure of a linear power spectrum multiplying an integral, involving

another linear power spectrum:

$$P_{b_1 b_2}(k) = 2 \int \mathrm{d}^3\mathbf{q}\, F_2(\mathbf{q}, \mathbf{k} - \mathbf{q}) P_{\mathrm{L}}(q) P_{\mathrm{L}}(|\mathbf{k} - \mathbf{q}|) \tag{2.94}$$

$$P_{b_1 b_{\mathcal{G}_2}}(k) = 4 \int \mathrm{d}^3\mathbf{q}\, F_2(\mathbf{q}, \mathbf{k} - \mathbf{q}) S(\mathbf{q}, \mathbf{k} - \mathbf{q}) P_{\mathrm{L}}(q) P_{\mathrm{L}}(|\mathbf{k} - \mathbf{q}|) +$$
$$+ 8 P_{\mathrm{L}}(k) \int \mathrm{d}^3\mathbf{q}\, F_2(\mathbf{k}, -\mathbf{q}) S(\mathbf{q}, \mathbf{k} - \mathbf{q}) P_{\mathrm{L}}(q) \tag{2.95}$$

$$P_{b_1 b_{\Gamma_3}}(k) = -\frac{16}{7} P_{\mathrm{L}}(k) \int \mathrm{d}^3\mathbf{q}\, S(\mathbf{q}, \mathbf{k} - \mathbf{q}) S(\mathbf{k}, \mathbf{q}) P_{\mathrm{L}}(q) \tag{2.96}$$

$$P_{b_2 b_2}(k) = \frac{1}{2} \int \mathrm{d}^3\mathbf{q}\, P_{\mathrm{L}}(q) P_{\mathrm{L}}(|\mathbf{k} - \mathbf{q}|) - \frac{1}{2} \int P_{\mathrm{L}}^2(q) \mathrm{d}^3 q \tag{2.97}$$

$$P_{b_2 b_{\mathcal{G}_2}}(k) = 2 \int \mathrm{d}^3\mathbf{q}\, S(\mathbf{q}, \mathbf{k} - \mathbf{q}) P_{\mathrm{L}}(q) P_{\mathrm{L}}(|\mathbf{k} - \mathbf{q}|) \tag{2.98}$$

$$P_{b_{\mathcal{G}_2} b_{\mathcal{G}_2}}(k) = 2 \int \mathrm{d}^3\mathbf{q}\, S^2(\mathbf{q}, \mathbf{k} - \mathbf{q}) P_{\mathrm{L}}(q) P_{\mathrm{L}}(|\mathbf{k} - \mathbf{q}|), \tag{2.99}$$

where the additional kernel $S(\mathbf{q}_1, \mathbf{q}_2) \equiv (\hat{\mathbf{q}}_1 \cdot \hat{\mathbf{q}}_2)^2 - 1$ has been introduced. It can be shown that the propagator type integral in $P_{b_1 b_{\mathcal{G}_2}}(k)$ and the term $P_{b_1 b_{\Gamma_3}}(k)$ are directly proportional to each other.

The reason why in $P_{b_2 b_2}(k)$ a constant term is subtracted is related to the concept of bias renormalization [93, 94, 59]. Whenever higher-order terms are added to the bias expansion, we require that physical bias parameters should not vary: if they depended on the perturbation theory order, than they would not be well-defined physical observables. If the constant term of $P_{b_2 b_2}(k)$ were not subtracted, this would lead to an extra constant contribution that would be absorbed by the constant shot-noise. This in turn would effectively change the value of α_P as defined in equation 2.93. Subtracting the large-scale limit of $P_{b_2 b_2}(k)$, however, allows for the α_P coefficient appearing in the power spectrum model to be a well-defined physical parameter. A similar discussion is valid also for other bias parameters, where other operators must be renormalized. In this text, every time a theoretical model is written down in terms of bias parameters, they must be intended as already renormalized through the procedure of bias renormalization. In this way, whenever bias parameters are fitted, we can trust the values of these parameters since they represent well-defined physical quantities.

Using the same procedure, it is possible to write as well the theoretical prediction at one-loop level of the galaxy-matter cross power spectrum in real space:

$$P_{gm}(k) = b_1 \left[P_{\mathrm{L}}(k) + \Delta P_{1-\mathrm{loop}}(k) \right] - \tilde{c}_0^{\times} k^2 P_{\mathrm{L}}(k) +$$
$$+ \frac{1}{2} \left[b_2 P_{b_1 b_2}(k) + b_{\mathcal{G}_2} P_{b_1 b_{\mathcal{G}_2}}(k) + b_{\Gamma_3} P_{b_1 b_{\Gamma_3}}(k) \right], \tag{2.100}$$

where it is assumed that $\tilde{c}_0^{\times} = b_1 c_s^2 + b_{\nabla^2 \delta}$, and further k^2 corrections due to stochastic terms can be added to this expression.

Similarly, one can define as well the tree-level galaxy bispectrum

$$B_g(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = b_1^2 \left[2 b_1 F_2(\mathbf{k}_1, \mathbf{k}_2) + b_2 - 2 b_{\mathcal{G}_2} S(\mathbf{k}_1, \mathbf{k}_2) \right] P_{\mathrm{L}}(k_1) P_{\mathrm{L}}(k_2) + 2 \text{ cyc.} +$$
$$+ b_1^2 (1 + \alpha_1) \left[P_{\mathrm{L}}(k_1) + P_{\mathrm{L}}(k_2) + P_{\mathrm{L}}(k_3) \right] P_{\mathrm{PSN}} + (1 + \alpha_2) P_{\mathrm{PSN}}^2 \tag{2.101}$$

where the parameters α_1 and α_2, generally assumed different from each other, describe deviations of the shot-noise part of the bispectrum from the Poisson prediction. While at tree-level there are no counterterms in the bispectrum, it is in theory possible to introduce the higher-derivative bias $b_{\nabla^2 \delta}$ to

the expression. By performing the substitution $b_1 \mapsto b_1 - k^2 b_{\nabla^2 \delta}$ in the tree-level model except for the shot-noise part, it is sufficient to add the following term in order to take into account higher-derivative bias in the bispectrum

$$
\begin{aligned}
\Delta B_{\text{h.d.}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = &- 2b_1^2 b_{\nabla^2}(k_1^2 + k_2^2 + k_3^3) F_2(\mathbf{k}_1, \mathbf{k}_2) P_{\text{L}}(k_1) P_{\text{L}}(k_2) + \text{perm.} \\
&- b_1 b_{\nabla^2}(k_1^2 + k_2^2) \left[b_2 + 2b_{\mathcal{G}_2} S(\mathbf{k}_1, \mathbf{k}_2) \right] P_{\text{L}}(k_1) P_{\text{L}}(k_2) + \text{perm.}.
\end{aligned}
\tag{2.102}
$$

While this is not expected to have an important effect in the large-scale bispectrum alone, when the bispectrum is fitted together with the power spectrum assuming a one-loop model, it is possible that this additional term could break the degeneracy between c_s^2 and $b_{\nabla^2 \delta}$.

2.6 Redshift-space distortions

The redshift of objects measured by galaxy surveys is actually the sum of the cosmological redshift due to the Hubble flow and the peculiar velocity of the object along the line of sight. Therefore, if one uniquely identifies the position of galaxies using the measured redshifts – meaning observing galaxies in redshift space – this will affect the shape of the observed distribution of galaxies with respect to the real distribution. The distortion in the observed galaxy distribution is an effect called redshift-space distortions (RSDs), and clearly this affects as well the correlation functions in Fourier space.

There are two main ways this phenomenon affects the observed galaxy distribution. At the largest scales, peculiar velocities are mainly due to a coherent inflow of objects toward the center of an halo: this causes an apparent squashing of the galaxy distributions along the line of sight, a feature called the Kaiser effect. At smaller scales, instead, due to the random motion of galaxies within dark matter halos, an elongation along the line of sight is observed, and this feature takes the name of Finger-of-God (FoG) effect.

The Kaiser effect [95] introduces an additional direction-dependent prefactor to the matter power spectrum, that even in linear theory is no longer isotropic:

$$
P_{\text{L}}^{(s)}(\mathbf{k}) = (1 + f\mu^2)^2 P_{\text{L}}(k),
\tag{2.103}
$$

where $\mu \equiv \hat{\mathbf{k}} \cdot \hat{\mathbf{n}}$ is the cosine of the angle between the Fourier mode $\mathbf{k}$ and the line-of-sight direction $\hat{\mathbf{n}}$, and where $f \equiv \frac{\mathrm{d}\log D}{\mathrm{d}\log a}$ is the growth rate of matter perturbations. In ΛCDM, one can directly relate f to the matter-content of the Universe, and by Taylor expanding around $\Omega_m = 1$, one finds

$$
f(z) \simeq \left[\Omega_m(z) \right]^{6/11}.
\tag{2.104}
$$

This means that RSDs are a way to directly probe dark energy, since it provides a dynamical test of the expansion of the Universe by relating the dependence of the linear growth of perturbations with redshift.

At one-loop level, the Kaiser term of the matter power spectrum in redshift-space can be written as

$$
P^{(s)}(\mathbf{k}) = P^{(s)}(k, \mu) = P_{\delta\delta}(k) + 2f\mu^2 P_{\delta\theta}(k) + (f\mu^2)^2 P_{\theta\theta}(k),
\tag{2.105}
$$

where $P_{\delta\delta}(k)$ is the power spectrum of density perturbations, $P_{\theta\theta}(k)$ is the velocity divergence power spectrum, and $P_{\delta\theta}(k)$ is the density perturbations-velocity divergence cross-power spectrum. At the

linear level, these three spectra are exactly the same and equal to the linear power spectrum $P_{\rm L}(k)$; at one-loop level, however, they start being different, and the loop-corrections for the two new spectra are

$$\Delta P^{\rm 1-loop}_{\theta\theta}(k) = 2 \int {\rm d}^3{\bf q}\, [G_2({\bf q},{\bf k}-{\bf q})]^2\, P_{\rm L}(q)P_{\rm L}(|{\bf k}-{\bf q}|) + 6P_{\rm L}(k)\int {\rm d}^3{\bf q}\, G_3({\bf k},{\bf q},-{\bf q})P_{\rm L}(q) \quad (2.106)$$

$$\Delta P^{\rm 1-loop}_{\delta\theta}(k) = 2 \int {\rm d}^3{\bf q}\, F_2({\bf q},{\bf k}-{\bf q})G_2({\bf q},{\bf k}-{\bf q})P_{\rm L}(q)P_{\rm L}(|{\bf k}-{\bf q}|)+$$

$$+ 3P_{\rm L}(k)\int {\rm d}^3{\bf q}\, [F_3({\bf k},{\bf q},-{\bf q}) + G_3({\bf k},{\bf q},-{\bf q})]\, P_{\rm L}(q). \quad (2.107)$$

Notice however that the full matter power spectrum at one-loop in redshift space contains many more contributions. It can be written down explicitly as in Appendix A.1 where b_1 has to be set to 1 and all other bias parameters are set to zero.

The FoG effect can be phenomenologically modelled by a damping of the anisotropic power spectrum at smaller scales; at linear level (for simplicity) one can write

$$P^{(s)}_{\rm L}({\bf k}) = (1 + f\mu^2)^2 P_{\rm L}(k) D_{\rm FoG}[k^2\mu^2\sigma_v^2], \quad (2.108)$$

where σ_v^2 is the velocity dispersion, and is usually treated as a free parameter. The shape of $D_{\rm FoG}[x]$ is usually assumed to be either a Gaussian, $D_{\rm FoG}[x] \sim \exp(-x^2)$, or a Lorentzian, $D_{\rm FoG}[x] \sim 1/(1+x^2)$; in both these cases, the Taylor series expansion at second order in x is the same, $D_{\rm FoG}[x] \sim 1 - x^2 + \mathcal{O}(x^4)$, and in this limit the FoG damping can be approximated with some counterterm-like terms $\sim k^2 P_{\rm L}(k)$, that thus will be degenerate with the usual EFT counterterms.

Other more accurate models that describe the matter power spectrum in redshift-space have been proposed as well, either derived from the pairwise velocity distribution or specifically accounting for BAOs [96, 97].

When writing the correlation functions for galaxies, it is more convenient to write the full anisotropic correlation functions as a function of generalized, redshift-space kernels; considering only contributions that enter in the predictions for the galaxy tree-level bispectrum and the one-loop power spectrum, the generalized kernels become [86]:

$$Z_1({\bf k}) = b_1 + f\mu^2 \quad (2.109)$$

$$Z_2({\bf k}_1,{\bf k}_2) = b_1 F_2({\bf k}_1,{\bf k}_2) + \frac{b_2}{2} + b_{\mathcal{G}_2} S({\bf k}_1,{\bf k}_2)+$$

$$+ f\mu^2 G_2({\bf k}_1,{\bf k}_2) + \frac{f\mu k}{2}\left[\frac{\mu_1}{k_1}Z_1({\bf k}_2) + \frac{\mu_2}{k_2}Z_1({\bf k}_1)\right] \quad (2.110)$$

$$Z_3({\bf k}_1,{\bf k}_2,{\bf k}_3) = b_1 F_3({\bf k}_1,{\bf k}_2,{\bf k}_3) + b_2 F_2({\bf k}_1,{\bf k}_2) + 2b_{\mathcal{G}_2} S({\bf k}_1,{\bf k}_{23})F_2({\bf k}_2,{\bf k}_3)+$$

$$- \frac{4}{7}b_{\Gamma_3} S({\bf k}_1,{\bf k}_{23})S({\bf k}_2,{\bf k}_3) + f\mu^2 G_3({\bf k}_1,{\bf k}_2,{\bf k}_3) + \frac{(f\mu k)^2}{2}Z_1({\bf k}_1)\frac{\mu_2}{k_2}\frac{\mu_3}{k_3}+$$

$$+ f\mu k Z_1({\bf k}_1)\frac{\mu_{23}}{k_{23}}G_2({\bf k}_2,{\bf k}_3) + \frac{b_2 f\mu k}{2}\frac{\mu_1}{k_1} + b_{\mathcal{G}_2} f\mu k\frac{\mu_1}{k_1}S({\bf k}_2,{\bf k}_3)+$$

$$+ f\mu k\frac{\mu_3}{k_3}\left[b_1 F_2({\bf k}_1,{\bf k}_2) + f\mu_{12}^2 G_2({\bf k}_1,{\bf k}_2)\right]. \quad (2.111)$$

Here, ${\bf k}$ and μ without subscripts in the n-th order kernel are ${\bf k} = {\bf k}_1 + \cdots + {\bf k}_n$ and $\mu = \hat{\bf k}\cdot\hat{\bf n}$, while ${\bf k}_{ij} = {\bf k}_i + {\bf k}_j$, $\mu_{ij} = \hat{\bf k}_{ij}\cdot\hat{\bf n}$, and finally $Z_3({\bf k}_1,{\bf k}_2,{\bf k}_3)$ as written here is still to be symmetrized with

respect to all the Fourier momenta. With these generalized kernels, the anisotropic galaxy power spectrum in redshift space at one-loop can be written as ([46])

$$P_g^{(s)}(\mathbf{k}) = Z_1^2(\mathbf{k})P_\mathrm{L}(k) + 2\int \mathrm{d}^3\mathbf{q}\, [Z_2(\mathbf{q}, \mathbf{k} - \mathbf{q})]^2 P_\mathrm{L}(q)P_\mathrm{L}(|\mathbf{k} - \mathbf{q}|)$$
$$+ 6Z_1(\mathbf{k})P_\mathrm{L}(k)\int \mathrm{d}^3\mathbf{q}\, Z_3(\mathbf{k}, \mathbf{q}, -\mathbf{q})P_\mathrm{L}(q) + P_\mathrm{ct}^{(s)}(\mathbf{k}) + P_\mathrm{noise}^{(s)}(\mathbf{k}), \qquad (2.112)$$

where $P_\mathrm{ct}^{(s)}(\mathbf{k})$ contains all EFT counterterms of the appropriate order, and where $P_\mathrm{noise}^{(s)}(\mathbf{k})$ contains the constant shot-noise term and some $\mathbf{k}$-dependent stochastic terms. The 1-loop correction terms can be written as a combination of 23 mode-coupling terms (7 of which are shared with the power spectrum model in real space), and 5 propagator terms (2 of which are shared with real-space model), for a total of 28 unique terms. In this way, all terms that are produced are at most $\mathcal{O}(\mu^8)$. While employing a different bias expansion than the one proposed in [60], the number of integrals in the one-loop galaxy power spectrum in redshift space is in perfect agreement between the treatments.

The full expression for the anisotropic power spectrum is given in Appendix A.1. From there, one can notice how, in the limit for $f \to 0$, the anisotropic power spectrum reduces to the prediction for the real-space galaxy power spectrum, equation 2.92.

In principle, one could measure directly the anisotropic power spectrum as a function of two variables, $P_g(\mathbf{k}) = P_g(k, \mu)$, or equivalently $P_g(\mathbf{k}) = P_g(k_\parallel, k_\perp)$, where $k_\parallel$ and $k_\perp$ are the components of $\mathbf{k}$ along the line-of-sight and perpendicular to it, respectively; however, this would require to work with a very large number of data points. Instead of working directly with the anisotropic power spectrum, it is more convenient to work with its Legendre multipoles:

$$P_\ell(k) \equiv \frac{2\ell + 1}{2}\int_{-1}^{1} \mathrm{d}\mu\, \mathcal{L}_\ell(\mu)P_g^{(s)}(\mathbf{k}), \qquad (2.113)$$

where $\mathcal{L}_\ell(\mu)$ is the Legendre polynomial of degree ℓ as a function of the cosine μ of the Fourier mode $\mathbf{k}$ with respect to the line-of-sight. In this way, since the bulk of information is contained in the first even multipoles – with them being the monopole ($\ell = 0$), the quadrupole ($\ell = 2$) and the hexadecapole ($\ell = 4$) – instead of working with one function in two variables, we can now consider three functions of one variable, that can be treated in an easier way in data analysis.

By looking at the full expression of the anisotropic power spectrum, it could be deemed trivial to write explicitly the multipoles, since the dependence onto the angle with the line of sight appears only as integer powers of μ, that can be integrated exactly once multiplied by the appropriate Legendre polynomials. However, infrared resummation techniques in redshift space introduce spurious μ-dependent contributions in the wiggly part of the power spectrum terms (see section 2.7 below). This makes non-trivial to write down an explicit form of the multipoles of the galaxy power spectrum in redshift space.

Using the same generalized kernels in redshift-space, one is able to write down as well an expression for the tree-level bispectrum in redshift space:

$$B_g^{(s)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) \equiv [2Z_1(\mathbf{k}_1)Z_1(\mathbf{k}_2)Z_2(\mathbf{k}_1, \mathbf{k}_2)P_\mathrm{L}(k_1)P_\mathrm{L}(k_2) + \text{cyc.}] + B_\mathrm{noise}^{(s)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3). \qquad (2.114)$$

2.7 Baryon acoustic oscillations and IR resummation

The matter 2-point correlation function in configuration space shows a peak at a physical scale around 150 Mpc. This peak is a consequence of the interaction of dark matter and baryonic matter in the early Universe. The physics behind this peak is well known from CMB observations, therefore the position of this peak can be used as a stantard ruler, a know physical distance that, when measured in the galaxy population at several redshifts, can be used to study the expansion history of the Universe.

Let us know discuss the physical process producing this feature. Take a point-like overdensity in all species at very early time. Before recombination, baryonic matter was coupled to radiation to form a tightly-coupled baryon-photon fluid. Because of the pressure gradient caused by the Thomson scattering between photons and electrons in this fluid, the overdensity moves radially at the fluid sound speed, away from the initial position. This leaves the cold dark matter perturbation at the center. At decoupling, photons are finally free-streaming, leaving a baryon overdensity around the distance travelled up to that point, $\sim 150\,\mathrm{Mpc}$. Under the action of gravitational instability, baryonic matter and dark matter are attracted by both overdensities. This mechanism, applied to the generic superposition of overdensities that characterizes the density perturbations in the early Universe, is the one responsible for the peak in the correlation function around 150 Mpc. In Fourier space, this produces oscillations in the power spectrum at very high redshift.

The peak in the galaxy 2PCF has also been detected in real data, with the first detection in [98]. By measuring this standard ruler at different redshifts, it is possible to determine the relation between the angular diameter distance relative to the peak and redshift, and thus the evolution of the Hubble parameter. This can complement measures of the expansion rate through the luminosity distance of standard candles [3, 4], to shed light on dark energy.

At lower redshift, however, bulk flows are responsible for perturbations displacements of about $\sim 10\,\mathrm{Mpc}$, and this effect contributes to a damping of the acoustic peak in the correlation function as predicted by linear theory. In Fourier space, this translates into a damping of the oscillatory features of the power spectrum [13]. This means that evolving the initial power spectrum taking into account only the linear growth factor of density perturbations leads to larger oscillations at lower redshift than the ones that would be observed in the actual matter power spectrum; this effect is also observed in the measured matter power spectrum in N-body simulations. Moreover, since loop-corrections are usually treated as functionals of the linear power spectrum of matter, this greater amplitude of the acoustic oscillations would propagate even at higher order in perturbation theory.

To solve these problems, we follow the approach of [99, 100] developed under the context of Time-Sliced PT [101], according to which the BAO feature in the power spectrum is resummed non-perturbatively. In order to achieve this, the first step is to split the linear power spectrum into a smooth no-wiggle part $P_{\mathrm{nw}}(k)$, capturing the full broadband shape of the power spectrum, and into a wiggly part $P_{\mathrm{w}}(k)$, describing the baryon acoustic oscillations:

$$P_{\mathrm{L}}(k) = P_{\mathrm{nw}}(k) + P_{\mathrm{w}}(k). \tag{2.115}$$

There are a series of different approaches to compute the broadband shape of the linear power spectrum; one common approach is to use the prescription from Eisenstein and Hu [102] to compute an approximate broadband power spectrum $P_{\mathrm{EH}}(k)$, and to use this approximate smooth power spectrum

as a way to filter out the wiggles by applying a Gaussian smoothing kernel [103],

$$P_{\mathrm{nw}}(k) = P_{\mathrm{EH}}(k)\mathcal{F}\left[\frac{P_{\mathrm{L}}(k)}{P_{\mathrm{EH}}(k)}\right];\qquad(2.116)$$

once $P_{\mathrm{nw}}(k)$ is computed, the wiggly part can be obtained by a simple subtraction of the smooth part from the full linear power spectrum.

The wiggly part is then suppressed by a μ-dependent damping factor, $\mathrm{e}^{-k^2\Sigma_{\mathrm{tot}}^2(\mu)}$, with

$$\Sigma_{\mathrm{tot}}^2(\mu) = \Sigma^2 + f\mu^2(2\Sigma^2 + f(\Sigma^2 - \delta\Sigma^2)) + (f\mu^2)^2\delta\Sigma^2,\qquad(2.117)$$

and where the quantities Σ^2 and $\delta\Sigma^2$ are defined by

$$\Sigma^2 = \frac{1}{6\pi^2}\int_0^{k_{\mathrm{S}}}\mathrm{d}q\,P_{\mathrm{nw}}(q)\left[1 - j_0\left(\frac{q}{k_{\mathrm{osc}}}\right) + 2j_2\left(\frac{q}{k_{\mathrm{osc}}}\right)\right]\qquad(2.118)$$

$$\delta\Sigma^2 = \frac{1}{6\pi^2}\int_0^{k_{\mathrm{S}}}\mathrm{d}q\,3P_{\mathrm{nw}}(q)j_2\left(\frac{q}{k_{\mathrm{osc}}}\right);\qquad(2.119)$$

in these expressions, k_{osc} is the BAO wavenumber, and k_{S} is the scale that separates the short and long modes, $k_{\mathrm{S}} = 0.2\,h\,\mathrm{Mpc}^{-1}$ [99]. With these, the leading order galaxy power spectrum in redshift space is

$$P_{\mathrm{LO}}(\mathbf{k}) = Z_1(\mathbf{k})^2\left[P_{\mathrm{nw}}(k) + \mathrm{e}^{-k^2\Sigma_{\mathrm{tot}}^2(\mu)}P_{\mathrm{w}}(k)\right].\qquad(2.120)$$

Loop-corrections would then be computed from the $P_{\mathrm{LO}}(\mathbf{k})$ instead of the linear power spectrum; in real space this does not introduce any complication whatsoever, since in the limit for $f \to 0$ one has $Z_1(\mathbf{k}) \to b_1$ and $\Sigma_{\mathrm{tot}}^2(\mu) \to \Sigma^2$, and therefore $P_{\mathrm{LO}}(k)$ is a function of just $k = |\mathbf{k}|$; in redshift space, however, this is not the case anymore, since loop-corrections then would have to be computed as more complicated integrals of the leading-order anisotropic power spectrum; an alternative approach is to compute loop-corrections from $P_{\mathrm{L}}(k)$ and $P_{\mathrm{nw}}(k)$, compute the wiggly part of the loop corrections, and then resum those [101, 99, 46, 104]. Schematically

$$\Delta P_{\mathrm{L}}^{1-\mathrm{loop}}(\mathbf{k}) = P_{1-\mathrm{loop}}[P_{\mathrm{L}}]\qquad(2.121)$$

$$\Delta P_{\mathrm{nw}}^{1-\mathrm{loop}}(\mathbf{k}) = P_{1-\mathrm{loop}}[P_{\mathrm{nw}}]\qquad(2.122)$$

$$\Delta P_{\mathrm{w}}^{1-\mathrm{loop}}(\mathbf{k}) = \Delta P_{\mathrm{L}}^{1-\mathrm{loop}}(\mathbf{k}) - \Delta P_{\mathrm{nw}}^{1-\mathrm{loop}}(\mathbf{k})\qquad(2.123)$$

$$\Delta P_{\mathrm{IR}}^{1-\mathrm{loop}}(\mathbf{k}) = \Delta P_{\mathrm{nw}}^{1-\mathrm{loop}}(\mathbf{k}) + \mathrm{e}^{-k^2\Sigma_{\mathrm{tot}}^2(\mu)}\Delta P_{\mathrm{w}}^{1-\mathrm{loop}}(\mathbf{k}).\qquad(2.124)$$

Then, it is possible to compute the next-to-leading order power spectrum as follows:

$$P_{\mathrm{NLO}}(\mathbf{k}) = Z_1(\mathbf{k})^2\left[P_{\mathrm{nw}}(k) + \mathrm{e}^{-k^2\Sigma_{\mathrm{tot}}^2(\mu)}\left(1 + k^2\Sigma_{\mathrm{tot}}^2(\mu)\right)P_{\mathrm{w}}(k)\right] + \Delta P_{\mathrm{IR}}^{1-\mathrm{loop}}(\mathbf{k}).\qquad(2.125)$$

However, with this, the theoretical prediction still lacks contributions of order $\mathcal{O}(P_{\mathrm{w}}^2)$, but they are usually deemed sub-leading with respect to all other terms.

Similarly, for the tree-level galaxy bispectrum in redshift space, one can write the leading-order

prediction in the IR-resummation formalism as [100]

$$B_{\mathrm{LO}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = Z_1(\mathbf{k}_1) Z_1(\mathbf{k}_2) Z_2(\mathbf{k}_1, \mathbf{k}_2) \big[P_{\mathrm{nw}}(k_1) P_{\mathrm{nw}}(k_2) +$$
$$+ P_{\mathrm{nw}}(k_1) \mathrm{e}^{-k_2^2 \Sigma_{\mathrm{tot}}^2(\mu_2)} P_{\mathrm{w}}(k_2) + P_{\mathrm{nw}}(k_2) \mathrm{e}^{-k_1^2 \Sigma_{\mathrm{tot}}^2(\mu_1)} P_{\mathrm{w}}(k_1) \big] + \mathrm{cyc.} \qquad (2.126)$$

This represents the theoretical framework that describes the galaxy power spectrum and bispectrum of the galaxy distribution in redshift space; the power spectrum model described here has been recently used to re-analyse the data from the *BOSS* galaxy survey with a full-shape analysis [48, 16], and was also compared to data from a set of large N-body simulations [47] with promising results.

The galaxy power spectrum and bispectrum in redshift space as described here are basically linear combinations of functionals of the leading-order matter power spectrum, which is on its own a functional of the linear matter power spectrum. The linear matter power spectrum can be computed starting from the cosmological parameters, and therefore is the function representing the main cosmological information content.

In order to extract this information, we have to compare these theoretical predictions to actual measurements of these correlators. This is the main objective of galaxy redshift surveys.

However, when considering real data, there are a few complications to take into account. One of such complications arises when considering the survey window function, that introduces non-trivial convolutions in all correlation functions in Fourier space, as well as in covariance matrices. This aspect, however, while being crucial to the analysis of actual data, will not be discussed in this dissertation since, for the time being, the data used for the analyses presented here are based on N-body simulations, and therefore we do not need to take into account any effect due to the survey geometry.

On the other hand, the way in which measurements of correlation functions are performed requires additional adjustments of our theoretical predictions; for this reason, let us now delve more deeply into the estimators of correlation functions in Fourier space.

2.8 Estimators of correlation functions in Fourier space

The main product of galaxy redshift surveys is a catalog of the angular positions and redshifts of the galaxies that have been observed by the survey. Starting from one of such catalogs, one can measure correlation functions in Fourier space, like power spectrum and bispectrum, by means of estimators defined on a large but finite volume enclosing the survey footprint. By assuming a fiducial cosmology, redshifts can be converted into distances, and therefore to each galaxy can be assigned a position in three-dimensional space.

In this dissertation, no actual data will be analysed but only outputs from N-body simulations. Data coming from galaxy redshift surveys are affected by the geometry of the survey, and the complications arising from the survey window function is beyond the scope of this text. For the moment, we will focus only on the distribution of matter and dark matter halos in periodic boxes, extracted from N-body simulations in the form of simple three-dimensional catalogs. In fact, the list of the positions of either the particles or dark matter halos of the simulation completely defines the number density

of the box,

$$n(\mathbf{x}) = \sum_{p=1}^{N_p} \delta_D(\mathbf{x} - \mathbf{x}_p) \tag{2.127}$$

where the sum is performed over all the N_p objects in the catalog, and the Dirac delta function represents the density profile of point-like particles at positions $\mathbf{x}_p$.

At this point, one can write directly the Fourier counterpart of the galaxy density field, defined over a discrete Fourier grid, using direct summation: given a box of size L, the Fourier transform of the density field on a Fourier grid of linear size N_G can be written as

$$\delta_{\mathbf{k}} = \frac{1}{k_f^3} \frac{1}{N_P} \sum_{p=1}^{N_P} e^{-i\mathbf{k}\cdot\mathbf{x}_p}, \tag{2.128}$$

where $k_f \equiv 2\pi/L$ is the fundamental frequency of the box – that is, the longest possible wavemode one has inside the box – and where the sum runs over all the N_P objects in the catalog, each with position $\mathbf{x}_p$; additionally, the Fourier modes $\mathbf{k}$ are defined only over the N_G^3 grid points. This direct summation technique, requiring to compute $\delta_{\mathbf{k}}$ at each grid point, is of course highly numerically expensive, scaling as $N_G^3 N_P$, and therefore alternative approaches are usually employed.

In fact, one can use one of the common mass assignment schemes, to assign each object in the catalog to a small fraction of the total N_G^3 grid points, with different weights depending on the relative position of each object to the grid points. This gives a discrete description of the overdensity field in configuration space $\delta(\mathbf{x}_G)$, where $\mathbf{x}_G$ can be seen as the discrete position on the grid. By means of a Fast Fourier Transform (FFT), the estimated overdensity field $\delta(\mathbf{x}_G)$ can be transformed into its Fourier counterpart $\delta_{\mathbf{k}}$, and from this field the different correlation functions can be estimated [105]. In this case, the computation time scales as the number of objects in the catalogs.

Given the density field in Fourier space $\delta_{\mathbf{k}}$, the real-space power spectrum estimator is

$$\hat{P}(k) \equiv \frac{k_f^3}{N_P(k)} \sum_{\mathbf{q}\in k} |\delta_{\mathbf{q}}|^2 ; \tag{2.129}$$

the notation $\mathbf{q} \in k$ means that the sum runs over all discrete wavevectors $\mathbf{q}$ in a k-bin of size Δk, or $k - \Delta k/2 \le |\mathbf{q}| < k + \Delta k/2$. With "discrete wavevector" here we mean any vector on the Fourier grid, that has coordinates that are integer multiples of the fundamental frequency: $\mathbf{q}/k_f \in \mathbb{Z}^3$. Finally, $N_P(k)$ represents the number of discrete wavevectors $\mathbf{q}$ in the k-bin with width Δk:

$$N_P(k) \equiv \sum_{\mathbf{q}\in k} 1. \tag{2.130}$$

Similarly, the real-space bispectrum estimator is defined as

$$\hat{B}(k_1, k_2, k_3) \equiv \frac{k_f^3}{N_B(k_1, k_2, k_3)} \sum_{\mathbf{q}_1\in k_1} \sum_{\mathbf{q}_2\in k_2} \sum_{\mathbf{q}_3\in k_3} \delta_K(\mathbf{q}_{123}) \delta_{\mathbf{q}_1} \delta_{\mathbf{q}_2} \delta_{\mathbf{q}_3} ; \tag{2.131}$$

differently from the power spectrum case, in the bispectrum the individual bins are actually triangle bins: the bin $t_i(k_1, k_2, k_3)$ is defined as the set of all closed fundamental triangles whose sides are each

in one of the bins of centers k_1, k_2, and k_3, or symbolically,

$$t(k_1, k_2, k_3) = \{(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) : k_j - \Delta k/2 \leq |\mathbf{q}_j| < k_j + \Delta k/2 \wedge \mathbf{q}_{123} = 0\}; \tag{2.132}$$

the Kronecker delta in the expression for the estimator, equation 2.131, is what forces only closed triangles, such that $\mathbf{q}_{123} = 0$, to be considered. The quantity $N_B(k_1, k_2, k_3)$ is the number of fundamental triangles in the triangle bin:

$$N_B(k_1, k_2, k_3) \equiv \#t(k_1, k_2, k_3) = \sum_{\mathbf{q}_1 \in k_1} \sum_{\mathbf{q}_2 \in k_2} \sum_{\mathbf{q}_3 \in k_3} \delta_K(\mathbf{q}_{123}). \tag{2.133}$$

Going back to the power spectrum, usually one considers one-dimensional bins that are defined as equidistant from each other and of equal width. In this specific case, the binning scheme can be defined completely by the first center c and by the width of the bin s, both in units of the fundamental frequency k_f, and by the total number of bins N_k; then, the center of each bin will be defined by

$$k_{c,i} \equiv [c + (i-1)s] \, k_f, \qquad i = 1, \ldots, N_k, \tag{2.134}$$

and the lower and higher limits of each bin will be simply given by $k_{c,i} - sk_f/2$ and $k_{c,i} + sk_f/2$ respectively. Thus, the numbers (s, c, N_k) completely define a binning scheme for the power spectrum.

Similarly for the bispectrum, with the additional *caveat* that not all triangle bins chosen by a random combination of centers k_1, k_2, k_3 are associated to non-empty triangle bins; as an example, assuming a binning scheme with $s = 1$ and $c = 1$, the triangle bins $(5, 4, 2)k_f$ and $(9, 7, 6)k_f$ contain fundamental triangles, while $(4, 1, 1)k_f$ and $(10, 4, 2)k_f$ do not. For the bispectrum case, one can assume $k_1 \geq k_2 \geq k_3$, and all non-empty triangle bins can be defined and ordered by setting:

$$k_1 = [c + (i_1 - 1)s] \, k_f \qquad i_1 = 1, \ldots, N_k \tag{2.135}$$

$$k_2 = [c + (i_2 - 1)s] \, k_f \qquad i_2 = \left\lceil \frac{i_1 - c/s}{2} \right\rceil, \ldots, i_1 \tag{2.136}$$

$$k_3 = [c + (i_3 - 1)s] \, k_f \qquad i_3 = \max\left[1, \lceil i_1 - i_2 - c/s \rceil\right], \ldots, i_2. \tag{2.137}$$

Notice that while some of the triangle bins defined in this way have centers that do not form a closed triangle, they still contain closed fundamental triangles on the grid: an example with the binning scheme $(s = 1, c = 1)$ is the triangle bin $(4, 2, 1)k_f$, whose centers do not form a closed triangle (since $4 \not\leq 2 + 1$), but still contains fundamental triangles, like the ones having sides $(\sqrt{13}, \sqrt{5}, \sqrt{2})k_f$ or $(\sqrt{14}, \sqrt{6}, \sqrt{2})k_f$. The number of triangle bins is given by

$$N_t \equiv \sum_{i_1=1}^{N_k} \sum_{i_2=\lceil(i_1-c/s)/2\rceil}^{i_1} \sum_{i_3=\max[1,\lceil i_1-i_2-c/s\rceil]}^{i_2} 1. \tag{2.138}$$

Simple extensions can be implemented for the redshift-space multipoles. In particular, for the power spectrum one has

$$\hat{P}_\ell(k) \equiv (2\ell + 1) \frac{k_f^3}{N_P(k)} \sum_{\mathbf{q} \in k} \mathcal{L}_\ell(\mu_{\mathbf{q}}) \left| \delta_{\mathbf{q}}^{(s)} \right|^2, \tag{2.139}$$

where $\mu_{\mathbf{q}}$ is the cosine of the angle between the wavevector $\mathbf{q}$ and the line of sight, ℓ is the order of the

multipole, and $\mathcal{L}_\ell(\mu)$ is the Legendre polynomial of degree ℓ. Similarly for the bispectrum multipoles,

$$\hat{B}_\ell(k_1, k_2, k_3) \equiv (2\ell + 1) \frac{k_f^3}{N_B(k_1, k_2, k_3)} \sum_{\mathbf{q}_1 \in k_1} \sum_{\mathbf{q}_2 \in k_2} \sum_{\mathbf{q}_3 \in k_3} \delta_K(\mathbf{q}_{123}) \mathcal{L}_\ell(\mu_{\mathbf{q}_1}) \delta_{\mathbf{q}_1} \delta_{\mathbf{q}_2} \delta_{\mathbf{q}_3}, \tag{2.140}$$

where $\mathbf{q}_1 \in k_1$ is assumed to be the longest Fourier momentum, and $\mu_{\mathbf{q}_1}$ being the cosine between $\mathbf{q}_1$ and the line-of-sight direction.

All estimators reported here contain also a contribution due to the auto-correlation of a particle with itself, the so-called shot-noise. Considering this, the estimator of the power spectrum in real space can be written as

$$\hat{P}(k) = P_{\text{tot}}(k) = P(k) + \frac{1}{(2\pi)^3 \bar{n}}, \tag{2.141}$$

where $P(k)$ is the physical power spectrum, and similarly for the bispectrum

$$\hat{B}(k_1, k_2, k_3) = B_{\text{tot}}(k_1, k_2, k_3) = B(k_1, k_2, k_3) + \frac{1}{(2\pi)^3 \bar{n}} \left[\tilde{P}(k_1) + \tilde{P}(k_2) + \tilde{P}(k_3) \right] + \frac{1}{(2\pi)^6 \bar{n}^2}. \tag{2.142}$$

Here, $\tilde{P}(k)$ is the power spectrum estimated on the triangle bin,

$$\tilde{P}(k_1) \equiv \frac{k_f^3}{N_B(k_1, k_2, k_3)} \sum_{\mathbf{q}_1 \in k_1} \sum_{\mathbf{q}_2 \in k_2} \sum_{\mathbf{q}_3 \in k_3} \delta_K(\mathbf{q}_{123}) |\delta_{\mathbf{q}_1}|^2; \tag{2.143}$$

however, it can be verified that $\hat{P}(k) \simeq \tilde{P}(k)$ and they can be replaced by one another[2]. However, the Poisson-shot noise assumption is not always true. In this case, one cannot subtract the shot noise exactly from the measurements, and is forced to model it instead.

Given the estimator of equation 2.129, the Gaussian power spectrum covariance can be written as

$$C_{ij}^P = \frac{2}{N_P(k)} \delta_{ij}^K P_{\text{tot}}^2(k_i). \tag{2.144}$$

Given the estimator in eq. (2.131) and introducing the notation $t_i = \{k_{i_1}, k_{i_2}, k_{i_3}\}$ for a generic triangle configuration, the bispectrum covariance can be written in general terms as

$$
\begin{aligned}
C_{ij}^B \simeq\; & \delta_{i_1 j_1}^K \delta_{i_2 j_2}^K \delta_{i_3 j_3}^K \frac{1}{N_B(t_i)} P_{\text{tot}}(k_{i_1}) P_{\text{tot}}(k_{i_2}) P_{\text{tot}}(k_{i_3}) + 5 \text{ perm.} \\
& + \delta_{i_1 j_1}^K \frac{k_f^6}{N_B(t_i) N_B(t_j)} \sum_{\mathbf{q}_1 \in k_{i_1}} \sum_{\mathbf{q}_2 \in k_{i_2}} \sum_{\mathbf{q}_3 \in k_{i_3}} \sum_{\mathbf{p}_2 \in k_{j_2}} \sum_{\mathbf{p}_3 \in k_{j_3}} \delta_K(\mathbf{q}_{123}) \delta_K(\mathbf{q}_1 - \mathbf{p}_{23}) \\
& \times B_{\text{tot}}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) B_{\text{tot}}(\mathbf{q}_1, \mathbf{p}_2, \mathbf{p}_3) + 2 \text{ perm.} \\
& + \delta_{i_1 j_1}^K \frac{k_f^6}{N_B(t_i) N_B(t_j)} \sum_{\mathbf{q}_1 \in k_{i_1}} \sum_{\mathbf{q}_2 \in k_{i_2}} \sum_{\mathbf{q}_3 \in k_{i_3}} \sum_{\mathbf{p}_2 \in k_{j_2}} \sum_{\mathbf{p}_3 \in k_{j_3}} \delta_K(\mathbf{q}_{123}) \delta_K(\mathbf{q}_1 - \mathbf{p}_{23}) \\
& \times P_{\text{tot}}(\mathbf{q}_1) T_{\text{tot}}(\mathbf{q}_2, \mathbf{q}_3, \mathbf{p}_2, \mathbf{p}_3) + 2 \text{ perm.} \\
& + \frac{k_f^9}{N_B(t_i) N_B(t_j)} \sum_{\mathbf{q}_1 \in k_{i_1}} \sum_{\mathbf{q}_2 \in k_{i_2}} \sum_{\mathbf{q}_3 \in k_{i_3}} \sum_{\mathbf{p}_1 \in k_{j_1}} \sum_{\mathbf{p}_2 \in k_{j_2}} \sum_{\mathbf{p}_3 \in k_{j_3}} \delta_K(\mathbf{q}_{123}) \delta_K(\mathbf{p}_{123}) \\
& \times T_{6,\text{tot}}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3),
\end{aligned}
\tag{2.145}
$$

---

[2]Of course, this statement is dependent on the error bars one is working with.

where we assumed the thin-shell approximation ($k_i \gg \Delta k$) to be valid for the Gaussian contribution on the r.h.s. and P_{tot}, B_{tot}, T_{tot} and $T_{6,\text{tot}}$ represent, respectively, the power spectrum, bispectrum, trispectrum and the connected 6-point function including shot-noise.

2.8.1 Discretization effects on theoretical predictions

Comparing theory and data in the most consistent way is crucial in analysing the power spectrum and bispectrum. While complications arise with actual data coming from surveys, the case of data extracted from N-body simulation is relatively simple to pose. Focusing in particular on the case of N-body simulations, it is important to have in mind the real- and redshift-space estimators for the correlation functions when computing theoretical predictions. In fact, theoretical predictions are defined over single Fourier momenta in the case of the power spectrum, or over single Fourier-space triangles in the case of the bispectrum, while measurements are defined over bins and triangle bins respectively.

There are several approaches one can take at this point. In the case of the power spectrum in real space, the most naive choice would be to compute theoretical predictions at the center of each bin, meaning that the power spectrum at the bin k is equal to the power spectrum evaluated at the center k_c of the bin:

$$P_{\text{c}}(k) = P_{\text{theory}}(k_c), \tag{2.146}$$

where $P_{\text{theory}}(k)$ is the theoretical prediction of the power spectrum, defined over all scales k. However, the moduli of Fourier momenta are not uniformly distributed in a single bin; in the continuous case, a single bin can be approximated as a spherical shell, and therefore one can define an average momentum size in the spherical shell:

$$k_{\text{eff}}^{\text{cont}}(k) = \frac{1}{\int_{\mathbf{q} \in k} \mathrm{d}^3\mathbf{q}} \int_{\mathbf{q} \in k} \mathrm{d}^3\mathbf{q} \, q, \tag{2.147}$$

that can be computed exactly. Since all estimators rely on a discrete Fourier grid, better yet would be to compute the average momentum sizes on a discrete grid, meaning

$$k_{\text{eff}}^{\text{grid}}(k) = \frac{1}{N_P(k)} \sum_{\mathbf{q} \in k} q. \tag{2.148}$$

Given either of the two effective momenta, one can compute an effective power spectrum for the k-bin by simply computing the power spectrum model $P(k)$ at the effective momenta:

$$P_{\text{eff}}^{\text{cont}}(k) = P_{\text{theory}}(k_{\text{eff}}^{\text{cont}}), \tag{2.149}$$

$$P_{\text{eff}}^{\text{grid}}(k) = P_{\text{theory}}(k_{\text{eff}}^{\text{grid}}). \tag{2.150}$$

However, the best way to compute a theoretical prediction to compare with the aforementioned estimators, would be to compute the average of the power spectrum model $P_{\text{theory}}(k)$ over each shell of the discrete grid

$$P_{\text{bin}}^{\text{grid}}(k) = \frac{1}{N_P(k)} \sum_{\mathbf{q} \in k} P_{\text{theory}}(q). \tag{2.151}$$
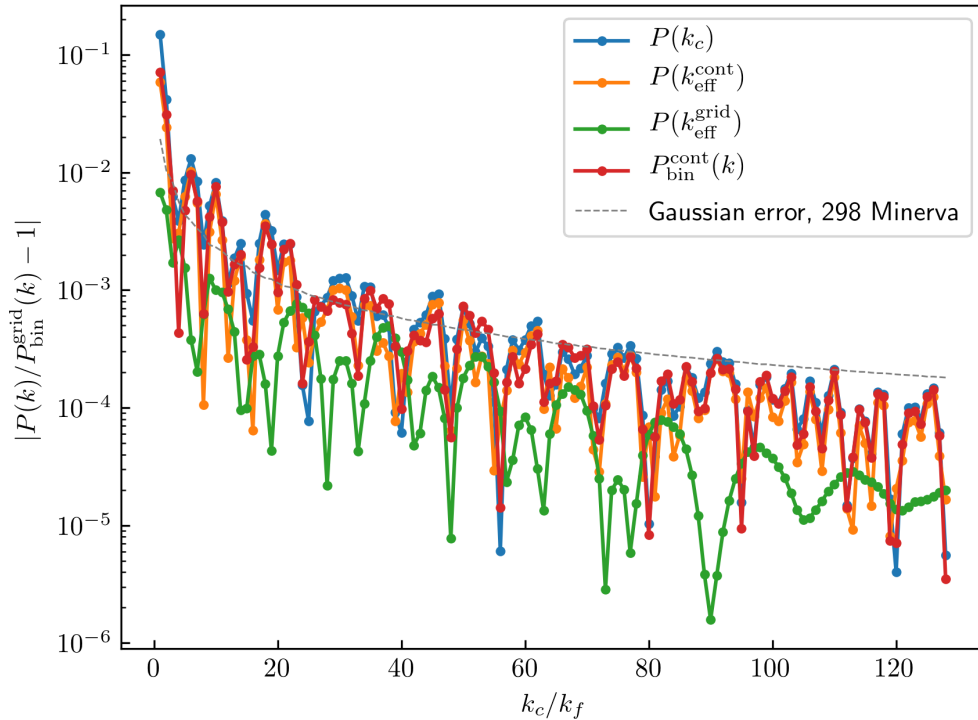
Figure 2.4: Comparison between all the different approaches to compute the theoretical predictions with respect to the discretization effects in the Fourier grid.

Sometimes, the discrete sum can be approximated with an integral

$$P_{\text{bin}}^{\text{cont}}(k) = \frac{1}{\int_{\mathbf{q} \in k} \mathrm{d}^3 \mathbf{q}} \int_{\mathbf{q} \in k} \mathrm{d}^3 \mathbf{q} \, P_{\text{theory}}(q) \tag{2.152}$$

A comparison of all these different ways to compute the theoretical predictions are shown in 2.4 for a model with the linear power spectrum computed from the Minerva cosmology (see section 3.1). Systematic deviations from the $P_{\text{bin}}^{\text{grid}}$ are here compared with the Gaussian prediction for the statistical error relative to a total volume of 298 boxes with $L = 1500 \, \text{Mpc} \, h^{-1}$ (equivalent to the full Minerva set).

Each approach has its own advantages and disadvantages. Evaluating the power spectrum at the bin centers is the fastest and doesn't depend on knowing exactly the binning scheme, since the bin centers are enough. However, this is generally the most inaccurate of all approaches, and can in theory produce systematic uncertainties that can lead to biased results. Effective predictions have the advantage of being fast and more accurate, with the only cost of some *a priori* computation of the effective momenta. This pre-computation makes the method slower, however notice how computing the effective momenta on the grid is more accurate than the approximation over the continuum. Using the full average over the discrete grid gives in general the best agreement, but at a computational cost, scaling with the cube of the size of the Fourier grid, and therefore with k_{max}^3. This can in theory represent a problem when the power spectrum is evaluated a very large number of times – e.g., in Monte Carlo Markov Chains when cosmological parameters are varied – but since this requires a triple sum over the grid coordinates, it is possible to implement it efficiently.

Similar approaches can be developed also in the case of the bispectrum in real space. Here only formulas defined over the discrete grid will be presented, but in theory it is always possible to generalize

them even in the continuum case. The fully averaged bispectrum is taken here as a reference, with

$$B_{\text{bin}}^{\text{grid}}(k_1, k_2, k_3) = \frac{1}{N_B(k_1, k_2, k_3)} \sum_{\mathbf{q}_1 \in k_1} \sum_{\mathbf{q}_2 \in k_2} \sum_{\mathbf{q}_3 \in k_3} \delta_K(\mathbf{q}_{123}) B_{\text{theory}}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3). \qquad (2.153)$$

Before going forward, consider that, under the assumption of isotropy, the bispectrum can be considered as a function of just the moduli of the three Fourier momenta. With this, one can evaluate the bispectrum directly at generic Fourier momenta of moduli given by the center of the triangle bin,

$$B_{\text{c}}(k_1, k_2, k_3) = B_{\text{theory}}(k_1, k_2, k_3), \qquad (2.154)$$

however this approach is not completely well defined, especially for bins whose centers do not form a closed triangle, because the bispectrum evaluated directly at those momenta cannot really be trusted.

A number of different definitions of effective triangles for the bispectrum are possible. One can consider to average the individual momenta over the triangle bin,

$$k_{\text{eff,n}} = \frac{1}{N_B(k_1, k_2, k_3)} \sum_{\mathbf{q}_1 \in k_1} \sum_{\mathbf{q}_2 \in k_2} \sum_{\mathbf{q}_3 \in k_3} \delta_K(\mathbf{q}_{123}) q_n, \qquad n = 1, 2, 3. \qquad (2.155)$$

With this definition, one can notice that an equilateral bin with $k_1 = k_2 = k_3$ will also have $k_{\text{eff,1}} = k_{\text{eff,2}} = k_{\text{eff,3}}$; however, an equilateral bin will not contain exclusively equilateral triangles, and thus one should rightfully expect that the effective momenta should be different. In order to take into account this effect, another proposal for the effective triangles is to sort the Fourier momenta such that

$$k_{\text{eff,l}} = \frac{1}{N_B(k_1, k_2, k_3)} \sum_{\mathbf{q}_1 \in k_1} \sum_{\mathbf{q}_2 \in k_2} \sum_{\mathbf{q}_3 \in k_3} \delta_K(\mathbf{q}_{123}) \max(q_1, q_2, q_3)$$

$$k_{\text{eff,m}} = \frac{1}{N_B(k_1, k_2, k_3)} \sum_{\mathbf{q}_1 \in k_1} \sum_{\mathbf{q}_2 \in k_2} \sum_{\mathbf{q}_3 \in k_3} \delta_K(\mathbf{q}_{123}) \text{med}(q_1, q_2, q_3)$$

$$k_{\text{eff,s}} = \frac{1}{N_B(k_1, k_2, k_3)} \sum_{\mathbf{q}_1 \in k_1} \sum_{\mathbf{q}_2 \in k_2} \sum_{\mathbf{q}_3 \in k_3} \delta_K(\mathbf{q}_{123}) \min(q_1, q_2, q_3). \qquad (2.156)$$

Consider now, for example, the tree-level model for the matter bispectrum in real space in SPT

$$B(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) = 2 F_2(\mathbf{q}_1, \mathbf{q}_2) P_{\text{L}}(q_1) P_{\text{L}}(q_2) + \text{cyc.};$$

one could go even further and consider, along with a definition for effective momenta, also some effective angle-dependent part, for example

$$F_2^{\text{eff}}(k_i, k_j) = \frac{1}{N_B(k_1, k_2, k_3)} \sum_{\mathbf{q}_1 \in k_1} \sum_{\mathbf{q}_2 \in k_2} \sum_{\mathbf{q}_3 \in k_3} \delta_K(\mathbf{q}_{123}) F_2(\mathbf{q}_i, \mathbf{q}_j), \qquad (2.157)$$

where $\mathbf{q}_i$ and $\mathbf{q}_j$ are such that $i \neq j$, and they can either be the sorted or unsorted momenta. In this case, the set of $\{k_{\text{eff,l}}, k_{\text{eff,m}}, k_{\text{eff,s}}, F_2^{\text{eff}}(k_l, k_m), F_2^{\text{eff}}(k_m, k_s), F_2^{\text{eff}}(k_l, k_s)\}$ defines yet another different effective approach for the bispectrum.

Discretization effects become all the more important in redshift space. To obtain the multipoles of the anisotropic power spectrum, one has to fully average the anisotropic power spectrum, weighted
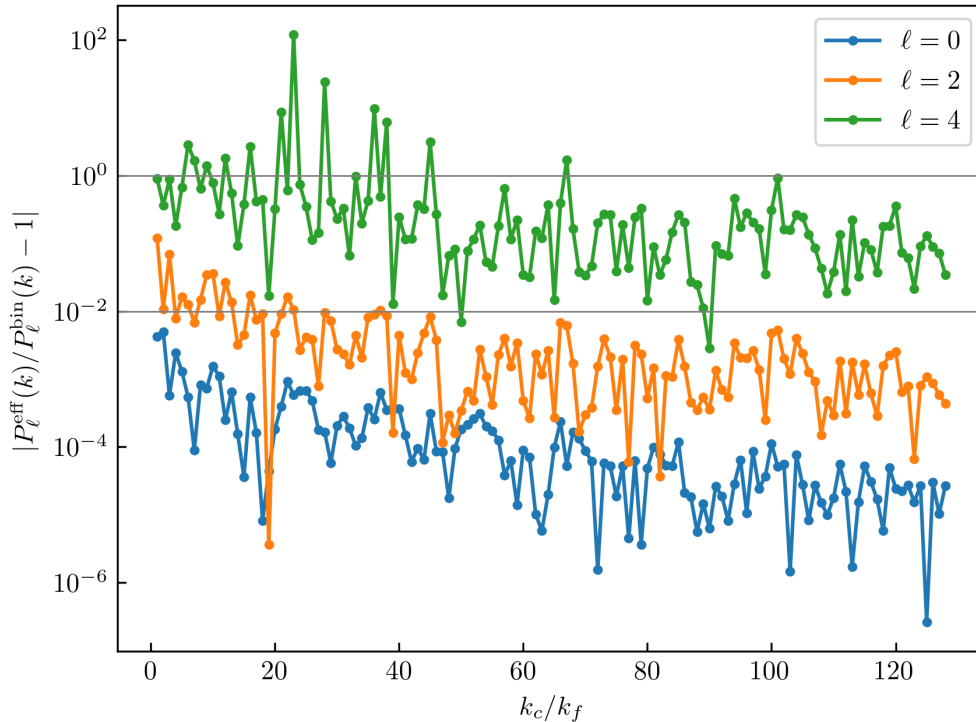
Figure 2.5: Comparison between the theoretical models for the power spectrum multipoles at linear level, evaluated at the effective Fourier momenta and averaged over the discrete Fourier grid.

by the appropriate Legendre polynomial, over the discrete Fourier grid

$$P_\ell^{\text{bin}}(k) = \frac{2\ell + 1}{N_P(k)} \sum_{\mathbf{q} \in k} \mathcal{L}_\ell(\mu_{\mathbf{q}}) P_{\text{theory}}(q, \mu_{\mathbf{q}}). \tag{2.158}$$

Comparing this prediction with the theoretical models for the multipoles evaluated at the effective momenta, figure 2.5, it is evident how for the quadrupole and the hexadecapole the effective prediction become respectively less accurate ($\mathcal{O}(1\%)$), and a very bad approximation, due to a lower amplitude of the multipoles and important discretization effects arising for the Legendre polynomials. Therefore, in redshift-space, performing the full average of the theoretical predictions over the discrete Fourier grid becomes crucial in order to have an accurate comparison of theory and data. Further approaches, similar to the effective evaluation of the prediction but with a gain in accuracy, are still possible; they are reviewed in Appendix B.

2.9   State of the art

While the theoretical framework we have described so far has been developed over the course of the past years, recently it has found renewed interest, especially considering that future missions will survey the redshift of several galaxies over very big volumes.

The upcoming survey *Euclid* will measure shapes and distances of galaxies over $\sim 15\,000\,\text{deg}^2$ of the sky, up to redshift $z \sim 2$, and will be able to measure ~ 30 million spectroscopic redshifts to study galaxy clustering, as well as ~ 2 billion photometric galaxy images for weak lensing observations. The spectroscopic survey will consist of four redshift bins, ranging from $z \simeq 0.9$ to $z \simeq 1.8$, with an observed number density of objects of a few 10^{-4} for each redshift bin [84].

41

The state of the art on the analysis of galaxy clustering actual measurements is represented by the analysis of the data from the Baryon Oscillation Spectroscopic Survey (*BOSS*, [18]).

One of the first assessments of the impact of bispectrum data onto the power spectrum analysis in the *BOSS* clustering measurements is presented in [33]; in that work, the authors use approximate theoretical prediction to model the monopole of the redshift-space power spectrum and of the bispectrum of galaxies, and assume a local-Lagrangian evolution of the tidal bias parameters. While not being able to completely break the degeneracy between the bias parameters, the growth rate f, and the amplitude of dark-matter density fluctuations σ_8, adding information either from different datasets (CMB experiments) or different correlators (the 2PCF in configuration space) can break the $f - \sigma_8$ degeneracy and provide cosmological information [34].

The analysis in [35] includes to the analysis the quadrupole of the galaxy power spectrum and more triangle configurations of the monopole of the galaxy bispectrum; the authors assume the same theoretical model of previous analyses (one-loop model for the power spectrum, tree-level for the bispectrum), and introduce the possibility of having one shot-noise parameter for the deviation from the Poisson prediction, which is the same between power spectrum and bispectrum (in the notation of this text, this means $\alpha_P = \alpha_1 = \alpha_2$).

Since bispectrum measurements consist of a large number of triangle bins, and thus covariance matrices are not easy to estimate from a sufficient number of mocks, compression methods have been introduced and compared to results from the standard bispectrum. In [106], different compression methods are compared and their performances seem to indicate that compressing the information contained in a large number of bins (that otherwise would not be possible to analyze) provides a great improvement in parameter constraints.

In [36], the detection of the BAO feature in the galaxy bispectrum from *BOSS* has been claimed: the fit to the bispectrum data was consistent with the fit to the power spectrum, and even a joint fit of the two correlators gave consistent results. However, in this work the power spectrum was modelled through a phenomenological fit, and the bispectrum prediction lacked of any non-tidal bias.

A more consistent analysis of the redshift-space power spectrum was performed in [107], where a full, gRPT model [108] for the one-loop anisotropic power spectrum was applied, along with a consistent bias expansion, albeit always considering a local-Lagrangian relation for the second-order tidal bias.

Finally, the first fully consistent analyses of the *BOSS* clustering data have been performed in [48, 16]. Both groups employ a fully consistent, perturbation-theory based model for the anisotropic, redshift-space power spectrum, based on the EFTofLSS, including infrared resummation, which is the same described in this text. In both cases, the cosmological parameter space has been sampled using the common MCMC integration method, with a full computation of the theoretical model at each step, while in most previous analysis the cosmology was fixed to a fiducial one and deviations from the fiducial cosmology were quantified using the Alcock-Paczynski parametrization [109]. In particular, [48] find cosmological parameters completely independently from CMB experiments – without assuming any prior that is based on CMB observations – with an uncertainty on the Hubble parameter H_0 comparable to that of Planck.

Our final goal is to employ both theoretical and technical frameworks that are *on par* with the analysis of [48, 16], but with the important inclusion of the galaxy bispectrum. This requires a full-modelling of the theoretical prediction of power spectrum and bispectrum in redshift-space, and in

order to compute the models consistently at each re-evaluation of the cosmological parameters a fast implementation of loop-corrections and binning of the theoretical predictions is needed.

Chapter 3

Data from N-body simulations and mock catalogs

In this chapter, I describe the data set used for the main likelihood analysis, consisting of measurements of power spectrum and bispectrum of dark-matter halos, extracted from N-body simulations. I also use power spectrum and bispectrum measurements of dark-matter halos from mock halo catalogs to estimate the covariance matrices used in the main analyses. I also consider measurements of the matter power spectrum and of the halo-matter cross-power spectrum in order to measure the values of some of the parameters fitted in the likelihood analysis, as a reference to assess the recovery of unbiased estimates of such parameters.

3.1   Minerva simulations

My collaborators and I base our work on the Minerva set of 298 N-body simulations, first presented in [64]. Each run evolves the positions and velocities of 1000^3 dark-matter particles in a periodic cubic box of side $L = 1500\,h^{-1}\,\mathrm{Mpc}$ using the `GADGET-2` code[1] [110]. The flat ΛCDM background cosmology is determined by the dimensionless Hubble parameter $h = 0.695$, the total matter density $\Omega_m = 0.285$, and the baryon density $\Omega_b = 0.046$. This choice corresponds to the best-fit of the combined analysis of the WMAP results and BOSS DR9 results presented in Table I of [111]. The particle mass in the simulations is thus $m_p \simeq 2.67 \times 10^{11}\,h^{-1}\,\mathrm{M}_\odot$. Initial density and velocity perturbations are generated at redshift $z_{\mathrm{in}} = 63$ by displacing the simulation particles from a regular grid using second-order Lagrangian perturbation theory (LPT). The transfer function for the Gaussian linear fluctuations in the matter density is computed with the `CAMB` code [87] assuming a primordial scalar spectral index of $n_s = 0.9632$ and a r.m.s. matter density fluctuation averaged over spheres of radius equal to $8\,h^{-1}\,\mathrm{Mpc}$ (linearly extrapolated to $z = 0$) of $\sigma_8 = 0.828$.

Dark-matter halos are identified using a standard friends-of-friends algorithm with a linking length of 0.2 times the mean one-dimensional interparticle separation. Unbound particles are removed using the `SUBFIND` code [112]. We only consider halos that contain at least 42 particles, corresponding to a minumum mass of $M \simeq 1.12 \times 10^{13}\,h^{-1}\,\mathrm{M}_\odot$. For simplicity, we limit our study to a single output at redshift $z = 1$, as this value is of particular relevance for upcoming spectroscopic galaxy surveys such as Euclid [20] or DESI [113]. The mean number density for the resulting halo population is of
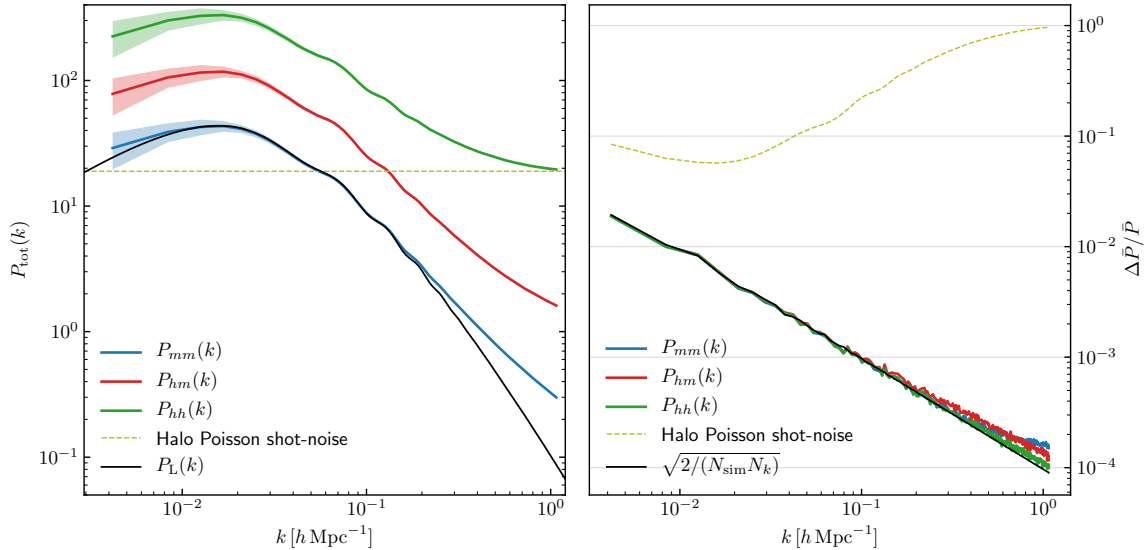
---

[1]http://www.gadgetcode.org/

Figure 3.1: Measurements of the power spectra in real space of the Minerva simulations at redshift $z = 1$. The left panel shows the averaged matter power spectrum (solid blue line), the halo-matter cross power spectrum (solid red), and the total halo power spectrum including shot noise (solid green). Colored shaded areas represent the standard deviations of the measurements for each value of k. Also shown are the linear power spectrum of the Minerva cosmology (solid black) and the level of the Poisson shot-noise for the halo power spectrum (dashed olive). The right panel shows the ratio between the standard error on the mean and the average measurement of each power spectrum. Also shown are the Gaussian prediction for this ratio (solid black line) and the ratio between the Poisson shot noise of the halo power spectrum and the average halo power spectrum (dashed olive line).

$2.13 \times 10^{-4} \, h^3 \, \mathrm{Mpc}^{-3}$.

We measure the matter, halo, and cross power spectra, as well as halo and matter bispectra from estimates of the Fourier-space overdensity $\delta_{\mathbf{k}}$ on a grid of linear size 256, obtained with the PowerI4 code[2] described in [105]. The power spectrum and bispectrum estimators used to perform the measurements are thoroughly described in section 2.8.

3.1.1 Power spectrum measurements

Figure 3.1 shows the measurements of the matter power spectrum $P_{mm}(k)$, the halo-matter cross power spectrum $P_{hm}(k)$, and the total halo power spectrum $P_{hh}(k)$ for 256 bins of width equal to the fundamental frequency of the box, $k_f = 2\pi/L \simeq 0.0042 \, h \, \mathrm{Mpc}$. In particular, we show the average measurements of the whole set of 298 Minerva simulations. The linear power spectrum is also shown as a reference.

The right panel shows the standard errors on the mean, of the three power spectra. Due to the large number of simulations we consider, the average of our measurements reaches subpercent precision in the determination of the halo power spectrum even on the relatively large scales we are interested in. Notice that this is of the same order of possible systematic effects introduced by the set-up and implementation of the N-body solver (see *e.g.* [114] for a quantification of systematic errors on the matter power spectrum). In the case of the halo power spectrum, the uncertainties are always much lower than the Poisson shot-noise. For this reason, in the following we consider the shot-noise as part of the signal and fit it together with the physical power spectrum.

In figure 3.2 we show a comparison of the halo power spectrum measured with three different
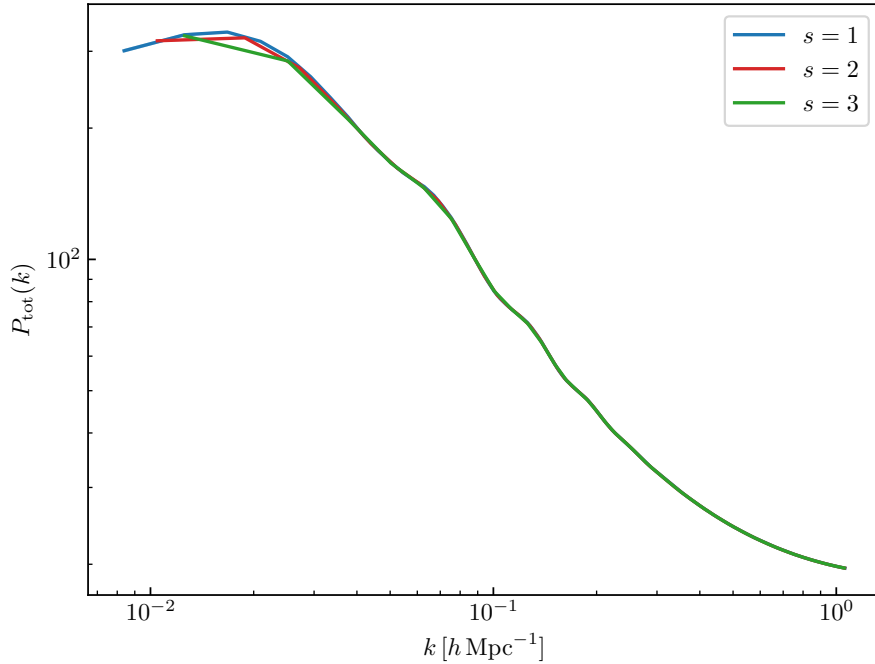

Figure 3.2: Comparison between measurements of the halo power spectrum of the Minerva simulations with different binning schemes.

binning schemes: following the notation of subsection 2.8, we take $(s = 1, c = 2)$ in blue, $(s = 2, c = 2.5)$ in red, $(s = 3, c = 3)$ in green. These binning schemes are defined in such a way that, when Fourier modes are binned in one-dimensional bins based on their modulus, there are values of k_{max} for which all the modes with $k < k_{\mathrm{max}}$ are accounted for, regardless of the scheme.

3.1.2 Direct measurements of bias and stochasticity

Using the measurements from the matter and halo auto-spectra and the halo-matter cross-spectrum, we can estimate the values of the linear bias b_1 and of the large-scale shot-noise correction parameter α_P.

In a very schematic way, starting from equations 2.78, 2.100, and 2.92, we can sketch the theoretical models of these three power spectra at one loop in the following way:

$$P_{mm}(k) = P_{\mathrm{L}}(k) + \Delta P^m_{1-\mathrm{loop}}(k) + P_{\mathrm{ct}}(k), \tag{3.1}$$

$$P_{hm}(k) = b_1 P_{mm}(k) + \Delta P^{\times}_{1-\mathrm{loop}}(k) - b_{\nabla^2} k^2 P_{\mathrm{L}}(k), \tag{3.2}$$

$$P_{hh}(k) = b_1^2 P_{mm}(k) + 2b_1 \left[\Delta P^{\times}_{1-\mathrm{loop}}(k) - b_{\nabla^2} k^2 P_{\mathrm{L}}(k) \right] + \Delta P^h_{1-\mathrm{loop}}(k) + \frac{1 + \alpha_P}{\bar{n}}. \tag{3.3}$$

We can then define our estimate of the linear bias $\hat{b}_1$ as the large-scale limit of the ratio between the halo-matter cross-spectrum and the matter auto-spectrum,

$$\hat{b}_1 = \lim_{k \to 0} \frac{P_{hm}(k)}{P_{mm}(k)} = \lim_{k \to 0} \left[b_1 + \frac{\Delta P^{\times}_{1-loop}(k) - b_{\nabla^2} k^2 P_{\mathrm{L}}(k)}{P_{mm}(k)} \right]; \tag{3.4}$$

therefore, starting from measurements, we can either fit the large scale limit with a constant in order to estimate the value of the linear bias, or we can make assumptions on the k-dependent behaviour of
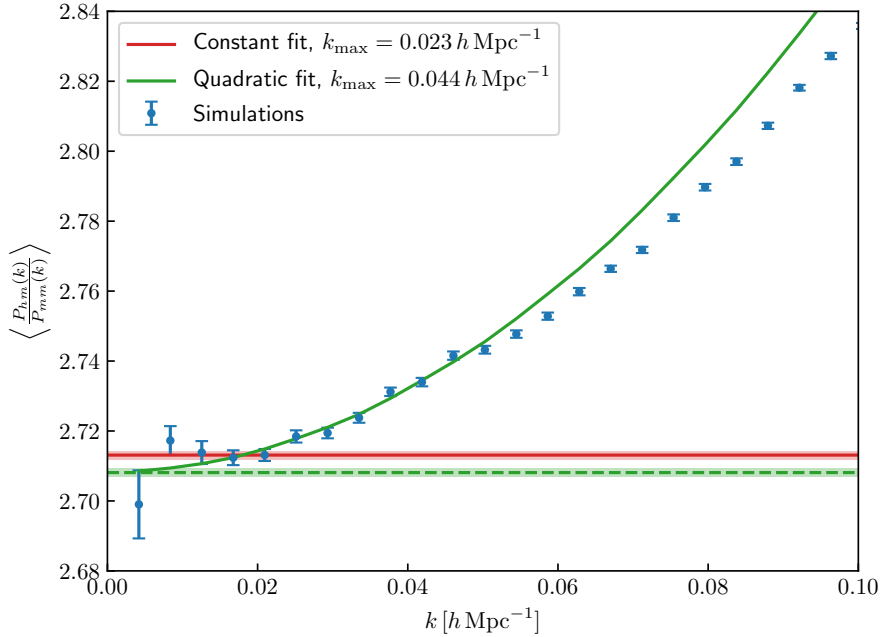
46

Figure 3.3: Average of the ratios of the cross halo-matter power spectrum and the matter power spectrum over the full set of Minerva simulations. The red line is the best fit of the data up to $k_{max} = 0.023\,h\,Mpc^{-1}$ with a constant function; the solid green line is the best fit of the data up to $k_{max} = 0.044\,h\,Mpc^{-1}$ with a constant plus a k^2-dependent term. The green dashed line shows the best-fit value for $\hat{b}_1$ with this second fit function.

the ratio of the higher-order terms in $P_{hm}(k)$ and the full $P_{mm}(k)$. In figure 3.3 we show the ratios in equation 3.4 averaged over the full Minerva set. A fit of these data points can be very sensitive to both the functional form of the fit and the k_{max} chosen for the fit, meaning that a direct measurement of the linear bias in this way is not extremely robust. This is mainly due to the important fact that, given the size of the statistical uncertainties coming from such a large set of N-body simulation, the scale-dependence of the ratio is evident at all scales considered, and this behaviour has to be accounted for in the modelling of this ratio, at least partially, as well as in the value of k_{max} used to perform the fit. In particular, we choose two different possible functions: a constant, that we have fitted up to $k_{max} = 0.023\,h\,\mathrm{Mpc}^{-1}$, and a constant plus a k^2-dependent term, that we have fitted up to $k_{max} = 0.044\,h\,\mathrm{Mpc}^{-1}$. To perform these fits, we use χ^2-minimization. The two different values of b_1 are clearly inconsistent with each other: in the case of the constant fit, we find $\hat{b}_1 = 2.7131 \pm 0.0012$, while with the quadratic fit we find $\hat{b}_1 = 2.7081 \pm 0.0012$, translating into a tension between the two measurements almost at the 3σ level.

Similarly, it is possible to write down α_P as a large-scale limit of the stochasticity defined as in [115]; using the sketched models of the equations above, and assuming that we have an estimate $\hat{b}_1$ of the linear bias, we have

$$P_{hh}(k) - 2\hat{b}_1 P_{hm}(k) + \hat{b}_1^2 P_{mm}(k) = \frac{1 + \alpha_P}{\bar{n}} + \Delta P^h_{1-\mathrm{loop}}(k). \tag{3.5}$$

The corresponding plot is shown in figure 3.4, where $\hat{b}_1 = 2.7081$ has been assumed; the best-fit values for the stochasticity are $\hat{\alpha}_P = -0.3020 \pm 0.0020$ for the constant fit and $\hat{\alpha}_P = -0.3052 \pm 0.0020$ for the quadratic fit. This time, the two values are actually consistent within 1σ.
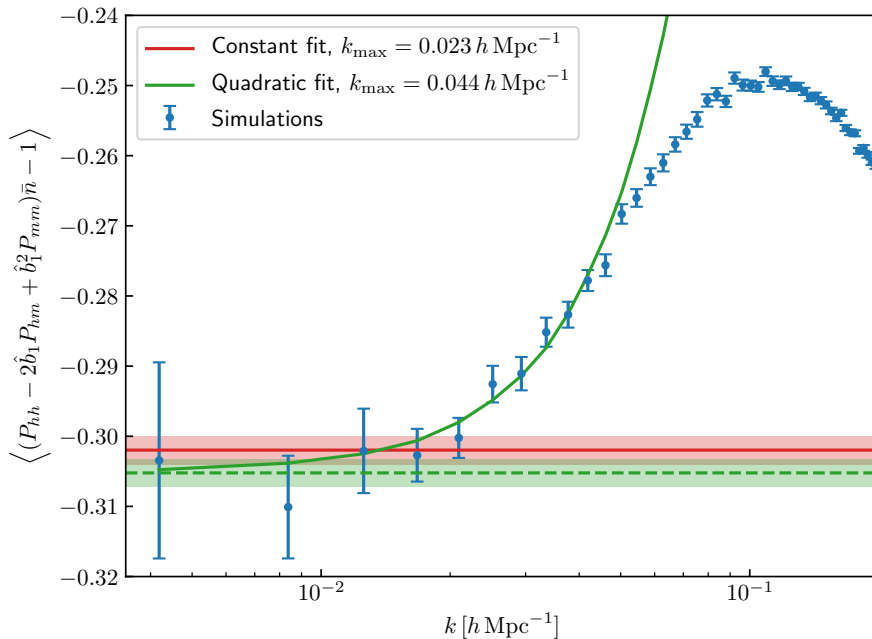
Figure 3.4: Average of the stochasticity over the full set of Minerva simulation assuming $b_1 = \hat{b}_1$ from the quadratic fit. The red line is the best fit of the data up to $k_{max} = 0.023\,h\,\mathrm{Mpc}^{-1}$ with a constant function; the solid green line is the best fit of the data up to $k_{max} = 0.044\,h\,\mathrm{Mpc}^{-1}$ with a constant plus a k^2-dependent term. The green dashed line shows the best-fit value for $\hat{\alpha}_P$ with this second fit function.

3.1.3 Bispectrum measurements

Figure 3.5 shows the mean of *all* measured triangle configurations and the relative error on the mean for the three binning schemes. The order of the triangles appearing in these plots (and in several others in the rest of the text) matches the ordering defined in section 2.8, and corresponds to increasing the value of the sides $\{k_1, k_2, k_3\}$ with the constraint $k_1 \geq k_2 \geq k_3$. The ticks on the horizontal axis and the corresponding vertical lines mark equilateral configurations where the value of k_1 changes. It follows that, in between two ticks, all points correspond to triangles with the same k_1, while k_2 and k_3 assume all allowed values. Again, the Poisson shot-noise contribution is shown with a dashed line in the upper half of each panel while its relative size appears in the bottom half.

3.2 Pinocchio mock catalogs

In addition to the full N-body simulations, we make use of a larger set of $10,000$ mock halo catalogs generated with the PINOCCHIO code [65, 66, 67] using the same cosmological model and simulation settings. PINOCCHIO uses third-order Lagrangian Perturbation Theory to shift matter particles and relies on a set of criteria, based on the ellipsoidal-collapse model, to group them into halos. Note that 298 PINOCCHIO realisations have been obtained using the same random seeds for the initial conditions of the Minerva simulations, thereby allowing a one-to-one comparison not affected by sample variance. Both the N-body simulations and the PINOCCHIO mocks have been introduced and used in a series of papers [116, 117, 118] aimed at comparing several methods for the production of approximate catalogs in terms of their predictions for the 2-point correlation function, power spectrum, and bispectrum, along with their respective covariance properties. We refer the reader to these works for a first assessment of the accuracy of the PINOCCHIO mocks and to [119] for a general review of approximate
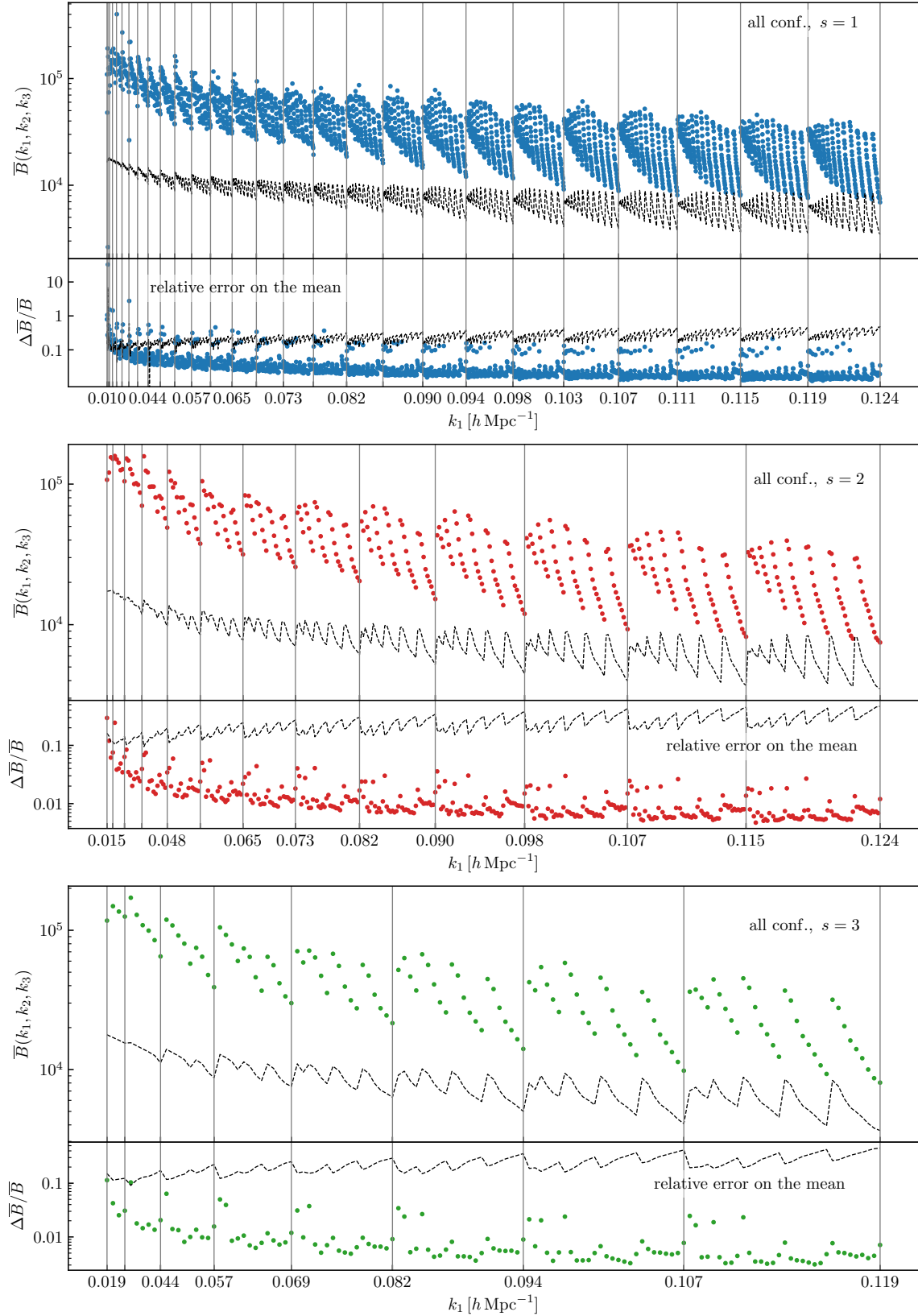
Figure 3.5: Measurements of the halo bispectrum from the N-body simulations for the three binning schemes. In each panel, the bottom plot shows the relative error on the mean. Dotted black lines show the absolute (top) and relative (bottom) Poisson shot-noise contribution to the measurements.

methods.

The main results of references [116, 117, 118] are derived defining the mock halo catalogs in terms of a mass threshold that provides the same halo number density, on average, as the reference N-body simulations. We will make here a different choice and set the mass threshold by matching the clustering amplitude of the halos in real space as determined by the *total* halo power spectrum, including shot noise, eq. (2.141), since this is the quantity determining the power spectrum and bispectrum Gaussian variance. The mass threshold, in this case, controls the overall amplitude via both the linear bias and the number density. This matching is crucial to minimize any systematic difference between the covariance matrices for the power spectrum and the bispectrum extracted from the PINOCCHIO mocks and the N-body simulations. In fact, as shown in eq. (2.144) and eq. (2.145), the Gaussian contribution to the error on the power spectrum and bispectrum measurements, representing the leading term at large scales, depends essentially on $P_{\mathrm{tot}}(k)$.

The solid lines in the left-hand-side panel of figure 3.6 show the ratio between the mean $P_{\mathrm{tot}}(k)$ estimated from the PINOCCHIO mocks and from the N-body simulations. This comparison is limited to the 298 realisations sharing the same initial conditions. Different binning schemes are represented with different colors and the shaded regions denote the corresponding error on the mean of the mocks measurements. The total power in the PINOCCHIO mocks matches the result from the N-body simulations to better than one percent up to at least $k \sim 0.12\,h\,\mathrm{Mpc}^{-1}$. These are the scales we are interested in for the bispectrum. Notice that the *shot-noise subtracted* power spectra show, instead, a discrepancy of about 2-3% at the largest scales, compensated by a similar discrepancy in the shot-noise contribution. The ratio between the power-spectrum variance estimated from the PINOCCHIO mocks and from the N-body simulations is shown in the right-hand-side panel of figure 3.6, again only using the 298 realisations with the same initial seeds. The ratio scatters around one, with a few percent deviations.

Figure 3.7 shows a similar comparison but in terms of the bispectrum and its variance. The left-hand-side panels show the ratio between the mean bispectra measured from the PINOCCHIO mocks and from the N-body simulations for the three binning schemes. In this case, we consider measurements corrected for Poisson shot-noise, as in eq. (2.142). Notice that the bispectrum from the mocks is suppressed with respect to the one from the simulations by about 6-7% with some dependence on the triangle shape that follows from the discrepancy in the large-scale power spectra mentioned above, in addition to the small-scale suppression due to LPT displacements. The right-hand-side panels show that the variance of the bispectrum in the PINOCCHIO mocks reproduces that in the N-body simulations with a scatter of about 10% but with no significant systematic error, except for a slight suppression at the few percent level visible in the $s = 3$ measurements.

Figure 3.8 shows the correlation matrix

$$r_{ij} \equiv \frac{C_{ij}}{\sqrt{C_{ii}\,C_{jj}}}\,, \tag{3.6}$$

of the power spectrum measurements of the PINOCCHIO mocks, as well as of the N-body simulations, for each of the binning schemes used: the left panel corresponds to the correlation matrix for $s = 1$, the middle panel to $s = 2$, and the right panel to $s = 3$. Considering the matrices estimated from the PINOCCHIO mocks, at smaller scales, cross-correlations between different Fourier bins become more relevant, mainly because of two effects that become more important in the non-linear regime: the
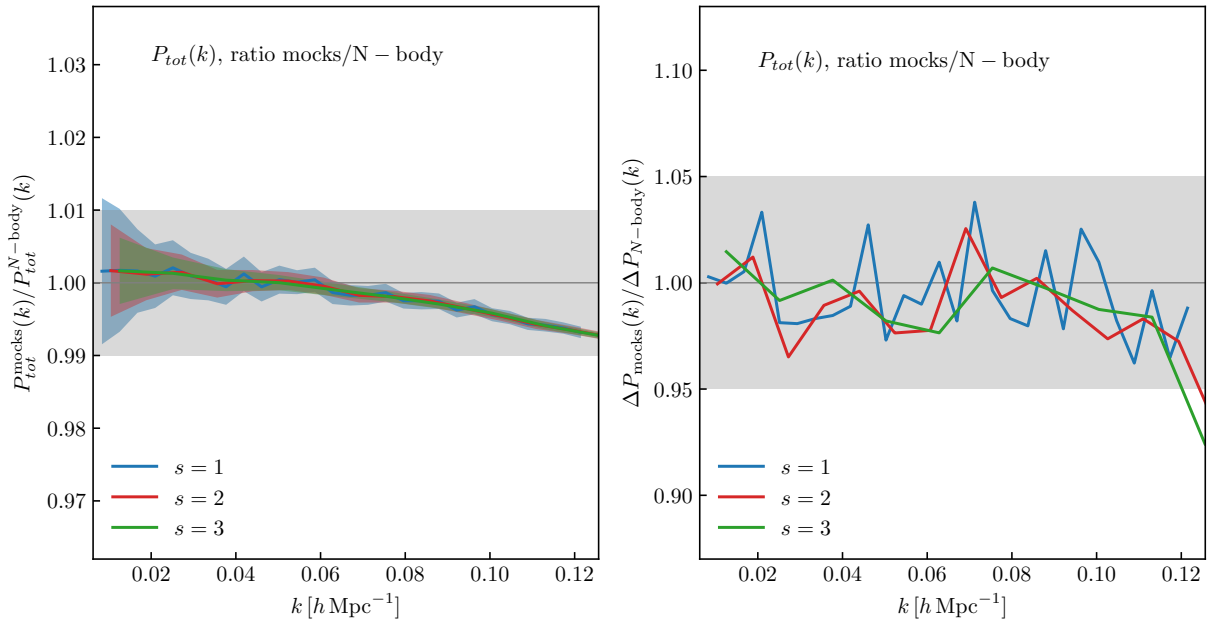
Figure 3.6: Left: ratio of the mean total halo power spectrum estimated from the PINOCCHIO mocks to the same quantity measured from the N-body simulations for the 298 realisations sharing the same initial conditions. Different colors represent the different binning schemes in k, with the shaded regions denoting the corresponding error on the mean. The total power spectra are consistent with each other at the percent level at least up to $0.12\,h\,\mathrm{Mpc}^{-1}$. Right: ratio between the standard deviations estimated from the PINOCCHIO mocks and from the N-body simulations, again limited to the 298 realisations sharing the initial seeds. The standard deviations are consistent at the ~ 5 percent level up to $0.12\,h\,\mathrm{Mpc}^{-1}$.

shot-noise contribution to the total power spectrum, and the non-Gaussian contribution to the power spectrum covariance, depending at leading order on the total halo trispectrum. The cross-correlations from the Minerva instead show a very noisy signal, due to the fact that covariance matrices have been estimated by a smaller number (by a factor ~ 30) of measurements with respect to the PINOCCHIO measurements.

Figure 3.9 shows the cross-correlation matrix of the bispectrum measurements extracted from the PINOCCHIO and the N-body simulations using bins with $s = 1$. In particular, the top four panels represent the four corners of the r_{ij} matrix as defined by the ordering described above in eq. (2.138), illustrating the correlation properties of large-scale and small-scale triangles. Each row and column corresponds to a triangle configuration whose sides are given in units of the fundamental wavenumber as a triplet of integers. The smallest-scale triangle in this case is $\{29, 29, 29\}k_f = \{0.12, 0.12, 0.12\}\,h\,\mathrm{Mpc}^{-1}$. Matrix elements above the diagonal are estimated from the 10,000 mocks, while those below are estimated from the 298 simulations. Clearly, both covariances appear to be dominated by diagonal terms but show very different noise levels for the off-diagonal elements. The bottom panels compare some rows of the matrix r_{ij} (that have been highlighted in the top plots with red borders) obtained from different datasets. They allow a more direct and quantitative comparison of the results from the 10,000 PINOCCHIO mocks (red lines), the 298 N-body simulations (blue lines), and the corresponding 298 mocks with matched initial conditions (green lines). Notice that the noise in the r_{ij} elements extracted from this last dataset reproduces quite closely the noise coming from simulations. On the other hand, the off-diagonal cross-correlations from the full set of mocks are very close to zero, with differences well below the 5% level.

It would be misleading, based on this simple inspection, to conclude that off-diagonal contributions
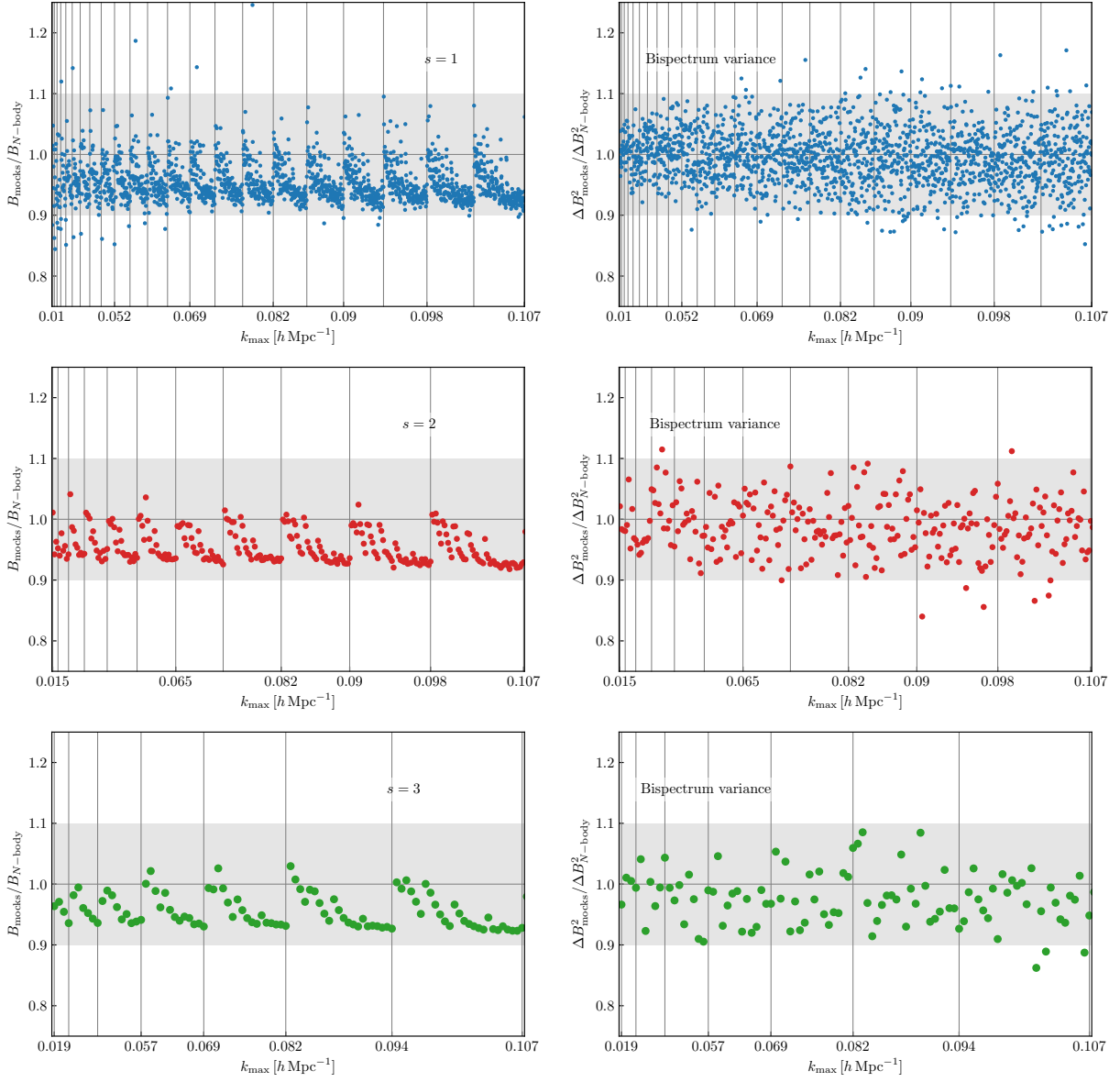
Figure 3.7: The left-hand-side panels display the ratio of the mean bispectrum measured from the PINOCCHIO mocks (with Poisson shot-noise subtracted) to the corresponding quantity measured from the N-body simulations for the three binning schemes adopted in this work ($s = 1, 2$ and 3 from top to bottom). The right-hand side panels show the ratio of the bispectrum variance measured in the mocks to the one measured from the simulations. Both comparisons consider only the 298 realisations with matched initial conditions.

to the covariance matrix of the bispectrum can be safely ignored. A counterexample is shown in figure 3.10 where we consider bins with $s = 3$. Here the top panels show again the four corners of the covariance matrix up to the same maximum wavenumber, with the last triangles being $\{8, 8, 8\}\Delta k = \{24, 24, 24\}k_f = \{0.1, 0.1, 0.1\}\, h\,\mathrm{Mpc}^{-1}$. Several off-diagonal elements of the matrix r_{ij} assume values of 10-20%. They correspond to different triangles that share one or more sides and it is expected that the bispectrum covariance receives contributions from the non-Gaussian terms in this case. These features are present as well in the covariance estimated from the 298 simulations, but the noise affecting them is also of the order of 10-20%. Correlations among different triangles are present for any binning scheme and in the same overall amount.

In order to quantify the importance of off-diagonal terms, let us consider the cumulative signal-to-
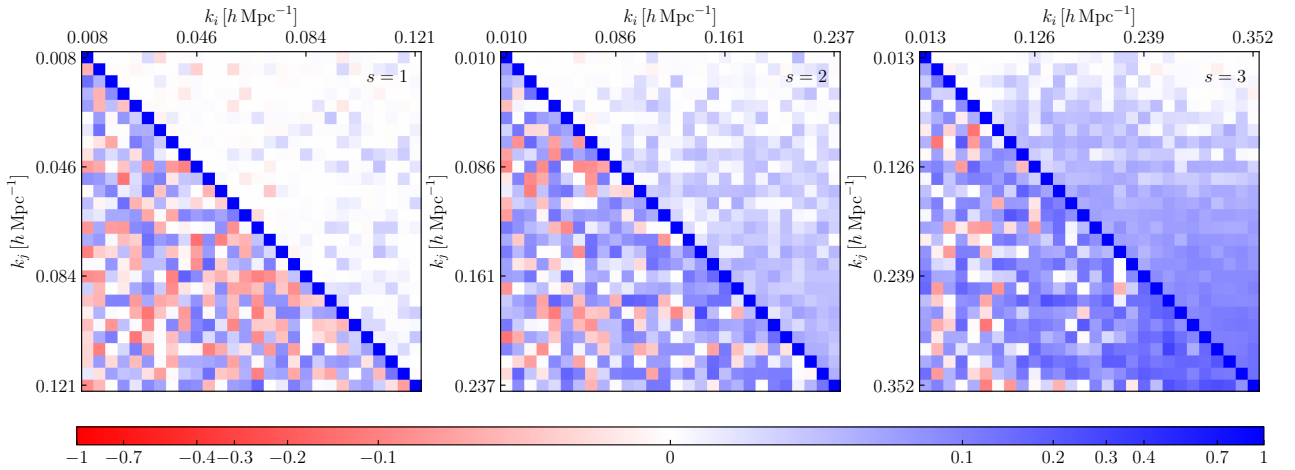
Figure 3.8: Correlation matrices for the power spectrum measurements from the mock catalogs (upper triangular matrices) and from the N-body simulations (lower triangular matrices). The three panels show the correlation for the three different binning schemes: $s = 1$ in the left panel, $s = 2$ in the middle panel, $s = 3$ in the right panel.

noise ratio for the bispectrum defined as

$$\left(\frac{S}{N}\right)^2 = \sum_{i,j}^{N_{\mathrm{t}}(k_{\max})} B_i \, C_{ij}^{-1} \, B_j \, , \tag{3.7}$$

where the indices i and j run over all triangle configurations having no sides larger than $k_{\max}$.

It is interesting to see how this quantity changes if one neglects the contribution from off-diagonal terms in the covariance. This test is performed in the left-hand-side panel of figure 3.11 where we use the "signal" and the "noise" extracted from the 10,000 mocks. In this case we can expect a sufficiently precise determination of the covariance matrix (and its inverse), with a residual statistical error on the order of a few percent. Here the solid and dashed lines represent, respectively, the results using the full covariance and the diagonal part alone. Note that off-diagonal terms become increasingly more important as $k_{\max}$ grows, causing a reduction of $(S/N)^2$ by a factor of two at $k_{\max} \sim 0.08 \, h \, \mathrm{Mpc}^{-1}$. In the right-hand-side panel, instead, we repeat the test by using the bispectrum measurements B_i and their variance ΔB_i from the N-body simulations. In this case, the covariance matrix is obtained using the relation $C_{ij} = \Delta B_i \, \Delta B_j \, r_{ij}$ where the cross-correlation matrix is estimated from the 10,000 mocks. The results are essentially the same as in the left-hand-side panel.

Figure 3.12 shows the correlation matrix for the measurements of power spectrum and bispectrum with $s = 2$ – limited here to $k_{\max,\mathrm{P}} = 0.24 \, h \, \mathrm{Mpc}^{-1}$ and $k_{\max,\mathrm{B}} = 0.1 \, h \, \mathrm{Mpc}^{-1}$ – including cross correlations between power spectrum Fourier bins and bispectrum triangle bins. The off-diagonal blocks show relatively large cross-correlations between the power spectrum Fourier modes and the bispectrum triangle bins, that for some configurations reach 30 or even 40%.

In order to quantify the importance of these cross-correlations, let us consider now the cumulative signal-to-noise ratio for the power spectrum and bispectrum measurements,

$$\left(\frac{S}{N}\right)^2_{P \oplus B} = \sum_{i,j} X_i \, \left[C_{ij}^X\right]^{-1} X_j \, , \tag{3.8}$$

where here X is the combination of power spectrum and bispectrum measurements, $\mathbb{C}^X$ fo components C_{ij}^X is the full covariance including also cross-correlations between the two correlation functions, and
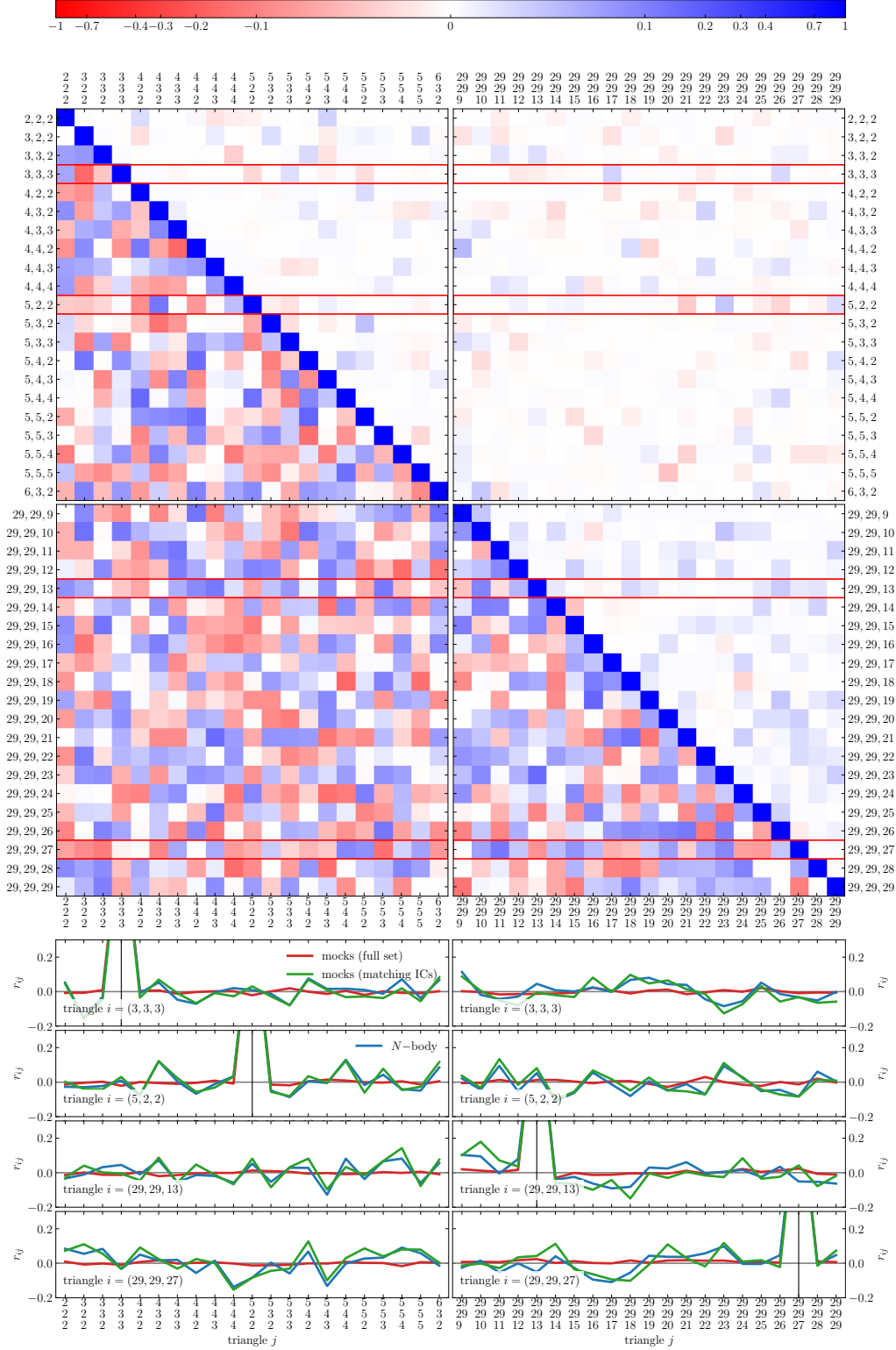
Figure 3.9: The cross-correlation coefficients r_{ij}, eq. (3.6), of the bispectrum covariance matrix estimated from the 10,000 PINOCCHIO mocks compared with those estimated from the 298 N-body simulations for the binning with $s = 1$. The top four panels represent the four corners of the matrix as defined by the ordering described above in eq. (2.138). Each row and column corresponds to a triangle whose sides are given in units of the fundamental wavenumber as a triplet of integers. The smallest-scale triangle in this case is $\{29, 29, 29\} k_f = \{0.12, 0.12, 0.12\} \, h \, \mathrm{Mpc}^{-1}$. Matrix elements above the diagonal are estimated from the 10,000 mocks, while those below are estimated from the 298 simulations. Bottom panels show the rows marked above with red borders, comparing the estimates from 298 simulations (blue), from the 298 mocks with matching seeds (green) and from the full set of 10,000 mocks (red).

54

Figure 3.10: As in figure 3.9 but for $s = 3$. Off-diagonal correlations become more evident with a binning scheme with larger s.
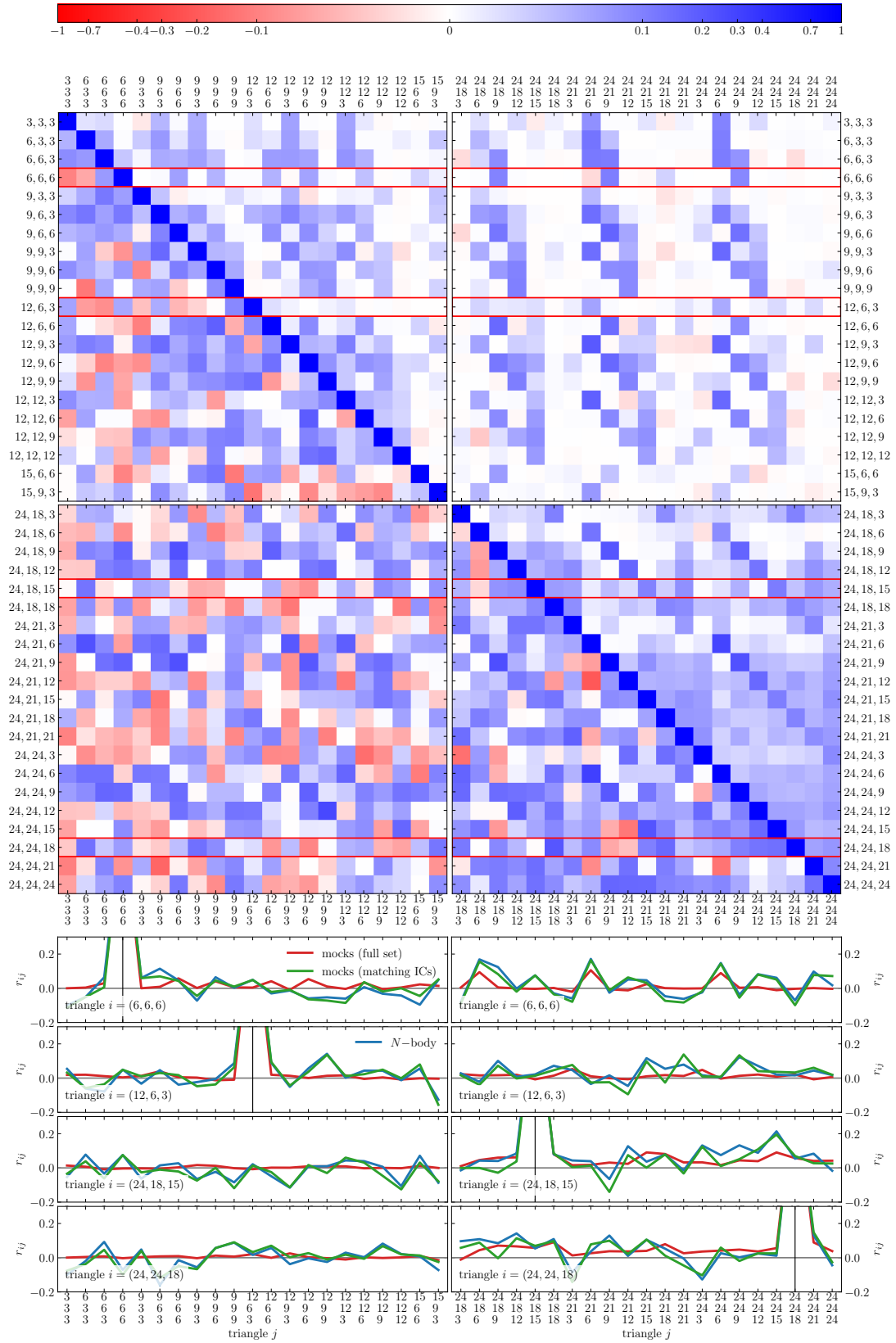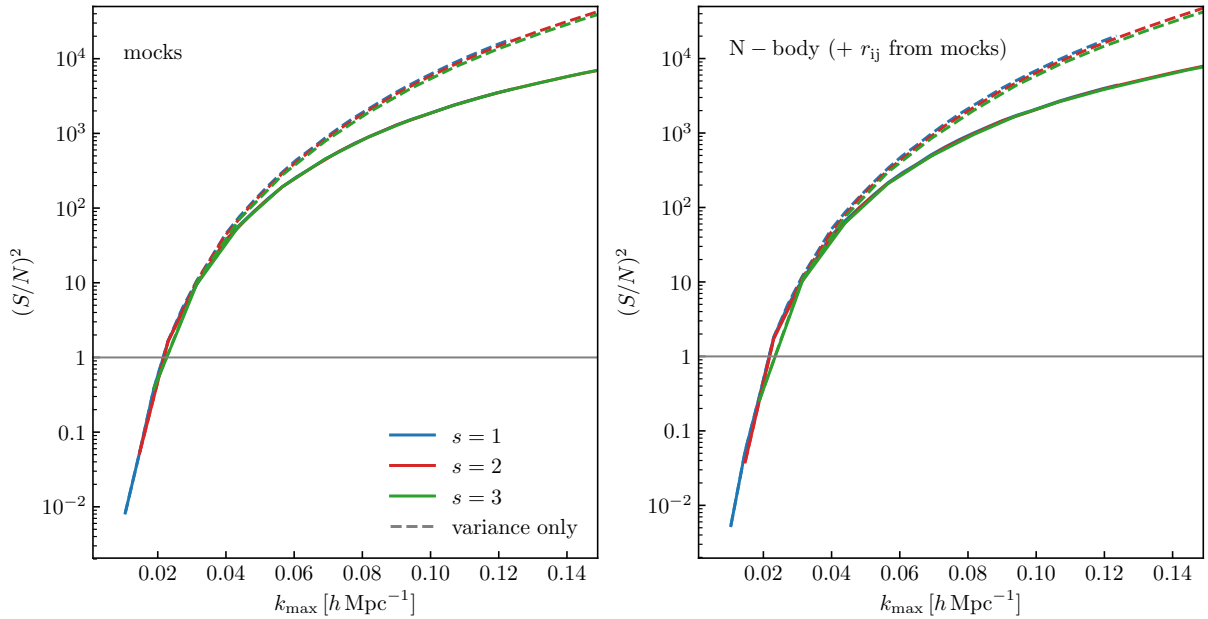
Figure 3.11: Cumulative signal-to-noise ratio for the bispectrum measured from the PINOCCHIO mocks (left panel) and from the N-body simulations (right panel). In the latter case, the full covariance matrix is estimated from the variance measured from the N-body runs and the cross-correlation coefficients r_{ij} estimated again from the PINOCCHIO mocks, since a direct measurement of the covariance matrix only from the N-body simulations would be too noisy. The solid lines are obtained considering the full covariance matrix (for the simulations we use the cross-correlation coefficients estimated from the 10,000 mocks). For comparison, we also show results obtained using only the diagonal part of the covariance matrix (dashed lines). Results for the three binning schemes we adopt are shown with different colors. As expected, they almost perfectly coincide.

where the sum runs over both power spectrum Fourier bins and bispectrum triangle bins. This signal-to-noise ratio is shown in figure 3.13, where we show also the case where the covariance is intended either diagonal, and the case where cross-correlations between power spectrum and bispectrum are instead set to zero – and in this case the signal-to-noise ratio is just the additive sum of the signal-to-noise ratios of each correlation function. In particular, we use the average of the measurements from the N-body simulations; the case with a diagonal covariance is shown in blue, the case where the covariance has no cross-correlations is shown in red, while the case with the full covariance is shown in green. The signal-to-noise ratio is shown as a function of the $k_{\max}$ of the measurements, assumed to be equal for power spectrum and bispectrum $k_{\max} = k_{\max,\mathrm{P}} = k_{\max,\mathrm{B}}$. Differences are important, especially comparing the case with the full covariance with the case without cross-correlations between power spectrum and bispectrum. This however is to be considered just as a preliminary analysis, and a full analysis of the effects causing these differences will be performed in due time.

The comparison between mocks and simulations over the subset of realisations sharing the initial seeds performed in terms of the bispectrum variance, figure 3.7, as well as in terms of cross-correlation coefficients r_{ij}, constitutes our main justification for using the covariance from the larger set of 10,000 PINOCCHIO runs in the analysis of the simulation measurements. In what follows, we will assume that any systematic error in the determination of the power spectrum, bispectrum, and joint covariance matrices based on the approximate halo mocks is negligible, but we will comment on those few instances where a small residual systematic error could affect our results.
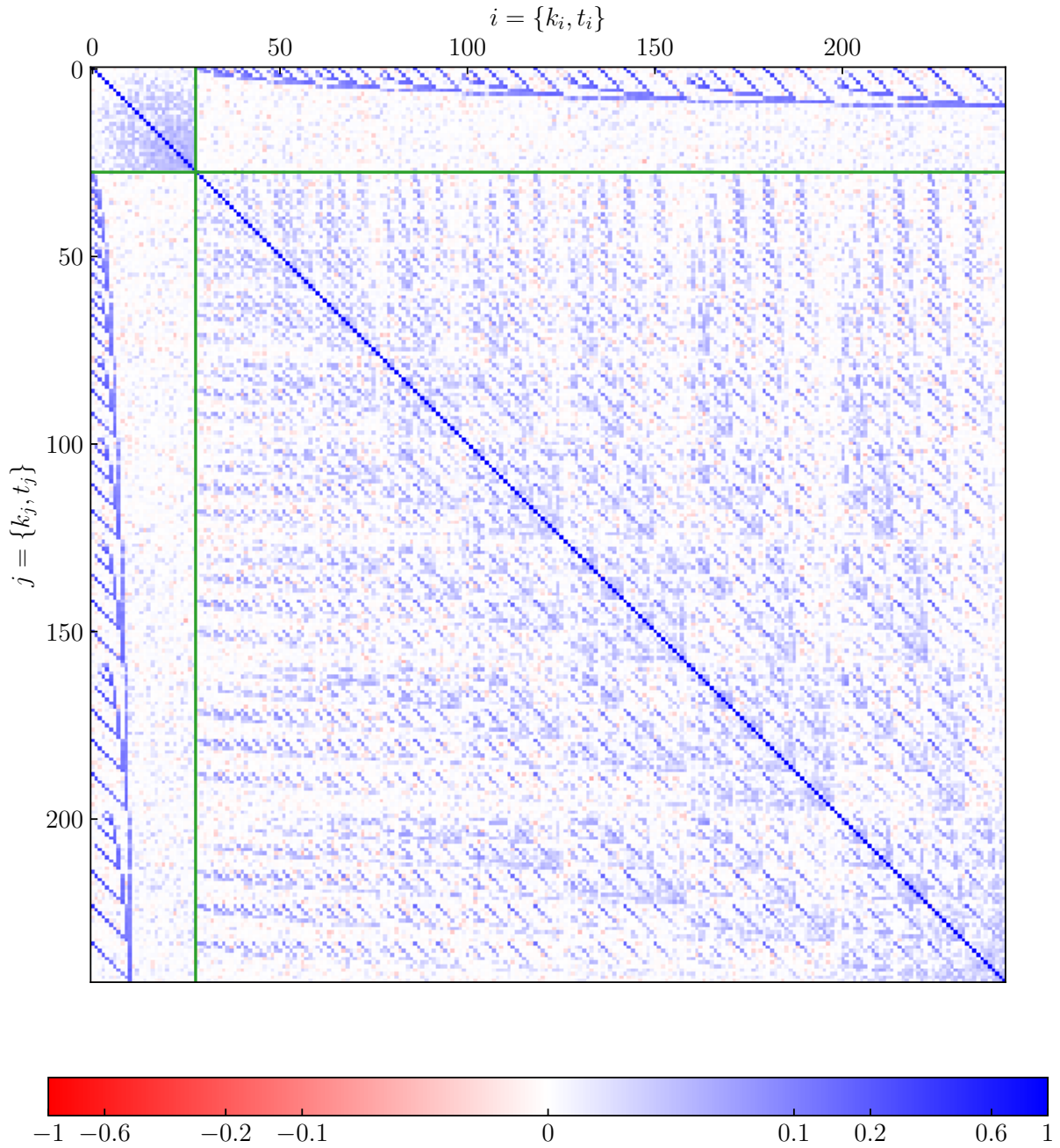
Figure 3.12: Correlation coefficients r_{ij} for the joint measurements of power spectrum and bispectrum of the mock halo catalogs. The indices i and j run over the power spectrum bins first, and then over the bispectrum triangle bins. Therefore, the upper left block is the power spectrum part of the correlation matrix, and the lower right block is the bispectrum part; the off-diagonal parts show the cross-correlations between power spectrum bins and bispectrum triangles. Bins have a width $\Delta k = 2k_f$, the power spectrum part is limited to $k_{\mathrm{max,P}} = 0.24\,h\,\mathrm{Mpc}^{-1}$, while the bispectrum measurements reach $k_{\mathrm{max,P}} = 0.10\,h\,\mathrm{Mpc}^{-1}$.
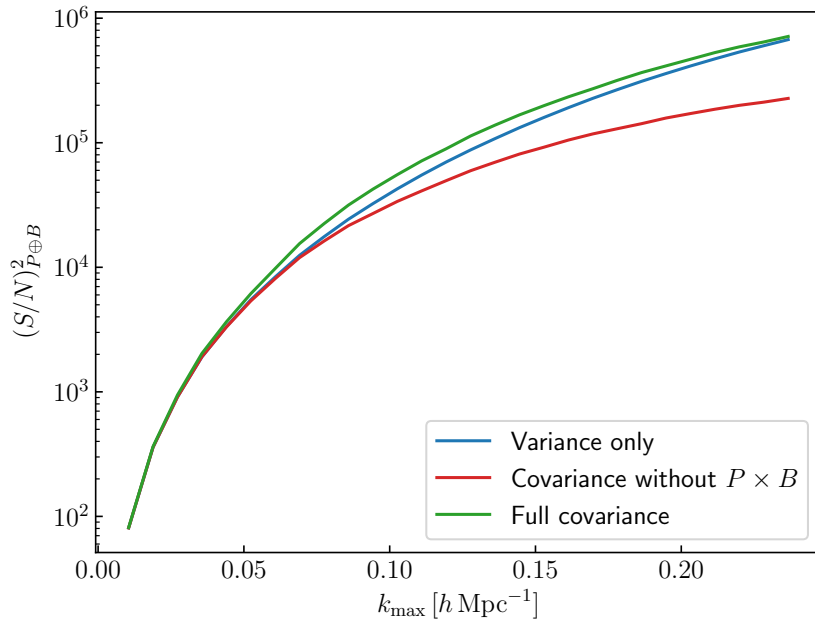
Figure 3.13: Cumulative signal-to-noise ratio for the joint measurements of power spectrum and bispectrum with $\Delta k = 2k_f$. This assumes a signal given by the average measurements of the full set of N-body simulations and a noise given by three different choices for the covariance: diagonal covariance (solid blue line), block-diagonal covariance, where cross-correlations between power spectrum an bispectrum are set to zero (solid red line), and full covariance (solid green line). In all three cases, the covariance has been estimated from the 10,000 PINOCCHIO mocks.

Chapter 4

Toward a robust inference method for the galaxy bispectrum in real space

The theoretical framework described in chapter 2 has several applications in the analysis of the Large Scale Structure. In the present chapter, I will focus on the analysis of the galaxy bispectrum in real space measured from the simulations described in chapter 3.1 using a tree-level model based on Standard Perturbation Theory. I report on the methodology employed in the likelihood analysis of the bispectrum and summarize the results of this analysis, already published in [73].

4.1 Model inference

4.1.1 Theoretical model

The theoretical model that will be employed in the following to describe the galaxy bispectrum at tree-level is the same as the one given in equation 2.101. However, here we use the slightly different bias expansion from [70, 59]. In terms of the tree-level bispectrum model, the only difference is notational, with the tidal bias γ_2 replacing $b_{\mathcal{G}_2}$, and where a comparison of the two bias basis implies $\gamma_2 = b_{\mathcal{G}_2}$. It is worth noticing that, in alternative perturbative approaches [44], additional terms appear in the tree-level bispectrum of biased tracers [120, 121, 122]. However, as they present a k^2 scaling similar to 1-loop corrections in SPT we neglect them in this work.

We fix the cosmological parameters to the actual values of the N-body simulations and only fit the five bias plus shot-noise parameters. This allows us to pre-compute the functions

$$B_{\mathrm{TL}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 2F_2(\mathbf{k}_1, \mathbf{k}_2)P_{\mathrm{L}}(k_1)P_{\mathrm{L}}(k_2) + \text{ cyc.} \tag{4.1}$$

$$\Sigma(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = P_{\mathrm{L}}(k_1)P_{\mathrm{L}}(k_2) + \text{ cyc.} \tag{4.2}$$

$$K(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = S(\mathbf{k}_1, \mathbf{k}_2)P_{\mathrm{L}}(k_1)P_{\mathrm{L}}(k_2) + \text{ cyc.} \tag{4.3}$$

as well as the fully averaged sum of the three linear power spectra in eq. (2.101), and thus minimize the time required to evaluate the model and the likelihood function. Note that, since we fit the model to measurements of the bispectrum from individual simulations, we use the halo number density $\bar{n}$ measured in each box to evaluate the Poissonian shot-noise contribution.

Although eq. (2.101) contains five tunable parameters, it is important to assess how many of them

59

can be justified by the data as a function of k_{max}. For this reason, we consider different models obtained by reducing the parameter space in eq. (2.101) as follows (see also table 4.1 for a compact summary):

- $\mathcal{M}_1$: The most basic model assumes a linear-bias relation ($b_2 = \gamma_2 = 0$) and Poissonian shot noise ($\alpha_1 = \alpha_2 = 0$);

- $\mathcal{M}_{1f}$: This is still a 1-parameter model where both quadratic parameters b_2 and γ_2,[1] are expressed as functions of the linear bias parameter, respectively through the fitting function $\tilde{b}_2(b_1) = 0.412 - 2.143\,b_1 + 0.929\,b_1^2 + 0.008\,b_1^3$ provided by [71] (see also [123, 124]) and assuming local Lagrangian biasing[2] [129], *i.e.* $\gamma_2 = -\frac{2}{7}(b_1 - 1)$ [70, 92]. For the shot-noise corrections we assume the Poisson prediction $\alpha_1 = \alpha_2 = 0$;

- $\mathcal{M}_{2b_2}$: 2-parameter model with b_1 and b_2 free to vary while $\gamma_2 = -\frac{2}{7}(b_1 - 1)$ and $\alpha_1 = \alpha_2 = 0$;

- $\mathcal{M}_{2\gamma_2}$: 2-parameter model with b_1 and γ_2 free to vary while $\tilde{b}_2(b_1)$ follows the fit in [71] and $\alpha_1 = \alpha_2 = 0$;

- $\mathcal{M}_{2\mathrm{loc}}$: 2-parameter, local model with b_1 and b_2 free to vary while $\gamma_2 = 0$ and $\alpha_1 = \alpha_2 = 0$.

- $\mathcal{M}_3$: 3-parameter model with all bias parameters free to vary while $\alpha_1 = \alpha_2 = 0$; this is used as reference model;

- $\mathcal{M}_4$: 4-parameter model with additional freedom in the description of shot noise; a single shot-noise correction parameter $\alpha_1 = \alpha_2$ is allowed to vary, as assumed for instance in [33, 35], in addition to the three bias parameters;

- $\mathcal{M}_5$: 5-parameter model where all parameters in eq. (2.101) are free to vary.

It is worth stressing that our main goal here is to assess the constraining power of the halo bispectrum as a function of k_{max} without considering other data that set additional constraints and break degeneracies among the model parameters. We will discuss the combination with the halo power spectrum in chapter 5. Note that forthcoming galaxy redshift surveys will span much smaller volumes than our N-body simulations but also deal with a substantially higher number-density of tracers ($\gtrsim 10^{-3}h^3$ Mpc^{-3}). A redshift bin of size $\Delta z = 0.2$, that is large enough to properly measure BAO features along the line of sight, corresponds to a volume at most of about 10-12 h^{-3} Gpc^3 both for DESI [113] or Euclid [84]. Therefore, our bispectrum measurements are subject to a smaller sample variance and larger shot-noise corrections with respect to what will be available from future galaxy samples.

4.1.2 Binning effects

As we have seen, there are different ways to take into account discretization effects of the estimators in the theoretical prediction. In particular, we want to study also the impact that these binning techniques have in the determination of the bias parameters.

---

[1]The definition of the tidal-bias operator adopted in [71] differs from ours. They use the "square" of the traceless tidal field to define it while we use the second-order Galileon operator. As a result, our second-order bias, b_2, relates to theirs, $\tilde{b}_2$, as $b_2 = \tilde{b}_2 + \frac{4}{3}\gamma_2$.

[2]Several recent studies provide evidence of small systematic deviations from this relation, with γ_2 being slightly more negative but still linearly related to b_1 [125, 126, 127]. Larger deviations have been measured by [128]

Model	b_1	b_2	γ_2	α_1	α_2
$\mathcal{M}_1$	✓	0	0	0	0
$\mathcal{M}_{1f}$	✓	$\tilde{b}_2(b_1) + \frac{4}{3}\gamma_2(b_1)$	$-\frac{2}{7}(b_1 - 1)$	0	0
$\mathcal{M}_{2\mathrm{loc}}$	✓	✓	0	0	0
$\mathcal{M}_{2b_2}$	✓	✓	$-\frac{2}{7}(b_1 - 1)$	0	0
$\mathcal{M}_{2\gamma_2}$	✓	$\tilde{b}_2(b_1) + \frac{4}{3}\gamma_2$	✓	0	0
$\mathcal{M}_3$	✓	✓	✓	0	0
$\mathcal{M}_4$	✓	✓	✓	✓	α_1
$\mathcal{M}_5$	✓	✓	✓	✓	✓

Table 4.1: Summary of the models analysed in this work. A checkmark ✓ highlights the parameters that are left free to vary. The remaining ones are set to the value indicated in the table. Here, $\tilde{b}_2(b_1)$ is the fitting formula by [71] for the alternative quadratic bias coefficient introduced in footnote 1.

In figure 4.1, we show how much the bispectrum evaluated at the effective configurations differs from the correct bin average given in eq. (2.153). We plot all the different contributions to the tree-level bispectrum separately. For $s = 1$, discrepancies are generally at the few percent level, with a small subset of configurations exceeding 5% only for the K term. As expected, such differences grow with the bin size but remain smaller than 5%, now with the exception of K and the matter contribution B_{TL}, *i.e.* both shape-dependent terms. For the specific case of squeezed isosceles triangles (*i.e.* those configurations with bin *centers* given by $k_1 = k_2$ and $k_3 = \Delta k$) we notice that using the averages of the sorted wavenumbers as in eq. (2.156) works much better than the other case. For completeness, in figure 4.1, we also consider theoretical predictions for the bispectrum directly evaluated at the bin centers $\{k_1, k_2, k_3\}$, clearly limited to those triangle bins whose triplet centers form a closed triangle ("closed bins"). In general, using the bin center performs much worse than any of the two effective solutions we introduced above, and moreover this treatment is not well defined for open bins. For instance, for collinear triangles with $k_1 = k_2 + k_3$, it always gives $K = 0$ for the $s = 1$ and $s = 3$ cases.

4.1.3 Likelihood function

The likelihood function $\mathcal{L}$ of a hypothesis given some data is proportional to the probability of obtaining the data under the assumption that the hypothesis is true. The simplest and most-commonly made assumption is to treat the data as generated by an unbiased estimator that produces Gaussian measurement errors. For a set of N_{t} bispectrum configurations, this gives

$$\ln \mathcal{L} = -\frac{1}{2} \sum_{i=1}^{N_{\mathrm{t}}} \sum_{j=1}^{N_{\mathrm{t}}} \delta B_i \, C_{ij}^{-1} \, \delta B_j \equiv -\frac{1}{2} \delta \mathbf{B}^T \cdot \mathbb{C}^{-1} \cdot \delta \mathbf{B} = -\frac{\chi^2}{2} \, , \tag{4.4}$$

with $\delta B_i = B_i^{\mathrm{obs}} - B_i^{\mathrm{th}}$ where B_i^{obs} denotes the observed (binned) bispectrum for the triangle configuration t_i and B_i^{th} is the corresponding theoretical prediction. The second expression on the r.h.s. adopts a compact notation expressed in terms of the bispectrum data vector $\mathbf{B}$ and the corresponding covariance matrix $\mathbb{C}$. In our analysis, however, the covariance matrix is numerically estimated from the 10,000 mocks described in section 3.2 using their sample covariance. Despite the large number of realisations, the resulting estimate, $\tilde{\mathbb{C}}$ is still plagued by statistical errors that generate two effects. First, on average, they produce a bias in the precision matrix [130]. Secondly, they lead to a loss of information in the parameter-inference process (*i.e.* the resulting posterior probability distributions of the model parameters are, on average, broader than in the absence of noise) [131, 132, 133, 134, 135].
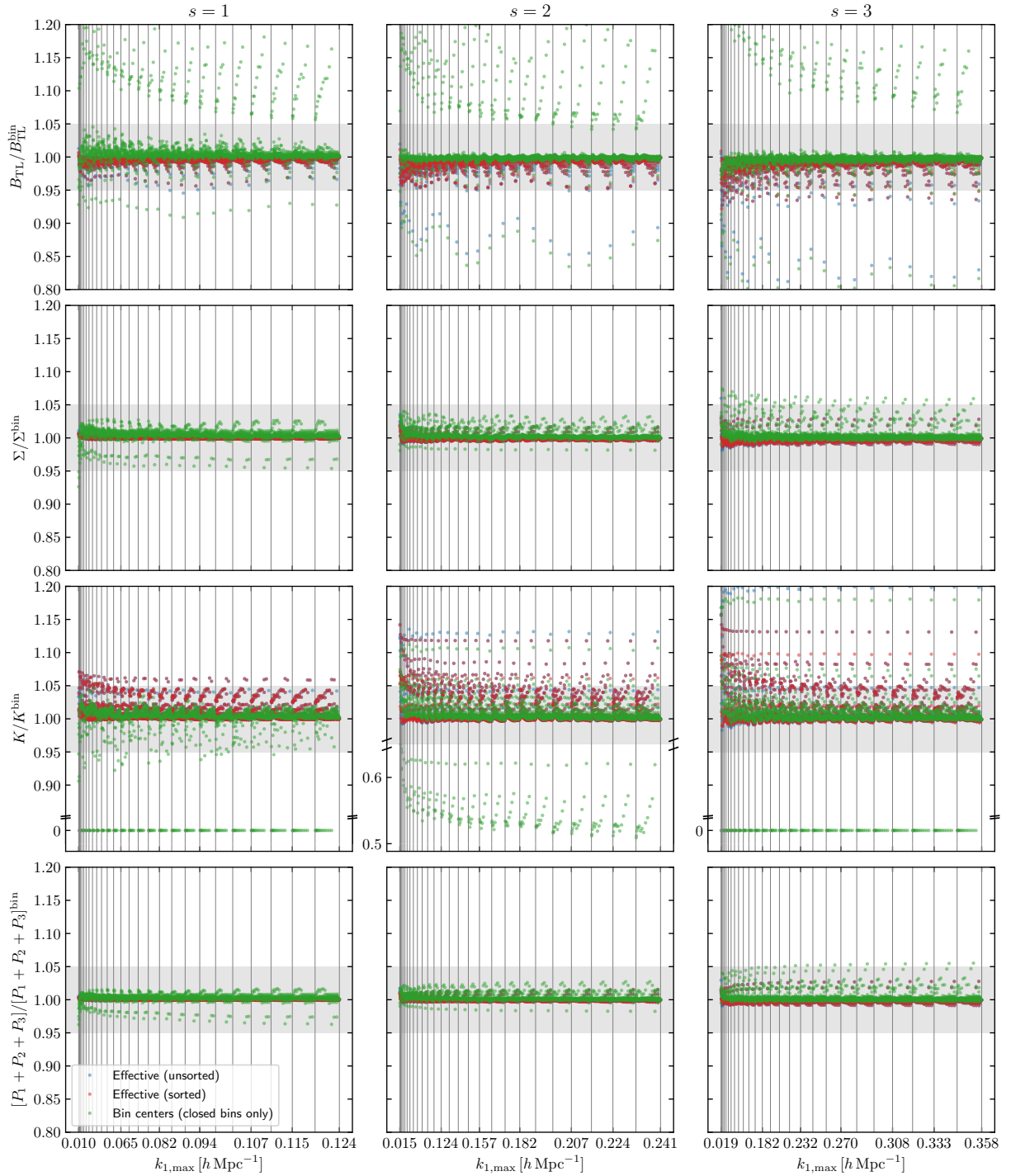
Figure 4.1: Comparison of the bispectrum model evaluated at particular triangle configurations of wavenumbers and its full average over the triangle bins, equation (2.153). We consider two different definitions of effective triangles obtained by sorting or not the sides of a fundamental triangle before averaging them over the bin, see equations (2.155) and (2.156). A third option is obtained by taking the bin center for each leg of the triangular bin, but this procedure is applied only for bins where the bin centers form a closed triangle ("closed bins"). Each column corresponds to a different binning scheme, $s = 1, 2, 3$ from left to right. Each row refers to a different contribution to the tree-level model for the bispectrum given in eq. (2.101). Notice the broken y-axes in the third row of panels, where the large deviations are due to the fact that the terms $[(\hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2)^2 - 1]$ become exactly zero for collinear triangle bins in binning schemes with $s = 1, 3$; this does not happen for the scheme with $s = 2$ since the bin centers defined there never form a collinear triangle with $k_1 = k_2 + k_3$. The effective-sorted method performs generally better, while evaluating the model at the bin centers performs the worst.

It has been shown that considering Gaussian errors and marginalizing over the (unknown) population covariance matrix (by assuming an independence Jeffreys prior) given the noisy estimate leads to the non-Gaussian likelihood [69]:

$$\ln \mathcal{L} = -\frac{N_M}{2} \ln \left[1 + \frac{\delta \mathbf{B} \cdot \widetilde{\mathbb{C}}^{-1} \cdot \delta \mathbf{B}}{N_M - 1} \right] + \ln \left(\frac{\bar{c}_p}{\sqrt{\det \widetilde{\mathbb{C}}}} \right), \tag{4.5}$$

where N_M denotes the number of simulations used to estimate the covariance matrix and $\bar{c}_p$ is a normalization constant that does not depend on the model parameters and we can neglect. This can be interpreted as a generalisation of the multivariate t-distribution. Note that eq. (4.5) has the same numerical complexity as the multivariate Gaussian in eq. (4.4) and can thus be easily used to run MCMC simulations. This is our reference method. Another approach frequently used in the literature is to adopt a Gaussian likelihood with an unbiased estimator of the precision matrix obtained by rescaling the inverse of the sample covariance matrix [130, 136]:

$$\hat{\mathbb{C}}^{-1} = \frac{N_M - N_{\rm t} - 2}{N_M - 1} \widetilde{\mathbb{C}}^{-1} \ . \tag{4.6}$$

We also consider this possibility. In our case, $N_M = 10,000$ and $N_{\rm t}$ varies with the binning strategy. For $s = 1$ and $k_{\max} \simeq 0.12 \, h \, {\rm Mpc}^{-1}$, we have $N_{\rm t} \sim 2,700$ and the correction is of the order of 27% and thus not negligible. For $s = 2$ and 3 and the same $k_{\max}$, instead, it amounts to a few percent or less. Note that the credibility intervals of the model parameters obtained using either eq. (4.5) or a Gaussian likelihood with the rescaled precision matrix given in eq. (4.6) need to be corrected in order to account for the loss of information due to the uncertainty in the covariance matrix, see *e.g.* [134, 135]. We will revisit this procedure in Section 4.3.

A related issue is assessing to what degree eq. (4.4) provides a good approximation to the actual likelihood function. Departures from the ideal case could originate for different reasons. In the first place, we can expect that bispectrum estimates at the largest scales have a non-Gaussian probability distribution simply because they behave like the third power of the Gaussian field $\delta_h(\mathbf{q})$ [38, 137]. In fact, while at small scales these measurements get contributions from a large number of Fourier modes and Gaussianity is recovered due to the central limit theorem, for k-bins close to the fundamental wavenumber k_f, only a few modes contribute, and a certain degree of non-Gaussianity could be present. In addition, one can expect a non-Gaussian distribution to arise from the small-scale non-linear evolution and non-linear bias. However, after inspecting the noisy probability density function of our 298 bispectrum measurements, as well as the values of skewness and kurtosis of the distributions of each individual triangle configuration normalised by their standard errors, we could not detect any significant departure from a Gaussian distribution at the scales that are relevant for this work.

Our ultimate goal is to assess the suitability of different theoretical models for the halo bispectrum on large scales. We thus fit the models to the measurements extracted from our $N_R = 298$ N-body simulations. Since the different realisations are statistically independent, we write the total likelihood function as the product of the likelihoods of the individual measurements or, equivalently,

$$\ln \mathcal{L}_{\rm tot} = \sum_{\alpha=1}^{N_R} \ln \mathcal{L}_\alpha \ . \tag{4.7}$$

For the partial likelihood functions, we use eq. (4.5), neglecting the normalisation constant, as our

Parameter	Broad prior (uniform)	Narrow prior (uniform)
b_1	$[0.5, 5]$	$[0.5, 5]$
b_2	$[-5, 5]$	$[-5, 5]$
γ_2	$[-5, 5]$	$[-5, 5]$
α_1	$[-10, 10]$	$[-1, 1]$
α_2	$[-100, 100]$	$[-1, 1]$

Table 4.2: Uniform prior intervals of the model parameters. For the shot-noise parameters, two different priors are used: the broader one does not influence the inference process while the narrower one makes sure that shot-noise corrections are positive and deviations from Poisson noise are small.

benchmark method. For comparison, we also use eq. (4.4) combined with eq. (4.6). Note that, in the Gaussian case, this procedure gives the same parameter constraints as fitting the mean of our measurements with a suitably rescaled covariance matrix (*e.g.* [138]). This does not hold true in general, though.

4.1.4 Prior probabilities and posterior distributions

We assume uniform priors for all the model parameters (see Table 4.2). More specifically, for the shot-noise parameters, we consider two possibilities. We first consider non-informative, very broad priors that do not influence the inference process (so that it is fully determined by the bispectrum data). In the second set of priors for the shot noise parameters, we force the shot noise to be positive (*i.e.* $\alpha_{1,2} > -1$) and make sure that corrections are not larger than the Poisson terms.

We evaluate the posterior distribution by means of Monte Carlo Markov Chain simulations using the Python affine-invariant sampler `emcee` [139]. For each run, we use 100 walkers to sample the parameter space. To make sure that they are sufficiently converged, we stop the simulations after 50 integrated autocorrelation times [140].

4.2 Statistical methods: goodness of fit and model comparison

In this section, we provide a concise introduction to the statistical tools we use to determine the goodness of fit between observed and predicted results as well as to compare the performance of models with a different number of parameters.

4.2.1 Model selection

We consider a set of candidate models $\mathcal{M}_k \in \mathcal{M}$ for the halo bispectrum (each characterized by the parameters set $\boldsymbol{\theta}_k$) and use different criteria to select which of them "better" agrees with the measurements from our simulations. Two major threads run through the statistical literature on Bayesian model selection that go under the names of "explanatory" and "predictive" modelling. Let us consider a particular phenomenon or mechanism that gives rise to noisy data (the Data Generating Process or DGP). A model is an often imperfect collection of mathematical and statistical rules giving rise to an output that resembles the actual data in important ways. Explanatory modelling tests theoretically motivated hypotheses and looks for the "true model" (*i.e.* the DGP or approximations thereof) by examining posterior probabilities given the observed data [141]. Predictive modelling, on the other hand, searches for the model that makes the best prediction of future observations

generated by the same DGP as the observed data. This is based on the posterior predictive probability distribution function which gives the probability of future measurements conditional on the observed data:

$$p(\text{new data}|\text{data}, \mathcal{M}_k) = \int p(\text{new data}|\boldsymbol{\theta}_k, \mathcal{M}_k) \, P(\boldsymbol{\theta}_k|\text{data}) \, \mathrm{d}\boldsymbol{\theta}_k \;, \qquad (4.8)$$

where P denotes the posterior distribution of $\boldsymbol{\theta}_k$ given the data. For finite and noisy measurements, the two concepts possibly lead to different conclusions. In this work, we make use of one model selection criterion from each class.

Bayes factors from the Savage-Dickey density ratio In order to evaluate the evidence that the data provide in favour of $\mathcal{M}_j$ with respect to $\mathcal{M}_k$, explanatory modelling generally relies on calculating the Bayes Factor (BF), *i.e.* the ratio between the probabilities of observing the data under the models,

$$\mathrm{BF}_{jk} = \frac{p(\text{data}|\mathcal{M}_j)}{p(\text{data}|\mathcal{M}_k)} = \frac{\int \mathcal{L}(\text{data}|\boldsymbol{\theta}_j) \, \pi(\boldsymbol{\theta}_j|\mathcal{M}_j) \, \mathrm{d}\boldsymbol{\theta}_j}{\int \mathcal{L}(\text{data}|\boldsymbol{\theta}_k) \, \pi(\boldsymbol{\theta}_k|\mathcal{M}_k) \, \mathrm{d}\boldsymbol{\theta}_k} \;, \qquad (4.9)$$

where $\pi(\boldsymbol{\theta}_j|\mathcal{M}_j)$ denotes the prior for the model parameters. This quantity coincides with the ratio between the marginal likelihoods, *i.e.* between the a priori predictions that the models make about the probability of observing the data in the experiment. The posterior odds of model $\mathcal{M}_j$ relative to model $\mathcal{M}_k$ are obtained by multiplying the prior odds times the Bayes factor BF_{jk}. Ref. [142] recommends that Bayes factors larger than 3, 10, and 100 should be used to speak of substantial, strong and decisive evidence in favour of model $\mathcal{M}_j$ against model $\mathcal{M}_k$. Note that Bayes factors account for model complexity and automatically penalise more complicated models with respect to simpler ones.

The main drawback of applying this machinery to our bispectrum study is that Bayes factors are notoriously challenging to compute from MCMC simulations. This is because the MCMC method sparsely samples regions of parameter space where the likelihood function is relatively low that might give important (if not dominant) contributions to the integrals in equation (4.9) [141]. However, the situation improves dramatically if we limit our analysis to properly nested models. Model $\mathcal{M}_j$ is said to be properly nested under $\mathcal{M}_k$ if: (i) the parameter set $\boldsymbol{\theta}_j \subset \boldsymbol{\theta}_k$ (i.e. $\mathcal{M}_j$ is obtained by fixing each additional parameter in $\mathcal{M}_k$ to a constant value, i.e. $\boldsymbol{\psi} = \mathbf{c}$ where $\boldsymbol{\psi}$ denotes the set of parameters that are free to vary in $\mathcal{M}_k$ and are fixed in $\mathcal{M}_j$); (ii) the prior distributions in the models satisfy $\lim_{\boldsymbol{\psi} \to \mathbf{c}} \pi(\boldsymbol{\theta}_k|\mathcal{M}_k) = \pi(\boldsymbol{\theta}_j|\mathcal{M}_j)$; and (iii) the likelihood functions in the models satisfy the relation $\mathcal{L}(\text{data}|\boldsymbol{\theta}_j, \mathcal{M}_j) = \mathcal{L}(\text{data}|\boldsymbol{\theta}_j, \mathbf{c}, \mathcal{M}_k)$. In this case, the Bayes factor is given by [143]

$$\mathrm{BF}_{jk} = \frac{\int P(\boldsymbol{\theta}_j, \boldsymbol{\psi} = \mathbf{c}|\mathcal{M}_k) \, \mathrm{d}\boldsymbol{\theta}_j}{\int \pi(\boldsymbol{\theta}_j, \boldsymbol{\psi} = \mathbf{c}, |\mathcal{M}_k) \, \mathrm{d}\boldsymbol{\theta}_j} \;, \qquad (4.10)$$

which is known as the Savage-Dickey density ratio and is relatively easy to compute from a MCMC simulation. See *e.g.* [144] for an application to cosmology.

Deviance Information Criterion Given a dataset, a model $\mathcal{M}_k$, and a particular set of values for the model parameters $\hat{\boldsymbol{\theta}}_k$, we define the deviance statistic as

$$D = 2 \log C(\text{data}) - 2 \log \mathcal{L}(\text{data}|\hat{\boldsymbol{\theta}}_k, \mathcal{M}_k) \;, \qquad (4.11)$$

where the fully specified function $C(\mathrm{data})$ does not depend on the candidate model. Frequentist model assessment is based on the difference of the log-likelihoods between a model and the saturated model (that perfectly fits all data) and can thus be also formulated in terms of the difference of the deviances (note that, apart from the constant $C(\mathrm{data})$, D coincides with the χ^2 statistic for measurements with Gaussian noise). Ref. [145] suggested to use the posterior distribution of D as a measure of goodness of fit in the Bayesian framework. [146] formalised this concept by introducing the Deviance Information Criterion (DIC) as a method for model selection. This is based on the following statistic:

$$\mathrm{DIC} = \langle D \rangle_{\mathrm{post}} + p_D \,, \tag{4.12}$$

where $\langle D \rangle_{\mathrm{post}}$ denotes the posterior expectation of the deviance and $p_D = \langle D \rangle_{\mathrm{post}} - D(\langle \boldsymbol{\theta} \rangle_{\mathrm{post}})$ is a Bayesian measure of model complexity that gives an estimate of the effective number of model parameters. Models associated with a lower DIC are better supported by the data. Starting from the definitions given above, we can write the DIC metric as $\mathrm{DIC} = D(\langle \boldsymbol{\theta} \rangle_{\mathrm{post}}) + 2p_D$. This notation makes it clear that the DIC is based on a trade-off between model accuracy and complexity. Fit quality is measured by plugging the posterior mean of the parameters in the deviance: the better the model fits the data, the larger are the values assumed by the likelihood function, and thus the smaller is $D(\langle \boldsymbol{\theta} \rangle_{\mathrm{post}})$. On the other hand, by adding $2p_D$ we penalise increasing model complexity in order to avoid overfitting. Note that $\langle D \rangle_{\mathrm{post}} = D(\langle \boldsymbol{\theta} \rangle_{\mathrm{post}}) + p_D$ already incorporates a penalty for complexity and should then be considered a measure of model adequacy rather than a pure measure of fit. The DIC can be interpreted as the Bayesian generalisation of the Akaike Information Criterion (AIC) used in the maximum-likelihood framework [147]. For non-hierarchical models and large data samples, the DIC asymptotically reduces to the AIC. For a number of reasons[3], the definition of p_D given above has been subject to much criticism. An alternative estimator for the effective number of parameters in non-hierarchical models (which is invariant under reparametrization and never negative but only gives meaningful estimates when priors are non-informative) is $p_V = \mathrm{Var}(D)/2$ where $\mathrm{Var}(D)$ denotes the posterior variance of the deviance [148]. In this work, we use this second estimator to build the DIC metric from the MCMC runs.

The DIC statistic for a model is very easy to calculate when the likelihood is available in closed form and the posterior distribution is obtained by MCMC simulation. The actual value of the DIC for a model has no particular meaning, only differences $\Delta\mathrm{DIC}$ between models matter. According to a commonly used rule of thumb, values of $\Delta\mathrm{DIC} < 2$ are insignificant while differences of 5 and 10 provide substantial and decisive evidence against the less supported model. However, the reliability of this scale depends on the application. Tests have shown that the DIC tends to select overfitting models if p_D is not small with respect to the sample size (*e.g.* [149]). Consistently, a number of

---

[3]The concept of effective number of degrees of freedom was originally introduced to deal with hierarchical Bayesian models. In this case, the parameters that regulate observations at the individual level depend on hyperparameters (which are assigned hyperpriors) that describe the group level. In complex, multi-level hierarchies, parameters are not independent and it is not obvious how to calculate their total number. The advantage of introducing the p_D estimator is that it uses MCMC results directly and straightforwardly. In general, the p_D statistic measures the constraining power of the data compared to the prior. Whenever the deviance is well approximated by a quadratic function around $\langle \boldsymbol{\theta} \rangle_{\mathrm{post}}$ (*i.e.* the likelihood of the model parameters is approximately Gaussian as expected for large datasets from the Bayesian central limit theorem), each model parameter contributes one to p_D if the posterior information about the parameter is dominated by the likelihood, it contributes zero if the information is dominated by the prior, and it contributes a number in between zero and one if both the prior and the likelihood are important to constrain its value [148]. However, in peculiar cases in which the Gaussian approximation for the likelihood fails, p_D could even assume negative values. Moreover, p_D (and thus the DIC metric) is not invariant under reparametrization of the model (while $\langle D \rangle_{\mathrm{post}}$ is).

theoretical studies suggest to increase the penalty for model complexity and use $\text{DIC}^* = \langle D \rangle_{\text{post}} + 2p_D$ [150, 149, 151, 152]. In fact, under mild regularity assumptions, the original DIC selects the model that asymptotically (for large samples) gives the smallest expected Kullback-Leibler divergence between the DGP and the plug-in predictive distribution evaluated at $\langle \boldsymbol{\theta} \rangle_{\text{post}}$. Instead, one would like to minimise the divergence with respect to the proper predictive distribution given in equation (4.8). We stress once again that the goal here is not to select the true model but rather to make a pragmatic choice that agrees with observations and provides good predictions for future datasets.

4.2.2 Goodness of fit and posterior predictive p-values

We would like to assess which models provide a good fit to the data and which do not. In classical statistics, the maximum-likelihood method is often used to determine the best-fit model parameters $\hat{\boldsymbol{\theta}}$ for a given dataset. A test statistic (*e.g.* the χ^2) is then selected to determine the "significance" of the fit [153]. Under the null hypothesis that the data are actually sampled from the model with $\hat{\boldsymbol{\theta}}$, one computes the conditional frequentist probability of obtaining as many or more extreme data of the test statistic. This p-value corresponds to the long-run frequency taken over the sampling distribution of the data under the null hypothesis (*i.e.* it is the probability that other unobserved data-sets would be more extreme than the one that was observed in terms of the test statistic). When the p-value for an experiment is small, then one has to assume that either an unusual event has occurred or that the null hypothesis is not true. Thus, the smaller the p value, the less it is plausible that the null hypothesis is true.

Concern about the interpretation of p-values is widespread in statistics. Well-known problems of this approach are (i) that it is not possible to consider nuisance parameters and (ii) that it depends on how the data acquisition process is terminated and thus violates the likelihood principle.

Several authors have developed a Bayesian motivated adaptation of the classical goodness-of-fit test based on the p-value [154, 155, 156, 148]. The method relies on the posterior predictive probability distribution function given in equation (4.8). This function gives the probability (conditional on the observed data) of replicated data that could have been observed or, to think in predictive terms, that would be observed in the future if the experiment is repeated. Note that what is kept fixed here is the observed data while the classical method relies on probabilities that are conditional on the parameters of the best-fit model $\hat{\boldsymbol{\theta}}$ (*i.e.* the set $\hat{\boldsymbol{\theta}}$ is kept fixed). The argument then proceeds as follows. A discrepancy variable $\Delta(\text{data}, \boldsymbol{\theta})$ is introduced to quantify the deviation of the model (with parameters $\boldsymbol{\theta}$) from the data. The posterior predictive p-value (ppp) of Δ is defined as

$$\text{ppp} = \text{Prob}\left[\Delta(\text{replicated data}, \boldsymbol{\theta}) \geq \Delta(\text{data}, \boldsymbol{\theta})|\text{data}\right] , \tag{4.13}$$

where the probability is taken over the joint distribution

$$P(\boldsymbol{\theta}, \text{replicated data}|\text{data}) = p(\text{replicated data}|\boldsymbol{\theta})\, P(\boldsymbol{\theta}|\text{data}) . \tag{4.14}$$

In practice, we compute the ppp of Δ using MCMC simulations by drawing one replica from the statistical model for each step of the chain. The estimated ppp corresponds then to the fraction of steps for which the discrepancy variable equals or exceeds its realised value. A ppp which is close to zero or one indicates that the realised data have a low probability of occurring under the postulated

model, *i.e.* that the model does not fit the data well. It is important to stress, however, that ppp's do not have in general a uniform distribution under the true model (meaning that the probability to find ppp > 0.95 is not necessarily 5 percent) as they often tend to have distributions that are more concentrated around 0.5. Therefore, if one wants to associate a precise statistical significance to them, they need to be calibrated.

In this work, we use the ppp as a measure of goodness of fit by adopting the log likelihood as the discrepancy variable. We perform the calibration of the ppp's by generating artificial data based on our reference model with added noise and by fitting them. We find that the distribution of our ppp's is remarkably close to uniform under the true model. Therefore, in the analysis of the N-body data, we interpret extreme values of ppp near zero or one as revealing a systematic misfit between the bispectrum measurements and the model predictions that cannot be ignored. However, in order to facilitate understanding to readers who are more familiar with frequentist goodness-of-fit tests, we also provide the value of the posterior averaged χ^2 statistic and the corresponding upper one-sided 95 per cent confidence limit as a function of $k_{\max}$ (see Section 4.3.1 for further details). Note that the number of degrees of freedom that should be associated to this statistic is the total number of data points (as taking the posterior average gives a larger value than minimizing the χ^2 as in equation (4.12) for the deviance).

Finally, in order to quantify the degree to which a model systematically deviates from the actual measurements at the level of single data points, we make use of a technique known as graphical posterior predictive checking [154, 156]. In line with the ppp, this concept adopts a frequentist-like approach in a fully Bayesian framework. The underlying idea is that, if a model provides a good fit, it could be used to generate replicated data that look like those that have been observed. In practice, we compute the difference between each realised data point (*i.e.* the measurements from the N-body simulations) and the mean of the replicated data sampled from the posterior probability distribution function (*i.e.* by considering one replica for each model sampled by the Markov chain). We then convert this difference into a standardized residual, R, by expressing it in units of the rms uncertainty of the data. Systematic deviations with $|R| \gg 1$ indicate potential shortcomings of the model.

4.3   Results

In this section, we present the results obtained using all the tools described above to fit the halo bispectrum with different versions of the tree-level model.

4.3.1   Benchmark analysis

As a starting point, we fit model $\mathcal{M}_3$ to our set of bispectrum measurements using the likelihood function given in eq. (4.5). Bin averages for the model predictions are evaluated exactly using eq. (2.153). We refer to this combination of choices as our reference study.

Figure 4.2, like many others that follow, displays our results in five complementary panels. In the top-left corner, we show the mean and the rms values of the posterior distribution[4] for each model

---

[4]We do not correct the width of the posteriors for the loss of information due to the uncertainty in the covariance matrix [134]. In fact, this is impossible to do in an exact way [135]. Moreover, for all scales at which we obtain a good fit, the expected size of the correction (a few percent) is several times smaller than the statistical uncertainty with which we measure the rms value of the model parameters in the MCMC simulations.

parameter as a function of $k_{\max}$. The goodness of fit is displayed in the top-right panel. Here, we show the ppp[5] and the posterior-averaged χ^2 statistic divided by the number of data points, both as a function of $k_{\max}$. The central panel presents contour plots for the joint posterior density of all parameter pairs at $k_{\max} = 0.082\,h\,\mathrm{Mpc}^{-1}$. This particular value has been selected for two reasons: i) it is an exact multiple of all the bin sizes we consider, thus ensuring that all measurements contain the same information; ii) it approximately coincides with the largest $k_{\max}$ for which the fits we present are good. Finally, in the bottom panel, we provide a direct comparison of the data and the fit (at $k_{\max} = 0.082\,h\,\mathrm{Mpc}^{-1}$) in terms of the standardized residuals of the posterior predictive checks (PPCs) described at the end of section 4.2.2. In practice, we compute the difference between the measured bispectrum and the posterior mean of the replicas. The result is then averaged over all the realizations and normalized to the standard deviation of the measurements. Equilateral (binned) configurations are highlighted with vertical lines, so that from one vertical line to the other we span all configuration shapes with $k_1 \geq k_2 \geq k_3$ at fixed k_1.

In order to test the reliability of the fits presented above, we also derive the linear-bias parameter by directly comparing the halo-matter cross power spectrum and the matter power spectrum of the N-body simulations at very large scales, as described in section 3.1.2. We find a value of $b_1^\times = 2.7131 \pm 0.0012$, where the superscript here indicates that this measurement comes from cross-correlating halos and matter. In figure 4.2, this result is represented as a narrow horizontal band in the top-left panel and as a vertical band in the center-right panel, both painted gray.

Coming to the specific results of our benchmark study, we note that the range of validity of model $\mathcal{M}_3$ does not extend beyond $k_{\max} = 0.1\,h\,\mathrm{Mpc}^{-1}$ and shows some dependence on the bin size. The ppp approaches one at $k_{\max}$ of, approximately, $0.082, 0.107$ and $0.119\,h\,\mathrm{Mpc}^{-1}$ for $\Delta k = k_f, 2k_f$ and $3k_f$, respectively. Similar conclusions can be drawn based on $\langle\chi^2\rangle_{\mathrm{post}}$. The values of $k_{\max}$ for which a good fit is achieved also correspond to consistent posterior probability distributions for the model parameters. All fits provide compatible results (within the statistical uncertainty) up to $k_{\max} \sim 0.08\,h\,\mathrm{Mpc}^{-1}$. When we include smaller scales in the analysis, instead, the location of the posterior for γ_2 starts running with $k_{\max}$ and we simultaneously obtain unsatisfactory values of the goodness-of-fit statistics. The posterior probabilities for $k_{\max} = 0.082\,h\,\mathrm{Mpc}^{-1}$ show marked degeneracies between the model parameters that cannot be individually constrained apart from b_1. The standardized residuals of the PPCs do not single out any particular triangle configuration for which the model systematically deviates from the data for any binning scheme. Note that the goodness of fit at large $k_{\max}$ could deteriorate either because the tree-level models become inaccurate on small scales but also because our estimate of the covariance matrix becomes more and more imprecise.

The relatively small difference among the results for different bin sizes are to some extent expected. In general, the ability of the bispectrum to break degeneracies between parameters depend on the different dependence on shape of the different contributions. This is clearly affected by the binning. In particular, it is reduced by a larger binning. At the same time, a larger binning also implies smaller error bars. The combination of such different factors is likely at the origin of the observed differences.

---

[5]Replicated data are generated by adding Gaussian noise (with covariance matrix $\widetilde{\mathbb{C}}$) to the theoretical models at each step of the MCMC simulations, meaning that the quantity $\delta\mathbf{B}_{\mathrm{repl}} \cdot \widetilde{\mathbb{C}}^{-1} \cdot \delta\mathbf{B}_{\mathrm{repl}}$ follows a χ^2 distribution with N_t degrees of freedom for each realisation.
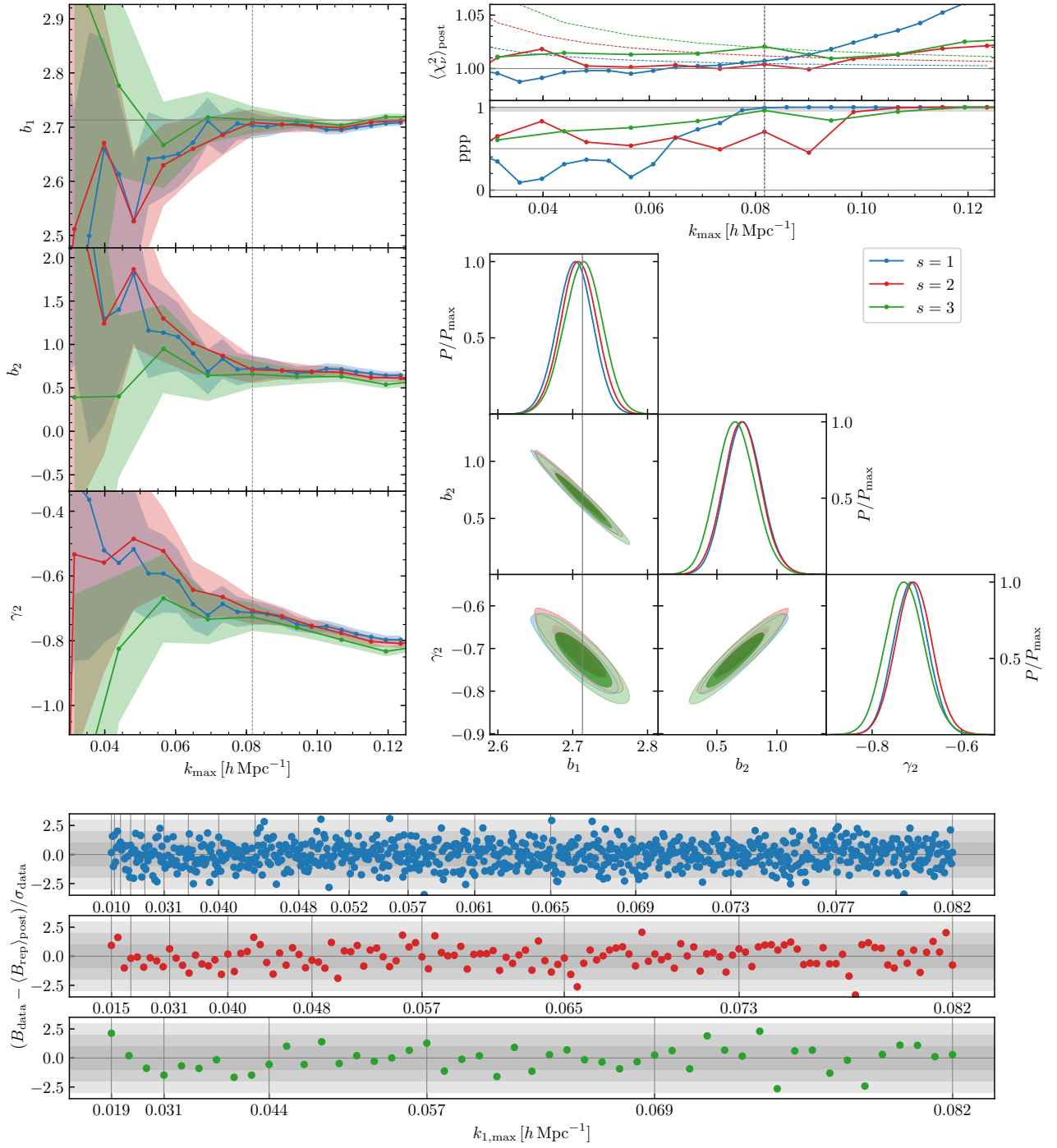
Figure 4.2: Fit of model $\mathcal{M}_3$ to the halo bispectrum measured using three different bin widths $\Delta k = k_f, 2k_f, 3k_f$. The top-left panels show the posterior mean (solid lines) and rms scatter (shaded areas) of the model parameters as a function of $k_{\max}$. The vertical dashed line highlights the reference scale of $k_{\max} = 0.082\,h\,\mathrm{Mpc}^{-1}$ for which we display contour plots for the joint posterior density of parameter pairs in the middle-right panel. Here, darker and lighter shaded areas represent the 68 and 95 percent joint credibility regions, respectively. The narrow gray bands indicate the constraints on the linear-bias parameters derived from the halo-matter cross power spectrum. The standardized residuals of the posterior predictive checks for the same $k_{\max}$ are shown in the large panel on the bottom. Two goodness-of-fit diagnostics are displayed in the top-right panel as function of $k_{\max}$: the reduced χ^2 statistic averaged over the posterior (top inset) and the ppp (bottom inset). As a reference, the dashed curves in the top inset indicate the upper one-sided 95 percent confidence limit in a frequentist χ^2 test (note that the number of datapoints included in the fit varies with $k_{\max}$).

4.3.2 Model selection: shot noise

We now fit models $\mathcal{M}_4$ and $\mathcal{M}_5$ to the bispectrum data. They extend model $\mathcal{M}_3$ to include corrections to Poissonian shot noise that are routinely included in the analysis of survey data, *e.g.* [35, 17]. Figure 4.3 compares the results for the three models for the measurements with $s = 1$ when using broad priors for the shot-noise parameters. Analogous results are obtained adopting one of the other binning schemes. The most important thing to notice is that adding extra shot-noise parameters does not lead to any significant improvement in the goodness of fit and thus does not extend the range of validity of the models in terms of $k_{\max}$. Further insight is obtained by looking at the PPCs for $k_{\max} = 0.082\, h\,\mathrm{Mpc}^{-1}$ (not shown): the residuals are virtually identical for the three models. In general, the fit results for $\mathcal{M}_4$ are very similar to $\mathcal{M}_3$ as, for sufficiently large $k_{\max}$, the additional parameter is always well constrained to be close to zero. On the other hand, large degeneracies between α_1 and α_2 (that, in this case, can assume opposite signs) as well as between the shot-noise and the bias parameters are present for $\mathcal{M}_5$. This leads to substantially larger marginalised posteriors for the bias parameters. It is also worth stressing that the loose constraints set by the data on α_1 and α_2 span a much larger range than expected from theoretical shot-noise models [157, 158]. We thus conclude that the large-scale bispectrum is insufficient to inform these models.

Figure 4.4 shows a comparison similar to the one shown in figure 4.3, but for the narrower priors on the shot-noise parameters. As in the previous case, both for $\mathcal{M}_4$ and $\mathcal{M}_5$, α_1 is well constrained and consistent with 0, while in $\mathcal{M}_5$, α_2 is completely unconstrained. On the other hand, the results for the bias parameters are essentially unaffected by the introduction of shot-noise parameters with such priors.

In figure 4.5, we apply model-selection techniques to the nested models $\mathcal{M}_3$, $\mathcal{M}_4$, and $\mathcal{M}_5$ considering both priors sets. We first look at the constraining power of our bispectrum data with $s = 1$ (similar results are found for the other bin sizes). The top-right and middle-right panels show the effective number of parameters of the fits as a function of $k_{\max}$, respectively for the broad and narrow shot-noise priors. While all parameters of $\mathcal{M}_3$ and $\mathcal{M}_4$ stop being prior dominated for $k_{\max} \simeq 0.04\, h\,\mathrm{Mpc}^{-1}$, the bispectrum data can fully constrain $\mathcal{M}_5$ only from $k_{\max} \simeq 0.06\, h\,\mathrm{Mpc}^{-1}$ for the broad priors, and from $k_{\max} \simeq 0.08\, h\,\mathrm{Mpc}^{-1}$ for the narrow priors. In the top-left and middle-left panels, instead, we show the DIC difference with respect to $\mathcal{M}_3$, again for broad and narrow priors respectively. In both cases, for small values of $k_{\max}$, $\mathcal{M}_3$ is slightly preferred by the DIC, although with low significance. A fair conclusion is that, for $k_{\max} < 0.09\, h\,\mathrm{Mpc}^{-1}$, the three models provide very similar DIC and cannot be ranked. On smaller scales, the DIC strongly prefers $\mathcal{M}_5$ but this is irrelevant as none of the models provides an acceptable goodness of fit. It is, in fact, quite possible that some non-linear effects that are not included in the tree-level model are partially accounted for by the shot-noise corrections. Stronger conclusions can be drawn based on the Bayes factors evaluated using the Savage-Dickey density ratio (bottom panels). In this case, for all binning schemes, $\mathcal{M}_3$ is decisively preferred over $\mathcal{M}_4$ at $k_{\max} \simeq 0.09\, h\,\mathrm{Mpc}^{-1}$ where the ratio exceeds 100 with broad shot-noise priors, and 10 for narrow shot-noise priors (although it suddenly drops for larger values of $k_{\max}$ where both models fail to properly fit the data). Similarly, $\mathcal{M}_4$ is preferred over $\mathcal{M}_5$ although to a lesser degree as the BF (not shown) only reaches a maximum value of about 20 at $k_{\max} \simeq 0.08\, h\,\mathrm{Mpc}^{-1}$. The bottom line of this section is that our data provide no evidence for non-Poissonian shot-noise corrections on the scales in which the tree-level model fits well, *i.e.* $k_{\max} < 0.08\, h\,\mathrm{Mpc}^{-1}$.
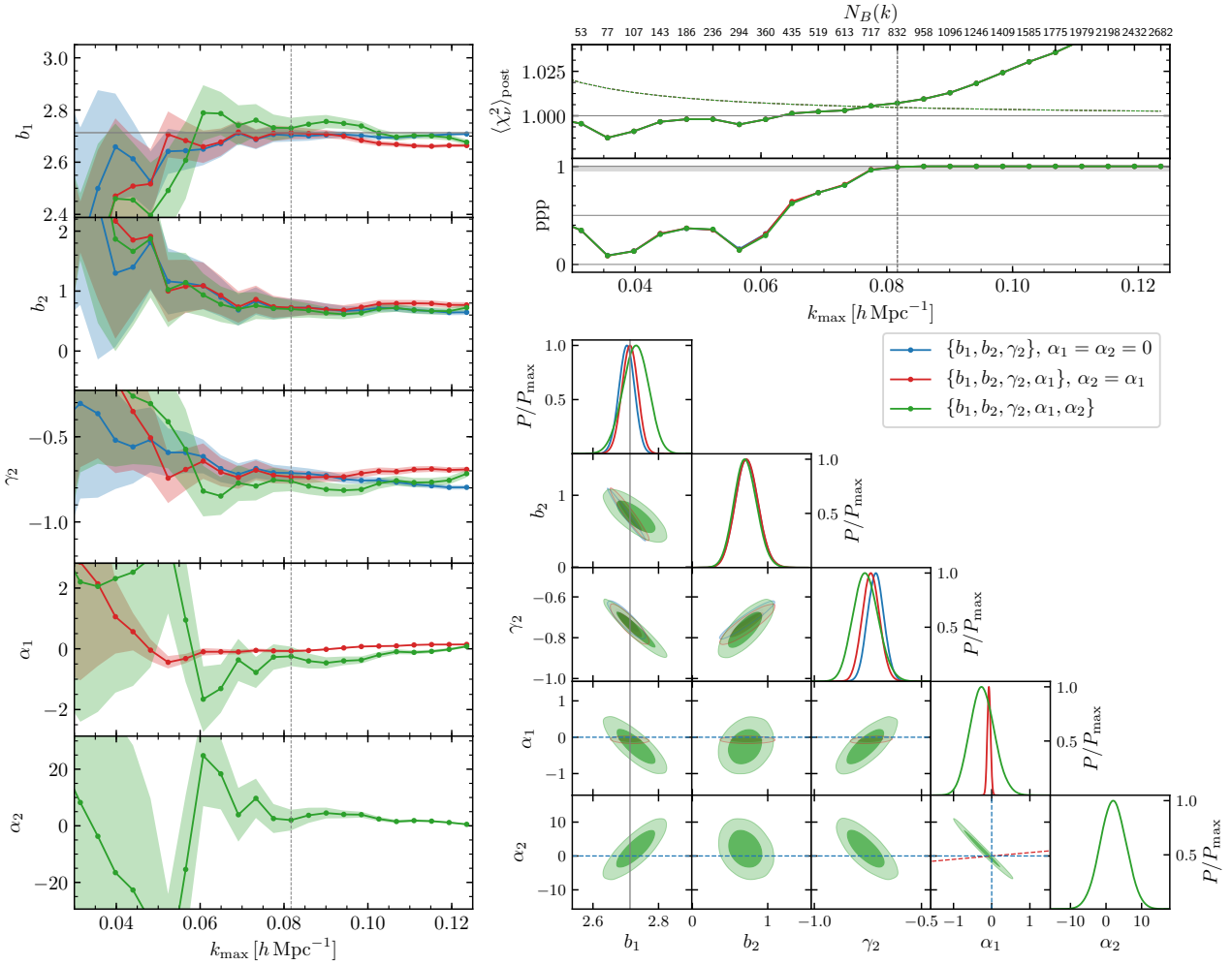
Figure 4.3: As in figure 4.2 but comparing models $\mathcal{M}_3$, $\mathcal{M}_4$ and $\mathcal{M}_5$ for the binned data with $s = 1$ and the broad priors for α_1 and α_2. The labels in the inset summarize the difference between the models: the parameters between curly brackets are let free to vary within the prior range while the others are kept fixed. The blue dashed lines show the relations $\alpha_1 = \alpha_2 = 0$ in $\mathcal{M}_3$, while the red dashed line shows the relation $\alpha_2 = \alpha_1$ in $\mathcal{M}_4$.

4.3.3 Model selection: reducing the number of bias parameters

In section 4.1.1, we have described several possibilities for reducing the freedom of our reference theoretical model $\mathcal{M}_3$. We now investigate whether these restricted models provide an accurate description of the bispectrum measurements extracted from our simulations and contrast them with $\mathcal{M}_3$.

As a first test, in figure 4.6, we compare models $\mathcal{M}_1$ and $\mathcal{M}_{2\mathrm{loc}}$ (which correspond to truncating a local Eulerian bias expansion at first and second order, respectively) with $\mathcal{M}_3$ (which also includes the tidal-bias term). In terms of goodness of fit, $\mathcal{M}_1$ fails around $k_{\max} \simeq 0.05\,h\,\mathrm{Mpc}^{-1}$, *i.e.* at significantly larger scales than the other models that, on the other hand, provide almost identical values for the ppp and $\langle \chi_\nu^2 \rangle$. The local models retrieve different values for the bias parameters with respect to $\mathcal{M}_3$ (see also [70, 92]). Without combining our results with other clustering statistics, it is impossible to say whether $\mathcal{M}_{2\mathrm{loc}}$ and $\mathcal{M}_3$ provide a realistic description of the data since, in terms of goodness of fit, they are practically equivalent. However, it is interesting to notice that the DIC shown in the bottom-left panel of figure 4.6 not only strongly disfavours $\mathcal{M}_1$ already at $k_{\max} \simeq 0.03\,h\,\mathrm{Mpc}^{-1}$, but also clearly indicates a preference for $\mathcal{M}_3$ at $k_{\max} \simeq 0.06\,h\,\mathrm{Mpc}^{-1}$. Equivalent conclusions can be drawn
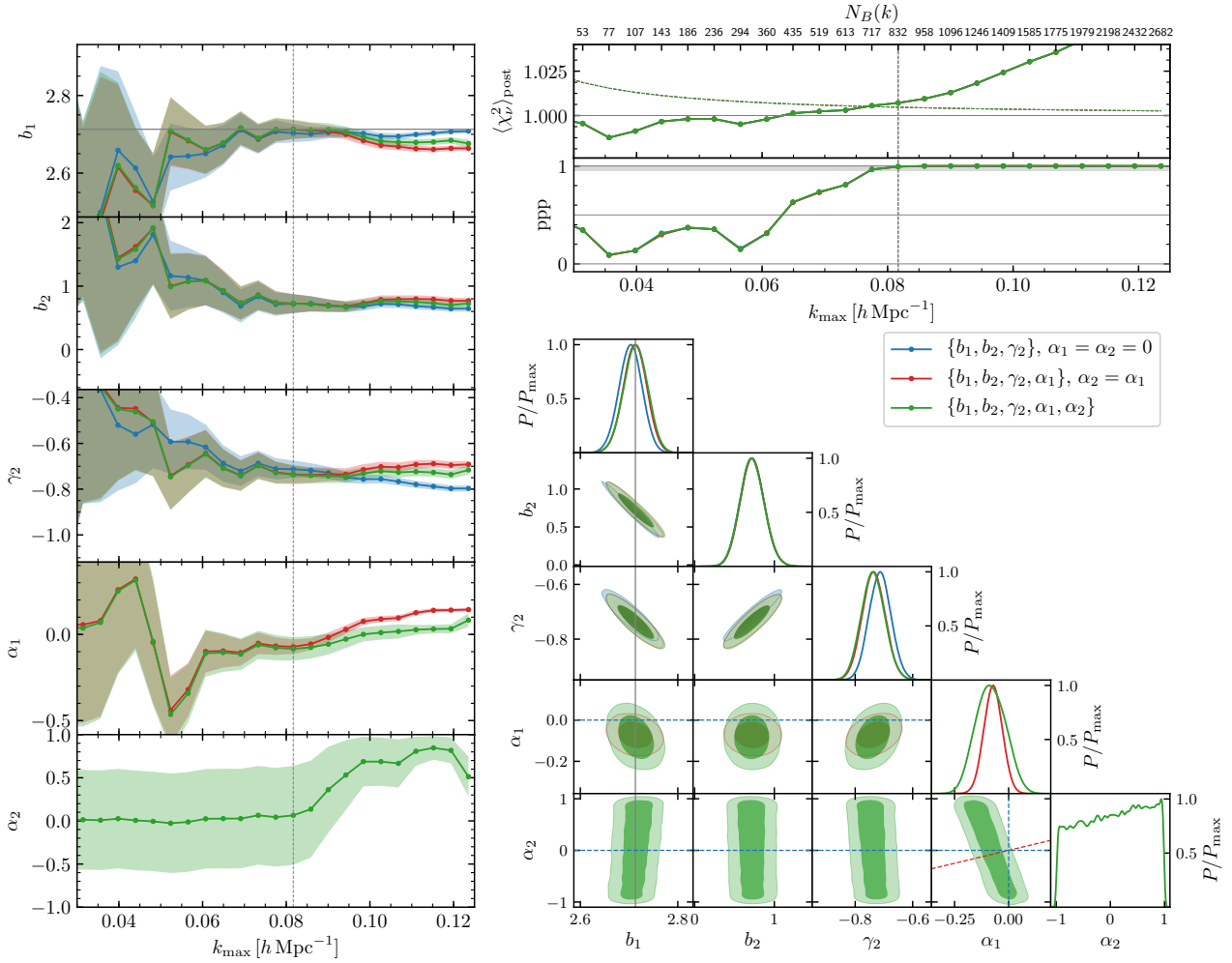
Figure 4.4: Same as figure 4.3 but assuming narrow priors on the shot-noise parameters, *i.e.* $\alpha_{1,2} \in [-1, 1]$. While α_2 is completely unconstrained where the model still gives a good fit to the data, the posteriors for the other parameters do not change sensibly when adding parameters.

from the Savage-Dickey ratios (not shown in the figure for the sake of brevity).

A second series of tests is conducted in figure 4.7. Here, we consider the models $\mathcal{M}_{1f}$, $\mathcal{M}_{2b_2}$ and $\mathcal{M}_{2\gamma_2}$ that have been obtained by mathematically relating the bias parameters of $\mathcal{M}_3$. It is important to remember that the relations we use have different origins. While the function $\tilde{b}_2(b_1)$ is a fit to N-body simulations [71], $\gamma_2(b_1)$ embodies an assumption characterizing a class of bias models that have been already ruled out by recent studies [126, 127]. We find that all models achieve essentially the same goodness of fit as our benchmark model. On the other hand, the posterior distributions of the parameters as a function of $k_{\max}$ show significant differences between $\mathcal{M}_3$ and all cases in which the local-Lagrangian relation for γ_2 is adopted. Only the results obtained with $\mathcal{M}_{2\gamma_2}$ are largely consistent with the reference model. One can also notice how assuming the $\tilde{b}_2(b_1)$ relation greatly reduces error bars with respect to what happens when assuming instead the $\gamma_2(b_1)$ relation. This is probably related to the fact that b_1 and b_2 are more strongly degenerate than b_1 and γ_2 in the fit for the benchmark model as evidenced in the contour plots at the reference scale. Also note that, when we impose one single relation among the bias coefficients, the joint posteriors for the parameters, including the derived ones, are fully consistent with the results obtained for $\mathcal{M}_3$ at the reference scale. On the other hand, if both relations are imposed simultaneously, we find significant tension in the joint posteriors
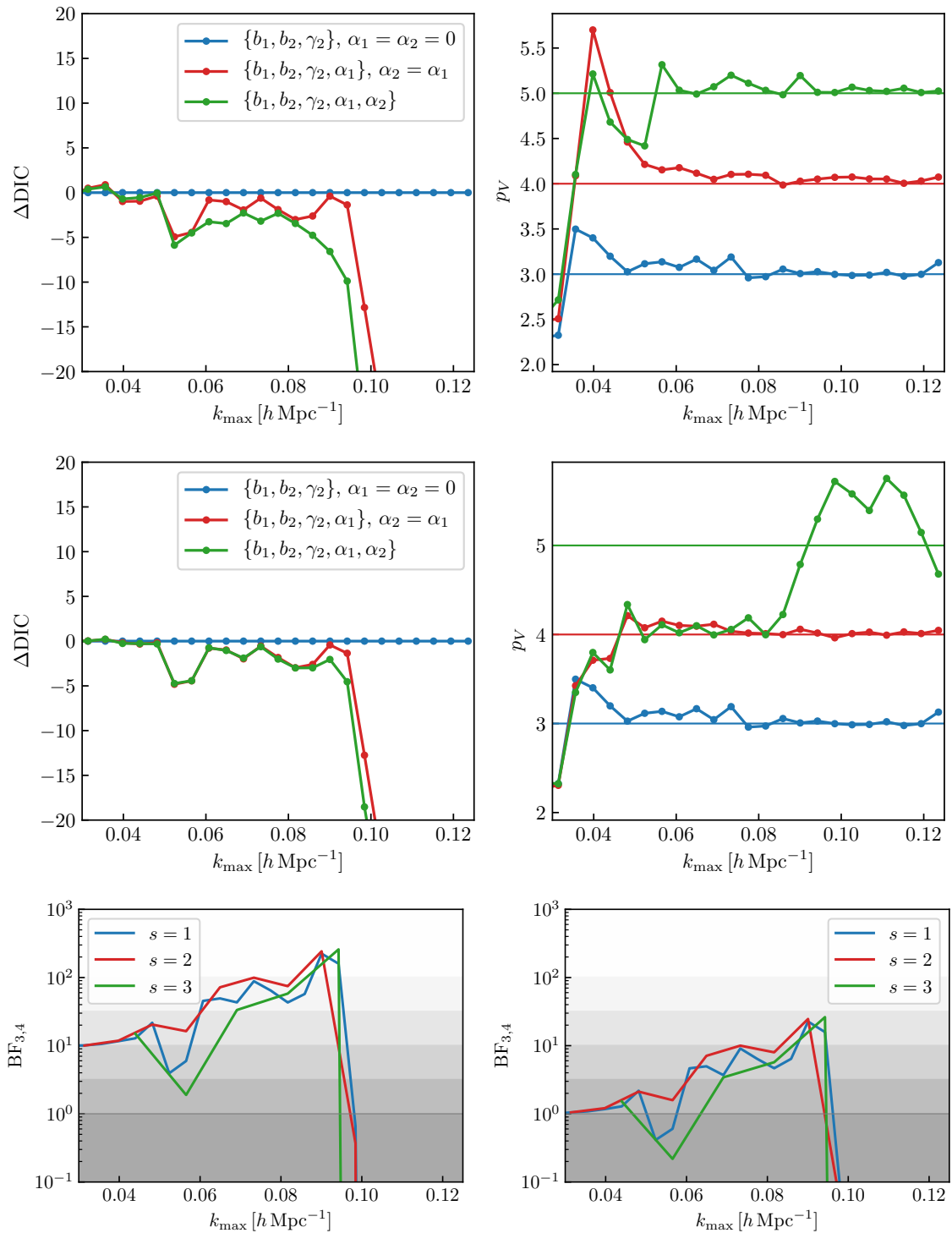
Figure 4.5: Model comparison between the fits presented in figure 4.3 and figure 4.4. The DIC difference with respect to model $\mathcal{M}_3$ is presented in the top-left panel as a function of $k_{\max}$ and the corresponding effective number of parameters p_V is displayed in the top-right panel for the case of broad priors, while corresponding quantities for narrow shot-noise priors are shown in the middle-left and middle-right panels. The bottom panels show the BF from the Savage-Dickey density ratio for the comparison between $\mathcal{M}_3$ and $\mathcal{M}_4$ for broad priors (bottom-left) and for narrow priors (bottom-right). In this case, results are shown for the three different bin widths, $s = 1, 2$ and 3 (solid blue, red and green respectively). The shaded areas represent the Jeffreys' scale for the BF and correspond to regions where the evidence for model i over model j is, from bottom to top, "negative", "barely worth mentioning", "substantial", "strong", "very strong" and "decisive". Model $\mathcal{M}_3$ appears to be preferred over $\mathcal{M}_4$ and $\mathcal{M}_5$ according to both the model selection tests (BF$_{4,5}$ not shown). Note that, while the DIC is slightly larger in the case of $\mathcal{M}_3$, the difference is not so large to favour the other models (that also have a larger number of parameters).
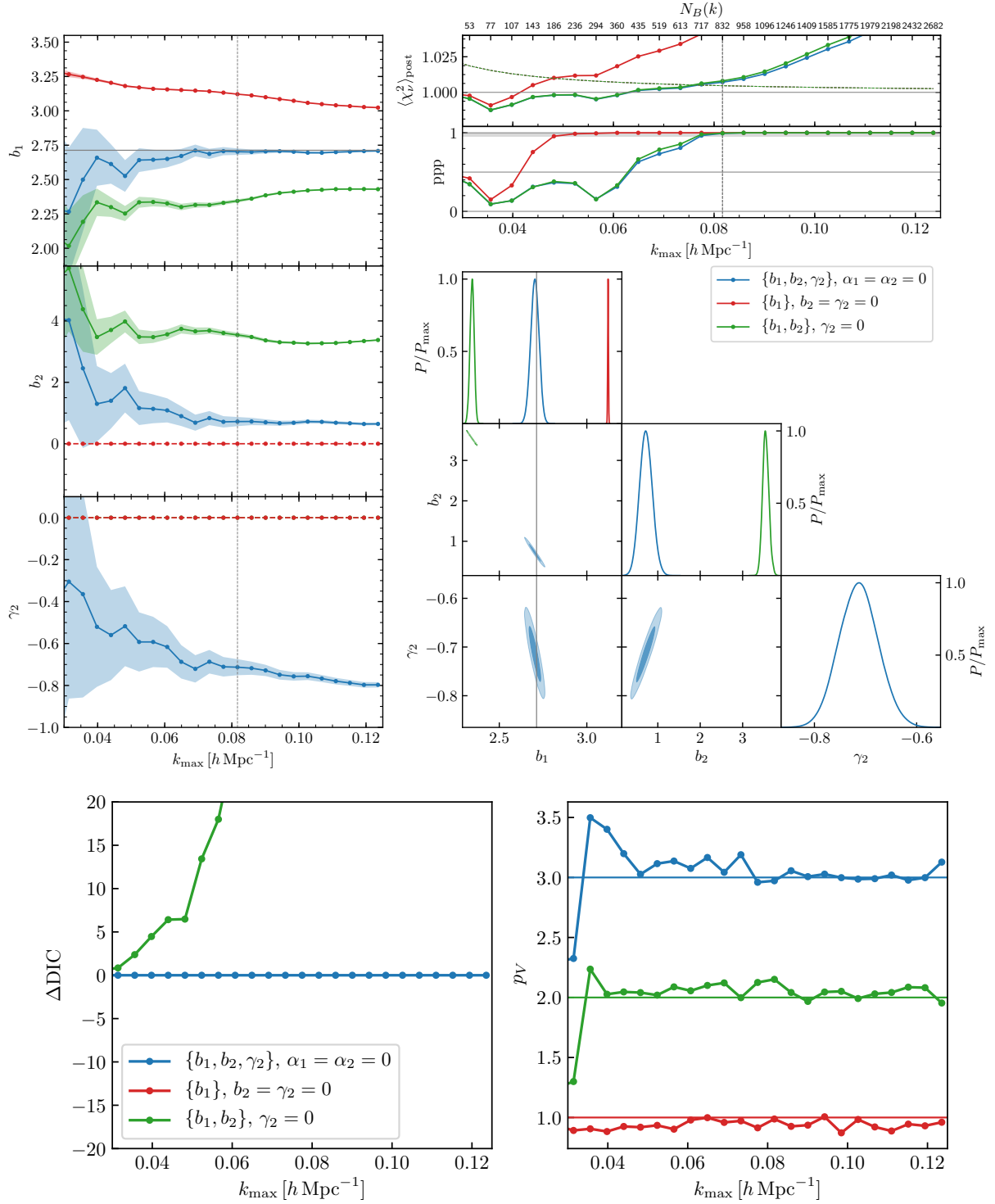
Figure 4.6: As in figures 4.3 and 4.5 but now comparing models $\mathcal{M}_1$, $\mathcal{M}_{2\text{loc}}$ and $\mathcal{M}_3$. With respect to $\mathcal{M}_3$, both $\mathcal{M}_1$ and $\mathcal{M}_{2\text{loc}}$ give biased values for the parameters that have been left free to vary; the data clearly prefers $\mathcal{M}_3$ over the other two models.

with respect to the benchmark model, thus suggesting that the local-Lagrangian approximation is not compatible with the $\tilde{b}_2(b_1)$ fit. In fact, the DIC clearly disfavours both $\mathcal{M}_{1f}$ and $\mathcal{M}_{2b_2}$ while $\mathcal{M}_{2\gamma_2}$ and $\mathcal{M}_3$ are essentially equivalent up to $0.1\,h\,\text{Mpc}^{-1}$. It is important to understand how the figure on the goodness of fit (in which all the models seem to fit the data equally well) can be reconciled with the conclusions we have drawn from the DIC. The key is to remember that we are fitting a very
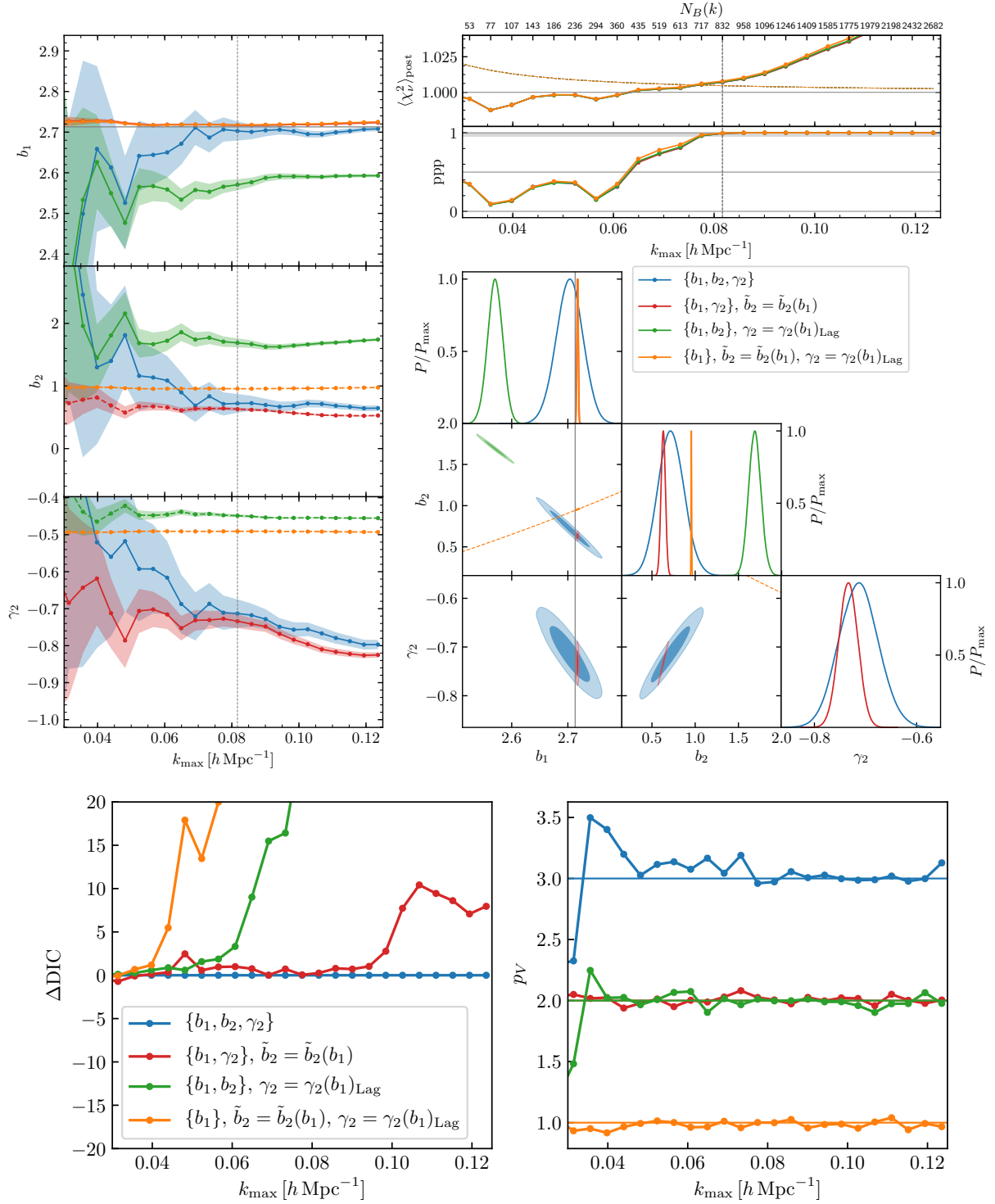
Figure 4.7: As in figures 4.3 and 4.5 but now comparing models $\mathcal{M}_{1f}$, $\mathcal{M}_{2b_2}$, $\mathcal{M}_{2\gamma_2}$ and $\mathcal{M}_3$. The orange dashed lines in the contour plot represent the combination of the fitting function $\tilde{b}_2(b_1)$ presented in [71] and the relation $\gamma_2(b_1)$ derived from the assumption of local-Lagrangian biasing. Model $\mathcal{M}_{2b_2}$ gives biased values for the fitted parameters as a function of $k_{\max}$ with respect to $\mathcal{M}_3$; while model $\mathcal{M}_{1f}$ gives values of b_1 consistent with respect to the ones found with $\mathcal{M}_3$, the values of the derived parameters are biased with respect to $\mathcal{M}_3$; model $\mathcal{M}_{2\gamma_2}$ appears to be consistent with $\mathcal{M}_3$, in posteriors for the parameters and also according to the DIC. Both $\mathcal{M}_{1f}$ and $\mathcal{M}_{2b_2}$ are disfavoured according to the DIC.

large number of data points, namely 247,936 at the reference scale of $k_{\max} = 0.082\,h\,\mathrm{Mpc}^{-1}$. The DIC is driven by the fact that the posterior average of $-\log\mathcal{L}$ for $\mathcal{M}_3$ is smaller than that obtained with $\mathcal{M}_{1f}$ by nearly 18. Since this difference is substantially larger than the number of extra parameters in

$\mathcal{M}_3$, the DIC prefers this model. On the other hand, the goodness-of-fit statistic $\langle\chi^2_\nu\rangle_{\mathrm{post}}$ essentially depends on the average log-likelihood per data point which is very similar for all models. Therefore, although all models considered in this section provide a statistically acceptable fit of the data, $\mathcal{M}_{2\gamma_2}$ and $\mathcal{M}_3$ are preferred as they better describe our ensemble of measurements on small scales and do not overfit.

In all fits we have analysed so far, the posterior distribution of b_1 was always in good agreement with the measurement of $b_1^\times$. However, this is not the case for many of the models presented in figure 4.7. For instance, $\mathcal{M}_{2b_2}$ gives a strongly biased estimate of b_1 (with respect to $b_1^\times$) for wavenumbers larger than $\sim 0.06\,h\,\mathrm{Mpc}^{-1}$. In addition, while the posteriors of b_1 from $\mathcal{M}_{2\gamma_2}$ and $\mathcal{M}_{1f}$ appear to be closer to the measured value $b_1^\times$, both of them are actually more than 3σ away from $b_1^\times$. Therefore, of the four models analysed here, only our benchmark model $\mathcal{M}_3$ gives values of b_1 that are fully consistent with $b_1^\times$, and thus it is the only model giving fully unbiased values of the parameters. The model-selection techniques we implemented allowed us to single out the reference model without prior knowledge of the actual values of the parameters. We envision that model-selection diagnostics will be particularly useful when considering more complex theoretical models that include bias loop corrections and depend on a much larger number of parameters.

4.3.4   Binning of theoretical predictions

All results presented so far rely on averaging the models for the bispectrum over all fundamental triangles that correspond to a given triangle bin following eq. (2.153). In this section, we investigate the impact of using simpler but less accurate theoretical predictions that require a single evaluation of the bispectrum at a triplet of effective wavenumbers defined either as in eq. (2.155) or (2.156). Figure 4.8 shows the influence of the different methods on the fit of model $\mathcal{M}_3$ to the bispectrum data with $\Delta k = 3\,k_f$, for which we expect the largest variations. For the $k_{\max}$ range in which the fit is good, the posteriors for b_1 obtained with the sorted effective wavenumbers lead to rather small differences with those obtained with exact binning, while b_2 and γ_2 show a deviation of $\sim 0.5\sigma$; using the unsorted effective wavenumbers introduces a slightly larger bias in all parameters.

We are not showing here the results corresponding to theoretical predictions evaluated on the triangle bin centers as this approach applies only "closed" triangle bins. A comparison limited to such configurations would show, as we can expect, substantially biased estimates for the model parameters.

4.3.5   Likelihood function

We now study how the shape of the likelihood function influences our results. To this purpose, we compare four options: i) the likelihood function introduced by [69] and presented in eq. (4.5) that we have used to derive all the results presented so far; ii) a Gaussian likelihood combined with the unbiased estimate of the precision matrix given in eq. (4.6) [130, 136]; iii) the same as the previous case but considering only the diagonal part of the uncorrected covariance matrix for the data (*i.e.* setting to zero all off-diagonal elements); iv) the same as the previous case, but considering a Gaussian variance estimated using the full non-linear halo power spectrum as measured from the N-body simulations. The corresponding fit results are displayed in figures 4.9 and 4.10 for the data with $\Delta k = k_f$ and $3k_f$, respectively. All cases refer to model $\mathcal{M}_3$ (fully averaged over the triangular bins). The first striking feature here is that, whenever the goodness of fit is acceptable, the posterior distributions for the
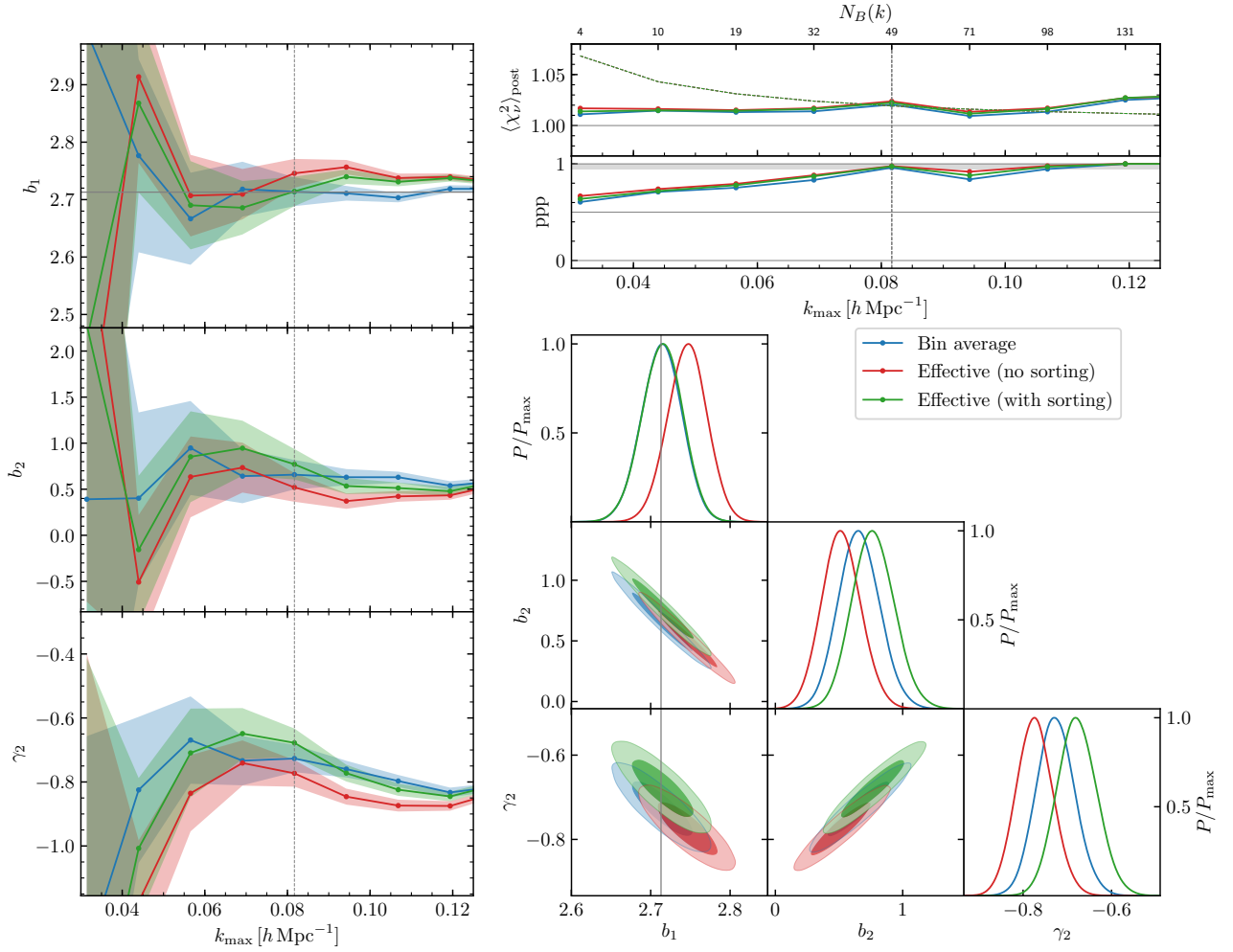
Figure 4.8: As in figure 4.3 with $\Delta k = 3k_f$ but now comparing the different methods to evaluate the theoretical model for the binned bispectrum presented in section 2.8.1 and figure 4.1. While the goodness-of-fit tests do not clearly prefer any of the different methods, the effective approaches give slightly biased values for the parameters, but still consistent with the ones found with a full bin average of the theoretical prediction.

model parameters are practically identical for all likelihood functions. This might be a consequence of the enormous compression involved in our exercise where we use tens of thousands of datapoints to measure only three parameters [6]. Also, the difference in the signal-to-noise ratio between the case of the full-covariance and its diagonal approximation shown in figure 3.11 does not directly reflect in a noticeable difference in the final constraints. Our results are obtained from measurements in periodic boxes at very large scales. Neglecting covariances is likely to produce larger biases in the presence of a window function and extending the analysis to more non-linear scales, when finite-volume effects and non-Gaussianity provide larger contributions to all off-diagonal elements of the covariance matrix [40].

Similarly, the goodness-of-fit statistics derived from the two likelihood functions that account for the off-diagonal covariances coincide almost perfectly. On the other hand, some deviations are noticeable when only the diagonal variances are considered. This approximation leads to rather optimistic estimates of the goodness of fit as a function of k_{max} for $\Delta k = k_f$ and to slightly pessimistic ones for $\Delta k = 3k_f$ when only the diagonal part of the full estimated covariance is considered. On

---

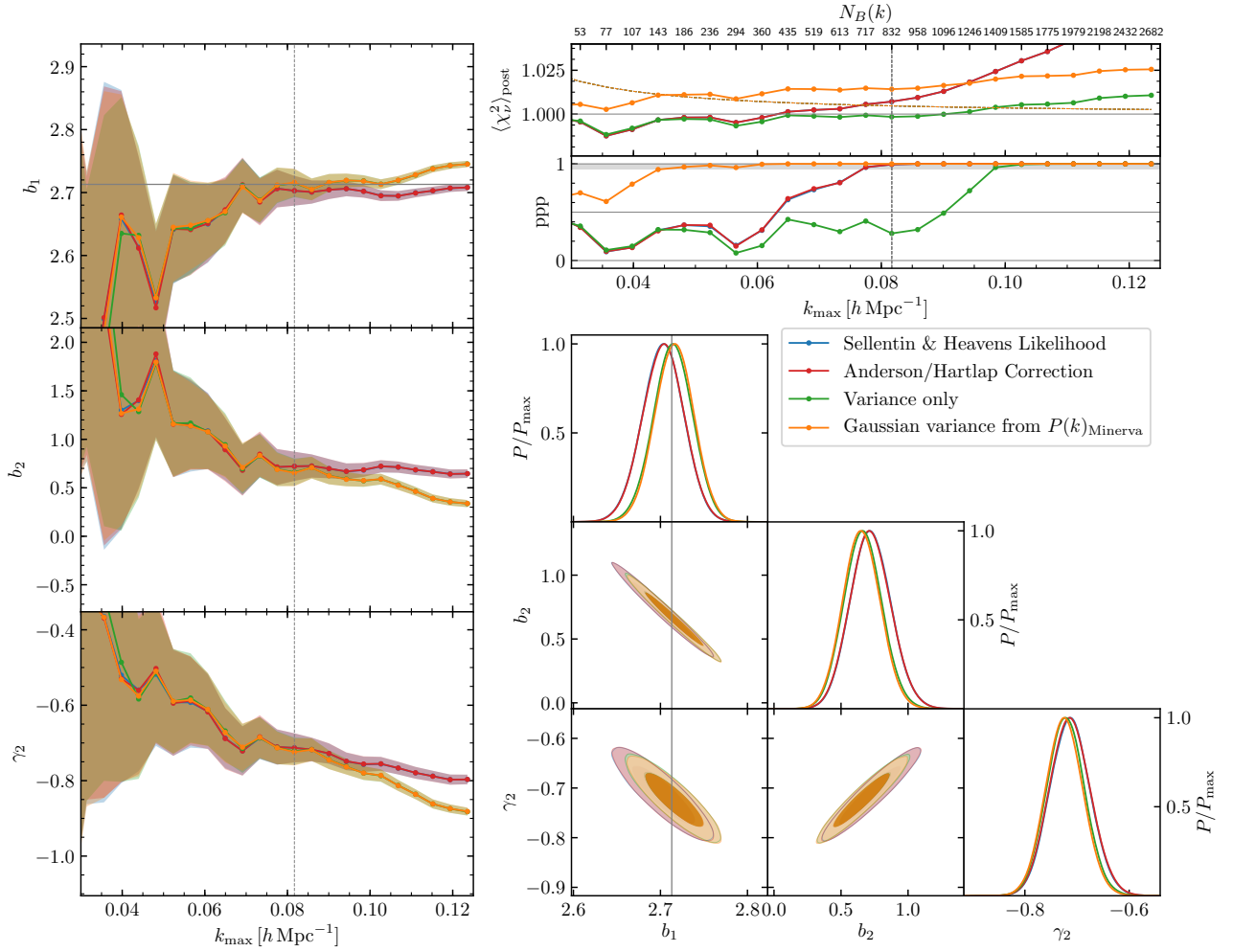[6]Elena Sellentin, private communication.

Figure 4.9: As in figure 4.3 with $\Delta k = k_f$ but now comparing the different functional forms for the likelihood presented in section 4.1.3 as well as the simple case of a Gaussian likelihood evaluated using only the diagonal part of the estimated covariance matrix. While the posteriors do not change in the range of validity of the model, the goodness-of-fit statistics give different results when using only the diagonal from the covariance estimated from the mocks or with the predicted Gaussian variance obtained in terms of the measured power spectrum.

the other hand, in the case of the Gaussian variance prediction obtained from the measured power spectrum, the goodness of fit becomes even worse. It is not easy to identify the precise origin of these effects, but it is important to stress that the variance for the wider bins collects contributions from off-diagonal elements of the covariance matrix for the narrower bins. This probably reduces, to some extent, the difference from the results obtained with the full covariance.

4.4   Conclusions

In this chapter, we discussed how to fit the real-space halo bispectrum at large scales ($k \lesssim 0.1\, h\, \mathrm{Mpc}^{-1}$) with a likelihood-based method. We consider dark-matter halos with a minimum mass of $10^{13}\, h^{-1}\, \mathrm{M}_\odot$ at redshift $z = 1$, extracted from a large set of about three-hundred N-body simulations corresponding to a total volume of $1000\, h^{-3}\, \mathrm{Gpc}^3$. This is much larger than the characteristic volume of current and forthcoming galaxy surveys: for instance, it is 100 times larger than the typical connected region that will be covered by the Euclid spectroscopic survey in a single redshift bin of, say, $\Delta z = 0.1$ [30]. Therefore, a model that fits our data well will generate systematic deviations below 10% of the
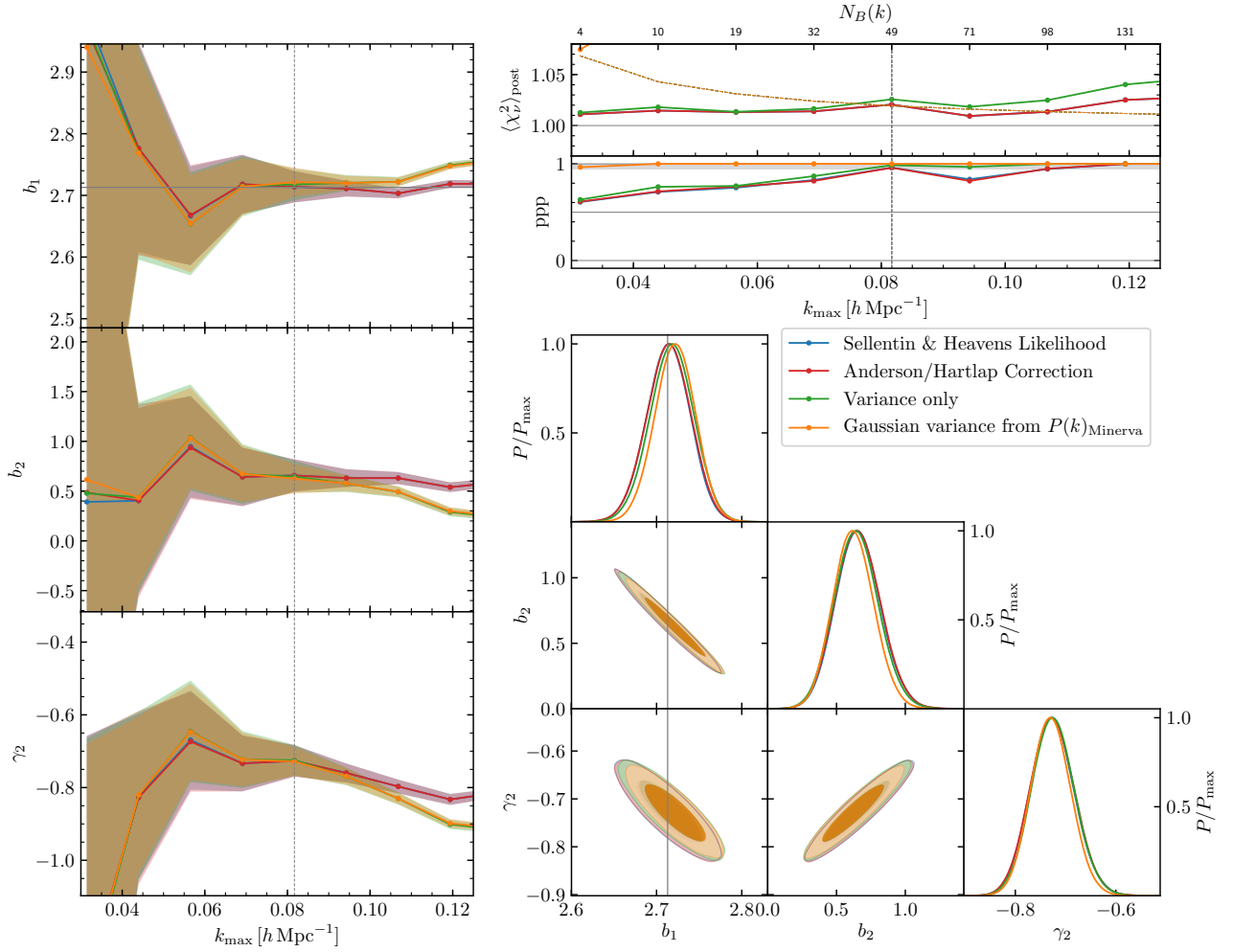
Figure 4.10: As in figure 4.9 but for $\Delta k = 3k_f$.

statistical errors expected for experiments such as Euclid or DESI. We precisely estimate the full, non-linear covariance for our data from an even larger set of 10,000 halo mock catalogs obtained with the Lagrangian PT-based PINOCCHIO code. Statistical errors in the precision matrix are accounted for at the likelihood level with two different methods. We also pay particular attention to reduce any systematic error on the bispectrum covariance by carefully selecting the mass threshold for the mock halo catalogs. This provides a percent-level match to the N-body power spectrum that controls the main Gaussian contribution to the error budget. We consider a tree-level perturbative model for the halo bispectrum that allows quick numerical evaluations and therefore provides us with two advantages: MCMC chains run fast and binned theoretical predictions can be computed exactly. Our benchmark model depends on two local bias parameters, b_1 and b_2, plus the quadratic tidal-field parameter γ_2 and assumes Poissonian shot noise.

With these tools at our disposal, we study in the first place the goodness of fit. This issue is usually neglected in the literature, mainly due to the difficulty in obtaining an accurate and precise estimate of the covariance matrix. We determine the goodness of fit by means of the ppp introduced in section 4.2.2. To connect it with a more familiar diagnostic test, we also compute the posterior average of the reduced χ^2, which gives consistent results. We find that the tree-level model provides a good fit to our data up to 0.08-0.10 $h\,\mathrm{Mpc}^{-1}$ depending on the binning of the Fourier modes adopted

for the bispectrum measurements, that is $\Delta k = s k_f$ with $s = 1$, 2 or 3. The deterioration of the goodness of fit for larger values of k does not necessarily imply that the model needs higher-order corrections as it might instead reflect that our estimate of the covariance matrix for the bispectrum becomes less accurate. The posterior distributions for the bias parameters are largely consistent across the different measurements. When plotted as a function of k_{max}, they begin to show a marked scale dependence that, in conjunction with unusually small errors, indicates a failure for the model for scales $k_{\mathrm{max}} \gtrsim 0.09\,h\,\mathrm{Mpc}^{-1}$, consistent with the results for the goodness of fit.

We investigate several variations of our benchmark model by considering, for instance, either non-Poissonian shot noise or some specific relations between the bias parameters. We then apply Bayesian model selection to determine the optimal number of parameters that are needed to describe the numerical data without overfitting. In particular, our comparison is based on the BF (computed through the Savage-Dickey density ratio for nested theoretical models) and the DIC, both introduced in section 4.2.1. Our results can be summarised as follows.

1. The bispectrum data do not support the introduction of non-Poissonian shot-noise corrections. The BF strongly favours our benchmark model with respect to more elaborated models including additional shot-noise parameters, while the DIC is indecisive in this respect.

2. Model-selection diagnostics clearly favour theoretical descriptions including a tidal bias term over a local Eulerian bias expansion with $\gamma_2 = 0$.

3. A local Lagrangian bias expansion in which $\gamma_2 = -\frac{2}{7}(b_1 - 1)$ is also disfavoured by the data with respect to models in which γ_2 is free to vary. On the other hand, using a fit from the literature to relate $\tilde{b}_2$ and b_1 (while keeping γ_2 free) gives a two-parameter model that is equivalent (in terms of DIC) to our benchmark description and leads to substantially smaller uncertainties for b_1, b_2 and γ_2. In a sense, using this fit is equivalent to combine the bispectrum with other data.

Some of these conclusions confirm what other authors have found with different methods. We find remarkable that model-selection techniques applied to the bispectrum end up preferring exactly those models that are supported by other independent studies even when no difference can be noted in terms of simple goodness-of-fit diagnostics. We envision that model selection will play an important role in the future as the number of parameters controlling loop models for the power spectrum and the bispectrum in redshift space will become particularly large with significant degeneracies among them [159]. Varying the cosmological parameters in order to analyze data from actual surveys will also introduce additional complications. We intend to address these issues in our future work.

Bispectrum measurements are invariably performed within finite bins collecting similar triangles of wavenumbers. Averaging the theoretical predictions over the same bins can be expensive if the model contains loop corrections. For this reason, we test several methods to reduce the number of model evaluations needed in order to fit binned data. We find that a single evaluation per bin (corresponding to a suitably defined effective triangle) generates only very minor systematic errors for $\Delta k = 3\,k_f$. On the other hand, naively using a triangle with sides corresponding to the bin centers leads to distinctly incorrect posteriors for the parameters of the fit.

As a final test, we consider different functional forms for the likelihood distribution. Two of them have been designed to account for statistical errors that plague the estimation of covariance matrices from mock catalogs. The first is obtained by marginalising over the unknown covariance matrix

conditioned to its estimate [69], the second simply combines a Gaussian likelihood with a re-scaled precision matrix [68]. Lastly, as done in many previous works, we use a Gaussian likelihood and only consider either the diagonal part of the estimated covariance or the Gaussian variance prediction obtained in terms of the measured halo power spectrum from the N-body simulations. In all cases, we find essentially the same posterior distribution for the model parameters. However, estimates of the goodness of fit get artificially inflated by neglecting covariances when fitting data within bins of $\Delta k = k_f$ (note that this does not happen for the wider bins). No differences, instead, are noticeable between the two approaches that correct for the noise in the covariance matrix. Based on this, we conclude that, for measurements in periodic boxes, inference performed considering only the variance leads to trustable posteriors while the range of scales over which a model provides an acceptable description of the data should be determined using the full covariance. We expect, however, that neglecting off-diagonal elements could not represent a viable option in an actual galaxy survey. In fact, window-function effects combined with non-Gaussianities in the galaxy distribution should generate significant contributions to the off-diagonal covariance terms.

As already mentioned, this work should be regarded as the first step towards establishing a solid inference method for the galaxy bispectrum. By taking advantage of a reliable estimation of the covariance matrix for the halo bispectrum, we have measured the influence of several, often overlooked, details. In the remainder of this dissertation and in our future work, we will investigate other key methodological aspects. Among them, we will consider theoretical models that include loop corrections and test intrinsically non-Gaussian likelihood functions. Our ultimate goal is to set up a robust pipeline for extracting cosmological information from the joint analysis of the galaxy power spectrum and the bispectrum, whose starting point is represented by the next chapter.

Chapter 5

Joint analysis of the galaxy power spectrum and bispectrum in real space

In this chapter, I present the ongoing effort my collaborators and I are putting forward, aimed to developing a robust likelihood analysis of power spectrum and bispectrum. Our reference model corresponds to the EFTofLSS one-loop prediction for the power spectrum, validated in [47], and the tree-level expression for the bispectrum. In particular, we explore constraints on bias parameters, we investigate possible, physically-motivated extensions of the model that we consider as reference, as well as the possibility to reduce the parameter space through bias relations. We then discuss the impact of different assumptions of the likelihood functions on parameters constraints. Finally, we perform a likelihood analysis to constrain cosmological parameters, consistently re-evaluating the theoretical model every time the cosmology is varied. All of this will be presented in [160].

5.1 Model inference

5.1.1 Theoretical model

As mentioned above, in the following analysis we use a tree-level model to fit the bispectrum data, equation 2.101 and a one-loop EFT model for the power spectrum data, equation 2.92, both in real space. The models will apply to a different range of scale for the two observables, as we expect the one-loop power spectrum model to reach larger values of k. The main goal, therefore, is to investigate how large-scale information from the bispectrum can constrain non-linear bias, and thereby non-linear corrections to the power spectrum. For both observables, we perform IR resummation, as detailed in section 2.7.

In a first stage, we fix the cosmological parameters. With this, the tree-level model for the galaxy bispectrum depends on 5 parameters, $\{b_1, b_2, b_{\mathcal{G}_2}, \alpha_1, \alpha_2\}$, while the 1-loop model for the galaxy power spectrum depends on 6 parameters, $\{b_1, b_2, b_{\mathcal{G}_2}, b_{\Gamma_3}, \tilde{c}_0, \alpha_P\}$, assuming a constant shot-noise contribution. In principle, terms depending on a higher-derivative bias contribution could also be added to the bispectrum model; since the power spectrum depends on the same higher-derivative bias coefficient, it is possible that the bispectrum is able to break the degeneracy between $b_{\nabla^2 \delta}$ and the effective sound speed c_s^2 that controls the amplitude of the EFT counterterm. However, the results presented in chapter 4 show that at the relatively large scales we consider for the bispectrum, $k_{\max} \sim 0.09 \, h \, \mathrm{Mpc}^{-1}$, no k-dependent correction to the linear bias b_1 is needed to correctly describe the measurements,

hence we do not deem these contributions as necessary in our reference analysis. We will test their relevance in section 5.3.2. Higher-derivative terms, however, might become relevant at one-loop, since loop corrections are expected to include k-dependent terms of the same order of the higher-derivative corrections. For this reason, in the following we neglect the dependence of the bispectrum model on higher-derivative bias in our reference case, but we explore the possibility of adding such parameter to our analysis to explore possible extensions of the reference models.

The full set of fitted parameters therefore would be $\{b_1, b_2, b_{\mathcal{G}_2}, b_{\Gamma_3}, \tilde{c}_0, \alpha_P, \alpha_1, \alpha_2\}$; however, in the following, we set $\alpha_2 = 0$, since in the bispectrum-only analysis we were not able to get any constraint on this parameter when assuming some narrow and more physically motivated prior. While the addition of the power spectrum could improve constraints on b_1 and, as a consequence, also on α_2 among the other parameters, we checked that this is, in fact, not the case. Notice that the choice of setting $\alpha_2 = 0$ is, in principle, different from what we assumed in chapter 4, where we assumed $\alpha_1 = \alpha_2$ in model $\mathcal{M}_4$.

Our reference case is therefore represented by the set of parameters $\{b_1, b_2, b_{\mathcal{G}_2}, b_{\Gamma_3}, \tilde{c}_0, \alpha_P, \alpha_1\}$; as already mentioned, however, we investigate the possibility of having also $b_{\nabla^2 \delta}$, α_2, as well as a k-dependent correction to the large-scale shot-noise of the power spectrum, ϵ_{k^2}, that is motivated by stochastic operators in the bias expansion, equation 2.91.

We also consider some relations between the bias parameters in order to reduce the dimensionality of the parameter space. While this is not *a priori* a problem in a joint fit of power spectrum and bispectrum, this is motivated by the fact that the large degeneracies between $b_{\mathcal{G}_2}$ and b_{Γ_3} in the fit of the power spectrum alone make difficult the determination of both bias and cosmological parameters. Reducing the dimensionality of the bias parameter space, moreover, can help significantly in obtaining tighter constraints on the cosmological parameters even in the simultaneous fit of power spectrum and bispectrum; this however can occur in a fully consistent way only if these relations between bias parameters hold true at every stage of the analysis, otherwise the assumption of an inconsistent bias relation could, in theory, lead to biased constraints. For example, in [48], in order to break the degeneracies, the authors set $b_{\Gamma_3} = 0$; however, our full analysis makes clear that this is not an assumption that we can make consistently for our data set, since a value of $b_{\Gamma_3} = 0$ is excluded by many sigmas when a joint fit of power spectrum and bispectrum is performed.

We tested the following relations between bias parameters with joint fits of power spectrum and bispectrum:

- the relation $b_2(b_1, b_{\mathcal{G}_2})$ from [71], namely

$$b_2 = 0.412 - 2.143\, b_1 + 0.929\, b_1^2 + 0.008\, b_1^3 + \frac{4}{3}\, b_{\mathcal{G}_2}; \tag{5.1}$$

  this is a fitting formula coming from measurements of separate universe simulations, and it has been obtained by fitting for values of b_1 in the range $(1, 10)$; notice that the $4/3\, b_{\mathcal{G}_2}$ term in the equation is not present in [71] because of the different bias expansion used in that work;

- the relation $b_{\mathcal{G}_2}(b_1)$ from [72],

$$b_{\mathcal{G}_2} = 0.524 - 0.547\, b_1 + 0.046\, b_1^2; \tag{5.2}$$

  this relation comes from the quadratic fit of the prediction of the tidal bias through the excursion set approach analysed in [161];

- the relation $b_{\Gamma_3}(b_1, b_{\mathcal{G}_2})$ from [59, 72], that with our bias expansion becomes

$$b_{\Gamma_3} = -\frac{1}{6}(b_1 - 1) - \frac{3}{2}b_{\mathcal{G}_2}, \tag{5.3}$$

and is derived by assuming co-evolution.

While this is not an extensive list of all possible relations present in the literature, they represent a good sample to explore. Here, we do not consider the standard local-Lagrangian limit relation between $b_{\mathcal{G}_2}$ and b_1 [70, 92] since different studies have shown the limits of this relation [126, 127], and even in our bispectrum-only analysis of chapter 4, this relation was giving constraints on the bias parameters with large deviations with respect to our reference analysis, where all bias parameters were left free to vary.

In a second stage, in section 5.3.7 we perform a joint analysis of power spectrum and bispectrum at fixed $k_{\max}$ with the objective of recovering the input cosmological parameters of the N-body simulations. In particular, we assume as known the scalar tilt n_s and the baryon content Ω_b – since in actual data analysis these parameters are very well constrained by CMB experiments – and vary the scalar amplitude A_s, the total matter density $\Omega_m h^2$, and the Hubble parameter h.

The theoretical predictions for both power spectrum and bispectrum are always fully averaged over the bins following equations 2.151 and 2.153. However, we also explore the impact on the determination of parameters of evaluating the theoretical predictions of both power spectrum and bispectrum at effective modes.

5.1.2   Likelihood function, priors, and posteriors

The data we fit are the measurements of power spectrum and bispectrum from the N-body simulations described in section 3.1, and we focus only on those measurements with binning scheme defined by $(s = 2, c = 2.5)$.

As for the analysis in chapter 4, we fit all measurements together assuming that they are independent. This means that the total log-likelihood we use to sample the parameter space is given by

$$\log \mathcal{L}_{\text{tot}}(\boldsymbol{\theta}) = \sum_{\alpha} \log \mathcal{L}_{\alpha}(\boldsymbol{\theta}|\mathbf{X}_{\alpha}), \tag{5.4}$$

where the subscript α runs over all realizations and $\mathbf{X}_{\alpha}$ is the dataset of realization α. The dataset $\mathbf{X}_{\alpha}$ represents either the data vector for the power spectrum, for the bispectrum, or the combination of the two. For the individual $\log \mathcal{L}_{\alpha}$, we use two different types of likelihood functions, depending on the type of covariance used.

When the covariance is chosen as the sample covariance of the measurements from the mock catalogues, we assume the Sellentin & Heavens likelihood [69] in order to account for additional statistical uncertainties due to the fact that the covariance is estimated with a limited number of mocks,

$$\log \mathcal{L}_{\alpha} = -\frac{N_M}{2} \ln \left[1 + \frac{\delta\mathbf{X}_{\alpha} \cdot \widetilde{\mathbb{C}}^{-1} \cdot \delta\mathbf{X}_{\alpha}}{N_M - 1} \right] + \ln \left(\frac{\bar{c}_p}{\sqrt{\det \widetilde{\mathbb{C}}}} \right), \tag{5.5}$$

where $\delta\mathbf{X}_{\alpha} = \mathbf{X}_{\alpha} - \mathbf{X}(\boldsymbol{\theta})$ is the vector of to the residuals between realization α and the theoretical model, N_M is the number of mocks, $\widetilde{\mathbb{C}}$ is the estimated covariance and $\bar{c}_p$ is a normalization constant that, since it is model-independent, can be neglected. In the case of a theoretical prediction for the

covariance, we use a Gaussian likelihood

$$\log \mathcal{L}_\alpha = -\frac{1}{2} \delta \mathbf{X}_\alpha \cdot \widetilde{\mathbb{C}}^{-1} \cdot \delta \mathbf{X}_\alpha. \tag{5.6}$$

We study different approximations for the covariance of the combinations of power spectrum and bispectrum, and for each of them we employ the relevant likelihood function.

Notice how, in both cases, the individual log-likelihoods are functions of the chi-square of the dataset of the single realization α, $\chi^2_\alpha = \delta \mathbf{X}_\alpha \cdot \widetilde{\mathbb{C}}^{-1} \cdot \delta \mathbf{X}_\alpha$. This allows us to have a fast implementation of the likelihood when only bias parameters are varied. In the case of the joint analysis between power spectrum and bispectrum, let $\hat{\mathbf{X}}_\alpha = (\hat{\mathbf{P}}_\alpha, \hat{\mathbf{B}}_\alpha)$ be the data vector of the measurements of the simulation α from the set of N_R N-body simulation, and let $\mathbf{X}(\boldsymbol{\theta}) = (\mathbf{P}(\boldsymbol{\theta}), \mathbf{B}(\boldsymbol{\theta}))$ be the theoretical prediction evaluated at a particular point in parameter space $\boldsymbol{\theta}$. The point in parameter space $\boldsymbol{\theta}$ has some components in the subspace of cosmological parameters space $\boldsymbol{\theta}_c$ and others in the subspace of bias parameters space $\boldsymbol{\theta}_b$, such that in general $\boldsymbol{\theta} = (\boldsymbol{\theta}_b, \boldsymbol{\theta}_c)$. When cosmological parameters are fixed, $\boldsymbol{\theta} = \boldsymbol{\theta}_b$, and with our specific choice of theoretical models both power spectrum and bispectrum are just linear combinations of templates, where the coefficients entering the linear combinations are products of parameters in $\boldsymbol{\theta}_b$; this allows us to write the theoretical predictions in our case as

$$\mathbf{P}(\boldsymbol{\theta}_b) = \sum_n \eta_n^P(\boldsymbol{\theta}_b) \boldsymbol{\mathcal{P}}_n \tag{5.7}$$

$$\mathbf{B}(\boldsymbol{\theta}_b) = \sum_n \eta_n^B(\boldsymbol{\theta}_b) \boldsymbol{\mathcal{B}}_n \tag{5.8}$$

where $\eta_n^P(\boldsymbol{\theta}_b)$ and $\eta_n^B(\boldsymbol{\theta}_b)$ are appropriate products of parameters in $\boldsymbol{\theta}_b$, while $\boldsymbol{\mathcal{P}}_n$ and $\boldsymbol{\mathcal{B}}_n$ are theoretical templates of power spectrum and bispectrum, averaged over the discrete Fourier bins. With this in mind, it is convenient to rewrite $\mathbf{X}(\boldsymbol{\theta})$ as

$$\mathbf{X}(\boldsymbol{\theta}) = \sum_n \eta_n^X(\boldsymbol{\theta}_b) \boldsymbol{\mathcal{X}}_n, \tag{5.9}$$

where $\eta_n^X(\boldsymbol{\theta}_b) = (\eta_n^P(\boldsymbol{\theta}_b), \eta_n^B(\boldsymbol{\theta}_b))$, and $\boldsymbol{\mathcal{X}}_n = (\boldsymbol{\mathcal{P}}_n, \boldsymbol{\mathcal{B}}_n)$. Given the estimated covariance matrix $\widetilde{\mathbb{C}}$, the chi-square of the single realization α, with cosmological parameters fixed, is given by

$$\begin{aligned}
\chi^2_\alpha &= (\hat{\mathbf{X}}_\alpha - \mathbf{X}(\boldsymbol{\theta}_b))^T \cdot \widetilde{\mathbb{C}}^{-1} \cdot (\hat{\mathbf{X}}_\alpha - \mathbf{X}(\boldsymbol{\theta}_b)) = \\
&= \hat{\mathbf{X}}_\alpha^T \cdot \widetilde{\mathbb{C}}^{-1} \cdot \hat{\mathbf{X}}_\alpha - 2\hat{\mathbf{X}}_\alpha^T \cdot \widetilde{\mathbb{C}}^{-1} \cdot \mathbf{X}(\boldsymbol{\theta}_b) + \mathbf{X}(\boldsymbol{\theta}_b)^T \cdot \widetilde{\mathbb{C}}^{-1} \cdot \mathbf{X}(\boldsymbol{\theta}_b) = \\
&= \hat{\mathbf{X}}_\alpha^T \cdot \widetilde{\mathbb{C}}^{-1} \cdot \hat{\mathbf{X}}_\alpha - 2\sum_n \eta_n^X(\boldsymbol{\theta}_b) \hat{\mathbf{X}}_\alpha^T \cdot \widetilde{\mathbb{C}}^{-1} \cdot \boldsymbol{\mathcal{X}}_n + \sum_n \sum_l \eta_n^X(\boldsymbol{\theta}_b) \eta_l^X(\boldsymbol{\theta}_b) \boldsymbol{\mathcal{X}}_n^T \cdot \widetilde{\mathbb{C}}^{-1} \cdot \boldsymbol{\mathcal{X}}_l.
\end{aligned} \tag{5.10}$$

If we define the quantities

$$M_\alpha = \hat{\mathbf{X}}_\alpha^T \cdot \widetilde{\mathbb{C}}^{-1} \cdot \hat{\mathbf{X}}_\alpha \tag{5.11}$$

$$M_{\alpha,n} = \hat{\mathbf{X}}_\alpha^T \cdot \widetilde{\mathbb{C}}^{-1} \cdot \boldsymbol{\mathcal{X}}_n \tag{5.12}$$

$$M_{n,l} = \boldsymbol{\mathcal{X}}_n^T \cdot \widetilde{\mathbb{C}}^{-1} \cdot \boldsymbol{\mathcal{X}}_l \tag{5.13}$$

that are independent from the bias parameters, the chi-square for the single realization with fixed

Parameter	Prior (uniform)
b_1	$[0.9, 3.5]$
b_2	$[-4, 4]$
$b_{\mathcal{G}_2}$	$[-4, 4]$
b_{Γ_3}	$[-10, 10]$
$\tilde{c}_0 \, [\, h^{-2} \, \mathrm{Mpc}^2 \,]$	$[-100, 100]$
$b_{\nabla^2 \delta} \, [\, h^{-2} \, \mathrm{Mpc}^2 \,]$	$[-100, 100]$
ϵ_{k^2}	$[-10, 10]$
α_P	$[-1, 1]$
α_1	$[-1, 1]$
α_2	$[-1, 1]$
A_s/A_s^{fid}	$[0.0004, 4.]$
h	$[0.4, 1.]$
$\Omega_m h^2$	$[0.07224, 0.2224]$

Table 5.1: Uniform prior intervals of the model parameters.

cosmological parameters and with our choice for theoretical models, becomes

$$\chi_\alpha^2 = M_\alpha - 2 \sum_n \eta_n^X(\boldsymbol{\theta}_b) \, M_{\alpha,n} + \sum_n \sum_l \eta_n^X(\boldsymbol{\theta}_b) \eta_l^X(\boldsymbol{\theta}_b) \, M_{n,l}; \tag{5.14}$$

therefore, if we compute M_α, $M_{\alpha,n}$, and $M_{n,l}$, the chi-square of realization α is given by a combination of quantities M, that can be precomputed, and known combinations of parameters η. This allows for a fast determination of the chi-square whenever cosmological parameters are fixed. However, whenever cosmological parameters are varied, it's numerically faster to compute the chi-square in the usual way. We did not explore the possibility to run chains with fast and slow parameters [162]: this will be considered in future works.

The parameters in $\boldsymbol{\theta}_b$ that we vary are $\boldsymbol{\theta}_b = \{b_1, b_2, b_{\mathcal{G}_2}, b_{\Gamma_3}, \tilde{c}_0, \alpha_P, \alpha_1\}$. Since α_2 is not constrained by our data, we set it equal to zero. The cosmological parameters we vary are $\boldsymbol{\theta}_c = \{A_s, h, \omega_\mathrm{m}\}$, while the other parameters are fixed to the values used to generate the initial power spectrum and to run the N-body simulations. In general, we choose uniform priors for all parameters as defined in Table 5.1.

As in chapter 4, we evaluate posterior distributions by means of Monte Carlo Markov Chains using the Python code `emcee` [139]. When $\boldsymbol{\theta} = \boldsymbol{\theta}_b$, we evaluate the posterior distribution by simulating 100 dependent walkers; moves are performed using the affine invariant "stretch move" ensemble method from [140] with parallelization, as described in [139]. We run chains for a number of steps equal to $\min(100\,000, 100\tau)$, where τ is the integrated autocorrelation time. Using equations 5.10, we can run chains of this type in $\mathcal{O}(10)$ minutes.

When we include cosmological parameters, we evaluate the posterior distribution by simulating 8 independent walkers; moves are performed using a Metropolis-Hastings sampler with steps defined by a Gaussian proposal function, with covariance determined iteratively running chains a few times. Chains are run until convergence determined by the Gelman-Rubin diagnostic [163], assuming a precision $\epsilon = 0.05$ and a confidence percentile of 95%. The change in sampling method is due to longer running times when sampling cosmological parameters. At each step, we call the Boltzmann solver `CAMB` [87] to compute the linear power spectrum, we compute loop corrections using a custom implementation based on the `FAST-PT` code [164], and also perform our IR-resummation routine; the power spectrum

is exactly averaged over the discrete Fourier grid. We also compute the tree-level bispectrum and we average it exactly over the discrete Fourier grid, using a fast code implemented in our likelihood. This allows us to have a likelihood evaluation (and thus one step) in ~ 1.5 s, and therefore to reach convergence in a relatively short time, $\mathcal{O}(10)$ hours.

Marginalized one-dimensional and two-dimensional posteriors distributions are shown in triangle density plots generated through the code `GetDist` [165].

5.2   Goodness-of-fit and model selection

As a way to compare the quality of the fits we perform, we compute the posterior predictive p-value ppp and the posterior-averaged reduced chi-square $\langle \chi_\nu^2 \rangle_{\text{post}}$. For details on the particular choice of these diagnostics, we redirect the reader to chapter 4; for the purpose of this analysis, suffice to say that we consider a value of ppp ≥ 0.95 to signal a failure of the model in reproducing a good fit to the data. Moreover, we compare $\langle \chi_\nu^2 \rangle_{\text{post}}$ to the corresponding upper one-sided 95 percent confidence limit associated to a number of degrees of freedom equal to the total number of data points fitted by the model; whenever $\langle \chi_\nu^2 \rangle_{\text{post}}$ is greater than this confidence level, we claim that the model is failing to reproduce the data.

However, our main goal is to extract *unbiased* values of the fitted parameters. When fitting cosmological parameters from measurements performed on N-body simulations, the real values of the parameters are known, and in this situation it is straightforward to compare the posteriors from the MCMC simulations and the input values. When only dealing with bias parameters we do not know, a *priori*, their value with great precision. For this reason, we consider measurements of the linear bias parameter b_1 and the Poisson shot-noise correction α_P from the halo-matter cross-power spectrum $P_{hm}(k)$ and the matter auto-power spectrum $P_m(k)$ from the N-body simulations, following the methodology described in section 3.1.2. Notice that while alternative methods exist that use the halo auto-power spectrum instead of the halo-matter cross power spectrum to compute the value of the linear bias, this can give in general biased values of b_1, because the large-scales limit of the halo power spectrum is $P_h(k \to 0) \simeq b_1^2 P_m(k) + P_{\text{shot}}$, where P_{shot} is the shot-noise. Estimating the shot-noise power spectrum *a priori* is in general non-trivial, since the only estimate one can make assumes a purely Poisson shot-noise. For high-mass halos, this is generally not the case [166, 167, 168, 158].[1] Therefore, in order to have a measure of b_1 which is independent from our likelihood, we fit the large-scale behaviour of $P_{hm}(k)/P_m(k)$ with a constant plus a k^2-correction as explained in section 3.1.2 and define b_1 as the fitted constant. This gives us a reference value of $b_1^\times = 2.7081 \pm 0.0012$, with the ×-superscript indicating that this value comes from the *cross*-power spectrum[2].

Measures of the stochasticity (equivalent to our α_P parameter) have been proposed in the literature (e.g. [115]), however given our small statistical uncertainty, we are already sensitive to scale-dependent corrections, ignored in these works, at scales as large as $0.02\,h\,\text{Mpc}^{-1}$. For the shot-noise correction

---

[1] An alternative to this, could be to fit the halo-matter power spectrum along with the halo power spectrum and the halo bispectrum; however, to have full control of that case, we would need the full covariance for all correlation functions including all cross-covariances between the three observables; in our case, this is not possible, since the matter density field is not available for the 10,000 Pinocchio mocks.

[2] Notice that a different value was given in chapter 4; that value was obtained fitting a constant to the large-scale behaviour of $P_{hm}(k)/P_m(k)$. However, due to the larger error-bars there, the posteriors are compatible even with this new value in the same range of validity of the model. On the other hand, the tighter posteriors from the joint analysis required a more accurate determination of the linear bias from the cross-power spectrum.

α_P, using the method described in section 3.1.2, we find a value of $\alpha_P^\times = 0.3052 \pm 0.0020$. However, since k^2 corrections to the stochastic noise, that are not included in our theoretical model, could in theory bias the recovered value of α_P, we do not assume this fiducial value as a relevant reference value to assess possible bias in the results from the likelihood analysis. In any case, we also explore the effect of having a k^2 correction to the shot-noise in our theoretical model and how this impacts the recovery of the shot-noise parameter.

Since in our analysis we explore the possibility to reduce the bias part of the parameter space by imposing some relations between the model parameters, we also want to assess which of these models works better in reproducing our data. In order to do so, we compute the DIC from the MCMC simulations as described in section 4.2 and we use it as model-selection statistics. Since these models are not nested with respect to our reference model, we cannot use the Savage-Dickey density ratio as an estimate of the Bayes Factor.

5.3 Results

In this section, I show the main results of the likelihood analysis of power spectrum and bispectrum. We assume a reference model and assess the validity range of this model, we then explore possible extensions of the reference model and study the possibility of reducing the number of bias parameters using relations between them. We investigate the impact of different covariance approximations and of binning effects on the parameters posteriors. Then we attempt to reduce the bispectrum constraints and extend the bispectrum model to smaller scales by selecting triangle configurations in different ways. Finally, we report on the results of the simultaneous fit of power spectrum and bispectrum including cosmological parameters and a consistent re-evaluation of the theoretical model every time the cosmological parameters are varied.

5.3.1 Reference model

Here we report the results on the analysis of our reference model, where the power spectrum model is assumed at one-loop level and depends on $\{b_1, b_2, b_{\mathcal{G}_2} b_{\Gamma_3}, \tilde{c}_0, \alpha_P\}$, and the bispectrum model is assumed at tree-level and depends on $\{b_1, b_2, b_{\mathcal{G}_2}, \alpha_1\}$. All other bias and shot-noise parameters are set to zero, while the cosmological parameters are fixed to the real values used to run the N-body simulations.

In figure 5.1 we show the parameters constraints of power spectrum alone, bispectrum alone, and joint fit of power spectrum and bispectrum assuming the full mocks covariance and the Sellentin and Heavens likelihood, equation 4.5. In the case of the joint fit, we fix $k_{\mathrm{max,B}} = 0.09\,h\,\mathrm{Mpc}^{-1}$ and show constraints as a function of $k_{\mathrm{max,P}}$. Most of the figures presented in this chapter have the same structure of the ones shown in chapter 4: the left panels show parameter constraints as a function of k_{max}, the upper right panels show the goodness-of-fit measures as a function of k_{max}, while the contour plot shows 1D and 2D marginalized posteriors of the fitted parameters at reference scales, defined below.

While the model used for the bispectrum is slightly different with respect to the 4-parameters model in chapter 4,[3] the goodness of fit statistics are still consistent, giving a range of validity up to $k_{\mathrm{max,B}} = 0.09\,h\,\mathrm{Mpc}^{-1}$, and within this range we also recover an unbiased estimate of the linear bias b_1.

---

[3]There we had $\alpha_1 = \alpha_2$, while here we have $\alpha_2 = 0$ and α_1 free to vary.

The power spectrum model appears to be good in fitting the data for all the considered values of $k_{\mathrm{max,P}}$ up to $0.24\,h\,\mathrm{Mpc}^{-1}$; considering the small error bar on the measurement of $b_1^{\times}$ from the cross power spectrum, we allow for a 2σ difference between $b_1^{\times}$ and the 1D marginalized posterior. Under this condition, the model can recover an unbiased value of b_1 up to $k_{\mathrm{max}} \sim 0.21\,h\,\mathrm{Mpc}^{-1}$. Similarly for the stochasticity α_P compared with the value measured with the combination of the three power spectra. However, inspecting the marginalized posteriors over the full parameter space sampled by the MCMC (in figure 5.1, only part of the parameter space is shown, and at fixed scale $k_{\mathrm{max,P}} = 0.207\,h\,\mathrm{Mpc}^{-1}$, in order to highlight the posteriors from the joint analysis), higher order biases show bimodal posteriors: this is probably due to the fact that we have very small error bars and that we are considering halos of large masses. In fact, for tracers this massive, b_2 is expected to be quite large, meaning that the loop correction proportional to b_2^2 becomes relevant, with positive and negative values of the quadratic bias that lead to equally good fits. Then, the bimodality spreads through degeneracies to the other higher-order biases, $b_{\mathcal{G}_2}$ and b_{Γ_3}.

The constraints for the joint analysis are shown in green in figure 5.1; here we set $k_{\mathrm{max,B}} = 0.09\,h\,\mathrm{Mpc}^{-1}$. While the goodness-of-fit statistics signal an accordance between model and data, the fact that the power spectrum model gives biased values of b_1 for $k_{\mathrm{max}} \gtrsim 0.21\,h\,\mathrm{Mpc}^{-1}$ does not let us trust the joint analysis for k larger than that value. The joint fit shows tighter constraints on all parameters, in particular for the higher-order bias parameters.

The contour plots show the 1D and 2D marginalized posteriors for all parameters; specifically, the power spectrum contours are at $k_{\mathrm{max,P}} = 0.207\,h\,\mathrm{Mpc}^{-1}$, while the bispectrum contours are at $k_{\mathrm{max,P}} = 0.090\,h\,\mathrm{Mpc}^{-1}$; the contours for the joint fit assumes both values of k_{max} at the same time.

Figure 5.2 shows the agreement between the posterior-averaged models of power spectrum and bispectrum in terms of the posterior predictive checks presented in section 4.2.2. The PPCs are shown in the case of the power spectrum only fit, of the bispectrum only fit, and of the joint fit of power spectrum and bispectrum; in all cases, values of $k_{\mathrm{max,P}} = 0.207\,h\,\mathrm{Mpc}^{-1}$ and $k_{\mathrm{max,B}} = 0.090\,h\,\mathrm{Mpc}^{-1}$ are assumed. The posterior-averaged models coming from the joint fit are very well consistent with the models of the fit of power spectrum and bispectrum considered individually.

In any case, it is evident how the inclusion of triangle configurations of the bispectrum even limited to relatively large scales is able to provide relevant information on the bias parameters in general, and on the linear bias in particular, even when the cosmological parameters are assumed to be known. When fitting the power spectrum alone, the linear bias constraints at the reference scale of $k_{\mathrm{max,P}} = 0.207\,h\,\mathrm{Mpc}^{-1}$ were $b_1^P = 2.6946 \pm 0.0082$, compatible with the fiducial value only within $\sim 1.6\sigma$; the addition of the triangle configurations from the bispectrum $k_{\mathrm{max,B}} = 0.090\,h\,\mathrm{Mpc}^{-1}$, constraints the linear bias at $b_1^{P \oplus B} = 2.7083 \pm 0.0020$, perfectly consistent with the value measured from the cross power spectrum, and with a posterior width that is 4 time tighter than the one of the power spectrum only case.

5.3.2 Extensions of the reference model

Here we explore some basic extensions of our reference model, namely the inclusion of the shot-noise parameter α_2 in the bispectrum modelling, of an additional scale-dependent correction to the shot-noise of the power spectrum, and of the higher-derivative bias $b_{\nabla^2\delta}$ to the bispectrum. All these parameters are physically motivated, and thus represent possible extensions of our reference model.

We have already introduced the second shot-noise parameter for the bispectrum. With respect
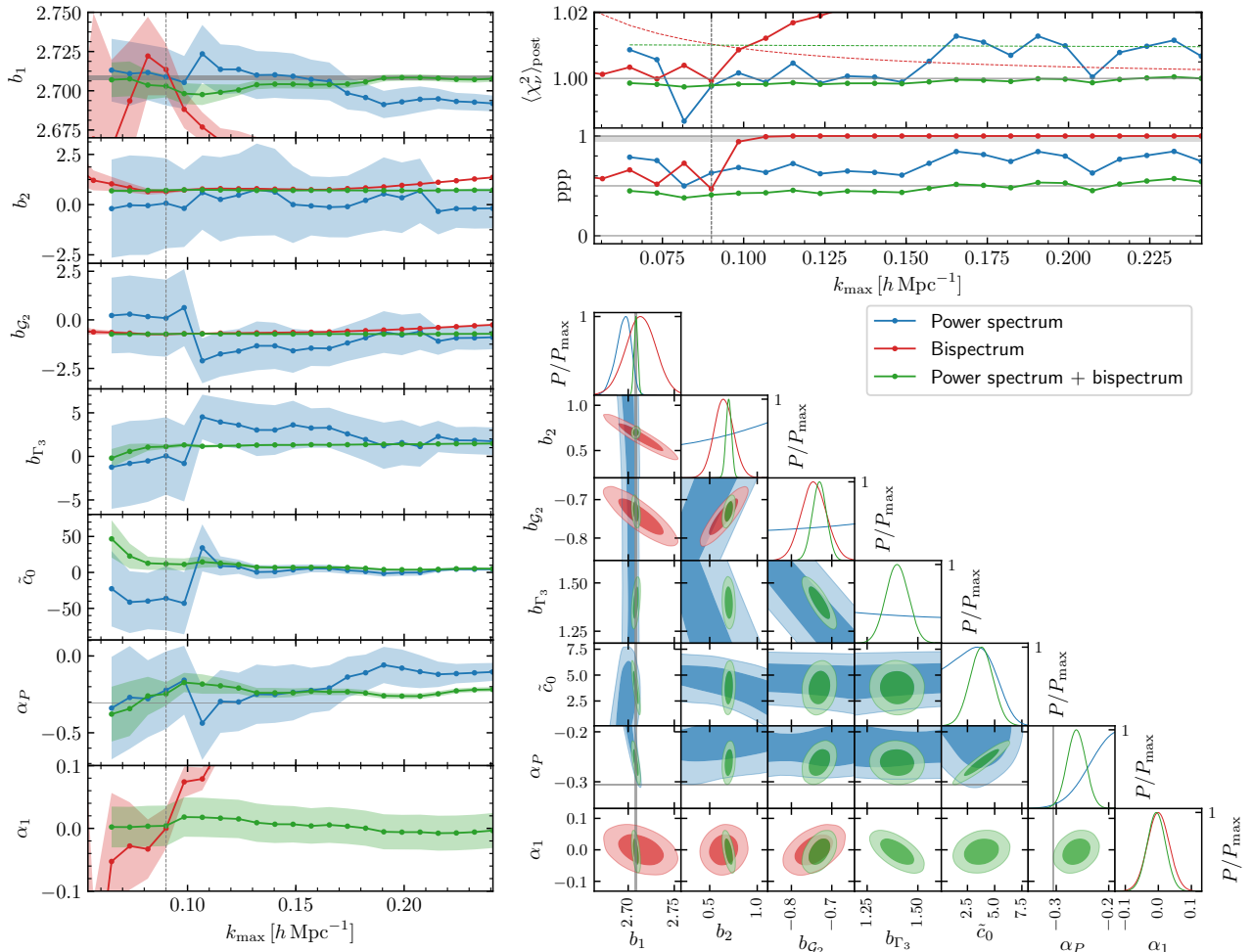
Figure 5.1: Comparison of parameters posteriors and goodness-of-fit statistics for fit of the power spectrum alone (blue), of the bispectrum alone (red), and of the joint fit of power spectrum and bispectrum (green). In the case of the joint fit, $k_{\mathrm{max,B}}$ is set to $0.090\,h\,\mathrm{Mpc}^{-1}$, and the results are plotted as a function of $k_{\mathrm{max,P}}$. The contour plot shows 1D and 2D marginalized constraints for the respective cases with $k_{\mathrm{max,P}} = 0.207\,h\,\mathrm{Mpc}^{-1}$ and $k_{\mathrm{max,B}} = 0.090\,h\,\mathrm{Mpc}^{-1}$

to the reference model considered in the previous section, the scale-independent shot-noise part of the bispectrum model is now let free to vary. We consider as well the addition of a k^2 contribution of the shot-noise of the power spectrum, ϵ_{k^2}, as in equation 2.93. We consider also corrections to the bispectrum model induced by the presence of a higher-derivative bias; for this reason, we add to the tree-level model of the galaxy bispectrum the term in equation 2.102, where corrections due to higher-derivative operators in the shot-noise part of the bispectrum are neglected. The parameter $b_{\nabla^2\delta}$ is the same that enters in the one-loop prediction for the galaxy power spectrum, equation 2.92, where the parameter is completely degenerate with the EFT counterterm. Thus we were considering them together, through the definition of the parameter $\tilde{c}_0 = b_1^2 c_s^2 + b_1 b_{\nabla^2\delta}$. Due to the appearance of the higher-derivative bias in the bispectrum, we could expect that, in a simultaneous fit of power spectrum and bispectrum, the degeneracy between the EFT counterterm and $b_{\nabla^2\delta}$ might be broken.

Figure 5.3 shows the comparison between the reference model and the "maximal" model including α_2, $b_{\nabla^2\delta}$, and ϵ_{k^2}, as well as all the parameters of the fiducial model, both fitted at $k_{\mathrm{max,P}} = 0.207\,h\,\mathrm{Mpc}^{-1}$ and $k_{\mathrm{max,B}} = 0.090\,h\,\mathrm{Mpc}^{-1}$. Keep in mind that introducing the higher-derivative bias has the effect of separating the contributions of $b_{\nabla^2\delta}$ and the effective sound speed of the matter power spectrum c_s^2, that were considered together in $\tilde{c}_0$. Therefore, the $\tilde{c}_0$ considered in the maximal model
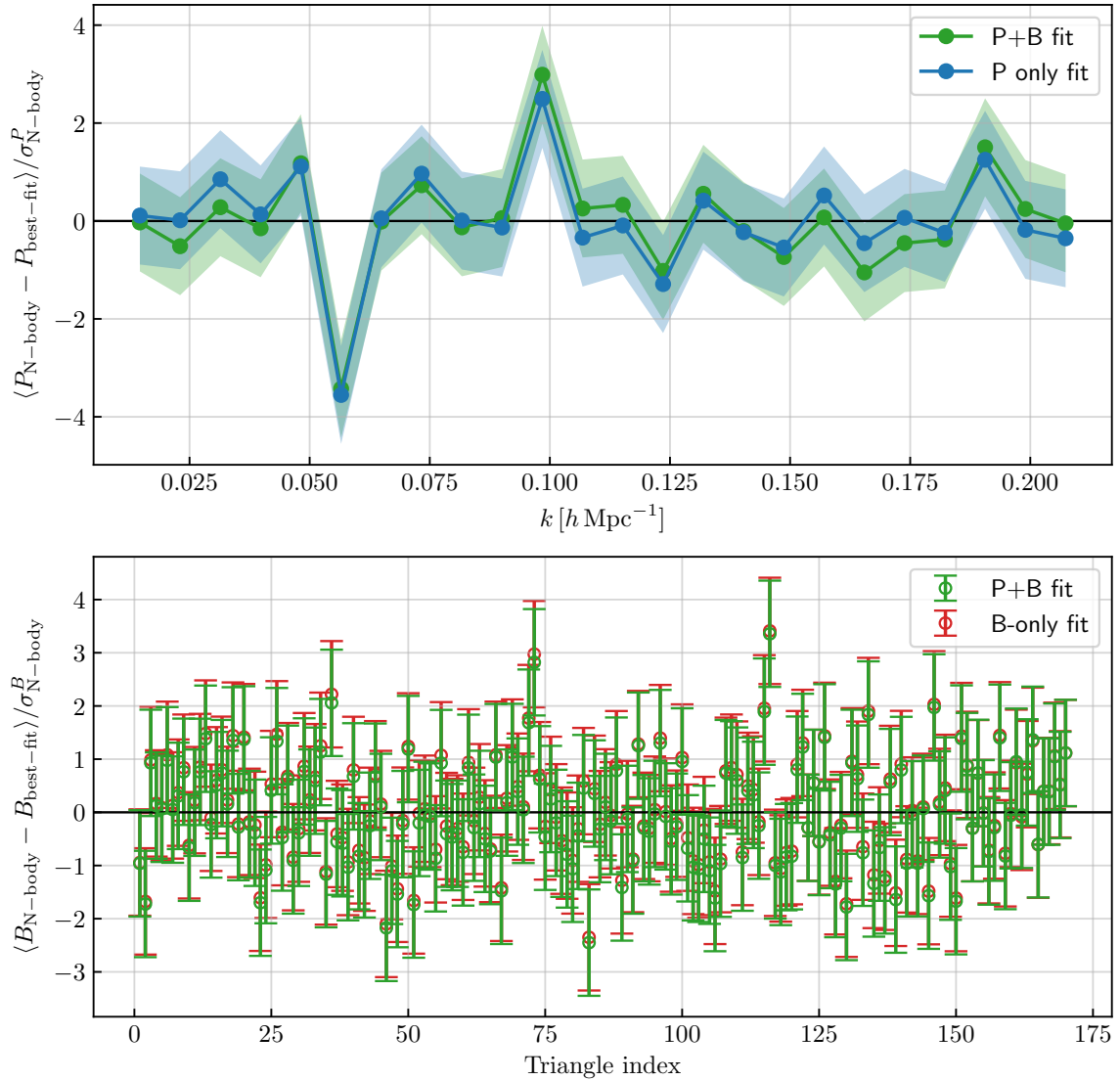
91

Figure 5.2: Posterior predictive checks for the power spectrum (upper panel) and bispectrum (lower panel) data coming from the Minerva N-body simulations. Three cases are shown: the PPCs for the power spectrum-only fit (in blue), the ones for the bispectrum-only fit (red), and the ones for the joint fit of power spectrum and bispectrum (green). The shaded areas in the upper panel and the error bars in the lower panel are relative to a $1-\sigma$ uncertainty in the PPCs. In all cases, values of $k_{\mathrm{max,P}} = 0.207\ h\,\mathrm{Mpc}^{-1}$ and $k_{\mathrm{max,B}} = 0.090\ h\,\mathrm{Mpc}^{-1}$ are assumed.

has to be considered simply as $b_1^2 c_s^2$.

The shot-noise parameter α_2 is non-degenerate with most of the other parameters, except for some mild degeneracy with b_2 and α_1; the stochastic parameter ϵ_{k^2} shows the largest degeneracy with b_{Γ_3} and some milder degeneracies with the linear bias b_1 and the EFT counterterm $\tilde{c}_0$. The higher-derivative bias $b_{\nabla^2\delta}$ is, as expected, very degenerate with the EFT counterterm, proving that the bispectrum can only reduce, to some extent, the degeneracy between these two parameters.

Notice that in the maximal model, the value of α_P measured from the the large scale limit of $P_{hh}(k)$ is within 1σ of the posterior, while this was not the case in the reference model. The fact that the posterior of α_P now agrees with the fiducial value is likely due to the degradation of the posterior caused by the large degeneracy between $b_{\nabla^2\delta}$ and $\tilde{c}_0$. To investigate this further, we perform the following test. First of all, we fit the matter power spectrum up to $k_{\mathrm{max}} = 0.15\ h\,\mathrm{Mpc}^{-1}$ with a one-loop level model with three parameters, $\{b_1, c_s^2, \alpha_P\}$; marginalising over the linear bias and the

shot-noise parameter, we find $c_s^2 = 0.65^{+0.03}_{-0.04}\ h^{-2}\,\mathrm{Mpc}^2$; from this, we define a Gaussian prior on c_s^2, with mean given by the posterior average ($0.65\ h^{-2}\,\mathrm{Mpc}^2$), and with a standard deviation roughly 3 times the width of the posterior ($0.1\ h^{-2}\,\mathrm{Mpc}^2$). Then, we run two different MCMC simulations: one in which, starting from the maximal model, $b_{\nabla^2\delta}$ is set to zero; another where we fit the maximal model while imposing the Gaussian prior on $c_s^2 = \tilde{c}_0/b_1^2$. In both of these two runs, α_P is consistent only at the 95% confidence level with the reference value, suggesting that the fact that the maximal model agrees with the fiducial value of α_P at 1σ level is likely due to the large degeneracy between $b_{\nabla^2\delta}$ and c_s^2, rather than because the power spectrum and bispectrum models need the extra parameters introduced in the maximal model in order to consistently describe the measurements. We are nevertheless considering a discrepancy at the 2σ level between the recovered value and the expected value not to be particularly worrying, in part due to the difficulty in determining the latter, as discussed in section 3.1.2. As a further consistency check, we also compute the difference in DIC of these two MCMC runs with respect to the fit performed with the maximal model; we find $\Delta\mathrm{DIC} = 13.30$ for the case with $b_{\nabla^2\delta} = 0$ and $\Delta\mathrm{DIC} = 16.59$ for the case with the Gaussian prior on c_s^2, in both cases disfavoring the maximal model. This is our main motivation for setting $b_{\nabla^2\delta} = 0$ in the reference model.

The shot-noise parameter α_2 is unconstrained and prior-dominated, and thus perfectly consistent with zero. Setting it to zero, gives a $\Delta\mathrm{DIC} = 4.7$ with respect to the maximal model with $b_{\nabla^2\delta} = 0$, favoring the model with $\alpha_2 = 0$.

Finally, from the fit using the maximal model, we can see that the marginalized 1D posterior of ϵ_{k^2}, while preferring values larger than zero, is still consistent with zero. Notice that setting this parameter to zero, as well as $\alpha_2 = 0$ and $b_{\nabla^2\delta} = 0$, leads us back to the reference model we analyzed in the previous section. With respect to the reference model, adding the k^2-dependent stochastic term increases the DIC by 2.4. This is quite a small difference (we assume differences of order $\Delta\mathrm{DIC} \gtrsim 5$ to be substantial). However, the fact that the reference model fits fewer parameters, leads us to prefer the reference model over the extension with ϵ_{k^2}

Notice also that all parameters constraints found fitting the data with the reference model are perfectly consistent with the constraints one can find by fitting the more general maximal model. All these reasons serve as our motivation to use the reference model with $b_1, b_2, b_{\mathcal{G}_2}$, b_{Γ_3}, $\tilde{c}_0$, α_P, and α_1 rather than the extensions considered here.

5.3.3 Bias relations

Here we explore the possibility of reducing the parameter space by introducing relations between bias parameters. Such relations, particularly in the case of halo bias parameters, are often physically motivated in the context of the Halo Model or are simply described by fits to numerical simulations. We explore three relations: equation 5.1 for $b_2(b_1, b_{\mathcal{G}_2})$ (from [71]), equation 5.2 for $b_{\mathcal{G}_2}(b_1)$ (from [72]), and equation 5.3 for $b_{\Gamma_3}(b_1, b_{\mathcal{G}_2})$ (from [59, 72]).

We consider these relations, in the first place, as means to reduce the parameter space in the case of the maximal model. If they fail in this case, we clearly do not expect them to be particularly useful in the more constrained case of the reference one. In any case, we compare posteriors obtained by imposing each of these relations with ones obtained with the reference model and with the ones given by the maximal model, while model-selection comparisons are performed only with respect to the maximal model.

Figure 5.4 shows the comparison of the parameter posteriors for the simultaneous fit of power
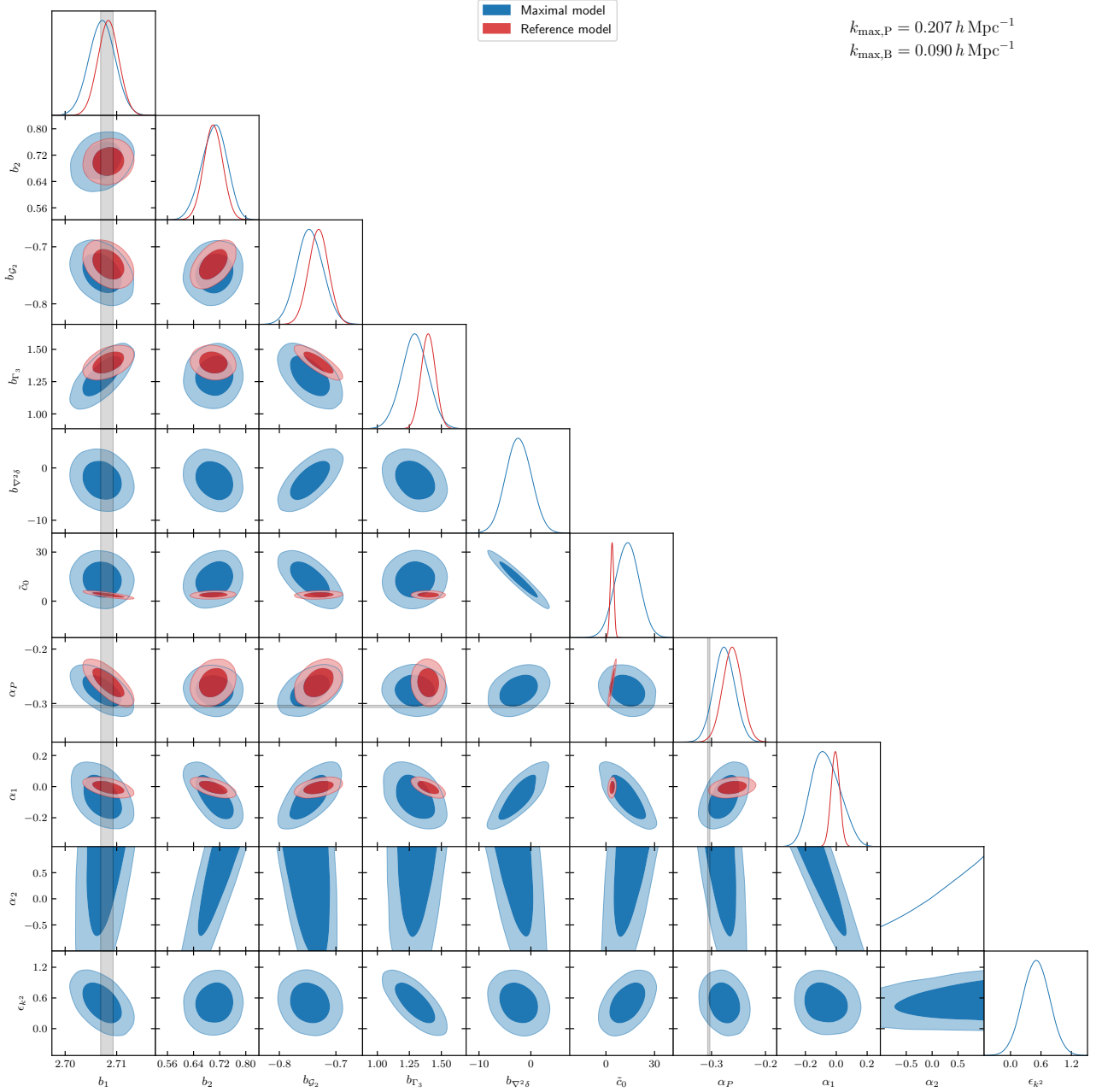
Figure 5.3: Comparison of 1D and 2D marginalized constraints between fits performed using either the reference model (red) or the maximal model (blue) that includes all the parameters of the reference model as well as $b_{\nabla^2\delta}$, α_2, and ϵ_{k^2}. The fit was performed assuming $k_{\mathrm{max,P}} = 0.207\,h\,\mathrm{Mpc}^{-1}$ and $k_{\mathrm{max,B}} = 0.090\,h\,\mathrm{Mpc}^{-1}$. Grey bands show fiducial values for the corresponding parameters coming from direct measurements (see section 3.1.2).

spectrum and bispectrum with different models at the fixed scales $k_{\mathrm{max,P}} = 0.207\,h\,\mathrm{Mpc}^{-1}$ and $k_{\mathrm{max,B}} = 0.09\,h\,\mathrm{Mpc}^{-1}$; it shows in blue the marginalized posteriors for the maximal model, in red the posteriors for the maximal model where the relation $b_2(b_1, b_{\mathcal{G}_2})$ of equation 5.1 has been imposed, and in green for the reference model. With the exception of α_1, all posteriors are consistent at 1σ level, while showing deviations of order $0.5 \div 1\sigma$. Even though in the maximal case α_2, while preferring positive values, is in any case unconstrained, imposing the relation on b_2 forces α_2 toward negative values. The panel showing the 2D posterior for (b_1, α_P) shows agreement with the fiducial values of those parameters.

Figure 5.5 shows a similar comparison, but in red we show the posteriors for the maximal model where the relation $b_{\mathcal{G}_2}(b_1)$, equation 5.2, has been imposed. In this case, the deviations of the posteriors
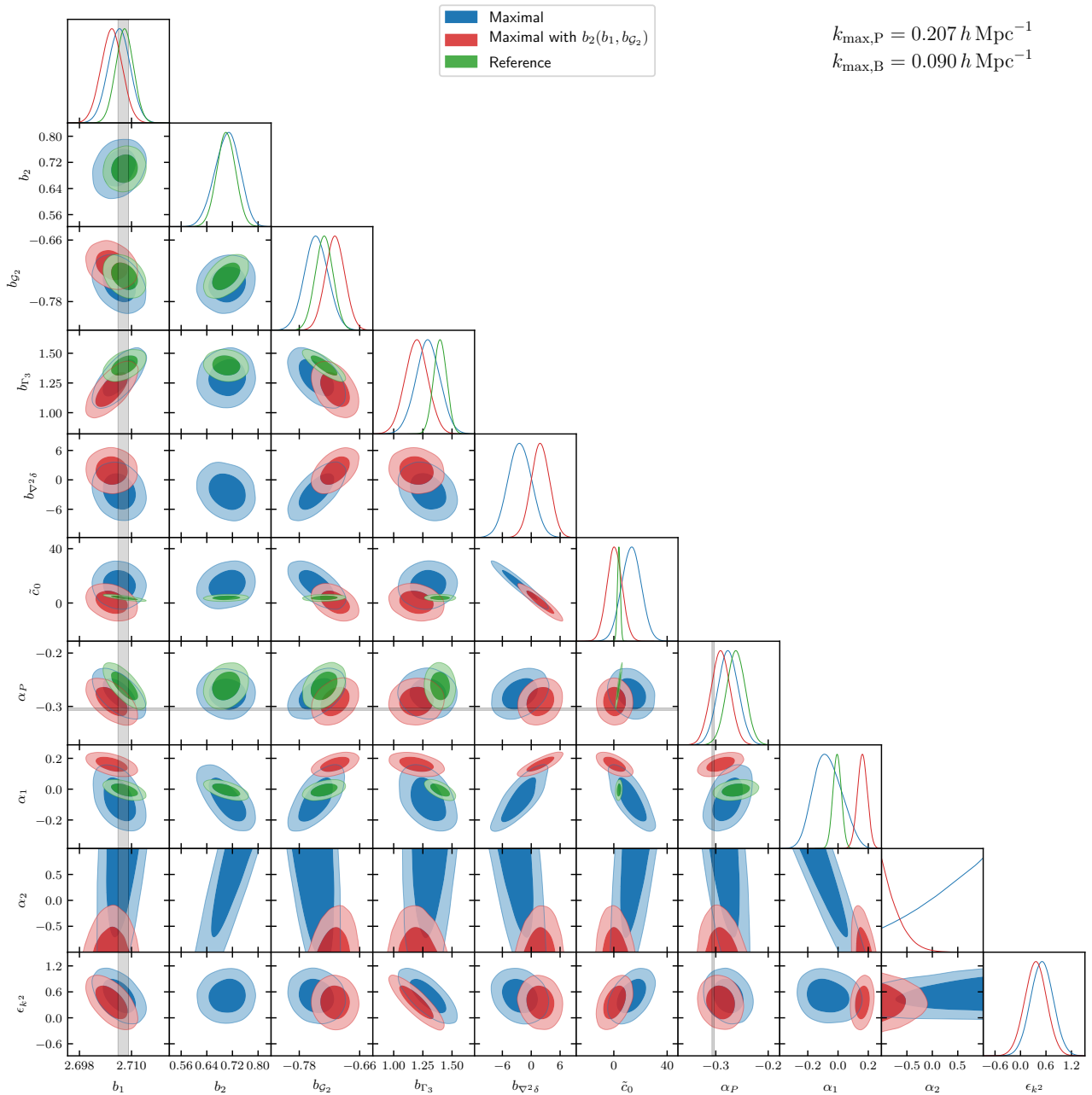
94

Figure 5.4: Comparison between the marginalized 1D and 2D posteriors of the fitted parameters of three different models to the power spectrum and bispectrum data. In blue the posteriors relative to the fit with the maximal model with 10 free parameters; in red, the ones relative to the maximal model with relation 5.1 imposed; in green the ones relative to the reference model with 7 free parameters.

are larger than in the previous one, with b_1 being consistent with the fiducial value only at 2σ, and with important inconsistencies in $b_{\nabla^2 \delta}$, $\tilde{c}_0$, and α_1. Also in this case, α_2 is pushed toward negative values. Notice how the 2D constraints in the (α_1, b_2) panel shows complete disagreement between the case with the maximal model, and the case with maximal model where the bias relation has been imposed.

Finally, figure 5.6 shows the same comparison but considering the bias relation $b_{\Gamma_3}(b_1, b_{\mathcal{G}_2})$, equation 5.3. This time, the posterior for b_1 is inconsistent by more than 2σ level with the fiducial value measured from the cross power spectrum. Important deviations can be seen as well in α_P and ϵ_{k^2}, with the latter showing the largest deviations with respect to the maximal case.
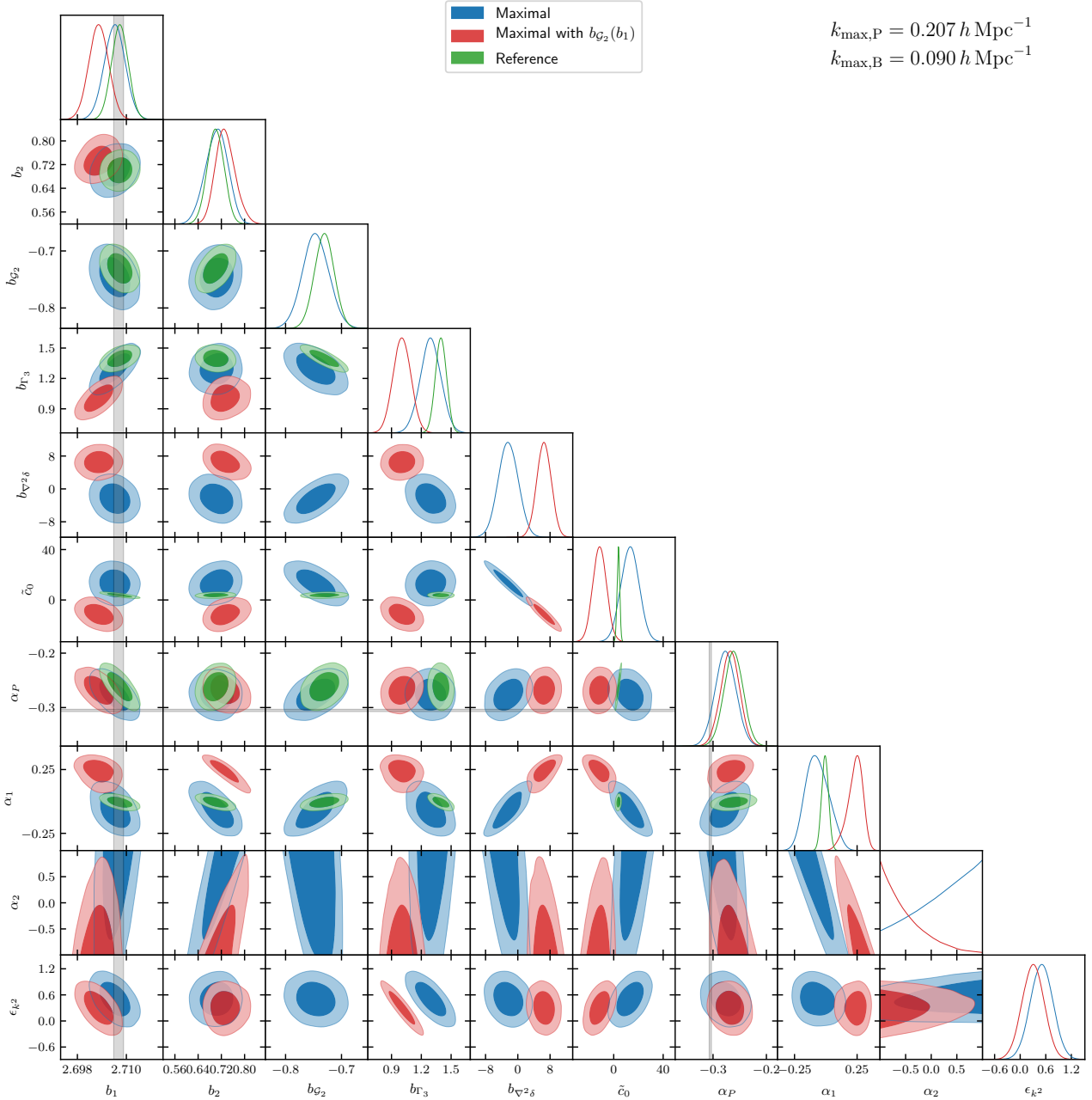
Figure 5.5: Similar to figure 5.4, but comparing the maximal model with relation 5.2 imposed.

Considering exclusively the posteriors, the only bias relation that seems to give relatively good results without biasing too much the other parameters seems to be equation 5.1. Let us now consider the value of the DIC for the maximal models where each bias relation is imposed, and compare these with the value of the DIC of the maximal model. Imposing relation 5.1 increases the DIC by $\Delta\mathrm{DIC}_{b_2} = 1.46$, and according to this value, we can say that the two models are equivalent; since by imposing the bias relation the number of free parameters is reduced, this would lead to a slightly preference toward the maximal model with the bias relation imposed. However, the reference model shows tighter constraints in all parameters, whose posteriors are generally more consistent with the case where the maximal model is used, and the DIC difference with respect to the maximal model is $\Delta\mathrm{DIC} = 23.42$ favoring the reference model. This still makes us prefer the reference model over the maximal model with the $b_2(b_1, b_{\mathcal{G}_2})$ relation imposed. The other bias relations cause the DIC to
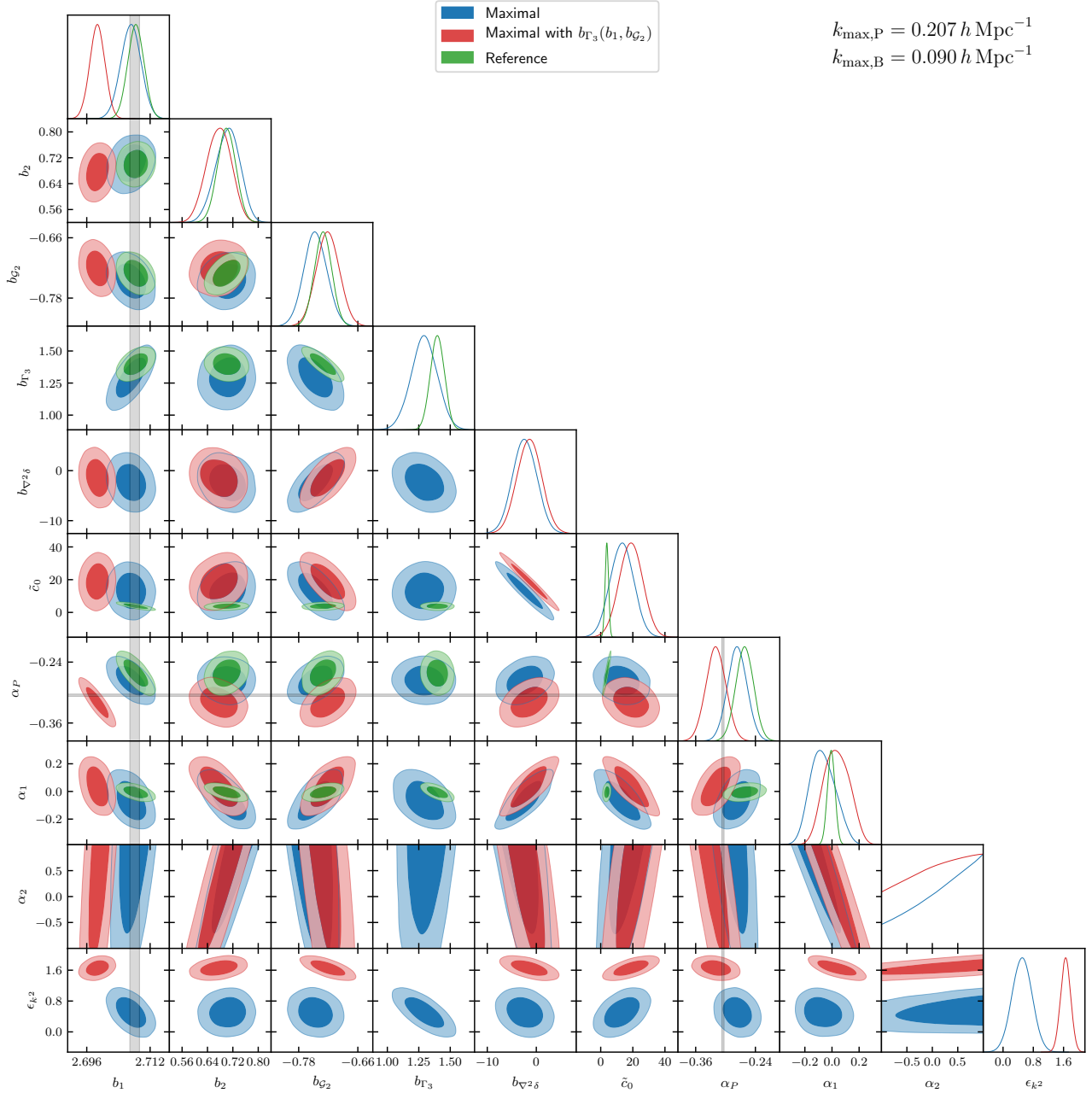
Figure 5.6: Similar to figure 5.4, but comparing the maximal model with relation 5.3 imposed.

increase significantly, with $\Delta\text{DIC}_{b_{\mathcal{G}_2}} = 23.07$ and $\Delta\text{DIC}_{b_{\Gamma_3}} = 17, 65$, suggesting a strong preference for the maximal model over the models with the bias relations imposed.

5.3.4 Covariance approximations

So far, we consistently assumed a covariance matrix for the power spectrum, bispectrum, and their combination estimated from the full set of 10,000 PINOCCHIO mocks. Now we study the effect of common approximations to the full covariance. In particular, we consider three different alternatives:

- a mocks covariance with cross-correlations between power spectrum and bispectrum set to zero.

- the diagonal mocks covariance;

- a theoretical Gaussian prediction corresponding to the diagonal variance for both power spectrum and bispectrum, evaluated in terms of the best-fit bias parameters determined in the previous analysis.

The main goal is to assess increasingly significant levels of approximation and, in particular, the importance of the cross-correlation between the two statistics. Figure 5.7 shows the comparison with our reference case. Again, here we fix $k_{\mathrm{max,B}} = 0.090\,h\,\mathrm{Mpc}^{-1}$ and plot results as a function of $k_{\mathrm{max,P}}$. All three alternatives to the full covariance agree more or less consistently at all values of $k_{\mathrm{max,P}}$; the posteriors agree also with the ones obtained with the full covariance from mocks at smaller scales, while at larger scales ($k_{\mathrm{max,P}} \lesssim 0.1\,h\,\mathrm{Mpc}^{-1}$) posteriors change slightly for most parameters, while remaining more or less consistent at the 1σ level (2σ for some specific cases).

The goodness of fit tests show that the case with the theoretical variance has much lower values of the posterior-averaged reduced chi-square, and the ppp is consistently 0. This is likely due to the fact that the theoretical variance is computed using the posterior-averaged values for the bias parameters, but it assumes the shot-noise contribution coming from the power spectrum to be Poisson ($\alpha_P = 0$). This causes a larger variance both for the power spectrum and the bispectrum. The other two approaches show similar goodness-of-fit statistics, however, with the mocks covariance without cross-correlations behaving slightly better than the mocks variance.

5.3.5   Binning effects

We explore the impact on the posteriors of the parameters of evaluating the theoretical models of the power spectrum and of the bispectrum at effective modes and triangles respectively. In particular, we consider the posteriors of the reference case, where the models are fully averaged, and the case where the power spectrum is estimated at the effective momenta defined in equation 2.148, while the bispectrum is estimated at the effective triangles defined in equation 2.156. Results are shown in figure 5.8 for a joint fit of power spectrum and bispectrum with $k_{\mathrm{max,P}} = 0.207\,h\,\mathrm{Mpc}^{-1}$ and $k_{\mathrm{max,B}} = 0.09\,h\,\mathrm{Mpc}^{-1}$ using the reference model; the case with the full bin-average of the theoretical model is in blue, while the case with the effective model is in red. All posteriors evaluated with the effective prediction show deviations at some degree: while $b_{\mathcal{G}_2}$ b_{Γ_3} and α_1 do not vary significantly, b_2 and $\tilde{c}_0$ show the largest deviations, of the order of 2σ. With the model evaluated at effective modes, the posterior of b_1 is still consistent with the fiducial value, but the comparison with the fiducial value of α_P shows large inconsistencies, at more than 2σ. The 2D posteriors in the $(\tilde{c}_0, \alpha_P)$ shows that the two cases are inconsistent with each other. Moreover, comparing the two models with the DIC, we find that evaluating the theoretical prediction at the effective modes instead of consistently averaging it over the Fourier grid increases the DIC by $\Delta \mathrm{DIC} = 31.47$. All of this seems to suggest that, with our error bars, evaluating the full bin-average of the theoretical predictions becomes crucial, and since in the bispectrum-only case the posteriors were not showing such large deviations, in this case this is possibly an effect due to the power spectrum.

5.3.6   Triangle selection criteria

It is customary in the analysis of bispectrum measurements involving all possible triangular configurations, to assess the validity of the theoretical models as a function of k_{max}, corresponding to the largest side of the triangle. However, it is clear that the validity of perturbative models can be extended to a
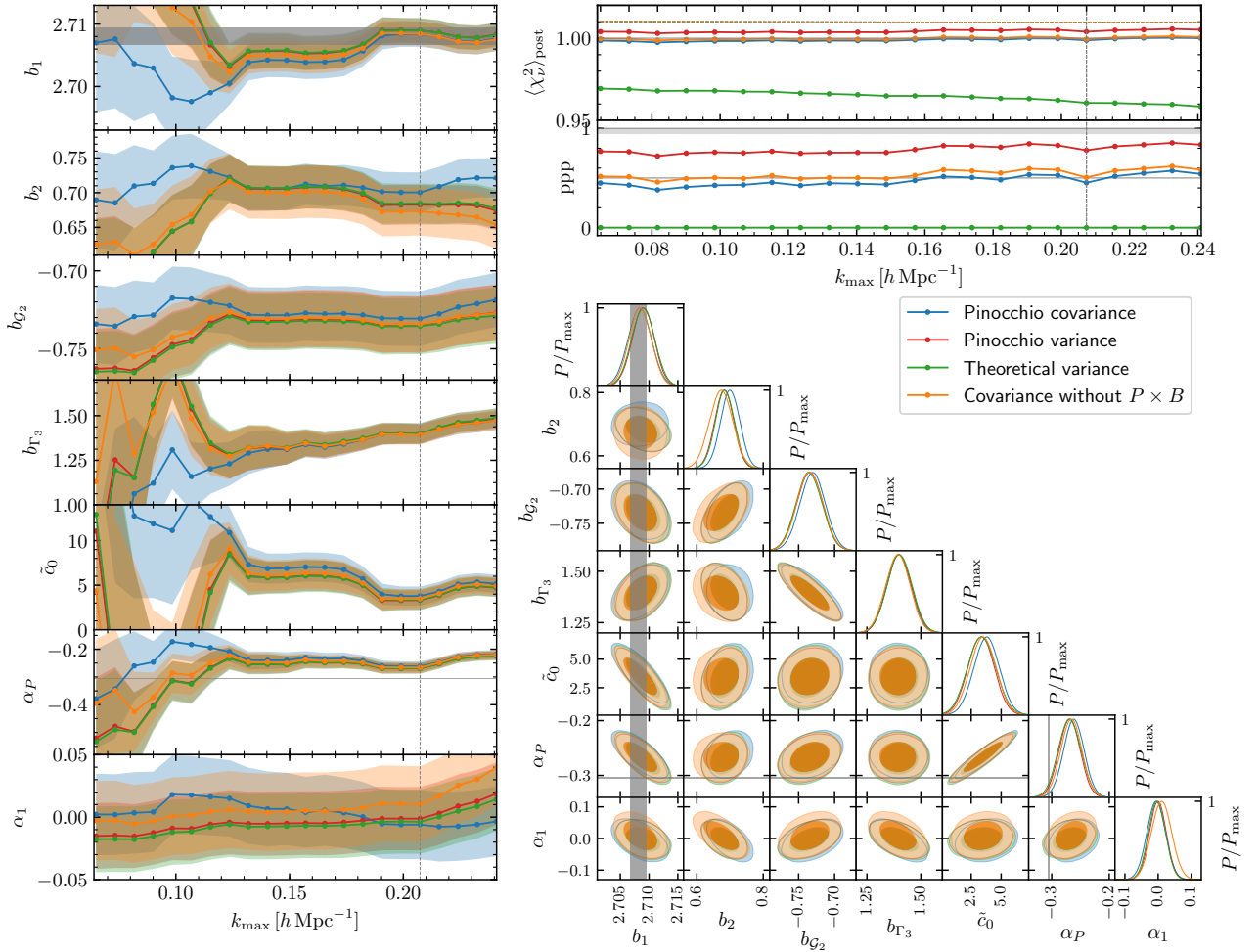
Figure 5.7: Comparison of parameter constraints and goodness-of-fit statistics between a fit done with the full mock covariance (blue), with the mock variance (red), with a theoretical variance (green), and with a covariance where cross-correlations between power spectrum and bispectrum have been set to zero (orange).

subset of triangles having sides that are larger than the given $k_{\max}$. For this reason, as an extension on the likelihood analysis of the bispectrum alone, we explore how different criteria for selecting triangles can change or improve constraints on the fitted parameters.

Consider figure 5.9: each panel shows the $(k_3/k_1, k_2/k_1)$ plane for a fixed value of k_1; triangle bins are plotted as squares, centered in the corresponding values of k_2/k_1 and k_3/k_1. The color of each triangle measures the absolute value of the posterior predictive checks (see section 4.2.2) of the Minerva measurements assuming a theoretical model averaged over the posterior, estimated from a joint fit with the setup of the reference case. It can be seen that, even though the maximum fitted Fourier mode for the bispectrum is $k_{\max,B} = 0.09\,h\,\mathrm{Mpc}^{-1}$, there are triangle bins with $k_1 > k_{\max,B}$ that are well described by the theoretical model. Looking at figure 5.9, it is clear that the model tends to fail first for those configurations that cluster in the right region of the $(k_3/k_1, k_2/k_1)$ plane, corresponding to the region where triangles close to equilateral are found.

It is possible, therefore, to imagine different criteria for the selection of triangles to include in the likelihood analysis in addition to the standard conditions $k_1 \leq k_{\max}$. We explore here 4 different ways to select triangles:

- given a fixed value of $k_{12,\mathrm{sum}}$, we select triangles such that $k_{12} \leq k_{12,\mathrm{sum}}$, and study constraints
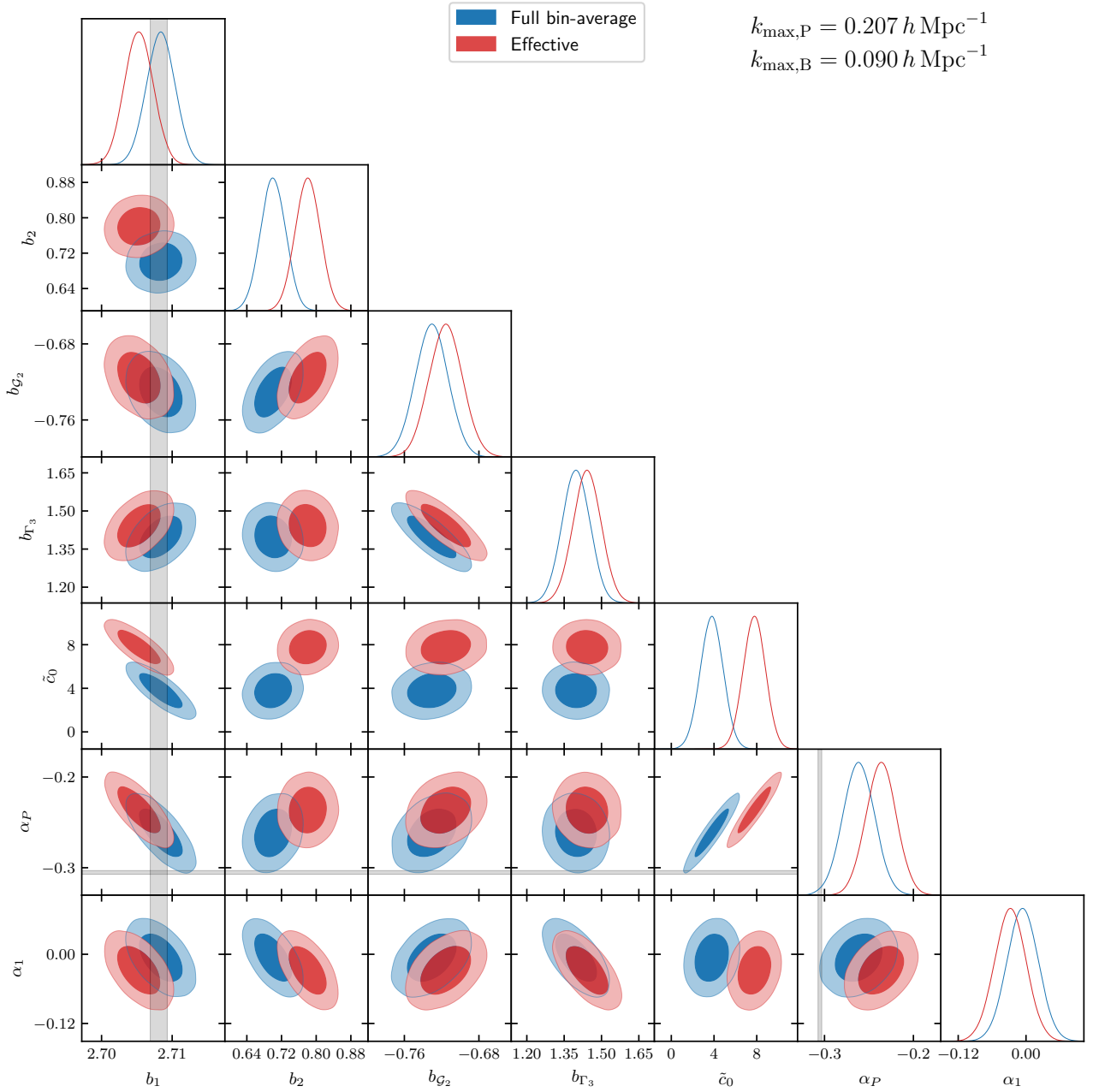
Figure 5.8: Comparison of the 1D and 2D marginalized posteriors from the joint fit of power spectrum and bispectrum where the models have been either fully averaged over the discrete Fourier bins (in blue) or evaluated over effective momenta and triangles (in red).

as a function of $k_{1,\max}$;

- given a fixed value of $k_{13,\mathrm{sum}}$, we select triangles such that $k_{13} \leq k_{13,\mathrm{sum}}$, and study constraints as a function of $k_{1,\max}$;

- we select triangles such that $k_{12} \leq k_{12,\mathrm{sum}}$, and study constraints as a function of $k_{12,\max}$;

- we select triangles such that $k_{13} \leq k_{13,\mathrm{sum}}$, and study constraints as a function of $k_{13,\max}$.

All these criteria tend to select fewer and fewer equilateral triangles at larger values of $k_{1,\max}$; for this reason we expect them, to some extent, to provide tighter constraints on the parameters, while keeping the posteriors unbiased with respect to the fiducial values of the parameters.
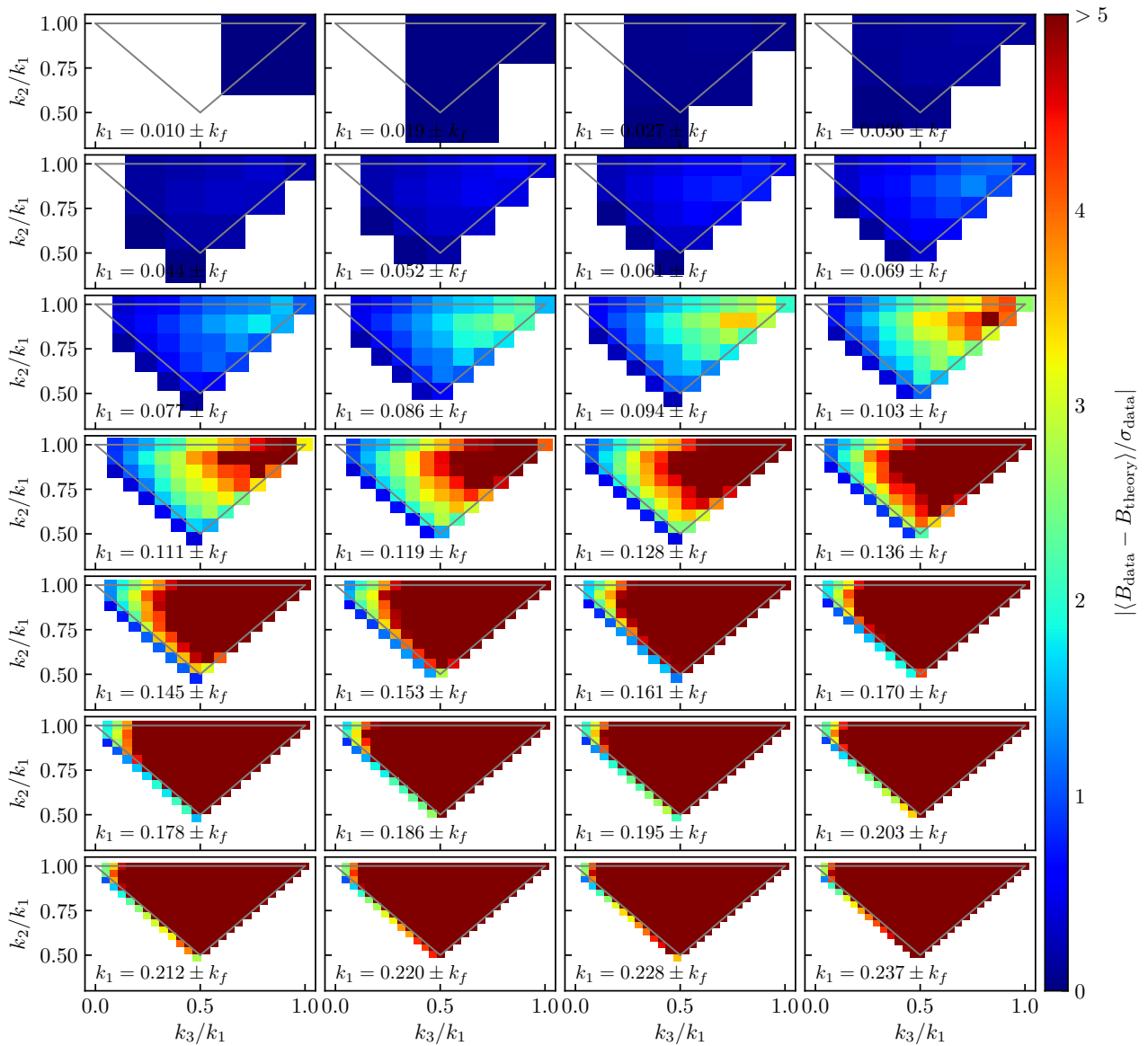
Figure 5.9: Each panel shows triangle bins in the (k_3, k_2) for a specific value of k_1; triangle bins are plotted as squares, centered in the corresponding values of k_2 and k_3. The color of each triangle measures the absolute value of the posterior predictive checks of the Minerva measurements assuming a theoretical model averaged over the posterior estimated from a joint fit of power spectrum and bispectrum with $k_{\mathrm{max,P}} = 0.21\,h\,\mathrm{Mpc}^{-1}$ and $k_{\mathrm{max,B}} = 0.09\,h\,\mathrm{Mpc}^{-1}$. Notice that the colorbar saturates at the 5σ level, in order to make more evident the deviations for triangles at the relevant values of k_1.

Figure 5.10 shows an example of one of these different selection criteria. In particular, the plot shows in blue the results of a fit where all triangles with sides smaller than a given k_{max} are fitted, and this will represent our reference case. In red we show the results of the fit where triangles are selected by imposing that all sides are smaller than a given k_{max} as well as the additional condition that $k_1 + k_3 < 0.14\,h\,\mathrm{Mpc}^{-1}$. This type of criterion selects all triangles up to a particular $\tilde{k}_1$, and then only a subset of triangles with sides ranging between $\tilde{k}_1$ and k_{max}, in such a way that triangles close to equilateral with $k_1 > \tilde{k}_1$ are excluded even when we consider larger values of k_{max}. In other terms, with this choice, the validity of the model of the tree-level bispectrum can be extended to the subset of triangles with $k_1 > \tilde{k}_1$ and $k_1 + k_3 < 0.14\,h\,\mathrm{Mpc}^{-1}$. The posteriors of the parameters fitted using this triangle selection criterion with $k_{\mathrm{max}} = 0.14\,h\,\mathrm{Mpc}^{-1}$ are well consistent with the reference case with $k_{\mathrm{max}} = 0.09\,h\,\mathrm{Mpc}^{-1}$; the widths of the posteriors of each parameter are either of the same order
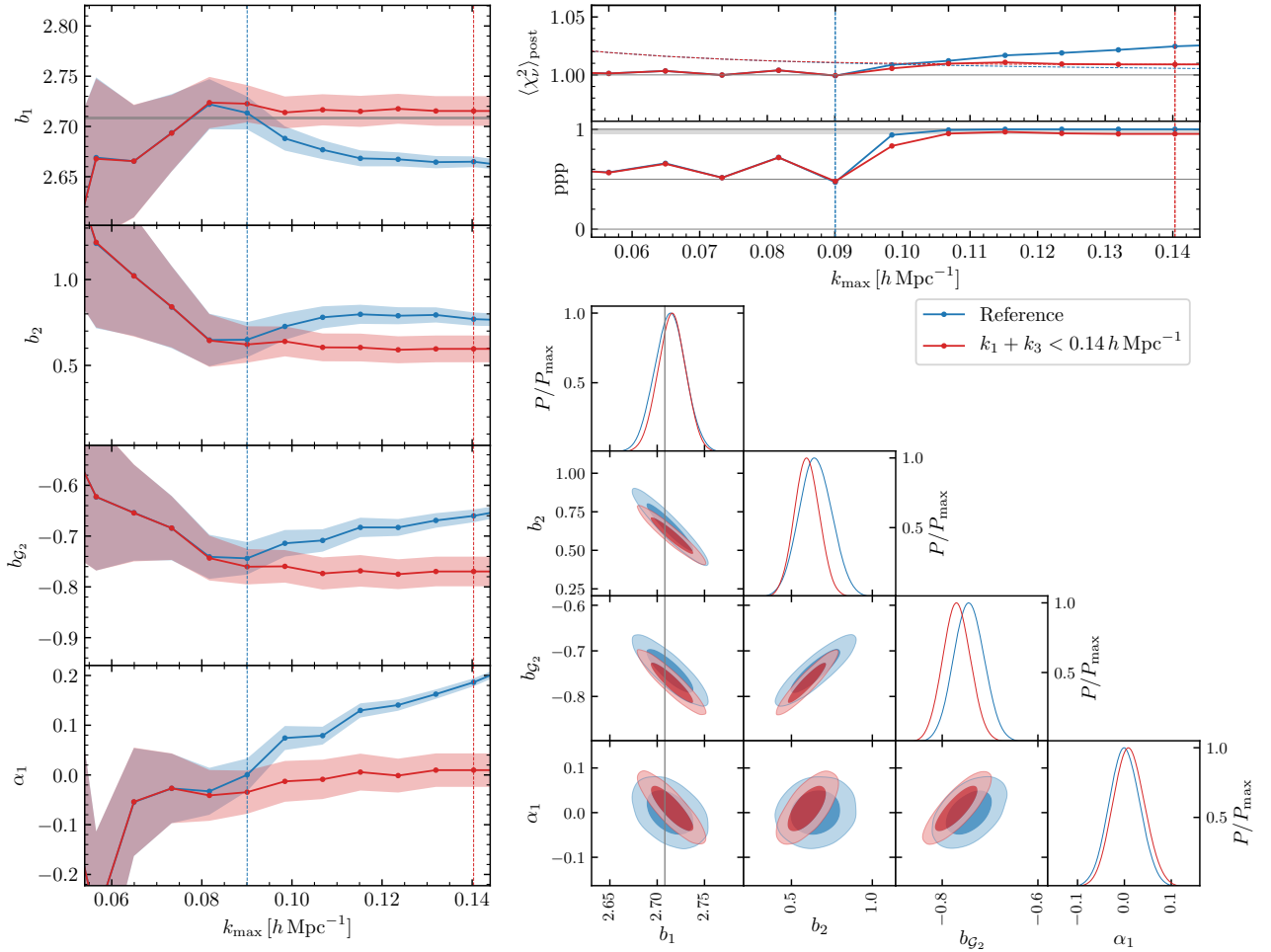
Figure 5.10: Comparison of parameter constraints and goodness-of-fit statistics between bispectrum fits either in the case where triangles are selected only in terms of $k_1 \leq k_{max}$ (in blue), or in the case where the additional criterion $k_1 + k_3 < 0.14\, h\,\mathrm{Mpc}^{-1}$ is imposed.

or slightly smaller, meaning that this selection criterion not only allows us to extend the tree-level bispectrum model to some smaller-scales triangles, albeit only a subset of them, but allows also to have parameter constraints that are similar, if not even better, than the reference case.

Figure 5.11 shows a similar comparison, but this time, apart from the condition that all sides have to be smaller than a given k_{max}, we impose the additional condition $k_1 + k_2 < k_{sum}$. Here we take two different values of k_{sum}. In red, we show results for $k_{sum} = 0.14\, h\,\mathrm{Mpc}^{-1}$. Notice that for $k_{max} > k_{sum}$, results are always the same as the ones at $k_{max} = k_{sum}$. While the posteriors for $k_{sum} = k_{max} = 0.14\, h\,\mathrm{Mpc}^{-1}$ are perfectly consistent with the reference case at $k_{max} = 0.09\, h\,\mathrm{Mpc}^{-1}$, the width of posteriors are much larger than in the reference case, meaning that, even though the tree-level bispectrum model can be extended to smaller scales, as shown by the goodness-of-fit statistics, the amount of information we can extract with this specific triangle selection criterion is not much compared to the reference case.

The situation changes slightly when the case with $k_{sum} = 0.18\, h\,\mathrm{Mpc}^{-1}$ is considered. With this condition, the posteriors of the fitted parameters improve with respect to the reference case, especially when looking at the contour plot with $k_{max} = k_{sum} = 0.18\, h\,\mathrm{Mpc}^{-1}$. The goodness-of-fit statistics also show that the validity of the model can really be extended to such smaller scales, always provided that only triangles with $k_1 + k_2 < 0.18\, h\,\mathrm{Mpc}^{-1}$ are selected. However, in this specific case, the fiducial

Figure 5.11: Similar to figure 5.10, but this time the traditional selection criterion $k_1 \leq k_{\max}$ (again in blue) is compared to the cases where either $k_1 + k_2 < 0.14\,h\,\mathrm{Mpc}^{-1}$ (in red) or $k_1 + k_2 < 0.182\,h\,\mathrm{Mpc}^{-1}$ (in green) is imposed.

value of b_1 measured from the cross-power spectrum is not consistent at 1σ with the posterior of this parameter, still the deviation is not very large, and the consistency is anyway at the level of 2σ.

Finally, figure 5.12 shows results for the analysis where all triangles are selected satisfying either $k_1 + k_2 < k_{12,\mathrm{sum}}$ or $k_1 + k_3 < k_{13,\mathrm{sum}}$, and the analysis is performed as a function of the value of either k_{sum}. However, in the reference case, we are actually showing results as a function of $k_{\max}$. The goodness-of-fit statistics show that selecting triangles using $k_1 + k_3 < k_{13,\mathrm{sum}}$ allows us to reach up to $k_{13,\mathrm{sum}} = 0.14\,h\,\mathrm{Mpc}^{-1}$, while using $k_1 + k_2 < k_{12,\mathrm{sum}}$ could allow us to reach even $k_{12,\mathrm{sum}} = 0.18\,h\,\mathrm{Mpc}^{-1}$; however, the comparison between the posteriors of b_1 and the value measured from the cross halo-matter power spectrum shows that when one uses $k_1 + k_2 < k_{12,\mathrm{sum}}$, posteriors with $k_{12,\mathrm{sum}} > 0.165\,h\,\mathrm{Mpc}^{-1}$ cannot be completely trusted.

Taking as a reference $k_{\mathrm{sum},12} = 0.165\,h\,\mathrm{Mpc}^{-1}$ and $k_{\mathrm{sum},13} = 0.14\,h\,\mathrm{Mpc}^{-1}$, the width of the posterior of b_1 is of the same order in the two cases, being 1.56×10^{-2} and 1.48×10^{-2} respectively, while in the reference case the width of the posterior of b_1 at $k_{\max} = 0.090\,h\,\mathrm{Mpc}^{-1}$ is 1.64×10^{-2}, meaning that the width of the posterior is reduced by $\sim 5\%$ and $\sim 10\%$ with the two criteria respectively. Consider that in the reference case with $k_{\max} = 0.090\,h\,\mathrm{Mpc}^{-1}$ the number of selected triangles per realization is 170, while $k_1 + k_2 < 0.165\,h\,\mathrm{Mpc}^{-1}$ selects 183 triangles, and $k_1 + k_3 < 0.14\,h\,\mathrm{Mpc}^{-1}$ selects 222. It appears that, judging in terms of the goodness of fit alone, using a selection criterion of $k_1 + k_2 < k_{\mathrm{sum}}$ allows to
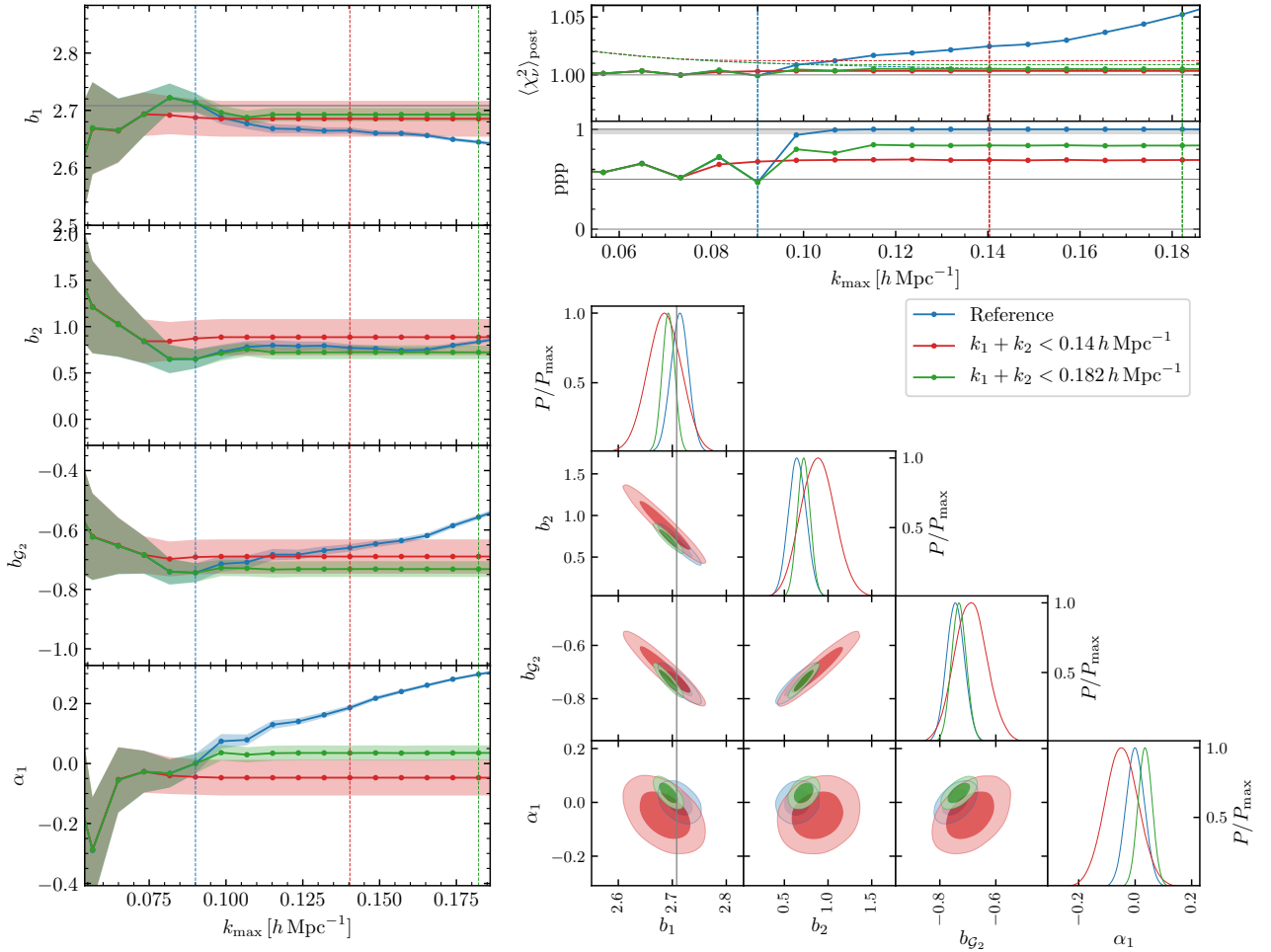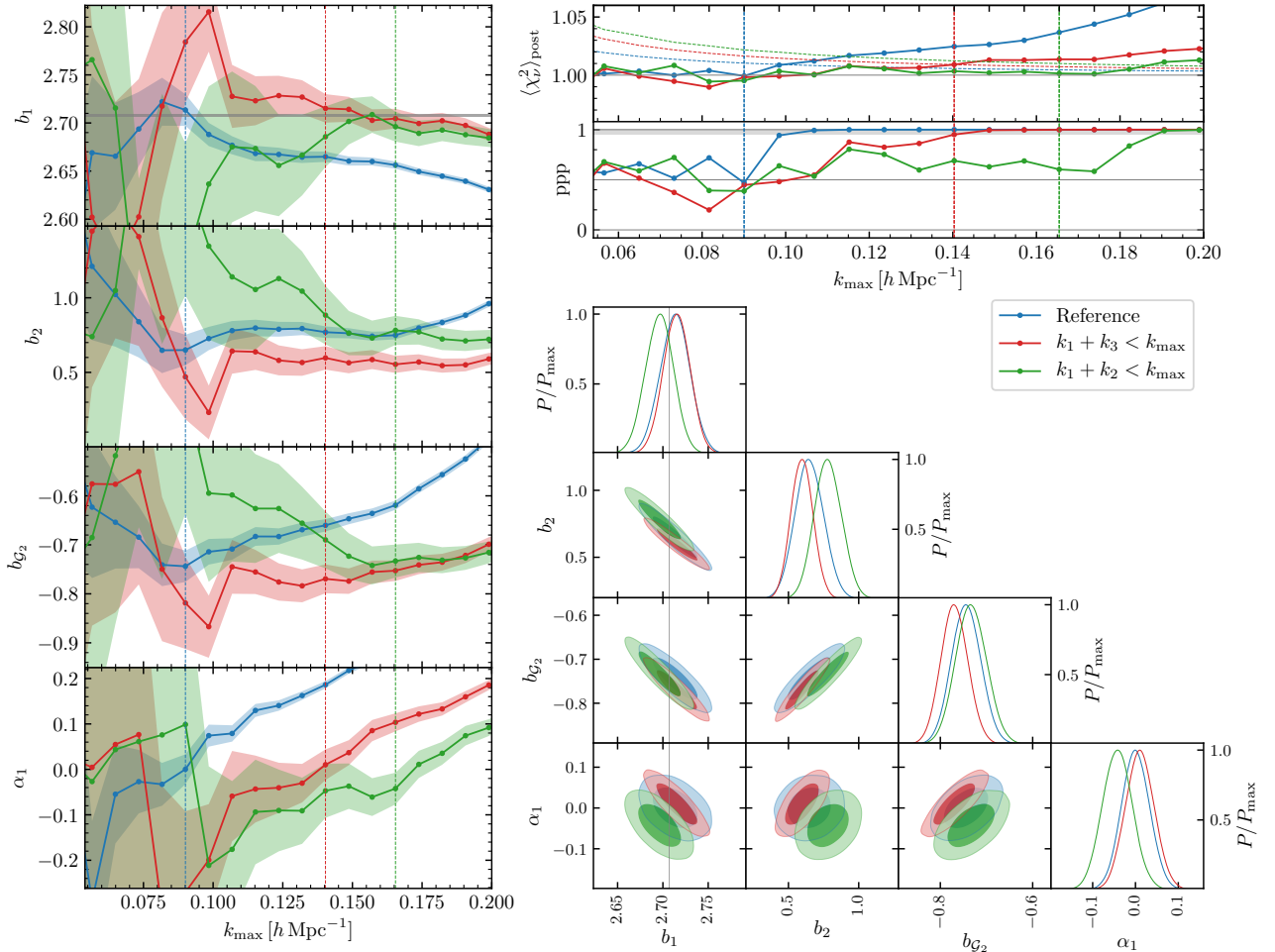
Figure 5.12: Similar to figure 5.10, but this time the traditional selection criterion $k_1 \leq k_{\max}$ (again in blue) is compared to the cases where either $k_1 + k_3 < k_{\max}$ (in red) or $k_1 + k_2 < k_{\max}$ (in green) is imposed.

reach smaller scales; on the other hand, the condition $k_1 + k_3 < k_{\text{sum}}$, while failing for a smaller k_{sum}, includes relatively more information, leading to better and less-biased constraints.

5.3.7   Cosmological parameters

In this section, we finally consider a set of parameters including cosmological parameters, in addition to the bias ones.

Figure 5.13 shows the constraints derived from the joint fit of power spectrum and bispectrum, using one-loop and tree-level models respectively, including three cosmological parameters, $\{A_s, h, \Omega_{\text{m}} h^2\}$; for convenience, we are normalizing A_s to the input value from the N-body simulations. The other cosmological parameters are set to their real value, used as an input in the N-body simulations. The fit was performed with $k_{\max,\text{P}} = 0.207\ h\,\text{Mpc}^{-1}$ and $k_{\max,\text{B}} = 0.090\ h\,\text{Mpc}^{-1}$, and using the full mocks covariance from the measurements on the mock catalogs. Grey lines (bands) show the real (measured) values of the corresponding parameters.

Table 5.2 shows the posterior-averaged values of the parameters, with errors being the standard deviation of the posteriors for each parameter, along with the fiducial values of the parameters (when available).

The 1D marginalized posteriors of b_1, α_P, and the cosmological parameters are all consistent at

Parameter	$\langle\theta\rangle \pm 1\sigma$	θ_{fid}	$|\langle\theta\rangle - \theta_{\mathrm{fid}}|/\sigma$
b_1	2.710 ± 0.037	2.7081 ± 0.0012	0.05
b_2	0.699 ± 0.107		
$b_{\mathcal{G}_2}$	-0.733 ± 0.021		
b_{Γ_3}	1.461 ± 0.080		
$\tilde{c}_0\,[\,h^{-2}\,\mathrm{Mpc}^2\,]$	2.252 ± 1.773		
α_P	-0.281 ± 0.035	-0.3052 ± 0.0020	0.69
α_1	-0.027 ± 0.035		
A_s/A_s^{fid}	1.0001 ± 0.0285	1.0000	0.0039
h	0.6953 ± 0.0008	0.695	0.38
$\Omega_m h^2$	0.1382 ± 0.0005	0.1377	1.03

Table 5.2: One-dimensional marginalized posteriors of the parameters from the joint fit of power spectrum and bispectrum as described in the main text. The first column lists all fitted parameters; the second column lists the mean values for the parameters posteriors along with their standard deviations. Non-empty entries in the third column quote the fiducial values of the parameters, coming either from direct measurements or from the input of the N-body simulations. The last column shows the distance in sigma of the fiducial value to the posterior-averaged parameter; where the fiducial value has an error, the sigma is intended as root sum squared of the uncertainty of the direct measurement and the standard deviation of the posterior.

less than 1σ level with their fiducial values (with $\Omega_m h^2$ being just above 1σ). Similar consistency is reached even in the 2D marginalized contours for pairs of these parameters, with the only exception being the ellipse for $(\Omega_m h^2, \alpha_P)$, where the consistency is only within 2σ.

It is clear that the significance of these results can only be assessed by a proper comparison with a power spectrum only analysis. This will appear in the published paper, of course, and the choice of the parameters and priors that can provide a consistent comparison is currently under discussion by the collaboration. It is nevertheless interesting to remark here that a fit of the power spectrum analysis run with the relevant subset of parameters and the same prior would not lead to convergent chains. In other words, with the model and the flat priors considered here, only the joint analysis of power spectrum and bispectrum can provide sensible constrain on cosmological parameters.

5.4 Conclusions

While the project presented in this chapter is very close to completion, I should remark that these conclusion are nevertheless preliminary to a certain extent.

The main goal of this work has been the analysis of the real-space halo bispectrum in conjunction with the real-space halo power spectrum at large scales ($k_{\max} \lesssim 0.1\,h\,\mathrm{Mpc}^{-1}$ for the bispectrum, $k_{\max} \lesssim 0.25\,h\,\mathrm{Mpc}^{-1}$ for the power spectrum) with a likelihood-based method. We consider dark-matter halos with a minimum mass of $\sim 10^{13} h^{-1}\,\mathrm{M}_\odot$ at redshift $z = 1$ from the Minerva set of N-body simulations consisting of about 300 boxes with side $L = 1500\,h^{-1}\,\mathrm{Mpc}$, for a total volume of roughly $1000\,h^{-3}\,\mathrm{Gpc}^3$. While this volume is much larger than what has been observed by current surveys and than what will be observed by forthcoming experiments, such a large volume allows us to reduce significantly the statistical uncertainties, so that any systematic error induced by the theoretical modelling is well within the strictest requirements of future observations.

We estimate the covariance matrix through estimates of power spectrum and bispectrum measurements from 10,000 mock halo catalogs with the same cosmological setup as the N-body simulations generated with the Lagrangian PT-based PINOCCHIO code [65, 66, 67]. Despite the large number of
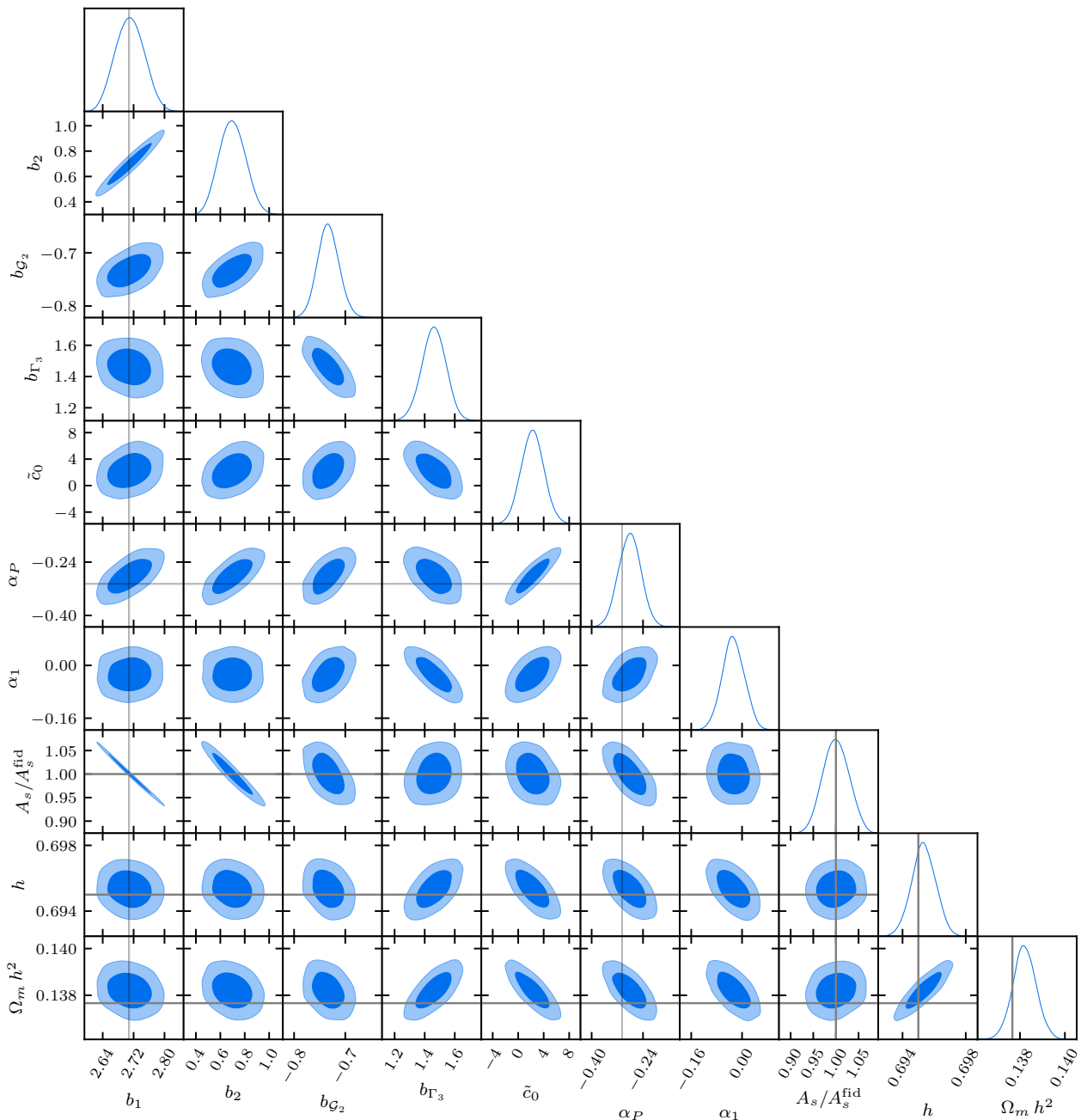
Figure 5.13: Contour plot showing the 1D and 2D marginalized posteriors for all parameters from the joint fit of power spectrum and bispectrum as described in the main text. Grey lines and bands show the fiducial values of the parameters, coming either from the input of the N-body simulations or from direct measurements.

mock simulations, we also account for statistical uncertainties in the precision matrix by employing the likelihood presented in [69]. As described in section 3.2, by a proper definition of the halo mass threshold, we ensure that the covariance recovered from the mocks provides as accurate description of the covariance of the N-body power spectrum and bispectrum. This allows us to extend to the power spectrum and bispectrum combination, the goodness-of-fit tools previously considered for the bispectrum alone.

For the bispectrum we use a tree-level perturbative model, while for the power spectrum we use a one-loop model. Our reference case is based on a model with a linear and a quadratic local bias parameters b_1 and b_2, with two non-local bias parameters $b_{\mathcal{G}_2}$ and b_{Γ_3}, with an effective sound speed

$\tilde{c}_0$ that accounts both for the EFT counterterm and higher-derivative bias, a constant shot-noise for the power spectrum that allows for deviation from the Poisson prediction, parametrized by α_P, and a single parameter regulating deviations from the Poisson prediction of the shot-noise of the bispectrum in the k-dependent part, α_1. This is slightly different with respect to what was considered in chapter 4, where α_1 was set to zero in the benchmark model: since the simultaneous likelihood analysis of power spectrum and bispectrum is aimed in reducing statistical uncertainties in the fitted parameters, we account for the possibility that the joint fit of the two correlation functions can constrain a non-Poissonian shot-noise in the bispectrum. The theoretical models take into account also for the damping of the respective BAO features arising due to bulk flows, and are thus non-perturbatively resummed in the infrared. We also perform additional adjustments to our theoretical predictions to account for discretization effects in the estimators of the correlation functions in Fourier space: both power spectrum and bispectrum models are fully averaged over one-dimensional bins and triangle bins respectively.

The goodness of fit is determined by two main criteria: good measures of the posterior predictive p-value (or equivalently, of the posterior-averaged reduced chi-square), and recovery of unbiased estimates of the fitted parameters. For this reason, we use measurements of the matter power spectrum and the halo-matter cross-power spectrum from the N-body simulations to directly measure the linear bias b_1 and the deviation from the Poisson prediction of the shot-noise α_P, and allow for the fitted parameter to be consistent with these measured values by no more than twice the width of the marginalized one-dimensional posteriors. We find, in addition to the already known fact that the tree-level model for the bispectrum is able to provide a good fit to our data up to $k_{\max} \simeq 0.09\,h\,\mathrm{Mpc}^{-1}$ (chapter 4), that the one-loop model for the power spectrum provides a good fit and is able to recover unbiased estimates of the fitted parameters up to $k_{\max} \simeq 0.21\,h\,\mathrm{Mpc}^{-1}$. Clearly, these results are specific to the halo population we considered, but while our intent is to develop a solid pipeline later to be applied to more realistic data sets, we notice that these numbers are not too far off from the range of validity of the perturbation theory model when compared to power spectrum measurements obtained from similarly large volume simulations [47].

We investigate variations of our reference model by considering, for instance, the addition of the higher-derivative bias in both the power spectrum and bispectrum models, of a second deviation to the Poisson prediction for the constant shot-noise contribution of the bispectrum, and of a k-dependent stochastic contribution for the power spectrum. We find that the second shot-noise parameter of the bispectrum α_2 is completely unconstrained and prior-dominated, and that both the higher-derivative bias and the stochastic contribution of the shot-noise of the power spectrum to be loosely constrained while still being consistent with zero. We use the DIC to compare the model extensions to our reference model and we find that the addition of these three parameters increases (and thus degrades) the value of the DIC, favouring our reference model over any considered extension.

We explore theoretically motivated and phenomenological relations between bias parameters [71, 59, 72] with the target of reducing the number of model parameters with respect to the maximal model. While expressing the quadratic local bias b_2 as a function of the linear bias b_1 and the tidal bias $b_{\mathcal{G}_2}$ leads to a good fit with relatively unbiased posteriors, the width of these posteriors is larger than what is found in the reference model; the other bias relations explored lead to both biased constraints and increase in the DIC with respect to the maximal model.

We study the impact of various approximations to the covariance matrix, as well as the specific

relevance of the cross-correlation between power spectrum and bispectrum, by comparing the posteriors with those obtained with the full covariance estimated from the mocks. We explore the cases of a diagonal covariance estimated from the mock catalogs, a fully theoretical diagonal covariance, and a covariance estimated from the mocks where cross-correlations between power spectrum modes and bispectrum triangle bins are set to zero. We find consistent posteriors in the mildly non-linear regime in the range $0.1\,h\,\mathrm{Mpc}^{-1} \div 0.2\,h\,\mathrm{Mpc}^{-1}$, with $\sim 1\sigma$ deviations at larger scales and with visible variations only in the goodness of fit.

We investigate the effect on the posteriors of the model parameters of evaluating theoretical models at effective Fourier momenta, in order to allow for faster evaluations of the theoretical predictions. We find large deviations in the recovered parameters also of order of 2σ or more.

We explore the possibility of extending the tree-level model for the bispectrum at smaller scales at least for a subset of triangle bins. For this reason, we consider different triangle selection criteria with respect to the standard $k_1 \le k_{\max}$. We find that selecting triangles based on the criterion $k_1 + k_2 \le k_{12,\mathrm{sum}}$ allows the extension of the tree-level model for less-equilateral triangles, up to $k_{12,\mathrm{sum}} \sim 0.165\,h\,\mathrm{Mpc}^{-1}$ – corresponding to triangles with values of $k_1 \lesssim 0.11\,h\,\mathrm{Mpc}^{-1}$ – without any improvement in the one-dimensional marginalized posteriors; on the other hand, selecting triangles based on the criterion $k_1 + k_3 \le k_{13,\mathrm{sum}}$ allows the extension of the tree-level model for a slightly different subset of less-equilateral triangles, up to $k_{13,\mathrm{sum}} \sim 0.14\,h\,\mathrm{Mpc}^{-1}$ – corresponding to triangles with values of $k_1 \lesssim 0.13\,h\,\mathrm{Mpc}^{-1}$ – with just a slight improvement in the posterior widths. In general, it appears that the additional triangles that can be allowed in the analysis beyond a given $k_{\max}$ do not provide sufficient signal to significantly improve the parameters constraints for the joint power spectrum and bispectrum case.

Finally, we perform a simultaneous fit of power spectrum and bispectrum using the reference model, with $k_{\max,\mathrm{B}} \simeq 0.09\,h\,\mathrm{Mpc}^{-1}$ and $k_{\max,\mathrm{P}} \simeq 0.21\,h\,\mathrm{Mpc}^{-1}$, varying as well three cosmological parameters, mainly the scalar amplitude A_s of the primordial power spectrum, the Hubble parameter h, and the relative abundance of total matter $\Omega_m h^2$. The full theoretical models, including linear power spectrum, loop-corrections, IR resummation, and full bin-averages, are recomputed consistently at each step of the MCMC. We recover unbiased values for all the fitted parameters of which we know either an independent measurement (b_1 and α_P) or the correct value (the cosmological parameter). In particular, the relative errors on the cosmological parameters are of order 3% for the scalar amplitude A_s, order 0.1% for the Hubble parameter, and order 0.4% for the matter density. The impact of the bispectrum in breaking degeneracies between parameters when fitted together with the power spectrum played a crucial role in this context: while in [47] similar error bars are reported, we did not set to a fixed value the parameter b_{Γ_3}, that in the case of the power spectrum alone has a very large degeneracy with $b_{\mathcal{G}_2}$. In our case, this degeneracy was broken by the addition of the bispectrum, and had we fixed the value of b_{Γ_3} in our fit, we would have likely found biased posteriors for the fitted parameters.

Following in spirit the approach in chapter 4, this work represents a further step into establishing a robust inference method for the analysis of the galaxy power spectrum and bispectrum. We take advantage of a reliable estimation of the covariance matrix, we consider the computation of an accurate theoretical model, and we use data sets characterized by large volumes to reduce both systematic and statistical uncertainties. With this current work, we already show that using the bispectrum in conjunction with the power spectrum can provide reliable constraints on the cosmological parameters.

In future works, more key aspects related to the analysis of actual data will be considered, mainly the impact of the survey geometry and of the addition of redshift-space distortions, that requires an accurate modelling both in the one-loop power spectrum and in the tree-level bispectrum.

Chapter 6

Further contributions

Several of the tools and data products developed for the two main projects described in this thesis found applications in a few additional projects where I am not first author. I will briefly describe in this section three of these projects in terms of their goals and results, as well as illustrating my specific contribution.

All these projects constitute, in different ways, extensions to the analysis developed in my work and rely on the same data sets, the Minerva simulations, and, with the exception of the first one described below, on the set of PINOCCHIO mocks.

6.1 Models comparison for the matter bispectrum

The first project is lead by Davit Alkhanishvili [169] and should appear on the `arXiv` in September 2020 or shortly thereafter. Its main goal is to assess the range of validity of theoretical predictions in PT for the matter bispectrum beyond the tree-level approximation.

An interesting extension of the work presented in this thesis, in fact, is considering the bispectrum model at 1-loop. Here we start with a comparison of 1-loop predictions for the matter bispectrum against N-body simulations. A direct comparison of the theoretical prediction of the matter bispectrum to measurements of dark-matter particles evolved by N-body simulations is a relevant test to assess the validity of the model free from other non-linear effects, such as galaxy biasing and redshift-space distortions, that depending on several free parameters, could in theory compensate for an inaccurate modelling.

We make use of the same power spectrum and bispectrum measurements of the matter density field of the full set of Minerva simulations described in section 3.1. In addition, we include similar measurements from a different set of N-body simulations run by Matteo Biagetti [170].

Loop corrections are evaluated through direct numerical integration. This task is numerically expensive, especially for the bispectrum. In addition, as discussed in section 2.8.1, these predictions should be averaged over the triangle bins in order to avoid unwanted systematic errors. In the case of the one-loop matter bispectrum, computing multidimensional integrals for all the triangles that can be defined over the discrete Fourier grid can become an almost impossible task, without the appropriate optimization techniques I have provided. The exact binning, extended to the loop predictions, constituted then a benchmark against which we have been able to assess various approximations as, for instance, limiting the exact binning to the sole tree-level component, that seems to be a promising

approach.

The large cumulative volume provided by the full set of the Minerva simulations allowed us to study the validity, and specifically the reach in terms of $k_{\max}$, of the different perturbative schemes considered as a function of the volume available, and therefore with increasingly smaller statistical uncertainties. We obtained this by fitting an increasing number of simulations up to the maximum available.

In order to validate the models, an analysis of the reduced chi-squared is performed; in particular, the reduced chi-squared can be written as a contribution due to the goodness-of-fit of the model, and a contribution that assesses how well the sample covariance is described by the real covariance. A proper estimate of the covariance properties of the matter bispectrum is therefore quite relevant and, in this case, cannot be provided by the halo mock catalogs.

In the case of the power spectrum, the covariance matrix can be well approximated with the Gaussian contribution, equation 2.144; for the bispectrum, the Gaussian contribution in equation 2.145, does not provide a good approximation of the sample variance. An accurate modeling of all the terms in the theoretical bispectrum covariance involves the computation of additional contributions depending on higher-order correlation functions (up to six-points), and the full bin-average of these correlation functions. While such task was beyond the scope of this work, a rough approximation for the bispectrum variance can be written as

$$
\begin{aligned}
\Delta B^2(k_1, k_2, k_3) \simeq {} & \frac{6}{N_B(k_1, k_2, k_3)} P(k_1) P(k_2) P(k_3) \\
& + 2 \left[\frac{1}{N_P(k_1)} + \frac{1}{N_P(k_2)} + \frac{1}{N_P(k_3)} \right] B^2(k_1, k_2, k_3),
\end{aligned}
\tag{6.1}
$$

with the second term being an extra contribution that approximates both the $\mathcal{O}(B^2)$ and $\mathcal{O}(TP)$ terms in equation 2.145, and that I verified works better than the usual $\mathcal{O}(P^3)$ term in describing the sample variance.

The EFT counterterm amplitudes are fitted to the data at fixed scales, and then the datasets are compared with the predictions of the different theoretical models, as a function of the maximum Fourier mode $k_{\max}$.

As a result, we find that the IR-resummed EFT models are the ones in the set of models analysed that give the best-fit to the data, with $k_{\max} \sim 0.15 - 0.18\,h\,\mathrm{Mpc}^{-1}$, for a total volume corresponding to the full set of simulations. For such models the range of validity is determined also accounting for the smallest scale providing a sensible determination of the EFT counterterms.

Our analysis extends previous results [53, 54] to a larger simulated volume and provides a more rigorous comparison with respect to several methodological aspects, including binning effects and the relevance of possible systematic errors in the numerical results.

6.2 Comparison with the modal estimator

This second project, about to be published [171], is lead by Joyce Byun. The main goal is a comparison of a standard bispectrum analysis as the one presented in section 4 to the analysis based on the modal bispectrum estimator introduced by [172, 173]. This comparison should extend a previous work [41] where a larger set of compressed statistics were considered but that was limited to a comparison in

terms of Fisher matrix forecasts. In this case, for the first time, we compare the bias parameters constraints obtained from the Minerva halo catalogs obtained in both cases by means of full Monte-Carlo analysis. Moreover, we do so with a robust estimate of the full covariance matrix for both methods.

The modal estimator, in fact, is a method to compress the non-Gaussian information contained in the bispectrum. As we have already mentioned, the bispectrum is a function of a large number (e.g., order $\mathcal{O}(10^{12})$ in a Fourier sphere of radius $85.5\,k_f$ defined on a discrete grid) of fundamental triangles in Fourier space; while the standard bispectrum estimator can compress the information contained in this large number of fundamental triangles in thousands or tens of thousands of bins (depending on the bin width), this represents still a large number of datapoints. In this case, estimating a non-singular covariance matrix from a number of mock simulations could become numerically prohibitive.

The modal decomposition method, initially introduced in the context of CMB analysis, [174] and refs. therein, is based on the assumption that the bispectrum $B(k_1, k_2, k_3)$, or more generally its product with a weight function $w(k_1, k_2, k_3)$, can be approximated by a linear combination of a number N_{m} of modes $Q_n(k_1, k_2, k_3)$, with $N_{\mathrm{m}} < N_{\mathrm{t}}$, such that

$$w(k_1, k_2, k_3)B(k_1, k_2, k_3) \simeq \sum_{n=1}^{N_{\mathrm{m}}} \beta_n^Q Q_n(k_1, k_2, k_3), \tag{6.2}$$

where β_n^Q are the modal coefficients relative to the basis Q_n.

Once an appropriate modal basis is chosen and an inner product between the modes of the basis is defined, one can define the matrix γ of the inner products of the set of basis functions,

$$\gamma_{nm} = \langle\!\langle Q_n | Q_m \rangle\!\rangle, \tag{6.3}$$

and the modal coefficients can be estimated by solving the linear matrix equation

$$\langle\!\langle Q_m | w\mathcal{B} \rangle\!\rangle = \sum_n \beta_n^Q \gamma_{nm}, \tag{6.4}$$

where here we indicate with $\mathcal{B}$ the bispectrum measurement on the individual fundamental triangle $(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)$. This procedure can be applied both to theory and measurements, since in the first case $\mathcal{B}$ is an analytical function of the fundamental triangles, and in the measurements $\mathcal{B} \propto \delta_{\mathbf{q}_1} \delta_{\mathbf{q}_2} \delta_{\mathbf{q}_3}$.

Since the tree-level galaxy bispectrum model can be written as the linear combination of triangle-dependent terms (B_{TL}, Σ, K,...) where the expansion coefficients are just a combination of the bias and shot-noise parameters, something on the same line can be done for the β_n^Q by selecting some custom modal functions, Q_n^{tree}. When these custom functions are selected, it is possible to write an exact expansion of the galaxy bispectrum depending on these modes,

$$wB_h(k_1, k_2, k_3) = \sum_n \beta_n^{\mathrm{tree}} Q_n^{\mathrm{tree}}(k_1, k_2, k_3). \tag{6.5}$$

Different generic modal bases can be chosen, however in that case this equality is only approximate. It is then possible to fit the modal coefficients computed from the theoretical model with the ones measured from the catalogs, and obtain constraints on the bias and shot-noise parameters. Notice that this procedure does not involve any binning procedure since the modal expansion and its coefficients

are defined directly in terms of the Fourier-space density field $\delta_\mathbf{q}$; however, the binning is implicitly accounted for in the specific definition of the inner product.

We measure the modal coefficients of the full set of 298 Minerva simulations presented in section 3.1, and we fit them to the coefficients obtained by the simple tree-level model for the galaxy bispectrum. We take into account several factors that can affect the parameter constraints, such as the choice of the modal basis functions and their numbers, of the algorithm to compute the inner products (that is analogous to the discretization effects of the standard bispectrum estimator), the choice of $k_{\max}$, and the impact of the covariance matrix, both in terms of the inclusion of off-diagonal elements, and in terms of the number of mocks needed to have a stable estimation of the mock covariance.

My work has been crucial in this aspect, since it provided a benchmark that could serve as a comparison for the parameter constraints obtained by a fit of the modal coefficients. As an example, we show in figure 6.1 parameter constraints on the bias and shot-noise parameters, obtained from a fit of 6 custom modes, and with 59 modes. The fit assumes a $k_{\max} \simeq 0.057\,h\,\mathrm{Mpc}^{-1}$, which is a scale that has been proven to be well within the range of validity of the standard bispectrum estimator. In both cases, the posteriors are largely consistent with each other, suggesting that using 6 custom modes is enough to perform data analysis. Moreover, the constraints obtained using the modal coefficients are fully consistent with the ones obtained from the fit of the standard bispectrum, regardless of the binning scheme.

6.3   Galaxy power spectrum and bispectrum in redshift space

The last project I mention in this brief review corresponds to the extension of the joint power spectrum and bispectrum analysis to redshift space. This is the most natural extension of my work and, indeed, entails a further development of the software I have written for the evaluation of the PT predictions and the evaluation of the power spectrum and bispectrum likelihood functions. Up to this point this has been my own work although it will be completed with contributions from my collaborators Chiara Moretti and Federico Rizzo, and will be presented in a few publications, likely by the end of 2020.

In practice, we started the implementation of the theoretical predictions for the anisotropic power spectrum in redshift space and, specifically, for its Legendre multipoles. Similarly, as a first step, we will include the tree-level prediction for the redshift-space bispectrum monopole and quadrupole, the latter being an observable not yet properly explored in the literature.

The EFT one-loop galaxy power spectrum in redshift space requires a full computation of 2.112; loop-corrections arising from the mode-coupling integral and the propagator integral can then be separated according to their μ dependence and to the powers of bias coefficients and growth rate. This allows the mode-coupling integral to be separated into 23 different k-dependent contributions, equation A.2 and following equations, and the propagator integral to be separated into 5 k-dependent contributions, equation A.20 and following equations.

As it is the case in real space, a full consistent analysis in redshift space, including cosmological parameters, requires the computation of all loop-corrections at every step of the MCMC. Computing 28 loop integrals using standard numerical integration can become numerically expensive, and for this reason an efficient implementation is required. This can be obtained by following the approach of [164], where all mode-coupling type integrals can be written as a function of just seven "building-blocks" integrals, linearly combined using some numerical coefficients, and where the propagator type integrals
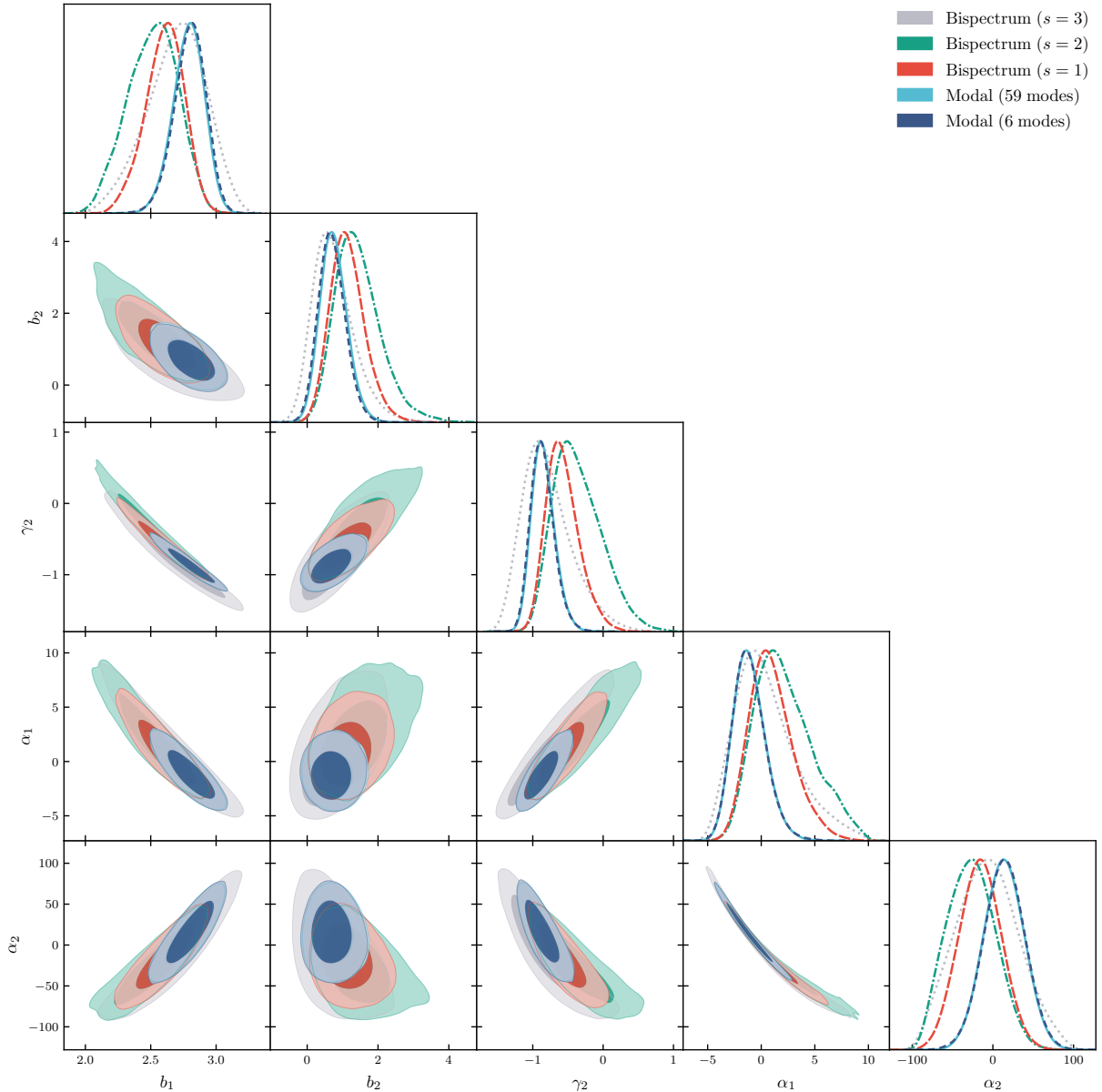
Figure 6.1: Comparison between the 1D and 2D marginalized posteriors obtained with the three different binning schemes of the standard bispectrum estimator ($s = 1, 2, 3$ respectively in grey, green, and red) with respect to the ones obtained with the modal estimator assuming either 59 or 6 custom modes (in cyan and blue respectively).

can be written with some fast Fourier convolution implementation. Notice that the mode-coupling integrals are all written as a function of the same 7 integrals used for the real-space loop corrections, and therefore the contributions exclusive to the redshift space power spectrum can be computed without virtually no extra computation time, apart from the appropriate linear combination of integrals and coefficients. With respect to real space, however, there are three extra propagator integrals, that being one-dimensional integrals can be anyway computed very quickly.

The full expressions for the anisotropic galaxy power spectrum, along with all 28 k-dependent contributions and their `FAST-PT` implementation is given in appendix A.

Since IR resummation has to be implemented as well, however, all these contributions have to be computed twice: first as a functional of the linear power spectrum $P_{\rm L}(k)$, and then a second time as a functional of the smooth power spectrum $P_{\rm nw}(k)$. Then one computes the wiggly part of the loop

Figure 6.2: Comparison between the power spectrum multipoles measurements averaged over the full set of N-body simulation ($\ell = 0, 2, 4$ in blue, red, and green respectively) and the theoretical models whose parameters are fixed to the best-fit parameters of the joint fit of power spectrum and bispectrum in real space, with $k_{\max,\mathrm{P}} = 0.207\ h\,\mathrm{Mpc}^{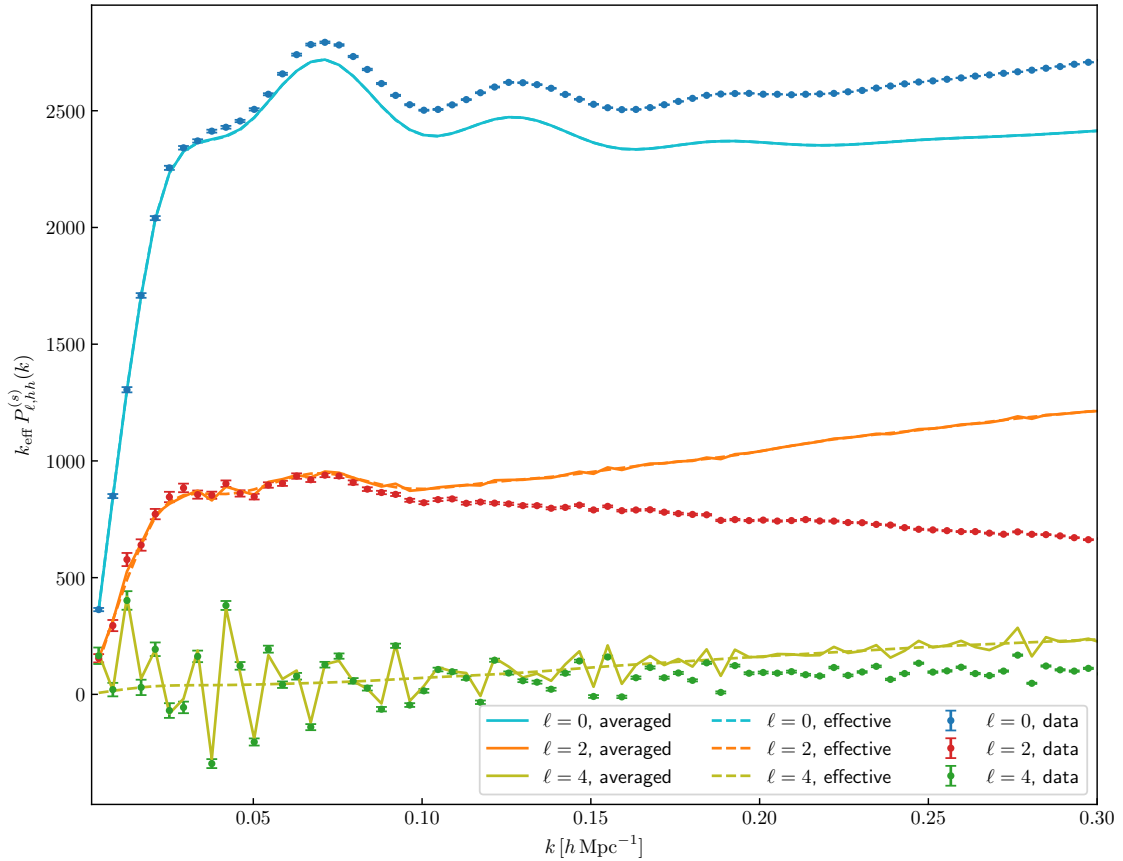-1}$ and $k_{\max,\mathrm{P}} = 0.09\ h\,\mathrm{Mpc}^{-1}$. The overlaid continuous lines represent the bin-averaged predictions of the multipoles, while the dashed lines show the models of the multipoles evaluated at the effective momenta; cyan, orange, and lime correspond respectively to $\ell = 0, 2, 4$.

corrections and the fully IR-resummed anisotropic power spectrum following the procedure described in section 2.7. After that, one can simply compute the power spectrum multipoles via equation 2.113.

Figure 6.2 shows the comparison between the average power spectrum multipoles measured from the set of 298 Minerva simulations compared to the theoretical prediction of the multipoles. The predictions shown here assume the bias parameters, one counterterm and the shot-noise parameter to be the same as the ones found by fitting together the power spectrum and bispectrum in real space, it assumes the growth rate f for the Minerva cosmology at redshift $z = 1$, and all other counterterms are assumed to be zero (while it is expected that they should be of the same order of $\tilde{c}_0$). Therefore, we assume that any systematic difference between data and model can be reconciled by an appropriate combination of counterterms coefficients and of small deviations of the bias parameters.

Two different ways to compute the prediction are considered here: one where the model is evaluated on the effective Fourier momenta of each bin, and one where the model is fully averaged over the discrete Fourier bin. Notice that, while for the monopole the difference between the two theoretical predictions are very small, differences become more important for the quadrupole (though they remain of the order of 1%), and for the hexadecapole, where a naive evaluation of the model at the effective momenta fails completely in describing the discretization effects introduced by the estimator.

The full modelling of the redshift-space galaxy bispectrum is still an ongoing effort.

Chapter 7

Conclusions

The main focus of this dissertation has been on the likelihood analysis of the galaxy bispectrum, both alone and in conjunction with the galaxy power spectrum. The inclusion in the standard cosmological analysis of higher-order correlation functions, in fact is particularly relevant for future surveys, such as *Euclid* [20, 84]. These surveys will map the distribution of tracers of the large-scale structure over comoving volumes of considerable size, and going beyond the traditional studies of two-point statistics will be a necessary effort to fully harness the information content of the LSS.

In the Baryon Oscillation Spectroscopic Survey (*BOSS*), the bispectrum has been already measured with enough signal, and analyses, albeit with not perfectly consistent models (that for example assumed local-Lagrangian predictions for tidal biases and common shot-noise corrections between power spectrum and bispectrum), have been performed [33, 34, 35]. Moreover, an analysis of the bispectrum from *BOSS* claimed a first detection of BAO features in this observable [36].

Methods to compress the information coming from the bispectrum – which is crucial in order to be able to estimate a covariance from a limited number of mock catalogues – have been proposed and compared, and improved constraints on the model parameters have been found when a large number of triangle bins (that would be otherwise impossible to analyse) is compressed into a small number of data points [106]. The modal estimator [172, 173, 174] has also been considered as an alternative to compress the information content of the galaxy bispectrum [41].

Fisher information analyses have been recently performed to assess the potential information content in the bispectrum that can be extracted from galaxy clustering when the bispectrum is analysed jointly with the power spectrum. It has been suggested that the redshift-space bispectrum contains additional information with respect to the power spectrum but that, however, using only the bispectrum monopole might lead to a loss of cosmological information with respect to the information content of the full redshift-space bispectrum [30]. Moreover, the bispectrum can significantly help in breaking degeneracies among cosmological and bias parameters (such as b_1, f and σ_8), and considering galaxy biasing and selection effects, a joint analysis of power spectrum and bispectrum can lead to a measurement of the linear growth rate of density perturbation f with a 10% precision [32]. Another forecast claims that neutrino masses can also be constrained with a $28\,\mathrm{meV}$ error, with an analysis of the power spectrum and bispectrum in redshift-space in a Euclid-like survey [25].

The information content of the bispectrum can also be used to constrain extensions of the concordance ΛCDM model, such as primordial non-Gaussianities [27, 28, 29], dynamical dark-energy [30], and modified theories of gravity [31].

With these motivations, my collaborators and I performed a detailed study of the statistical tools required for a rigorous analysis of the galaxy bispectrum. We considered several assumptions that are often made in the construction of the bispectrum likelihood function, and quantified the associated potential systematic errors on the determination of bias and cosmological parameters.

To this end, we considered a very large number of N-body simulations, constituting the Minerva set first presented in [64]; this is a set of almost 300 N-body simulations, evolving dark matter particles in boxes of size of $1500 \, h^{-1}$ Mpc, for a total volume of roughly $1000 \, h^{-3}$ Gpc^3. While such volumes will never be observed by present or future galaxy surveys, this allows for a relevant reduction of statistical uncertainties. Measurement of the power spectrum and the bispectrum were performed on dark-matter halos of minimum mass of the order of $10^{13} h^{-1} \mathrm{M}_\odot$ at redshift $z = 1$; moreover, measurements of the total matter power spectrum and of the halo-matter cross-power spectrum have been performed, with the aim of having independent measurements of some of the fitted parameters in our likelihood analysis as, for instance, the linear bias.

An even larger set of 10,000 mock halo catalogs, generated with the 3LPT-based code PINOCCHIO [65, 66, 67] was used to estimate covariance matrices. Since the number of triangle configurations can be of the order of $\sim 1,000$ or more, depending on the binning scheme employed in measuring correlation functions, this allowed for a robust estimate of a non-singular covariance matrix including also off-diagonal cross-correlation between the power spectrum and bispectrum. Still, a covariance estimated through such a large number of catalogs is a random object sampled from a distribution: we employed then the likelihood function from [69], to marginalize over this statistical uncertainty.

The theoretical models for power spectrum and bispectrum that have been compared to the measurements are all PT-based predictions: we employed a one-loop model for the power spectrum including an EFT counterterm, whose coefficient we fitted together with the other parameters; and a tree-level model for the bispectrum. We also performed IR resummation of both of our theoretical predictions, in order to account for the smoothing of the oscillatory features in the non-linear correlation functions. The power spectrum and bispectrum model we considered represent the state of the art of the models used to perform likelihood analyses [35, 48, 175]. In both models for power spectrum and bispectrum, we allowed for corrections to the shot-noise from the Poisson predictions. This is motivated by the fact that high-mass tracers (such as the dark-matter halos we are considering) are expected to have a sub-Poisson shot-noise in the large-scale limit [166, 167, 168, 158], and by direct measurements of the large-scale stochasticity coming from the combination of the auto- and cross-power spectra of the Minerva simulations. Finally, we accounted for the discretization effects introduced in the measurements by the commonly-used Fourier estimators [105, 176] by performing full bin-averages of the theoretical predictions using the same Fourier grids used in the measurements; we also studied the impact on the fitted parameters if this choice is not made, and the theoretical predictions were then evaluated over some effective momenta (for the power spectrum) and effective triangles (for the bispectrum).

We studied the posteriors of the parameters fitted by our likelihood pipeline as a function of the length of the maximum Fourier mode considered k_{max}, and used Bayesian techniques in order to measure the goodness of fit, namely the posterior-averaged reduced chi-square $\langle \chi^2_\nu \rangle$, and the posterior predictive p-value ppp.

In a first stage, we assumed that the cosmological parameters were known and we fixed them to their correct values, namely the ones used in the N-body simulations, and we fitted only for

bias, shot-noise, and counterterm parameters. In the bispectrum-only fits, the reference case was represented by a three-model parameter, including only the linear bias b_1, the quadratic bias b_2, and tidal bias $b_{\mathcal{G}_2}$ (or γ_2, depending on the specific bias expansion; however, in the bases used here, $b_{\mathcal{G}_2} = \gamma_2$), while the shot-noise parameters regulating deviations from the Poisson prediction were set to zero [1]. We found that the tree-level model for the bispectrum has been able to reproduce the measurements up to $k_{\max} \simeq 0.08 \div 0.1\, h\,\mathrm{Mpc}^{-1}$, depending on the binning scheme considered. We fitted the power spectrum alone using a model including, in addition to the three bias parameters of the bispectrum, also the bias b_{Γ_3} that appears at third-order, an effective counterterm coefficient $\tilde{c}_0$, and the parameter α_P measuring the shot-noise deviation from the Poisson prediction. The goodness of fit tests we employed have signaled a good fit of the model to the data at all scales considered, up to $k_{\max} \simeq 0.24\, h\,\mathrm{Mpc}^{-1}$, but the posteriors of the fitted parameters were found to be unbiased (at the 2σ level) only up to $k_{\max} \simeq 0.207\, h\,\mathrm{Mpc}^{-1}$. Moreover, the higher-order bias parameters were badly constrained. The joint fit of power spectrum and bispectrum was performed by fixing the maximum Fourier scale for the bispectrum at $k_{\max,\mathrm{B}} \simeq 0.09\, h\,\mathrm{Mpc}^{-1}$, and as a function of the minimum scale for the power spectrum, $k_{\max,\mathrm{P}}$. Again, we found good fits and unbiased values of the fitted parameters up to $k_{\max} \simeq 0.207\, h\,\mathrm{Mpc}^{-1}$; moreover, the addition of the information coming from the bispectrum at relatively large scales has improved significantly the width of the posteriors (e.g., the posterior width of the linear bias b_1 was reduced by a factor of 4 at the considered $k_{\max}$), has broken degeneracies between parameters, and has been able to greatly improve the constraints of the higher-order biases.

We also explored the possibility of extending the theoretical models by adding further parameters, and compared model performances using the DIC in all cases, and the Bayes Factor (estimated with the Savage-Dickey density ratio) for cases with properly nested model. A second shot-noise correction to the bispectrum has been found to be completely unconstrained and prior dominated; considering a higher-derivative bias both in the power spectrum and bispectrum models has given relatively good constraints on it, but still compatible with zero, and in turns re-introduced a large degeneracy with the EFT counterterm, showing that the addition of the bispectrum cannot fully break the degeneracy between the two parameters; a k^2 correction to the power-spectrum shot-noise, motivated by stochastic bias operators, has given good constraints and compatible with zero, but according to the DIC, its addition does not improve the fit.

We investigated as well physically-motivated or phenomenological relations between parameters [71, 177, 178, 92, 129, 70, 59, 72] with the objective of reducing the parameter space. In the case of the bispectrum-alone fits, the quadratic and tidal bias coefficients have been found to be definitely necessary to obtain a good fit of the model to the data. Imposing the local-Lagrangian prediction for $b_{\mathcal{G}_2}$ [177, 178, 129, 70, 92] however resulted in biased posteriors and was disfavored by the DIC; on the other hand, the prediction for b_2 from [71] was consistent according to the DIC, while recovering only slightly biased values of b_1. When fitting power spectrum and bispectrum together, however, no one of the relations we tested appeared to give unbiased estimates of all the model parameters. Still, the bias relation from [71] provided good estimates of the local and tidal bias parameters.

We studied the impact of different assumptions for the covariance matrix. In the case of a bispectrum-only fit, we considered – apart from the covariance estimated from mocks – also Gaussian diagonal predictions for the variance (using either the diagonal of our estimate of the diagonal, or

---

[1] When comparing results with the power spectrum, and when performing a simultaneous fit of power spectrum and bispectrum, we also allow for one correction to the shot-noise Poisson prediction to the bispectrum.

a Gaussian variance estimated from the measured power spectrum). Up to the $k_{\max}$ in which the tree-level model remains consistent to describe the data, the posteriors of the parameters did not vary, but the goodness of fit did, suggesting that correlations in the covariance matrix are not important in estimating fitted parameters. When fitting power spectrum and bispectrum simultaneously, however, things change, and estimate of the parameters fitted with a variance, or with a covariance without cross-correlations between power spectrum and bispectrum, were found to be different at the largest scales, while they converged to the same values when including more and more information from the smaller scales of the power spectrum. Our results partially justified the Gaussian approximation for the covariance matrices, that are here compared for the first time thanks to a robust numerical estimation based on a large number of mocks. Clearly, the extension to the bispectrum covariance in the case of proper survey measurements is not immediate since it is complicated by the effect of the window function (see [179] for a perturbative analysis).

We quantified, this as well for the first time, the systematic error on the recovered bias parameters induced by an approximate evaluation of the theoretical predictions that does not account for their full bin-average. In the case of the bispectrum alone, using either definition of effective triangles produced parameters posteriors that, while being consistent with the ones obtained by performing the full bin-average, showed slight deviations of order $\sim 0.5\sigma$. This became more dramatic in the case of the joint fit of power spectrum and bispectrum, where the deviations observed between the posteriors obtained with theoretical prediction evaluated at effective momenta and triangles have been found to be of order $1 - 2\sigma$, depending on the parameters.

We investigated different criteria to select triangle configurations in the bispectrum, and compared them with the standard criterion $k_1 \leq k_{\max}$. In particular, we studied combinations of the triangle sides that reduce the number of equilateral configurations (for which the model works worse in reproducing the data) while possibly extending the minimum scale assumed to smaller values, such as $k_1 + k_2 \leq k_{12,\max}$ and $k_1 + k_3 \leq k_{13,\max}$. We found that, while being effectively able to describe some of the triangles with $k_1 \geq 0.1\,h\,\mathrm{Mpc}^{-1}$, the constraints on the fitted parameters do not improve sensibly. A simultaneous fit of power spectrum and bispectrum, considering these triangles selection criteria, is currently under study.

Finally, we ran a joint fit of power spectrum and bispectrum in real space in order to recover unbiased estimates of the input cosmological parameters used to run the N-body simulations. We varied three cosmological parameters, $\{A_s, \Omega_m h^2, h\}$, while the others were assumed to be known and were fixed to their respective real values; at the same time, we varied seven other parameters, $\{b_1, b_2, b_{\mathcal{G}_2}, b_{\Gamma_3}, \tilde{c}_0, \alpha_P, \alpha_1\}$. The full theoretical predictions was recomputed at each step of the MCMC: the linear power spectrum was computed by the Boltzmann code `CAMB` [87], and was then IR-resummed using customized routines, loop corrections were computed with a `FAST-PT` implementation [164], and the final theoretical predictions were consistently bin-averaged accounting for the discretization effects introduced by the Fourier-space estimators. This required for fast implementations of each single step of the evaluation of the theoretical predictions, which was possible in my code; with my implementation, the bottle-neck in this computation is due to the call to the Boltzmann code that requires $\sim 1\,\mathrm{s}$, with one full likelihood evaluation requiring $\sim 1.5\,\mathrm{s}$. For this analysis, we considered the full covariance estimated from the mock halo catalogs, and therefore used the likelihood from [69]. The recovered posteriors for the cosmological parameters were all found to be well consistent with the real values, with only $\Omega_m h^2$ being at slightly above 1σ level, and at the same time both the

one-dimensional marginalized posteriors of the linear bias b_1 and the shot-noise correction α_P were found to be perfectly consistent with the measurements using the cross-power spectrum and the matter power spectrum. Given our large volume, we have been able to measure the cosmological parameters with great precision, with a 3% uncertainty in the case of the power spectrum amplitude A_s, a 0.1% uncertainty in the Hubble parameter h, and a 0.4% uncertainty in $\Omega_m h^2$.

The technology developed in this projects has been valuable also to other projects, currently ongoing: in [169], the implementation of the full bin-average of the bispectrum model has been used, and it has been applied to the one-loop matter bispectrum predictions based on different flavours of perturbation theory. Computing loop-corrections for the bispectrum, in fact, requires the evaluation of non-trivial multi-dimensional integrals, thus a fast computation of the full bin-average of this integral has been helpful, and will be helpful also in the future, when loop-corrections for the galaxy bispectrum will have to be implemented and re-evaluated every time the cosmological parameters are varied. My contribution also provided some insights on the fully-theoretical prediction of the bispectrum variance.

The results of this project have represented the benchmark to which the results of [171] have been compared to; this is of particular importance for the modal estimator, that allows to extract the information of the bispectrum and compress it in a small number of modes, and for which an estimate of the covariance matrix can require far less mock catalogs.

The current effort my collaborators and I are putting forward is aimed to the joint analysis of power spectrum and bispectrum in redshift space. This requires an adequate modelling of the redshift-space power spectrum multipoles (monopole, quadrupole, and hexadecapole) at one-loop level and of the redshift-space bispectrum multipoles (monopole and quadrupole) at tree-level. The power spectrum multipoles involve a total of 28 loop-correction terms, all of which have been implemented by us extending the `FAST-PT` code. This allows for a very fast computation of the power spectrum multipoles, including IR resummation, that in our implementation requires less than one-tenth of a second, crucial when fitting cosmological parameters. For power spectrum multipoles, moreover, accounting for discretization effects in the theoretical predictions becomes crucial, since a prediction using effective momenta would introduce in the model systematic deviations of order 1-10% for the quadrupole, and a much more dramatic effect would be observed in the hexadecapole. The theoretical prediction of the tree-level bispectrum multipoles is still under-way; also for the bispectrum multipoles we expect binning effects to be important. In appendix B, I discuss some further implementations that take into account binning effects in the power spectrum and bispectrum models in real space, and for the power spectrum model in redshift space, that can be computed in a time as fast as the effective predictions and that, approximating the full bin-average, can provide much better agreement with the full bin-averaged predictions. With all of this implemented, it will then be able to perform a full likelihood analysis of power spectrum and bispectrum multipoles in redshift space, and have a better idea on how to exploit the bispectrum to obtain better constraints on the cosmological parameters.

This will not be enough to analyse actual data, however: the survey geometry has an effect on the measured correlation functions, and this has to be taken into account in the theoretical predictions as well. Therefore, the models have to be convolved with the window function describing the survey geometry in an accurate and fast way. Moreover, the survey window function has an effect on the covariance matrix as well, therefore it is important to account for it even when estimating the covariance from mock catalogs, or when predicting it with a theoretical model. These are all aspects that have to be taken into account when analysing actual data and that still lack (either partially or

completely) of some theoretical framework and practical implementation.

In this context, my project represents a starting point toward a full likelihood analysis of the galaxy power spectrum and bispectrum, that will be of great relevance especially for future surveys, like *Euclid*. Analysing the bispectrum along with the power spectrum allows to break parameter degeneracies, and this can help in having much tighter constraints on parameters that are important in modern cosmology, such as the relative abundance of cold dark matter, or the growth rate as a function of redshift, that provides a dynamical test of dark energy. But it will be possible as well to study extensions of the standard ΛCDM model, considering a dynamical dark energy component, or warm dark matter, neutrino masses, primordial non-Gaussianities, and extensions beyond General Relativity. For this reason, this research can be seen as crucial in the direction toward a complete cosmological analysis of galaxy clustering measurements from spectroscopic redshift surveys.

Appendix A

The anisotropic galaxy power spectrum

In this section, I write down the full expression for the anisotropic galaxy power spectrum at one-loop level separating all 28 individual integrals entering in the prediction, as well as the full `FAST-PT` implementation of the 23 mode-coupling integrals.

A.1 Full expression of the anisotropic galaxy power spectrum at one-loop

The anisotropic galaxy power spectrum at one-loop in redshift space (neglecting IR resummation) is

$$P_g(\mathbf{k}) = Z_1(\mathbf{k})P_{\mathrm{L}}(k) + P_{22}^{(s)}(\mathbf{k}) + P_{13}^{(s)}(\mathbf{k}) + P_{\mathrm{ct}}^{(s)}(\mathbf{k}) + P_{\mathrm{noise}}^{(s)}(\mathbf{k}), \tag{A.1}$$

with $Z_1(\mathbf{k})P_{\mathrm{L}}(k)$ representing the leading-order contribution, $P_{22}^{(s)}(\mathbf{k}) + P_{13}^{(s)}(\mathbf{k})$ being the one-loop corrections, $P_{\mathrm{ct}}^{(s)}(\mathbf{k})$ containing all EFT counterterms, and $P_{\mathrm{noise}}^{(s)}(\mathbf{k})$ being the shot-noise, including k-dependent stochastic corrections.

The term $P_{22}^{(s)}(\mathbf{k})$ can be written as the linear combination of 23 different integrals, all of which have to be intended as already regularized following bias renormalization (see section 4.3.3):

$$
\begin{aligned}
P_{22}^{(s)}(\mathbf{k}) = 2\int \mathrm{d}^3\mathbf{q}[Z_2(\mathbf{q},\mathbf{p})]^2 P_{\mathrm{L}}(q)P_{\mathrm{L}}(p) = \\
= b_1^2 P_{b_1 b_1}(k) + b_1 b_2 P_{b_1 b_2}(k) + \left[b_1 b_{\mathcal{G}_2} + \frac{1}{4}f^2\mu^2 b_1\right] P_{b_1 b_{\mathcal{G}_2}}(k) + b_2^2 P_{b_2 b_2}(k) + \\
+ \left[b_2 b_{\mathcal{G}_2} + \frac{1}{4}f^2\mu^2 b_2\right] P_{b_2 b_{\mathcal{G}_2}}(k) + \left[b_{\mathcal{G}_2}^2 + \frac{1}{2}f^2\mu^2 b_{\mathcal{G}_2} + \frac{3}{32}f^4\mu^4\right] P_{b_{\mathcal{G}_2} b_{\mathcal{G}_2}}(k) + \\
+ \mu^2\left\{f^2 b_1^2 P_{\mu^2 f^2 b_1^2}(k) + f b_1 P_{\mu^2 f b_1}(k) + f b_2 P_{\mu^2 f b_2}(k) + \left[f b_{\mathcal{G}_2} + \frac{1}{8}f^3\mu^2\right] P_{\mu^2 f b_{\mathcal{G}_2}}(k) + \right. \\
\left. + f b_1^2 P_{\mu^2 f b_1^2}(k) + f b_1 b_2 P_{\mu^2 f b_1 b_2}(k) + f b_1 b_{\mathcal{G}_2} P_{\mu^2 f b_1 b_{\mathcal{G}_2}}(k)\right\} + \\
+ \mu^4\left\{f^2 P_{\mu^4 f^2}(k) + f^3 b_1 P_{\mu^4 f^3 b_1}(k) + f^2 b_2 P_{\mu^4 f^2 b_2}(k) + \right. \\
\left. + f^2 b_{\mathcal{G}_2} P_{\mu^4 f^2 b_{\mathcal{G}_2}}(k) + f^2 b_1 P_{\mu^4 f^2 b_1}(k) + f^2 b_1^2 P_{\mu^4 f^2 b_1^2}(k)\right\} + \\
+ \mu^6\left\{f^4 P_{\mu^6 f^4}(k) + f^3 b_1 P_{\mu^6 f^3 b_1}(k) + f^3 P_{\mu^6 f^3}(k)\right\} + \mu^8 f^4 P_{\mu^8 f^4}(k);
\end{aligned}
\tag{A.2}
$$

apart from $P_{b_1 b_1}(k) = P_{22}(k)$ in real space, and from the contributions also present in the real-space one-loop galaxy power spectrum, 17 more terms can be written. By setting $\mathbf{k} - \mathbf{q} = \mathbf{p}$ and $\eta = \hat{\mathbf{q}} \cdot \hat{\mathbf{p}}$, and writing $\int_{\mathbf{q}} \equiv \int \mathrm{d}^3 \mathbf{q}$, these integrals are defined by:

$$P_{\mu^2 f^2 b_1^2}(k) = \frac{1}{4} \int_{\mathbf{q}} S(\mathbf{q}, \mathbf{p}) \left[2 - \left(\frac{q^2}{p^2} + \frac{p^2}{q^2} \right) \right] P_{\mathrm{L}}(q) P_{\mathrm{L}}(p) \tag{A.3}$$

$$P_{\mu^2 f b_1}(k) = 4 \int_{\mathbf{q}} F_2(\mathbf{q}, \mathbf{p}) G_2(\mathbf{q}, \mathbf{p}) P_{\mathrm{L}}(q) P_{\mathrm{L}}(p) \tag{A.4}$$

$$P_{\mu^2 f b_2}(k) = 2 \int_{\mathbf{q}} G_2(\mathbf{q}, \mathbf{p}) P_{\mathrm{L}}(q) P_{\mathrm{L}}(p) \tag{A.5}$$

$$P_{\mu^2 f b_{\mathcal{G}_2}}(k) = 4 \int_{\mathbf{q}} S(\mathbf{q}, \mathbf{p}) G_2(\mathbf{q}, \mathbf{p}) P_{\mathrm{L}}(q) P_{\mathrm{L}}(p) \tag{A.6}$$

$$P_{\mu^2 f b_1^2}(k) = 2 \int_{\mathbf{q}} F_2(\mathbf{q}, \mathbf{p}) \left[2 + \eta \left(\frac{q}{p} + \frac{p}{q} \right) \right] P_{\mathrm{L}}(q) P_{\mathrm{L}}(p) \tag{A.7}$$

$$P_{\mu^2 f b_1 b_2}(k) = \int_{\mathbf{q}} \left[2 + \eta \left(\frac{q}{p} + \frac{p}{q} \right) \right] P_{\mathrm{L}}(q) P_{\mathrm{L}}(p) \tag{A.8}$$

$$P_{\mu^2 f b_1 b_{\mathcal{G}_2}}(k) = 2 \int_{\mathbf{q}} S(\mathbf{q}, \mathbf{p}) \left[2 + \eta \left(\frac{q}{p} + \frac{p}{q} \right) \right] P_{\mathrm{L}}(q) P_{\mathrm{L}}(p) \tag{A.9}$$

$$P_{\mu^4 f^2}(k) = 2 \int_{\mathbf{q}} G_2^2(\mathbf{q}, \mathbf{p}) P_{\mathrm{L}}(q) P_{\mathrm{L}}(p) \tag{A.10}$$

$$P_{\mu^4 f^3 b_1}(k) = \frac{1}{2} \int_{\mathbf{q}} S(\mathbf{q}, \mathbf{p}) \left[4 + \eta \left(\frac{q}{p} + \frac{p}{q} \right) - \left(\frac{q^2}{p^2} + \frac{p^2}{q^2} \right) \right] P_{\mathrm{L}}(q) P_{\mathrm{L}}(p) \tag{A.11}$$

$$P_{\mu^4 f^2 b_2}(k) = \frac{1}{2} \int_{\mathbf{q}} \left[3 + 2\eta \left(\frac{q}{p} + \frac{p}{q} \right) + \eta^2 \right] P_{\mathrm{L}}(q) P_{\mathrm{L}}(p) \tag{A.12}$$

$$P_{\mu^4 f^2 b_{\mathcal{G}_2}}(k) = \int_{\mathbf{q}} S(\mathbf{q}, \mathbf{p}) \left[3 + 2\eta \left(\frac{q}{p} + \frac{p}{q} \right) + \eta^2 \right] P_{\mathrm{L}}(q) P_{\mathrm{L}}(p) \tag{A.13}$$

$$P_{\mu^4 f^2 b_1}(k) = \int_{\mathbf{q}} \left[F_2(\mathbf{q}, \mathbf{p})(3 + \eta^2) + 2\eta [F_2(\mathbf{q}, \mathbf{p}) + G_2(\mathbf{q}, \mathbf{p})] \left(\frac{q}{p} + \frac{p}{q} \right) + 4 G_2(\mathbf{q}, \mathbf{p}) \right] P_{\mathrm{L}}(q) P_{\mathrm{L}}(p) \tag{A.14}$$

$$P_{\mu^4 f^2 b_1^2}(k) = \frac{1}{4} \int_{\mathbf{q}} \left[(3\eta^2 - 1) \left(\frac{q^2}{p^2} + \frac{p^2}{q^2} \right) + 8\eta \left(\frac{q}{p} + \frac{p}{q} \right) + 2(5 + \eta^2) \right] P_{\mathrm{L}}(q) P_{\mathrm{L}}(p) \tag{A.15}$$

$$P_{\mu^6 f^4}(k) = \frac{1}{8} \int_{\mathbf{q}} S(\mathbf{q}, \mathbf{p}) \left[\eta^2 + 11 + 4\eta \left(\frac{q}{p} + \frac{p}{q} \right) - 2 \left(\frac{q^2}{p^2} + \frac{p^2}{q^2} \right) \right] P_{\mathrm{L}}(q) P_{\mathrm{L}}(p) \tag{A.16}$$

$$P_{\mu^6 f^3 b_1}(k) = \frac{1}{2} \int_{\mathbf{q}} \left[8 + 4\eta^2 + (\eta^3 + 7\eta) \left(\frac{q}{p} + \frac{p}{q} \right) + (3\eta^2 - 1) \left(\frac{q^2}{p^2} + \frac{p^2}{q^2} \right) \right] P_{\mathrm{L}}(q) P_{\mathrm{L}}(p) \tag{A.17}$$

$$P_{\mu^6 f^3}(k) = \int_{\mathbf{q}} G_2(\mathbf{q}, \mathbf{p}) \left[\eta^2 + 2\eta \left(\frac{q}{p} + \frac{p}{q} \right) + 3 \right] P_{\mathrm{L}}(q) P_{\mathrm{L}}(p) \tag{A.18}$$

$$P_{\mu^8 f^4}(k) = \frac{1}{16} \int_{\mathbf{q}} \left[3(\eta^2 + 3)^2 + 8(\eta^3 + 3\eta) \left(\frac{q}{p} + \frac{p}{q} \right) + 4(3\eta^2 - 1) \left(\frac{q^2}{p^2} + \frac{p^2}{q^2} \right) \right] P_{\mathrm{L}}(q) P_{\mathrm{L}}(p). \tag{A.19}$$

The term $P_{13}^{(s)}(\mathbf{k})$ instead is the linear combination of 5 integrals:

$$
\begin{aligned}
P_{13}^{(s)}(\mathbf{k}) = 6Z_1(\mathbf{k})P_{\mathrm{L}}(k) \int \mathrm{d}^3\mathbf{q}\, Z_3(\mathbf{k},\mathbf{q},-\mathbf{q})P_{\mathrm{L}}(q) = \\
= Z_1(\mathbf{k})P_{\mathrm{L}}(k)\Bigg\{ b_1 P_{Z_1 b_1}(k) - \left[b_{\Gamma_3} + \frac{5}{2}b_{\mathcal{G}_2} + \frac{3}{8}f^2\mu^2 \right] P_{Z_1 b_{\Gamma_3}}(k) + \\
+ f\mu^2 P_{Z_1 f\mu^2}(k) + f\mu^2 b_1 P_{Z_1 f\mu^2 b_1}(k) + f^2\mu^4 P_{Z_1 f^2\mu^4}(k) \Bigg\}, \quad \text{(A.20)}
\end{aligned}
$$

where $P_{Z_1 b_1}(k)P_{\mathrm{L}}(k) = P_{13}(k)$, $P_{Z_1 b_{\Gamma_3}}(k)P_{\mathrm{L}}(k) = P_{b_1 b_{\Gamma_3}}(k)$, and

$$
P_{Z_1 f\mu^2}(k) = 6 \int_{\mathbf{q}} G_3(\mathbf{k},\mathbf{q},-\mathbf{q})P_{\mathrm{L}}(q) \tag{A.21}
$$

$$
P_{Z_1 f\mu^2 b_1}(k) = \frac{1}{42} \int_{\mathbf{q}} \left[2 \cdot \frac{24k^2 q^2 - 19k^4 - 9q^4}{k^2 q^2} - 9 \cdot \frac{(k^2 - q^2)^3}{k^3 q^3} \log\left|\frac{k+q}{k-q}\right| \right] P_{\mathrm{L}}(q) \tag{A.22}
$$

$$
P_{Z_1 f^2\mu^4}(k) = \frac{1}{168} \int_{\mathbf{q}} \left[2 \cdot \frac{9k^6 - 109k^4 q^2 + 63k^2 q^4 - 27q^6}{k^2 q^4} - 9 \cdot \frac{(k^2 - q^2)^3(k^2 + 3q^2)}{k^3 q^5} \log\left|\frac{k+q}{k-q}\right| \right] P_{\mathrm{L}}(q) \tag{A.23}
$$

The EFT counterterms can be written as

$$
P_{\mathrm{ct}}(\mathbf{k}) = -2\left[c_0 + f\mu^2 c_2 + f^2\mu^4 c_4 \right] k^2 P_{\mathrm{L}}(k) + c_{\nabla^4} f^4 \mu^4 k^4 Z_1(\mathbf{k})^2 P_{\mathrm{L}}(k), \tag{A.24}
$$

while the stochastic contribution to the redshift-space anisotropic power spectrum is given by

$$
P_{\mathrm{noise}}(\mathbf{k}) = \frac{1 + \alpha_P}{\bar{n}} + \left[\alpha_P^{(0)} + \mu^2 \alpha_P^{(2)} \right] k^2. \tag{A.25}
$$

With this, the full model depends on four bias parameters, four counterterm coefficients, three shot-noise corrections, and the growth rate, for a total of 12 free parameters.

A.2 `FAST-PT` implementation of the $P_{1-\mathrm{loop}}^{(s)}(k)$ integrals

By following the procedure presented in [164], the integrals entering in the definition of $P_{22}^{(s)}(\mathbf{k})$ can be computed in a more efficient way. Let us define the integrals

$$
J_{\alpha\beta\ell}(k) = \int \mathrm{d}^3\mathbf{q}\, q^\alpha p^\beta \mathcal{P}_\ell(\eta) P_{\mathrm{L}}(q) P_{\mathrm{L}}(p), \tag{A.26}
$$

and with these, it is possible to write all the contributions to the $P_{22}^{(s)}(\mathbf{k})$ as a function of these 7 integrals,

$$
\left\{ J_{0,0,0}(k),\ J_{0,0,2}(k),\ J_{0,0,4}(k),\ J_{2,-2,0}^{\mathrm{reg}}(k),\ J_{2,-2,2}(k),\ J_{1,-1,1}(k),\ J_{1,-1,3}(k) \right\} \tag{A.27}
$$

In particular, the $P_{22}^{(s)}(k)$ contribution is given by a linear combination of these $J_{\alpha\beta\ell}(k)$ integrals, that have to be combined with the appropriate coefficients shown in Table A.1; this gives each one of the 23 integrals defined above.

Combining all these terms with the appropriate combinations of bias parameters, growth rate, and powers of the cosine to the line of sight, gives the full expression of the anisotropic $P_{22}^{(s)}(\mathbf{k})$ in redshift

	$J_{0,0,0}(k)$	$J_{0,0,2}(k)$	$J_{0,0,4}(k)$	$J_{2,-2,0}^{\mathrm{reg}}(k)$	$J_{2,-2,2}(k)$	$J_{1,-1,1}(k)$	$J_{1,-1,3}(k)$
$P_{b_1 b_1}(k)$	1219/735	1342/1029	64/1715	1/3	2/3	124/35	16/35
$P_{b_1 b_2}(k)$	34/21	8/21	0	0	0	2	0
$P_{b_1 b_{\mathcal{G}_2}}(k)$	-72/35	88/49	64/245	0	0	-8/5	8/5
$P_{b_2 b_2}(k)$	1/2	0	0	0	0	0	0
$P_{b_2 b_{\mathcal{G}_2}}(k)$	-4/3	4/3	0	0	0	0	0
$P_{b_{\mathcal{G}_2} b_{\mathcal{G}_2}}(k)$	16/15	-32/21	16/35	0	0	0	0
$P_{\mu^2 f b_1}(k)$	2006/735	3212/1029	256/1715	2/3	4/3	232/35	48/35
$P_{\mu^2 f b_2}(k)$	26/21	16/21	0	0	0	2	0
$P_{\mu^2 f b_{\mathcal{G}_2}}(k)$	-152/105	136/147	128/245	0	0	-8/5	8/5
$P_{\mu^2 f^2 b_1^2}(k)$	-1/3	1/3	0	1/3	-1/3	0	0
$P_{\mu^2 f b_1^2}(k)$	82/21	44/21	0	2/3	4/3	264/35	16/35
$P_{\mu^2 f b_1 b_2}(k)$	2	0	0	0	0	2	0
$P_{\mu^2 f b_1 b_{\mathcal{G}_2}}(k)$	-8/3	8/3	0	0	0	-8/5	8/5
$P_{\mu^4 f^3 b_1}(k)$	-4/3	4/3	0	2/3	-2/3	-2/5	2/5
$P_{\mu^4 f^2 b_2}(k)$	5/3	1/3	0	0	0	2	0
$P_{\mu^4 f^3 b_{\mathcal{G}_2}}(k)$	-32/15	40/21	8/35	0	0	-8/5	8/5
$P_{\mu^4 f^2 b_1}(k)$	98/15	794/147	16/245	4/3	8/3	498/35	62/35
$P_{\mu^4 f^2 b_1^2}(k)$	8/3	1/3	0	0	1	4	0
$P_{\mu^4 f^2}(k)$	851/735	1742/1029	256/1715	1/3	2/3	108/35	32/35
$P_{\mu^6 f^4}(k)$	-14/15	19/21	1/35	1/3	-1/3	-2/5	2/5
$P_{\mu^6 f^3}(k)$	292/105	454/147	32/245	2/3	4/3	234/35	46/35
$P_{\mu^6 f^3 b_1}(k)$	14/3	4/3	0	0	2	38/5	2/5
$P_{\mu^8 f^4}(k)$	21/10	6/7	3/70	0	1	18/5	2/5

Table A.1: Coefficient entering in the `FAST-PT` implementation of the P_{22} terms in redshift space.

space, equation A.2.

The integrals contributing to $P_{13}^{(s)}(\mathbf{k})$, however, cannot be written as a function of the $J_{\alpha\beta\ell}(k)$ integrals, but reference [164] describes a technique to have a fast implementation of propagator-like integrals, that we use as well for the remaining one-loop contributions to the anisotropic power spectrum.

Appendix B

Approximating full bin-averages with a perturbative effective approach

As discussed in the main text, the evaluation of the theoretical model should account for the discretization effects introduced by the Fourier-space estimators of power spectrum and bispectrum. However, for higher-order statistics, such as the bispectrum, or for correlators in redshift-space, where anisotropies have to considered, this can become numerically expensive, and when the theoretical model has to be repeatedly re-computed several times, the time spent in computing full bin-averages of the theoretical predictions can be too prohibitive.

A commonly used alternative to the full bin-average approach is to compute the theoretical models in averaged, effective momenta (for the power spectrum) and triangles (for the bispectrum), that can be precomputed and that offer a large improvement in computation time. However, this approach can introduce systematic deviations in the theoretical models, and if the statistical uncertainties of the measurements are comparable to these deviations, parameters inferred with a prediction evaluated in this way can show deviations themselves.

I propose here an approach that, while allowing a computation of the model that is almost as fast as the one employed in the effective approach, can provide much more accurate theoretical predictions with respect to the full bin-average. This is based on a simple Taylor expansion of the power spectrum around some fixed k-modes, that for convenience can coincide with the standard effective momenta. Moreover, the approximation defined here provides a natural definition of the commonly employed effective momenta for the power spectrum. Finally, for the power spectrum in real-space and for the power spectrum multipoles in redshift space, this approach provides a model-independent approximation.

In the following sections, I discuss this perturbative effective approximation and the application to the real space power spectrum, the real space bispectrum, and the multipoles of the power spectrum in redshift space.

B.1 Power spectrum in real space

The power spectrum estimator in real space is

$$\hat{P}(k) \equiv \frac{k_f^3}{N_P(k)} \sum_{\mathbf{q} \in k} |\delta_{\mathbf{q}}|^2 , \qquad (\text{B.1})$$

where $k_f = 2\pi/L$ is the fundamental frequency of a cubic box of side L and

$$N_P(k) \equiv \sum_{\mathbf{q} \in k} 1. \qquad (\text{B.2})$$

A consistent way to evaluate the theoretical model of the power spectrum $P(k)$ – defined over continuous values of k – accounting for discretization effects of the measurements, is therefore

$$P_{\text{bin}}(k) = \frac{1}{N_P(k)} \sum_{\mathbf{q} \in k} P(q), \qquad (\text{B.3})$$

meaning that the power spectrum is averaged over each bin/shell in discrete Fourier space. For a Fourier grid with $k_{\max} \gg k_f$, a full computation can be time consuming. To avoid this, the common approach employed in the literature is to precompute the effective momenta of each spherical shell

$$\bar{k} = \frac{1}{N_P(k)} \sum_{\mathbf{q} \in k} q \qquad (\text{B.4})$$

and to evaluate the power spectrum at these effective momenta, $P_{\text{eff}}(k) = P(\bar{k})$.

With the advent of precision cosmology, one seeks precision of less than 0.1%; however, at the largest scales, $P_{\text{eff}}(k)$ deviates from the fully averaged $P_{\text{bin}}(k)$ even by order $\sim 1\%$.

Assume now that the theoretical model can be expanded in Taylor series around each $\bar{k}$:

$$P(q) = P(\bar{k}) + P'(\bar{k})(q - \bar{k}) + \frac{1}{2} P''(\bar{k})(q - \bar{k})^2 + \dots; \qquad (\text{B.5})$$

we can then write the fully averaged power spectrum in series:

$$P_{\text{bin}}(k) = \frac{1}{N_P(k)} \sum_{\mathbf{q} \in k} \left[P(\bar{k}) + P'(\bar{k})(q - \bar{k}) + \frac{1}{2} P''(\bar{k})(q - \bar{k})^2 + \dots \right]. \qquad (\text{B.6})$$

The zeroth order (leading-order) term reduces to the standard effective power spectrum,

$$P_{\text{bin}}^{\text{LO}}(k) = P_{\text{eff}}(k) = P(\bar{k}); \qquad (\text{B.7})$$

the contribution to be added to the leading-order to get the first, next-to-leading order approximation gives

$$\frac{1}{N_P(k)} \sum_{\mathbf{q} \in k} P'(\bar{k})(q - \bar{k}) = \frac{1}{N_P(k)} P'(\bar{k}) \left[\sum_{\mathbf{q} \in k} q - \bar{k} \sum_{\mathbf{q} \in k} 1 \right] = 0, \qquad (\text{B.8})$$

which is zero because of the definitions of $\bar{k}$ and $N_P(k)$. Notice that while the Taylor expansion of the power spectrum can be performed around any fixed value $\tilde{k} \in k$, the first-derivative contribution is zero only when the expansion is performed around the effective momenta, since in that case the expression

in the square brackets is identically zero. This provides a natural definition of the effective momenta, as those momenta that have a first-derivative contribution to the bin-average approximation exactly equal to zero. Thus, evaluating the theoretical model of the power spectrum at the effective momenta provides an approximation that is exact up to second order in k,

$$P_{\text{bin}}^{\text{NLO}}(k) = P_{\text{bin}}^{\text{LO}}(k). \tag{B.9}$$

Moving on with the perturbative approximation, the second-derivative term becomes:

$$\frac{1}{N_P(k)} \sum_{\mathbf{q} \in k} \frac{1}{2} P''(\bar{k})(q - \bar{k})^2 = \frac{P''(\bar{k})}{2N_P(k)} \sum_{\mathbf{q} \in k} (q^2 - 2q\bar{k} + \bar{k}^2) = \frac{P''(\bar{k})}{2N_P(k)} \left[\sum_{\mathbf{q} \in k} q^2 - N_P(k)\bar{k}^2 \right]. \tag{B.10}$$

If we define

$$\sigma_k^2 = \frac{1}{N_P(k)} \sum_{\mathbf{q} \in k} q^2 - \bar{k}^2, \tag{B.11}$$

we can approximate the fully averaged power spectrum at next-to-next-to-leading order as

$$P_{\text{bin}}^{\text{NNLO}}(k) \simeq P(\bar{k}) + \frac{1}{2} P''(\bar{k})\sigma_k^2. \tag{B.12}$$

This means that we can approximate with higher precision the full average of the power spectrum by adding to the power spectrum evaluated at the effective momenta a contribution depending on the second derivative of the power spectrum. In order to have a well-behaved, continuous n-th derivative, it is required that the starting power spectrum is interpolated with at least an $(n+1)$-th order spline, so that the n-th derivative corresponds to a linear, and therefore continuous, interpolation. This means that it is sufficient to interpolate the theoretical model of the power spectrum with a cubic spline to have a linearly interpolated second derivative.

The quantities $\bar{k}$ and σ_k^2 are precomputed only once, and therefore the computational cost of this approach is of the same order of the usual effective approach, while providing a great improvement in accuracy, as shown in figure B.1.

B.2   Bispectrum in real space

Something similar can be written also for the bispectrum; in real space, its estimator is given by

$$\hat{B}(k_1, k_2, k_3) \equiv \frac{k_f^3}{N_B(k_1, k_2, k_3)} \sum_{\mathbf{q}_1 \in k_1} \sum_{\mathbf{q}_2 \in k_2} \sum_{\mathbf{q}_3 \in k_3} \delta_K(\mathbf{q}_{123}) \delta_{\mathbf{q}_1} \delta_{\mathbf{q}_2} \delta_{\mathbf{q}_3}, \tag{B.13}$$

where the quantity $N_B(k_1, k_2, k_3)$ represents the number of fundamental triangles in the triangle bin of centers k_1, k_2, and k_3:

$$N_B(k_1, k_2, k_3) \equiv \sum_{\mathbf{q}_1 \in k_1} \sum_{\mathbf{q}_2 \in k_2} \sum_{\mathbf{q}_3 \in k_3} \delta_K(\mathbf{q}_{123}). \tag{B.14}$$

The full bin-average of the bispectrum model, therefore, can be written as

$$B_{\text{bin}}(k_1, k_2, k_3) = \frac{1}{N_B(k_1, k_2, k_3)} \sum_{\mathbf{q}_1 \in k_1} \sum_{\mathbf{q}_2 \in k_2} \sum_{\mathbf{q}_3 \in k_3} \delta_K(\mathbf{q}_{123}) B(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3). \tag{B.15}$$
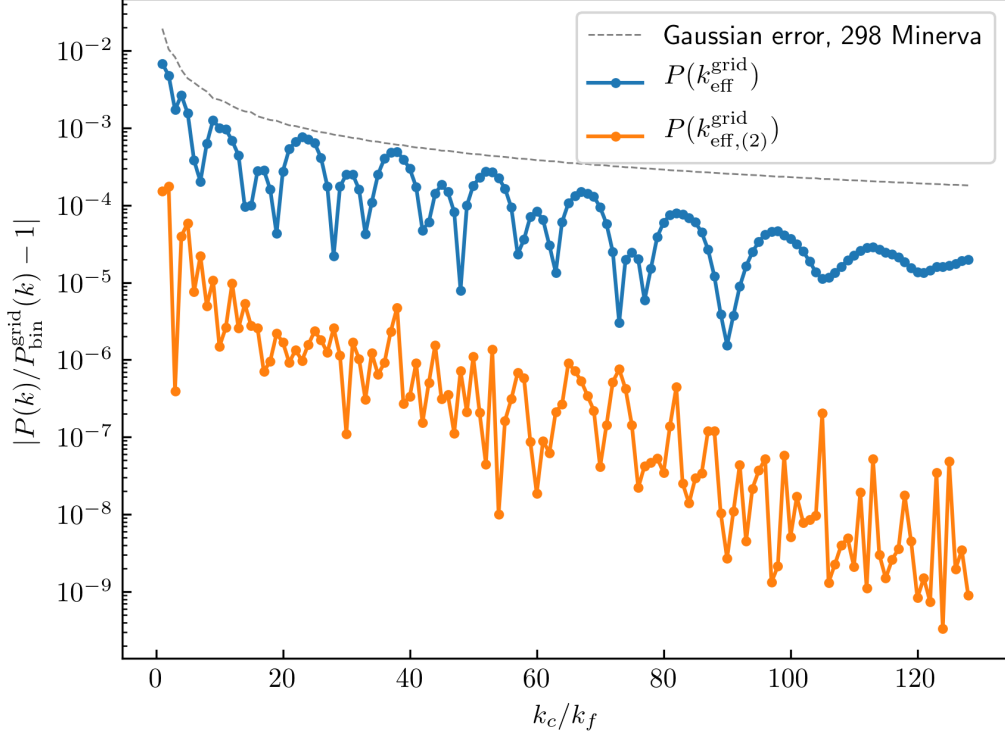
Figure B.1: Relative residuals from the bin-averaged theoretical model of a linear power spectrum of either the model evaluated to the effective momenta (in blue) or the model evaluated using equation B.12. The dashed line represent the prediction for the Gaussian error relative to 298 Minerva boxes.

Similarly as is the case for the power spectrum, it is possible to define some effective momenta also for the bispectrum; however, in this case the definition is not unique. Here, for convenience, we will define the "sorted effective momenta"

$$\bar{k}_l = \frac{1}{N_B(k_1, k_2, k_3)} \sum_{\mathbf{q}_1 \in k_1} \sum_{\mathbf{q}_2 \in k_2} \sum_{\mathbf{q}_3 \in k_3} \delta_K(\mathbf{q}_{123}) \max(q_1, q_2, q_3) \tag{B.16}$$

$$\bar{k}_m = \frac{1}{N_B(k_1, k_2, k_3)} \sum_{\mathbf{q}_1 \in k_1} \sum_{\mathbf{q}_2 \in k_2} \sum_{\mathbf{q}_3 \in k_3} \delta_K(\mathbf{q}_{123}) \text{med}(q_1, q_2, q_3) \tag{B.17}$$

$$\bar{k}_s = \frac{1}{N_B(k_1, k_2, k_3)} \sum_{\mathbf{q}_1 \in k_1} \sum_{\mathbf{q}_2 \in k_2} \sum_{\mathbf{q}_3 \in k_3} \delta_K(\mathbf{q}_{123}) \min(q_1, q_2, q_3), \tag{B.18}$$

and the effective bispectrum in this case can be written as

$$B_{\text{eff}}(k_1, k_2, k_3) = B(\bar{k}_l, \bar{k}_m, \bar{k}_s). \tag{B.19}$$

In the case of the bispectrum, due to the shape-dependence, it is more convenient to assume a specific model; let us consider the tree-level model for the galaxy bispectrum in real space, that has the structure of some kernel, depending only on the triangle shape, multiplied by two linear power spectra. If we call this kernel K, we have

$$B(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) = K(\mathbf{q}_1, \mathbf{q}_2) P_L(q_1) P_L(q_2) + \text{cyc.} \tag{B.20}$$

In the following, we will use the following notation

$$\frac{1}{N_B(k_1, k_2, k_3)} \sum_{\mathbf{q}_1 \in k_1} \sum_{\mathbf{q}_2 \in k_2} \sum_{\mathbf{q}_3 \in k_3} \delta_K(\mathbf{q}_{123}) f(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) = \langle f(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) \rangle_\triangle \qquad \text{(B.21)}$$

In order to proceed, we expand the product of the power spectra in Taylor series, keeping terms up to (and including) the second-derivative order,

$$P_L(q_1)P_L(q_2) = P_L(\bar{k}_l)P_L(\bar{k}_m) + P_L'(\bar{k}_l)P_L(\bar{k}_m)(q - \bar{k}_l) + P_L(\bar{k}_l)P_L'(\bar{k}_m)(q - \bar{k}_m) +$$
$$+ \frac{1}{2}P_L''(\bar{k}_l)P_L(\bar{k}_m)(q - \bar{k}_l)^2 + \frac{1}{2}P_L(\bar{k}_l)P_L''(\bar{k}_m)(q - \bar{k}_m)^2 +$$
$$+ P_L'(\bar{k}_l)P_L'(\bar{k}_m)(q - \bar{k}_l)(q - \bar{k}_m) + \mathcal{O}((\Delta q)^3). \qquad \text{(B.22)}$$

At the zero-th order, we have that the leading order approximation to the full bin-average of the bispectrum is simply

$$B_{\text{bin}}^{\text{LO}}(k_1, k_2, k_3) \simeq \langle K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle P_L(\bar{k}_l)P_L(\bar{k}_m) + \text{cyc.} \qquad \text{(B.23)}$$

At next-to-leading order, new terms appear, where the kernel has to be multiplied by each of the two Fourier momenta before performing the bin-average over the triangle bin:

$$B_{\text{bin}}^{\text{NLO}}(k_1, k_2, k_3) \simeq \langle K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle P_L(\bar{k}_l)P_L(\bar{k}_m) +$$
$$+ \left[\langle q_1 K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle - \bar{k}_l \langle K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle \right] P_L'(\bar{k}_l)P_L(\bar{k}_m) +$$
$$+ \left[\langle q_2 K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle - \bar{k}_m \langle K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle \right] P_L(\bar{k}_l)P_L'(\bar{k}_m) + \text{cyc.} \qquad \text{(B.24)}$$

Finally, adding terms from the second-order Taylor expansion introduces both second derivative terms and products of first derivative terms, and with these contributions the next-to-next-to-leading order becomes:

$$B_{\text{bin}}^{\text{NNLO}}(k_1, k_2, k_3) \simeq \langle K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle P_L(\bar{k}_l)P_L(\bar{k}_m) +$$
$$+ \left[\langle q_1 K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle - \bar{k}_l \langle K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle \right] P_L'(\bar{k}_l)P_L(\bar{k}_m) +$$
$$+ \left[\langle q_2 K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle - \bar{k}_m \langle K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle \right] P_L(\bar{k}_l)P_L'(\bar{k}_m) +$$
$$+ \frac{1}{2} \left[\langle q_1^2 K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle - 2\bar{k}_l \langle q_1 K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle + \bar{k}_l^2 \langle K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle \right] P_L''(\bar{k}_l)P_L(\bar{k}_m) +$$
$$+ \frac{1}{2} \left[\langle q_2^2 K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle - 2\bar{k}_m \langle q_2 K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle + \bar{k}_m^2 \langle K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle \right] P_L(\bar{k}_l)P_L''(\bar{k}_m) +$$
$$+ \left[\langle q_1 q_2 K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle - \bar{k}_l \langle q_2 K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle + \right.$$
$$\left. - \bar{k}_m \langle q_1 K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle + \bar{k}_l \bar{k}_m \langle K(\mathbf{q}_1, \mathbf{q}_2) \rangle_\triangle \right] P_L'(\bar{k}_l)P_L'(\bar{k}_m) + \text{cyc.} \qquad \text{(B.25)}$$

For each kernel, this approximation requires to precompute 18 averages involving kernels, on top of which 3 effective momenta have to be computed. For the galaxy bispectrum at tree-level in real space, one has three kernels, $K = F_2, 1, S$, and therefore a total of 57 numbers must be precomputed per triangle bin. Notice that the shot-noise part depending on the linear power spectra does not require
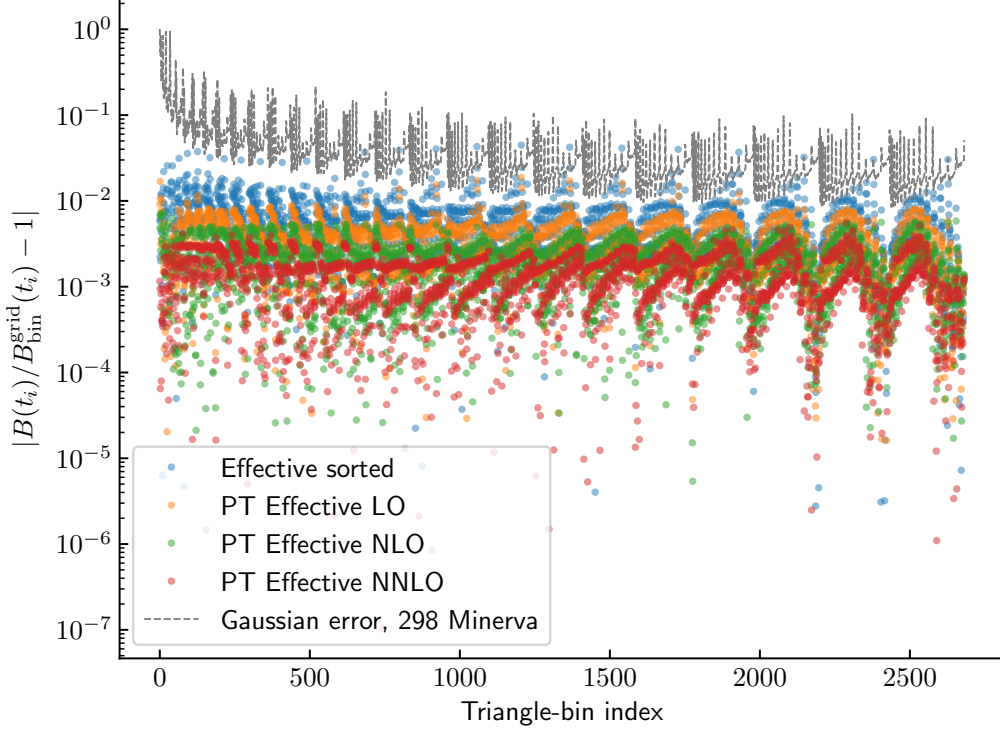
Figure B.2: Relative residuals from the bin-averaged model for the tree-level matter bispectrum of the model evaluated at the effective triangles with sorted sides (blue), with the Leading Order approximation of equation B.23 (red), with the NLO of equation B.24 (green), and with the NNLO of equation B.25. The Gaussian prediction of the error relative to 298 Minerva boxes is also shown (dashed line).

the computation of any extra term:

$$B_{\text{shot-noise}}^{\text{bin}}(k_1, k_2, k_3) \simeq \frac{1 + \alpha_1}{\bar{n}} b_1^2 \left[P_{\text{L}}(\bar{k}_{\text{l}}) + \frac{1}{2} P_{\text{L}}''(\bar{k}_{\text{l}}) \left(\langle q_1^2 \rangle_{\triangle} - \bar{k}_{\text{l}}^2 \right) \right] + \text{cyc.} + \frac{1 + \alpha_2}{\bar{n}^2}, \qquad (B.26)$$

since $\langle q_1^2 \rangle_{\triangle}$ is already computed when considering the second-order terms for $K = 1$.

A comparison is shown in figure B.2. While the gain in accuracy is not at the same level of the power spectrum case, this lets us compute a more accurate value of the theoretical prediction in a much shorter time. Each extra order increases the accuracy of the effective prediction of about half order of magnitude; the NNLO of the perturbative effective prediction reaches ~ 0.5% accuracy, to be compared to the almost 10% of the effective sorted and the few percent of the LO perturbative effective. The computation time for the tree-level matter bispectrum using the NNLO is of a few milliseconds using a naive `Python` implementation, to be compared to few $0.1\,\text{s}$ for the fastest (and most memory consuming) implementation in `C++` for triangles up to ~ $30k_f$.

B.3   Power spectrum in redshift space

Keeping in mind the estimator for the power spectrum in redshift space, we can write, for a particular multipole ℓ,

$$P_{\text{bin}}^{\ell}(k) = \frac{2\ell + 1}{N_P(k)} \sum_{\mathbf{q} \in k} P^{(s)}(\mathbf{q}) \mathcal{L}_{\ell}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}}), \qquad (B.27)$$

where $\hat{\mathbf{n}}$ is the line-of-sight direction; recalling the multipoles expansion of the redshift-space anisotropic power spectrum

$$P^{(s)}(\mathbf{q}) = \sum_{\ell'} P_{\ell'}(q)\mathcal{L}_{\ell'}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}}), \tag{B.28}$$

we obtain the expression for the bin-averaged multipole as

$$P_{\text{bin}}^{\ell}(k) = \frac{2\ell + 1}{N_P(k)} \sum_{\mathbf{q} \in k} \sum_{\ell'} P_{\ell'}(q)\mathcal{L}_{\ell'}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}})\mathcal{L}_{\ell}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}}). \tag{B.29}$$

Assume now a Taylor expansion for the multipole of the power spectrum around the average mode in the k-bin, $\bar{k}$:

$$P_{\ell}(q) = P_{\ell}(\bar{k}) + P_{\ell}'(\bar{k})(q - \bar{k}) + \frac{1}{2}P_{\ell}''(\bar{k})(q - \bar{k})^2 + \ldots; \tag{B.30}$$

at leading order, $P_{\ell}(q) = P_{\ell}(\bar{k})$, and thus

$$P_{\text{bin}}^{\ell}(k) \simeq \frac{2\ell + 1}{N_P(k)} \sum_{\mathbf{q} \in k} \sum_{\ell'} P_{\ell'}(\bar{k})\mathcal{L}_{\ell'}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}})\mathcal{L}_{\ell}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}}) =$$
$$= \frac{2\ell + 1}{N_P(k)} \sum_{\ell'} P_{\ell'}(\bar{k}) \sum_{\mathbf{q} \in k} \mathcal{L}_{\ell'}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}})\mathcal{L}_{\ell}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}}). \tag{B.31}$$

It is then possible to define the bin-average of the product of the two Legendre polynomials as an object with indices ℓ and ℓ', and with a k-bin dependence:

$$\Pi_{\ell\ell'}(k) \equiv \frac{1}{N_P(k)} \sum_{\mathbf{q} \in k} \mathcal{L}_{\ell'}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}})\mathcal{L}_{\ell}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}}), \tag{B.32}$$

and with this the zero-th order approximation of the bin-average of the multipole ℓ of the anisotropic power spectrum becomes

$$P_{\text{bin}}^{\ell}(k) \simeq (2\ell + 1) \sum_{\ell'} P_{\ell'}(\bar{k}) \, \Pi_{\ell\ell'}(k). \tag{B.33}$$

In the case of the anisotropic power spectrum of galaxies at one-loop, the bulk of the information comes from the first few multipoles, $\ell = 0, 2, 4$. Considering that $\Pi_{\ell\ell'}(k)$ is symmetric in the exchange of ℓ and ℓ', this means that $\Pi_{\ell\ell'}(k)$ represents a set of 6 unique k-dependent terms. By precomputing the $\bar{k}$ and as well as the $\Pi_{\ell\ell'}(k)$ for each k bin, it is possible to approximate in a fast way the generic ℓ-order multipole of the anisotropic power spectrum.

At next-to-leading order, the first derivative of the power spectrum multipole appears in the expression,

$$P_{\text{bin}}^{\ell}(k) \simeq (2\ell + 1) \sum_{\ell'} P_{\ell'}(\bar{k}) \, \Pi_{\ell\ell'}(k) + \frac{2\ell + 1}{N_P(k)} \sum_{\mathbf{q} \in k} \sum_{\ell'} P_{\ell'}'(\bar{k})(q - \bar{k})\mathcal{L}_{\ell'}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}})\mathcal{L}_{\ell}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}}) =$$
$$= (2\ell + 1) \sum_{\ell'} \left[P_{\ell'}(\bar{k}) - \bar{k}P_{\ell'}'(\bar{k}) \right] \Pi_{\ell\ell'}(k) + \frac{2\ell + 1}{N_P(k)} \sum_{\ell'} P_{\ell'}'(\bar{k}) \sum_{\mathbf{q} \in k} q\mathcal{L}_{\ell'}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}})\mathcal{L}_{\ell}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}}), \tag{B.34}$$

and by defining

$$\Xi_{\ell\ell'}(k) = \frac{1}{N_P(k)} \sum_{\mathbf{q} \in k} q \, \mathcal{L}_{\ell'}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}})\mathcal{L}_{\ell}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}}), \tag{B.35}$$

the next-to-leading order approximation of the bin-average of the generic multipole ℓ of the anisotropic

power spectrum becomes

$$P^\ell_{\text{bin}}(k) = (2\ell + 1) \sum_{\ell'} \left[P_{\ell'}(\bar{k})\Pi_{\ell\ell'}(k) + P'_{\ell'}(\bar{k}) \left(\Xi_{\ell\ell'}(k) - \bar{k}\Pi_{\ell\ell'}(k) \right) \right]. \tag{B.36}$$

The next-to-next-to-leading order term gives a contribution

$$
\begin{aligned}
\frac{1}{2} \frac{2\ell + 1}{N_P(k)} \sum_{\mathbf{q} \in k} \sum_{\ell'} P''_{\ell'}(\bar{k})(q - \bar{k})^2 \mathcal{L}_{\ell'}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}}) \mathcal{L}_\ell(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}}) &= \\
= \frac{1}{2} \frac{2\ell + 1}{N_P(k)} \sum_{\ell'} P''_{\ell'}(\bar{k}) \sum_{\mathbf{q} \in k} (q^2 - 2q\bar{k} + \bar{k}^2) \mathcal{L}_{\ell'}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}}) \mathcal{L}_\ell(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}}) &= \\
= \frac{2\ell + 1}{2} \sum_{\ell'} P''_{\ell'}(\bar{k}) \left\{ \frac{1}{N_P(k)} \sum_{\mathbf{q} \in k} q^2 \, \mathcal{L}_{\ell'}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}}) \mathcal{L}_\ell(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}}) - 2\bar{k}\Xi_{\ell\ell'}(k) + \bar{k}^2\Pi_{\ell\ell'}(k) \right\} &= \\
= \frac{2\ell + 1}{2} \sum_{\ell'} P''_{\ell'}(\bar{k}) \left[\Psi_{\ell\ell'}(k) - 2\bar{k}\Xi_{\ell\ell'}(k) + \bar{k}^2\Pi_{\ell\ell'}(k) \right], &
\end{aligned}
\tag{B.37}
$$

where we defined the quantity $\Psi_{\ell\ell'}(k)$ as

$$\Psi_{\ell\ell'}(k) \equiv \frac{1}{N_P(k)} \sum_{\mathbf{q} \in k} q^2 \, \mathcal{L}_{\ell'}(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}}) \mathcal{L}_\ell(\hat{\mathbf{q}} \cdot \hat{\mathbf{n}}) \tag{B.38}$$

and therefore one can write the NNLO approximation of the bin-average of the multipoles of the anisotropic power spectrum as

$$
\begin{aligned}
P^\ell_{\text{bin}}(k) \simeq (2\ell + 1) \sum_{\ell'} \Bigg[& P_{\ell'}(\bar{k})\Pi_{\ell\ell'}(k) + P'_{\ell'}(\bar{k}) \left(\Xi_{\ell\ell'}(k) - \bar{k}\Pi_{\ell\ell'}(k) \right) + \\
& + \frac{1}{2} P''_{\ell'}(\bar{k}) \left(\Psi_{\ell\ell'}(k) - 2\bar{k}\Xi_{\ell\ell'}(k) + \bar{k}^2\Pi_{\ell\ell'}(k) \right) \Bigg].
\end{aligned}
\tag{B.39}
$$

Figure B.3 shows the accuracy that this effective perturbative approach can reach in the approximation of the bin-average of the different multipoles of the anisotropic power spectrum of galaxies. For the monopole, the "naive" effective approach performs basically as well as the perturbative approach at LO and at NLO. The NNLO however has a sensible gain in accuracy in approximating the bin-average of the monopole. For the quadrupole, the perturbative effective approach at LO and at NLO has similar performances, and in both cases they perform better than the naive effective approach. Finally, for the hexadecapole, going to higher orders in the perturbative effective approach, gives better and better results.

The NNLO of the perturbative effective approach gives a very good approximation of the full bin-average of the multipoles, and a naive `Python` implementation takes $\mathcal{O}(1)$ ms to perform a cubic spline interpolation of the multipoles, the computation of the derivatives, and the computation of approximated bin-averages, provided that the effective momenta $\bar{k}$ and the bin-averages of the products of Legendre polynomials $\Pi_{\ell\ell'}(k)$, $\Xi_{\ell\ell'}(k)$, and $\Psi_{\ell\ell'}(k)$ have been precomputed.
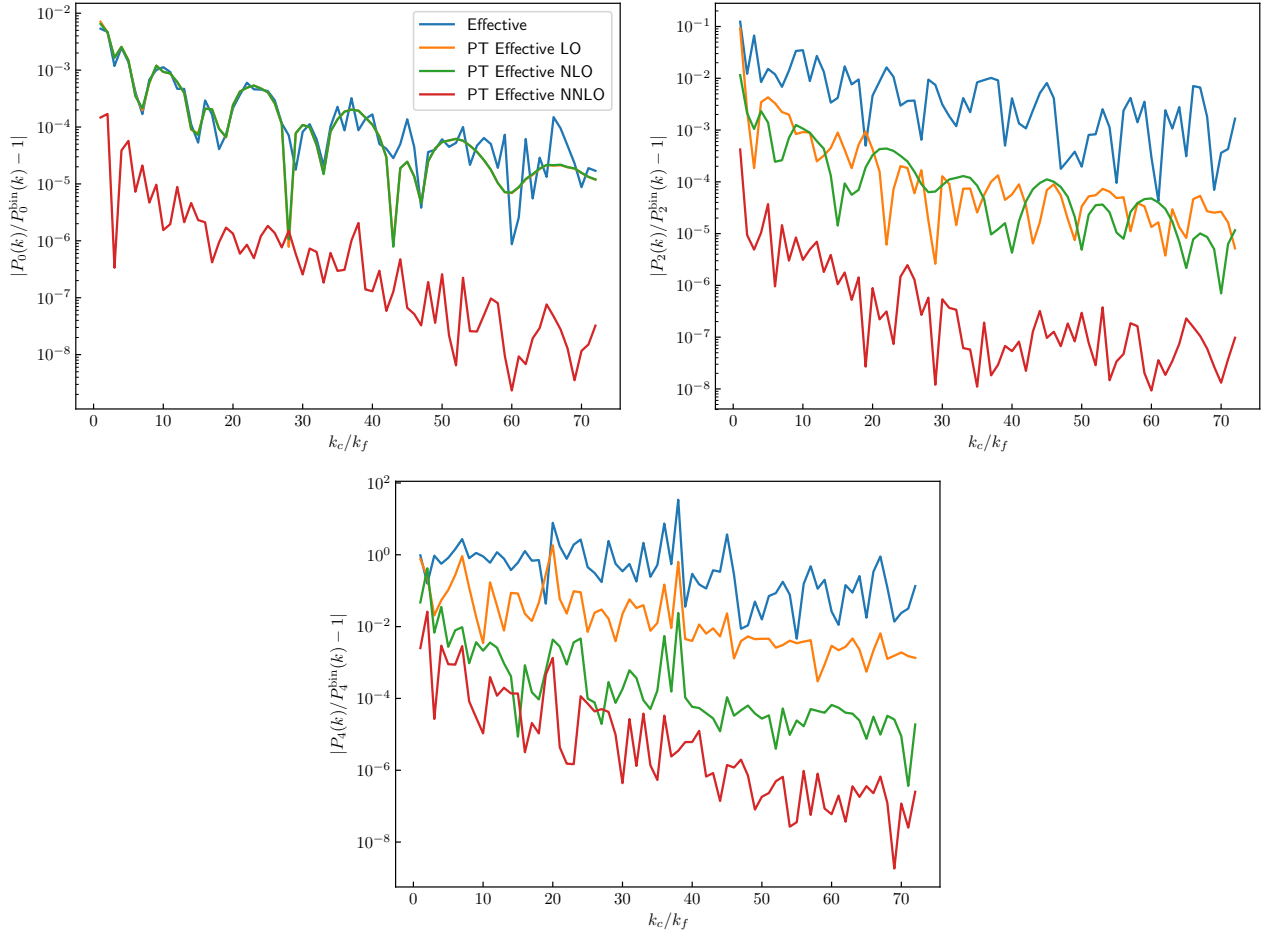
Figure B.3: Similar to the previous figures, but for the power spectrum multipoles in redshift space: $\ell = 0$ is on the upper left panel, $\ell = 2$ is on the upper right panel, while $\ell = 4$ is on the lower panel.

Bibliography

[1] Planck Collaboration, Y. Akrami, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi et al., *Planck 2018 results. I. Overview and the cosmological legacy of Planck*, *arXiv e-prints* (2018) arXiv:1807.06205 [`1807.06205`].

[2] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi et al., *Planck 2018 results. VI. Cosmological parameters*, *arXiv e-prints* (2018) arXiv:1807.06209 [`1807.06209`].

[3] A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P. M. Garnavich et al., *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*, *Astron. J.* **116** (1998) 1009 [`astro-ph/9805201`].

[4] S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro et al., *Measurements of Omega and Lambda from 42 High-Redshift Supernovae*, *Astrophys. J.* **517** (1999) 565 [`astro-ph/9812133`].

[5] D. J. Eisenstein, M. R. Blanton, I. Zehavi, N. A. Bahcall, J. Brinkmann, J. Loveday et al., *The Small-Scale Clustering of Luminous Red Galaxies via Cross-Correlation Techniques*, *Astrophys. J.* **619** (2005) 178 [`arXiv:astro-ph/0411559`].

[6] eBOSS Collaboration, S. Alam, M. Aubert, S. Avila, C. Balland, J. E. Bautista et al., *The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Cosmological Implications from two Decades of Spectroscopic Surveys at the Apache Point observatory*, *arXiv e-prints* (2020) arXiv:2007.08991 [`2007.08991`].

[7] A. G. Sánchez, J. N. Grieb, S. Salazar-Albornoz, S. Alam, F. Beutler, A. J. Ross et al., *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: combining correlated Gaussian posterior distributions*, *Mon. Not. R. Astron. Soc.* **464** (2017) 1493 [`1607.03146`].

[8] J. N. Grieb, A. G. Sánchez, S. Salazar-Albornoz, R. Scoccimarro, M. Crocce, C. Dalla Vecchia et al., *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: Cosmological implications of the Fourier space wedges of the final sample*, *Mon. Not. R. Astron. Soc.* **467** (2017) 2085 [`1607.03143`].

[9] F. Beutler, H.-J. Seo, A. J. Ross, P. McDonald, S. Saito, A. S. Bolton et al., *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Fourier space*, *Mon. Not. R. Astron. Soc.* **464** (2017) 3409 [`1607.03149`].

[10] V. C. Rubin and J. Ford, W. Kent, *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*, *Astrophys. J.* **159** (1970) 379.

[11] F. Zwicky, *Die Rotverschiebung von extragalaktischen Nebeln*, *Helvetica Physica Acta* **6** (1933) 110.

[12] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley et al., *Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results*, *ArXiv e-prints* (2012) [1212.5226].

[13] D. J. Eisenstein, H.-J. Seo and M. White, *On the Robustness of the Acoustic Scale in the Low-Redshift Clustering of Matter*, *Astrophys. J.* **664** (2007) 660 [astro-ph/0604361].

[14] M. Viel, K. Markovič, M. Baldi and J. Weller, *The non-linear matter power spectrum in warm dark matter cosmologies*, *Mon. Not. R. Astron. Soc.* **421** (2012) 50 [arXiv: 1107.4094].

[15] A. J. Cuesta, V. Niro and L. Verde, *Neutrino mass limits: Robust information from the power spectrum of galaxy surveys*, *Physics of the Dark Universe* **13** (2016) 77 [1511.05983].

[16] G. d'Amico, J. Gleyzes, N. Kokron, K. Markovic, L. Senatore, P. Zhang et al., *The cosmological analysis of the SDSS/BOSS data from the Effective Field Theory of Large-Scale Structure*, *Journal of Cosmology and Astro-Particle Physics* **2020** (2020) 005 [1909.05271].

[17] E. Castorina, N. Hand, U. Seljak, F. Beutler, C.-H. Chuang, C. Zhao et al., *Redshift-weighted constraints on primordial non-Gaussianity from the clustering of the eBOSS DR14 quasars in Fourier space*, *Journal of Cosmology and Astro-Particle Physics* **2019** (2019) 010 [1904.08859].

[18] K. S. Dawson, D. J. Schlegel, C. P. Ahn, S. F. Anderson, É. Aubourg, S. Bailey et al., *The Baryon Oscillation Spectroscopic Survey of SDSS-III*, *Astron. J.* **145** (2013) 10 [1208.0022].

[19] K. S. Dawson, J.-P. Kneib, W. J. Percival, S. Alam, F. D. Albareti, S. F. Anderson et al., *The SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Overview and Early Data*, *Astron. J.* **151** (2016) 44 [1508.04473].

[20] R. Laureijs, J. Amiaux, S. Arduini, J. . Auguères, J. Brinchmann, R. Cole et al., *Euclid Definition Study Report*, *ArXiv: 1110.3193* (2011) [1110.3193].

[21] M. Levi, C. Bebek, T. Beers, R. Blum, R. Cahn, D. Eisenstein et al., *The DESI Experiment, a whitepaper for Snowmass 2013*, *ArXiv: 1308.0847* (2013) [1308.0847].

[22] Y.-S. Song, A. Taruya and A. Oka, *Cosmology with anisotropic galaxy clustering from the combination of power spectrum and bispectrum*, *Journal of Cosmology and Astro-Particle Physics* **8** (2015) 007 [1502.03099].

[23] K. C. Chan and L. Blot, *Assessment of the information content of the power spectrum and bispectrum*, *Phys. Rev. D* **96** (2017) 023528 [1610.06585].

[24] P. Gagrani and L. Samushia, *Information Content of the Angular Multipoles of Redshift-Space Galaxy Bispectrum*, *Mon. Not. R. Astron. Soc.* **467** (2017) 928 [1610.03488].

[25] A. Chudaykin and M. M. Ivanov, *Measuring neutrino masses with large-scale structure: Euclid forecast with controlled theoretical error*, Journal of Cosmology and Astro-Particle Physics **2019** (2019) 034 [1907.06666].

[26] C. Hahn, F. Villaescusa-Navarro, E. Castorina and R. Scoccimarro, *Constraining M_ν with the bispectrum. Part I. Breaking parameter degeneracies*, Journal of Cosmology and Astro-Particle Physics **2020** (2020) 040 [1909.11107].

[27] M. Tellarini, A. J. Ross, G. Tasinato and D. Wands, *Galaxy bispectrum, primordial non-Gaussianity and redshift space distortions*, Journal of Cosmology and Astro-Particle Physics **6** (2016) 014 [1603.06814].

[28] D. Yamauchi, S. Yokoyama and K. Takahashi, *Multitracer technique for galaxy bispectrum: An application to constraints on nonlocal primordial non-Gaussianities*, Phys. Rev. D **95** (2017) 063530 [1611.03590].

[29] D. Karagiannis, A. Lazanu, M. Liguori, A. Raccanelli, N. Bartolo and L. Verde, *Constraining primordial non-Gaussianity with bispectrum and power spectrum from upcoming optical and radio surveys*, Mon. Not. R. Astron. Soc. **478** (2018) 1341 [1801.09280].

[30] V. Yankelevich and C. Porciani, *Cosmological information in the redshift-space bispectrum*, Mon. Not. R. Astron. Soc. (2018) [1807.07076].

[31] D. Yamauchi, S. Yokoyama and H. Tashiro, *Constraining modified theories of gravity with the galaxy bispectrum*, Phys. Rev. D **96** (2017) 123516 [1709.03243].

[32] N. Agarwal, V. Desjacques, D. Jeong and F. Schmidt, *Information content in the redshift-space galaxy power spectrum and bispectrum*, arXiv e-prints (2020) arXiv:2007.04340 [2007.04340].

[33] H. Gil-Marín, J. Noreña, L. Verde, W. J. Percival, C. Wagner, M. Manera et al., *The power spectrum and bispectrum of SDSS DR11 BOSS galaxies - I. Bias and gravity*, Mon. Not. R. Astron. Soc. **451** (2015) 539 [1407.5668].

[34] H. Gil-Marín, L. Verde, J. Noreña, A. J. Cuesta, L. Samushia, W. J. Percival et al., *The power spectrum and bispectrum of SDSS DR11 BOSS galaxies - II. Cosmological interpretation*, Mon. Not. R. Astron. Soc. **452** (2015) 1914 [1408.0027].

[35] H. Gil-Marín, W. J. Percival, L. Verde, J. R. Brownstein, C.-H. Chuang, F.-S. Kitaura 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*, Mon. Not. R. Astron. Soc. **465** (2017) 1757 [1606.00439].

[36] D. W. Pearson and L. Samushia, *A Detection of the Baryon Acoustic Oscillation Features in the SDSS BOSS DR12 Galaxy Bispectrum*, ArXiv e-prints (2017) [1712.04970].

[37] Z. Slepian, D. J. Eisenstein, J. R. Brownstein, C.-H. Chuang, H. Gil-Marín, S. Ho et al., *Detection of baryon acoustic oscillation features in the large-scale three-point correlation function of SDSS BOSS DR12 CMASS galaxies*, Mon. Not. R. Astron. Soc. **469** (2017) 1738 [1607.06097].

[38] R. Scoccimarro, *The bispectrum: From theory to observations*, *Astrophys. J.* **544** (2000) 597 [astro-ph/0004086].

[39] E. Sefusatti and R. Scoccimarro, *Galaxy bias and halo-occupation numbers from large-scale clustering*, *Phys. Rev. D* **71** (2005) 063001 [astro-ph/0412626].

[40] E. Sefusatti, M. Crocce, S. Pueblas and R. Scoccimarro, *Cosmology and the bispectrum*, *Phys. Rev. D* **74** (2006) 023522 [arXiv: astro-ph/0604505].

[41] J. Byun, A. Eggemeier, D. Regan, D. Seery and R. E. Smith, *Towards optimal cosmological parameter recovery from compressed bispectrum statistics*, *Mon. Not. R. Astron. Soc.* **471** (2017) 1581 [1705.04392].

[42] D. Gualdi, M. Manera, B. Joachimi and O. Lahav, *Maximal compression of the redshift-space galaxy power spectrum and bispectrum*, *Mon. Not. R. Astron. Soc.* **476** (2018) 4045 [1709.03600].

[43] J. J. M. Carrasco, M. P. Hertzberg and L. Senatore, *The effective field theory of cosmological large scale structures*, *Journal of High Energy Physics* **9** (2012) 82 [1206.2926].

[44] L. Senatore, *Bias in the effective field theory of large scale structures*, *Journal of Cosmology and Astro-Particle Physics* **11** (2015) 007 [1406.7843].

[45] L. Senatore and M. Zaldarriaga, *Redshift Space Distortions in the Effective Field Theory of Large Scale Structures*, *ArXiv e-prints* (2014) [1409.1225].

[46] L. Senatore and M. Zaldarriaga, *The IR-resummed Effective Field Theory of Large Scale Structures*, *Journal of Cosmology and Astro-Particle Physics* **2** (2015) 013 [1404.5954].

[47] T. Nishimichi, G. D'Amico, M. M. Ivanov, L. Senatore, M. Simonović, M. Takada et al., *Blinded challenge for precision cosmology with large-scale structure: results from effective field theory for the redshift-space galaxy power spectrum*, *arXiv e-prints* (2020) arXiv:2003.08277 [2003.08277].

[48] M. M. Ivanov, M. Simonović and M. Zaldarriaga, *Cosmological Parameters from the BOSS Galaxy Power Spectrum*, *JCAP* **05** (2020) 042 [1909.05277].

[49] T. Tröster, A. G. Sánchez, M. Asgari, C. Blake, M. Crocce, C. Heymans et al., *Cosmology from large-scale structure. Constraining ΛCDM with BOSS*, *Astron. Astrophys.* **633** (2020) L10 [1909.11006].

[50] R. Scoccimarro, S. Colombi, J. N. Fry, J. A. Frieman, E. Hivon and A. Melott, *Nonlinear evolution of the bispectrum of cosmological perturbations*, *Astrophys. J.* **496** (1998) 586 [astro-ph/9704075].

[51] E. Sefusatti, *One-loop perturbative corrections to the matter and galaxy bispectrum with non-gaussian initial conditions*, *Phys. Rev. D* **80** (2009) 123002 [0905.0717].

[52] E. Sefusatti, M. Crocce and V. Desjacques, *The matter bispectrum in N-body simulations with non-Gaussian initial conditions*, *Mon. Not. R. Astron. Soc.* **406** (2010) 1014 [1003.0007].

[53] T. Baldauf, L. Mercolli, M. Mirbabayi and E. Pajer, *The bispectrum in the Effective Field Theory of Large Scale Structure*, *Journal of Cosmology and Astro-Particle Physics* **5** (2015) 007 [1406.4135].

[54] R. E. Angulo, S. Foreman, M. Schmittfull and L. Senatore, *The one-loop matter bispectrum in the Effective Field Theory of Large Scale Structures*, *Journal of Cosmology and Astro-Particle Physics* **10** (2015) 039 [1406.4143].

[55] A. Lazanu, T. Giannantonio, M. Schmittfull and E. P. S. Shellard, *Matter bispectrum of large-scale structure: Three-dimensional comparison between theoretical models and numerical simulations*, *Phys. Rev. D* **93** (2016) 083517 [1510.04075].

[56] N. McCullagh, D. Jeong and A. S. Szalay, *Toward accurate modelling of the non-linear matter bispectrum: standard perturbation theory and transients from initial conditions*, *Mon. Not. R. Astron. Soc.* **455** (2016) 2945 [1507.07824].

[57] I. Hashimoto, Y. Rasera and A. Taruya, *Precision cosmology with redshift-space bispectrum: A perturbation theory based model at one-loop order*, *Phys. Rev. D* **96** (2017) 043526 [1705.02574].

[58] A. Lazanu and M. Liguori, *The two and three-loop matter bispectrum in perturbation theories*, *Journal of Cosmology and Astro-Particle Physics* **4** (2018) 055 [1803.03184].

[59] A. Eggemeier, R. Scoccimarro and R. E. Smith, *Bias Loop Corrections to the Galaxy Bispectrum*, *arXiv e-prints* (2018) arXiv:1812.03208 [1812.03208].

[60] V. Desjacques, D. Jeong and F. Schmidt, *The galaxy power spectrum and bispectrum in redshift space*, *Journal of Cosmology and Astro-Particle Physics* **12** (2018) 035 [1806.04015].

[61] C. Rampf and T. Buchert, *Lagrangian perturbations and the matter bispectrum I: fourth-order model for non-linear clustering*, *Journal of Cosmology and Astro-Particle Physics* **6** (2012) 21 [1203.4260].

[62] C. Rampf and Y. Y. Y. Wong, *Lagrangian perturbations and the matter bispectrum II: the resummed one-loop correction to the matter bispectrum*, *Journal of Cosmology and Astro-Particle Physics* **6** (2012) 18 [1203.4261].

[63] M. Simonović, T. Baldauf, M. Zaldarriaga, J. J. Carrasco and J. A. Kollmeier, *Cosmological perturbation theory using the FFTLog: formalism and connection to QFT loop integrals*, *Journal of Cosmology and Astro-Particle Physics* **4** (2018) 030 [1708.08130].

[64] J. N. Grieb, A. G. Sánchez, S. Salazar-Albornoz and C. Dalla Vecchia, *Gaussian covariance matrices for anisotropic galaxy clustering measurements*, *Mon. Not. R. Astron. Soc.* **457** (2016) 1577 [1509.04293].

[65] P. Monaco, T. Theuns and G. Taffoni, *The pinocchio algorithm: pinpointing orbit-crossing collapsed hierarchical objects in a linear density field*, *Mon. Not. R. Astron. Soc.* **331** (2002) 587 [arXiv:astro-ph/0109323].

[66] P. Monaco, E. Sefusatti, S. Borgani, M. Crocce, P. Fosalba, R. K. Sheth et al., *An accurate tool for the fast generation of dark matter halo catalogues*, Mon. Not. R. Astron. Soc. **433** (2013) 2389 [1305.1505].

[67] E. Munari, P. Monaco, E. Sefusatti, E. Castorina, F. G. Mohammad, S. Anselmi et al., *Improving fast generation of halo catalogues with higher order Lagrangian perturbation theory*, Mon. Not. R. Astron. Soc. **465** (2017) 4658 [1605.04788].

[68] J. Hartlap, T. Schrabback, P. Simon and P. Schneider, *The non-Gaussianity of the cosmic shear likelihood or how odd is the Chandra Deep Field South?*, Astron. Astrophys. **504** (2009) 689 [0901.3269].

[69] E. Sellentin and A. F. Heavens, *Parameter inference with estimated covariance matrices*, Mon. Not. R. Astron. Soc. **456** (2016) L132 [1511.05969].

[70] K. C. Chan, R. Scoccimarro and R. K. Sheth, *Gravity and large-scale nonlocal bias*, Phys. Rev. D **85** (2012) 083509 [1201.3614].

[71] T. Lazeyras, C. Wagner, T. Baldauf and F. Schmidt, *Precision measurement of the local bias of dark matter halos*, Journal of Cosmology and Astro-Particle Physics **2** (2016) 018 [1511.01096].

[72] A. Eggemeier, R. Scoccimarro, M. Crocce, A. Pezzotta and A. G. Sánchez, *Testing one-loop galaxy bias – I. Power spectrum*, arXiv e-prints (2020) arXiv:2006.09729 [2006.09729].

[73] A. Oddo, E. Sefusatti, C. Porciani, P. Monaco and A. G. Sánchez, *Toward a robust inference method for the galaxy bispectrum: likelihood function and model selection*, Journal of Cosmology and Astro-Particle Physics **2020** (2020) 056 [1908.01774].

[74] A. Einstein, *Die Feldgleichungen der Gravitation*, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin (1915) 844.

[75] F. W. Dyson, A. S. Eddington and C. Davidson, *A Determination of the Deflection of Light by the Sun's Gravitational Field, from Observations Made at the Total Eclipse of May 29, 1919*, Philosophical Transactions of the Royal Society of London Series A **220** (1920) 291.

[76] A. Einstein, *Die Grundlage der allgemeinen Relativitätstheorie*, Annalen der Physik **354** (1916) 769.

[77] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley et al., *Observation of Gravitational Waves from a Binary Black Hole Merger*, Physical Review Letters **116** (2016) 061102 [1602.03837].

[78] Event Horizon Telescope Collaboration, K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay et al., *First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole*, Astrophys. J. Lett. **875** (2019) L1 [1906.11238].

[79] H. P. Robertson, *Kinematics and World-Structure*, Astrophys. J. **82** (1935) 284.

[80] A. G. Walker, *On Milne's Theory of World-Structure*, *Proceedings of the London Mathematical Society* **42** (1937) 90.

[81] A. Einstein, *Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie*, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin* (1917) 142.

[82] A. Friedmann, *Über die Krümmung des Raumes*, *Zeitschrift fur Physik* **10** (1922) 377.

[83] A. Friedmann, *Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes*, *Zeitschrift fur Physik* **21** (1924) 326.

[84] Euclid Collaboration, A. Blanchard, S. Camera, C. Carbone, V. F. Cardone, S. Casas et al., *Euclid preparation: VII. Forecast validation for Euclid cosmological probes*, *arXiv e-prints* (2019) arXiv:1910.09273 [`1910.09273`].

[85] Planck Collaboration, Y. Akrami, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi et al., *Planck 2018 results. IX. Constraints on primordial non-Gaussianity*, *arXiv e-prints* (2019) arXiv:1905.05697 [`1905.05697`].

[86] F. Bernardeau, S. Colombi, E. Gaztañaga and R. Scoccimarro, *Large-scale structure of the universe and cosmological perturbation theory*, Phys. Rep. **367** (2002) 1 [`astro-ph/0112551`].

[87] A. Lewis, A. Challinor and A. Lasenby, *Efficient computation of cosmic microwave background anisotropies in closed friedmann-robertson-walker models*, *Astrophys. J.* **538** (2000) 473 [`astro-ph/9911177`].

[88] J. Lesgourgues, *The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview*, *ArXiv: 1104.2932* (2011) [`1104.2932`].

[89] M. H. Goroff, B. Grinstein, S.-J. Rey and M. B. Wise, *Coupling of modes of cosmological mass density fluctuations*, *Astrophys. J.* **311** (1986) 6.

[90] V. Desjacques, D. Jeong and F. Schmidt, *Large-scale galaxy bias*, Phys. Rep. **733** (2018) 1 [`1611.09787`].

[91] J. E. Pollack, R. E. Smith and C. Porciani, *Modelling large-scale halo bias using the bispectrum*, *Mon. Not. R. Astron. Soc.* **420** (2012) 3469 [`1109.3458`].

[92] T. Baldauf, U. Seljak, V. Desjacques and P. McDonald, *Evidence for quadratic tidal tensor bias from the halo bispectrum*, *Phys. Rev. D* **86** (2012) 083540 [`1201.4827`].

[93] P. McDonald, *Clustering of dark matter tracers: Renormalizing the bias parameters*, *Phys. Rev. D* **74** (2006) 103512 [`astro-ph/0609413`].

[94] V. Assassi, D. Baumann, D. Green and M. Zaldarriaga, *Renormalized halo bias*, *Journal of Cosmology and Astro-Particle Physics* **8** (2014) 056 [`1402.5916`].

[95] N. Kaiser, *Clustering in real space and in redshift space*, *Mon. Not. R. Astron. Soc.* **227** (1987) 1.

[96] R. Scoccimarro, *Redshift-space distortions, pairwise velocities, and nonlinearities*, *Phys. Rev. D* **70** (2004) 083007 [arXiv:astro-ph/0407214].

[97] A. Taruya, T. Nishimichi and S. Saito, *Baryon acoustic oscillations in 2D: Modeling redshift-space power spectrum from perturbation theory*, *Phys. Rev. D* **82** (2010) 063522 [1006.0699].

[98] D. J. Eisenstein, I. Zehavi, D. W. Hogg, R. Scoccimarro, M. R. Blanton, R. C. Nichol et al., *Detection of the baryon acoustic peak in the large-scale correlation function of sdss luminous red galaxies*, *Astrophys. J.* **633** (2005) 560 [astro-ph/0501171].

[99] D. Blas, M. Garny, M. M. Ivanov and S. Sibiryakov, *Time-sliced perturbation theory II: baryon acoustic oscillations and infrared resummation*, *Journal of Cosmology and Astro-Particle Physics* **7** (2016) 028 [1605.02149].

[100] M. Ivanov and S. Sibiryakov, *Infrared resummation for biased tracers in redshift space*, *Journal of Cosmology and Astro-Particle Physics* **2018** (2018) .

[101] D. Blas, M. Garny, M. M. Ivanov and S. Sibiryakov, *Time-sliced perturbation theory for large scale structure I: general formalism*, *Journal of Cosmology and Astro-Particle Physics* **2016** (2016) 052 [1512.05807].

[102] D. J. Eisenstein and W. Hu, *Baryonic features in the matter transfer function*, *Astrophys. J.* **496** (1998) 605 [arXiv:astro-ph/9709112].

[103] Z. Vlah, U. Seljak, M. Yat Chu and Y. Feng, *Perturbation theory, effective field theory, and oscillations in the power spectrum*, *Journal of Cosmology and Astro-Particle Physics* **2016** (2016) 057 [1509.02120].

[104] T. Baldauf, M. Mirbabayi, M. Simonović and M. Zaldarriaga, *Equivalence principle and the baryon acoustic peak*, *Phys. Rev. D* **92** (2015) 043514 [1504.04366].

[105] E. Sefusatti, M. Crocce, R. Scoccimarro and H. M. P. Couchman, *Accurate estimators of correlation functions in Fourier space*, *Mon. Not. R. Astron. Soc.* **460** (2016) 3624 [1512.07295].

[106] D. Gualdi, H. Gil-Marín, R. L. Schuhmann, M. Manera, B. Joachimi and O. Lahav, *Enhancing BOSS bispectrum cosmological constraints with maximal compression*, *Mon. Not. R. Astron. Soc.* **484** (2019) 3713 [1806.02853].

[107] A. G. Sánchez, R. Scoccimarro, M. Crocce, J. N. Grieb, S. Salazar-Albornoz, C. Dalla Vecchia et al., *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: Cosmological implications of the configuration-space clustering wedges*, *Mon. Not. R. Astron. Soc.* **464** (2017) 1640 [1607.03147].

[108] M. Crocce, D. Blas and R. Scoccimarro, *In prep.*, .

[109] C. Alcock and B. Paczynski, *An evolution free test for non-zero cosmological constant*, *Nature* **281** (1979) 358.

[110] V. Springel, *The cosmological simulation code GADGET-2*, Mon. Not. R. Astron. Soc. **364** (2005) 1105 [`astro-ph/0505010`].

[111] A. G. Sánchez, E. A. Kazin, F. Beutler, C.-H. Chuang, A. J. Cuesta, D. J. Eisenstein et al., *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological constraints from the full shape of the clustering wedges*, Mon. Not. R. Astron. Soc. **433** (2013) 1202 [`1303.4396`].

[112] V. Springel, N. Yoshida and S. D. M. White, *GADGET: a code for collisionless and gasdynamical cosmological simulations*, New Astronomy **6** (2001) 79 [`astro-ph/0003162`].

[113] DESI Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, L. E. Allen et al., *The DESI Experiment Part I: Science,Targeting, and Survey Design*, ArXiv e-prints (2016) [`1611.00036`].

[114] A. Schneider, R. Teyssier, D. Potter, J. Stadel, J. Onions, D. S. Reed et al., *Matter power spectrum and the challenge of percent accuracy*, Journal of Cosmology and Astro-Particle Physics **4** (2016) 047 [`1503.05920`].

[115] M. Biagetti, T. Lazeyras, T. Baldauf, V. Desjacques and F. Schmidt, *Verifying the consistency relation for the scale-dependent bias from local primordial non-Gaussianity*, Mon. Not. R. Astron. Soc. **468** (2017) 3277 [`1611.04901`].

[116] M. Lippich, A. G. Sánchez, M. Colavincenzo, E. Sefusatti, P. Monaco, L. Blot et al., *Comparing approximate methods for mock catalogues and covariance matrices - I. Correlation function*, Mon. Not. R. Astron. Soc. **482** (2019) 1786 [`1806.09477`].

[117] L. Blot, M. Crocce, E. Sefusatti, M. Lippich, A. G. Sánchez, M. Colavincenzo et al., *Comparing approximate methods for mock catalogues and covariance matrices II: Power spectrum multipoles*, Mon. Not. R. Astron. Soc. (2019) [`1806.09497`].

[118] M. Colavincenzo, E. Sefusatti, P. Monaco, L. Blot, M. Crocce, M. Lippich et al., *Comparing approximate methods for mock catalogues and covariance matrices - III: bispectrum*, Mon. Not. R. Astron. Soc. **482** (2019) 4883 [`1806.09499`].

[119] P. Monaco, *Approximate Methods for the Generation of Dark Matter Halo Catalogs in the Age of Precision Cosmology*, Galaxies **4** (2016) 53 [`1605.07752`].

[120] R. Angulo, M. Fasiello, L. Senatore and Z. Vlah, *On the statistics of biased tracers in the Effective Field Theory of Large Scale Structures*, Journal of Cosmology and Astro-Particle Physics **9** (2015) 029 [`1503.08826`].

[121] T. Fujita, V. Mauerhofer, L. Senatore, Z. Vlah and R. Angulo, *Very Massive Tracers and Higher Derivative Biases*, ArXiv e-prints (2016) [`1609.00717`].

[122] E. O. Nadler, A. Perko and L. Senatore, *On the bispectra of very massive tracers in the Effective Field Theory of Large-Scale Structure*, Journal of Cosmology and Astro-Particle Physics **2** (2018) 058 [`1710.10308`].

[123] K. Hoffmann, J. Bel and E. Gaztañaga, *Comparing halo bias from abundance and clustering*, *Mon. Not. R. Astron. Soc.* **450** (2015) 1674 [1503.00313].

[124] K. Hoffmann, J. Bel and E. Gaztañaga, *Linear and non-linear bias: predictions versus measurements*, *Mon. Not. R. Astron. Soc.* **465** (2017) 2225 [1607.01024].

[125] S. Saito, T. Baldauf, Z. Vlah, U. Seljak, T. Okumura and P. McDonald, *Understanding higher-order nonlocal halo bias at large scales by combining the power spectrum with the bispectrum*, *Phys. Rev. D* **90** (2014) 123522 [1405.1447].

[126] T. Lazeyras and F. Schmidt, *Beyond LIMD bias: a measurement of the complete set of third-order halo bias parameters*, *Journal of Cosmology and Astro-Particle Physics* **9** (2018) 008 [1712.07531].

[127] M. M. Abidi and T. Baldauf, *Cubic halo bias in Eulerian and Lagrangian space*, *Journal of Cosmology and Astro-Particle Physics* **2018** (2018) 029 [1802.07622].

[128] C. Modi, E. Castorina and U. Seljak, *Halo bias in Lagrangian space: estimators and theoretical predictions*, *Mon. Not. R. Astron. Soc.* **472** (2017) 3959 [1612.01621].

[129] P. Catelan, C. Porciani and M. Kamionkowski, *Two ways of biasing galaxy formation*, *Mon. Not. R. Astron. Soc.* **318** (2000) L39 [astro-ph/0005544].

[130] T. W. Anderson, *An introduction to multivariate statistical analysis*. Wiley New York, 1958.

[131] S. Dodelson and M. D. Schneider, *The effect of covariance estimator error on cosmological parameter constraints*, *Phys. Rev. D* **88** (2013) 063537 [1304.2593].

[132] A. N. Taylor, B. Joachimi and T. D. Kitching, *Putting the precision in precision cosmology: How accurate should your data covariance matrix be?*, *Mon. Not. R. Astron. Soc.* **432** (2013) 1928 [1212.4359].

[133] A. Taylor and B. Joachimi, *Estimating cosmological parameter covariance*, *Mon. Not. R. Astron. Soc.* **442** (2014) 2728 [1402.6983].

[134] W. J. Percival, A. J. Ross, A. G. Sánchez, L. Samushia, A. Burden, R. Crittenden et al., *The clustering of Galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: including covariance matrix errors*, *Mon. Not. R. Astron. Soc.* **439** (2014) 2531 [1312.4841].

[135] E. Sellentin and A. F. Heavens, *Quantifying lost information due to covariance matrix estimation in parameter inference*, *Mon. Not. R. Astron. Soc.* **464** (2017) 4658 [1609.00504].

[136] J. Hartlap, P. Simon and P. Schneider, *Why your model parameter confidences might be too optimistic. Unbiased estimation of the inverse covariance matrix*, *Astron. Astrophys.* **464** (2007) 399 [astro-ph/0608064].

[137] C. Hahn, F. Beutler, M. Sinha, A. Berlind, S. Ho and D. W. Hogg, *Likelihood non-Gaussianity in large-scale structure analyses*, *Mon. Not. R. Astron. Soc.* **485** (2019) 2956 [1803.06348].

[138] E. Sefusatti, M. Crocce and V. Desjacques, *The halo bispectrum in N-body simulations with non-Gaussian initial conditions*, *Mon. Not. R. Astron. Soc.* **425** (2012) 2903 [1111.6966].

[139] D. Foreman-Mackey, D. W. Hogg, D. Lang and J. Goodman, *emcee: The MCMC Hammer*, *Publications of the Astronomical Society of the Pacific* **125** (2013) 306 [`1202.3665`].

[140] J. Goodman and J. Weare, *Ensemble samplers with affine invariance*, *Communications in Applied Mathematics and Computational Science, Vol. 5, No. 1, p. 65-80, 2010* **5** (2010) 65.

[141] R. E. Kass and A. E. Raftery, *Bayes factors*, *Journal of the American Statistical Association* **90** (1995) 773 [`https://www.tandfonline.com/doi/pdf/10.1080/01621459.1995.10476572`].

[142] H. Jeffreys, *Theory of Probability*, The International series of monographs on physics. Clarendon Press, 1961.

[143] J. M. Dickey, *The weighted likelihood ratio, linear hypotheses on normal location parameters*, *Ann. Math. Statist.* **42** (1971) 204.

[144] R. Trotta, *Applications of Bayesian model selection to cosmological parameters*, *Mon. Not. R. Astron. Soc.* **378** (2007) 72 [`astro-ph/0504022`].

[145] M. .-. . Proceedings of Conference on Foundational Questions in Statistical Inference, Aarhus, ed., *The direct use of likelihood for significance testing*, Department of Theoretical Statistics, Institute of Mathematics, University of Aarhus, 1974.

[146] D. J. Spiegelhalter, N. G. Best, B. P. Carlin and A. Van Der Linde, *Bayesian measures of model complexity and fit*, *Journal of the Royal Statistical Society Series B* **64** (2002) 583.

[147] H. Akaike, *Information Theory and an Extension of the Maximum Likelihood Principle*, pp. 199–213. Springer New York, New York, NY, 1998.

[148] A. Gelman, J. B. Carlin, H. S. Stern and D. B. Rubin, *Bayesian Data Analysis*. Chapman and Hall/CRC, 2nd ed. ed., 2004.

[149] M. Plummer, *Penalized loss functions for Bayesian model comparison*, *Biostatistics* **9** (2008) 523 [`http://oup.prod.sis.lan/biostatistics/article-pdf/9/3/523/17742541/kxm049.pdf`].

[150] A. van der Linde, *Dic in variable selection*, *Statistica Neerlandica* **59** (2005) 45 [`https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9574.2005.00278.x`].

[151] A. van der Linde, *A bayesian view of model complexity*, *Statistica Neerlandica* **66** (2012) 253 [`https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9574.2011.00518.x`].

[152] T. Ando, *Predictive bayesian model selection*, *American Journal of Mathematical and Management Sciences* **31** (2011) 13 [`https://doi.org/10.1080/01966324.2011.10737798`].

[153] R. Fisher, *Statistical Methods for Research Workers*, Biological monographs and manuals. Oliver and Boyd, 1925.

[154] D. B. Rubin, *Bayesianly justifiable and relevant frequency calculations for the applied statistician*, *Ann. Statist.* **12** (1984) 1151.

[155] X.-L. Meng, *Posterior predictive p-values*, *Ann. Statist.* **22** (1994) 1142.

[156] A. Gelman, X. Meng and H. Stern, *Posterior predictive assessment of model fitness via realized discrepancies*, *Statistica Sinica* **6** (1996) 733.

[157] T. Baldauf, U. Seljak, R. E. Smith, N. Hamaus and V. Desjacques, *Halo stochasticity from exclusion and nonlinear clustering*, *Phys. Rev. D* **88** (2013) 083507 [`1305.2917`].

[158] D. Ginzburg, V. Desjacques and K. C. Chan, *Shot noise and biased tracers: A new look at the halo model*, *Phys. Rev. D* **96** (2017) 083528 [`1706.08738`].

[159] K. F. Werner and C. Porciani, *Renormalisation of linear halo bias in N-body simulations*, *arXiv e-prints* (2019) arXiv:1907.03774 [`1907.03774`].

[160] A. Oddo, F. Rizzo, E. Sefusatti and C. Porciani, *In prep.*, .

[161] R. K. Sheth, K. C. Chan and R. Scoccimarro, *Nonlocal Lagrangian bias*, *Phys. Rev. D* **87** (2013) 083002 [`1207.7117`].

[162] A. Lewis, *Efficient sampling of fast and slow cosmological parameters*, *Phys. Rev. D* **87** (2013) 103529 [`1304.4473`].

[163] A. Gelman and D. B. Rubin, *Inference from Iterative Simulation Using Multiple Sequences*, *Statistical Science* **7** (1992) 457.

[164] J. E. McEwen, X. Fang, C. M. Hirata and J. A. Blazek, *FAST-PT: a novel algorithm to calculate convolution integrals in cosmological perturbation theory*, *Journal of Cosmology and Astro-Particle Physics* **9** (2016) 015 [`1603.04826`].

[165] A. Lewis, *GetDist: a Python package for analysing Monte Carlo samples*, `1910.13970`.

[166] R. Casas-Miranda, H. J. Mo, R. K. Sheth and G. Boerner, *On the distribution of haloes, galaxies and mass*, *Mon. Not. R. Astron. Soc.* **333** (2002) 730 [`astro-ph/0105008`].

[167] N. Hamaus, U. Seljak, V. Desjacques, R. E. Smith and T. Baldauf, *Minimizing the stochasticity of halos in large-scale structure surveys*, *Phys. Rev. D* **82** (2010) 043515 [`1004.5377`].

[168] U. Seljak, N. Hamaus and V. Desjacques, *How to Suppress the Shot Noise in Galaxy Surveys*, *Physical Review Letters* **103** (2009) 091303 [`0904.2963`].

[169] D. Alkhanishvili, M. Biagetti, A. Lazanu, A. Oddo, C. Porciani and E. Sefusatti, *In prep.*, .

[170] M. Biagetti, *In prep.*, .

[171] J. Byun, A. Oddo, C. Porciani and E. Sefusatti, *In prep.*, .

[172] J. R. Fergusson, D. M. Regan and E. P. S. Shellard, *Rapid separable analysis of higher order correlators in large-scale structure*, *Phys. Rev. D* **86** (2012) 063511 [`1008.1730`].

[173] D. M. Regan, M. M. Schmittfull, E. P. S. Shellard and J. R. Fergusson, *Universal non-Gaussian initial conditions for N-body simulations*, *Phys. Rev. D* **86** (2012) 123524 [`1108.3813`].

[174] M. Liguori, E. Sefusatti, J. R. Fergusson and E. P. S. Shellard, *Primordial non-gaussianity and bispectrum measurements in the cosmic microwave background and large-scale structure*, Advances in Astronomy **2010** (2010) [`1001.4707`].

[175] G. D'Amico, M. Musso, J. Noreña and A. Paranjape, *An improved calculation of the non-gaussian halo mass function*, Journal of Cosmology and Astro-Particle Physics **2** (2011) 1 [`1005.1203`].

[176] R. Scoccimarro, *Fast estimators for redshift-space clustering*, Phys. Rev. D **92** (2015) 083532 [`1506.02729`].

[177] J. N. Fry, *The evolution of bias*, Astrophys. J. Lett. **461** (1996) L65.

[178] P. Catelan, F. Lucchin, S. Matarrese and C. Porciani, *The bias field of dark matter haloes*, Mon. Not. R. Astron. Soc. **297** (1998) 692 [`arXiv:astro-ph/9708067`].

[179] D. Wadekar and R. Scoccimarro, *The Galaxy Power Spectrum Multipoles Covariance in Perturbation Theory*, arXiv e-prints (2019) arXiv:1910.02914 [`1910.02914`].