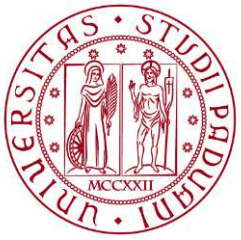

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Head Office: Università degli Studi di Padova

Department: Dipartimento di Fisica e Astronomia "Galileo Galilei"

Ph.D. COURSE IN: Physics

SERIES XVIII

Toward accurate and precise cosmology with galaxy redshift surveys

Coordinator: Prof. Giulio Monaco

Supervisor: Prof. Alvise Raccanelli

Co-Supervisor: Dott. Nicola Bellomo

Ph.D. student: Francesco Spezzati

List of Publications

- **Equivalence of the field level inference and conventional analyses on large scales**
  Francesco Spezzati, Marco Marinucci, Marko Simonović
  e-Print: 2507.05378 [astro-ph.CO]
  Submitted to PRD

- **Observed unequal time power spectrum**
  Francesco Spezzati, Eleonora Vanzan, Alvise Raccanelli, Zvonimir Vlah, Daniele Bertacca
  e-Print: 2503.09759 [astro-ph.CO]
  Published in: Phys.Rev.D 112 (2025) 2, 023530

- **The constraining power of the marked power spectrum: an analytical study**
  Marco Marinucci, Gabriel Jung, Michele Liguori, Andrea Ravenni, Francesco Spezzati, Adam Andrews, Marco Baldi, William R. Coulton, Dionysios Karagiannis, Francisco Villaescusa-Navarro, Benjamin Wandlet
  e-Print: 2411.14377 [astro-ph.CO]
  Published in: JCAP 09 (2025) 036

- **3D-radial correlation function**
  Francesco Spezzati, Alvise Raccanelli
  e-Print: 2408.01495 [astro-ph.CO]
  Published in: Phys.Rev.D 111 (2025) 12, 123539

Abstract

Maximizing the extraction of cosmological information while ensuring accurate modeling is essential in the era of precision galaxy surveys. This thesis contributes to both aspects. First, we examine the accuracy of common approximations—such as the plane-parallel and flat-sky assumptions—showing that they can bias cosmological parameter estimates, and we propose alternative models, including a Radial 3D approach and an unequal-time formalism, that preserve three-dimensional clustering information with minimal complexity. Second, we investigate innovative methods and statistics such as field level inference, the galaxy trispectrum, and the marked power spectrum, aimed at extracting additional information on top of the conventional analyses. Although these capture higher-order correlations, we find that the improvement beyond standard analyses is limited. Taken together, these results help to clarify the strengths and limitations of current modeling strategies and suggest directions for the analysis of forthcoming high-precision survey data.

Contents

I Introduction and Motivation

Since the dawn of time, humanity has gazed at the night sky looking for answers, striving to uncover the origins, composition, and ultimate fate of the Universe. These pursuits have transcended cultures and epochs, shaping philosophical and scientific thought for thousands of years. Despite this long-standing curiosity — and the emergence of modern scientific tools and models — many aspects of the cosmos remain deeply enigmatic. This thesis aims to be one more small step — one more grain of sand — in the long-standing endeavor to seek answers to these profound questions that continue to inspire the fundamental sense of awe of humankind. Nowadays, cosmology has transformed from philosophical speculation into a data-driven precision science thanks to both theoretical and observational efforts carried out in the last centuries. The current understanding of the Universe can be schematized as follows.

Our Universe is homogeneous and isotropic on large scales ($\gtrsim 100$ Mpc) [1], it is spatially flat [2], the contribution to the total energy budget is about 5% from baryons, 25% from Cold Dark Matter, and the remaining from dark energy in the form of a cosmological constant Λ that causes the present-day accelerated expansion of the Universe [2, 3]. This is why the current standard cosmological model is called ΛCDM. Initial conditions are set by inflation [4–7], a phase of accelerated expansion that naturally led to isotropy, homogeneity, and spatial flatness, while also generating the primordial seeds of cosmic structure. The subsequent evolution of these density fluctuations is governed by gravity as described by General Relativity. [3]

These information was mostly obtained thanks to observations of the Cosmic Microwave Background [1, 8, 9] which allowed cosmologists to obtain precise measurements of the fundamental cosmological parameters governing our Universe, such as the baryon and cold dark matter density, the amplitude of the primordial fluctuations, the value of the expansion rate of the universe today, the level of non-Gaussianity of the initial conditions, etc. However, even though new CMB experiments are planned for the near future and an $O(1)$ improvement on the constraining power on cosmological parameters is expected [10] (especially coming from the increase in the precision in the measurement of the polarization E-modes) the information content of CMB data is almost saturated and it is reaching the cosmic variance limit. Nevertheless, unique information about the early universe can be obtained by the measurement of the B-mode polarizations of the CMB, as they would constitute compelling evidence for the presence of a stochastic background of gravitational waves generated during inflation [11, 12]. Detecting or constraining this signal is therefore of central importance for testing inflationary models and alternatives. Considerable efforts in this direction are being carried out in the form of future space-based experiments like LiteBIRD [13] or ground-based observatories such as CMB-S4 [14].

On the other hand, the study of the Large-Scale Structure (LSS) of the Universe offers the potential for significant improvements in the precision of cosmological parameter measurements, owing to the vast cosmic volume probed by current and future galaxy surveys. Even more importantly, the study of galaxy clustering can provide unique information about the late-time evolution of the Universe, directly testing gravity and dark energy, shedding light on some of the most intriguing puzzles of modern cosmology. Is General Relativity the correct theory of gravity at all scales? What is dark energy? Is it really a cosmological constant? What is dark matter made of?

Modern galaxy surveys such as BOSS [15] have cataloged millions of galaxy positions with increasing precision and their data have been analyzed by many independent groups and proven to be an invaluable source of cosmological information. This observational effort is expanding rapidly: the Dark Energy

[1] While isotropy can be proven by observations, homogeneity is an underlying philosophical assumption of the model which has not been tested yet.

[2] A clarification is needed here. The curvature of our Universe is defined by the cosmological parameter $\Omega_k = 0.0007 \pm 0.0019$ [1] and even tough this value is consistent with a spatially flat universe ($\Omega_k = 0$) the probability for Ω_k to be exactly 0 is null as it is the probability to chose a fixed number in a set of infinite continuous values.

[3] Our Universe is not invariant under time translation as energy is in not conserved in General Relativity.

Spectroscopic Instrument (DESI) [16] has already collected data on tens of millions of galaxies across a huge portion of the sky and analyzed Data Releases 1-2. The Euclid satellite [17] and the SPHEREx [18] mission are now collecting data and upcoming missions such as Vera Rubin Observatory [19], the Nancy Grace Roman Space Telescope [20], and MegaMapper [21] will extend this work to even larger volumes of the Universe. The outcome of these efforts will be the most precise three-dimensional map of the Universe in the history of human civilization. Such a map represents the cosmic web: a multitude of dense galaxy clusters, elongated filaments, and vast empty regions. However, the interpretation of these structures and the extraction of cosmological information from their statistical patterns — measured with exceptional precision by current and forthcoming surveys — requires a commensurate effort on the theoretical side. In fact, an increase in precision in the measurements can be fruitful if and only if theoretical predictions are accurate enough to analyze and interpret the data. This thesis focuses on this aspect: the development of more accurate and precise theoretical models for the LSS and their possible application to current and future galaxy surveys.

In the first part of the thesis, after laying down the formalism describing the observed linear galaxy overdensity field, we examine the validity of several commonly employed approximations in the geometrical and mathematical modeling of galaxy statistics, demonstrating that some of them can introduce biases in the cosmological analyses of modern galaxy surveys, and we develop new, more accurate models designed to overcome these limitations. In the opening of the second part, we provide a concise review of the theoretical framework employed to extend the modeling of the galaxy overdensity field into mildly non-linear scales. Then, we asses the effectiveness of recently proposed methods — such as field level inference, higher order statistics and non-standard summary statistics — that aim to extract more precise cosmological information from galaxy survey data compared to conventional analyses. Finally, in the concluding part of the thesis, we present a case study of a realistic data analysis using mock galaxy survey datasets designed to emulate the observations expected from the Nancy Grace Roman Telescope.

II Galaxy clustering at linear order

A Conventions and background evolution

Let us start by schematizing some basic notions and conventions that will be used throughout this thesis. Unless otherwise noted, we will always use natural units in which $c = 1$ and the following Fourier transform convention:

$$f(\boldsymbol{x}) = \int \frac{d\boldsymbol{q}}{(2\pi)^3} e^{i\boldsymbol{q}\cdot\boldsymbol{x}} f(\boldsymbol{q}), \quad f(\boldsymbol{q}) = \int d\boldsymbol{x} e^{-i\boldsymbol{q}\cdot\boldsymbol{x}} f(\boldsymbol{x}) \,. \tag{1}$$

In all that follows, we will assume that the Universe can be modeled by a spatially flat, homogeneous, and isotropic background, plus some small perturbations. The metric describing the background is given by the Friedmann-Lemaître-Robertson-Walker (FLRW) line element:

$$ds^2 = -dt^2 + a^2(t) \left[dx^2 + x^2 d^2\Omega \right] \,, \tag{2}$$

where t is cosmic time, x is the radial comoving distance and $a(t)$ is the scale factor (which depends only on time due to homogeneity and isotropy) and accounts for the expansion of the Universe. The expansion rate is given by the Hubble function:

$$H = \frac{\dot{a}}{a} \,. \tag{3}$$

The time variable used in this thesis will often be the redshift, which is related to the difference between the wavelength of the emitted photon λ_{source} and the observed one λ_{obs}, due to the Doppler effect

$$1 + z = \frac{\lambda_{\text{obs}}}{\lambda_{\text{source}}} = \frac{a_0}{a(t)} \,, \tag{4}$$

where a_0 is the value of the scale factor today and we will use the normalization $a_0 = 1$. We further assume that the background spacetime is described by a perfect fluid with stress-energy tensor given by

$$T_{\mu\nu} = u^\mu u^\nu (\rho + P) + g_{\mu\nu} P \,, \tag{5}$$

where u^μ is the 4-velocity of the fluid $g_{\mu\nu} u^\mu u^\nu = -1$, ρ and P are the energy density and isotropic pressure of the background. Plugging the FLRW metric and the expression for the stress-energy tensor of the perfect fluid into the Einstein field equations

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu} \,, \tag{6}$$

where Λ is the cosmological constant, one obtains the first two Friedmann equations for a spatially flat Universe

$$H^2 = \frac{8\pi G}{3} \rho + \frac{\Lambda}{3} \,,$$
$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3} (\rho + 3P) + \frac{\Lambda}{3}$$

where $\dot{} \equiv d/dt$ As a consequence of the Bianchi identity, the left-hand side of Einstein's equations must satisfy $T^{\mu\nu}{}_{,\mu} = 0$ (where $_{,\mu}$ indicates covariant derivatives) so that we get the continuity equation

$$\dot{\rho} + 3H(\rho + P) = 0\,. \tag{7}$$

Only two out of the three previous equations are independent. In order to close the system of equations and describe the background dynamics of each component of our Universe, we must specify their equation of state.

$$w = \frac{P}{\rho} \tag{8}$$

We have that $w = 0$ for matter, $w = 1/3$ for radiation, and $w = -1$ for the cosmological constant. Plugging the equation of state into the continuity equation, we get the time evolution of each component.

$$\rho_i(t) = \rho_{0,i} a(t)^{-3(1+w_i)}\,, \tag{9}$$

where i stands for each component of the Universe and ρ_0 is the background energy density of each species today. This means that matter is diluted by the volume expansion of the universe $\rho_m \propto a^{-3}$, radiation experiences a further dilution due to redshift so that $\rho_r \propto a^{-4}$ and the energy density of the dark energy component is constant. One can then define the energy density of each species in terms of the critical energy density today as

$$\Omega_i = \frac{\rho_i}{\rho_c}, \quad \rho_c = \frac{3H_0^2}{8\pi G}\,, \tag{10}$$

and the physical densities are defined as

$$\omega_i = \Omega_i h^2\,, \tag{11}$$

where h is defined in terms of the Hubble parameter today as $H_0 = 100h\text{km Mpc}^{-1}\text{ s}^{-1}$.

Having obtained the time evolution of the energy density of the species populating our Universe, we can get the evolution of the Hubble parameter.

$$H(z) = H_0\sqrt{\Omega_m(1+z)^3 + \Omega_r(1+z)^4 + \Omega_\Lambda}\,. \tag{12}$$

Then, imposing $ds^2 = 0$ as photons travel on null geodesics and considering only radial distances we have that comoving radial distance (in Mpc) from the observer to a source is

$$x(z) = \int dz' \frac{1}{H(z)}\,, \tag{13}$$

while the angular diameter distance is

$$D_A(z) = (1+z)^{-1} x(z)\,. \tag{14}$$

.

B *Growth of structure and gravitational instability*

Let us now study the evolution of dark matter perturbations at late times i.e., matter-dominated and later epoch. The formalism schematized here is extensively treated in classic references and books like [22–25].

While General Relativity governs the growth of structures in the Universe, it is well known that on scales that are

$$aH \ll k \ , \tag{15}$$

one can work in the Newtonian limit in order to compute theoretical predictions for the matter overdensity field. Moreover, while gravity is in general a non-linear theory, on large scales, matter perturbations are very small (i.e., their variance is $\ll 1$) so that they can be treated as linear. The scale at which non-linear corrections start to become non-negligible is often taken to be around $k_{\rm lin} \sim 0.1h/{\rm Mpc}$ at $z = 0$ and in general it increases with redshift. It should be emphasized that this represents only a rough estimate. In practical applications, careful attention must be paid to the determination of this quantity, and the validity of the linear model for the relevant scales should always be verified, typically through comparison with numerical simulations or by explicitly assessing the magnitude of non-linear contributions.

In the first part of the thesis, we will assume that matter is a non-relativistic perfect fluid. While this simplification does not impact linear theory results, we will see in the second part of the thesis that the assumption of perfect fluid is intrinsically wrong and corrections arising from fluid viscosity arise naturally when one goes beyond linear order.

We start by defining the relations between the physical coordinates $\boldsymbol{r}$ and the comoving coordinate $\boldsymbol{x}$. We have

$$\boldsymbol{r} = a(t)\boldsymbol{x}, \quad \nabla_r = \frac{1}{a}\nabla_x \ , \quad \frac{\partial \boldsymbol{r}}{\partial t} = \boldsymbol{v} = H\boldsymbol{r} + \boldsymbol{u} \ , \tag{16}$$

where $\boldsymbol{u} = a\partial\boldsymbol{x}/\partial t$ is the peculiar velocity. For a generic function, the Lagrangian (or convective derivative) is defined as:

$$\frac{Df(\boldsymbol{r},t)}{Dt} = \left.\frac{\partial f(\boldsymbol{r},t)}{\partial t}\right|_{\boldsymbol{r}} + \frac{\partial f(\boldsymbol{r},t)}{\partial \boldsymbol{r}}\frac{\partial \boldsymbol{r}}{\partial t} \ , \tag{17}$$

and since the two Lagrangian derivatives must be equal in both coordinates

$$\begin{aligned} \frac{Df(\boldsymbol{r},t)}{Dt} &= \frac{Df(\boldsymbol{x},t)}{Dt} \ , \\ \left.\frac{\partial f(\boldsymbol{r},t)}{\partial t}\right|_{\boldsymbol{r}} &= \left.\frac{\partial f(\boldsymbol{x},t)}{\partial t}\right|_{\boldsymbol{x}} - H(\boldsymbol{x}\cdot\nabla_x)f(\boldsymbol{x},t) \ . \end{aligned} \tag{18}$$

Starting from the continuity, Euler, and Poisson equations for the dark matter fluid:

$$\begin{aligned} \left.\frac{d\rho}{dt}\right|_{\boldsymbol{r}} &= -\nabla_r(\rho\boldsymbol{v}) \ , \\ (\partial_t|_r + \boldsymbol{v}\cdot\nabla_r)\boldsymbol{v} &= -\frac{1}{\rho}\nabla_r p - \nabla_r\Phi \ , \\ \nabla_r^2\Phi &= 4\pi G\rho \ , \end{aligned} \tag{19}$$

we express all quantities in comoving coordinates and introduce small perturbations, such as:

$$\begin{aligned} \rho(\boldsymbol{x},t) &= \bar{\rho}(t) + \delta\rho(\boldsymbol{x},t) \ , \\ \Phi(\boldsymbol{x},t) &= \Phi_b(\boldsymbol{x},t) + \phi(\boldsymbol{x},t) \ , \end{aligned} \tag{20}$$

where

$$\nabla^2\phi = 4\pi G\delta\rho(\boldsymbol{x},t)a^2 \ , \tag{21}$$

and ϕ is called the peculiar gravitational potential and the last equation is called the cosmological Poisson equation. The source $\delta\rho$ can have both positive and negative signs. We have that the perturbed continuity equations read:

$$\left.\frac{\partial\rho}{\partial t}\right|_x = -\frac{1}{a}\nabla_x(\rho\boldsymbol{u}) - 3H\rho \ . \tag{22}$$

For the Euler equation, we have:

$$\left.\frac{\partial Ha\boldsymbol{x}}{\partial t}\right|_x + \left.\frac{\partial\boldsymbol{u}}{\partial t}\right|_x + H\boldsymbol{u} + \frac{1}{a}\boldsymbol{u}\cdot\nabla_x\boldsymbol{u} = -\frac{1}{\rho a}\nabla_x p - \frac{1}{a}\nabla_x\Phi_b - \frac{1}{a}\nabla_x\phi \ , \tag{23}$$

and since background quantities satisfy Friedmann equations, we have that the equation of motion for the peculiar velocity reduces to:

$$\left.\frac{\partial\boldsymbol{u}}{\partial t}\right|_x + H\boldsymbol{u} + \frac{1}{a}\boldsymbol{u}\cdot\nabla_x\boldsymbol{u} = -\frac{1}{a\rho}\nabla_x p - \frac{1}{a}\nabla_x\phi \ . \tag{24}$$

We can now linearize the perturbed continuity and Euler equations such that $\rho = \bar{\rho}(1+\delta)$ where $\delta << 1$. Then it is convenient to go to Fourier space and study the evolution for each mode $\boldsymbol{k}$. We have that:

$$\begin{aligned} \dot{\delta} &= -\frac{1}{a}i\boldsymbol{k}\boldsymbol{u} \ , \\ \dot{\boldsymbol{u}} + H\boldsymbol{u} &= -\frac{1}{a}i\boldsymbol{k}(\phi + c_s^2\delta) \ , \\ \phi &= \frac{1}{k^2}4\pi G\bar{\rho}(t)\delta a^2 \ , \end{aligned} \tag{25}$$

where $c_s^2 = \partial p/\partial\rho$ is used.

The velocity is pure gradient since the component of the velocity $\boldsymbol{u}$ that is orthogonal to the $\boldsymbol{k}$-mode obeys the equation:

$$\begin{aligned} \dot{\boldsymbol{u}}_\perp + H\boldsymbol{u}_\perp &= 0 \ , \\ \boldsymbol{u}_\perp &\propto a^{-1} \ . \end{aligned} \tag{26}$$

Physically, this result is a consequence of Kelvin's circulation theorem that states that vorticity is conserved along the fluid stream line in the absence of dissipative forces. If there is no initial vorticity, the velocity is pure gradient at late times. Differentiating the continuity equation w.r.t. time, replacing it in the Euler equation and using the Poisson equation, one obtains a second order in time differential equation:

$$\ddot{\delta} + 2H\dot{\delta} + \left(\frac{c_s^2k^2}{a^2} - 4\pi G\bar{\rho}\right)\delta = 0 \ , \tag{27}$$

and one scales larger than the Jeans wavelength:

$$k_j = \frac{a}{c_s}\sqrt{4\pi G\bar{\rho}} \ , \tag{28}$$

we have the approximate equation:

$$\ddot{\delta} + 2H\dot{\delta} - 4\pi G\bar{\rho}\delta \sim 0 \ . \tag{29}$$

The physical interpretation of this equation is that matter perturbations evolve and grow according to a source term that involves the matter density and are diluted by the expansion of the Universe. Therefore,

since at linear order the differential equation for δ is only time dependent, we can separate the time and space dependence of the growing solution of Equation (29):

$$\delta(\boldsymbol{k}, t) = D(t)\delta(\boldsymbol{k}) \ , \tag{30}$$

where $D(t)$ is the growth factor. This solution must satisfy the continuity equation:

$$\frac{i\boldsymbol{k}}{k^2} f\delta(\boldsymbol{k}, t) = \frac{1}{aH}\boldsymbol{u} \ , \tag{31}$$

where we have defined the growth rate:

$$f = \frac{d\ln D}{d\ln a} \ . \tag{32}$$

C *From matter to galaxies: linear theory*

Having obtained the equations of motion for the linear matter overdensity field, we have now to relate it to the galaxy overdensity, defined as

$$\delta_g(\boldsymbol{x}.z) = \frac{\rho(\boldsymbol{x}, z)}{\bar{\rho}(\boldsymbol{x}, z)} - 1 \ . \tag{33}$$

At linear scales it turns out that a simple ansatz of the form [26–29]

$$\delta_g(\boldsymbol{x}, z) = b(z)\delta(\boldsymbol{x}, z) \ , \tag{34}$$

is enough to make accurate predictions. To build an intuitive schematic understanding of how this relation arises, consider the following. Galaxies form within dark matter halos, which themselves originate from highly non-linear peaks of the matter overdensity field—regions where fluctuations exceed the threshold required for gravitational collapse. Crucially, the formation of dark matter halos, takes place over relatively small regions characterized by their typical Lagrangian radius R. This naturally suggests a separation between short-wavelength fluctuations, relevant for halo formation, and long-wavelength "background" modes with $\lambda \gg R$.

In the absence of long-wavelength fluctuations, each small region would have had the same probability of crossing the collapse threshold and thus forming a dark matter halo. However, the presence of a long-wavelength background modulates the statistics of the small-scale fluctuations: regions embedded within a large-scale overdensity have an enhanced probability of forming halos compared to those residing in underdense environments. The response of short-scale modes to the long-wavelength background can only depend on locally observable scalar quantities in the long modes, consistent with the symmetries of our Universe. By the Equivalence Principle, neither ϕ nor $\nabla\phi$ are locally observable, and therefore the leading-order contribution must be proportional to $\nabla^2\phi$. On large scales, this yields a relation of the type

$$\delta_g \propto \nabla^2\phi = b(t)\,\delta, \tag{35}$$

which motivates an ansatz of the form given in Equation 34, with $b(t)$ acting as a free coefficient encoding the physics of halo and galaxy formation.

While this parametrization can be approximated as local in space—since galaxy and halo formation occur on scales smaller than R—the bias coefficient b must necessarily depend on time. This follows from the fact that the Universe is not invariant under time translations, and thus the properties of galaxies at a given epoch depend on their full formation history. Within linear theory, however, we can always

separate the time and space dependence of the matter density perturbations as Equation (30), which allows one to absorb the relevant time dependence into a redefined effective bias parameter,

$$\tilde{b}(t) = \int dt'\, b(t,t')\, D(t,t'), \tag{36}$$

which is equivalent to assuming the linear bias relation of Equation 34. It should be stressed that this is a simplified treatment, valid only within the regime of linear theory. A more sophisticated bias expansion is required in order to accurately describe the clustering of galaxies on mildly non-linear scales. We will return to this issue in detail in the second part of the thesis.

Let us now introduce the concept of Redshift Space Distortions (RSD) and briefly describe how to model them in linear theory. The total velocity of the galaxies that we observe in the sky is given by a cosmological component due to the Hubble flow and the peculiar velocity of each galaxy falling in the gravitational potential wells. So, the *observed* position of a galaxy in the sky, if not at rest w.r.t. the background is, at linear order in velocity [24, 30]

$$\boldsymbol{x}_{\rm obs} = \boldsymbol{x} + \frac{1}{aH} u_{\parallel}(\boldsymbol{x})\hat{\boldsymbol{n}}\ , \tag{37}$$

This mapping defines two spaces that are called real space ($\boldsymbol{x}$) in which the galaxies are located in reality and redshift space ($\boldsymbol{x}_{obs}$) in which galaxies are observed. The number of galaxies in a particular region is conserved when we going from real to redshift space

$$n_{\rm g,obs}\left(\boldsymbol{x}_{\rm obs}\right) d^3x_{\rm obs} = n_{\rm g}(\boldsymbol{x}) d^3x = n_{\rm g}(\boldsymbol{x})\boldsymbol{J}\ , \tag{38}$$

and the Jacobian of the coordinate transformation is

$$J \equiv \left|\frac{d^3x}{d^3x_{\rm obs}}\right| = \left|\frac{dx}{dx_{\rm obs}}\right| \frac{x^2}{x^2_{\rm obs}} = \left(1 + \frac{u_{\parallel}}{aHx}\right)^{-2} \left|1 + \frac{1}{aH}\frac{\partial}{\partial x} u_{\parallel}\right|^{-1}\ . \tag{39}$$

At linear order we can express the density as $\bar{n}_g[1+\delta_g]$. In general $\bar{n}(\boldsymbol{x}_{obs}) \neq \bar{n}(\boldsymbol{x})$ and $\delta_{obs}(\boldsymbol{x}_{obs}) \neq \delta(\boldsymbol{x})$ and so we have that:

$$1 + \delta_{obs}(\boldsymbol{x}_{obs)} = [1+\delta(\boldsymbol{x})]\frac{\bar{n}(\boldsymbol{x})}{\bar{n}(\boldsymbol{x}_{obs})} \left(1 + \frac{u_{\parallel}}{aHx}\right)^{-2} \left|1 + \frac{1}{aH}\frac{\partial}{\partial x} u_{\parallel}\right|^{-1}\ . \tag{40}$$

Now we can linearize the previous equation to obtain:

$$\delta_{obs}(\boldsymbol{x}) = \delta(\boldsymbol{x}) - \left(\frac{\partial}{\partial x} + \frac{\alpha(\boldsymbol{x})}{x}\right)\frac{\boldsymbol{u}_{\parallel}}{aH}\ , \tag{41}$$

where we have defined $\alpha = \frac{d\ln x^2 \bar{n}(\boldsymbol{x})}{d\ln x}$ and we have used the fact that to keep everything at linear order in perturbations we have that $\delta_{obs}(\boldsymbol{x}_{obs}) \sim \delta_{obs}(\boldsymbol{x})$.

For what concerns the velocity, we assume that on large linear scales where gravity dominates we have

$$\boldsymbol{u}_m = \boldsymbol{u}_g\ . \tag{42}$$

guaranteed by the Equivalence principle.

D *Primordial non-Gaussianity of the local type*

The inflationary paradigm is today the leading framework for describing the physics of the very early Universe. It postulates a phase of rapid expansion with an approximately constant Hubble parameter, during which quantum fluctuations of light fields are stretched to super-horizon scales and subsequently freeze as classical perturbations. After horizon re-entry, these perturbations provide the initial conditions for the formation of the anisotropies observed in the cosmic microwave background and for the large-scale structure of the Universe.

In its simplest realization, namely single-field slow-roll inflation (assuming Bunch-Davis vacuum and the attractor solution), the curvature perturbation is predicted to be nearly Gaussian, adiabatic, and statistically homogeneous and isotropic. In this scenario, the statistical properties of the primordial perturbations are completely specified by the two-point correlation function, and higher-order connected correlators are suppressed by slow-roll parameters and thus are expected to be extremely small. As a consequence, the detection of primordial non-Gaussianity (PNG) at an observable level would constitute compelling evidence for physics beyond the simplest inflationary scenarios, offering powerful constraints on the nature of the fields and interactions operative in the early Universe. In this section, we will briefly schematize one possible way to extract information about PNG using galaxy clustering.

One particular manner to parametrize deviations from Gaussian initial conditions is by the local (in configuration space) relation [31–33]:

$$\Phi_{\rm NG} = \phi + f_{\rm NL}\left[\phi^2 - \langle\phi^2\rangle\right] \ , \tag{43}$$

where Φ denotes Bardeen's gauge-invariant potential, which, on sub-horizon scales, reduces to the usual Newtonian gravitational potential and the dimensionless parameter $f_{\rm NL}$ (or more in general the dimensionless function) describes deviations from perfectly Gaussian initial conditions. Here ϕ is the Gaussian random field and the second term accounts for deviations from Gaussianity.

Being local in configuration space, Equation (43) couples long and short modes in Fourier space. For this reason, to investigate local-type non-Gaussianity, the bispectrum of the primordial potential in the squeezed limit (one of the wavenumbers is much smaller than the other two) is often studied. Simplest models of inflation in which the inflaton is the only dynamical field (single clock inflation) predict a small three-point function in this limit, the amplitude being fixed by the well-known consistency relations [34–36]. This is because in such models, the effects of the long mode on the short scales can be reabsorbed in a rescaling of coordinates. In the particular case of single field slow roll inflation, such consistency relations predict a level of local Primordial non-Gaussianity of order of the slow roll parameters ϵ, η[37, 38].[4] On the other hand, sizable deviations from Gaussian initial conditions of the form of (43) are expected for all the models in which the fluctuations of an additional light field, different from the inflaton, contribute to the curvature perturbations we observe [39].[5] Examples of this general scenario are the curvaton models and multi-field inflationary models.

However, working at the linear level in δ_g, we are interested in how deviations from Gaussian initial conditions affect the galaxy two point correlation function. As previously stated, the primordial bispectrum generated by Primordial non-Gaussianity (PNG) of the local type couples small scale density perturbations with large scale potential modes. This affects dark matter halo clustering at late time so that a non-zero $f_{\rm NL}$ in Equation (43) introduces a scale-dependent modification of the halo bias [33, 43, 44];

---

[4] In standard single field slow roll inflation the expression for the bispectrum of the curvature perurbations is not of the local form as it includes terms that are important for different configurations of triangles.

[5] A bispectrum peaking at equilateral or flattened triangles is a peculiar feature of interactions in single-field inflation where the inflaton is the only degree of freedom [40–42].

we can write the total non-Gaussian bias in the case of local shape PNG as [43, 45]:

$$b_{\rm NG}(z,k) = b_G(z) + b_\phi f_{\rm NL} \frac{3\Omega_{0m}H_0^2}{2k^2T(k)g(z)} \ , \tag{44}$$

where $b_G(z)$ is the usual scale-independent bias, calculated assuming Gaussian initial conditions, $T(k)$ is the matter transfer function and $g(z)$ is the growth factor normalized to the scale factor during the matter-dominated era. A scale dependence in halo bias can also arise from non-local types of primordial non-Gaussianity. However, the resulting scale dependence is significantly weaker than the characteristic k^{-2} scaling predicted by the local model [46]. For this reason, constraints on primordial non-Gaussianity from the galaxy power spectrum typically focus on the local type [47]. Moreover, Equation (44) tells us that measurements of $f_{\rm NL}$ strongly depend on the knowledge of b_ϕ , which represents the response of the galaxy number density n_g to a rescaling of the normalization amplitude of the power spectrum σ_8 i.e., $b_\phi = 2\frac{dlnn_g}{dln\sigma_8}$ [48]. Since measuring b_ϕ directly from the definition is a non-trivial task, the perfect degeneracy between b_ϕ and $f_{\rm NL}$ represents a problem for reaching high-precision measurements of local PNG (see e.g., [49, 50]). Some efforts have been made in this direction see e.g., [51–55] but since this topic is currently the subject of ongoing discussions, and for the sake of simplicity, here we assume the so-called universality relation and a fair sample of halos so that

$$b_\phi = 2\delta_c(b_1 - 1) \ , \tag{45}$$

and $\delta_{\rm c}$ is the critical value of the matter overdensity for spherical collapse in an Einstein-de Sitter universe taken to be $\delta_{\rm c} = 1.686$ [56]. In general, a modification $\delta_c \to q\delta_c$ is often assumed, in which q has the role of a fudge factor that accounts, for example, for elliptical collapse. For simplicity in all that follows, we assume that $q = 1$.

E Galaxy 2-point correlation function

Since the Universe we observe today is a particular realization of a stochastic process occurring during inflation, the only way to extract cosmological information from the large-scale structure is through summary statistics, which compress the total information available in a given cosmic volume on some observable statistical quantities that can be theoretically modeled. However, observations can only access one realization of the Universe i.e., the portion of it which falls inside the observer's past lightcone. Given this, a statistical approach requires two ingredients: the ergodic theorem and the "fair sample" hypothesis. The idea behind ergodicity is that ensemble averages can be traded for spatial averages over one realization of a random field – provided that spatial correlations decay sufficiently rapidly with separation, so that there are many statistically independent volumes in one realization. The fair sample hypothesis states that samples from well-separated spots are uncorrelated, and the collection of such samples is a statistical ensemble.

Assuming ergodicity and the fair sample hypothesis we can further proceed in our statistical approach to galaxy clustering. If gravity was a linear theory, in the absence of PNG, all the information would be included just in the 2-point correlation function. However, since General Relativity is a non-linear theory, even for perfectly Gaussian initial conditions, δ_g exhibits deviations from Gaussianity which are encoded in higher-order correlation functions. However, as we will show, the SNR of n-point correlation function decreases very steeply with n and the only higher-order correlation functions that one usually computes are the bispectrum and, sometimes, the trispectrum. In any case, since in the first part of this thesis we are working at linear level, we will study only the galaxy 2-point correlation function (2PCF) $\xi(s)$, which is defined as the excess probability compared to random of finding two galaxies at a given linear separation s. Notice that due to statistical homogeneity and isotropy, in real space the galaxy 2PCF depends only on the modulus s. In redshift space however, since redshift space

distortions break isotropy, as it is clear from Equation (41) (RSD are maximized along the line of sight), in general the 2PCF depends on both s and the orientation angle ϕ, defined by $\cos\phi = \hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{\epsilon}}_z$. We will also consider the Fourier counterpart of the galaxy 2PCF, the power spectrum which is defined as:

$$\langle \delta_g(\boldsymbol{k}) \delta_g(\boldsymbol{k}') \rangle = (2\pi)^3 \delta_D^3(\boldsymbol{k} + \boldsymbol{k}') P_g(k) \, . \tag{46}$$

where the Dirac delta imposes homogeneity and, due to the assumed isotropy, in real space $P(k)$ depends only on the modulus k while in redshift space we have to keep the $\boldsymbol{k}$ dependence since it also depends on the orientation angle, analogously to the 2PCF.

Let us further comment on the invariance under translation and rotation of the redshift space observables. As previously said, unlike in real space, the redshift space galaxy 2PCF is not isotropic (due to the ϕ dependence of the RSD kernel). What about the translational invariance?

In general it is also broken by the presence of an observer. In fact, in order to identify uniquely the triangle formed by a pair of galaxies in the sky and the observer, the 2PCF must depend on 3 degrees of freedom: for example the linear separation between the sources s, the orientation angle ϕ *and* the angular distance of the galaxies (other choices are possible and all of them lead to consistent results). Since the observer is fixed it is clear that this system is not invariant under space translation.[6]

All these concepts will be explained and illustrated in detail in Sect.III but the important point that we want to make here is that statistical isotropy and homogeneity in general are not satisfied by the 2PCF in redshift space. In fact, the definition of the power spectrum of Equation(46) is mathematically ill in redshift space. In other words, in Fourier space, the covariance of the galaxy overdensity field, which is guaranteed to be diagonal by homogeneity in real space, is not diagonal any more when going to redshift space and different k modes are, in general, coupled[24, 57]. This implies that in redshift space the power spectrum defined in Equation (46) is not the Fourier counterpart of the galaxy 2PCF.

This complicates a lot the problem of predicting observables in redshift space, and approximations are often used to overcome these issues. The most common is the 'distant observer' approximation (we will discuss it extensively in Sec.III) which requires that the distance between the sources is way smaller than the distance between the galaxies and the observer. In this approximation, translational invariance is recovered and the mapping between the galaxy overdensity in redshift space and real space can be written in a simple form. However, lots of efforts have been made in order to drop the assumptions and to test its regime of validity. These issues are the subject of the following two chapters.

[6] The system preserves the rotational symmetry about the observer and this justifies the often used spherical harmonics expansion of the correlation function. Spherical harmonics modes remain orthogonal with respect to the angular indices even in redshift space.

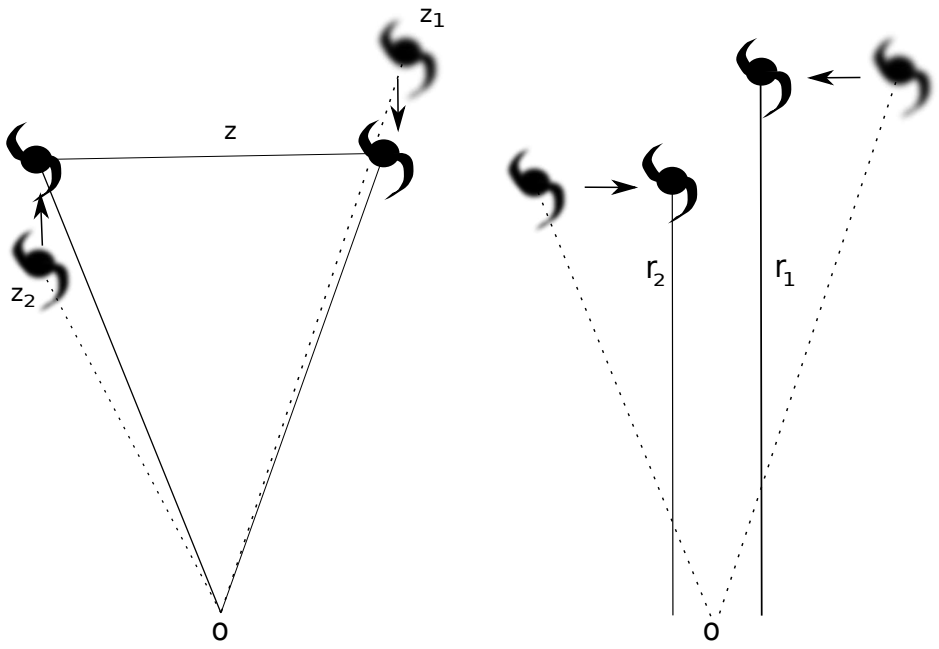

Figure 1: (*Left*) Flat sky approximation and (*Right*) plane parallel approximation

III 3D galaxy clustering: configuration space

We are now ready to go through the first part of the research work outlined in this thesis, which is focused on the mathematical and geometrical modeling of the galaxy two-point correlation function at linear scales. A key issue in the interpretation of galaxy clustering measurements from current and forthcoming surveys is whether the modeling must preserve the full three-dimensional structure of the observer–source system. Approximations that project this information onto a two-dimensional representation may discard essential information, potentially reducing the accuracy of cosmological parameter estimation, introducing systematic biases in our understanding of the underlying physics. In this chapter, we investigate this problem in configuration (position) space. When modeling the galaxy two point correlation function, one can safely use the plane parallel or flat sky approximations in the cases where the area covered by a galaxy survey is very small compared to the distance between the observer and the sources. The two approximations are similar but not strictly the same. The plane parallel (or distant observer) approximation, shown in the right panel of Fig. 1, assumes that since the linear separation between two galaxies is much smaller than their distance from us, we can treat the two lines of sight connecting them to the observer as parallel to each other, and we can approximate to 0 the opening angle between them (for the geometry of the problem treated here, see Figure 2). The flat sky limit, in the left panel of Fig. 1, implies that we are flattening the sky in the radial direction (i.e., we see all the galaxies at the same redshift), failing to account for the full 3-dimensionality of the problem. The flat sky approximation can be extended to include radial separations as recently investigated in [58–60] in the context of angular and Fourier space correlations. We will see later how this can be implemented in our configuration space treatment in order to improve upon the flat sky modeling, without necessarily implementing the fully exact calculation. In what follows, we refer to the flat sky model as a formalism that employs both the plane-parallel and the flat sky approximations.

Given that the forthcoming generation of galaxy surveys will probe larger cosmological volumes, both in area and in redshift range, it is worth investigating if these approximations fail to model with enough accuracy the statistical properties of galaxy clustering. In fact, for wide surveys such as SPHEREx [18],

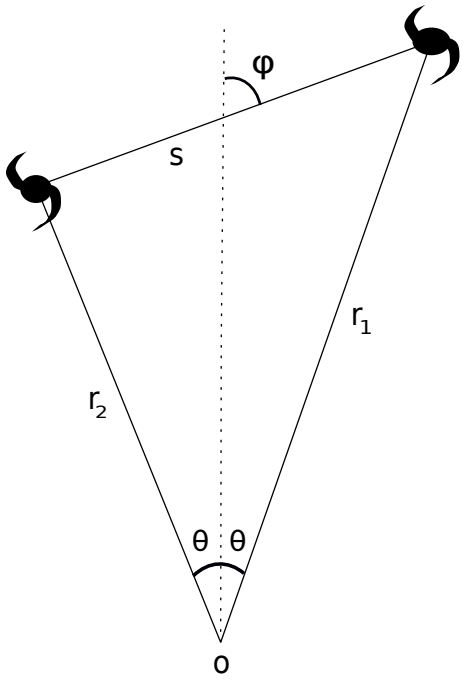

Figure 2: Coordinates system adopted in this section.

Euclid [17, 61, 62], DESI [16, 63] and the Vera Rubin Observatory [19], that will observe galaxy pairs with very large angular separations, it will be crucial to check the accuracy of the modeling employed. Scales probed by such surveys have not been accessible until now, so, in order to correctly extract all the cosmological information from the data, it will be necessary to pair newly achieved precision with matching accuracy.

Similarly, very deep surveys like the proposed SIRMOS [64] and MegaMapper [21], require to account for the full 3D geometry of the system, including galaxy pairs with large radial separation; moreover, high-z galaxies have larger redshift uncertainties compared to low-z ones, and this can result in having very thick bins in the radial direction, making the flat sky approximation less accurate. As we have seen, at linear order, in the flat sky approximation, the relation between the observed galaxy overdensity and the real space one can be written, in Fourier space, as:

$$\delta_{obs,g}(\mathbf{k},z) = D(z)\left[b(z) + \mu^2 f(z) - i\alpha(z)\mu f(z)k^{-1}\right]\delta_m(\mathbf{k}) \,, \tag{47}$$

where $\mu = \cos\phi$, with ϕ the angle with the line of sight and f is the growth rate parameter. In this work, as we want to evaluate best fit shifts on some example parameters, we parameterize the growth of structure as [65]:

$$f(z) = \Omega_m(z)^\gamma \,, \tag{48}$$

which has been found to be a good fit for many models; γ is called the growth index and in the standard ΛCDM+GR scenario $\gamma \sim 0.55$.

While the first two terms in Equation (47) are the usual Kaiser term [66, 67], the contribution proportional to $\alpha(z)$ is related to the fact that local overdensities around each galaxy can affect the apparent movement of galaxies when going from real- to redshift-space, and can be thought of as a

Doppler velocity term [68–73] . In the Newtonian limit we have [71]:

$$\alpha(z) = \frac{2}{\chi} + H(z)\frac{d\ln N(z)}{dz} \ , \tag{49}$$

where $N(z)$ is the galaxy redshift distribution and we called $\chi(z)$ the comoving distance of the source. A more general expression for the α-term is needed if one wants to include local general relativistic effects that can be non-negligible at very large scales [71, 74–76], but including such corrections is beyond the scope of this thesis.

The simplest model for the 2PCF employs both the flat sky and the plane parallel approximations, and galaxies are treated as they are all at the same redshift. In this scenario the 2PCF will depend only on the modulus of the linear separation between the two sources s, the orientation angle ϕ between $\mathbf{s}$ and the radial direction $\mathbf{n}$ and the effective redshift at which the 2PCF is calculated. In this work, we neglect Doppler terms in what we call the "flat sky" model in order the study the impact of neglecting such terms in the modelling (see Sec. **??** for details). In this scenario we have the standard result [67]:

$$\begin{aligned}\xi_{pp}(s,\phi,z) = D(z)^2\Bigg[\left(b(z)^2 + \frac{2b(z)f(z)}{3} + \frac{f(z)^2}{5}\right)\xi_0^0(s) + \\ \left(-\frac{4}{3}b(z)f(z) - \frac{4}{7}f(z)^2\right)\xi_2^0(s)\mathcal{L}_2(\cos(\phi)) + \frac{8}{35}f(z)^2\xi_2^0(s)\mathcal{L}_4(\cos(\phi))\Bigg] \ ,\end{aligned} \tag{50}$$

where $\mathcal{L}_\backslash$ are the the Legendre polynomials of order n, with the added non-Gaussian contributions:

$$\xi_{pp,ng}(s,\phi,z) = D(z)^2\left[\left(\Delta b(z)^2\xi^4_{0,ng}(s) + 2\Delta b(z)\left(b(z) + \frac{f(z)}{3}\right)\xi^2_{0.ng}(s) - \frac{4}{3}\Delta b(z)f(z)\xi^2_{2,ng}(s)\mathcal{L}(\cos(\phi))\right)\right] . \tag{51}$$

In the flat sky approximation, the RSD kernels of the observed 2PCF in the i-th redshift bin are often computed at a fixed $z_{\mathrm{eff},i}$. Here for simplicity we assume $z_{\mathrm{eff},i} = (z_{i,max} + z_{i,min})/2$, while in real data analyses the effective redshift is calculated applying specific weights to each galaxy (see e.g., [77–81]).

In order to drop the flat sky and plane parallel approximation, one can develop a formalism that allows to retain the full 3-dimensionality of the physics involved. Following the approach proposed in [68] and developed in [69, 70, 82], we define the spherical transforms of the matter overdensity based on Legendre polynomials $\mathcal{L}_\ell$:

$$\mathcal{A}^n_\ell(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3}(ik)^{-n}\mathcal{L}_\ell(\mu)\exp(i\mathbf{k}\cdot\mathbf{x})\delta(\mathbf{k}) \ . \tag{52}$$

For the non Gaussian term we have:

$$\tilde{\mathcal{A}}^2_0(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3}(ik)^{-2}\frac{1}{T(k)}\mathcal{L}_0(\mu)\exp(i\mathbf{k}\cdot\mathbf{x})\delta(\mathbf{k}) \ . \tag{53}$$

We can therefore write, in configuration space:

$$\delta_{g,obs}(\mathbf{x},z) = D(z)\left[\left(b(z) + \frac{1}{3}f(z)\right)\mathcal{A}^0_0(\mathbf{x}) + \frac{2}{3}f(z)\mathcal{A}^0_2(\mathbf{x}) + \alpha(z)f(z)\mathcal{A}^1_1(\mathbf{x}) - \Delta b(z)\tilde{\mathcal{A}}^2_0(\mathbf{x})\right] \ , \tag{54}$$

so that the 2-point correlator of the spherically transformed matter overdensity is:

$$\left\langle \mathcal{A}^{n_1}_{\ell_1}(\mathbf{x}_1,z_1)\,\mathcal{A}^{n_2}_{\ell_2}(\mathbf{x}_2,z_2)\right\rangle = D(z_1)D(z_2)\times(-1)^{\ell_2}\int\frac{d^3k}{(2\pi)^3}(ik)^{-(n_1+n_2)}\mathcal{L}_{\ell_1}(\hat{\mathbf{k}}\cdot\mathbf{n}_1)\mathcal{L}_{\ell_2}(\hat{\mathbf{k}}\cdot\mathbf{n}_2)e^{i\mathbf{k}\cdot\mathbf{s}}P(k) \ , \tag{55}$$

where s is the comoving separation between the two galaxies and $P(k)$ is the matter power spectrum.

Expanding $\mathcal{L}_\ell$ and $\exp(i\mathbf{k}\cdot\mathbf{s})$ in spherical harmonics and applying the Gaunt integral, we obtain:

$$\left\langle \mathcal{A}_{\ell_1}^{n_1}(\mathbf{x}_1, z_1)\, \mathcal{A}_{\ell_2}^{n_2}(\mathbf{x}_2, z_2)\right\rangle = D(z_1)D(z_2) \times \sum_L (-1)^{\ell_2} i^{L-n_1-n_2}(2L+1) \begin{pmatrix} \ell_1 & \ell_2 & L \\ 0 & 0 & 0 \end{pmatrix} \tag{56}$$

$$\times\, S_{\ell_1\ell_2 L}(\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_{12})\, \xi_L^{n_1+n_2}(s)\ , \tag{57}$$

for $|\ell_1 - \ell_2| \leq L \leq \ell_1 + \ell_2$. In this way, we have expanded the correlator in the tripolar basis functions [68]:

$$S_{\ell_1\ell_2 L}(\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_{12}) = \left[\frac{(4\pi)^3}{(2\ell_1+1)(2\ell_2+1)(2L+1)}\right]^{1/2} \sum_{m_1,m_2,M} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & M \end{pmatrix} \tag{58}$$

$$\times\, Y_{\ell_1 m_1}(\mathbf{n}_1)\, Y_{\ell_2 m_2}(\mathbf{n}_2)\, Y_{LM}(\mathbf{n}_{12})\ , \tag{59}$$

for $\ell_1 \leq m_1 \leq \ell_1, \ell_2 \leq m_2 \leq \ell_2$ and $L \leq M \leq L$ and with coefficients:

$$\xi_L^n(s) = \int \frac{dk}{2\pi^2} k^{2-n} j_L(sk) P(k)\ . \tag{60}$$

In the non-Gaussian case, the additional contribution is:

$$\xi_{L,ng}^n(s) = \int \frac{dk}{2\pi^2} \frac{k^{2-n}}{T(k)^m} j_L(sk) P(k)\ , \tag{61}$$

where $m = 2$ for $n = 4$ and $m = 1$ in the other cases.

The $k-$behaviour of each correlator is governed by the index n. The correlators $\langle \mathcal{A}_1^1, \mathcal{A}_1^1\rangle$,$\langle \tilde{\mathcal{A}}_0^2, \mathcal{A}_0^0\rangle$, $\langle \mathcal{A}_0^0, \tilde{\mathcal{A}}_0^2\rangle$ have the same $\propto k^{-2}$ behaviour. This means that in principle the contributions coming from the Doppler terms can mimic an effective non-Gaussianity [71, 72] in the galaxy 2-point correlators.

The tripolar decomposition of the observed two point galaxy correlation function is then:

$$\xi(s, \theta, \phi, z_1, z_2) = D(z_1)D(z_2) \sum_{\ell_1,\ell_2,L,n} B_n^{\ell_1\ell_2 L}(z_1, z_2) S_{\ell_1\ell_2 L}(\theta, \phi) \xi_L^n(s)\ . \tag{62}$$

The complete set of coefficients $B_n^{l_1,l_2,L}(z_1, z_2)$ for the Gaussian case can be found in literature, and for completeness we report them here [75, 83, 84]:

$$\begin{aligned}
B_0^{000} &= (b_1 + \frac{1}{3}f_1)(b_2 + \frac{1}{3}f_2) & B_2^{112} &= -\frac{\sqrt{30}}{3} f_1\alpha_1 f_2\alpha_2 \\
B_1^{011} &= \sqrt{3}(b_1 + \frac{1}{3}f_1) f_2\alpha_2 & B_1^{121} &= \frac{2\sqrt{30}}{15}\alpha_1 f_2 \\
B_0^{022} &= -\frac{2\sqrt{5}}{3}(b_1 + \frac{1}{3}f_1) f_2 & B_1^{213} &= -\frac{2\sqrt{105}}{15} f_1 f_2\alpha_2 \\
B_1^{101} &= -\sqrt{3} f_1\alpha_1(b_2 + \frac{1}{3}f_2) & B_0^{222} &= \frac{4\sqrt{70}}{63} f_1 f_2 \\
B_2^{110} &= -\frac{\sqrt{3}}{3} f_1\alpha_1 f_2\alpha_2 & B_1^{123} &= \frac{2\sqrt{105}}{15}\alpha_1 f_2 \\
B_1^{211} &= -\frac{2\sqrt{30}}{15} f_1 f_2\alpha_2 & B_0^{202} &= \frac{2\sqrt{5}}{3} f_1(b_2 + \frac{1}{3}f_2) \\
B_0^{220} &= \frac{4\sqrt{5}}{45} f_1 f_2 & B_0^{224} &= \frac{4\sqrt{70}}{35} f_1 f_2\ .
\end{aligned} \tag{63}$$

For the non-Gaussian contributions, we have:

$$\begin{aligned}
&\xi_{ng}(s,\theta,\phi,z_1,z_2) = \\
&D(z_1)D(z_2)\Bigg\{\Delta b(z_1)\Delta b(z_2)\xi^4_{0,ng}(s) + \left[\Delta b(z_1)\left(b(z_2)+\frac{f(z_2)}{3}\right)+\Delta b(z_2)\left(b(z_1)+\frac{f(z_1)}{3}\right)\right]\xi^2_{0.ng}(s) \\
&-\frac{2\sqrt{5}}{3}\left(S_{022}(\theta,\phi)\Delta b(z_1)f(z_2)+S_{202}(\theta,\phi)\Delta b(z_2)f(z_1)\right)\xi^2_{2,ng}(s) \\
&+\xi^3_{1,ng}(s)\sqrt{3}\left(-S_{011}(\theta,\phi)\Delta b(z_1)\alpha(z_2)f(z_2)+S_{101}(\theta,\phi)\Delta b(z_2)\alpha(z_1)f(z_1)\right)\Bigg\}\,.
\end{aligned} \tag{64}$$

We see that the coefficients ξ^l_m depend only on the linear separation between the two sources s and on the cosmological parameters; the full 3D geometry of the system is encoded in the tripolar spherical harmonics basis $S_{\ell_1\ell_2 L}(\theta,\phi)$, while the terms that describe deviations from Gaussianity, the growth of structure, and all the projection effects are included in the coefficients $B_{\ell_1\ell_2 L}(z_1,z_2)$.

A Observed multipoles of the galaxy 2PCF in redshift space

Since the three unit vectors that define the position of the two galaxies w.r.t. the observer are on the same plane, the redshift space correlation function depends only on the shape and size of the triangle formed by the observer and the two sources. This triangle is uniquely determined by specifying any three of its variables, such as three sides, three angles, two angles and one side, etc. In this work we express our calculations using the linear separation between the two galaxies, $\{s,\theta\}$ that is the half of the opening angle, and ϕ, the angle between the bisector of the opening angle and s. The $\hat{z}$ axis is aligned with the bisector of θ that defines the $\phi=0$ direction. The system of coordinates is shown in Fig. 2.

In order to fully describe the observable with only three degrees of freedom (s,ϕ,θ), we have to remove the redshift dependence (z_1,z_2) in Equation (62). We can write the redshift dependence of the 2PCF inverting the following relations [69]:

$$r_1 = \frac{\sin(\phi+\theta)}{\sin(2\theta)}s, \quad r_2 = \frac{\sin(\phi-\theta)}{\sin(2\theta)}s\,, \tag{65}$$

in order to express $z_1(r_1(s,\theta,\phi))$ and $z_2(r_2(s,\theta,\phi))$. At a given linear separation s and orientation angle ϕ, galaxy pairs might be observed with different opening angles θ. For instance, pairs that are at a certain distance from the observer are seen under a smaller θ w.r.t. pairs that are closer. The resulting observed 2PCF is then averaged over all pairs with fixed s and ϕ, as in [85]:

$$\xi_{\rm obs}(s,\phi) = \frac{\int_{\theta_{min}}^{\theta_{max}} d\theta N(s,\theta,\phi)\xi(s,\theta,\phi)}{\int_{\theta_{min}}^{\theta_{max}} d\theta N(s,\theta,\phi)}\,, \tag{66}$$

where $N(s,\theta,\phi)$ is a weight function that accounts for the distribution of pairs (see Sect. **??** for more details).

Since we perform observations on the spherical sky, galaxy clustering analyses are often carried out by expanding the 2PCF in Legendre polynomials [67] as each multipole contains different physical information about the underlying clustering, so it is convenient to focus on them for our modeling. We can expand the 2PCF with Legendre polynomials as [67]:

$$\xi_L(s) = (2L+1)\int_0^{\pi/2}\xi_{\rm obs}(s,\phi)\sin(\phi)\mathcal{L}_L(\cos(\phi))d\phi\,, \tag{67}$$

where we used the fact that for single tracer analyses the 2PCF is symmetric w.r.t the transformation

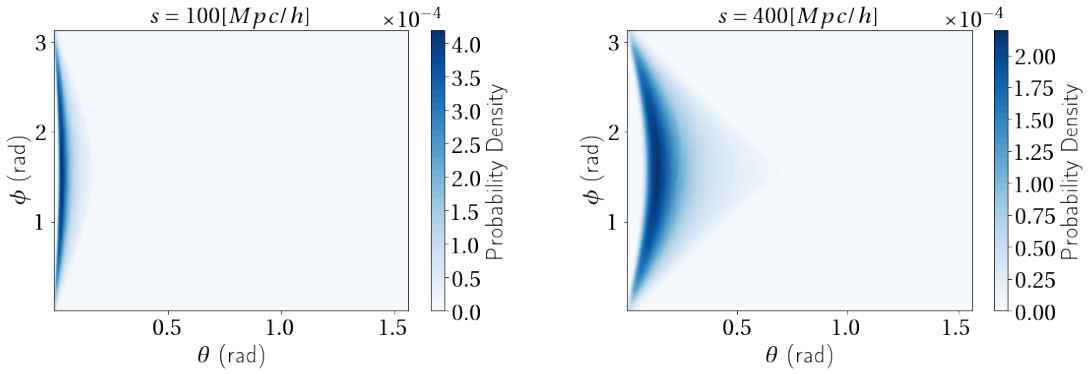

Figure 3: $N(s, \theta, \phi)$ distributions for fixed values of linear separations s. The distribution is normalized as $\int N(s, \theta, \phi) d\theta d\phi = 1$ for each value of s.

$\phi \to \pi - \phi$.

The integral in Equation (66) over the opening angle θ allows to account for the full 3D structure of the problem. However, we have to consider the full distributions of galaxies in $\{s, \theta, \phi\}$, which in general can be not uniform (see e.g., [79]). This needs to be considered by applying a weight function, $\xi_{\rm obs}(s, \theta, \phi) = N(s, \theta, \phi)\xi(s, \theta, \phi)$. In a real data analysis, those distributions will be extracted from the observed catalog. In a theoretical forecast as the one we present here, the weightings depend on the probability to find a galaxy in a particular position, which is connected to the specific redshift distribution of the survey, and we need to account for the survey geometry. The finiteness of the bin width translates into a limited range of θ under which the different pairs can be observed in that bin, for each linear separation s and orientation ϕ. In order to predict $N(s, \theta, \phi)$, we simulate mock catalogs of galaxies by populating the survey volume with particles distributed according to the specific $N(z)$ of that survey and tabulate the probability distribution as function of the coordinates $\{s,\theta,\phi\}$, seen by the observer. Of course this is an idealized scenario, because in practical surveys the observed region of the sky is discontinuous, the mask is non trivial and the angular selection function is not uniform.

In this work we are interested in theoretical forecasts, so in order to get physical insights we limit ourselves to the idealized scenario where the probability distributions of a triangle configuration is uniquely specified, given $N(z)$ and $f_{\rm sky}$. As an example mock survey we use a SPHEREx-like survey ($f_{sky} = 0.75$), with two different binning schemes:

-*Thick bins* scenario: $0 < z < 4.6$, bins with $0 < z < 1$ have $\Delta z = 0.2$, while bins with $1 < z < 4.6$ have $\Delta z = 0.6$ (this is the standard binning scheme for the SPHEREx survey [18])

-*Thin bins* scenario: $0 < z < 2$, where all bins have equal width $\Delta z = 0.2$.

We use the SPHEREx galaxy sample with $\sigma(z)/(1 + z) \leq 0.03 - 0.1$, in order to minimize the shot noise contribution; we do not account for the error introduced by the redshift uncertainty, and we checked that they do not propagate differently to different modelings. In Fig. 3 we show the probability distribution $N(s, \theta, \phi)$ obtained with the procedure outlined above for two fixed values of the linear separations. We see that the majority of the galaxies are seen under a small opening angle even for large separations; this fact will be important to understand the results of Section III B.

If we are expanding the 2PCF in Legendre polynomials, they must be an orthonormal basis independent from the chosen coordinate system; this condition is broken in the 3D wide angle case, as the triangle formed by the observer and the two sources closes (and so the galaxy pair is actually observable) only when ϕ is larger than θ (see also [84]). In order to deal with this issue in the full 3D modeling, one can either follow [84] and modify the expansion, or limit the maximum linear separation in order to allow the integral on ϕ to be performed over the entire $\mu = \{-1, 1\}$ range. Following the latter, this limitation requires imposing an upper limit on the maximum linear separation $s_{{\rm max},i}$ observed in the

i-th redshift bin. We checked that, for every bin, correlations with separation larger than $s_{\max}$ are noise dominated and so they will not add any further significant information to the analysis. Therefore, we limit the linear separations we consider to values that satisfy this condition.

Moreover, in the wide angle 3D model we can drop the approximation that the galaxies are located at the same (effective) redshift $z_{\rm eff}$. In fact, using two angles, i.e., the orientation ϕ and the opening angle θ (as in Fig. 2) allows us to account properly for the 3D geometry of the system and to model correctly the position of each galaxy inside the redshift bin and the distributions of the sources with respect to the observer. For this reason, in the flat sky scenario, the more pairs we are "mis-measuring" (wrongly identifying their redshift with $z_{\rm eff}$), the worse we are modeling the correlation function.

We perform a Fisher forecast analysis [86, 87] to investigate how wide angle 3D corrections affect measurements of the growth index and Primordial non-Gaussianity, for some example future galaxy surveys. We then study how incorrectly modeling the 2PCF with the (standard) flat sky approximation will cause a mis-estimation of the inferred best fit values of the growth index and $f_{\rm NL}$. In what follows, we assume perfect knowledge of the cosmological parameters at their fiducial values obtained by the Planck Collaboration [1], as we are interested not in the exact forecasting of cosmological parameters, but on the impact of using the flat sky approximation. We, instead, leave the bias parameters in each bin free to vary. The Fisher matrix element is, for each redshift bin:

$$F_{a,b} = \sum_{i,j} \frac{\partial \tilde{\xi}(s_i)}{\partial \theta_a} \mathrm{Cov}^{-1}\left(s_i, s_j\right) \frac{\partial \tilde{\xi}(s_j)}{\partial \theta_b} \,, \tag{68}$$

where $\tilde{\xi}(s) = \{\xi_0(s_i), \xi_0(s_{i+1})...\xi_0(s_N), \xi_2(s_i), \xi_2(s_{i+1})...\xi_2(s_N)\}$ is the vector containing the theoretical monopole and quadrupole of the 2PCF.

Since we use only linear scales in the analysis, we adopt a conservative approach similar to the one proposed by [47] to determine the minimum linear separation at each redshift. We fix the maximum scale $k = 0.15h/\mathrm{Mpc}$ at $z = 0$ and we keep the variance of the matter perturbations constant at each redshift to find $k_{NL}(z)$ imposing the condition that at $z = 0$, $\sigma^2(z) = 0.36$. We fixed and absolute maximum value of $k_{NL} = 0.3$, and checked that a less conservative approach does not significantly change the results of this work.

Using the flat sky approximation in the theoretical model will lead to a wrong estimation for the parameters that best fit the data. We can model the shift in this best fit value of the parameter Θ as [88, 89]:

$$\Delta\Theta = \left(\sum_m F_m\right)^{-1} \left[\sum_{m,i,j} \nabla_\Theta \tilde{\xi}^{pp}(s_i)_m \left(\mathrm{Cov}^{pp}(s_i, s_j)^{-1}\right)_m \left(\tilde{\xi}^{wa}(s_j)_m - \tilde{\xi}^{pp}(s_j)_m\right)\right] \,, \tag{69}$$

where the sum is performed over the m redshift bins; this is valid since we assume that measurements in different non-overlapping redshift bins are not correlated; this is valid in the case where integrated projection effects are neglected (see e.g., [83]).

We estimate the theoretical covariance for the multipoles of the correlation function for each redshift bin following the procedure outlined in [90, 91]:

$$\begin{aligned}\mathrm{Cov}_{\ell,\ell',i}(s,s') = \frac{i^{\ell-\ell'}}{V_i}\Bigg[&\frac{(2\ell+1)^2}{2\pi\bar{n}_i{}^2 s^2}\delta(s-s')\delta_{\ell\ell'} + D(z_i)^2 \sum_L \begin{pmatrix} \ell & \ell' & L \\ 0 & 0 & 0 \end{pmatrix}^2 \times \\ &\times \left(\frac{1}{\bar{n}_i}\int c_L(k,z_i)\mathcal{G}_{\ell\ell'}(k,s,s')dk + D(z_i)^2 \int \tilde{c}_L(k,z_i)\mathcal{D}_{\ell\ell'}(k,s,s')dk\right)\Bigg] \,,\end{aligned} \tag{70}$$

where $\bar{n}_i$ is the average galaxy number density in the i-th redshift bin centered at redshift z_i of volume V_i and:

$$\begin{aligned}\mathcal{G}_{\ell\ell'}(s,s') &= \frac{2(2\ell+1)(2\ell'+1)}{\pi^2}k^2P(k)j_\ell(ks)j_{\ell'}(ks')\ ,\\ \mathcal{D}_{\ell\ell'}(s,s') &= \frac{(2\ell+1)(2\ell'+1)}{\pi^2}k^2P^2(k)j_\ell(ks)j_{\ell'}(ks')\ ,\end{aligned} \tag{71}$$

and the coefficients c_L (where $L = 0, 2, 4$) are obtained by expanding the galaxy power spectrum in redshift space in Legendre polynomials $P(k,\mu) = c_0P(k) + c_2P(k)\mathcal{L}_2(\mu) + c_4P(k)\mathcal{L}_4(\mu)$, with [67]:

$$\begin{aligned}c_0 &= b^2 + \frac{2}{3}bf + \frac{f^2}{5}\ ,\\ c_2 &= \frac{4}{3}bf + \frac{4}{7}f^2\ ,\\ c_4 &= \frac{8}{35}f^2\ ,\end{aligned} \tag{72}$$

where the redshift and k dependencies are omitted for brevity. The coefficients $\tilde{c}_L$ obtained for the expansion of $(P(k,\mu)/P(k))^2$ in Legendre polynomials (i.e., $(c_0 + c_2\mathcal{L}_2(\mu) + c_4\mathcal{L}_4(\mu))^2 = \sum_L \tilde{c}_L\mathcal{L}(\mu)$ are:

$$\begin{aligned}\tilde{c}_0 &= c_0^2 + \frac{c_2^2}{5} + \frac{c_4^2}{9},\\ \tilde{c}_2 &= \frac{2}{7}c_2(7c_0 + c_2) + \frac{4}{7}c_2c_4 + \frac{100}{693}c_4^2,\\ \tilde{c}_4 &= \frac{18}{35}c_2^2 + 2c_0c_4 + \frac{40}{77}c_2c_4 + \frac{162}{1001}c_4^2\ .\end{aligned} \tag{73}$$

As we can see from Equation (70), theoretical errors are calculated assuming deviations from Gaussianity only at the power spectrum level (i.e., in the scale dependent bias) neglecting contributions coming from the connected 4-point correlation function.

Since the expression for the theoretical covariance is given in the flat sky approximation the constraints obtained for the wide angle model are only an estimation as for a proper exact calculations one need to substitute in Equation 68 the wide angle expression of the covariance. On the other hand the results obtained for the best fit shift are exact as the covariance in Equation (69) is the one calculated assuming a wrong modelling i.e., the flat sky approximation.

Following all the definitions and considerations above, we now define the different models we will test for the analyses of galaxy clustering. After having introduced the two extremes of full 3D wide angle and the Kaiser, flat-sky plane-parallel ones, we here define some intermediate models that promise better accuracy but with considerably simpler expressions, and test how accurately they can retrieve some cosmological parameters. For simplicity, in the following we will refer to the "flat sky" model for the case where we use the flat sky, the plane parallel and the effective redshift (i.e., all the sources inside a redshift bin are evaluated at the same redshift) approximations.

In Fig. 4 we plot the difference between the wide angle 3D model and the flat sky one as a function of the width of the redshift bin, for three fixed linear separations; we see that the difference between the two models is larger for thick bins, where flattening all the galaxies to an average redshift is less accurate, as expected.

One can try to improve the modeling of the 2PCF by dropping some of the approximations explained above, still without using the full wide angle 3D formalism, in order to minimize the theoretical systematics that will lead to biased results in the final measurements.

One can argue that fixing the radial position of the sources inside a redshift bin to some $z_{\rm eff}$ will lead us to a loss of information on the radial clustering of galaxies, especially if the bin is very thick. Such a

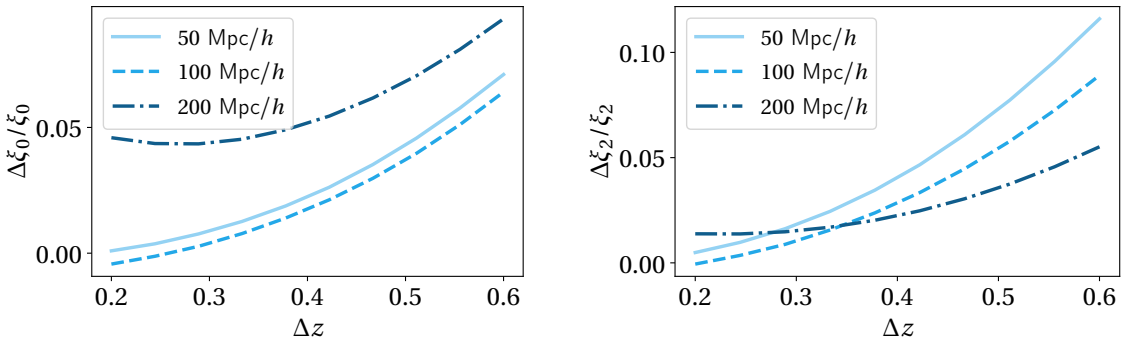

Figure 4: Difference between the wide angle 3D model and the flat sky approximation as a function of the width if the redshift bin for three linear separation for the monopole (left) and the quadrupole (right).

compression of the full 3D information is in principle unjustified as quantities such as the bias, the growth factor evolve within the bin range, and the clustering along the radial directions can be non negligible. In fact, results from previous sections show that this can be one of the main sources of inaccuracy.

To overcome this issue, one can average the flat sky model over each redshift bin and weighting for the redshift distribution of the sources. The observed correlation function in this case reads:

$$\xi_{\rm obs}(s,\phi) = \frac{\int_{z_{\rm min}}^{z_{\rm max}} N(z)\xi_{pp}(s,\phi,z)dz}{\int_{z_{\rm min}}^{z_{\rm max}} N(z)dz} \,. \tag{74}$$

However, this procedure allows us to account for all the equal time correlations (i.e., correlations of sources at the same redshift) in a specific redshift bin, but the flatness of the sky, by definition, prevents us to model correctly the unequal time correlations that are present in the wide angle 3D model. Such correlations have been recently modelled by [58, 59], where it is showed that there can be non negligible corrections to the equal time galaxy power spectrum for radial correlations.

In Fig. 5 we look again at the comparison between the full 3D wide angle case and the flat sky one, but this time for the full shape of the correlation function, and including the radial averaging case. We can notice already an improvement with respect to the flat sky case, especially for thick bins, again confirming expectations. The Figure shows the monopole and quadrupole of the galaxy 2PCF in redshift space for the full wide angle 3D model (dashed lines), the flat sky averaged one (Equation (74)) (continuous line) and the flat sky modeling (dotted lines) for the mock surveys used in this work: thick bin scenario (upper panels), thin bin scenario (lower panels).

B Radial 3D model

The next model we investigate, in the context of a simplified version of the full 3D wide angle that retains as much accuracy as possible, is what we call the "radial 3D" model.

As shown in Fig. 3, in practice galaxies are too far away w.r.t. the observer to be seen under large aperture angle θ, so one can study what happens if we take the limit $\theta = 0$ in the tripolar basis coefficients $S_{\ell_1,\ell_2,L}$ in Equation (62). At the same time, however, we keep the dependence on θ in the redshift dependent quantities (i.e., in the coefficients $B_n^{\ell_1,\ell_2,\ell_L}$) in order to maintain the 3-dimensionality of the problem. In this scenario the 2PCF in redshift space can be expressed with a simpler formula:

$$\xi(s,\theta,\phi) = D_1 D_2 \sum_n F_{L,n}(s,\theta,\phi)\xi_L^n(s)\mathcal{L}_L(\cos\phi) \,. \tag{75}$$

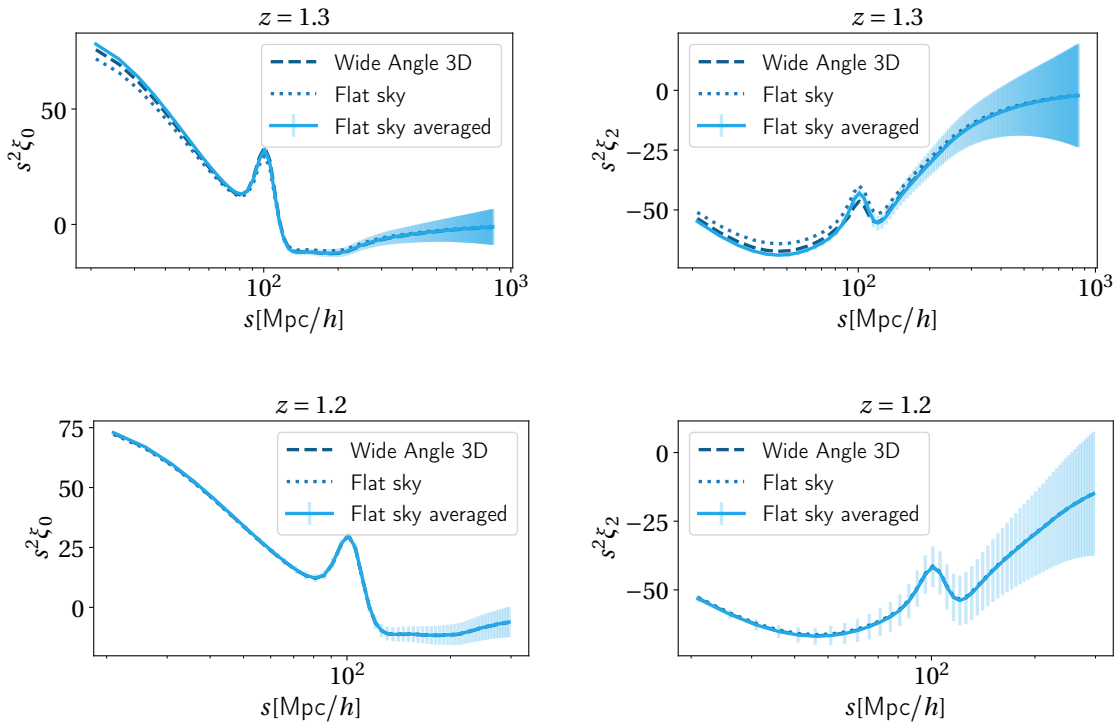

Figure 5: Monopole and quadrupole of the galaxy 2PCF in redshift space for different models in the thick bins scenario (upper panels) and in the thin bin scenario (lower panels). Continuous lines: flat sky averaged model. Dashed lines: wide angle 3D model. Dotted lines: flat sky model

The $F_{L,n}$ coefficients are:

$$
\begin{aligned}
F_{00} &= \mathcal{F}_1\mathcal{F}_2 + \frac{4}{45} f_1 f_2 \,, \\
F_{20} &= -\frac{2}{3}\left(\mathcal{F}_1 f_2 + \mathcal{F}_2 f_1\right) - \frac{8}{63} f_1 f_2 \,, \\
F_{40} &= \frac{8}{35} f_1 f_2 \,, \\
F_{11} &= \alpha_1 f_1 \mathcal{F}_2 - \alpha_2 f_2 \mathcal{F}_1 + \frac{4}{15} f_1 f_2 \left(\alpha_1 - \alpha_2\right) \,, \\
F_{31} &= \frac{2}{5} f_1 f_2 (\alpha_2 - \alpha_1) \,, \\
F_{02} &= \frac{\alpha_1 \alpha_2 f_1 f_2}{3} \,, \\
F_{22} &= -\frac{2\alpha_1 \alpha_2 f_1 f_2}{3} \,,
\end{aligned}
\tag{76}
$$

where $\mathcal{F}_i = b_i + 1/3 f_i$ and the subscript $i = 1, 2$ identifies different galaxies, and the θ dependence of the observable is only encoded in the redshift dependent quantities $\mathcal{F}_i$, f_i,b_i and α_i since the redshift of each source can be rewritten as a function of $\{s, \theta, \phi\}$.

This expression, while obviously more complicated than the Kaiser formula, is still way simpler that then full 3D including wide angle terms.

For the non Gaussian terms we have:

$$\begin{aligned}\tilde{F}_{02} &= \Delta b_1 \mathcal{F}_2 + \Delta b_2 \mathcal{F}_1 \\ \tilde{F}_{04} &= \Delta b_1 \Delta b_2 \\ \tilde{F}_{22} &= -\frac{2}{3}(\Delta b_1 f_2 + \Delta b_2 f_1) \\ \tilde{F}_{13} &= \alpha_2 f_2 \Delta b_1 - \alpha_1 f_1 \Delta b_2 \ ,\end{aligned} \tag{77}$$

where in this case the ξ_L^n functions in Equation (75) must be replaced by $\xi_{L,ng}^n$. In this expression, all the quantities that depend on redshift ($\{f, b, \alpha\}$) will carry a dependence on the three variables $\{s, \theta, \phi\}$, and therefore require the use of the three degrees of freedom that are necessary to define the triangle formed by the observer and the galaxy pair.

This simplified model is still able to capture the full 3-dimensionality of the problem, as the peculiar position of each sources inside the redshift bin is modelled correctly. In fact, once we identify a box of cosmic volume defined by a given redshift range, since in the radial 3D model we preserve the three degrees of freedom of the full wide angle 3D formalism ($\{s, \theta, \phi\}$), we can correctly identify every triangle configuration formed by the observer and the sources in the box, while being able to weight every configuration for the probability of being observed. In this way, we can properly account for unequal time correlators as the full radial information is not lost as it is in the flat sky model. This allows the radial 3D model to be more accurate than the flat sky models while remaining simpler than the full wide angle 3D one also in terms of computational cost. The radial 3D model is different w.r.t. the plane parallel (distant observer) one as in the latter case the aperture angle is not considered at all in the modeling. In the radial 3D model, instead, the line of sights are still treated as non parallel to each other and the θ coordinate is fundamental in order to correctly identify every triangle configuration, as explained above.

C *Impact of different approximations*

In this work we present an estimation of the impact of taking some approximations used in galaxy clustering analyses instead of using a fully correct modeling. In particular, the results of our work quantify the impact of using the plane parallel and flat sky approximations in measurements of the growth of structure and Primordial non-Gaussianity.

To interpret the results correctly, we must understand that the difference between the standard, so-called Kaiser flat sky model and the full, wide angle 3D one is twofold. In the wide angle 3D formalism there are terms proportional to the opening angle θ (which are not present in the standard modeling), which increase with θ and encode the dependence on the aperture angle between the correlated sources. These terms are present even if all the redshift dependent quantities in the 2PCF are evaluated at a fixed mean $z_{\rm eff}$ in each redshift bin (see [68]).

In Fig. 6 we show the first of our main results, which is the predicted best fit and limits for different modelings. We show this for the γ and $f_{\rm NL}$ parameters, comparing results when using the full 3D wide angle model or different approximations, namely the flat sky modeling (averaged and not) and a radial 3D model. In dashed grey and star is the forecast for the standard Kaiser modeling, using flat sky and one single effective redshift, while in solid red and circles we show the full 3D wide angle modeling. In dot-dashed blue and triangles we show the radial 3D model, and in dotted black and squares we show the flat sky averaged one. The left panel shows results for the SPHEREx-like thick bins scenario, while the right panel shows results for the thin bins configuration.

We can see that, while constraints do not dramatically change for different models (the FoM changes by less than 10%), the expected best fits change considerably for the thick bins. For the growth rate parameter, the expected high precision in future measurements and the fact that its redshift evolution is not properly captured when assuming all sources are at the same effective redshift, makes it that the

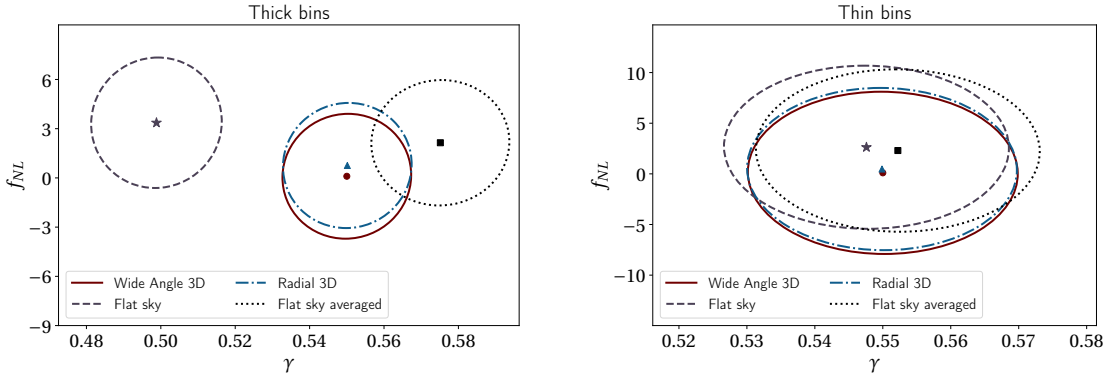

Figure 6: Best fits and 1σ limits of γ and $f_{\rm NL}$ in the case of the wide angle 3D model (red continuous line and circles), flat sky averaged (black dotted line and squares), flat sky (grey dashed line and stars) and the radial 3D model (blue dot-dashed line and triangles). Left panel is for the thick bins configuration, the right panel shows the thin bin scenario.

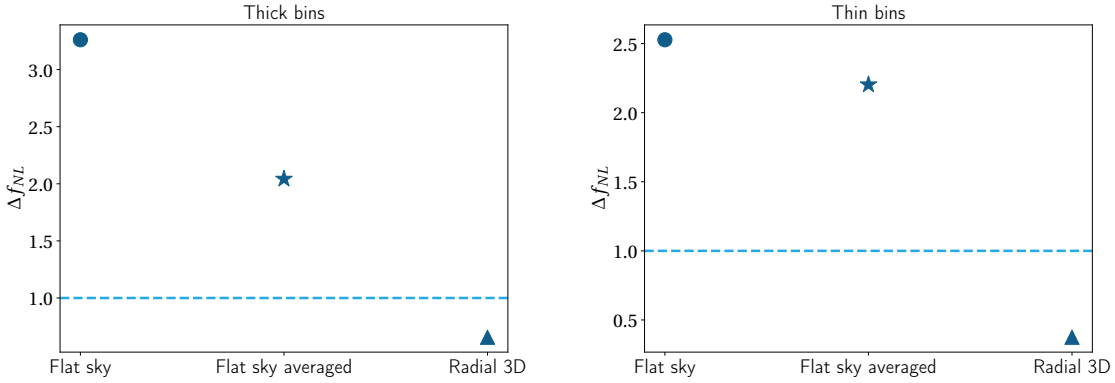

Figure 7: Shift in the best fit for $f_{\rm NL}$, for the different approximation models considered in this work. The blue dashed line indicates $f_{\rm NL} = 1$, a target value for many future galaxy surveys.

best fit is considerably affected in the thick bins case. As expected, the effect becomes smaller when we consider thinner redshift bins. The flat sky model, even when averaged over z within the bin, will get a $\sim 1\sigma$ shift in the best fit value of γ in the case of the thick bin scenario.

PNG, on the other hand, is less affected by the radial terms; however, its effect is larger at larger scales, and it is degenerate with geometry effects. For this reason, there is a shift also for the non-Gaussianity parameter. However, for the same reason, paired with the fact that the error bars are larger at large scales, the shift is approximately within one sigma for the thick bins case, and below that for the thin bins case. Such a shift, for future surveys, could be more relevant, especially when adding the bispectrum.

In Figure 7 we focus on the shift of the best fit for $f_{\rm NL}$. The plot shows the shift, for both cases and the various models considered here, induced by incorrect modeling. In dashed blue we show $f_{\rm NL} = 1$, as this is the target value for many future surveys, given that it is generally the discriminant between single- and multi- field inflationary models [92]. As we can see, while technically the shift on the inferred PNG parameter is below the $1-\sigma$-level, for theoretical reasons, and in anticipation of more precise measurements, it will still be very relevant for the flat sky models. The Radial 3D model, on the other hand, will only marginally bias the best fit.

The fact that the Radial 3D model retains almost the same level of accuracy of the full wide angle 3D one is due to two main factors. The magnitude of the θ-terms in the tripolar spherical harmonics coefficients increase at very large scales, where the error budget is large, and therefore the final impact

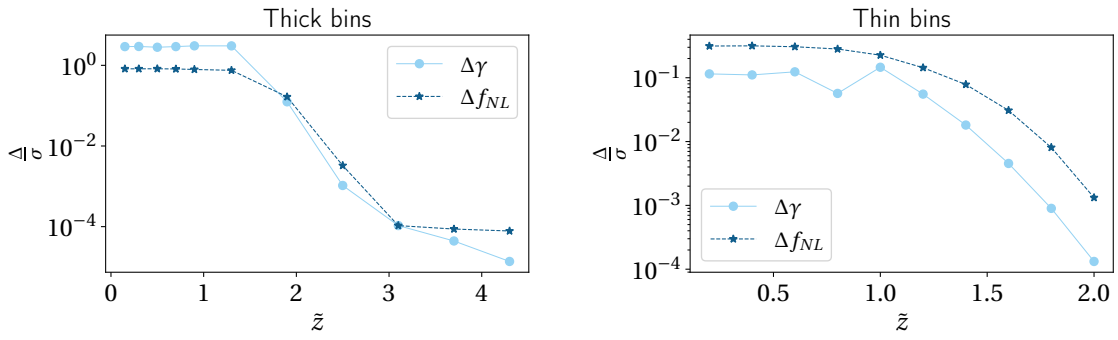

Figure 8: Best fit shift of γ and $f_{\rm NL}$ obtained if we model with the flat sky limit only bins above $\bar{z}$.

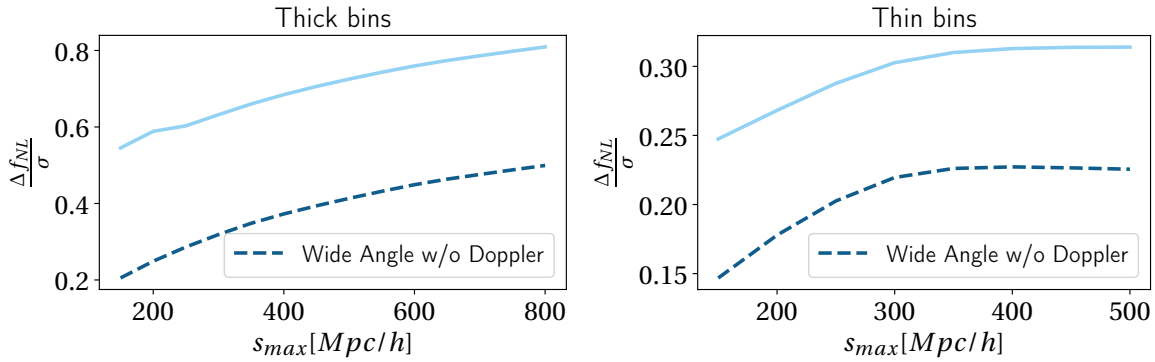

Figure 9: Best fit shift of $f_{\rm NL}$ as a function of the maximum scale included in the analysis. The dashed line shows the shift induced by neglecting Doppler terms.

on parameter estimation is reduced. As shown in Fig. 3, for a realistic survey scenario, galaxy pairs are far away from us so that there is little statistics for widely separated pairs.

It is important to recognize, however, that this does not imply that it will be correct to use the plane parallel (or distant observer) approximation for modeling the observed position of each galaxy pair. The correct approach to describe the system comprising the observer and the two galaxies always requires three degrees of freedom (in this work they are taken to be s, θ, ϕ), rather than the two degrees offered by the plane parallel case, in order to account for the full 3-dimensionality of the problem.

Here, as an additional test, we look at the importance of wide angle 3D corrections as a function of the redshift and scale. To do so, and to give a general guidance for when it is safe to use the standard Kaiser approach, in Fig. 8 we plot the best fit shifts as a function of the maximum redshift until which we use the full 3D wide angle modeling, which we call $\bar{z}$. We see that the shift becomes negligible (and therefore one can use the standard analysis after $\bar{z}$) right after $z \sim 1$, due to the combination of the fact that a large part of the shift due to radial effects comes from the very thick bin at $z = 1.3$, and the precision in the measurement decreases as we go to higher redshifts. In the thin bin scenario while the magnitude shift decreases with redshift as the precision in the measurements decreases in general it remains negligible for all redshift as expected.

In Fig. 9 we show the best fit shift as function of the maximum scale included in the analysis. As PNG contributions from the scale dependent bias are maximized at large scales due to their k^{-2} behaviour, the shift in $f_{\rm NL}$ grows if we analyze very large scales. The shift is due to two contributions: radial effects due to the poor modeling of the redshift evolution achieved by the flat sky approximation and the effect of Doppler terms that introduce an 'effective PNG' at large scales due to their k^{-2} dependence. Thus, in Fig. 10 we show the best fit values and 1σ constraints for the full wide angle 3D model (solid red), the wide angle 3D model without Doppler terms (blue dot-dashed) and flat sky averaged model (grey

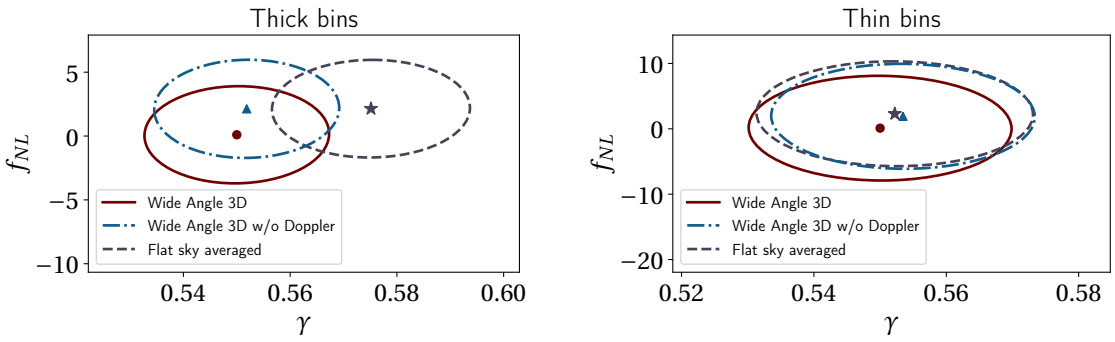

Figure 10: 1σ constraints and best fit values of γ and $f_{\rm NL}$ in the case of the wide angle 3D model (red continuous line), wide angle 3D model but neglecting Doppler terms (blue dot-dashed line) and flat sky averaged (grey dashed line).

dashed). We see that in general the best fit shift induced by Doppler terms remains always $< 1\sigma$, for the standard cases we discussed here, and therefore we neglect them for the purposes of this work. However, as mentioned above, the Doppler terms can become relevant, especially for some galaxy populations and in the multi-tracer case (see e.g., [84]); we will investigate in detail these situations and the impact of Doppler terms in our hybrid modeling in future work.

D nDGP

As a further application of our investigation, we look at how cosmological tests can be biased when using the flat sky approximation; in particular, we focus on tests of DGP [93], an alternative cosmological model. In the original DGP model, the Universe is described by a 4D brane embedded in a 5D Minkowski spacetime. Matter is confined in the brane, while gravity propagates in the extra dimension above the cross-over scale r_c, which is the fundamental extra parameter of this model. General relativity is recovered by taking the cross-over scale to be much larger than the current Hubble scale, i.e., for $H_0 r_c \gg 1$. Constraints on this parameter are traditionally expressed in terms of the dimensionless quantity:

$$\Omega_{\rm rc} \equiv \frac{1}{4 r_c^2 H_0^2} \,. \tag{78}$$

In the following we consider the "normal branch" of the DGP model (nDGP), where the background expansion history reproduces the ΛCDM one in a spatially-flat universe [94] , since the so-called "self accelerating branch" which do not require the presence of a dark energy component to explain the expansion of the universe has been found to be unstable [95, 96]. This is obtained by considering a dark energy component whose background dynamics exactly compensates the modified background evolution of the nDGP model so that deviations from ΛCDM take place only in the late-time evolution of perturbations. Constraints on the value of $\Omega_{\rm rc}$ has been recently investigated for example in [97–99] and in [100], where they obtained an upper bound $\Omega_{\rm rc} \lesssim 0.2$, analysing BOSS data.

At linear order in perturbation theory, we can write the differential equation for the growth factor $D(a)$ in the nDGP model as [94, 97, 100]:

$$\frac{d^2 D(a)}{d \ln a^2} + \left(2 + \frac{d \ln H}{d \ln a}\right) \frac{d D(a)}{d \ln a} - \frac{3}{2} \nu(a) \Omega_{{\rm m},a}(a) D(a) = 0 \,. \tag{79}$$

Assuming a ΛCDM background expansion rate, the difference with the ΛCDM case is captured by the

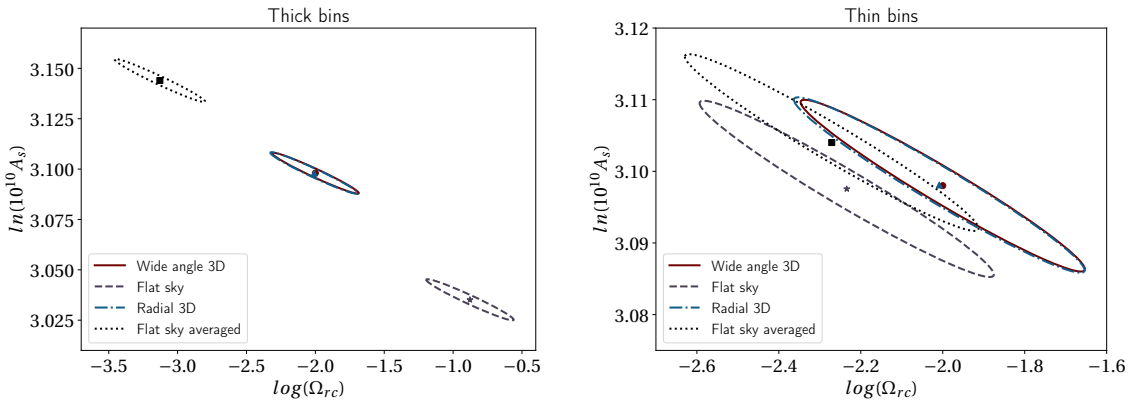

Figure 11: 1σ constraints and best fit values of $\log \Omega_{\rm rc}$ and $\ln\left(10^{10}A_s\right)$ and in the case of the wide angle 3D model (red continuous line), flat sky averaged (black dotted line), flat sky (grey dashed line) and the radial 3D model (blue dot-dashed line). The 3D radial and the full wide angle 3D cases are basically overlapping in both panels.

function $\nu(a)$ in the last term, defined as:

$$\nu(a) = 1 + \frac{1}{3\beta(a)}\,, \qquad \beta(a) \equiv 1 + \frac{H(a)}{H_0}\frac{1}{\sqrt{\Omega_{\rm rc}}}\left(1 + \frac{aH'(a)}{3H(a)}\right)\,. \tag{80}$$

which modifies the strength of the gravitational interaction. In the following we assume that $\nu \to 1$ (i.e., GR is recovered) at early times.

We perform a similar Fisher analysis of what done in previous sections and check whether the inferred value of the Ω_{rc} parameter is affected when taking simplifying assumptions in the modeling of the 2PCF. As before, we keep the cosmological parameters fixed to their Planck 2018 [1] values, but we let the primordial amplitude A_s vary, since it is degenerate with Ω_{rc}, as pointed out by [100], using however the Planck prior to mitigate the degeneracy with the bias parameter. Results are showed in Fig. 11, where we show the confidence ellipses and the best fit values of the parameters when using the wide angle 3D model, the radial 3D one, the flat sky approximation and the flat sky averaged model.

As for the ΛCDM case, the results are more relevant for the thick bin case than for the thin bins one. When the bins are large, we see that fixing all the sources at the same redshift will induce a relevant shift in the parameters ($\sim 3.5\sigma$ for $\log \Omega_{rc}$ and $\sim 6\sigma$ for $\ln\left(10^{10}A_s\right)$), while the Radial 3D model assures a similar level of accuracy that of the full wide angle 3D one.

IV 3D galaxy clustering: Fourier space

A *Unequal time correlators in redshift space*

As we have seen in the previous section, the radial information is important in order to retain the intrinsic 3-dimensionality of the system formed by the observer and the galaxies and in order to produce accurate and unbiased theoretical predictions. Another way to tackle the importance of the radial part of the galaxy correlators is to investigate their impact when predicting observables in Fourier space i.e., the galaxy power spectrum.

Recently, a new formalism was developed to account for this important but often overlooked aspect of galaxy correlations: sources being correlated are generally not located at the same redshift, introducing what can be called an "unequal-time" effect. In other words, unequal time contributions are meant to break the so-called "effective redshift" approximation which is often assumed in galaxy clustering analyses when calculating correlators in a given redshift bin. This correction can be incorporated in Fourier space through a Taylor expansion in the separation between tracers, allowing us to capture radial information that would otherwise be lost when galaxies within a redshift bin are approximated as lying at the same mean redshift.

In this section, we briefly review the basics of the unequal-time galaxy power spectrum; we refer the reader to [101–103] for more details. The "true" power spectrum is the ensemble average of two galaxy overdensity fields, in general evaluated at different times,

$$\langle \delta(\boldsymbol{k}, t)\delta^*(\boldsymbol{k}', t')\rangle = (2\pi)^3\delta_D^{(3)}(\boldsymbol{k}-\boldsymbol{k}')\mathcal{P}(\boldsymbol{k}, t, t')\,. \tag{81}$$

This object is in principle fully accessible only to a hypothetical meta-observer, i.e., one that could observe the full evolution of the overdensity field in 4D (in what we call the Hyperuranion space [101]). When we observe the sky, we only have access to information on the light-cone, thus we are limited to observables that can be constructed from measured quantities, e.g., angular positions and redshifts. While a meta-observer can access the full 4D information, for and the observer on the light-cone there is a degeneracy between cosmic time and the radial direction component of each Fourier mode (for a visualization of the setup, see [101, 102].

The most straightforward and correct way to maintain all the radial information provided by a survey consists in slicing the galaxy catalog in many thin redshift bins and compute the angular power spectra including all cross-correlations. While correct, this has the cost of having to compute a prohibitively large number of cross-bin correlations, with non-diagonal covariance for cross-correlations. Alternatively, the spherical Fourier-Bessel formalism naturally includes all radial modes, but these again come at extreme computational costs, on top of the problem of having the information spread over series of angular and Fourier modes (see e.g., [104–107]). While in general, at the level of the angular power spectrum or correlation function, radial correlations can be accounted for by including cross-bin terms in the analysis, it is interesting to investigate how these type of correlations can be modeled in Fourier space.

Recently, [101, 102] proposed a novel way to retain the unequal-time information in the power spectrum, while still maintaining a relatively low computational cost, both in Fourier space and with angular correlations. While the true dynamics of galaxy clustering happens in the 4D Hyperuranion space, where the meta-observer has access to 3 spatial dimensions at every instant in time, observations happen on the past lightcone of an observer. To build a proper observable, one then needs to connect the two, folding the 4D observable into a 3D one on the lightcone that includes the effect of radial/unequal in time modes. The procedure consists of two steps: first the galaxy overdensity field is projected on the sky to build the full unequal-time angular power spectrum, then the information on the unequalness-in-time between the two galaxies is folded into the radial Fourier modes.

We start by constructing the unequal-time galaxy angular power spectrum, including projection effects such as the standard Kaiser term, Doppler terms, and general relativistic corrections. To obtain an accurate picture of the importance of UT effects, it is necessary to include in the modeling all the effects that are expected to impact the large scales, such as Doppler term and general relativistic effects [108–116]. Moreover, we will include the effect of local-type Primordial non-Gaussianity, which induces a long-short mode coupling that affects halo clustering at late time, through a scale-dependent modification of the halo bias which also impacts the large scales [43, 45, 48, 117, 118]. In this section we focus on redshifts larger than $z \gtrsim 1$, where the global plane parallel approximation assumed so far is valid. In what follows ,we will refer to our assumption as "flat sky", in the sense that we assume the sky to be flattened, and not curved, as it was done in [103]. For a revisited treatment of the flat sky angular power spectrum see also [119]. Lower redshifts can be analyzed using one of the currently available approaches in angular [109, 110, 120, 121], configuration [83, 84, 111, 112], or spherical [104–107] spaces, or with an extension of this approach that includes sky curvature.

The unequal-time approach promises to provide a useful and more accurate way to analyze galaxy clustering, and as such there will be more progress needed for it to be applied to real data. A slightly different approach based on field level analyses is presented in [122], where unequal-time corrections are computed also in the case where one projects the fields individually before correlating them, and including cross-bin correlations. Here we focus on extending the theoretical modeling for the correlator approach.

In this chapter, we will drop the Newtonian limit and we consider General Relativistic corrections to the galaxy overdensity.[7] In General Relativity the definition of a perturbation is not unique. It involves a comparison between the physical (perturbed) spacetime and the smooth FLRW background spacetime. To perform such a comparison between physical quantities, living in the perturbed spacetime that actually describes our Universe, and unperturbed quantities in the background, it is necessary to introduce a mapping between the two manifolds. A gauge choice is the choice of this map. This mapping has not to be confused with a simple change of coordinates, which is nothing but a relabeling of points in the background spacetime and in the physical spacetime. A gauge choice, instead, changes the correspondence between points in the physical spacetime and in the background spacetime. There is not a unique choice of gauge. In what follows, we adopt the synchronous comoving gauge in which the line element reads

$$ds^2 = a(\tau)^2\left\{-d\tau^2 + \left[(1-2\mathcal{R})\delta_{ij} + 2\partial_i\partial_j E\right]dx^i dx^j\right\}, \tag{82}$$

where τ is the conformal time defined as $d\tau = dt/a(t)$, $\mathcal{R}$ and E are scalar perturbations, and $\mathcal{R}' = 0$ in ΛCDM, where the prime denotes derivative with respect to conformal time. The observed galaxy overdensity at redshift z and direction $\hat{\boldsymbol{n}}$ is [111, 123]

$$\begin{aligned}\delta_g(\hat{\boldsymbol{n}}, z) =& b\delta + \left[b_e - 1 - 2Q + \frac{1+z}{H}\frac{dH}{dz} - \frac{2(1-Q)(1+z)}{H\chi}\right](\partial_\parallel E' + E'') \\ &- \frac{1+z}{H}\partial_\parallel^2 E' - \frac{2}{\chi}(1-Q)(\chi\mathcal{R} + E'),\end{aligned} \tag{83}$$

where $\partial_\parallel = \hat{n}^i\partial_i$, $\chi(z)$ is the comoving distance, $b(z)$ is the linear bias,

$$b_e(z) = -(1+z)\frac{\partial \ln N_g(z)}{\partial z}$$

[7] In what follows we do not account for integrated general relativistic effects such as lensing and time delays. Including these type of terms in the unequal-time angular power spectrum presented in [101, 102] is the focus of another ongoing work.

is the evolution bias, and

$$Q(z) = \left.\frac{\partial \ln N_g}{\partial \ln L}\right|_{L=L_{\rm min}}$$

is the magnification bias, with N_g the number of galaxies with luminosity $L > L_{\rm min}$ at redshift z. The metric perturbations can be related to the matter density contrast in the synchronous gauge [111, 123] In ΛCDM+GR this leads to

$$\delta_g(\hat{\boldsymbol{n}}, z) = \left[b + f\partial_\parallel^2\nabla^{-2} - \frac{H^2}{(1+z)^2}\frac{3}{2}\Omega_m\mathcal{B}\nabla^{-2} + \frac{f\alpha}{\chi}\partial_\parallel\nabla^{-2}\right]\delta\,, \tag{84}$$

where $f = d\ln D/\ln a$ is the growth rate and the redshift dependent quantities α and $\mathcal{B}$ read

$$\alpha(b_e, Q, z) = -\frac{H(z)\chi(z)}{1+z}[b_e(z) - 1 - 2Q(z) + \frac{3}{2}\Omega_m(z) - \frac{2(1-Q(z))(1+z)}{H(z)\chi(z)}]\,, \tag{85}$$

$$\mathcal{B}(b_e, Q, z) = b_e(z)\left(1 - \frac{2f(z)}{3\Omega_m(z)}\right) + 1 + \frac{2f}{\Omega_m(z)} + \frac{3}{2}\Omega_m(z) - \frac{2(1-Q(z))(1+z)}{H(z)\chi(z)} - 4Q(z) - f(z)\,. \tag{86}$$

Let $\boldsymbol{\theta}$ be the transverse vector with respect to the line of sight $\hat{\boldsymbol{n}}$, so that $\boldsymbol{x} = \chi(\hat{\boldsymbol{n}} + \boldsymbol{\theta})$. The projected observed overdensity field is

$$\delta_g(\boldsymbol{\theta}, z[\bar{\chi}]) = \int d\chi W(\chi, \bar{\chi})\delta_g(\boldsymbol{x} = \chi(\hat{\boldsymbol{n}} + \boldsymbol{\theta}), z[\chi])\,, \tag{87}$$

where $W(\chi, \bar{\chi})$ is a window function centered at $\bar{\chi}$. For now we are keeping the explicit dependence on the window function, but in the following we will be interested in the case where it is infinitely thin, so that $\chi = \bar{\chi}$. In Fourier space,

$$\begin{aligned}\delta_g(\boldsymbol{\ell}, z[\bar{\chi}]) = \int d^2\boldsymbol{\theta} e^{-i\boldsymbol{\ell}\cdot\boldsymbol{\theta}} \int d\chi W(\chi, \bar{\chi}) \\ \times D(z[\chi])\left[b + f\partial_\parallel^2\nabla^{-2} - \frac{H^2}{(1+z)^2}\frac{3}{2}\Omega_m\mathcal{B}\nabla^{-2} + \frac{f\alpha}{\chi}\partial_\parallel\nabla^{-2}\right]\delta(\boldsymbol{x}, z[\chi])\,.\end{aligned} \tag{88}$$

The operators acting on the matter overdensity become

$$\begin{aligned}\partial_\parallel^a\nabla^{-2}\delta(\boldsymbol{x}, z[\chi]) &= \partial_\parallel^a\nabla^{-2}\int\frac{d^3\boldsymbol{k}}{(2\pi)^3}e^{i\chi\boldsymbol{k}\cdot(\hat{\boldsymbol{n}}+\vec{\theta})}\delta(\boldsymbol{k} = k_{\hat{n}}\hat{\boldsymbol{n}} + \boldsymbol{k}_\perp, z[\chi]) \\ &= i^{(a+2)}\int\frac{d^3\boldsymbol{k}}{(2\pi)^3}e^{i\chi\boldsymbol{k}\cdot(\hat{\boldsymbol{n}}+\vec{\theta})}\frac{(\boldsymbol{k}\cdot\hat{\boldsymbol{n}})^a}{k^2}\delta(\boldsymbol{k} = k_{\hat{n}}\hat{\boldsymbol{n}} + \boldsymbol{k}_\perp, z[\chi])\,,\end{aligned} \tag{89}$$

therefore, performing the angular integration, we get

$$\begin{aligned}\delta_g(\boldsymbol{\ell}, z[\bar{\chi}]) = \int d\chi W(\chi, \bar{\chi})\int\frac{dk_{\hat{n}}}{2\pi}\frac{d^2\boldsymbol{k}_\perp}{(2\pi)^2}e^{i\chi\boldsymbol{k}\cdot\hat{\boldsymbol{n}}}(2\pi)^2\delta_D^{(2)}(\boldsymbol{\ell} - \chi\boldsymbol{k}_\perp) \\ D(z[\chi])\left[b + f\frac{(\boldsymbol{k}\cdot\hat{\boldsymbol{n}})^2}{k^2} + \frac{H^2}{(1+z)^2}\frac{3}{2}\Omega_m\mathcal{B}\frac{1}{k^2} - i\frac{f\alpha}{\chi}\frac{(\boldsymbol{k}\cdot\hat{\boldsymbol{n}})}{k^2}\right]\delta(\boldsymbol{k} = k_{\hat{n}}\hat{\boldsymbol{n}} + \boldsymbol{k}_\perp, z[\chi])\,,\end{aligned} \tag{90}$$

where the Dirac delta relates the transverse Fourier mode to the angular scale, and integrating over $\boldsymbol{k}_\perp$ we obtain

$$\begin{aligned}\delta_g(\boldsymbol{\ell}, z[\bar{\chi}]) = \int\frac{d\chi}{\chi^2}W(\chi, \bar{\chi})D(z[\chi])\int\frac{dk_{\hat{n}}}{2\pi}e^{ik_{\hat{n}}\chi}\left[b + f\mu^2 + \frac{1}{k^2}\frac{H^2}{(1+z)^2}\frac{3}{2}\Omega_m\mathcal{B} - i\frac{f\mu\alpha}{k\chi}\right] \\ \times\delta(\boldsymbol{k} = k_{\hat{n}}\hat{\boldsymbol{n}} + \boldsymbol{\ell}/\chi, z[\chi])\,,\end{aligned} \tag{91}$$

where $\mu(k_{\hat{n}}, \ell/\chi)$ and $k(k_{\hat{n}}, \ell/\chi)$. The two-point unequal-time correlator of the projected overdensity field, Equation (91), for two different dark matter tracers A and B at redshifts $z[\bar{\chi}_1]$ and $z[\bar{\chi}_2]$ respectively, is

$$\begin{aligned}
&\left\langle \delta_{g,A}(\boldsymbol{\ell}_1, z_1[\bar{\chi_1}]) \delta^*_{g,B}(\boldsymbol{\ell}_2, z_2[\bar{\chi}_2]) \right\rangle \\
&= \int \frac{d\chi_1}{\chi_1^2} \frac{d\chi_2}{\chi_2^2} W(\chi_1, \bar{\chi}_1) W(\chi_2, \bar{\chi}_2) D(z[\chi_1])]) D(z[\chi_2]) \int \frac{dk_{1,\hat{n}}}{2\pi} \frac{dk_{2,\hat{n}}}{2\pi} e^{ik_{1,\hat{n}}\chi_1} e^{-ik_{2,\hat{n}}\chi_2} \\
&\quad \left[b_{1,A} + f_1\mu_1^2 + \frac{\mathcal{A}_{1,A}}{k_1^2} - i\frac{f_1\mu_1\alpha_{1,A}}{k_1\chi_1} \right] \left[b_{2,B} + f_2\mu_2^2 + \frac{\mathcal{A}_{2,B}}{k_2^2} + i\frac{f_2\mu_2\alpha_{2,B}}{k_2\chi_2} \right] \\
&\quad \left\langle \delta(k_{1,\hat{n}}\hat{\boldsymbol{n}}, \boldsymbol{\ell}_1/\chi_1, z_1[\chi_1]) \delta^*(k_{2,\hat{n}}\hat{\boldsymbol{n}}, \boldsymbol{\ell}_2/\chi_2, z_2[\chi_2]) \right\rangle \\
&= \int \frac{d\chi_1}{\chi_1^2} \frac{d\chi_2}{\chi_2^2} W(\chi_1, \bar{\chi}_1) W(\chi_2, \bar{\chi}_2) (2\pi)^2 \delta_D^{(2)} \left(\frac{\boldsymbol{\ell}_1}{\chi_1} - \frac{\boldsymbol{\ell}_2}{\chi_2} \right) D(z[\chi_1])]) D(z[\chi_2]) \\
&\quad \int \frac{dk_{\hat{n}}}{2\pi} e^{ik_{\hat{n}}(\chi_1 - \chi_2)} \mathcal{P}(k_{\hat{n}}\hat{\boldsymbol{n}}, \boldsymbol{k}_\perp, \chi_1, \chi_2) K_A(k, \mu_1, \chi_1) K_B(k, \mu_2, \chi_2) \,.
\end{aligned} \tag{92}$$

We have defined the redshift space kernel including the standard Kaiser term [24, 30], Doppler [112], and relativistic contribution [123] for the tracer X evaluated at χ_i, where $i = 1, 2$,

$$K_X(k, \mu, \chi_i) = \left[b_{X,i} + f_i\mu^2 + \frac{\mathcal{A}_{X,i}}{k^2} \mp i\frac{f_i\mu\alpha_{X,i}}{k\chi_i} \right] , \tag{93}$$

and

$$\mathcal{A}_{X,i}(b_{e,X}, Q_X, z_i) = \frac{H(z_i)^2}{(1+z_i)^2} \frac{3}{2} \Omega_m(z_i) \mathcal{B}(b_{e,X}, Q_X, z_i) \,. \tag{94}$$

When building the two-point correlator, we assume the plane-parallel approximation. An investigation on the range of validity of this approximation is discussed in Sec IV C, where we show that for the redshift range considered in Sect. **??** we can fully trust this approach. Introducing the mean distance $\chi = (\chi_1 + \chi_2)/2$ and the unequalness-in-time $\delta\chi = \chi_1 - \chi_2$, and changing variables from (χ_1, χ_2) to $(\chi, \delta\chi)$, Equation (92) becomes

$$\begin{aligned}
&\left\langle \delta_{g,A}(\boldsymbol{\ell}_1, z_1[\bar{\chi_1}]) \delta^*_{g,B}(\boldsymbol{\ell}_2, z_2[\bar{\chi}_2]) \right\rangle = \\
&\int \frac{d\chi}{\chi^2} d\delta\chi W(\chi_1(\chi, \delta\chi), \bar{\chi}_1) W(\chi_2(\chi, \delta\chi), \bar{\chi}_2) (2\pi)^2 \delta_D^{(2)} \left(\boldsymbol{\ell}_1 - \boldsymbol{\ell}_2 - \frac{\delta\chi}{2\chi} (\boldsymbol{\ell}_1 + \boldsymbol{\ell}_2) \right) \\
&\times D(z[\chi_1(\chi, \delta\chi)])) D(z[\chi_2(\chi, \delta\chi)])) \int \frac{dk_{\hat{n}}}{2\pi} e^{ik_{\hat{n}}\delta\chi} K_A(k, \mu, \chi_1(\chi, \delta\chi)) K_B(k, \mu, \chi_2(\chi, \delta\chi)) \\
&\times \mathcal{P}(k_{\hat{n}}\hat{\boldsymbol{n}}, \boldsymbol{k}_\perp, \chi, \delta\chi) \,.
\end{aligned} \tag{95}$$

If $\chi_1 = \chi_2$, we recover the perfect equal-time case, where the angular power spectrum can be easily calculated using the well-known properties of the Bessel functions (see [103] for a detailed discussion, and [101] for an equal-time calculation). Moreover, the argument of the Dirac delta function picks up an extra term $-\delta\chi(\boldsymbol{\ell}_1 + \boldsymbol{\ell}_2)/2\chi$ with respect to the equal-time case, where it is only $\boldsymbol{\ell}_1 - \boldsymbol{\ell}_2$. Refs. [101, 102] show that the off-diagonal part of the Dirac delta gives rise to negligible corrections to the equal-time angular power spectrum, with respect to the other unequal-time corrections that arise from the power spectrum and the overdensity kernels,therefore we will drop them in the following, and we will use the standard $\delta_D^{(2)}(\boldsymbol{\ell}_1 - \boldsymbol{\ell}_2)$. [8]

[8] After a somewhat tedious calculation, one can show that this assumption remains true also when including Doppler terms and

Assuming that the window functions are narrow enough, so that they can be approximated by Dirac deltas centered at the mean values $\bar{\chi}_1$ and $\bar{\chi}_2$, we have

$$\mathbb{C}(\ell, \bar{\chi}, \delta\chi) = \frac{1}{\bar{\chi}^2} \int \frac{dk_{\hat{n}}}{2\pi} e^{ik_{\hat{n}}\delta\chi} K_A(k_{\hat{n}}\hat{\boldsymbol{n}}, \boldsymbol{\ell}/\bar{\chi}, \bar{\chi}, \delta\chi) K_B(k_{\hat{n}}\hat{\boldsymbol{n}}, \boldsymbol{\ell}/\bar{\chi}, \bar{\chi}, \delta\chi) \mathcal{P}(k_{\hat{n}}\hat{\boldsymbol{n}}, \boldsymbol{\ell}/\bar{\chi}, \bar{\chi}, \delta\chi)\,, \tag{96}$$

which is the flat-sky version of the full-sky unequal-time angular power spectrum. We refer the reader to Ref. [103] for a detailed discussion of the mathematical connection between flat-sky and full-sky, in terms of asymptotic behavior of the spherical Bessel functions.

The equal-time, observed power spectrum is defined as [102]

$$P_{\rm obs}\left(k_{\hat{n}}, \ell/\bar{\chi}, \bar{\chi}\right) = \bar{\chi}^2 \int d\delta\chi\, e^{-i\delta\chi k_{\hat{n}}} \mathbb{C}(\ell, \bar{\chi}, \delta\chi)\,. \tag{97}$$

where $k_{\hat{n}}$ represents the Fourier counterpart of the radial separation between the two tracers $\delta\chi$. In this fashion we can transfer the unequal-time information contained in the angular power spectrum $\mathbb{C}(\ell, \bar{\chi}, \delta\chi)$ to the observed power spectrum in Fourier space. In this transformation, not only the spatial separation but also any contribution arising from time evolution—through redshift dependence —is effectively mapped into comoving spatial coordinates. This follows from using light-cone coordinates, which ensure that time evolution and projection effects are incorporated into the spatial structure of the observed power spectrum.

Since the distance between the observed system and the observer is much larger than the radial distance between the sources (i.e., $\delta\chi/\bar{\chi} \ll 1$), and since radial correlations die off very steeply as one departs from the equal-time case, we can now expand the matter power spectrum and the kernels of the galaxy overdensity field around the equal-time case $\delta\chi = 0$, as

$$\begin{aligned} &K_A(k_{\hat{n}}\hat{\boldsymbol{n}}, \boldsymbol{\ell}/\bar{\chi}, \bar{\chi}, \delta\chi) K_B(k_{\hat{n}}\hat{\boldsymbol{n}}, \boldsymbol{\ell}/\bar{\chi}, \bar{\chi}, \delta\chi) \mathcal{P}(k_{\hat{n}}\hat{\boldsymbol{n}}, \boldsymbol{\ell}/\bar{\chi}, \bar{\chi}, \delta\chi) \\ &\quad = D\left[z(\bar{\chi})\right]^2 \sum_{n=0}^{\infty} c_n(k, \mu, \bar{\chi}) H\left[z(\bar{\chi})\right]^n (\delta\chi)^n \mathcal{P}_0(k)\,, \end{aligned} \tag{98}$$

where $D\left[z(\bar{\chi})\right]$ is the growth factor at the mean comoving distance between the two sources (i.e., $(\chi_1 + \chi_2)/2$), the coefficient c_0 contains the equal-time kernels, and the c_n embed the unequal-time (UT) corrections, whose detailed expressions are computed in Section IV B. The quantity $\mathcal{P}_0(k)$ is the theoretical dark matter equal-time linear power spectrum defined in Equation (81) at redshift $z = 0$. Then

$$\begin{aligned} P_{\rm obs}\left(k_{\hat{n}}, \ell/\bar{\chi}, \bar{\chi}\right) &= \int d\delta\chi \int \frac{dk'_{\hat{n}}}{2\pi} e^{i\delta\chi(k'_{\hat{n}} - k_{\hat{n}})} D\left[z(\bar{\chi})\right]^2 \sum_{n=0}^{\infty} c_n(k', \mu, \bar{\chi}) H\left[z(\bar{\chi})\right]^n \delta\chi^n \mathcal{P}_0(k') \\ &= D\left[z(\bar{\chi})\right]^2 \sum_{n=0}^{\infty} \int d\delta\chi \int \frac{dk'_{\hat{n}}}{2\pi} \left[\left(-iH\left[z(\bar{\chi})\right] \frac{d}{dk'_{\hat{n}}}\right)^n e^{i\delta\chi(k'_{\hat{n}} - k_{\hat{n}})}\right] c_n(k', \mu, \bar{\chi}) \mathcal{P}_0(k') \\ &= D\left[z(\bar{\chi})\right]^2 \sum_{n=0}^{\infty} \left(iH\left[z(\bar{\chi})\right] \frac{d}{dk_{\hat{n}}}\right)^n \left[c_n(k, \mu, \bar{\chi}) \mathcal{P}_0(k)\right]\,. \end{aligned} \tag{99}$$

Equation (99) is our master equation, that will be used in the following to derive the unequal-time corrections to the power spectrum.

General Relativistic effects (see [124] for the explicit computation).

B General Relativistic and Doppler unequal-time coefficients

While the simplest c_n coefficients were derived in [101, 102], here we add the contributions to $c_n(k,\mu,\bar{\chi})$ coming from Doppler terms and local general relativistic corrections; we start by expanding α and $\mathcal{A}$ around the equal-time case, $\delta\chi = 0$. At first order,

$$\frac{\mathcal{A}(z[\chi(\bar{\chi},\delta\chi)])}{\mathcal{A}(z[\chi(\bar{\chi},0)])} = 1 + H\left[2\left(\frac{H'}{H} - \frac{1}{1+z}\right) + \frac{\Omega_m'}{\Omega_m} + \frac{\mathcal{B}'}{\mathcal{B}}\right]\left.\frac{d\chi}{d\delta\chi}\right|_{\delta\chi=0}\delta\chi + O(\delta\chi^2)\,, \tag{100}$$

where $\mathcal{A}$ depends on the evolution bias b_e and magnification bias Q, and

$$\frac{\alpha(z[\chi(\bar{\chi},\delta\chi)])}{\alpha(z[\chi(\bar{\chi},0)])} = 1 + H\left(\frac{H'}{H} - \frac{1}{1+z} + \frac{\tilde{\alpha}'}{\tilde{\alpha}} + \frac{1}{H\bar{\chi}}\right)\left.\frac{d\chi}{d\delta\chi}\right|_{\delta\chi=0}\delta\chi + O(\delta\chi^2)\,, \tag{101}$$

with α being the Doppler term coefficient, and where a prime denotes derivative with respect to redshift. We have rewritten the Doppler term in Equation (85) as $\alpha = -H\chi\tilde{\alpha}/(1+z)$, with

$$\tilde{\alpha} = b_e(z) - 1 - 2Q(z) + \frac{3}{2}\Omega_m(z) - \frac{2(1-Q(z))(1+z)}{H(z)\chi(z)}\,. \tag{102}$$

Explicitly,

$$\tilde{\alpha}' = b_e' - 2Q' + \frac{3}{2}\Omega_m' + 2[\frac{Q'(1+z)}{H\bar{\chi}} + \frac{(1-Q)(1+z)}{(H\bar{\chi})^2} - \frac{1-Q}{H\bar{\chi}} + \frac{(1-Q)(1+z)}{H\bar{\chi}}\frac{H'}{H}]\,, \tag{103}$$

and

$$\begin{aligned}\mathcal{B}' =& b_e'\left(1 - \frac{2f}{3\Omega_m}\right) - \frac{2b_e f}{3\Omega_m}\left(\frac{f'}{f} - \frac{\Omega_m'}{\Omega_m}\right) + \frac{2f}{\Omega_m}\left(\frac{f'}{f} - \frac{\Omega_m'}{\Omega_m}\right) + \frac{3}{2}\Omega_m' - 4Q' - f' \\ &+ \frac{2}{H\bar{\chi}}\left[Q'(1+z) - (1-Q)\right] + \frac{2(1-Q)(1+z)}{H\bar{\chi}}\left(\frac{H'}{H} - \frac{1}{H\bar{\chi}}\right).\end{aligned} \tag{104}$$

Isolating the dependence on the wave-vector and orientation angle, the first-order coefficient in Equation (99) is

$$\begin{aligned}c_1(k,\mu,\bar{z}) =& c_{100}(\bar{z}) + \mu^2 c_{120}(\bar{z}) + \mu^4 c_{140}(\bar{z}) + \left(\frac{H}{k}\right)^2 c_{102}(\bar{z}) + \mu^2\left(\frac{H}{k}\right)^2 c_{122}(\bar{z}) + \left(\frac{H}{k}\right)^4 c_{104}(\bar{z}) \\ &+ i\mu\left(\frac{H}{k}\right) c_{111}(\bar{z}) + i\mu^3\left(\frac{H}{k}\right) c_{131}(\bar{z}) + i\mu\left(\frac{H}{k}\right)^3 c_{113}(\bar{z})\,.\end{aligned} \tag{105}$$

At this stage, different choices of the mean distance are possible – see Appendix B of [101] for more details. As in the previous section, we choose the arithmetic mean, $\bar{\chi} = (\chi_A + \chi_B)/2$, so that $\delta\chi = \chi_A - \chi_B$

and $d\chi_{A,B}/d\delta\chi = \pm 1/2$. The coefficients are then

$$c_{100} = \frac{1}{2} b_A b_B \left(\frac{b'_A}{b_A} - \frac{b'_B}{b_B} \right) , \tag{106a}$$

$$c_{120} = \frac{f}{2} \left[b'_A - b'_B - \frac{f'}{f} (b_A - b_B) \right] , \tag{106b}$$

$$c_{140} = 0 , \tag{106c}$$

$$c_{102} = \frac{1}{2} \left[\left(b'_A \frac{\mathcal{A}_B}{H^2} - b_A \frac{\mathcal{A}'_B}{H^2} \right) - \left(b'_B \frac{\mathcal{A}_A}{H^2} - b_B \frac{\mathcal{A}'_A}{H^2} \right) \right] , \tag{106d}$$

$$c_{122} = \frac{f}{2} \left[\frac{\mathcal{A}'_A}{H^2} - \frac{\mathcal{A}'_B}{H^2} - \frac{f'}{f} \left(\frac{\mathcal{A}_A}{H^2} - \frac{\mathcal{A}_B}{H^2} \right) + \frac{f \alpha_A \alpha_B}{(H\chi)^2} \left(\frac{\alpha'_A}{\alpha_A} - \frac{\alpha'_B}{\alpha_B} \right) \right] , \tag{106e}$$

$$c_{104} = \frac{1}{2} \frac{\mathcal{A}_A}{H^2} \frac{\mathcal{A}_B}{H^2} \left(\frac{\mathcal{A}'_A}{\mathcal{A}_A} - \frac{\mathcal{A}'_B}{\mathcal{A}_B} \right) ; \tag{106f}$$

$$c_{111} = \frac{f}{2H\chi} \left[b'_A \alpha_B + b'_B \alpha_A - b_A \alpha'_B - b_B \alpha'_A + (b_A \alpha_B + b_B \alpha_A) \left(\frac{1}{H\chi} - \frac{f'}{f} \right) \right] , \tag{107a}$$

$$c_{131} = \frac{f^2}{2H\chi} \left(\frac{\alpha_A + \alpha_B}{H\chi} - \alpha'_A - \alpha'_B \right) , \tag{107b}$$

$$c_{113} = \frac{f}{2H\chi} \left[\left(\alpha_A \frac{\mathcal{A}_B}{H^2} + \alpha_B \frac{\mathcal{A}_A}{H^2} \right) \left(\frac{1}{H\chi} - \frac{f'}{f} \right) + \alpha_A \frac{\mathcal{A}'_B}{H^2} + \alpha_B \frac{\mathcal{A}'_A}{H^2} - \alpha'_A \frac{\mathcal{A}_B}{H^2} - \alpha'_B \frac{\mathcal{A}_A}{H^2} \right] . \tag{107c}$$

The first set of coefficients, the c_{1ij} with $\{i, j\}$ being even, gives rise to an imaginary unequal-time correction. [9]The second set, the c_{1ij} with $\{i, j\}$ being odd, gives rise to a real correction. All redshift-dependent quantities in Equations (106)– (107) are evaluated at $\delta\chi = 0$, by construction. With the arithmetic choice of mean distance, the coefficient c_{140} vanishes since it only receives contributions from the $f\mu^2$ terms of the kernels.

The terms proportional to $\mathcal{A}'$ and α' depend on the time derivatives of the evolution and magnification biases, b'_e and Q', that are related to the galaxies' redshift and luminosity distribution.

In the absence of Doppler and general relativistic terms, as noted in [101, 102], the first order unequal-time correction would vanish in the single-tracer case, while in the multi-tracer case it would be purely imaginary, and sourced by the first set of coefficients. When including Doppler terms and general relativistic corrections, instead, the second set of coefficients is non-vanishing, even in the single-tracer case, and it gives rise to a real first order correction as well. Explicitly, in the single-tracer case, the first set of coefficients, Equations (106), vanishes, and Equations (107) become

$$c_{111} = \frac{fb\alpha}{H\chi} \left[\frac{b'}{b} - \frac{\alpha'}{\alpha} + \left(\frac{1}{H\chi} - \frac{f'}{f} \right) \right] , \tag{108a}$$

$$c_{131} = \frac{f^2 \alpha}{H\chi} \left(\frac{1}{H\chi} - \frac{\alpha'}{\alpha} \right) , \tag{108b}$$

$$c_{113} = \frac{f\alpha}{H\chi} \frac{\mathcal{A}}{H^2} \left[\left(\frac{1}{H\chi} - \frac{f'}{f} \right) - \frac{\alpha'}{\alpha} + \frac{\mathcal{A}'}{\mathcal{A}} \right] . \tag{108c}$$

In general, the largest deviations coming from UT contributions are imaginary, as illustrated in Figure 12. There we show the real and imaginary part of the redshift-space galaxy power spectrum in the equal-time and unequal-time cases, for $\mu = 1$ (as unequal-time corrections are expected to be maximized

[9] As shown explicitly in [101], coefficients arising only from the Taylor expansion of the growth factor $D(z)$ vanish at first order in $\delta\chi$ (i.e., unequal time corrections for dark matter in real space are $O(\delta\chi^2)$)

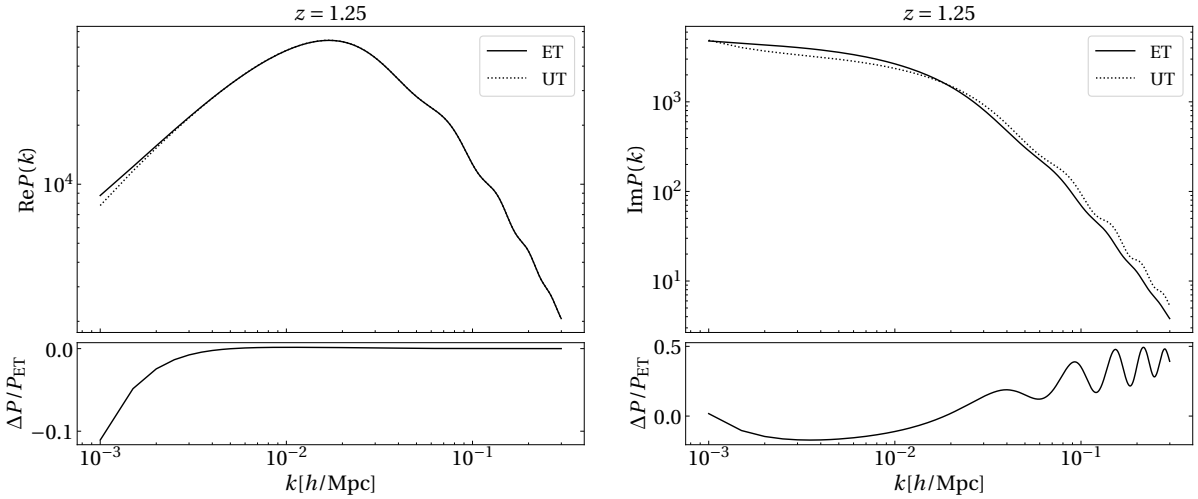

Figure 12: Real (*left*) and imaginary (*right*) part of the galaxy power spectrum in the equal-time and unequal-time case for $\mu = 1$. The residuals for UT corrections ΔP normalized to the ET power spectrum are shown in the lower panels.

for radial correlators).[10] As an example case, we used $b_A = (2+z)^{0.5}$, $b_B = (0.5+z)^{1.2}$, and number counts distributions $\frac{dN_X}{dz} \propto z^{\gamma_X} e^{-(z/0.5)^{1.5}}$ with $\gamma_A = 4$, $\gamma_B = 1.1$. The relative importance of UT corrections (i.e., $(P_{\mathrm{UT}} - P_{\mathrm{ET}})/P_{\mathrm{ET}}$), is shown in the lower panels, where one can see that the bulk of the UT contribution is in the imaginary part of the multi-tracer power spectrum. Such imaginary corrections arise because in general the multi-tracer power spectrum is not invariant under the transformation $\mu \to -\mu$.

Since UT corrections can give a contribution to the power spectrum at large scales, it is natural to wonder whether there might be degeneracies with the imprints of local-type Primordial non-Gaussianity (PnG). A similar investigation was carried out in the context of Doppler and relativistic equal-time corrections, in [112, 125]). In order to do this, the first step is to derive the unequal-time contributions in the presence of Primordial non-Gaussianity. Expanding the non-Gaussian bias around $\delta\chi = 0$ leads to

$$\frac{b_{\mathrm{NG}}(z[\chi(\bar{\chi}, \delta\chi)])}{b_{\mathrm{NG}}(z[\chi(\bar{\chi}, 0)])} = 1 + H\left(\frac{b'_\phi}{b_\phi} + \frac{f}{1+z}\right) \frac{d\chi}{d\delta\chi}\bigg|_{\delta\chi=0} \delta\chi + O(\delta\chi^2)\,, \tag{109}$$

which gives the additional UT correction

$$c_{1,\mathrm{NG}}(k,\mu,\bar{z}) = \frac{3 f_{\mathrm{NL}} \Omega_{m,0} H_0^2}{2T(k) D_{\mathrm{md}}(z)} \left[\left(\frac{H}{k}\right)^2 c_{102,\mathrm{NG}}(\bar{z}) + \mu^2 \left(\frac{H}{k}\right)^2 c_{122,\mathrm{NG}}(\bar{z}) + \left(\frac{H}{k}\right)^4 c_{104,\mathrm{NG}}(\bar{z}) + i\mu \left(\frac{H}{k}\right)^3 c_{113,\mathrm{NG}}(\bar{z})\right], \tag{110}$$

[10] The imaginary part of the equal-time power spectrum is zero in the single-tracer case, but when cross-correlating two different tracers, the equal-time power spectrum picks up a non-vanishing imaginary part (only when Doppler terms are included).

where the coefficients are

$$c_{102,\rm NG} = \frac{1}{2H^2}\left[\frac{f}{1+z}(b_{\phi,A}b_B - b_{\phi,B}b_A) + b_{\phi,B}b'_A - b_{\phi,A}b'_B + b'_{\phi,A}b_B - b'_{\phi,B}b_A\right], \tag{111a}$$

$$c_{122,\rm NG} = \frac{f}{2H^2}\left[\frac{f'}{f}(b_{\phi,B} - b_{\phi,A}) + b'_{\phi,A} - b'_{\phi,B} + \frac{f}{1+z}(b_{\phi,A} - b_{\phi,B})\right], \tag{111b}$$

$$c_{104,\rm NG} = \frac{1}{2H^4}\left[b_{\phi,B}\mathcal{A}'_A - b_{\phi,A}\mathcal{A}'_B + \frac{f}{1+z}(\mathcal{A}_B b_{\phi,A} - \mathcal{A}_A b_{\phi,B}) + \mathcal{A}_B b'_{\phi,A} - \mathcal{A}_A b'_{\phi,B}\right], \tag{111c}$$

$$c_{113,\rm NG} = \frac{f}{2H^3\chi}\Bigg[(\alpha_B b'_{\phi,A} + \alpha_A b'_{\phi,B}) + \frac{f}{1+z}(\alpha_A b_{\phi,B} + \alpha_B b_{\phi,A}) \tag{111d}$$

$$+ \frac{1}{H\chi}(\alpha_A b_{\phi,B} + \alpha_B b_{\phi,A}) + \frac{f'}{f}(-\alpha_A b_{\phi,B} - \alpha_B b_{\phi,A}) - \alpha'_A b_{\phi,B} - \alpha'_B b_{\phi,A}\Bigg]. \tag{111e}$$

For simplicity, we assume the universality relation, so that $b_\phi = 2\delta_c(b_{\rm G} - 1)$, where $\delta_c = 1.686$ is the critical value of the matter overdensity for spherical collapse (for some generalizations see, e.g., [51, 56, 126, 127]).

Even if the leading unequal time corrections are imaginary, one can see from Figure 12 that the real part of the UT contributions can give an $\sim 10\%$ correction to the standard equal time power spectrum at large scales. It is interesting to investigate whether this is degenerate with contributions coming from the scale-dependent bias sourced by a non-zero $f_{\rm NL}$. To proceed, let us write the main real UT contributions to the ET power spectrum. Recall that in general

$$\Delta_{\rm UT} \propto \frac{d}{dk_{\hat{n}}}\left[c_1(k,\mu,\bar{\chi})\mathcal{P}_0(k)\right] \propto \frac{d}{dk_{\hat{n}}}\left[\sum_{a,b} H^b \frac{\mu^a}{k^b} c_{1ab}(\bar{\chi})\mathcal{P}_0(k)\right], \tag{112}$$

where c_{1ab} are given by Equations (106) and Equations (107). Since k and μ both depend on $k_{\hat{n}}$,

$$\frac{d}{dk_{\hat{n}}}\left[\frac{\mu^a}{k^b}c_{1ab}(\bar{\chi})\mathcal{P}_0(k)\right] = \left\{\frac{1}{k^{b+1}}\left[-b\mu^{a+1} + a\mu^{a-1}(1-\mu^2)\right]c_{1ab}(\bar{\chi})\mathcal{P}_0(k) + c_{1ab}(\bar{\chi})\frac{\mu^{a+1}}{k^b}\frac{d\mathcal{P}_0(k)}{dk}\right\} \equiv C_{1ab}\,. \tag{113}$$

On very large scales $d\mathcal{P}_0(k)/dk \sim \mathcal{P}_0(k)/k$; as shown explicitly in Sec. IV D, the dominant real UT contributions are sourced by the coefficients c_{111}, c_{131}, therefore

$$\Delta_{UT,\rm real} \sim H^2\left(1-\mu^2\right)\left(c_{111}(\bar{\chi}) + 3\mu^2 c_{131}(\bar{\chi})\right)\frac{\mathcal{P}_0(k)}{k^2}\,. \tag{114}$$

One can notice that the real part of the UT corrections exhibits the same k^{-2} behavior of the scale dependent non-Gaussian bias of Equation (44) indicating that these kinds of corrections in principle can mimic an effective $f_{\rm NL}$ on large scales. In general the precise value of the effective $f_{\rm NL}$ depends on the values of the dimensionless coefficients c_{111}, c_{131}, which are quantities of order one and depend on $\{b_A, b_B, \alpha_A, \alpha_B, f\}$ and their derivatives. As a first investigation, in Figure 13 we present some values of c_{111}, c_{131} which can mimic different values of $f_{\rm NL}$ at different redshift, for $\mu = 0.4$ as an example (in general the net effect exhibits a weak dependence on μ). Non-Gaussian power spectra are shown for $f_{\rm NL} = \hat{f}_{\rm NL}$ where $\hat{f}_{\rm NL}$ is the effective $f_{\rm NL}$ mimicked by UT corrections. Notably, ET power spectra with $f_{\rm NL} = \hat{f}_{\rm NL}$ (red dashed) overlap with the UT models with zero $f_{\rm NL}$ (dashed black) showing that $\hat{f}_{\rm NL}$ is indeed the effective $f_{\rm NL}$ mimicked by unequal-time corrections. This effects varies non-trivially with z depending on the specific values of the biases and α of the tracers considered. While we leave a detailed investigation of the impact of this for future surveys, we can notice that the effective $f_{\rm NL}$ mimicked by the UT corrections seems to be relevant for the targets of current and future galaxy surveys measuring PnG.

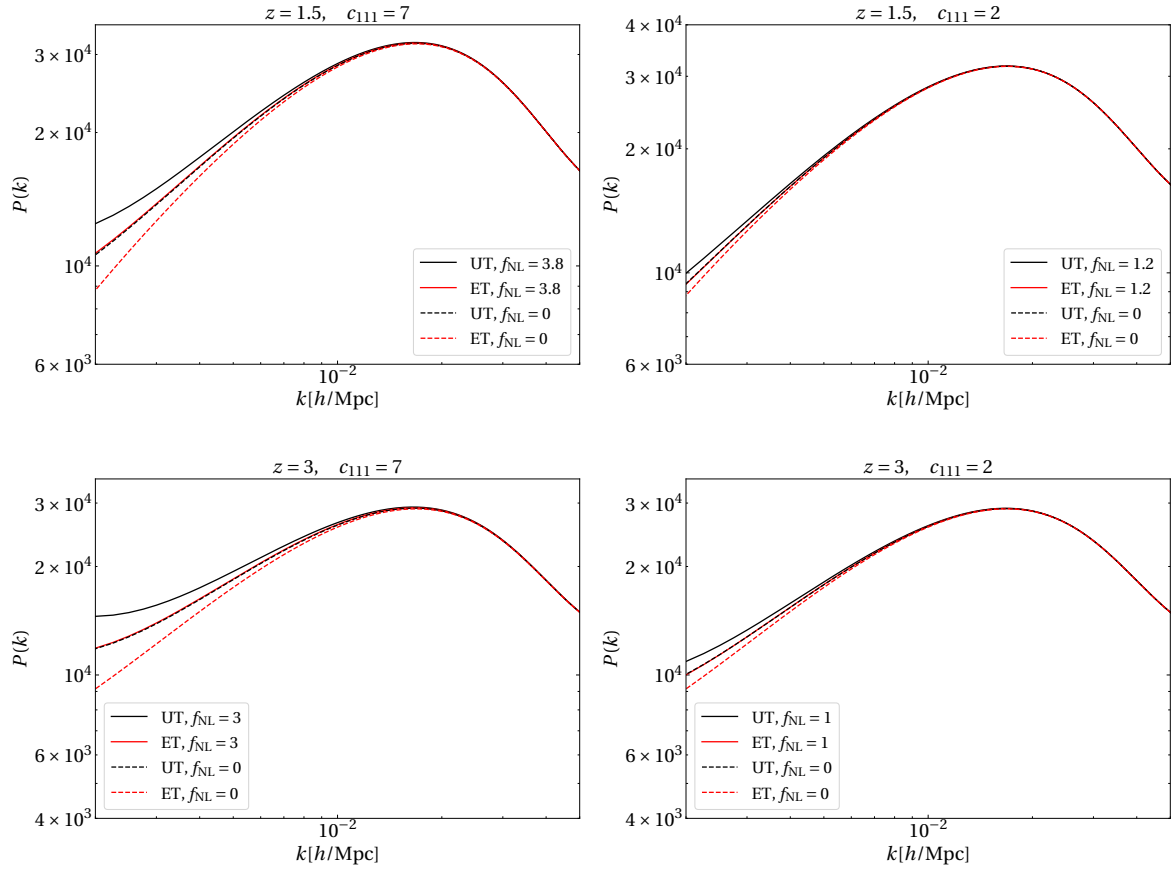

Figure 13: Equal-time power spectrum and unequal-time power spectrum for different values of the coefficient c_{111} at redshift $z = 1.5$ (*upper panels*) and $z = 3$ (*lower panels*) at $\mu = 0.4$. Here $c_{131} = 1$.

C *Validity of the flat sky approximation*

So far, we always adopted the flat sky approximation for the angular power spectrum $\mathbb{C}(\ell, \bar{\chi}, \delta\chi)$ It is interesting to see what happens if we substitute in Equation (97) the full-sky angular power spectrum $C_\ell(\bar{\chi}, \delta\chi)$. This comparison can give us a hint of the validity of the flat sky approximation. Then the two models to be compared are

$$\begin{aligned} P_{\text{fullsky}}\left(k_{\hat{n}}, \ell/\bar{\chi}, \bar{\chi}\right) &= \bar{\chi}^2 \int d\delta\chi \, e^{-i\delta\chi k_{\hat{n}}} C_\ell(\bar{\chi}, \delta\chi) \,, \\ P_{\text{flatsky}}\left(k_{\hat{n}}, \ell/\bar{\chi}, \bar{\chi}\right) &= D\left[z(\bar{\chi})\right]^2 \sum_{n=0}^{\infty} \left(iH\left[z(\bar{\chi})\right] \frac{d}{dk_{\hat{n}}}\right)^n \left[c_n(k_{\hat{n}}, \ell/\bar{\chi}, z(\bar{\chi})) \mathcal{P}_0(k_{\hat{n}}, \ell/\bar{\chi}, z(\bar{\chi}))\right] . \end{aligned} \tag{115}$$

To properly compare the two models, one needs to express everything in terms of the variables $\{k_{\hat{n}}, \ell/\bar{\chi}\}$. In these variables, the equal-time power spectrum covariance reads, neglecting shot noise terms,

$$\mathrm{Cov}(k_{\hat{n}}, \ell/\bar{\chi}) = \bar{\chi}^2 \frac{4\pi^2}{V \ell \Delta\ell \Delta k_{\hat{n}}} \left(P_{AB}^2 + P_{AA} P_{BB}\right) . \tag{116}$$

In Figure 14 and Figure 15 we compare respectively the real and imaginary part of the models for different redshifts for $k_{\hat{n}} \sim 0.003\, h/\mathrm{Mpc}$, $\Delta k_{\hat{n}} = 0.01\, h/\mathrm{Mpc}$, $\Delta\ell = 1$ as an example. The redshift half-width is kept fixed at $\Delta z = 0.2$, and $f_{\mathrm{sky}} \sim 0.75$. Errors taken from Equation (116) are also shown in grey.

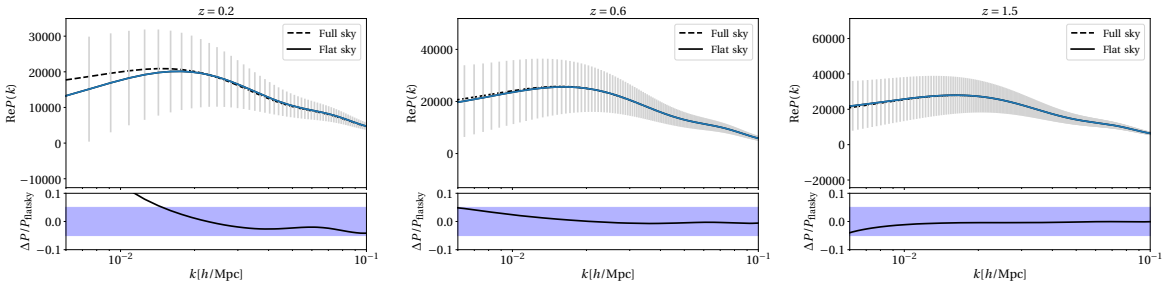

Figure 14: Comparison between the real part of the full-sky and flat-sky model in redshift space. Errors drawn from Equation (116) are also shown in grey. Residuals $(P_{\text{fullsky}} - P_{\text{flatsky}})/P_{\text{flatsky}}$ are shown in the lower panels, where the shaded region highlights deviations $\lesssim 5\%$.

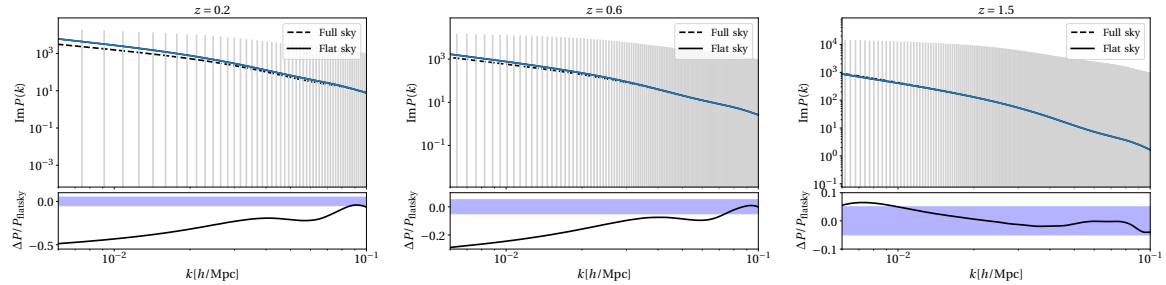

Figure 15: Same as Figure 14, but for the imaginary part of the power spectrum.

Residuals $(P_{\text{fullsky}} - P_{\text{flatsky}})/P_{\text{flatsky}}$ are shown in the lower panels, where the shaded region highlights deviations $\lesssim 5\%$.

For $z \lesssim 0.6$ the flat-sky model starts deviating considerably from the full-sky one at large scales, and this effect gets bigger at lower redshift, where the sky curvature becomes non-negligible for large scales. While these kinds of deviations seem to be well within the cosmic variance limit, in this work we decided to focus on the redshift range $z \gtrsim 1$, where one can safely assume that the flat-sky approximation is very accurate.

D *Relevance of unequal-time coefficients*

It can be instructive to investigate the relative importance of the various unequal-time corrections. The quanity $C_{1ab}(k, \mu)$ defined in Equation(113) gives the relative importance of various unequal-time corrections as a function of scales and orientation angle. They are plotted in Figure 16 and Figure 17 for $\mu = 1$ and in Figure 18 and Figure 19 for $\mu = 0$. As one can see, the bulk of the unequal-time coefficients comes from C_{100} which is imaginary and $\propto b_B b'_A - b_A b'_B$ and for $\mu = 1$. UT corrections arising from Doppler and GR terms, which contribute to the real part of unequal-time corrections, can be important at very large scales. In these case the dominant real UT correction is sourced by the coefficient c_{111} and it is maximized for $\mu = 0$.

Then, we present here a first investigation of the relevance of first-order unequal-time corrections to the galaxy power spectrum. Throughout the analysis, we use *Planck* 2018 fiducial values for the cosmological parameters [1] and a fiducial of $f_{\text{NL}} = 1$, and checked that a different choice of the fiducial value of f_{NL} in the range allowed by the latest CMB constraints [9] would not greatly impact the following results.

In order to estimate quantitatively the relevance of all UT corrections, we compute their total signal-to-noise ratio (SNR). In the following forecast we assume that our data-vector is made only by the cross-power spectrum between two different biased tracers. While in real data analysis one usually combines the cross-power spectrum with auto power spectra in order to increase the precision

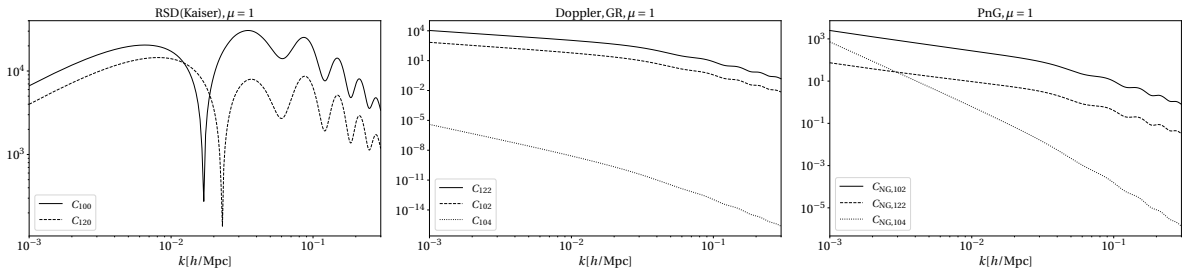

Figure 16: First order imaginary unequal-time corrections sourced by the RSD Kaiser term, for $z = 1.5$ for radial correlators.

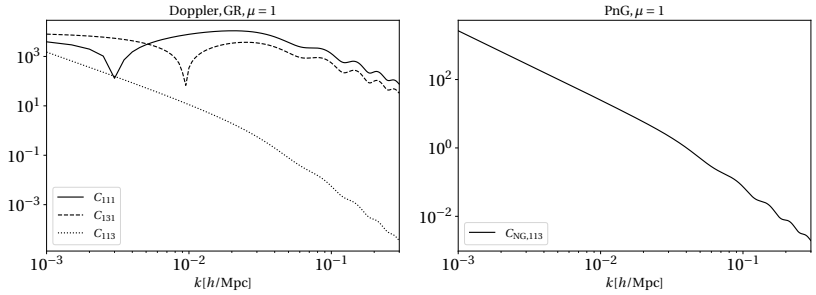

Figure 17: Same as Figure 16 but for real unequal-time corrections sourced by Doppler and GR terms.

in the measurement, our simplified treatment allows to obtain a first back of the envelope estimation of the relevance of unequal time terms. Furthermore, we will show explicitly, UT terms are in general subdominant for galaxy auto-correlations. While this work aims to be a proof of principle of the potential importance of UT corrections a more refined forecast on the impact of such contributions on cosmological parameters inference for realistic galaxy surveys is the subject of an ongoing work. With these assumptions only the cross-power spectrum contributes to the covariance which assumes the following form [128, 129]

$$\mathrm{Cov}(\boldsymbol{k}, \bar{z}) = \frac{4\pi^2}{V_s(\bar{z})k^2\Delta k\Delta\mu}\left[P^{\rm tot}_{{\rm obs},AA}(\boldsymbol{k}, \bar{z})P^{\rm tot}_{{\rm obs},BB}(\boldsymbol{k}, \bar{z}) + \left(P^{\rm tot}_{{\rm obs},AB}(\boldsymbol{k}, \bar{z})\right)^2\right], \tag{117}$$

where $P^{\rm tot}_{{\rm obs},XY} = P_{{\rm obs},XY} + \delta^K_{XY}/\bar{n}_X$, and we neglected cross-shot noise terms since we assume that the stochastic components of two galaxy samples are statistically independent. We considered a Gaussian diagonal form for our theoretical covariance (different k-bins are uncorrelated) since we are focusing on large scales where off-diagonal trispectrum-like contributions are shown to be subdominant (see for

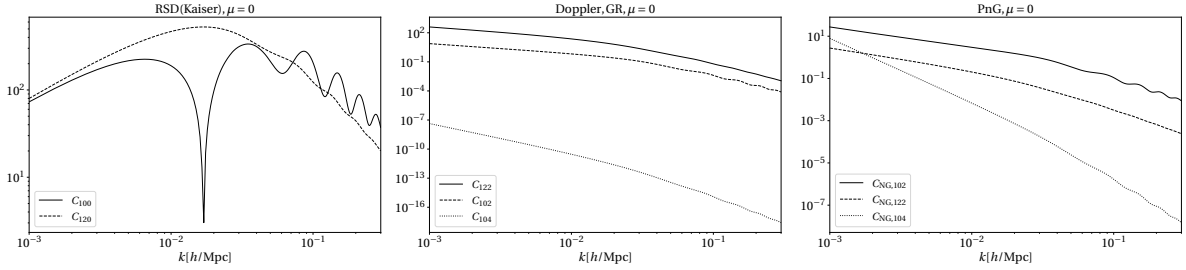

Figure 18: Same as Figure 16 but for transverse correlations.

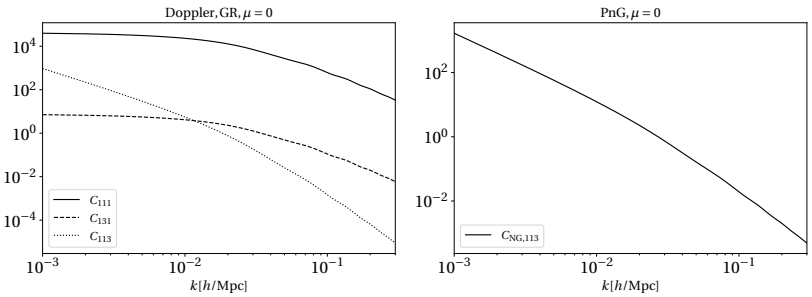

Figure 19: Same as Figure 17 but for transverse correlations.

example [130–134] and references therein). At first order in $\delta\chi$,

$$\begin{aligned}\mathrm{Cov}(\boldsymbol{k},\bar{z}) \simeq &\frac{4\pi^2}{V_s(\bar{z})k^2\Delta k\Delta\mu}\left\{P_{\mathrm{ET},AB}(\boldsymbol{k},\bar{z})^2+\left(P_{\mathrm{ET},AA}(\boldsymbol{k},\bar{z})+\frac{1}{\bar{n}_A}\right)\left(P_{\mathrm{ET},BB}(\boldsymbol{k},\bar{z})+\frac{1}{\bar{n}_B}\right)\right.\\ &\left.+2P_{\mathrm{ET},AB}(\boldsymbol{k},\bar{z})D(\bar{z})^2iH(\bar{z})\partial_{k_{\hat{n}}}\left[c_{1,AB}(k,\mu,\bar{z})\mathcal{P}_0(k)\right]\right\}.\end{aligned} \tag{118}$$

where $P_{\mathrm{ET},XY}(\boldsymbol{k},\bar{z}) = D(\bar{z})^2 c_{0,XY}(k,\mu,\bar{z})\mathcal{P}_0(k)$, and $V_s(\bar{z})$ is the volume of the redshift bin. We neglected UT contributions coming from the auto-galaxy power spectrum. The resulting signal-to-noise ratio is therefore

$$\mathrm{SNR}^2(\bar{z}) = \int_{k,\mu}\frac{\left|D(\bar{z})^2iH(\bar{z})\partial_{k_{\hat{n}}}\left[c_{1,AB}(k,\mu,\bar{z})\mathcal{P}_0(k)\right]\right|^2}{\mathrm{Cov}(\boldsymbol{k},\bar{z})}\,. \tag{119}$$

The UT covariance contains imaginary parts, but their contribution averages to zero when we perform the integration over the orientation angle, since the integrands are proportional to odd powers of μ. It can be shown that, for the redshift range considered in this work, employing the equal-time version of the covariance gives an almost identical SNR.

Furthermore, unequal-time corrections become more important at high redshift. As previously stated (and explicitly shown in Sec. IV D), the dominant UT corrections are imaginary, and come from the coefficients c_{100}. So, to get a first idea on the behavior of the dominant contribution, UT corrections can be approximated as

$$\Delta_{\mathrm{UT}} \sim i\mu H(\bar{z})c_{100}(\bar{z})\frac{d\mathcal{P}_0(k)}{dk}\,, \tag{120}$$

where recall that $c_{100} = \frac{1}{2}b_Ab_B\left(\frac{b'_A}{b_A}-\frac{b'_B}{b_B}\right)$. The scale dependence of the imaginary corrections is governed by the derivative of the linear power spectrum $d\mathcal{P}_0/dk$; this leads to a sign change around the matter-radiation equality scale, as one can see in the right panel of Figure 12. As an illustrative toy model, one can approximate the biases of the two tracers as some power law, i.e., $b^*_A(1+z)^\alpha$, $b^*_B(1+z)^\beta$; then

$$\Delta_{\mathrm{UT}} \sim \frac{i}{2}\mu aH(\bar{z})\frac{d\mathcal{P}_0(k)}{dk}b^*_Ab^*_B(\alpha-\beta)\,, \tag{121}$$

which increases for correlations along the line of sight – as expected, since UT effects are enhanced when the radial separation between the two sources is maximized – and at high redshifts.

This is confirmed in Figure 20, where the signal-to-noise ratio of unequal-time corrections is shown for different redshift bins centered at redshift z, for two different fixed galaxy number densities $\bar{n} \sim 10^{-3}\,(\mathrm{Mpc}/h)^{-3}$ (left panel) and $\bar{n} \sim 10^{-4}\,(\mathrm{Mpc}/h)^{-4}$ (right panel). We assume $f_{\mathrm{sky}} \sim 0.7$, biases $b_A =$

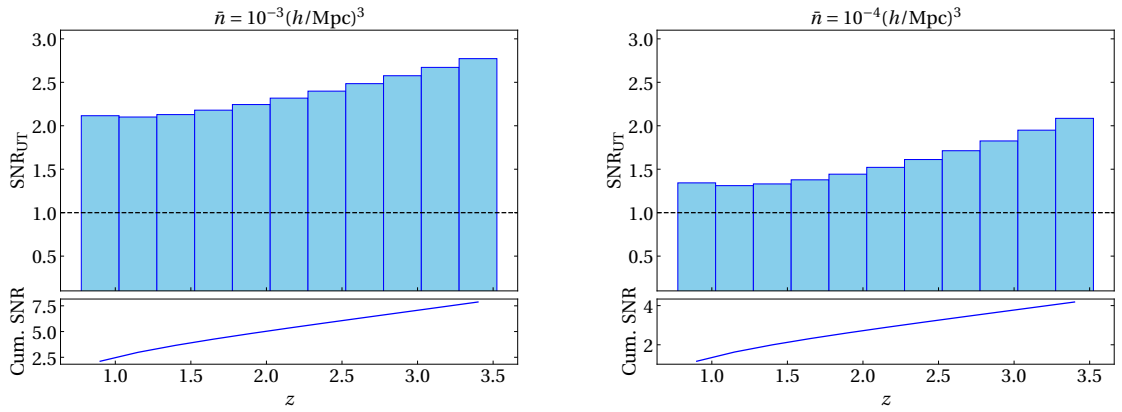

Figure 20: SNR of UT contributions for bins centered at mean redshifts z, for $\bar{n} = 10^{-3}\,(h/\mathrm{Mpc})^3$ (*left*) and $\bar{n} = 10^{-4}\,(h/\mathrm{Mpc})^3$ (*right*), using $b_A = (2+z)^{0.5}$ and $b_B = (0.5+z)^{1.2}$. The bottom panel shows the cumulative SNR when including higher z bins.

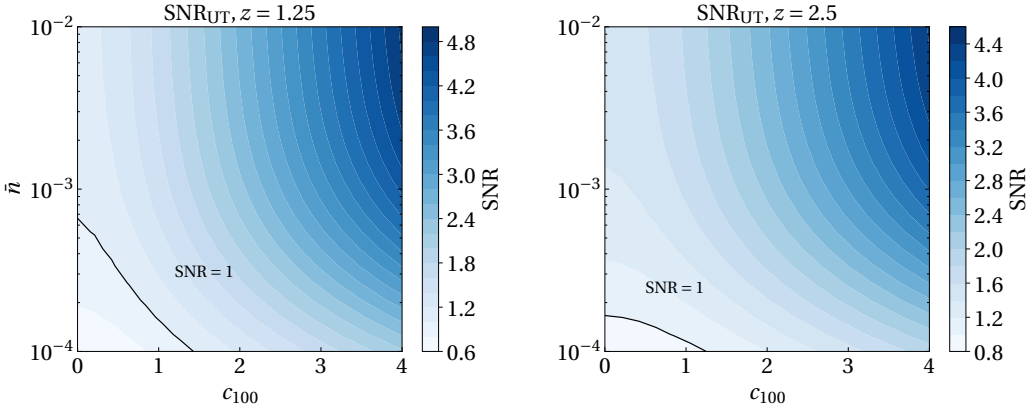

Figure 21: SNR of UT corrections, at $\bar{z} = 1.25$ (*left*) and $\bar{z} = 2.5$ (*right*), as a function of the mean number density of galaxies $\bar{n}$ and of the c_{100} coefficient, which depends on the difference between the biases and on their derivatives.

$(2+z)^{0.5}$ and $b_B = (0.5+z)^{1.2}$, and bins with the same comoving volume of $V = (3800\,\mathrm{Mpc}/h)^3$.

As an example, one can investigate a Stage IV-*like* scenario, combining two redshift bins in the range $1 < z < 2$; in this case, one obtains a signal-to-noise-ratio for the UT corrections SNR ~ 3. As shown explicitly in Sec. IV C, our flat-sky model can be fully trusted at $z \gtrsim 1$ where the curvature of the sky can be neglected in the theoretical modeling. Furthermore, UT corrections are sensitive to the biases of the two tracers and their redshift evolution through the coefficient c_{100}, which is in general a quantity of order one. This is illustrated in Figure 21, where the dependence of the SNR on c_{100} and on the number density of galaxies is shown, for redshift bins centered at $z = 1.25$ and $z = 2.5$. Finally, Doppler terms source a first-order real unequal-time contribution to the standard equal-time power spectrum, which is not vanishing also in the single-tracer case. The expression of this correction can be found in Sec. **??**. However, it can be shown that UT corrections in the single-tracer case are negligible for any realistic Stage IV-*like* scenario.

If one considers only the Kaiser RSD kernel, UT corrections vanish at first order in the single-tracer scenario. This is not the case when Doppler terms are also included; the first-order (for the Gaussian

case) UT correction for a single tracer is purely real and reads

$$c_{1,\mathrm{ST}}(k,\mu,\bar{z}) = \frac{i\mu f b\alpha}{k\chi}\left[\frac{b'}{b} - \frac{\alpha'}{\alpha} + \left(\frac{1}{H\chi} - \frac{f'}{f}\right)\right] + \frac{i\mu^3 f^2\alpha}{k\chi}\left(\frac{1}{H\chi} - \frac{\alpha'}{\alpha}\right) + \frac{i\mu f\alpha\mathcal{A}}{k^3\chi}\left[\left(\frac{1}{H\chi} - \frac{f'}{f}\right) - \frac{\alpha'}{\alpha} + \frac{\mathcal{A}'}{\mathcal{A}}\right], \tag{122}$$

while the non-Gaussian part is

$$c_{1,\mathrm{ST,NG}}(k,\mu,\bar{z}) = \frac{3f_{\mathrm{NL}}\Omega_{m,0}H_0^2}{2T(k)D_{\mathrm{md}}(z)}\frac{f\alpha b_\phi}{H^3\chi}\left[\frac{b'_\phi}{b_\phi} + \frac{f}{1+z} + \left(\frac{1}{H\chi} - \frac{f'}{f}\right) - \frac{\alpha'}{\alpha}\right]. \tag{123}$$

The SNR of the UT correction in the single-tracer case can be calculated in the same fashion as in Equation (119). However one can show that for Stage IV-like surveys the SNR $\ll 1$ making such corrections completely negligible.

V Beyond Linear Theory

In the introduction, we emphasized that one of the key advantages of studying the large-scale structure of the Universe is the access it provides to an extensive number of modes, which in turn can significantly enhance the precision of cosmological parameter measurements. It is evident, then, that if we consider only linear scales (like we did in previous chapters) we are wasting a lot of information encoded in LSS observables, so that a proper theoretical treatment of nonlinearities is needed.

Since Zel'dovich proposed a first model for nonlinear gravitational evolution of cosmological fluctuations in 1970 [135], lots of efforts have been carried out in order to refine the treatment of non-nonlinearities in galaxy clustering. In the following, we will briefly review some of the modern developments of this field and schematize the formalism that will be needed in order to proceed with the following chapter: the so-called Effective Field Theory of Large Scale Structure (EFTofLSS).[11]. For a more extensive treatment, see e.g., [25, 139–141]. Similar approaches to perturbation theory were also proposed in [142–147].

The description of the dark matter distribution as a fluid is not, in principle, justified a priori. Cold dark matter (CDM) is collisionless, with a mean free path that is effectively infinite, and therefore does not naturally exhibit the coherent motion characteristic of a single-stream regime. Gravitational instability further leads to shell crossing, where multiple dark matter streams overlap at the same spatial location, clearly signaling the breakdown of the fluid approximation.

Nevertheless, the fluid description of CDM is widely employed, supported by two main arguments. The first is that the typical displacement of dark matter particles remains finite as a consequence of their mutual gravitational interaction. This fact can be demonstrated schematically as follows.

Consider the initial (very deep in the matter domination era) Lagrangian coordinate of a dark matter particle $\boldsymbol{q}$. The mapping to its final Eulerian coordinate at time t is:

$$\boldsymbol{x}(t) = \boldsymbol{q} + \Phi(\boldsymbol{x}, t)\,, \tag{124}$$

where $\Phi(\boldsymbol{x}, t)$ is the displacement field. In the Zel'dovich approximation, we have that, in Fourier space

$$\delta(\boldsymbol{k}, t) \sim i\boldsymbol{k}/k^2\Phi(\boldsymbol{k}, t)\,, \tag{125}$$

so that the variance of the displacement field is:

$$\left\langle \Phi(\boldsymbol{x}, z)^2 \right\rangle \sim \int \frac{dq}{2\pi^2} P(q, z) \sim D(z)^2 36(\mathrm{Mpc}/h)^2\,, \tag{126}$$

where in the last equality we assumed a ΛCDM universe. This means that in our universe dark matter particles have been displaced by structure growth by roughly $\left\langle \Phi(\boldsymbol{x}, z)^2 \right\rangle^{1/2} \lesssim 6\mathrm{Mpc}/h$ in comoving coordinates on average. This means that if we smooth the dark matter field over the typical displacement scale we will see a coherent motion which can be reasonably treated as a fluid. The second reason is that dark matter particles form structures like dark matter halos and filaments. Inside these structures, clearly the fluid description breaks down due to very complicated non-linear gravitational interactions. However, if one smooths the dark matter field over the typical size of such structures then the dynamics of the smoothed field simplify and can be modeled with the fluid equations of motion.[12] These two observations seem to suggest that the correct way to proceed is to model the *smoothed* dark matter field,

[11] In this thesis we consider only Eulerian Perturbation Theory. For a Lagrangian description see for example [136–138] As we will show more in detail in later sections, the two formulations provides equivalent results for the scale of interest.

[12] The typical size of dark matter halos is often taken to be their Lagrangian radius, which for halos of mass $10^8 - 10^{14} M_\odot$ is $\lesssim 10$Mpc. For filaments the situations is more complicated as the definition of a "typical size" is a non trivial task as, for example, their are not spherically symmetric structures and their shape strongly depends on redshift. However, even if the precise number depends on the web-finder algorithm, their typical radius is often taken to be ~ 1Mpc at $z = 0$ [148, 149].

and not the field itself. To corroborate this intuition even more, in what follows we provide two simple examples showing that if the dark matter field is studied without applying a smoothing, one encounters some theoretical inconsistencies.

In the first part of this thesis, we have derived the perturbed continuity, Euler, and Poisson equations in an expanding universe for an ideal fluid, which can be written in comoving coordinates $\boldsymbol{x}$ and conformal time as

$$\begin{aligned} \partial_\tau \delta(\boldsymbol{x},\tau) + \nabla\cdot[(1+\delta(\boldsymbol{x},\tau))\boldsymbol{u}(\boldsymbol{x},\tau)] = 0\,, \\ \partial_\tau \boldsymbol{u}(\boldsymbol{x},\tau) + \boldsymbol{u}(\boldsymbol{x},\tau)\cdot\nabla\boldsymbol{u}(\boldsymbol{x},\tau) + \mathcal{H}(\tau)\boldsymbol{u} = \nabla\phi(\boldsymbol{x},\tau)\,, \\ \nabla^2\phi(\boldsymbol{x},\tau) = \frac{3}{2}\mathcal{H}^2(\tau)\Omega_m(\tau)\delta(\boldsymbol{x},\tau)\,. \end{aligned} \tag{127}$$

Going to Fourier space, we have

$$\frac{\partial\delta(\mathbf{k},\tau)}{\partial\tau} + \theta(\mathbf{k},\tau) = -\int\frac{d^3q}{(2\pi)^3}\alpha(\mathbf{q},\mathbf{k}-\mathbf{q})\theta(\mathbf{q},\tau)\delta(\mathbf{k}-\mathbf{q},\tau)\,, \tag{128}$$

$$\frac{\partial\theta(\mathbf{k},\tau)}{\partial\tau} + \mathcal{H}\theta(\mathbf{k},\tau) + \frac{3}{2}\Omega_m\mathcal{H}^2\delta(\mathbf{k},\tau) = -\int\frac{d^3q}{(2\pi)^3}\beta(\mathbf{q},\mathbf{k}-\mathbf{q})\theta(\mathbf{q},\tau)\theta(\mathbf{k}-\mathbf{q},\tau)\,, \tag{129}$$

where α and β are the mode-coupling kernels defined as:

$$\alpha(\mathbf{q}_1,\mathbf{q}_2) = \frac{(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{q}_1}{q_1^2}\,, \tag{130}$$

$$\beta(\mathbf{q}_1,\mathbf{q}_2) = \frac{|\mathbf{q}_1+\mathbf{q}_2|^2(\mathbf{q}_1\cdot\mathbf{q}_2)}{2q_1^2q_2^2}\,, \tag{131}$$

and we defined the peculiar velocity divergence $\theta = \nabla\cdot\boldsymbol{u}$. Recall that $\boldsymbol{u}$ is pure gradient in GR in the absence of dissipative forces or sources of primordial vorticity, as explained in earlier sections. One possible way to model mildly non-linear scales is via a perturbative expansion of the dark matter overdensity field δ and velocity divergence θ. For perturbation theory to make sense, the quantities which are expanded need to be small. However, if we calculate the variance of δ we get

$$\left\langle\delta^2\right\rangle = \int\frac{dk}{2\pi^2}k^2P(k)\,, \tag{132}$$

which is clearly divergent. This is the first inconsistency that one encounters in Standard Perturbation Theory (SPT). Let us ignore this and proceed in our perturbative expansion. One expands the density and velocity fields as

$$\delta(\mathbf{k},\tau) = \sum_{n=1}^{\infty}\delta^{(n)}(\mathbf{k},\tau) = \sum_{n=1}^{\infty}a^n(\tau)\delta^{(n)}(\mathbf{k})\,, \tag{133}$$

$$\theta(\mathbf{k},\tau) = -\sum_{n=1}^{\infty}\theta^{(n)}(\mathbf{k},\tau) = -\mathcal{H}f(\tau)\sum_{n=1}^{\infty}a^n(\tau)\theta^{(n)}(\mathbf{k})\,, \tag{134}$$

where the n-th order solution contains n copies of the linear solution, $\delta^{(1)}$. In the second equality, we have assumed separability in time and space: this turns out to be an excellent approximation in practice (and exact for Einstein de-Sitter universes where $f = \Omega_m = 1$) as in a ΛCDM universe $f^2 \sim \Omega_m$ at all

times. Plugging Equations (133) in the equations of motion one obtains the n-th order solution

$$\begin{aligned}\delta^{(n)}(\mathbf{k}) &= \int \frac{d^3q_1\cdots d^3q_n}{(2\pi)^{3(n-1)}}\delta_D(\mathbf{k}-\mathbf{q}_{1\cdots n})F_n(\mathbf{q}_1,\ldots,\mathbf{q}_n)\prod_{i=1}^{n}\delta^{(1)}(\mathbf{q}_i)\,,\\ \theta^{(n)}(\mathbf{k}) &= \int \frac{d^3q_1\cdots d^3q_n}{(2\pi)^{3(n-1)}}\delta_D(\mathbf{k}-\mathbf{q}_{1\cdots n})G_n(\mathbf{q}_1,\ldots,\mathbf{q}_n)\prod_{i=1}^{n}\theta^{(1)}(\mathbf{q}_i)\,,\end{aligned} \tag{135}$$

where δ_D is the Dirac delta function ensuring momentum conservation, and F_n, G_n are the symmetrized perturbation theory kernels that encode the non-linear interactions. The kernels up to third order are given by

$$F_2(\mathbf{p}_1,\mathbf{p}_2) = \frac{5}{7}\alpha(\mathbf{p}_1,\mathbf{p}_2) + \frac{2}{7}\beta(\mathbf{p}_1,\mathbf{p}_2)\,, \tag{136}$$

$$G_2(\mathbf{p}_1,\mathbf{p}_2) = \frac{3}{7}\alpha(\mathbf{p}_1,\mathbf{p}_2) + \frac{4}{7}\beta(\mathbf{p}_1,\mathbf{p}_2)\,, \tag{137}$$

$$\begin{aligned}F_3(\mathbf{p}_1,\mathbf{p}_2) = {}& \frac{1}{18}\left[7\alpha(\mathbf{p}_1,\mathbf{p}_{23})F_2(\mathbf{p}_2,\mathbf{p}_3) + 2\beta(\mathbf{p}_1,\mathbf{p}_{23})G_2(\mathbf{p}_2,\mathbf{p}_3)\right]\\ &+ \frac{G_2(\mathbf{p}_1,\mathbf{p}_2)}{18}\left[7\alpha(\mathbf{p}_{12},\mathbf{p}_3) + 2\beta(\mathbf{p}_{12},\mathbf{p}_3)\right]\,,\end{aligned} \tag{138}$$

$$\begin{aligned}G_3(\mathbf{p}_1,\mathbf{p}_2) = {}& \frac{1}{18}\left[3\alpha(\mathbf{p}_1,\mathbf{p}_{23})F_2(\mathbf{p}_2,\mathbf{p}_3) + 6\beta(\mathbf{p}_1,\mathbf{p}_{23})G_2(\mathbf{p}_2,\mathbf{p}_3)\right]\\ &+ \frac{G_2(\mathbf{p}_1,\mathbf{p}_2)}{18}\left[3\alpha(\mathbf{p}_{12},\mathbf{p}_3) + 6\beta(\mathbf{p}_{12},\mathbf{p}_3)\right]\,.\end{aligned} \tag{139}$$

Equipped with the explicit form of the kernels one can compute the 1-loop corrections to the linear power spectrum:

$$\begin{aligned}&\left\langle\delta^{(2)}(\boldsymbol{k}_1,\tau)\delta^{(2)}(\boldsymbol{k}_2,\tau)\right\rangle = \\ &= D^4(\tau)\int\frac{d^3q_{1\ldots4}}{(2\pi)^9}\delta_D(\boldsymbol{k}_1-\boldsymbol{q}_{12})\delta_D(\boldsymbol{k}_2-\boldsymbol{q}_{34})F_2(\boldsymbol{q}_1,\boldsymbol{q}_2)F_2(\boldsymbol{q}_3,\boldsymbol{q}_4)\left\langle\delta^{(1)}(\boldsymbol{q}_1)\delta^{(1)}(\boldsymbol{q}_2)\delta^{(1)}(\boldsymbol{q}_3)\delta^{(1)}(\boldsymbol{q}_4)\right\rangle\\ &= 2D^4(\tau)\int\frac{d^3q_{1\ldots4}}{(2\pi)^3}\delta_D(\boldsymbol{k}_1-\boldsymbol{q}_{12})\delta_D(\boldsymbol{k}_2-\boldsymbol{q}_{34})F_2(\boldsymbol{q}_1,\boldsymbol{q}_2)F_2(\boldsymbol{q}_3,\boldsymbol{q}_4)P_{11}(q_1)P_{11}(q_2)\delta_D(\boldsymbol{q}_{13})\delta_D(\boldsymbol{q}_{24})\\ &= 2D^4(\tau)\int\frac{d^3q}{(2\pi)^3}F_2^2(\boldsymbol{q},|\boldsymbol{k}-\boldsymbol{q}|)P(q)P(|\boldsymbol{k}-\boldsymbol{q}|)\delta_D(\boldsymbol{k}_1+\boldsymbol{k}_2) = P_{22}(\boldsymbol{k},\tau)\delta_D(\boldsymbol{k}_1+\boldsymbol{k}_2)\,,\end{aligned} \tag{140}$$

and in the same fashion

$$\begin{aligned}&2\left\langle\delta^{(1)}(\boldsymbol{k}_1,\tau)\delta^{(3)}(\boldsymbol{k}_2,\tau)\right\rangle = \\ &= 2D^4(\tau)\int\frac{d^3q_{1\ldots4}}{(2\pi)^9}\delta_D(\boldsymbol{k}_1-\boldsymbol{q}_1)\delta_D(\boldsymbol{k}_2-\boldsymbol{q}_{234})F_3(\boldsymbol{q}_2,\boldsymbol{q}_3,\boldsymbol{q}_4)\left\langle\delta^{(1)}(\boldsymbol{q}_1)\delta^{(1)}(\boldsymbol{q}_2)\delta^{(1)}(\boldsymbol{q}_3)\delta^{(1)}(\boldsymbol{q}_4)\right\rangle\\ &= 6D^4(\tau)P_{11}(k)\int\frac{d^3q}{(2\pi)^3}F_3(\boldsymbol{q},-\boldsymbol{q},\boldsymbol{k})P(q)\delta_D(\boldsymbol{k}_1+\boldsymbol{k}_2) = P_{13}(\boldsymbol{k},\tau)\delta_D(\boldsymbol{k}_1+\boldsymbol{k}_2)\,.\end{aligned} \tag{141}$$

Before proceeding, one can notice that, for example, the form of the P_{13} loop corrections is such that in the UV limit:

$$P_{13}(k)_{q\gg k} \propto k^2P_{11}(k)\int_{q\gg k}dqP_{11}(q)\,. \tag{142}$$

This expression involves an integral of P_{11} over very small nonlinear scales, where perturbation theory clearly breaks down. This is another inconsistency of the theory, and even if in the actual ΛCDM Universe the integral turns out to converge, this uncontrolled UV behavior does not guarantee that it converges to

the right answer.

A Smoothing a fluid

The inconsistencies that we have encountered so far can be trivially solved by smoothing the field at some large scale, so that the integrals involved in the computation of the variance of the overdensity field and in the loop-integrals can be truncated and perturbativity enforced. Therefore, in order for SPT to make sense the correct way to proceed is definitely to consider *smoothed* fields. Let us now show that, even if we consider a perfect fluid, when we smooth the field on some scale Λ an anisotropic stress energy tensor arises in the equations of motion due to back-reaction of small scale physics. To explicitly prove this, for simplicity, let us consider a perfect fluid in Minkowski space where $\boldsymbol{x}$ and t are space and time coordinates. Let us write down the equations of motion for the smoothed fields. To do this, we simply smooth the equation of motion of the field itself. The smoothed density, gravitational potential and momentum are defined as :

$$\begin{aligned}\rho_\ell(\boldsymbol{x}) &= \int d^3x W_\Lambda(\boldsymbol{x}-\boldsymbol{x}_1)\rho(\boldsymbol{x}_1)\,,\\ \phi_\ell(\boldsymbol{x}) &= \int d^3x W_\Lambda(\boldsymbol{x}-\boldsymbol{x}_1)\phi(\boldsymbol{x}_1)\,,\\ \rho_\ell \boldsymbol{v}_\ell(\boldsymbol{x}) &= \int d^3x W_\Lambda(\boldsymbol{x}-\boldsymbol{x}_1)\rho(\boldsymbol{x}_1)\boldsymbol{v}(\boldsymbol{x}_1)\,,\end{aligned} \tag{143}$$

where we drop the time dependence for simplicity. W_Λ is a smoothing window function and in what follows we use a Gaussian $W_\Lambda \propto e^{-\Lambda^2(x-x')^2/2}$ for simplicity, but the following results remain valid for every choice of window function. The idea is to take Λ so that one avoids strong non-linearities and the fields can be treated as fluids, e.g., $\Lambda \sim 10^{-1}h/\mathrm{Mpc}$. Since we are smoothing the field on large scales, we can truncate the collisionless Boltzmann equation after the two higher-order moments. For the continuity equation we have simply:

$$\partial_t \rho_\ell + \nabla\cdot[(\rho_\ell v^i_\ell)] = 0\,. \tag{144}$$

For the Euler equation, things are a bit more complicated but way more interesting. Even if we start the computation assuming the ideal fluid approximation for the un-smoothed field, an anisotropic energy tensor will arise in the final Euler equation for the smoothed field. This tensor encapsulates the back-reaction of small scale physics into long-wavelength modes, which is unavoidable. This is a direct consequence of the fact that the Euler equation includes products of fields evaluated at the same point in space, and the smoothing operation and multiplication do not commute, i.e., $(XY)_\ell \neq X_\ell Y_\ell$. To understand this more clearly, let us write down the Euler equation of the un-smoothed field, assuming that the stress-energy tensor can be neglected $\sigma_{i,j} \sim 0$.

$$\rho(\partial_t v^i + v^j_l \partial_j v^i) + \rho\partial_i\phi = 0\,. \tag{145}$$

Let us smooth the first two terms

$$\int_{x'} W_\Lambda(|\boldsymbol{x}-\boldsymbol{x}'|)\left[\rho\partial_t v^i(\boldsymbol{x}') + \rho v^j(\boldsymbol{x}')\partial_{j'}v^i(\boldsymbol{x}')\right] = \partial_t(\rho_\ell v^i_\ell) + \int_{x'} W_\Lambda v^i \partial_{j'}(\rho v^j) + \int_{x'} W_\Lambda \rho v^j \partial_{j'} v^i\,, \tag{146}$$

where we used that

$$\rho\partial_t v^i = \partial_t(\rho v^i) - v^i\partial_t\rho = \partial_t(\rho v^i) + v^i\partial_j(\rho v^j)\,, \tag{147}$$

using the continuity equation. Then we integrate the last term by parts (notice that the constant terms vanish at the boundaries for the property of the window function) and using $\partial_j W_\Lambda = -\partial_j W_\Lambda$, we find

$$\partial_t(\rho_\ell v^i_\ell) + \int_{x'} W_\Lambda v^i \partial_{j'}(\rho v^j) + \int_{x'} W_\Lambda \rho v^j \partial_{j'} v^i = \partial_t(\rho_\ell v^i_\ell) + \partial_j \int_{x'} W_\Lambda \rho v^i v^j \,. \tag{148}$$

Let us focus on the second term which involves the smoothing of a product of fields. Each field X can be split into long and short wavelength part:

$$X = X_\ell + X_s \,, \tag{149}$$

so that we have

$$\int_{x'} W_\Lambda \rho v^i v^j = \int_{x'} W_\Lambda \rho v^i_\ell v^j_\ell + 2\int_{x'} W_\Lambda \rho v^i_s v^j_\ell + \int_{x'} W_\Lambda \rho v^i_s v^j_s \,. \tag{150}$$

Since the long wavelength part of the velocity v^i is a slowly-varying function of space, it can be Taylor expanded around $\boldsymbol{x}$,

$$v^i_\ell(\boldsymbol{x}') = v^i_\ell(\boldsymbol{x}) + \partial_j v^i_\ell(\boldsymbol{x}) \cdot (\boldsymbol{x} - \boldsymbol{x}')^j + \frac{1}{2}\partial_k \partial_j v^i_\ell(\boldsymbol{x}) \cdot (\boldsymbol{x} - \boldsymbol{x}')^j (\boldsymbol{x} - \boldsymbol{x}')^k + \cdots \,. \tag{151}$$

To proceed, we use the fact that

$$\partial_{j'} W_\Lambda = -\partial_j W_\Lambda = \Lambda^2 (x - x')^j W_\Lambda \,, \tag{152}$$

and

$$\partial_{i'}\partial_{j'} W_\Lambda = \partial_i \partial_j W_\Lambda = -\Lambda^2 \delta_{ij} W_\Lambda + \Lambda^4 (x - x')^i (x - x')^j W_\Lambda \,. \tag{153}$$

So, considering only terms up to Λ^{-2}, the first term of Equation150 is

$$\int_{x'} W_\Lambda \rho v^i_\ell v^j_\ell \sim \rho_\ell v^i_\ell v^j_\ell + C^{ij}_1 \,, \tag{154}$$

where we defined

$$C^{ij}_1 \equiv \rho_\ell \frac{\partial_k v^i_\ell \partial_k v^j_\ell}{\Lambda^2} - \frac{2\partial_k \rho_\ell \partial_\ell v^{(i}_\ell}{\Lambda^2} v^{j)}_\ell + \rho_\ell v^{(i}_\ell \frac{\partial^2 v^{j)}_\ell}{\Lambda^2} \,. \tag{155}$$

For the second term, we have

$$\int_{x'} W_\Lambda \rho v^i_s v^j_\ell \sim \frac{\partial_k \rho_\ell \partial_k v^{(i}_\ell}{\Lambda^2} v^{j)}_l - \frac{1}{2}\rho_\ell \frac{\partial^2 v^{(i}_\ell}{\Lambda^2} v^{j)}_l = C^{(i,j)} \,, \tag{156}$$

and putting all together we obtain

$$\int_{x'} W_\Lambda \rho v^i v^j \sim \rho_\ell v^i_\ell v^j_\ell + \int_{x'} W_\Lambda \rho v^i_s v^j_s + C^{ij} \,, \tag{157}$$

where

$$C^{ij}_\ell \equiv \rho_\ell \frac{\nabla v^i_\ell \cdot \nabla v^j_\ell}{\Lambda^2} \,, \tag{158}$$

so that the first part of the smoothed Euler Equation becomes

$$\int_{x'} W_\Lambda(|\boldsymbol{x}-\boldsymbol{x}'|)\left[\rho\partial_t v^i(\boldsymbol{x}')+\rho v^j(\boldsymbol{x}')\partial_{j'}v^i(\boldsymbol{x}')\right]=\partial_t(\rho_\ell v^i_\ell)+\partial_j(\rho_\ell v^i_\ell v^j_\ell)+\partial_j[\rho_m v^i_s v^j_s]_\Lambda+\partial_j C^{ij}_\ell\,. \tag{159}$$

Expanding the time derivative using the smoothed continuity equation,

$$\int_{x'} W_\Lambda[\partial_t\rho_m+\partial_j(\rho_m v^j)]=\partial_t\rho_\ell+\partial_j(\rho_\ell v^j_\ell)=0\,, \tag{160}$$

we finally find

$$\partial_t(\rho_\ell v^i_\ell)+\partial_j(\rho_\ell v^i_\ell v^j_\ell)+\partial_j[\rho_m v^i_s v^j_s]_\Lambda+\partial_j C^{ij}_\ell=\rho_\ell\partial_t v^i_\ell+\rho_\ell v^j_\ell\partial_j v^i_\ell+\partial_j[\rho v^i_s v^j_s]_\Lambda+\partial_j C^{ij}_\ell\,. \tag{161}$$

In the right hand side of this equation we have three types of terms. Let us discuss them. The first two terms depend only on the long-wavelength modes and have the same structure as in the standard (un-smoothed) Euler Equation (145). The second term depends directly on the short-scale modes of the field. These modes are very large, strongly coupled, and therefore impossible to keep under control analytically. The third term is a higher derivative term (suppressed by powers of $1/\Lambda\boldsymbol{x}$), which depends only on long wavelength fluctuations. One can show that smoothing the last term $\rho\partial_i\phi$ of the Euler equation gives a similar structure of terms. The crucial point we want to point out in this simple example is that, even if we started assuming the ideal fluid approximation (neglecting any anisotropic stress), when smoothing the equations of motion a non-vanishing stress-energy tensor appears unavoidably in the Euler equation. Such energy tensor stems from the back-reaction of the short scale modes on the long wavelength ones and is formed by a dominant part which depends explicitly on short scale perturbations plus a subdominant part due to higher derivative corrections of long-mode fluctuations.

$$\sigma^{ij}=[\sigma^{ij}_s]_\Lambda+\sigma^{ij}_{\ell,\partial^2} \tag{162}$$

This proves the fact that, when considering smoothed fields, it makes no sense to use the ideal fluid approximation $\sigma_{ij}=0$.

Let us summarize the result we obtained until now. We have seen that to treat dark matter as a fluid we have to consider only large scales. This should be done in order to avoid short modes where very complicated non-perturbative nonlinear effects take place. Such a qualitative intuition is explicitly manifested through some inconsistencies that are present in standard perturbation theory. Hence, in order to stick to large scale only we have to smooth the field over some large scale Λ. We have shown that this process unavoidably implies the presence of an anisotropic stress energy tensor in the equations of motion even if we started with the naive assumption that dark matter can be treated as an ideal fluid. This stress energy tensor encapsulates the back-reaction of short scale fluctuations to long wavelength ones.

However, the explicit form of the anisotropic stress energy tensor depends directly on short scale physics which is very hard to model analytically. To proceed, we can ask ourselves what is the *average* contribution of this short scale dynamics to the long wavelength equations of motion. This strategy is justified since when we compute correlation functions of long wavelength fluctuations, we are taking expectation values and since short wavelength fluctuations are not observed, we can take the expectation value over short-distances directly. Crucially, such expectation value (and so the anisotropic stress-energy tensor) will depend on the long mode since a long wavelength fluctuation changes the particle geodesics on small-scales. Nevertheless, the only observable changes caused by the long wavelength mode on short scales must be proportional only to locally observable quantities given the symmetries of he Universe. So, given all this, what is the form of this averaged stress energy tensor which is allowed by the known symmetries of our Universe (i.e., translational and rotational invariance, Equivalence Principle)? At

lowest order it must be of the form :[13]

$$\left\langle [\sigma_s^{ij}]_\Lambda \right\rangle_\ell = c_1(t)\partial_i\partial_j\phi_\ell + c_2(t)\delta^{i,j}\nabla^2\phi_\ell + ...\mathcal{O}(\delta^2, \partial/\Lambda) + \Delta\sigma^{ij}_{stoch} \tag{163}$$

This expansion has two expansion parameters: $\delta_\ell \sim \nabla^2\phi_\ell$ and k/Λ which suppress the higher derivative terms like $\nabla^2(\partial_i\partial_j\phi_\ell)/\Lambda^2 \sim (k/\Lambda)^2\delta$. Since time translations are spontaneously broken in our Universe, the coefficients $c_1(t), c_2(t)$ are time-dependent. In full generality, the theory is non-local in time, as the short modes evolve with time scales of order H, as do the long modes. Thus, in the most general case, the stress tensor is:

$$\left\langle [\sigma_s^{ij}]_\Lambda \right\rangle_\ell = \int dt' T_1(t,t')\tilde{c}_1(t,t')\partial_i\partial_j\phi(\boldsymbol{x},t,t')_\ell + T_2(t,t')\tilde{c}_2(t,t')\delta^{i,j}\nabla^2\phi(\boldsymbol{x},t,t')_\ell + ... + \Delta\sigma^{ij}_{stoch} \tag{164}$$

where $T_n(t,t')$ are some time dependent unknown kernels. However, since we assume time-space separability for the perturbative solutions δ^n one can always take $D(t)^n$ out of the integrals over momenta and perform the time integral separately. In such a way one redefines:

$$c^n_{1,2}(t) = \int dt' T_{1,2}(t,t')D(t,t')^n\tilde{c}_{1,2}(t,t') \tag{165}$$

and we recover Equation (163), with the caveat that now one has different coefficients for each order in δ. Since, as we will see, the lowest order contribution ($\propto \delta^{(1)}$) is enough to capture the effect of the short scales back-reaction at the level of the 1-loop power spectrum, the time non-locality of the theory is irrelevant for our purpose. While the first two terms in Equation (163) depend only on the long modes, the last term encapsulates the stochasticity of the short modes i.e., due to random realizations of initial conditions short wavelength modes vary between different patches of the universe. Let us neglect it for now, we will return to this term later.

Let us now rewrite the Euler equation in Fourier space in terms of the comoving wavenumber and conformal time and keeping only the lowest order contributions to the anisotropic stress tensor. We have

$$\frac{\partial\theta_\ell(\mathbf{k},\tau)}{\partial\tau} + \mathcal{H}\theta_\ell(\mathbf{k},\tau) + \frac{3}{2}\Omega_m\mathcal{H}^2\delta_\ell(\mathbf{k},\tau) = -\int \frac{d^3q}{(2\pi)^3}\beta(\mathbf{q},\mathbf{k}-\mathbf{q})\theta_\ell(\mathbf{q},\tau)\theta_\ell(\mathbf{k}-\mathbf{q},\tau) - c_s^2(t)k^2\delta^(1)_\ell \tag{166}$$

where since the $c_1(t)$-term and $c_2(t)$-term are both proportional to $\partial_i\partial_j\nabla^2\phi_\ell$ they can be reabsorbed in a unique parameter. Notably, we see that the fact that we are working with a smoothed non-ideal fluid can be accounted for by just introducing a counterterm in the Euler equation. The precise value of the amplitude of this counterterm is unknown and must be fitted from simulations or real data but its $\boldsymbol{k}$ dependence is fixed by the symmetries of the problem. The one loop power spectrum is then:

$$P_{\rm tot} = P_{11} + P_{13} + P_{22} - c_s^2k^2P_{11} \tag{167}$$

Now let us turn back to the stochastic term $\Delta\sigma^{ij}_{\rm stoch}$. We can guess its form by just applying mass and momentum conservation. Consider a region of space of size R and suppose that we have measured $\delta(\boldsymbol{x})$ in this region from some N body simulation and we want to compare it with the one predicted via perturbation theory $\delta_{\rm PT}(\boldsymbol{x})$. If our theory is correct in predicting the long mode then the only difference is due to the stochasticity of the short modes back-reacting on the long one. Let us study the effect of

---

[13] Recall that ϕ and $\partial_i\phi$ are not local obervables due to the Equivalence Principle.

stochasticity on long wave modes in Fourier space. We have

$$\delta_{\rm stoch}(\boldsymbol{k}) = \int d^3x e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \left(\delta(\boldsymbol{x}) - \delta_{\rm PT}(\boldsymbol{x})\right) \sim \int d^3x (1 - i\boldsymbol{k}\cdot\boldsymbol{x} + \boldsymbol{O}(\boldsymbol{k})^2 + ...) \left(\delta(\boldsymbol{x}) - \delta_{\rm PT}(\boldsymbol{x})\right) \propto \boldsymbol{k}^2 \tag{168}$$

where we used the fact that we can expand the exponential as we are interested in long wavelength modes $k \ll 1/R$. We see that the first two terms vanish for mass and momentum conservation ($\boldsymbol{x}\delta$ measures the center of mass of the system). The power spectrum of the leading order stochastic contribution is then

$$P_{\rm stoch} = Ak^4 \tag{169}$$

where again the value of the amplitude needs to be fitted from simulations or data.

Let us now focus on the fact that, while until now we have considered a fluid smoothed on a particular scale Λ the final theory is independent on the precise value of this smoothing scale. To see this, note that the two counterterms discussed so far have the same structure of the problematic UV part of the loop integrals

$$\begin{aligned} P_{13,q\gg k} &= c_1(\Lambda)k^2 P_{11}(k) \\ P_{22,q\gg k} &= c_2(\Lambda)k^4 \end{aligned} \tag{170}$$

where c_1 and c_2 are just some amplitudes that depend on the smoothing scale Λ. This implies the following fact. As we vary Λ, the amplitudes c_1 and c_2 change, but also the counterterms will change accordingly since we are including fewer (or more) small-scale modes in the anisotropic tensor. This means that inside the UV counterterms there are parts that nicely cancel the problematic UV part of the loops, regardless of the precise value of Λ. This makes the theory Λ-*independent* as one should expect.

Finally, let us now briefly schematize the relative importance of the two counterterms. To do this, we imagine that our Universe is described by a power law linear power spectrum $P_{11} \sim \frac{1}{k_{\rm NL}^3}\left(\frac{k}{k_{\rm NL}}\right)^n$ where we take $k_{\rm NL} \sim 0.3h/{\rm Mpc}$. Then the dimensionless power spectrum, which is related to the variance of the overdensity field in configuration space, is:

$$\Delta^2_{\rm lin} = \frac{k^3}{2\pi^2} P_{11} \sim \left(\frac{k}{k_{\rm NL}}\right)^{3+n} \tag{171}$$

Then the loop terms scale like

$$\Delta^2_{\rm L-loop} \sim \left(\frac{k}{k_{\rm NL}}\right)^{(3+n)(1+L)} \tag{172}$$

while the c_s^2 counterterm

$$\Delta^2_{c_s^2} \sim \left(\frac{k}{k_{\rm NL}}\right)^{(5+n)} \tag{173}$$

and the stochastic counterterm:

$$\Delta^2_{\rm stoch} \sim \left(\frac{k}{k_{\rm NL}}\right)^{7} \tag{174}$$

Of course we know that our Universe is not a power law universe, but at mildly non-linear scales, where the variance gets the bulk of the contribution, it turns out that in ΛCDM, the power spectrum can be well approximated by $P \propto \left(\frac{k}{k_{\rm NL}}\right)^{-1.5}$. So substituting $n = -1.5$ into the previous scaling relations and for

$k < k_{\rm NL}$ we have the hierarchy

$$\Delta^2_{\rm lin} \gg \Delta^2_{1-\rm loop} > \Delta^2_{c_s^2} \gg \Delta^2_{2-\rm loop} \gg \Delta^2_{\rm stoch} \tag{175}$$

This implies that at the level of the one-loop power spectrum, the stochastic term can be safely neglected so that our theoretical model for the matter power spectrum is given by Equation (167). Obviously, as we go to higher k, 2-loop contributions and stochastic terms start to become important so that the 1-loop model is not reliable anymore in describing the dark matter overdensity field. However, in all that follows, we will always analyze a range of scales for which there is no need to go beyond 1-loop order , given the level of precision of all practical applications in this thesis.

B *Bulk flows and IR modes*

Let us now consider the IR limit of our theory. Here, we follow the treatment outlined in [150, 151] but equivalent results can be obtained using the formalism described in [152, 153]. The IR limit of the two separate contributions of the one-loop is

$$\begin{aligned} P_{13}(k)|_{q\ll k} &\sim \frac{k^2}{3}\int_{q\ll k}\frac{P_{11}}{2\pi^2}(q) \\ P_{22}(k)|_{q\ll k} &\sim -\frac{k^2}{3}\int_{q\ll k}\frac{P_{11}}{2\pi^2}(q) \end{aligned} \tag{176}$$

We see that both terms are proportional to the variance of the displacement fields in the IR domain (recall that $\left\langle \Phi^2 \right\rangle = \int P(q,t)/2\pi^2$ and that the integral is dominated by the large scale contributions). In fact, what Equations (176) physically represent, are the responses of each individual term in the 1-loop power spectrum to large scale bulk flows, which displace particles along the gravitational potential. This displacement is essentially due to the free fall of the dark matter overdensity along the long mode ($q \ll k$) gravitational potential. Thus, due to the equivalence principle, such displacement cannot have an impact on locally observable quantities and this is the physical reason why the contributions in Equations (176) exactly cancel in the final power spectrum, which is the actual observable quantity.

However, as we will show, this result holds true only if the linear power spectrum is purely scale invariant. In fact, the presence of the BAO scale in the 2PCF leads to IR corrections due to large scale bulk flows, which need to be kept resummed in the model for the power spectrum. Intuitively, this fact can be understood as follows. In the initial conditions, the BAO feature is a sharp oscillation induced by the preferred clustering of matter on distances of $\ell_{BAO} \sim 100{\rm Mpc}/h$. As non-linear gravitational evolution occurs, galaxies shift relative to their initial positions due to large-wavelength bulk flows; while the scale invariant part of the power spectrum remains unchanged due to the equivalence principle, any formerly sharp phenomenon gets blurred out on the scale of displacement scale Φ. Such displacement scale, as we have seen, is of the same order as the width of the BAO peak ($\sigma_{BAO} \sim 10{\rm Mpc}/h$) so the effect of the non-linear gravitational evolution on the late times BAO amplitude is sizable and must be taken into account in the modeling. Importantly, such blurring of the peak comes only from modes q which are

$$\ell^{-1}_{BAO} \lesssim q \ll \sigma^{-1}_{BAO} \tag{177}$$

In fact, modes that are much larger than σ^{-1}_{BAO} are just UV contributions coming from small scale physics which are not interesting to investigate in our case as we are interested in the effect of large scale bulk flows on the BAO peak. In any case, their contributions will be subdominant as their wavelength is way smaller than the size of the width of the BAO. On the other hand, very long wavelength modes with $q \ll \ell^{-1}_{BAO}$ do not have a locally observable imprint on the BAO peak. In fact, in this situation we can imagine our system (represented by the region of the BAO peak) as a free-falling system along a

long wavelength gravitational potential mode. Due to the equivalence principle, this reduces only to a coherent shift of all dark matter particles along the gravitational potential, so the 2PCF, which is the local observable, remains untouched. In order to derive the corrections to the power spectrum due to these large bulk flows, let us imagine two dark matter particles at point $\boldsymbol{x}/2$ and $-\boldsymbol{x}/2$;a gravitational potential mode ϕ_L of the type described in (177) will displace the two particles respectively by $\delta\boldsymbol{x}_1 = \Phi_1$ and $\delta\boldsymbol{x}_2 = \Phi_2$. In the Zel'dovich approximation we have $\Phi_{1,2} \sim i\boldsymbol{k}/k^2\delta(\boldsymbol{k})$. Then, we can calculate the leading order response of the 2PCF correlation function to ϕ_L:

$$\langle\delta(\boldsymbol{x}/2)\delta(-\boldsymbol{x}/2)\rangle\,|_{\phi_L} \simeq \int \frac{d\boldsymbol{k}}{(2\pi^3}P(k)e^{i\boldsymbol{k}\boldsymbol{x}}e^{i\boldsymbol{k}(\Phi_1-\Phi_2)} \tag{178}$$

We are interested in the response of our observable to all possible contributions of the type ϕ_L. So we can calculate the expectation value over L of the previous equation to obtain:

$$\left\langle\langle\delta(\boldsymbol{x}/2)\delta(\boldsymbol{x}/2)\rangle_{\phi_L}\right\rangle_L \simeq \int \frac{d\boldsymbol{k}}{(2\pi^3}P(k)e^{i\boldsymbol{k}\boldsymbol{x}}\left\langle e^{i\boldsymbol{k}(\Phi_1-\Phi_2)}\right\rangle \simeq \int \frac{d\boldsymbol{k}}{(2\pi^3}P(k)e^{i\boldsymbol{k}\boldsymbol{x}}e^{-\frac{1}{2}\left\langle k^ik^j(\Phi_1-\Phi_2)_i(\Phi_1-\Phi_2)_j\right\rangle} \tag{179}$$

where in the last equality we have used the cumulant theorem

$$\langle e^{-iX}\rangle = \exp\left(\sum_{n=0}^{\infty}\frac{(-i)^n}{n!}\langle X^n\rangle_c\right)\,. \tag{180}$$

Let us evaluate the argument of the exponential.

$$\begin{aligned}-\frac{1}{2}\left\langle k^ik^j(\Phi_1-\Phi_2)_i(\Phi_1-\Phi_2)_j\right\rangle &= \frac{1}{2}k^ik^j\int\frac{d\boldsymbol{q}_{1,2}}{(2\pi)^6}\frac{q_{1,i}q_{2,j}}{q_1^2q_2^2}\left\langle\delta(\boldsymbol{q}_1)\delta(\boldsymbol{q}_2)\right\rangle\times\\ &\left(e^{i(\boldsymbol{q}_1+\boldsymbol{q}_2)\boldsymbol{x}/2}+e^{i(-\boldsymbol{q}_1-\boldsymbol{q}_2)\boldsymbol{x}/2}-e^{i(\boldsymbol{q}_1-\boldsymbol{q}_2)\boldsymbol{x}/2}-e^{i(-\boldsymbol{q}_1+\boldsymbol{q}_2)\boldsymbol{x}/2}\right)\\ &= -k^ik^j\int\frac{d\boldsymbol{q}}{(2\pi)^3}\frac{q_iq_j}{q^4}P_{11}(q)\left(1-e^{i\boldsymbol{q}\boldsymbol{x}}\right)\end{aligned} \tag{181}$$

now using the following relations

$$\begin{aligned}e^{i\boldsymbol{q}\boldsymbol{x}} &= 4\pi\sum_{\ell m}Y_{\ell m}(\hat{\boldsymbol{q}})Y^*_{\ell m}(\hat{\boldsymbol{x}})j_\ell(qx)i^\ell\\ L_\ell(\hat{\boldsymbol{q}}\hat{\boldsymbol{x}}) &= \frac{4\pi}{2\ell+1}\sum_{m=-\ell}^{\ell}Y(\hat{\boldsymbol{q}})_{\ell m}Y^*_{\ell m}(\hat{\boldsymbol{x}})\end{aligned} \tag{182}$$

we can write

$$\begin{aligned}&-k^2\int\frac{dq}{(2\pi)^3}P_{11}(q)\int d\Omega_q(\hat{\boldsymbol{k}}\hat{\boldsymbol{q}})^2(1-e^{i\boldsymbol{q}\boldsymbol{x}})\\ &= -k^2\int\frac{dq}{(2\pi)^3}P_{11}(q)\left[\frac{4\pi}{3}-\int\left(\frac{1}{3}+\frac{2}{3}L_2(\hat{\boldsymbol{q}}\boldsymbol{k})\right)4\pi\sum_{\ell m}Y_{\ell m}(\hat{\boldsymbol{q}})Y^*_{\ell m}(\hat{\boldsymbol{x}})j_\ell(qx)i^\ell d\Omega_k\right]\\ &= -k^2\int\frac{dq}{(2\pi)^3}P_{11}(q)\left[\frac{4\pi}{3}-\frac{4\pi}{3}j_0(qx)-\int\frac{2}{3}\frac{4\pi}{5}\sum_{m=-2}^{2}Y(\hat{\boldsymbol{q}})_{20}Y^*_{20}(\hat{\boldsymbol{k}})4\pi\sum_{\ell m}Y_{\ell m}(\hat{\boldsymbol{q}})Y^*_{\ell m}(\hat{\boldsymbol{x}})j_\ell(qx)i^\ell d\Omega_k\right]\\ &= -k^2\int\frac{dq}{(2\pi)^2}P_{11}(q)\left[\frac{2}{3}-\frac{2}{3}j_0(qx)+\frac{4}{3}L_2(\hat{\boldsymbol{k}}\hat{\boldsymbol{x}})j_2(qx)\right]\\ &\simeq -k^2\int\frac{dq}{6\pi^2}P_{11}(q)\left[1-j_0(q\ell_{BAO})+2j_2(q\ell_{BAO})\right]\end{aligned} \tag{183}$$

where in the last equality we assumed that corrections due to relative displacements are relevant only on the BAO scales (so $\boldsymbol{x} \sim \ell_{BAO}$). Before writing the full formula for the response of the correlation function let us make two more comments. Again, since the corrections due to large bulk flows cancel exactly on the scale invariant part of the power spectrum (called the broadband) due to the equivalence principle, such contributions must act only on the "wiggly" part of the power spectrum $P_w(k)$. Moreover, since in this whole argument we are considering modes that satisfy (177), we must cut the integral over q in the above equation to some short scale modes Λ. In practice, the result of the integral is not so sensitive to the precise value of Λ, yielding $\sim 10\%$ difference, but we will return to this point in next sections. Having said all of this, the final formula for the correlation function corrected for large bulk flow displacement is:

$$\tilde{\xi}(\boldsymbol{x}) \simeq \int \frac{d\boldsymbol{k}}{(2\pi)^3} e^{i\boldsymbol{k}\boldsymbol{x}} (P_{nw}(k) + P_w(k) e^{-k^2\Sigma^2}) \tag{184}$$

where P_{nw} is the smooth part of the power spectrum and

$$\Sigma^2 = \int^{\Lambda} \frac{dq}{6\pi^2} P_{11}(q) \left[1 - j_0(q\ell_{BAO}) + 2j_2(q\ell_{BAO})\right] \tag{185}$$

If we substitute to P the full 1-loop model, going to Fourier space and defining $P_{1-\text{loop}} = P_{13} + P_{22} - c_s^2 k^2 P_{11}$, one gets

$$P_{\text{tot}}^{IR}(k) = P_{11,nw}(k) + P_{1-\text{loop},nw}(k) + P_{11,w}(k)(1 + \Sigma^2 k^2) e^{-k^2\Sigma^2} + P_{1-\text{loop},w}(k) e^{-k^2\Sigma^2} \tag{186}$$

where the extra term $\Sigma^2 k^2 P_{11,w}$ arises from the fact that replacing P_w with the 1-loop power spectrum double-counts the effect of the long modes. In fact, one can show that in the IR limit ($q \ll \Lambda < k$) the response of the wiggly part of the 1-loop power spectrum to a long wavelength mode already contains a term $\sim \Sigma^2 k^2 P_{11}^w$. Equation(186) represents the IR-resummed formula for the matter power spectrum at one loop order which we will assume as the theoretical model for the matter power spectrum henceforth.

C *Biased tracers in real space*

Until this point we have described the behavior of the dark matter overdensity field up to mildly nonlinear scales. However, the simple ansatz $\delta_g = b\delta_m$ to relate the galaxy overdensity to the dark matter one is valid only at linear order. If one wants to trust the theoretical modeling up to mildly-nonlinear scales, a better formula for the mapping between the observed biased tracer and the underlying dark matter field is needed. Such a mapping is called the bias expansion. In this section, we will briefly schematize the process to obtain the formula for the bias expansion up to third order in perturbation theory, needed for the computation of the summary statistics used in this thesis. For a more extensive treatment see [29, 154, 155] and references therein. In what follows we assume that the dark energy component of our universe is represented by the cosmological constant Λ and perfectly Gaussian initial conditions. In this thesis we will also neglect the effect of massive neutrino on structure growth and galaxy bias.[14]

Regardless of how complicated galaxy formation, evolution and baryonic physics are, we know that such phenomena take place at relatively small scales, say inside a region of size R. For scales much larger than this region the only force influencing the galaxy overdensity field is gravity. So, analogously to what we did for the matter field, to stick only to these large scales we can smooth the galaxy field on a scale R. However, for biased tracers, the choice of the precise values of the smoothing radius turns

[14] In the presence of massive neutrinos, the clustering of matter is reduced above the neutrino free streaming wavenumber, leaving a distinct damping feature in the CDM+baryons power spectrum at small scales (see [156] for a review). On the other hand, one can assume that halos form out of cold dark matter and baryons only, i.e., galaxies are biased tracers of the cold dark matter+baryons (CDM+b) field. In this way, the halo mass function is closer to universal and the bias expansion (in terms of $\delta_{\text{CDM+b}}$ has the same form as in the $m_\nu = 0$ case[29, 157, 158]. However, while accurate at linear order, this assumption is not completely correct, as a scale-dependent growth leads to a scale dependent feature in halo bias[159]. While this effect is predicted to reduce the effect of the neutrino damping in the CDM+baryons power spectrum by 20%-40%[160], it is very difficult to detect in N-body simulations due its small amplitude [89, 161–163].

out to be a non-trivial task. In the case of dark matter halos R can be reasonably considered to be the typical Lagrangian size of the halos since their evolution is only governed by gravity. For example, in the case of galaxies things become more complicated. If galaxy properties are entirely dictated by the characteristics of their host dark matter halos—as it is in halo occupation distribution (HOD) models and abundance matching frameworks—then the relevant smoothing scale for these galaxies corresponds directly to that of the host halos. Conversely, if the local galaxy formation rate is significantly influenced by the surrounding radiation environment—such as the ionizing UV background—then the effective smoothing scale may extend to the radiation's absorption length, potentially reaching scales of several hundred megaparsecs[164]. Similarly, thermal feedback from high-energy cosmic-ray cascades, which possess large mean free paths, can also mediate long-range effects on galaxy formation [165]. So one must be careful in choosing the precise value of R depending on the specifications of the tracer analyzed.

As previously stated, in most scenarios, galaxy physics happens in a relatively small region of space such that the bias expansion can be considered local in space. However, there are no reasons to think that such expansion can be local in time. In fact, the properties of a given dark matter halo or galaxy depend on the physics of all its past light-cone through galaxy formation and evolution, mergers, etc. Hence, a general expansion of the galaxy overdensity field can be written as:

$$\delta_g(\boldsymbol{x},t) = \int dt' \sum_N \frac{b_N(t,t')}{N!} \mathcal{O}_N(\boldsymbol{x},t',t,R) \tag{187}$$

where $b_N(t,t')$ are unknown time-dependent bias coefficients. The question now is what are the operators $\mathcal{O}_N(\boldsymbol{x},t',t,R)$. Since we are considering only large scales, the galaxy overdensity field at these scales is governed only by gravitational interaction. Thus, the galaxy field can depend only on local observables in a gravitational field. The symmetries of the problem, which are the Equivalence principle, rotational and translational (in space) invariance, motivate us to write a general expansion up to cubic order:

$$\begin{aligned}\delta_g(\boldsymbol{x}) = \int dt' [&\tilde{b}_1(t,t')\delta(\boldsymbol{x},t,t',R) \\ &+ \frac{\tilde{b}_2(t,t')}{2}\delta^2(\boldsymbol{x},t,t',R) + \tilde{b}_{\mathcal{G}_2}(t,t')\mathcal{G}_2(\boldsymbol{x},t,t',R) \\ &+ \frac{\tilde{b}_3(t,t')}{6}\delta^3(\boldsymbol{x},t,t',R) + \tilde{b}_{\mathcal{G}_3}(t,t')\mathcal{G}_3(\boldsymbol{x},t,t',R) + \tilde{b}_{\mathcal{G}_2\delta}(t,t')\mathcal{G}_2(\boldsymbol{x},t,t',R)\delta(\boldsymbol{x},t,t',R) \\ &+ \tilde{b}_{\Gamma_3}(t,t')\Gamma_3(\boldsymbol{x},t,t',R) - \tilde{c}(t,t')R^2\nabla^2\delta(\boldsymbol{x},t,t',R) + \text{h.o.}] \,,\end{aligned} \tag{188}$$

where δ is the nonlinear matter field (to be computed in perturbation theory up to cubic order) and we have defined the so-called Galileon operators

$$\mathcal{G}_2(\phi) = (\partial_i\partial_j\phi)^2 - (\partial^2\phi)^2 \,, \tag{189}$$

$$\mathcal{G}_3(\phi) = -\partial_i\partial_j\phi\partial_j\partial_k\phi\partial_i\partial_k\phi - \frac{1}{2}(\partial^2\phi)^2 + \frac{3}{2}(\partial_i\partial_j\phi)^2\partial^2\phi \,, \tag{190}$$

$$\Gamma_3(\phi,\phi_v) = \mathcal{G}_2(\phi) - \mathcal{G}_2(\phi_v) \,. \tag{191}$$

where ϕ_v is the velocity potential so that $\nabla^2\phi_v = \theta = \nabla\cdot\boldsymbol{u}$. Let us comment on this expansion. We can see that all operators are constructed via combination and contraction of $\partial_i\partial_j\phi$ which is the local observable due to the symmetries of the problem. Each power n of ϕ includes n copies of δ via Poisson equation, so that the expansion is organized in a perturbative order in δ. Equation(188) contains all possible terms up to third order in δ. However, Equation(188) can be further simplified. In fact, if we assume time

and space separability of the perturbative solution for δ, like it is done in practice using Einstein de Sitter approximation, the N-order solution includes N powers of the growth factor D which can be factorized out of the integral over momenta and reabsorbed in the bias expansion coefficients. Thus, the time integration can be performed separately leading to a simple redefinition of a new unknown bias coefficient at each perturbative order.[15] This is equivalent to writing, up to third order:

$$\begin{aligned}\delta_g(\boldsymbol{x}) = & b_1(t)\delta(\boldsymbol{x},t,R) \\ & + \frac{b_2(t)}{2}\delta^2(\boldsymbol{x},t,R) + b_{\mathcal{G}_2}(t)\mathcal{G}_2(\boldsymbol{x},t,R) \\ & + \frac{b_3(t)}{6}\delta^3(\boldsymbol{x},t,R) + b_{\mathcal{G}_3}(t)\mathcal{G}_3(\boldsymbol{x},t,R) + b_{\mathcal{G}_2\delta}(t)\mathcal{G}_2(\boldsymbol{x},t,R)\delta(\boldsymbol{x},t,R) \\ & + b_{\Gamma_3}(t)\Gamma_3(\boldsymbol{x},t,R) + c(t)R^2\nabla^2\delta(\boldsymbol{x},t,R)\ ,\end{aligned} \tag{192}$$

Let us now focus on the last term of the bias expansion $\propto R^2\nabla^2\delta$. While such terms arise naturally if one naively includes all possible local observable quantities given by the symmetry of the problem, they encode a different physical intuition from the other terms. In fact, they play the role of higher derivative terms that arise from the fact that we are smoothing the field on a scale R in order to consider only large scales where the bias expansion can be treated as local in space. As we showed for the dark matter case, the process of smoothing unavoidably gives rise to higher derivative terms suppressed by powers of k/Λ where Λ is the smoothing scale of the problem. Such terms govern the perturbativity of the theory and the validity of the locality (in space) assumption of the bias expansion. In fact, as the scale at which we are calculating the correlator approaches the smoothing scale the series does not converge anymore and the perturbative treatment breaks down. Interestingly , one can notice that the momentum dependence (k^2) of the leading order higher derivative term is the same as the leading order dark matter counterterm. In principle, these two terms must be kept separated as the smoothing scale is different and they encode two different physical effects (recall that in the dark matter case the expansion was made in terms of powers of k/k_{NL}). However, in practice, since in realistic analyses one considers the amplitudes of the counterterms as free parameters, the two terms can be reabsorbed in a single counterterm $\propto k^2$ in the galaxy power spectrum at one loop.

One simple way to show that such term arises naturally to keep under control non-local effects (in space) is the following. Let us imagine that the theory is not local in space. The operators in the bias expansion must be replaced by functionals. For example, at linear order we have:

$$\delta_g(\boldsymbol{x},t) = \int d\boldsymbol{y}\mathcal{F}(\boldsymbol{y},t)\delta(\boldsymbol{x}+\boldsymbol{y},t) + ... \tag{193}$$

Where we have used the homogeneity of the Universe which dictates that $\mathcal{F}$ must be $\boldsymbol{x}$-independent. Now, if we assume that the main contributions from the integrand come from a small region around $\boldsymbol{x}$ (which eventually includes galaxy formation, baryonic physics processes, etc.) we can expand δ around $\boldsymbol{x}$. This yields

$$\begin{aligned}\delta_g(\boldsymbol{x},t) &= \left(\int d\boldsymbol{y}\mathcal{F}(\boldsymbol{y},t)\right)\delta(\boldsymbol{x},t) + \frac{1}{6}\left(\int d\boldsymbol{y}y^2\mathcal{F}(\boldsymbol{y},t)\right)\nabla^2_{\boldsymbol{x}}\delta(\boldsymbol{x},t) + ... \\ &= b_1\delta + b_{\nabla^2\delta}\nabla^2\delta + ...\ ,\end{aligned} \tag{194}$$

where we took advantage of the isotropy of the kernels. Thus, with this simple example, we showed that higher derivative terms proportional to the Laplacian of δ arise naturally to keep under control

[15] It can be shown that even if one assumes ΛCDM or a quintessence model in the evaluation of the PT kernels, the bias expansion looks local in time up to third order and terms arising from the intrinsic non-locality in time of the theory arise only from the fourth perturbative order[29].

non-local small scale processes in the bias expansion. Furthermore, higher derivative terms include also the contribution from possible velocity bias at small scales (smaller or comparable to R). In fact, the assumption that galaxy and dark matter comove with the same velocity is only valid at large scales where they are both influenced by gravity only and the equivalence principle holds. At small scales, galaxies are strongly affected by baryonic physics so that in general they will have different peculiar velocities compared to the dark matter field.

Having discussed the role of higher derivative terms, there is still another problem to be solved in the bias expansion of Equation192 . To see this, let us consider the following example. The variance of the galaxy overdensity field is (neglecting Galileon operators for simplicity) up to second order:

$$\left\langle \delta_g \right\rangle = b_1 \left\langle \delta \right\rangle + \frac{b_2}{2} \left\langle \delta^2 \right\rangle = \frac{b_2}{2} \int \frac{dq}{2\pi^2} q^2 W_R(q) P(q) = \frac{b_2}{2} \sigma^2(R) \,. \tag{195}$$

This is unsatisfactory for two reasons. Firstly, we know that it must be that $\left\langle \delta_g \right\rangle = 0$ by definition. Secondly, this shows that our observables depend on R in an uncontrolled way. This is clearly unphysical. In order to avoid this R dependence, we can redefine the second order operator as

$$\delta^2 \to \delta^2 - \sigma^2(R) \,. \tag{196}$$

In this fashion, the Λ dependence of the δ^2 operator is reabsorbed by a counterterm, so that the new operator is independent of the smoothing scale. However, the story does not end here. Let us consider the galaxy-matter power spectrum. On large scales this correlator is the simplest way to measure the bias since

$$\left\langle \delta_g(\boldsymbol{k}) \delta(\boldsymbol{k'}) \right\rangle |_{k\to 0} = b_1 \left\langle \delta^1(\boldsymbol{k}) \delta^1(\boldsymbol{k'}) \right\rangle \,. \tag{197}$$

However, while this is true at first order in δ, going at second order leads to a leading contribution of the type

$$\begin{aligned} \frac{b_2}{2} \left\langle \delta^2 \delta^1 \right\rangle &= b_2 \left\langle (\delta^{(1)} \delta^{(2)}) \delta^{(1)} \right\rangle \\ &= \int \frac{d\boldsymbol{q}_{1..4}}{(2\pi)^{12}} \delta_D(\boldsymbol{k} - \boldsymbol{q}_{123}) \delta_D(\boldsymbol{k'} - \boldsymbol{q}_4) W_R(q_1) W_R(|\boldsymbol{q}_{23}|) F_2(\boldsymbol{q}_2, \boldsymbol{q}_3) \left\langle \delta^{(1)}(\boldsymbol{q}_1) \delta^{(1)}(\boldsymbol{q}_2) \delta^{(1)}(\boldsymbol{q}_3) \delta^{(1)}(\boldsymbol{q}_4) \right\rangle \\ &= 4 \int_{q \gg k} \frac{d\boldsymbol{q}}{(2\pi)^6} W_R(q) W_R(|\boldsymbol{p} + \boldsymbol{q}|) F_2(\boldsymbol{k}, \boldsymbol{q}) P(q) P(k) \\ &= \frac{68}{21} \sigma_R^2(k) P(k) \neq b_1 \,, \end{aligned} \tag{198}$$

where we assumed $q \gg \boldsymbol{k}$ since we are interested in very large k while the main contributions to the integral come from short modes $q \sim R$. We see that in order to ameliorate this issue, the operator needs to be further redefined as

$$\delta^2 \to \delta^2 - \sigma_R^2 - \frac{68}{21} \sigma_R^2 \delta \,. \tag{199}$$

Crucially, now a term proportional to δ appears in the renormalized operator so that, in order to be consistent in the perturbative treatment of the galaxy bias expansion oreder by order we have also to redefine a new basis of bias parameters in terms of the renormalized operators. These biased coefficients are the physical ones that actually fit the data. The bias coefficients defined in terms of un-normalized operator are called the "bare" bias coefficients and will never appear in practical analyses as they are un-physical since they depend on the smoothing scale.

One can apply the renormalization arguments discussed so far for higher order correlation functions

and going to higher order in PT to obtain the full set of renormalized operators. However, since lot of contributions are redundant and lots of counterterms have the same momentum dependence so that they can be combined in a single one, it can be shown that at third order in perturbation theory the bias expansion in terms of renormalized operators and renormalized bias parameters is simply given by:[16]

$$\begin{aligned}\delta_g(\boldsymbol{x},t) = b_1(t)\delta(\boldsymbol{x},t) + \frac{b_2(t)}{2}(\delta^2 - \sigma^2(t)) \\ + \frac{b_3(t)}{6}\delta^3(\boldsymbol{x},t) + b_{\mathcal{G}_3}(t)\mathcal{G}_3(\boldsymbol{x},t) + b_{\mathcal{G}_2\delta}(t)\mathcal{G}_2(\boldsymbol{x},t)\delta(\boldsymbol{x},t) \\ + b_{\Gamma_3}(t)\Gamma_3(\boldsymbol{x},t) + b_{\nabla^2\delta}R^2\nabla^2\delta(\boldsymbol{x},t)\ ,\end{aligned} \tag{200}$$

This bias expansion will be used in all that follows to calculate the 1-loop power spectrum of biased tracers in real space. A residual R-dependence is still left in the bias expansion, through the higher derivative terms. However, this is natural since R is a "physical" scale above which the locality in space assumption makes sense. This scale, as previously stated, controls the importance of non-local contributions coming from short modes through the higher derivative terms. However, this R dependence is kept under control in a perturbative sense. This is similar to what happens for the dark matter case where higher derivative contributions to the anisotropic stress-energy tensor were controlled by the $k/k_{\rm NL}$ parameter.

Finally, let us now briefly discuss a potential subtle issue of the galaxy bias expansion discussed so far. Until now, we have assumed that on large scales, where non-gravitational forces can be neglected, baryons and CDM comove and, therefore, can be treated as a single fluid and that the density and velocity differences due to pressure forces and feedback are fully taken into account through the higher-derivative terms in the bias expansion. However, as outlined in [166] this is not completely true for the following reason. At the time of recombination, there is a supersonic relative velocity between baryons and cold dark matter, of order 30km/s due to the fact that baryons suffer Thomson scattering as long as they are tied to the photons, while dark matter travels along geodesics and starts to form gravitational potential wells. When baryons are no longer coupled to the photons, their sound speed drops, but they do not immediately fall into the dark matter potential wells: instead, the relative motion allows baryons to escape out of the wells and suppresses the growth of structures on scales. This fact can in principle generate additional terms in the galaxy bias expansion as the relative velocity between the dark matter and the baryons is a local observable. This effect which can be studied using a multi-fluid description of the underlying matter field and which affects mostly the BAO peak, turns out to give a sub-percent contribution to the total galaxy power spectrum for $k \leq 0.3h$/Mpc and redshift ≤ 2, which represents the main range of interests of current and Stage IV galaxy surveys (see[29] and reference therein).

[16] Note that we used the same notation as in "bare" bias parameters case, but we stress again that now the bias coefficents are the renormalized ones.

VI Conventional analyses in real space and equivalence with field level inference

A Power spectrum, bispectrum and trispectrum

We have now all the ingredients to compute the deterministic part of our model for the halo/galaxy overdensity field. It is given by

$$\delta_g^{\rm model}(\boldsymbol{x}) = b_1\delta(\boldsymbol{x}) + \frac{b_2}{2}(\delta^2(\boldsymbol{x}) - \langle\delta^2\rangle) + b_{\mathcal{G}_2}\mathcal{G}_2(\boldsymbol{x}) + \frac{b_3}{6}\delta^3(\boldsymbol{x}) + b_{\mathcal{G}_3}\mathcal{G}_3(\boldsymbol{x}) + b_{\mathcal{G}_2\delta}\mathcal{G}_2(\boldsymbol{x})\delta(\boldsymbol{x}) \\ + b_{\Gamma_3}\Gamma_3(\boldsymbol{x}) - c_s^2\frac{\nabla^2}{k_{\rm NL}^2}\delta(\boldsymbol{x}) \ , \tag{201}$$

where recall that δ is the nonlinear matter field (to be computed in Eulerian perturbation theory up to cubic order). These terms are sufficient to describe the one-loop power spectrum, tree-level bispectrum and tree-level trispectrum. The model has eight unknown parameters that depend on the halo or galaxy properties and have to be measured from the data. These are

$$\theta_{\rm bias} = \{b_1, b_2, b_{\mathcal{G}_2}, b_3, b_{\mathcal{G}_3}, b_{\mathcal{G}_2\delta}, b_{\Gamma_3}, c_s^2\} \ . \tag{202}$$

Biases and counterterms are (strongly) degenerate with some cosmological parameters and measuring multiple correlation functions is crucial to break these degeneracies. In our analysis, we will focus only on one cosmological parameter which measures the amplitude of the linear density fluctuations A. We define the following parameter

$$\alpha \equiv \frac{A}{A^{\rm fid}} = \frac{\sigma_8}{\sigma_8^{\rm fid}} \ , \tag{203}$$

where the second equality holds if all other cosmological parameters are fixed.

We will do all calculations and the analysis in Fourier space. Given the real space model and parameters, we can write

$$\delta_g^{\rm model}(\boldsymbol{k}, z) = \sum_n \alpha^n D^n(z)\int_{\boldsymbol{q}_1}\cdots\int_{\boldsymbol{q}_n}(2\pi)^3\delta^D(\boldsymbol{k}-\boldsymbol{q}_{1\ldots n})K_n(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n;\theta)\delta_1(\boldsymbol{q}_1)\cdots\delta_1(\boldsymbol{q}_1) \\ - \alpha c_s^2\frac{k^2}{k_{\rm NL}^2}D(z)\delta_1(\boldsymbol{k}) \ , \tag{204}$$

where $\int_{\boldsymbol{q}} \equiv \int d^3\boldsymbol{q}/(2\pi)^3$, $\boldsymbol{q}_{1\ldots n} \equiv \boldsymbol{q}_1 + \cdots \boldsymbol{q}_n$, $D(z) \equiv D_+(z)/D_+(0)$ is the normalized growth factor at redshift z such that $D(0) = 1$, $\delta_1(\boldsymbol{k})$ is the linear density field and K_n are perturbation theory kernels that contain all information about the nonlinear evolution. These kernels are well-known in the literature (see for example [154, 167–169]) and here we just write them down for completeness. They are given as

$$K_1(\boldsymbol{q};\theta) = b_1 \ , \tag{205}$$

$$K_2(\boldsymbol{q}_1,\boldsymbol{q}_2;\theta) = b_1F_2(\boldsymbol{q}_1,\boldsymbol{q}_2) + \frac{b_2}{2} + b_{\mathcal{G}_2}\left(\frac{(\boldsymbol{q}_1\cdot\boldsymbol{q}_2)^2}{q_1^2q_2^2} - 1\right) \ , \tag{206}$$

$$K_3(\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3;\theta) = b_1 F_3(\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3) + \frac{b_2}{3}(F_2(\boldsymbol{q}_1,\boldsymbol{q}_2)+F_2(\boldsymbol{q}_1,\boldsymbol{q}_3)+F_2(\boldsymbol{q}_2,\boldsymbol{q}_3)) + 2b_{\mathcal{G}_2}F_{\mathcal{G}_2}(\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3) + \frac{b_3}{6} + b_{\mathcal{G}_3}F_{\mathcal{G}_3}(\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3) + b_{\Gamma_3}F_{\Gamma_3}(\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3) + \frac{b_{\mathcal{G}_2\delta}}{3}\left(\frac{(\boldsymbol{q}_1\cdot\boldsymbol{q}_2)^2}{q_1^2q_2^2}+\frac{(\boldsymbol{q}_1\cdot\boldsymbol{q}_3)^2}{q_1^2q_3^2}+\frac{(\boldsymbol{q}_2\cdot\boldsymbol{q}_3)^2}{q_2^2q_3^2}-3\right). \tag{207}$$

In these expressions F_2 and F_3 are the standard perturbation theory kernels, and additional terms are defined by

$$F_{\mathcal{G}_2}(\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3) \equiv \frac{1}{3}F_2(\boldsymbol{q}_1,\boldsymbol{q}_2)\left(\frac{((\boldsymbol{q}_1+\boldsymbol{q}_2)\cdot\boldsymbol{q}_3)^2}{|\boldsymbol{q}_1+\boldsymbol{q}_2|^2q_3^2}-1\right) + 2\text{ perm}\,. \tag{208}$$

$$F_{\mathcal{G}_3}(\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3) \equiv -\frac{1}{2}\left(1-\frac{(\boldsymbol{q}_1\cdot\boldsymbol{q}_2)^2}{q_1^2q_2^2}-\frac{(\boldsymbol{q}_1\cdot\boldsymbol{q}_3)^2}{q_1^2q_3^2}-\frac{(\boldsymbol{q}_2\cdot\boldsymbol{q}_3)^2}{q_2^2q_3^2}+2\frac{(\boldsymbol{q}_1\cdot\boldsymbol{q}_2)(\boldsymbol{q}_1\cdot\boldsymbol{q}_3)(\boldsymbol{q}_2\cdot\boldsymbol{q}_3)}{q_1^2q_2^2q_3^2}\right), \tag{209}$$

$$F_{\Gamma_3}(\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3) \equiv -\frac{4}{21}\left(\frac{(\boldsymbol{q}_1\cdot\boldsymbol{q}_2)^2}{q_1^2q_2^2}-1\right)\left(\frac{((\boldsymbol{q}_1+\boldsymbol{q}_2)\cdot\boldsymbol{q}_3)^2}{|\boldsymbol{q}_1+\boldsymbol{q}_2|^2q_3^2}-1\right) + 2\text{ perm}\,. \tag{210}$$

The kernels K_n can be used to compute the one-loop power spectrum, tree-level bispectrum and tree-level trispectrum. We will do that in the next section. Before that, we discuss the second important ingredient which is the stochastic noise.

The galaxy density field in perturbation theory is not fully deterministic, but it also contains stochastic noise which comes from our ignorance of the small-scale modes that determine galaxy formation and that are not a part of observations (or perturbation theory). This noise does not correlate with long-wavelength density fluctuations, but it can be modulated by the long modes. Therefore, the noise field can be written in a similar manner as the bias expansion

$$\epsilon(\boldsymbol{x}) = \epsilon_0(\boldsymbol{x}) + \epsilon_1(\boldsymbol{x})\delta(\boldsymbol{x}) + \frac{1}{2}\epsilon_2(\boldsymbol{x})(\delta^2(\boldsymbol{x})-\langle\delta^2\rangle) + \epsilon_{\mathcal{G}_2}(\boldsymbol{x})\mathcal{G}_2(\boldsymbol{x}) + \dots\,. \tag{211}$$

Note that $\epsilon_i(\boldsymbol{x})$ are non-Gaussian fields that can have significant three-point functions. It is clear from this expression that $\langle\delta\epsilon\rangle = 0$ as expected. The auto power spectrum of the noise does not inherit the "shape" of the cross spectra from δ, δ^2 or $\mathcal{G}_2$ due to the fact that the stochastic fields correlate only in the same spatial point

$$\langle\epsilon_i(\boldsymbol{x})\epsilon_j(\boldsymbol{x}')\rangle \sim c_{ij}\delta^D(\boldsymbol{x}-\boldsymbol{x}')\,. \tag{212}$$

On large scales, the noise power spectrum is approximately constant, with the amplitude approximately equal to the Poisson prediction

$$P_{\epsilon\epsilon}(k) = \frac{c_1}{\bar{n}_g}\,. \tag{213}$$

where c_1 is a free dimensionless parameter of order one. When computing higher order n-point functions, there are several additional terms that may appear, and we discuss them in the next section.

Equipped with the perturbative model for the galaxy overdensity field and the noise model, we are now ready to compute all relevant correlation functions. Since we are working at third order in PT, those are the one-loop power spectrum,tree-level bispectrum and tree-level trispectrum. The power spectrum

of biased tracers in real space up to one-loop corrections is given by

$$P_g(k,z) = b_1^2\alpha^2 P_{\rm lin}(k,z) + b_1^2\alpha^4 P_{13}(k,z) + b_1^2\alpha^4 P_{22}(k,z) - b_1^2 c_s^2\alpha^2 k^2 P_{\rm lin}(k,z) \\ + b_1 b_2 \alpha^4 \mathcal{I}_{\delta^2}(k,z) + \frac{b_2^2}{4}\alpha^4 \mathcal{I}_{\delta^2\delta^2}(k,z) + b_{\mathcal{G}_2}^2\alpha^4 \mathcal{I}_{\mathcal{G}_2\mathcal{G}_2}(k,z) + \frac{b_2 b_{\mathcal{G}_2}}{2}\alpha^4 \mathcal{I}_{\delta^2\mathcal{G}_2}(k,z) \\ + \left(2b_1 b_{\mathcal{G}_2} + \frac{4}{5}b_1 b_{\Gamma_3}\right)\alpha^4 \mathcal{F}_{\mathcal{G}_2}(k,z) + 2b_1 b_{\mathcal{G}_2}\alpha^4 \mathcal{I}_{\mathcal{G}_2}(k,z) + \frac{c_1}{\bar{n}_g}\,, \tag{214}$$

The time dependence of all momentum integrals is $D(z)^4$. The explicit shapes are:

$$\mathcal{I}_{\delta^2}(k) = 2\int_{\boldsymbol{q}} F_2(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q}) P_{\rm lin}(q) P_{\rm lin}(|\boldsymbol{k}-\boldsymbol{q}|) \tag{215}$$

$$\mathcal{I}_{\mathcal{G}_2}(k) = 2\int_{\boldsymbol{q}} \sigma^2(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q}) F_2(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q}) P_{\rm lin}(q) P_{\rm lin}(|\boldsymbol{k}-\boldsymbol{q}|)$$

$$\mathcal{F}_{\mathcal{G}_2}(k) = 4P_{\rm lin}(k)\int_{\boldsymbol{q}} \sigma^2(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q}) F_2(\boldsymbol{k}, -\boldsymbol{q}) P_{\rm lin}(q)$$

$$\mathcal{I}_{\delta^2\delta^2}(k) = 2\int_{\boldsymbol{q}} P_{\rm lin}(q) P_{\rm lin}(|\boldsymbol{k}-\boldsymbol{q}|)$$

$$\mathcal{I}_{\mathcal{G}_2\mathcal{G}_2}(k) = 2\int_{\boldsymbol{q}} (\sigma^2(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q}))^2 P_{\rm lin}(q) P_{\rm lin}(|\boldsymbol{k}-\boldsymbol{q}|)$$

$$\mathcal{I}_{\delta_2\mathcal{G}_2}(k) = 2\int_{\boldsymbol{q}} \sigma^2(\boldsymbol{q}, \boldsymbol{k}-\boldsymbol{q}) P_{\rm lin}(q) P_{\rm lin}(|\boldsymbol{k}-\boldsymbol{q}|)$$

where $\sigma^2(\boldsymbol{k}_1, \boldsymbol{k}_2) = (\boldsymbol{k}_1 \cdot \boldsymbol{k}_2/k_1 k_2)^2 - 1$. The only cubic bias that appears in the galaxy 1-loop power spectrum is b_{Γ_3} as the other third order biases give rise to shapes degenerate with lower order contributions so that they can be renormalized.

Let us further comment on Equation (214). We have explicitly written the dependence of all terms in the loop expansion on the amplitude parameter $\alpha = \frac{\sigma_8}{\sigma_{8,fid}}$ as it will be the cosmological parameter of interest in this section. All nuisance parameters in this expression are functions of time, but for the fixed redshift they can be treated as constants. This form of the power spectrum is particularly convenient since all nuisance and cosmological parameters enter the model just as simple multiplicative factors. At the level of the linear theory, b_1 and α are perfectly degenerate and only the combination $b_1\alpha$ can be measured with very high precision. This degeneracy can be in principle broken only by the loop contributions, but those have much smaller signal-to-noise, suppressed by the variance of the density field at the maximum wavenumber used in the analysis. In addition, since there are several new nuisance parameters at the loop level, the amplitude α is expected to be poorly constrained from the P alone. We will verify this explicitly below.

Let us stress that the correct model for the galaxy power spectrum contains the infrared-resummation, which correctly accounts for the broadening of the BAO peak, as explained in previous sections. However, in this section we will not use it in our baseline analysis for several reasons. First, as we will show we expect that the infrared-resummation cannot significantly affect the power spectrum on scales and redshifts that we use. Second, we will test this simple model against large-volume PT challenge simulations suit and explicitly show the validity of our approximation. Finally, even if this approximation was to miss a small, percent level effects, that would be irrelevant for the conclusions regarding the expected errors form a joint P+B and P+B+T analyses.

Let us now comment on the implementation of the theoretical model for the 1-loop power spectrum. Equation (214) may seem simple and straightforward but in practice if one computes the integrals via brute-force evaluation, the computational cost turns out to be prohibitive for real MCMC analyses. A theoretically clean procedure based on FFTlog methods to optimize the evaluation of the integrals in

the 1-loop power spectrum was presented in [170]. In this thesis, we will always use the mentioned procedure in the implementation of the loop integrals, and we briefly schematize the main ideas behind this method in the following. The starting point is the expansion of the linear power spectrum as

$$P_{lin}(k) = \sum_{m=-N/2}^{N/2} c_m k^{\nu+i\eta_m}\,, \tag{216}$$

where the coefficients and the frequencies are given by

$$c_m = \frac{1}{N}\sum_{l=0}^{N-1} P_{lin}(k_l) k_l^{-\nu} k_{min}^{-i\eta_m} e^{-2\pi i m l/N}\,, \quad \eta_m = \frac{2\pi m}{\log(k_{max}/k_{min})}\,. \tag{217}$$

All the loop integrals in Equations (215) can be rewritten as

$$\int_{\boldsymbol{q}} X(\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q}) P_{\rm lin}(|\boldsymbol{k}-\boldsymbol{q}|) P_{\rm lin}(q) \quad \text{and} \quad \int_{\boldsymbol{q}} Y(\boldsymbol{k},\boldsymbol{q},-\boldsymbol{q}) P_{\rm lin}(q) \tag{218}$$

It can be shown that they can be written, using the FFTLog expansion, as a simple matrix multiplication. For the "22-type" integrals we have

$$P_X(k) = k^3 \sum_{m_1,m_2} c_{m_1} k^{-2\nu_1} \cdot M_X(\nu_1,\nu_2) \cdot c_{m_2} k^{-2\nu_2}\,, \tag{219}$$

while for the "13-type"

$$P_Y(k) = k^3 P_{\rm lin}(k) \sum_{m_1} k^{-2\nu_1} M_Y(\nu_1) \tag{220}$$

for the "13-type" integrals, where the matrices are written above. The matrices are given by

$$M_{22}(\nu_1,\nu_2) = \frac{(\frac{3}{2}-\nu_{12})(\frac{1}{2}-\nu_{12})[\nu_1\nu_2(98\nu_{12}^2-14\nu_{12}+36)-91\nu_{12}]^2+3\nu_{12}+58]}{196\nu_1(1+\nu_1)(\frac{1}{2}-\nu_1)\nu_2(1+\nu_2)(\frac{1}{2}-\nu_2)} I(\nu_1,\nu_2)\,, \tag{221}$$

$$M_{13}(\nu_1) = \frac{1+9\nu_1}{4}\frac{\tan(\nu_1\pi)}{28\pi(\nu_1+1)\nu_1(\nu_1-1)(\nu_1-2)(\nu_1-3)}\,, \tag{222}$$

$$M_{\mathcal{I}_{\delta^2}}(\nu_1,\nu_2) = \frac{(3-2\nu_{12})(4-7\nu_{12})}{17\nu_1\nu_2} I(\nu_1,\nu_2)\,, \tag{223}$$

$$M_{\mathcal{I}_{\mathcal{G}_2}}(\nu_1,\nu_2) = -\frac{(3-2\nu_{12})(1-2\nu_{12})(6+7\nu_{12})}{28\nu_1(1+\nu_1)\nu_2(1+\nu_2)} I(\nu_1,\nu_2)\,, \tag{224}$$

$$M_{\mathcal{F}_{\mathcal{G}_2}}(\nu_1) = -\frac{15\tan(\nu_1\pi)}{28\pi(\nu_1+1)\nu_1(\nu_1-1)(\nu_1-2)(\nu_1-3)}\,, \tag{225}$$

$$M_{\mathcal{I}_{\delta^2\delta^2}}(\nu_1,\nu_2) = 2I(\nu_1,\nu_2)\,, \tag{226}$$

$$M_{\mathcal{I}_{\mathcal{G}_2\mathcal{G}_2}}(\nu_1,\nu_2) = \frac{(3-2\nu_{12})(1-2\nu_{12})}{\nu_1(1+\nu_1)\nu_2(1+\nu_2)} I(\nu_1,\nu_2)\,, \tag{227}$$

$$M_{\mathcal{I}_{\delta^2\mathcal{G}_2}}(\nu_1,\nu_2) = \frac{3-2\nu_{12}}{\nu_1\nu_2} I(\nu_1,\nu_2)\,. \tag{228}$$

where $I(\nu_1,\nu_2)$ is the master integral used in the FFTLog expansion

$$I(\nu_1,\nu_2) = \frac{1}{8\pi^{3/2}}\frac{\Gamma\left(\frac{3}{2}-\nu_1\right)\Gamma\left(\frac{3}{2}-\nu_2\right)\Gamma\left(\nu_{12}-\frac{3}{2}\right)}{\Gamma(\nu_1)\Gamma(\nu_2)\Gamma(3-\nu_{12})}\,. \tag{229}$$

With the outlined procedure we are able to compute the 1-loop power spectrum model for any given

cosmology in almost no time.

Next, we describe the higher order n-point functions that can be consistently computed using the cubic model. The tree-level bispectrum in real space is given by [25]:

$$B_g(k_1,k_2,k_3,z) = 2b_1^2\alpha^4 K_2(\boldsymbol{k}_1,\boldsymbol{k}_2)P_{\rm lin}(k_1,z)P_{\rm lin}(k_2,z) + 2\,\text{perms.} + \frac{c_2}{\bar{n}_g}b_1\alpha^2 P_{\rm lin}(k_1,z) + 2\,\text{perms.} + \frac{c_3}{\bar{n}_g^2}\,. \tag{230}$$

Note that there are two additional coefficients in the noise model compared to the power spectrum. These three free parameters c_1, c_2 and c_3 exactly match the noise parameters used in the analysis of [171].

The only additional n-point correlation function that can be fully computed using cubic order Eulerian kernels is the connected tree-level trispectrum, which is given by

$$\begin{aligned} T_g^c(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4,z) = {} & 4b_1^2\alpha^6 P_{\rm lin}(k_1,z)P_{\rm lin}(k_2,z)\big(K_2(\boldsymbol{k}_1,-\boldsymbol{k}_{13})K_2(\boldsymbol{k}_2,\boldsymbol{k}_{13})P_{\rm lin}(k_{13},z) \\ & + K_2(\boldsymbol{k}_1,-\boldsymbol{k}_{14})K_2(\boldsymbol{k}_2,\boldsymbol{k}_{14})P_{\rm lin}(k_{14},z)\big) + 5\,\text{perm.} \\ & + 6\,b_1^3\alpha^6 K_3(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)P_{\rm lin}(k_1,z)P_{\rm lin}(k_2,z)P_{\rm lin}(k_3,z) + 3\,\text{perm.} \\ & + \frac{c_2}{\bar{n}_g}B_{mgg}(\boldsymbol{k}_{12},\boldsymbol{k}_3,\boldsymbol{k}_4,z) + 5\,\text{perm.}\,, \end{aligned} \tag{231}$$

where we have defined

$$\begin{aligned} B_{mgg}(\boldsymbol{k}_{12},\boldsymbol{k}_3,\boldsymbol{k}_4,z) \equiv {} & \langle\delta_m(\boldsymbol{k}_{12},z)\delta_g(\boldsymbol{k}_3,z)\delta_g(\boldsymbol{k}_4,z)\rangle' = \\ = {} & 2b_1^2\alpha^4 F_2(\boldsymbol{k}_3,\boldsymbol{k}_4)P_{\rm lin}(k_3,z)P_{\rm lin}(k_4,z) \\ & + 2b_1\alpha^4 K_2(\boldsymbol{k}_{12},\boldsymbol{k}_4)P_{\rm lin}(k_4,z)P_{\rm lin}(k_{12},z) + 2b_1\alpha^4 K_2(\boldsymbol{k}_{12},\boldsymbol{k}_3)P_{\rm lin}(k_3,z)P_{\rm lin}(k_{12},z)\,. \end{aligned} \tag{232}$$

Note that here we keep only the largest noise contribution which at leading order has the same parameter c_2 as the bispectrum. We do this for two reasons: (a) in order to be able to have a fair comparison to the field-level analysis of [171] (see the following section) which does not use the most general noise model, and (b) in order to get the most optimistic results for additional constraining power of the trispectrum. Other (subdominant) noise contributions to the trispectrum can be derived from the noise model in Equation (211). Let us derive them. The total galaxy field can indeed be expressed as a sum of the perturbative solution and the stochastic contribution

$$\delta_g(\boldsymbol{x}) = \delta_g^{\rm PT}(\boldsymbol{x}) + \epsilon(\boldsymbol{x})\,. \tag{233}$$

We expand the noise field using

$$\epsilon(\boldsymbol{x}) = \epsilon_0(\boldsymbol{x}) + \epsilon_1(\boldsymbol{x})\delta(\boldsymbol{x}) + \frac{1}{2}\epsilon_2(\boldsymbol{x})(\delta^2(\boldsymbol{x}) - \langle\delta^2\rangle) + \epsilon_{\mathcal{G}_2}(\boldsymbol{x})\mathcal{G}_2(\boldsymbol{x}) + \dots\,, \tag{234}$$

and, since it is well-known that the spectra of the noise in the Poissonian limit satisfy

$$B_\epsilon = P_\epsilon^2 = \frac{1}{\bar{n}_g^2}\,, \tag{235}$$

we define the correlation among different noise operators as

$$\langle \epsilon_i(\boldsymbol{x})\epsilon_j(\boldsymbol{x}')\rangle \equiv \frac{c_{ij}}{\bar{n}_g}\delta^D(\boldsymbol{x}-\boldsymbol{x}')\,, \tag{236}$$

$$\langle \epsilon_i(\boldsymbol{x}_1)\epsilon_j(\boldsymbol{x}_2)\epsilon_k(\boldsymbol{x}_3)\rangle \equiv \frac{c_{ijk}}{\bar{n}_g^2}\delta^D(\boldsymbol{x}_1-\boldsymbol{x}_2)\delta^D(\boldsymbol{x}_2-\boldsymbol{x}_3)\,, \tag{237}$$

$$\langle \epsilon_i(\boldsymbol{x}_1)\epsilon_j(\boldsymbol{x}_2)\epsilon_k(\boldsymbol{x}_3)\epsilon_l(\boldsymbol{x}_4)\rangle \equiv \frac{c_{ijkl}}{\bar{n}_g^3}\delta^D(\boldsymbol{x}_1-\boldsymbol{x}_2)\delta^D(\boldsymbol{x}_2-\boldsymbol{x}_3)\delta^D(\boldsymbol{x}_3-\boldsymbol{x}_4)\,. \tag{238}$$

As a warm-up, we compute the power spectrum and bispectrum stochastic terms: the power spectrum of Equation (233) is

$$P_{gg}(\boldsymbol{k}) = \langle \delta_g(\boldsymbol{k})\delta_g(\boldsymbol{k}')\rangle' + \langle \epsilon(\boldsymbol{k})\epsilon(\boldsymbol{k}')\rangle'\,, \tag{239}$$

where we have exploited $\langle \delta\epsilon\rangle = 0$. The leading stochastic contribution for the galaxy power spectrum is then

$$\langle \epsilon_0(\boldsymbol{k})\epsilon_0(\boldsymbol{k}')\rangle' = \frac{c_{00}}{\bar{n}_g}\,, \tag{240}$$

with c_{00} being defined in Equation (236). Notice that additional terms in the expansion could arise that will eventually renormalize the c_{00} coefficient, e.g. the term $\langle \epsilon_1(\boldsymbol{x})\delta(\boldsymbol{x})\epsilon_1(\boldsymbol{x}')\delta(\boldsymbol{x}')\rangle'$. For the bispectrum the leading shot noise terms are

$$\langle \delta_g^{\rm PT}(\boldsymbol{k}_1)\epsilon(\boldsymbol{k}_2)\epsilon(\boldsymbol{k}_3)\rangle' + 2\text{ perms.} = \frac{c_{01}}{\bar{n}_g}b_1\big(P(k_1)+P(k_2)+P(k_3)\big)\,, \tag{241}$$

and

$$\langle \epsilon(\boldsymbol{k}_1)\epsilon(\boldsymbol{k}_2)\epsilon(\boldsymbol{k}_3)\rangle' = \frac{c_{000}}{\bar{n}_g^2}\,. \tag{242}$$

For the trispectrum we can write all the possible terms in the perturbative expansion as

$$\begin{aligned} T_g(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) =& \langle \delta_g(\boldsymbol{k}_1)\delta_g(\boldsymbol{k}_2)\delta_g(\boldsymbol{k}_3)\delta_g(\boldsymbol{k}_4)\rangle' \\ =& \langle \delta_g^{\rm PT}(\boldsymbol{k}_1)\delta_g^{\rm PT}(\boldsymbol{k}_2)\delta_g^{\rm PT}(\boldsymbol{k}_3)\delta_g^{\rm PT}(\boldsymbol{k}_4)\rangle' \\ &+ \langle \delta_g^{\rm PT}(\boldsymbol{k}_1)\delta_g^{\rm PT}(\boldsymbol{k}_2)\delta_g^{\rm PT}(\boldsymbol{k}_3)\epsilon(\boldsymbol{k}_4)\rangle' + 3\text{ perms} \\ &+ \langle \delta_g^{\rm PT}(\boldsymbol{k}_1)\delta_g^{\rm PT}(\boldsymbol{k}_2)\epsilon(\boldsymbol{k}_3)\epsilon(\boldsymbol{k}_4)\rangle' + 5\text{ perms} \\ &+ \langle \delta_g^{\rm PT}(\boldsymbol{k}_1)\epsilon(\boldsymbol{k}_2)\epsilon(\boldsymbol{k}_3)\epsilon(\boldsymbol{k}_4)\rangle' + 3\text{ perms} \\ &+ \langle \epsilon(\boldsymbol{k}_1)\epsilon(\boldsymbol{k}_2)\epsilon(\boldsymbol{k}_3)\epsilon(\boldsymbol{k}_4)\rangle'\,. \end{aligned} \tag{243}$$

The first line of Equation (243) is simply the tree-level perturbative expression for the galaxy trispectrum. The second line is automatically zero, since $\langle \epsilon\rangle = 0$. The third line gives rise to a number of possible

stochastic terms (up to 1-loop order):

$$\begin{aligned}
&\langle \delta_g^{\rm PT}(\boldsymbol{k}_1)\delta_g^{\rm PT}(\boldsymbol{k}_2)\epsilon(\boldsymbol{k}_3)\epsilon(\boldsymbol{k}_4)\rangle' = \langle b_1\delta(\boldsymbol{k}_1)b_1\delta(\boldsymbol{k}_2)\epsilon_0(\boldsymbol{k}_3)\epsilon_0(\boldsymbol{k}_4)\rangle' \\
&+ \langle b_1\delta(\boldsymbol{k}_1)b_1\delta(\boldsymbol{k}_2)\epsilon_0(\boldsymbol{k}_3)\int_p \epsilon_1(\boldsymbol{p})\delta(\boldsymbol{k}_4-\boldsymbol{p})\rangle' + \langle b_1\delta(\boldsymbol{k}_1)b_1\delta(\boldsymbol{k}_2)\epsilon_0(\boldsymbol{k}_3)\int_p \epsilon_2(\boldsymbol{p})\delta^2(\boldsymbol{k}_4-\boldsymbol{p})\rangle' \\
&+ \langle b_1\delta(\boldsymbol{k}_1)b_1\delta(\boldsymbol{k}_2)\epsilon_0(\boldsymbol{k}_3)\int_p \epsilon_{\mathcal{G}_2}(\boldsymbol{p})\mathcal{G}_2(\boldsymbol{k}_4-\boldsymbol{p})\rangle' + \langle b_1\delta(\boldsymbol{k}_1)b_1\delta(\boldsymbol{k}_2)\int_p \epsilon_1(\boldsymbol{p})\delta(\boldsymbol{k}_3-\boldsymbol{p})\int_p \epsilon_1(\boldsymbol{p})\delta(\boldsymbol{k}_4-\boldsymbol{p})\rangle' \\
&\langle b_1\delta(\boldsymbol{k}_1)b_2\delta^2(\boldsymbol{k}_2)\epsilon_0(\boldsymbol{k}_3)\epsilon_0(\boldsymbol{k}_4)\rangle' + \langle b_1\delta(\boldsymbol{k}_1)b_{\mathcal{G}_2}\mathcal{G}_2(\boldsymbol{k}_2)\epsilon_0(\boldsymbol{k}_3)\epsilon_0(\boldsymbol{k}_4)\rangle' \\
&+ \langle b_1\delta(\boldsymbol{k}_1)b_2\delta^2(\boldsymbol{k}_2)\epsilon_0(\boldsymbol{k}_3)\int_p \epsilon_1(\boldsymbol{p})\delta(\boldsymbol{k}_3-\boldsymbol{p})\rangle' + \langle b_1\delta(\boldsymbol{k}_1)b_{\mathcal{G}_2}\mathcal{G}_2(\boldsymbol{k}_2)\epsilon_0(\boldsymbol{k}_3)\int_p \epsilon_1(\boldsymbol{p})\delta(\boldsymbol{k}_3-\boldsymbol{p})\rangle' \,,
\end{aligned} \tag{244}$$

that eventually give the $\propto \bar{n}_g^{-1}$ contribution

$$\langle \delta_g^{\rm PT}(\boldsymbol{k}_1)\delta_g^{\rm PT}(\boldsymbol{k}_2)\epsilon(\boldsymbol{k}_3)\epsilon(\boldsymbol{k}_4)\rangle' = \frac{c_{01}}{\bar{n}_g}B_{mgg}(\boldsymbol{k}_{12},\boldsymbol{k}_3,\boldsymbol{k}_4) + \frac{2b_1^2}{\bar{n}_g}\left(c_{02}+c_{0\mathcal{G}_2}\sigma^2(\boldsymbol{k}_1,\boldsymbol{k}_2)+c_{11}\right)P(k_1)P(k_2)\,. \tag{245}$$

The fourth line of Equation (243) produces contributions of order $\sim \bar{n}_g^2$

$$\langle \delta_g^{\rm PT}(\boldsymbol{k}_1)\epsilon(\boldsymbol{k}_2)\epsilon(\boldsymbol{k}_3)\epsilon(\boldsymbol{k}_4)\rangle' \simeq \frac{b_1 c_{001}}{\bar{n}_g^2}P(k_1) + 3\,\text{perms.}\,. \tag{246}$$

at leading perturbative order. The last contribution is given by the noise trispectrum

$$\langle \epsilon(\boldsymbol{k}_1)\epsilon(\boldsymbol{k}_2)\epsilon(\boldsymbol{k}_3)\epsilon(\boldsymbol{k}_4)\rangle' = \frac{c_{0000}}{\bar{n}_g^3}\,. \tag{247}$$

Even though the impact of the sub-leading noise term in the trispectrum is not expected to be large (given reasonable priors on the amplitude of the noise), they would lead to slightly larger error bars in the trispectrum analysis.

B Estimators and covariances

Equipped with the theoretical model, in this section we present the likelihoods for the power spectrum, bispectrum, trispectrum used in the following applications. Working in the ideal world of Fisher forecasts or simulation boxes allows us to use simple estimators and derivation of corresponding covariance matrices is relatively straightforward. For simplicity we will start with the continuum limit and derive the χ^2 in terms of the standard binning.

We begin by considering the power spectrum. Let us assume that we have a simulation box with volume V and that the estimate of the galaxy density field in this box is given by $\hat{\delta}_g$. The standard power spectrum estimator is then given by

$$\hat{P}_g(k) = \frac{1}{VV_s}\int_{\boldsymbol{q}\in k}\hat{\delta}_g(\boldsymbol{q})\hat{\delta}_g(-\boldsymbol{q})\,, \tag{248}$$

where the volume of a spherical shell in the k-bin of thickness Δk is $V_s = 4\pi k^2\Delta k$. Note that $\int_{\boldsymbol{q}\in k}$ indicates that the integration is done only in the shell of radius k and thickness Δk. By taking the expectation value and using $V = (2\pi)^3\delta^D(0)$ we can see that the estimator is unbiased

$$\langle \hat{P}_g(k)\rangle = \frac{1}{VV_s}\int_{\boldsymbol{q}\in k}P_g(q)(2\pi)^3\delta^D(0) \approx P_g(k)\,. \tag{249}$$

The last equality is approximate, since we assume that $P(k)$ is approximately constant within the bin k (or in other words that the bins are narrow). The power spectrum covariance matrix Cov_P in the Gaussian approximation is given by

$$\mathrm{Cov}_P = \langle \hat{P}_g(k)\hat{P}_g(k')\rangle - \langle \hat{P}_g(k)\rangle\langle \hat{P}_g(k')\rangle = \frac{(2\pi)^3}{V}\frac{P_g^2(k)}{2\pi k^2\Delta k}\delta^K_{kk'} \,. \tag{250}$$

This is a well-known formula where the covariance matrix is diagonal (different bins are uncorrelated).

Let us now turn to the bispectrum. The bins in the data vector correspond to triangles of different shape labeled by the three sides $T = \{k_1, k_2, k_3\}$. Since each bin has a unique shape, the momenta are ordered $k_1 \geq k_2 \geq k_3$. The standard estimator is

$$\hat{B}(k_1, k_2, k_3) = \frac{1}{VV_{123}}\int_{\boldsymbol{q}_i\in k_i} \hat{\delta}_g(\boldsymbol{q}_1)\hat{\delta}_g(\boldsymbol{q}_2)\hat{\delta}_g(\boldsymbol{q}_3)\,\delta^D(\boldsymbol{q}_{123}) \,. \tag{251}$$

Using the well-known result for the volume V_{123}

$$V_{123} = \int_{\boldsymbol{q}_i\in k_i} \delta^D(\boldsymbol{q}_{123}) = 8\pi^2 k_1 k_2 k_3 \Delta k^3 \,, \tag{252}$$

one can check that the estimator is unbiased $\langle \hat{B}_g\rangle = B_g$. In the Gaussian approximation, the covariance matrix for the bispectrum estimator is given by

$$\begin{aligned}\mathrm{Cov}_B &= \langle \hat{B}_g(k_1, k_2, k_3)\hat{B}_g(k_1', k_2', k_3')\rangle - \langle \hat{B}_g(k_1, k_2, k_3)\rangle\langle \hat{B}_g(k_1', k_2', k_3')\rangle \\ &= \frac{(2\pi)^3}{V}\frac{\pi s_{123}}{k_1 k_2 k_3 \Delta k^3} P_g(k_1)P_g(k_2)P_g(k_3) \,,\end{aligned} \tag{253}$$

where s_{123} equals 1, 2 or 6 for general, isosceles and equilateral triangles respectively.

While the form of the power spectrum and bispectrum likelihoods in the Gaussian approximation is well-known, the trispectrum is much less used and explored in the comparisons to simulations and data (for some recent progress, see [172, 173]). First of all, the connected trispectrum is a function of six scalar variables. It can be parametrized in different ways, but we will use four magnitudes of the wavenumbers $k_1, \ldots, k_4$ and two magnitudes of the diagonals, D_1 and D_2, defined as

$$\boldsymbol{D}_1 \equiv -\boldsymbol{k}_1 - \boldsymbol{k}_2 \,, \qquad \text{and} \qquad \boldsymbol{D}_2 \equiv -\boldsymbol{k}_2 - \boldsymbol{k}_3 \,. \tag{254}$$

so that a quadrilateral $\boldsymbol{Q}$ is defined by $\boldsymbol{Q} = \{k_1, k_2, k_3, k_4, D_1, D_2\}$. In order to avoid the double counting, the sides of the quadrilaterals are ordered as $k_4 \leq k_3 \leq k_2 \leq k_1$, while the diagonals must satisfy the relations $|k_1 - k_2| \leq D_1 \leq |k_1 + k_2|$ and $|k_2 - k_3| \leq D_2 \leq |k_2 + k_3|$. The full data vector that depends on all momenta and diagonals is very long and has very large dimensional covariance matrix. For these reasons the compressed trispectrum is sometimes used in practice. One such example is the integrated trispectrum, where in the estimator one averages over all D_1 and D_2 for the fixed momenta $k_1, \ldots, k_4$. Using this estimator significantly reduces the length of the data vector and allows for a simple evaluation of the covariance matrix. The integrated trispectrum has been used both in real and redshift space, for comparison of perturbation theory predictions to simulations [172, 173].

One of the goal of this chapter is to estimate the impact of the trispectrum on analyses on large scales and find the most optimal possible improvement. For this reason we cannot use the simplified estimators such as integrated trispectrum, where some information is lost due to integration over the diagonals. This is particularly relevant for understanding the impact of the trispectrum on constraining all cubic bias parameters, where the only hope to do it comes from using the full shape keeping all dependence on different angles.

Given all of this the estimator for the connected trispectrum can be then written as

$$\hat{T}_c(k_1,k_2,k_3,k_4,D_1,D_2) = \tag{255}$$
$$\frac{1}{VV_{1234}}\int_{\boldsymbol{q}_i\in k_i}\int_{\boldsymbol{p}_i\in D_i}\hat{\delta}_g(\boldsymbol{q}_1)\hat{\delta}_g(\boldsymbol{q}_2)\hat{\delta}_g(\boldsymbol{q}_3)\hat{\delta}_g(\boldsymbol{q}_4)\,\delta^D(\boldsymbol{q}_{1234})\delta^D(\boldsymbol{q}_{12}+\boldsymbol{p}_1)\delta^D(\boldsymbol{q}_{23}+\boldsymbol{p}_2)$$
$$-\text{ disconnected terms}\,, \tag{256}$$

where the "disconnected terms" are chosen such that the estimator is unbiased $\langle\hat{T}_c\rangle = T_c$. Their explicit form is not of interest in the Gaussian likelihood, since they cancel in the covariance matrix.

Let us first compute the volume for the thin momentum shell containing all quadrilaterals with given momenta and diagonals. It is defined as

$$V_{1234} = \int_{p_i\in D_i} d^3\boldsymbol{p}_i \int_{q_j\in k_j} d^3\boldsymbol{q}_j\,\delta^D(\boldsymbol{q}_1+\boldsymbol{q}_2+\boldsymbol{q}_3+\boldsymbol{q}_4)\delta^D(\boldsymbol{q}_1+\boldsymbol{q}_2+\boldsymbol{p}_1)\delta^D(\boldsymbol{q}_2+\boldsymbol{q}_3+\boldsymbol{p}_2)\,. \tag{257}$$

Note that the symmetry among six edges of the tetrahedron is broken the moment these particular delta functions are chosen. What complicates the evaluation of the integral is that $\boldsymbol{q}_2$ appears three times as an argument of delta functions. When doing angular integrals this will lead to Wigner 3j symbols. Therefore, it is more convenient to rewrite the previous formula as follows

$$V_{1234} = \int_{p_i\in D_i} d^3\boldsymbol{p}_i \int_{q_j\in k_j} d^3\boldsymbol{q}_j\,\delta^D(\boldsymbol{q}_1+\boldsymbol{q}_4-\boldsymbol{p}_2)\delta^D(\boldsymbol{q}_1+\boldsymbol{q}_2+\boldsymbol{p}_1)\delta^D(\boldsymbol{q}_2+\boldsymbol{q}_3+\boldsymbol{p}_2)\,, \tag{258}$$

which we can always do using the delta functions. This form is simpler, since any of the momenta enter the arguments of the delta functions at most twice. Next we can write

$$V_{1234} = \int_{p_i\in D_i} d^3\boldsymbol{p}_i \int_{q_j\in k_j} d^3\boldsymbol{q}_j \int \frac{d^3\boldsymbol{r}}{(2\pi)^3}\frac{d^3\boldsymbol{x}}{(2\pi)^3}\frac{d^3\boldsymbol{y}}{(2\pi)^3}\,e^{-i\boldsymbol{r}\cdot(\boldsymbol{q}_1+\boldsymbol{q}_4-\boldsymbol{p}_2)-i\boldsymbol{x}\cdot(\boldsymbol{q}_1+\boldsymbol{q}_2+\boldsymbol{p}_1)-i\boldsymbol{y}\cdot(\boldsymbol{q}_2+\boldsymbol{q}_3+\boldsymbol{p}_2)}\,. \tag{259}$$

Angular integration over $\boldsymbol{p}_1$, $\boldsymbol{q}_4$ and $\boldsymbol{q}_3$ can be done easily.

$$V_{1234} = \frac{1}{8\pi^6}\int_{p_i\in D_i} p_i^2 dp_i \int_{q_j\in k_j} q_j^2 dq_j \int_0^\infty r^2 dr \int_0^\infty x^2 dx \int_0^\infty y^2 dy\; j_0(p_1x)j_0(q_4r)j_0(q_3y)$$
$$\times\int d\Omega_{p_2}\int d\Omega_{q_1}\int d\Omega_{q_2}\int d\Omega_r\int d\Omega_x\int d\Omega_y\,e^{-i\boldsymbol{r}\cdot(\boldsymbol{q}_1-\boldsymbol{p}_2)-i\boldsymbol{x}\cdot(\boldsymbol{q}_1+\boldsymbol{q}_2)-i\boldsymbol{y}\cdot(\boldsymbol{q}_2+\boldsymbol{p}_2)}\,. \tag{260}$$

In order to perform the other angular integrals, we will use the expansion of plane waves in spherical Bessel functions and spherical harmonics

$$e^{\pm i\boldsymbol{r}\cdot\boldsymbol{q}} = 4\pi\sum_{\ell m}(\pm i)^\ell j_\ell(qr)Y_{\ell m}(\hat{\boldsymbol{r}})Y^*_{\ell m}(\hat{\boldsymbol{q}})\,. \tag{261}$$

It follows that

$$\int d\Omega_{q_1}e^{-i\boldsymbol{r}\cdot\boldsymbol{q}_1-i\boldsymbol{x}\cdot\boldsymbol{q}_1} = (4\pi)^2\sum_{\ell_1m_1}\sum_{\ell_2m_2}(-i)^{\ell_1+\ell_2}j_{\ell_1}(q_1r)j_{\ell_2}(q_1x)Y_{\ell_1m_1}(\hat{\boldsymbol{r}})Y^*_{\ell_2m_2}(\hat{\boldsymbol{x}})\int d\Omega_{q_1}Y^*_{\ell_1m_1}(\hat{\boldsymbol{q}}_1)Y_{\ell_2m_2}(\hat{\boldsymbol{q}}_1)\,. \tag{262}$$

By orthonormality of the spherical harmonic this integral leads to

$$\int d\Omega_{q_1}e^{-i\boldsymbol{r}\cdot\boldsymbol{q}_1-i\boldsymbol{x}\cdot\boldsymbol{q}_1} = (4\pi)^2\sum_{\ell m}(-1)^\ell j_\ell(q_1r)j_\ell(q_1x)Y_{\ell m}(\hat{\boldsymbol{r}})Y^*_{\ell m}(\hat{\boldsymbol{x}})\,. \tag{263}$$

We can to the other angular integrals in the similar way. In particular,

$$\int d\Omega_{q_2} e^{-i\boldsymbol{x}\cdot\boldsymbol{q}_2 - i\boldsymbol{y}\cdot\boldsymbol{q}_2} = (4\pi)^2 \sum_{\ell m} (-1)^\ell j_\ell(q_2 x) j_\ell(q_2 y) Y_{\ell m}(\hat{\boldsymbol{x}}) Y^*_{\ell m}(\hat{\boldsymbol{y}}) \,, \tag{264}$$

$$\int d\Omega_{p_2} e^{+i\boldsymbol{r}\cdot\boldsymbol{p}_2 - i\boldsymbol{y}\cdot\boldsymbol{p}_2} = (4\pi)^2 \sum_{\ell m} j_\ell(p_2 r) j_\ell(p_2 y) Y_{\ell m}(\hat{\boldsymbol{y}}) Y^*_{\ell m}(\hat{\boldsymbol{r}}) \,. \tag{265}$$

Once all the angular integrals are done, the volume of the shell is a complicated radial integral of product of spherical Bessel functions

$$\begin{aligned} V_{1234} = 8^3 \int_{p_i\in D_i} p_i^2 dp_i \int_{q_j\in k_j} q_j^2 dq_j \int_0^\infty r^2 dr \int_0^\infty x^2 dx \int_0^\infty y^2 dy \sum_\ell (2\ell+1) \\ \times j_0(p_1 x) j_0(q_4 r) j_0(q_3 y) j_\ell(q_1 r) j_\ell(q_1 x) j_\ell(q_2 x) j_\ell(q_2 y) j_\ell(p_2 r) j_\ell(p_2 y) \,. \end{aligned} \tag{266}$$

Luckily, the radial integrals can be also done relatively straightforwardly. First, we can use

$$\int_0^\infty r^2 dr\, j_0(q_4 r) j_\ell(q_1 r) j_\ell(p_2 r) = \frac{\pi}{4 q_1 q_4 p_2} \mathcal{P}_\ell\left(\frac{q_1^2 + p_2^2 - q_4^2}{2 q_1 p_2}\right) \,, \tag{267}$$

where $\mathcal{P}_\ell(x)$ are the Legendre polynomials. The volume of the shell then becomes

$$\begin{aligned} V_s = (2\pi)^3 \int_{p_i\in D_i} p_i^2 dp_i \int_{q_j\in k_j} q_j^2 dq_j \sum_\ell (2\ell+1) \\ \times \frac{1}{q_1 q_4 p_2} \mathcal{P}_\ell\left(\frac{q_1^2 + p_2^2 - q_4^2}{2 q_1 p_2}\right) \frac{1}{q_1 q_2 p_1} \mathcal{P}_\ell\left(\frac{q_1^2 + q_2^2 - p_1^2}{2 q_1 q_2}\right) \frac{1}{q_2 q_3 p_2} \mathcal{P}_\ell\left(\frac{q_2^2 + p_2^2 - q_3^2}{2 q_2 p_2}\right) \,. \end{aligned} \tag{268}$$

Finally, the sum over ℓ can be done using

$$\sum_\ell (2\ell+1) \mathcal{P}_\ell(\mu_1) \mathcal{P}_\ell(\mu_2) \mathcal{P}_\ell(\mu_3) = \frac{2}{\pi} \frac{1}{\sqrt{1 + 2\mu_1\mu_2\mu_3 - \mu_1^2 - \mu_2^2 - \mu_3^2}} \,. \tag{269}$$

This equation has a nice geometric interpretation. It is related to the volume of the tetrahedron with three edges q_1, q_2 and p_2

$$V_{\text{tetra}} = \frac{q_1 q_2 p_2}{6} \sqrt{1 + 2\mu_1\mu_2\mu_3 - \mu_1^2 - \mu_2^2 - \mu_3^2} \,, \tag{270}$$

where μ_i are cosines of plane angles between three edges meeting in a common vertex. In our formula for V_{1234} we have exactly the same form after closer inspections of all angles. Therefore, we can finally write (assuming that all bins are Δk wide)

$$V_s = \frac{8}{3}\pi^2 \frac{k_1 k_2 k_3 k_4 D_1 D_2}{V_{\text{tetra}}} \Delta k^6 \,. \tag{271}$$

The volume V_{tetra} can be computed in terms of six edges and this expression is manifestly symmetric in $k_1, \ldots, k_4$ and D_1, D_2. Now we are ready to compute the trispectrum covariance matrix in the Gaussian

approximation:

$$\begin{aligned}
&\mathrm{Cov}_T(k_1,k_2,k_3,k_4,D_1,D_2) = \langle \hat{T}(k_1,k_2,k_3,k_4,D_1,D_2)\hat{T}(k_a,k_b,k_c,k_d,D_a,D_b)\rangle - \\
&\langle \hat{T}(k_1,k_2,k_3,k_4,D_1,D_2)\rangle\langle \hat{T}(k_a,k_b,k_c,k_d,D_a,D_b)\rangle \\
&= \frac{(2\pi)^{12}}{V^2 V_{1234} V_{abcd}} P(k_1)P(k_2)P(k_3)P(k_4) \prod_{i=1,a}^{4,d} \prod_{j=1,a}^{2,b} \int_{\boldsymbol{q}_i\in\boldsymbol{k}_i} d^3q_i \int_{\boldsymbol{p}_j\in\mathbf{D}_j} d^3p_j \\
&\times \delta_D(\boldsymbol{q}_{1234})\delta_D(\boldsymbol{q}_{12}+\boldsymbol{p}_1)\delta_D(\boldsymbol{q}_{23}+\boldsymbol{p}_2)\delta_D(\boldsymbol{q}_{abcd})\delta_D(\boldsymbol{q}_{ab}+\boldsymbol{p}_a)\delta_D(\boldsymbol{q}_{bc}+\boldsymbol{p}_b)\delta_D(\boldsymbol{q}_{1a})\delta_D(\boldsymbol{q}_{2b})\delta_D(\boldsymbol{q}_{3c})\delta_D(\boldsymbol{q}_{4d}) \\
&+23 \text{ perm.} \\
&= \frac{(2\pi)^9}{V V_{1234} V_{abcd}} P(k_1)P(k_2)P(k_3)P(k_4)\delta^K_{\boldsymbol{k}_1,\boldsymbol{k}_a}\delta^K_{\boldsymbol{k}_2,\boldsymbol{k}_b}\delta^K_{\boldsymbol{k}_3,\boldsymbol{k}_c}\delta^K_{\boldsymbol{k}_4,\boldsymbol{k}_d}\delta^K_{\boldsymbol{p}_1,\boldsymbol{p}_a}\delta^K_{\boldsymbol{p}_2,\boldsymbol{p}_b} \\
&\prod_{i=a}^{d}\prod_{j=a}^{b} \int_{\boldsymbol{q}_i\in\boldsymbol{k}_i} d^3q_i \int_{\boldsymbol{p}_j\in\mathbf{D}_j} d^3p_j \delta_D(\boldsymbol{q}_{abcd})\delta_D(\boldsymbol{q}_{ab}+\boldsymbol{p}_a)\delta_D(\boldsymbol{q}_{bc}+\boldsymbol{p}_b) + 23 \text{ perm.} \\
&= \frac{(2\pi)^9}{V V_{1234}} P(k_1)P(k_2)P(k_3)P(k_4)\delta^K_{\boldsymbol{k}_1,\boldsymbol{k}_a}\delta^K_{\boldsymbol{k}_2,\boldsymbol{k}_b}\delta^K_{\boldsymbol{k}_3,\boldsymbol{k}_c}\delta^K_{\boldsymbol{k}_4,\boldsymbol{k}_d}\delta^K_{\boldsymbol{p}_1,\boldsymbol{p}_a}\delta^K_{\boldsymbol{p}_2,\boldsymbol{p}_b} + 23 \text{ perm.} \\
&= \frac{(2\pi)^9 R_{1234}}{V V_{1234}} P(k_1)P(k_2)P(k_3)P(k_4)\ ,
\end{aligned} \tag{272}$$

where V_{1234} is given by Equation (271).

In the Gaussian approximation some PT cross-terms survive. We have

$$\begin{aligned}
&\langle T(k_1,k_2,k_3,k_4,D_1,D_2)P(k)\rangle - \langle T(k_1,k_2,k_3,k_4,D_1,D_2)\rangle\langle P(k)\rangle = \\
&\frac{1}{V^2 V_{1234} V_P} \prod_{i=1}^{4}\prod_{j=1,2}\prod_{n=a,b} \int_{q_i\in k_i} d^3q_i \int_{p_i\in D_j} d^3p_j \int_{q_n\in k_n} d^3q_n \delta_D(\boldsymbol{q}_{1234})\delta_D(\boldsymbol{q}_{ab})\delta_D(\boldsymbol{q}_{12}+\boldsymbol{p}_1)\delta_D(\boldsymbol{q}_{23}+\boldsymbol{p}_2)\times \\
&\langle(\delta(\boldsymbol{q}_1)\delta(\boldsymbol{q}_2)\delta(\boldsymbol{q}_3)\delta(\boldsymbol{q}_4)\delta(\boldsymbol{q}_a)\delta(\boldsymbol{q}_b)\rangle_c \\
&= 2\frac{(2\pi)^9}{V^2 V_{1234} V_P} \prod_{i=1}^{4}\prod_{j=1,2}\prod_{n=a,b} \int_{q_i\in k_i} d^3q_i \int_{p_i\in D_j} d^3p_j \int_{q_n\in k_n} d^3q_n \delta_D(\boldsymbol{q}_{1234})\delta_D(\boldsymbol{q}_{ab})\delta_D(\boldsymbol{q}_{12}+\boldsymbol{p}_1)\delta_D(\boldsymbol{q}_{23}+\boldsymbol{p}_2)\times \\
&\delta_D(\boldsymbol{q}_{1a})\delta_D(\boldsymbol{q}_{3b})\delta_D(\boldsymbol{q}_{24})P(q_1)P(q_3)P(q_2) \\
&= 2\frac{(2\pi)^6}{V V_{1234} V_P} P^2(k)P(k_2)\delta_{\boldsymbol{k},\boldsymbol{k}_1}\delta_{\boldsymbol{k},\boldsymbol{k}_3}\delta_{\boldsymbol{k}_2,\boldsymbol{k}_4} \prod_{i=3,4}\prod_{j=1,2} \int_{q_i\in k_i} d^3q_i \int_{p_i\in D_j} d^3p_j \delta_D(-\boldsymbol{q}_3-\boldsymbol{q}_4+\boldsymbol{p}_1)\delta_D(\boldsymbol{q}_3-\boldsymbol{q}_4+\boldsymbol{p}_2)\ ,
\end{aligned} \tag{273}$$

all the other contractions will give $\boldsymbol{p}_{1,2} = 0$ which do not fall in any bin. Now we want to evaluate:

$$\begin{aligned}
&\prod_{i=3,4}\prod_{j=1,2} \int_{q_i\in k_i} d^3q_i \int_{p_i\in D_j} d^3p_j \delta_D(-\boldsymbol{q}_3-\boldsymbol{q}_4+\boldsymbol{p}_1)\delta_D(\boldsymbol{q}_3-\boldsymbol{q}_4+\boldsymbol{p}_2) \\
&= \frac{1}{(2\pi^2)^2} \int d^3q_3 \int d^3q_4 \int dp_1 p_1^2 \int dp_2 p_2^2 \int d^3x \int d^3y e^{i\boldsymbol{q}_3(\boldsymbol{x}-\mathbf{y})} e^{-i\boldsymbol{q}_4(\boldsymbol{x}+\mathbf{y})} j_0(p_1 y) j_0(p_2 x)\ ,
\end{aligned} \tag{274}$$

where we performed angular integrations over $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$. Using Equation (261) leads to

$$2^6 \sum_\ell (2\ell+1) \int dq_3 dq_4 dp_1 dp_2 dx dy q_3^2 q_4^2 p_1^2 p_2^2 x^2 y^2 j_\ell(q_3 x) j_\ell(q_3 y) j_\ell(q_4 x) j_\ell(q_4 y) j_0(p_1 x) j_0(p_2 y)\ , \tag{275}$$

and using Equation (267) we get:

$$4\pi^2 \sum_\ell (2\ell+1) \int p_1 p_2 dq_3 dq_4 dp_1 dp_2 \mathcal{P}_\ell\left(\frac{q_3^2+q_4^2-p_1^2}{2q_3 q_4}\right)\mathcal{P}_\ell\left(\frac{q_3^2+q_4^2-p_2^2}{2q_3 q_4}\right)\ . \tag{276}$$

Now it is known that

$$\sum_{\ell} \frac{(2\ell+1)}{2} \mathcal{P}_\ell(\mu_1)\mathcal{P}_\ell(\mu_2) = \delta_D(\mu_1 - \mu_2) \,. \tag{277}$$

So Equation (276) reads

$$2(4\pi^2)\delta_{D_1,D_2} \int p_1^2 dp_1 \int dq_3 \int dq_4 = 8\pi^2 \delta_{D_1,D_2} D_1^2 \Delta k^3 \,. \tag{278}$$

So PT terms are non-zero only for quadrilaterals which have $k_1 = k_3$, $k_2 = k_4$, $D_1 = D_2$; given that we are considering quadrilaterals with ordered sides, the only possible configurations are the equilateral shapes, i.e. $k_1 = k_2 = k_3 = k_4$ and $D_1 = D_2$. They are $\sim 0.1\%$ of the total number so we will neglect them in what follows.

We are now ready to write down the likelihood for the power spectrum bispectrum and trispectrum. The power spectrum likelihood on large scales can be approximated by the well-known equation

$$\ln \mathcal{L}_P = -\frac{1}{2} \sum_{k=k_{\rm min}}^{k=k_{\rm max}} \frac{(P_g(k,\theta) - \hat{P}_g(k))^2}{{\rm Cov}_P(k)} \,. \tag{279}$$

We assume the Gaussian form of the likelihood which is appropriate on large scales. The sum runs over all k-bins up to a given $k_{\rm max}$. The set of variables θ contains all nuisance and cosmological parameters in the analysis, while the measured power spectrum in a given k-bin is $\hat{P}_g(k)$. For simplicity we neglect the explicit redshift dependence. Furthermore, we neglect the cosmological parameter dependence in the covariance matrix, since it carries negligible signal in the limit when the number of Fourier modes is much larger than the number of bins in the data vector. We will use this approximation for all other statistics as well.

On large scales used in our analysis, it is sufficient to use the linear theory in the computation of the covariance. All nonlinear corrections and off-diagonal elements in the covariance matrix are small on large scales and suppressed by the variance of the density field $\Delta^2(k_{\rm max})$. Furthermore, since we are interested in posterior on cosmological parameters after marginalization over all nuisance parameters, the final variance of cosmological parameters is not impacted only by the data covariance, but also the effective covariance due to the marginalization. This was explicitly demonstrated in some simple examples in [174]. This additional uncertainty is usually much bigger than any mistake made in the data covariance. For this reason, one recovers identical error bars on cosmological parameters on large scales in analyses based on perturbation theory, regardless of the choice of the data covariance. The explicit demonstration of this fact can be found in [175].

Similarly to the power spectrum, we assume a simple Gaussian likelihood with the diagonal covariance matrix for the bispectrum as well

$$\ln \mathcal{L}_B = -\frac{1}{2} \sum_{\mathcal{T}} \frac{(B_g(k_1,k_2,k_3,\theta) - \hat{B}_g(k_1,k_2,k_3))^2}{{\rm Cov}_B(k_1,k_2,k_3)} \,, \tag{280}$$

where the sum now runs over all triangles in Fourier space.

For reasons similar to the power spectrum case, one can justify the simple approximation for the bispectrum likelihood in perturbation theory. Corrections to the Gaussian approximation are expected to be of order $\Delta^2(k_{\rm max})$ and cannot significantly change the results in the full analysis. We will comment more on this and the cross-covariance with the power spectrum below.

For the full connected trispectrum we assume a Gaussian likelihood as well, given by

$$\ln \mathcal{L}_T = -\frac{1}{2}\sum_{\boldsymbol{Q}} \frac{(T_g^c(\boldsymbol{Q},\theta) - \hat{T}_g^c(\boldsymbol{Q}))^2}{\mathrm{Cov}_T(\boldsymbol{Q})} \,, \tag{281}$$

where the sum runs over all nonequivalent quadrilaterals .

In addition to the data covariance, it is often useful in perturbation theory based analysis to consider theoretical error covariance. Theoretical error takes into account realistic uncertainties in the perturbative modeling which serves two purposes. One is to prevent the overfitting by choosing aggressive data cuts. The other is to prevent biases in cosmological inference. Details of the formalism and applications to galaxy clustering and lensing can be found in [174, 176–179].

The elements (ij) of the theoretical error covariance matrix have the following form

$$(C_e)_{ij} = E_i E_j \rho_{ij} \,, \tag{282}$$

where the index i labels a bin in the data vector, which can be a k bin or a bispectrum or trispectrum configuration. Indices on the right hand side are not summed over. The quantity E_i is the measure of the absolute error for the theoretical prediction in a given bin i and can be obtained from estimating the higher-order perturbation theory contributions neglected in the calculation. The matrix ρ_{ij} measures the correlation between the two different bins and it is crucial to ensure that one effectively marginalizes only over sufficiently smooth theoretical predictions [176]. For the power spectrum it is chosen to be

$$\rho_{ij}^{\mathrm{P}} = \exp\left[-\frac{1}{2}\frac{(k_i - k_j)^2}{\Delta k^2}\right] \,, \tag{283}$$

where Δk is the parameters that determines the smoothness scale, over which the theoretical prediction vary very smoothly. In particular, we choose $\Delta k = 0.1\ h/\mathrm{Mpc}$. If the two k-bins are separated by more than Δk they are effectively uncorrelated and the theoretical prediction can vary within the theoretical uncertainty given by E_i and E_j. If the bins are within Δk, they are constrained to effectively fluctuate coherently. In the case of the bispectrum, the correlation matrix is a bit more complicated and takes into account the shape of the triangle as follows

$$\rho_{ij}^{\mathrm{B}} = \exp\left[-\frac{1}{2}\frac{(k_{i,1} - k_{j,1})^2}{\Delta k^2}\right]\exp\left[-\frac{1}{2}\frac{(k_{i,2} - k_{j,2})^2}{\Delta k^2}\right]\exp\left[-\frac{1}{2}\frac{(k_{i,3} - k_{j,3})^2}{\Delta k^2}\right] \,. \tag{284}$$

In conclusion, the theoretical error covariance matrix has two parameters. One is the coherence length Δk and the other one is the estimate of the absolute theoretical error in perturbation theory. Following [176] we use the following estimates for the two-loop power spectrum

$$E_i^{\mathrm{P,2-loop}} = D^4(z) P_g(k_i, z)\left(\frac{k_i}{0.45\ h/\mathrm{Mpc}}\right)^{3.3} \,, \tag{285}$$

and the one-loop bispectrum

$$E_i^{\mathrm{B,1-loop}} = D^2(z) B_g(k_{i,1}, k_{i,2}, k_{i,3}, z)\left(\frac{(k_{i,1} + k_{i,2} + k_{i,3})/3}{0.3\ h/\mathrm{Mpc}}\right)^{1.8} \,. \tag{286}$$

Note that requiring the theoretical error to be of the order of a given loop size is the other way to impose the perturbativity prior, where the total amplitude of fluctuations in the galaxy number density on large scales in always small.

Once the theoretical error covariance matrix is computed, it can be straightforwardly included in the analysis by computing the Gaussian likelihood with the total covariance

$$C_{ij} = C_{ij}^{\rm data} + (C_e)_{ij} \ , \tag{287}$$

where $C_{ij}^{\rm data}$ is the usual data covariance that contains cosmic variance and the noise. The new total covariance matrix is not diagonal and one has to generalize the formulas for the likelihood above accordingly. Nevertheless, the likelihood is still Gaussian.

Assuming the Gaussian likelihood (with or without the theoretical error) for the power spectrum, bispectrum and trispectrum, we will combine them in our analysis in a simple way as follows

$$\ln \mathcal{L} = \ln \mathcal{L}_P + \ln \mathcal{L}_B + \ln \mathcal{L}_T \ . \tag{288}$$

At first sight, this assumption may seem too optimistic, since we neglect all cross-covariance terms that may be relevant. In what follows we justify this choice.

At $k < k_{\rm max}$, the simple form of the total likelihood in Equation (288) is justified by arguments in perturbation theory. Since we are focusing on very large scales, the nonlinear contributions to the covariance and cross-covariance among different observables is small. In particular, these contributions are smaller than the effective error on cosmological parameters that comes from marginalization over many nuisance parameters [134]. This has been explicitly checked and verified for the power spectrum and bispectrum analyses on relevant scales (e.g., [131, 133, 134, 180, 181]. For example, the power spectrum analysis based on perturbation theory leads to identical results regardless of the choice of the covariance (Gaussian, non-linear or simulation-based), in a volume similar to the one we use in this work [175, 182]. The same conclusion holds in the joint P+B analysis as shown in many works which studied this combination in various setups (see for example [130, 183–185] and references therein).

Finally, regarding the trispectrum, we expect that the similar arguments hold and that Equation (288) is a very good approximation to the truth. Even if the analysis was optimistic, this is in line with our previous choices that maximize the possible information content of the trispectrum.

C Comparison with field level inference

Field-level Bayesian inference (FBI) of cosmological parameters has recently emerged as an alternative to conventional n-point function analyses. This approach offers several advantages, such as being optimal and allowing for a simpler inclusion of data systematics. However, in the context of galaxy clustering, field-level inference which relies on detailed cosmological simulations of the observed galaxy density field remains infeasible, due to very high computational cost. A practical alternative is FBI based on perturbative forward modeling. Within the framework of the Effective Field Theory of Large-Scale Structure (EFTofLSS) [139–141, 143, 154, 167, 186–188], perturbative forward modeling [189–198] provides a universal and reliable description of the galaxy density field on large scales with a handful of free parameters to be fitted from the data. Considerable effort has been made to test FBI with perturbative forward modeling and apply it in practice [182, 199–207] (see also [208] for a recent application of the LSS bootstrap [169, 209–211] to FBI).

A notable recent cosmological analysis using FBI was presented in [171], focusing on comparison of the field-level and conventional n-point function based approaches. By using the exact same perturbative forward model in both FBI and the simulation-based[17] inference (SBI) for the power spectrum (P) and bispectrum (B), the authors found that FBI achieves a 3-5 times more precise measurement of the

[17] Here "simulation" refers not to the real cosmological simulations, but generation of nonlinear realizations based on the same perturbative forward model as in FBI. For more details, see [171].

amplitude of the initial density field A than the standard P+B combination. The leading hypotheses for explaining this dramatic difference is that the FBI uses more information from higher order n-point functions and exploits at the field level some non-perturbative features of the nonlinear evolution which come with no new free parameters (e.g., large displacements). However, due to complexity of both FBI and SBI even in the perturbative framework, this has not been rigorously proven and a clear understanding of where would the additional information in FBI exactly come from is still lacking.

Similar analysis was presented in the data challenge for beyond the two-point galaxy clustering statistics [212]. For example, Fig. 22 in [212] shows the constraints on the amplitude of the linear field for the FBI and the standard EFT-based joint power spectrum and bispectrum analysis for the same volume as in [171], although with less dense sample. In this case the difference between the two methods is less dramatic (see the Sec. 6.3 of [212] for more details about the comparison), but it is hard to straightforwardly compare them. This is mainly due to fact that the $k_{\rm max}$ used in the P+B analysis is larger from the scale cuts in the FBI. In addition, for this larger $k_{\rm max}$ in the "restricted" EFT P+B analysis only the errors are presented (the posterior is centered at the best-fit values for the analysis with lower $k_{\rm max}$). This leaves several open questions. Is the P+B analysis biased for smaller $k_{\rm max}$? Are the error bars different for smaller $k_{\rm max}$? And how do they compare to the FBI results and the SBI errors from [171]? Providing clear answers is very important, both to get a better understanding of theoretical aspects of FBI and to get firmer basis for its applications in practice and comparison to the conventional analyses.

The goal of this chapter is to shed new light on these questions, focusing on the usual perturbation theory P+B analysis in the extremely simplified but realistic settings. We use a simple, tractable and intuitive Eulerian perturbative model (without infrared resummation), with the simple Gaussian likelihood and covariance. Validity of such model and likelihood on large scales can be easily verified in practice and we will provide several arguments for it. In a nutshell, such simplified approach is possible given that the analysis presented in [171] follows a particularly simple setup. It uses a simulation box with a relatively small volume of $V = (2\ {\rm Gpc}/h)^3$, which implies relatively large statistical error bars. The studied sample consists of dark matter halos in real space with a very high number density. Additionally, it focuses only on relatively high redshifts $z \geq 0.5$ and large scales, with the maximum wavenumber used in the analysis being $k_{\rm max} \leq 0.12\ h/{\rm Mpc}$. Such setup with regular periodic box, low shot noise, simple tracers without redshift-space distortions or other observational effects and small fluctuations on very large scales, is ideal for understanding FBI using simple analytical methods. Importantly, some of our results then can be easily derived in a very clear and transparent way, without the need to relay on MCMC analyses only. Our approach and results crucially rely on several key assumptions that simplify the description of the system significantly. We would like to clearly state them at the very beginning, given that we will use them throughout the following.

First, we assume that the field of dark matter halos in real space for given data cuts and redshifts is well described by Eulerian perturbation theory. Although the forward model used in [171] is based on Lagrangian perturbation theory, one can show that the two schemes are practically indistinguishable on large scales, given the statistical error bars of the specified simulation box with $V = (2\ {\rm Gpc}/h)^3$. We provide justification and more details in Sec. VI D. An important consequence of this fact is that the large displacements beyond those contained in Eulerian kernels cannot play a major role in explaining the results of [171].

Assuming that the Eulerian description is sufficient, we further assume that all nonlinearities are controlled by only one parameter—the variance of the density field Δ^2 at the scale k—and that this

parameter is small.[18] It is given by

$$\Delta^2(k) = \frac{1}{2\pi^2}\int_0^k q^2 dq P_{\rm lin}(q)\ , \tag{289}$$

where $P_{\rm lin}(k)$ is the linear matter power spectrum. In a ΛCDM-like cosmology, for $k_{\rm max} = 0.12\ h/{\rm Mpc}$ and $z = 0.5$, the variance of the density field is approximately $\Delta^2 \approx 0.13$, and it gets smaller at higher redshifts and for smaller values of $k_{\rm max}$. Importantly, we do not only assume that the relative size of each individual term in perturbative calculation is given by $\Delta^2(k_{\rm max})$, but that the *total* loop contribution to the correlation functions has the same size. This is a nontrivial assumption, since the total one-loop contributions to the higher order n-point function can have a large number of terms multiplied by a large number of different nuisance parameters. In other words, we assume that while all bias parameters and counterterms can be numbers of order $O(1)$, for a realistic tracers they cannot conspire to give large loop corrections. We will explicitly check that this is the case for a typical sample using large-volume numerical simulations and show that deviations of measured correlation functions from the leading order perturbative predictions are roughly given by $\Delta^2(k)$ on large scales. Even though our assumption puts some constraints on the allowed values of nuisance parameters, it is essential for the perturbation theory to make sense and it is sometimes referred to as the perturbativity prior [132].

With the Eulerian forward model where all nonlinearities are controlled by a single parameter Δ^2 it is easy to compute a few leading order correlation functions. Since we are interested in comparison to the FBI analysis of [171], we use the cubic model. Within this model, one can consistently compute the one-loop power spectrum, the tree-level bispectrum and the tree-level trispectrum (T). Higher-order n-point functions and additional loop corrections require quartic and higher-order nonlinearities. Furthermore, since we are focused on large scales only, their impact on cosmological constraints should be much less significant. To estimate this, we can approximate the signal-to-noise ratio (SNR) in a connected n-point function of the galaxy density field δ_g as follows

$$({\rm SNR})_n^2 \approx \int^{k_{\rm max}} \frac{d^3\boldsymbol{k}_1}{(2\pi)^3}\cdots\int^{k_{\rm max}} \frac{d^3\boldsymbol{k}_n}{(2\pi)^3}\frac{\langle\delta_g(\boldsymbol{k}_1)\cdots\delta_g(\boldsymbol{k}_n)\rangle_c^2}{P_g(k_1)\cdots P_g(k_n)} \sim N_{\rm pix.}\Delta^{2(n-2)}(k_{\rm max})\big[1+O(\Delta^2(k_{\rm max}))\big]\ , \tag{290}$$

where $P_g(k)$ is the galaxy power spectrum, $N_{\rm pix.} \equiv V\int^{k_{\rm max}} d^3\boldsymbol{k}/(2\pi)^3$ is the number of pixels or Fourier modes in a survey and $\Delta^2(k_{\rm max}) \lesssim 0.1$ for the range of redshifts and scales of interest. This estimate is correct up to $O(1)$ overall multiplicative factor which depends on fiducial values of bias parameters which we neglect for simplicity. For example, for the power spectrum ($n = 2$), the correct estimate is $({\rm SNR})_2^2 = N_{\rm pix.}/2$, up to small corrections of order $\Delta^2(k_{\rm max})$.[19] One can compute the exact SNR using the standard Gaussian likelihood that we will describe below and check that for higher-order n-point functions the agreement with the estimate above is rather good for small values of nonlinear bias parameters. This is shown in Fig. 22 where the exact calculations for the SNR of P, B and T are compared to Equation (290). Note that the signal-to-noise in each higher order n-point function is roughly one order of magnitude smaller than in the previous one, given the value of $\Delta^2(k_{\rm max})$. This implies that, in the absence of almost perfect degeneracies, the information content of the nonlinear field on large scales is almost saturated by the leading summary statistics. One well-known strong degeneracy for biased tracers in real space is between the amplitude of the linear fluctuations A and the linear bias b_1. This degeneracy is broken by combining the power spectrum and bispectrum, and hence the total constraining power of real space data is suppressed by Δ^2 compared to the total signal in the power spectrum.

[18] Other parameters, such as the Lagrangian radius of dark matter halos, are irrelevant, given the large number density of the sample, which implies low halo mass.

[19] The same parameter Δ^2 controls both the SNR for the n-point functions and the corrections to the SNR. This is expected in perturbation theory where all nonlinearities are controlled by Δ^2. Small corrections to Equation (290) are captured by the non-Gaussian covariance matrix. The smallness of Δ^2 justifies why using the simple Gaussian covariance matrix is sufficient on large scales.

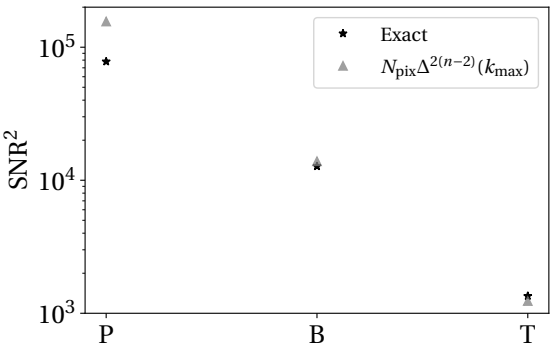

Figure 22: SNR for the power spectrum, the bispectrum and the trispectrum. The stars correspond to the exact evaluation using the standard Gaussian likelihood described in Sec. VI B (using fiducial bias parameters from Sec. VI F), while the triangles correspond to the estimate in Equation (290). The agreement between the two for the higher-order n-point functions is quite good. A factor of 2 difference for the power spectrum is expected, since the exact result for the SNR is given by $(\text{SNR})_2^2 = N_{\text{pix.}}/2$.

What do our assumptions imply for the FBI analysis? Given the Eulerian description, perturbativity prior and the negligible shot noise, one can compute the posterior for the FBI perturbatively. It is then possible to argue that FBI is equivalent, order by order in perturbation theory, to the analysis based on all n-point correlation functions that can be reliably computed up to the given order in the perturbative expansion used for the nonlinear field in the FBI. On the level of equations this equivalence has been explicitly demonstrated in [213] at leading order in Δ^2 and more recently in [214] at subleading order in Δ^2 and for the full cubic bias model. We refer the reader to these references for further details. Showing this connection using simulation data is more challenging and beyond the scope of this thesis. Here we will ask two concrete simple questions that can still shed some light on the connection between n-point functions and FBI. First, we can ask how well the amplitude A can be measured with P+B combination using a simple model and the usual Gaussian likelihood. Second, using the Fisher matrix, we can ask if addition of the trispectrum to the P+B analysis significantly improves the errors.

Before we move on to the verification of our model and approximations in simulations and the main analysis, let us first give a simple estimate of the results that we expect. The main advantage of using Eulerian perturbation theory is that the model is a simple polynomial in nuisance parameters and the amplitude of the linear density field. This means that the logarithm of the full likelihood in Equation (288) takes particularly simple form and one can easily estimate the precision with which different parameters can be measured.

First of all, the highest signal to noise is in the power spectrum. On large scales the most dominant term is the linear theory contribution proportional to $b_1^2\alpha^2$. Therefore, the galaxy power spectrum in real space measures the combination $b_1\alpha$ very precisely. The relative error can be estimated as follows

$$\frac{\sigma(b_1\alpha)}{b_1\alpha} = \frac{1}{\sqrt{2N_{\text{pix.}}}}\,, \tag{291}$$

where the number of pixels in practice can be estimated as $N_{\text{pix.}} \approx V k_{\text{max}}^3/(2\pi)^3$. Note that this error is different from the error on the amplitude of the power spectrum, since the amplitude of the power spectrum is proportional to α^2. In practice, this is the most-well measured number in the real space analysis and we can essentially treat it as fixed. The degeneracy between the linear bias and α is broken at the level of the power spectrum only very mildly by the loop contributions. However, given several free higher order bias parameters, the amplitude α is essentially unconstrained by the power spectrum alone.

Moving to higher order statistics, assuming $b_1\alpha$ is fixed, the amplitude of the tree-level bispectrum

is proportional to a single power of α. While the bispectrum also contains quadratic biases as free parameters, the angular dependence allows for degeneracy with α to be partially broken. Therefore, we can assume that the error on the amplitude of the linear field from the joint power spectrum and bispectrum analysis is given by the bispectrum SNR. We can write

$$\frac{\sigma(\alpha)}{\alpha} \approx \text{few} \times \frac{1}{\Delta(k_{\rm max})\sqrt{N_{\rm pix.}}} \,, \tag{292}$$

where we have used Equation (290) and the numerical factor of a few depends on how much information is lost by marginalizing over b_2 and $b_{\mathcal{G}_2}$, which can be easily obtained using the Fisher matrix.

Using $V = (2\ \text{Gpc}/h)^3$ and $k_{\rm max} = 0.1\ h/\text{Mpc}$, we can estimate

$$\frac{\sigma(b_1\alpha)}{b_1\alpha} \approx 0.004 \,, \qquad \text{and} \qquad \frac{\sigma(\alpha)}{\alpha} \approx \text{few} \times 0.02 \,. \tag{293}$$

These numbers give an idea of how well the P+B analysis in real space should perform. In reality, the constraints are a bit weaker, due to the noise and nonlinearities we neglected in these arguments. However, we can already see that the baseline expectation for the error on α is roughly $O(5-10\%)$, which agrees with the field level result of [171]. Furthermore, the simple estimates above give us two additional important predictions.

First, since there are no as strong degeneracies in the bispectrum as the $b_1 - \alpha$ degeneracy in the power spectrum, the trispectrum is not expected to significantly improve the errors. This is due to much smaller SNR in the trispectrum as well as existence of new cubic bias parameters that have to be measured as well. We expect $\Delta^2(k_{\rm max}) \approx O(10\%)$ improvements from inclusion of the trispectrum in the analysis. Second, in a P+B analysis we get a scaling of the error with $k_{\rm max}$ which is given by

$$\frac{\sigma(\alpha)}{\alpha} \sim \frac{k_{\rm max}^{-3/2}}{\Delta(k_{\rm max})} \,. \tag{294}$$

Note that the typical amplitude of fluctuations $\Delta(k_{\rm max})$ grows with $k_{\rm max}$, which implies that the error on α shrink faster than the naive $k_{\rm kmax}^{-3/2}$ scaling. This makes sense, since the information comes from the bispectrum whose SNR grows as the field becomes more non-Gaussian on large scales, in addition to the simple mode counting. Note that the higher order n-point functions would lead to different scaling with $\Delta(k_{\rm max})$. This is yet another way to estimate where does the information come from, by carefully measuring the dependence of the inferred errors on $k_{\rm max}$.

In what follows, we will show that all three predictions (the size of the error bars, importance of the trispectrum and scaling with $k_{\rm max}$) are confirmed in a real analysis.

D Why is Eulerian description sufficient on large scales?

Let us now justify why an Eulerian treatment of the galaxy overdensity is reliable on large scales. It is well known that Lagrangian perturbation theory (LPT) reproduces very well the nonlinear fields of dark matter or biased tracers, in real and redshift space. Even when using the Zel'dovich approximation only, one can generate realizations with familiar filamentary structures, which correlate well with the simulations. For this reason, LPT is often employed when making perturbative maps or doing detailed comparisons of simulations and perturbation theory. Furthermore, LPT automatically accounts for effects such as broadening of the BAO peak (or other features) and can be easily implemented in practice on a grid. Given all this, it is natural that the field-level analyses use forward modeling based on LPT to evolve a given initial linear density field (provided cosmological, noise and bias parameters). LPT up to cubic order with all relevant bias operators in the initial field was used in [171], as well as in the previous works [182, 197, 198, 205–207].

However, most of the field-level inference has been done so far on very large scales. For example, the largest $k_{\rm max}$ cut used in [171] is $k_{\rm max} = 0.12\ h/{\rm Mpc}$. In this regime, there is a significant simplification in the nonlinear dynamics and Eulerian perturbation theory (EPT) is sufficient to take into account all relevant nonlinearities in the field. While all theoretical calculations can be done in LPT as well, the advantage of EPT is that it is easier to work with, it expresses the nonlinear field in terms of the initial conditions in a simple way and allows to more easily disentangle different physical effects that can affect the analysis (such as displacements and broadening of the BAO peak, versus the growth of structure). Since one of the goals of this section is to better analytically understand the error bars for the field-level and conventional analyses, we will use EPT.

Why is EPT sufficiently good on large enough scales? There are two main reasons:

- First, the nonlinear fields generated using cubic EPT and LPT up to $k_{\rm max} = 0.12\ h/{\rm Mpc}$ and at redshifts $z = 0.5$ and $z = 1$ are almost indistinguishable, i.e., their cross-correlation coefficient is close to 1. This has been verified explicitly at the field level [215] and can be easily understood analytically. We will review the main arguments below.

- Second, even though the FBI is based on comparing the fields predicted in theory and data, it eventually boils down to measuring some summary statistics due to the averaging over all possible initial conditions, which is equivalent to computing the expectation values [213, 214]. It is well known that in EPT and LPT the expectation values are in much closer agreement than what the cross-correlation coefficient indicates, making EPT even more adequate in approximating LPT.

Let us review these two arguments in detail. We begin by estimating the cross-correlation coefficient between EPT and LPT. Let us imagine that in the initial (Lagrangian) coordinates $\boldsymbol{q}$ we have a linear field $\delta_1(\boldsymbol{q})$. We can imagine that we evolve this field in two different ways. In EPT, if we use just the linear evolution, the final field is just proportional to $\delta_1(\boldsymbol{q})$. On the other hand, in LPT we have to displace particles from their initial coordinates $\boldsymbol{q}$ to their final Eulerian positions $\boldsymbol{x}$, such that $\boldsymbol{x} = \boldsymbol{q} + \boldsymbol{\psi}(\boldsymbol{q})$. The leading contribution to the nonlinear displacement field $\boldsymbol{\psi}(\boldsymbol{q})$ is given by the Zel'dovich approximation $\boldsymbol{\psi}_1(\boldsymbol{q}) = -\boldsymbol{\nabla}/\nabla^2\delta_1(\boldsymbol{q})$, such that the leading order (LO) LPT field is given by

$$\delta^{\rm LO}_{\rm LPT}(\boldsymbol{k}) = \int d^3\boldsymbol{q} e^{-i\boldsymbol{k}\cdot\boldsymbol{\psi}_1(\boldsymbol{q})} e^{-i\boldsymbol{k}\cdot\boldsymbol{q}}\ . \tag{295}$$

Note that here we assume that all quantities are evaluated in linear theory at the final redshift and we suppress explicit time dependence.

Starting from this equation one can compute the cross-correlation coefficient between the leading order LPT and EPT fields. It is given by

$$\begin{aligned}\langle \delta^{\rm LO}_{\rm EPT}(\boldsymbol{k}_1)\delta^{\rm LO}_{\rm LPT}(\boldsymbol{k}_2)\rangle &= \int d^3\boldsymbol{q}_1 d^3\boldsymbol{q}_2 \left\langle \delta_1(\boldsymbol{q}_1) e^{-i\boldsymbol{k}_2\cdot\boldsymbol{\psi}_1(\boldsymbol{q}_2)}\right\rangle e^{-i\boldsymbol{k}_1\cdot\boldsymbol{q}_1 - i\boldsymbol{k}_2\cdot\boldsymbol{q}_2} \\ &= i\frac{d}{d\lambda}\int d^3\boldsymbol{q}_1 d^3\boldsymbol{q}_2 \left\langle e^{-i(\lambda\delta_1(\boldsymbol{q}_1) + \boldsymbol{k}_2\cdot\boldsymbol{\psi}_1(\boldsymbol{q}_2))}\right\rangle e^{-i\boldsymbol{k}_1\cdot\boldsymbol{q}_1 - i\boldsymbol{k}_2\cdot\boldsymbol{q}_2}\Big|_{\lambda=0}\ .\end{aligned} \tag{296}$$

The expectation value can be computed using the cumulant theorem. For a random variable X, we have

$$\langle e^{-iX}\rangle = \exp\left(\sum_{n=0}^{\infty}\frac{(-i)^n}{n!}\langle X^n\rangle_c\right)\ , \tag{297}$$

where the cumulants $\langle X^n\rangle_c$ in our case represent the connected n-point functions. Since at LO all fields in the exponent are Gaussian, the only nontrivial cumulant is the power spectrum. After taking the

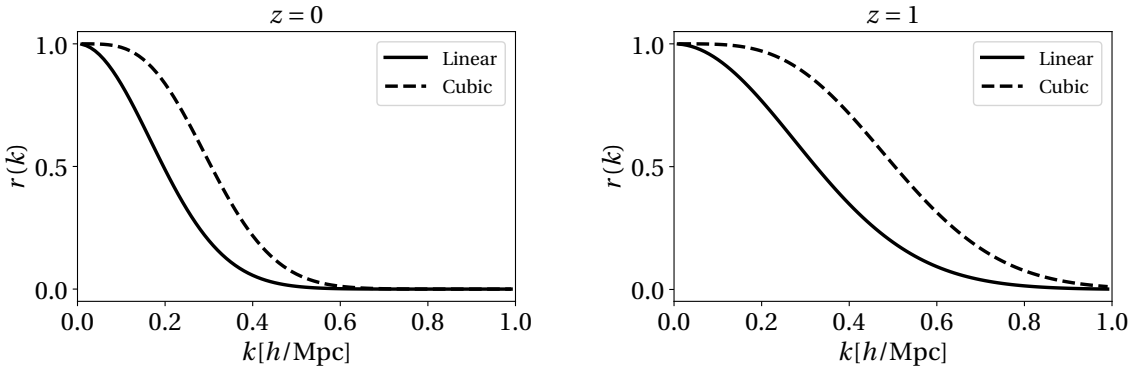

Figure 23: Estimate of the cross correlation coefficient between the EPT and LPT fields when both theories are computed up to same order in perturbation theory (linear or cubic model), at $z = 0$ (left panel) and $z = 1$ (right panel). These curves have the usual "s" shape and are in excellent agreement with the findings from the explicit grid computations of EPT and LPT (for example, see Fig. 10 in [215]).

derivative and setting $\lambda = 0$ we get

$$
\begin{aligned}
i\frac{d}{d\lambda}\langle e^{-i(\lambda\delta_1(\boldsymbol{q}_1)+\boldsymbol{k}_2\cdot\boldsymbol{\psi}_1(\boldsymbol{q}_2))}\rangle\Big|_{\lambda=0} &= i\frac{d}{d\lambda}e^{-\frac{1}{2}\left(\lambda^2\langle\delta_1^2(\boldsymbol{q}_1)\rangle+2\lambda\boldsymbol{k}_2\cdot\langle\delta_1(\boldsymbol{q}_1)\boldsymbol{\psi}_1(\boldsymbol{q}_2)\rangle+k_2^ik_2^j\langle\psi_{1i}(\boldsymbol{q}_2)\psi_{1j}(\boldsymbol{q}_2)\rangle\right)}\Big|_{\lambda=0} \\
&= -i\boldsymbol{k}_2\cdot\langle\delta_1(\boldsymbol{q}_1)\boldsymbol{\psi}_1(\boldsymbol{q}_2)\rangle e^{-\frac{1}{2}k_2^2\sigma_v^2}\,. \qquad (298)
\end{aligned}
$$

Using this in the Equation (296) gives the following exact result for the LO cross correlation

$$
\langle\delta^{\rm LO}_{\rm EPT}(\boldsymbol{k})\delta^{\rm LO}_{\rm LPT}(-\boldsymbol{k})\rangle' = P_{\rm lin}(k)e^{-\frac{1}{2}k^2\sigma_v^2}\,. \qquad (299)
$$

Assuming that at LO all power spectra are approximately equal to the linear power spectrum on large scales, we get the following LO cross-correlation coefficient

$$
r^{\rm LO}(k) \equiv \frac{\langle\delta^{\rm LO}_{\rm EPT}(\boldsymbol{k}')\delta^{\rm LO}_{\rm LPT}(\boldsymbol{k})\rangle'}{\sqrt{P_{\rm EPT}(k)P_{\rm LPT}(k)}} \approx e^{-\frac{1}{2}k^2\sigma_v^2}\,, \qquad (300)
$$

where prime on the correlation function indicates that $(2\pi)^3\delta^D(\boldsymbol{k}+\boldsymbol{k}')$ has been removed, $P_{\rm EPT}(k)$ and $P_{\rm LPT}(k)$ are leading order power spectra and the velocity dispersion is given by

$$
\sigma_v^2 = \frac{1}{6\pi^2}\int_0^\infty dq P_{\rm lin}(q)\,. \qquad (301)
$$

Let us here provide an intuitive justification for it. The major difference between EPT and LPT is the displacement field. Therefore, on scales larger than the typical variance of the displacement field, given by σ_v^2, the EPT and LPT fields are the same. On smaller scales they rapidly decorrelate and $r^{\rm LO}(k)$ goes to zero. The value of σ_v^2 in a ΛCDM-like cosmology is approximately

$$
\sigma_v^2(z) = 36\left(\frac{D_+(z)}{D_+(0)}\right)^2\ ({\rm Mpc}/h)^2\,. \qquad (302)
$$

This implies a significant decorrelation between the leading order EPT and LPT fields, even at $k_{\rm max} = 0.1\ h/{\rm Mpc}$.

However, in our analysis we will go beyond the leading order approximation. If we use cubic models for both EPT and LPT the agreement between the two improves. In this case a part of the nonlinearities induced by the large displacements is taken into account in EPT. Since the two theories match at one-loop

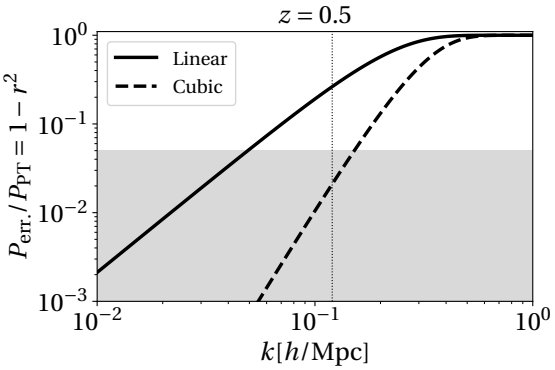

Figure 24: Power spectrum of the difference between LPT and EPT fields at $z = 0.5$ in linear and cubic models, normalized by $P_{\rm PT}$, see Equation (305). The shaded region corresponds to a relative error smaller than 5%. Note that $P_{\rm err.}/P_{\rm PT} \approx 2\Delta A/A$. Thin vertical line marks the maximum wavenumber used in the analysis, $k_{\rm max} = 0.12\ h/{\rm Mpc}$. At higher redshifts the agreement between EPT and LPT is even better.

order on large scales, the cross-correlation coefficient at next-to-leading order (NLO) will be better and must scale as

$$r^{\rm NLO}(k) = 1 + O(k^4\sigma_v^4)\ , \tag{303}$$

in the low-k limit, where $O(k^4\sigma_v^4)$ is the leading contribution from the unaccounted two-loop terms. While deriving the exact expression is more involved in this case, we can provide a simple estimate of the cross-correlation coefficient at NLO. In a way identical to what happens in the so-called infrared resummation, we can write

$$r^{\rm NLO}(k) \approx \left(1 + \frac{1}{2}k^2\sigma_v^2\right) e^{-\frac{1}{2}k^2\sigma_v^2}\ , \tag{304}$$

which is guaranteed to produce the correct low-k limit for the cubic EPT and LPT models, while keeping the same exponential suppression on small scales. In Fig. 23 we plot the cross-correlation coefficient for linear and cubic model, corresponding to Equation (300) and Equation (304) respectively. Our estimates agree very well with measurements of the cross-correlation coefficient from explicit EPT and LPT maps built at different order in perturbation theory (for example, see Fig. 10 in [215]). Going to cubic order in EPT and having the correct one-loop power spectrum, makes the fields correlate much better.

The cross-correlation coefficient is related to the error that one makes approximating LPT with EPT. If we define the "error" field as the difference $\delta_{\rm err} \equiv \delta_{\rm LPT} - \delta_{\rm EPT}$, then the power spectrum of this difference can be approximated as follows

$$P_{\rm err.} = \langle|\delta_{\rm LPT} - \delta_{\rm EPT}|^2\rangle' = P_{\rm LPT} + P_{\rm EPT} - 2\sqrt{P_{\rm LPT}P_{\rm EPT}}r \approx P_{\rm PT}(1 - r^2)\ , \tag{305}$$

where in order to make the last approximation we used that $P_{\rm LPT} \approx P_{\rm EPT} = P_{\rm PT}$ and that $r \approx 1$ on large scales, such that $2(1 - r) \approx (1 + r)(1 - r) = 1 - r^2$. In other words, the quantity $1 - r^2$ measures the relative size of the error power spectrum compared to the $P_{\rm PT}$. In Fig. 24 we plot $1 - r^2$ for redshift $z = 0.5$. We can see that the relative error of approximating LPT with EPT is at the percent level for the scale cuts used in the analyses, $k_{\rm max} = 0.1\ h/{\rm Mpc}$ and $k_{\rm max} = 0.12\ h/{\rm Mpc}$. This is much smaller than the $O(5\%)$ relative error on the amplitude of the linear field that can be obtained from a $(2\ {\rm Gpc}/h)^3$ volume box of biased tracers in real space [171]. This implies that on large scales and within given statistical errors, Eulerian and Lagrangian perturbative forward models are practically indistinguishable.

One may object that the arguments we presented so far apply only to the dark matter field. However,

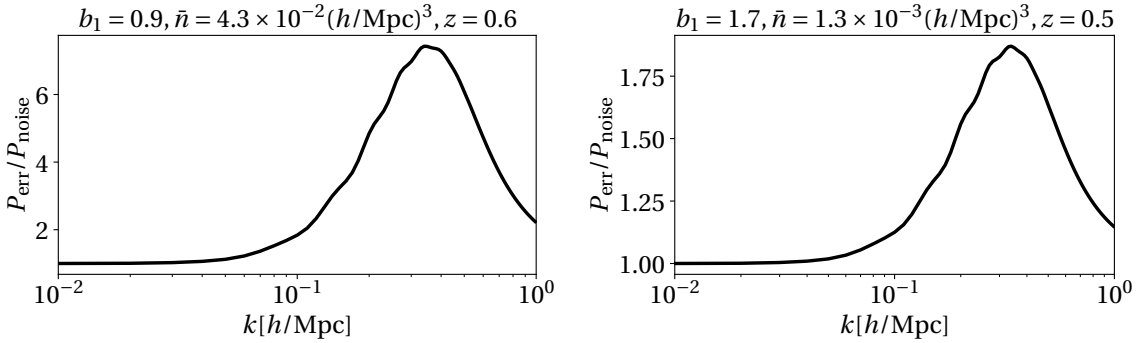

Figure 25: $P_{\mathrm{err}}/P_{\mathrm{noise}}$ as function of scale for two biased tracers similar the ones considered in [191] (*Left)* and in [171] (*Right*).

the situation for the biased tracers is in some sense even better, thanks to the additional source of the error due to the discrete sampling of the halo density field. Assuming constant shot noise power spectrum on large scales with the amplitude P_{noise}, we can write the ratio of the total error power spectrum to the noise as follows[20]

$$\frac{P_{\mathrm{err.}}}{P_{\mathrm{noise}}} \approx 1 + \left(1 - r^2\right) \frac{b_1^2 P_{\mathrm{lin}}(k)}{P_{\mathrm{noise}}} \,. \tag{306}$$

As expected, for biased tracers with large enough shot noise, all modeling errors are negligible compared to the shot noise P_{noise} on large scales, when r goes to one. In the left panel of Fig. 25 we plot the ratio from the previous formula for a very dense tracers with $\bar{n} \sim 4.3 \times 10^{-2} (h/\mathrm{Mpc})^3$. This agrees very well with measurements from simulations (for example, see the dark orange line on the left panel of Fig. (20) in [191]). In the right panel of Fig. 25 we show the same ratio but for the tracer considered in [171]. While the modeling error is roughly 10% larger than the sampling noise at $k_{\mathrm{max}} = 0.12\ h/\mathrm{Mpc}$, the sampling noise itself is just a small fraction of the total power spectrum on these scales. Therefore, the cross-correlation coefficient between EPT and LPT can be very easily understood both for dark matter field and biased tracers. In both cases the two perturbative approaches agree at the percent level on relevant scales.

Let us also briefly discuss the second reason for why EPT is a good approximation for LPT on large scales. So far we estimated the cross correlation coefficient and the error power spectrum and showed that on scales relevant for our analysis the EPT and LPT fields match at the percent level. However, in the FBI one performs the marginalization over all initial conditions, resulting in effective measurement of some summary statistics from the maps. In the perturbative regime these statistics are just the leading n-point functions [213]. Regardless of the details, we want to argue that the estimate of the mismatch based on the cross-correlation coefficient is the most pessimistic one. It is dominated by the velocity dispersion induced by the large bulk flows. However, it is well known that due to the equivalence principle the effects of these bulk flows exactly cancel in all summary statistics (with the exception of features such as BAO, which is irrelevant on large scales) [150–152, 216–219]. Therefore, the agreement between EPT and LPT predictions can be only better in practice than the $1 - r^2$ estimate indicates. For example, even though the Zel'dovich and linear Eulerian fields strongly decorrelate for $k \gtrsim 1/\sigma_v$, the two power spectra do not differ dramatically. The same is true for any higher order correlation function or other statistical quantities.

In conclusion, for the range of scales and redshifts used in [171] and our analysis in this thesis,

[20] For galaxies we assume $\delta_{g,\mathrm{LPT}} - \delta_{g,\mathrm{EPT}} = \delta_{\mathrm{err.}} + \epsilon_g$, where $\langle \epsilon_g \epsilon_g \rangle' = P_{\mathrm{noise}}$.

the standard Eulerian perturbation theory (without the IR resummation) and Lagrangian perturbation theory describe the halo density field equally well. Since there is no distinction between the two nonlinear fields on the level of summary statistics, all conclusions based on EPT must be valid for LPT as well.

E Validation on simulations

The goal of this section is to validate our model and assumptions using large-volume numerical simulations. Our results are not new, this has been already done in a large body of work on the power spectrum and bispectrum and comparison to large-volume simulations with a similar setup [130, 180, 183, 220–224]. In particular, we use the set of 10 boxes of the PT challenge simulation suite [225] (we refer the reader to this reference for details on the simulations). With the total volume of 566 $(\mathrm{Gpc}/h)^3$ at redshift $z = 0.6$, these simulations allow for very precise comparison of data and theory even on very large scales. We use the measured power spectrum and bispectrum of dark matter halos in real space, with number density and redshifts similar to the one analyzed in [171].

In particular, we want to test the following:

- Eulerian perturbation theory predicts the power spectrum and the bispectrum on large scales well enough and leads to unbiased inference of α.
- Large displacements on large scales do not produce any effect beyond the one captured by Eulerian perturbation theory. In particular, one does not have to do the infrared resummation. The difference between the infrared resummed and baseline model appears only on smaller scales, beyond $k_{\rm max}$ used in our analysis.
- Typical deviation of the best-fit model and the data for $k > k_{\rm max}$ is within the theoretical error estimate. This means that the realistic field of biased tracers has small fluctuations in agreement with perturbativity prior. Correspondingly, any combination of higher order bias and nonlinear correction has to be small on the scales considered, even though all bias coefficients may be of order $O(1)$.
- The error bars on the amplitude α are roughly given by the simple estimate in Equation (292). The error bars do not change significantly when the theoretical error is included in the analysis, indicating that the higher bias and loop corrections are not very relevant on very large scales used in the analysis.

In order to do this, we will use the measurements from the full volume of 566 $(\mathrm{Gpc}/h)^3$, but fit it with the covariance matrix rescaled to match the box of 8 $(\mathrm{Gpc}/h)^3$ which has the same volume as in [171]. We will also use the same two cuts, $k_{\rm max} = 0.1\ h/\mathrm{Mpc}$ and $k_{\rm max} = 0.12\ h/\mathrm{Mpc}$. We perform the joint power spectrum and bispectrum analysis using the theoretical model and likelihoods described above. We vary the following parameters

$$\theta = \{\alpha, b_1, b_2, b_{\mathcal{G}_2}, b_{\Gamma_3}, c_s^2, c_1, c_2, c_3\}\ , \tag{307}$$

that include all the necessary bias parameters to compute the one-loop power spectrum and the tree-level bispectrum. It is important to emphasize that we also make an appropriate binning of the model, integrating the theoretical prediction in each bin with the appropriate weight to take into account the number of Fourier modes for each momentum k. This is important for the lowest k-bins, but it does not significantly affect the final best fit and the error bars, since they are predominantly determined by the highest k modes in the analysis.

The main results are shown in Fig. 26, Fig. 27 and Fig. 28 and summarized in Tab. I. The first two figures show the residual plots for the power spectrum and the bispectrum obtained from the joint analysis up to $k_{\rm max} = 0.1\ h/\mathrm{Mpc}$. In all plots the statistical error bars are shown for the full PT Challenge

	P+B
$k_{\rm max} = 0.1\ h/{\rm Mpc}$	
α	0.960±0.085
$b_1/b_{1,\rm fid}$	1.051±0.094
$k_{\rm max} = 0.12\ h/{\rm Mpc}$	
α	1.026±0.052
$b_1/b_{1,\rm fid}$	0.979±0.051

Table I: 1σ constraints and best fit parameter for the PT Challenge P+B analysis using the covariance with the volume of $V = (2\ {\rm Gpc}/h)^2$.

volume, which is roughly 70 times larger than the volume of $V = 8\,({\rm Gpc}/h)^3$ that we are interested in. For this reason all error bars are extremely small. For instance, the relative statistical errors in each k-bin in the power spectrum are much smaller than 1%. This exquisite precision allows us to do the stringent test of the model.

Let us first focus on the power spectrum residuals shown in Fig. 26. The left panel shows the residuals together with the estimate of the 2-loop theoretical error. We can see that the best fit reproduces the data well within the theoretical uncertainties of perturbation theory. The two points with statistically significant deviation (2-3 sigma) are likely a random fluctuation, since they cannot be fitted by a sufficiently smooth curve. To verify this, we also fit the same data with the nonlinear `PyBird` code [226] and find almost identical results.[21] Importantly, the higher k-bins beyond $k_{\rm max} = 0.1\ h/{\rm Mpc}$ used in the analysis, lay within the estimate of the 2-loop theoretical error, indicating that the total galaxy field indeed follows expectations from perturbation theory. It is instructive to do the extrapolation of the best-fit model with $k_{\rm max} = 0.1\ h/{\rm Mpc}$ to smaller scales. This is shown on the right panel of Fig. 26. The residuals remain within the theoretical error which grows very fast with k. More importantly, one can see that using the Eulerian model which is the baseline for our analysis is insufficient on small scales to capture the BAO wiggles appropriately. However, the residual BAO wiggles appear only for $k \gtrsim 0.2\ h/{\rm Mpc}$, which is never used in our baseline analysis. This is an important confirmation that large displacements on large scales are captured well by Eulerian perturbation theory.

Similar residual plots for the bispectrum are shown in Fig. 27 for two different triangle configurations. One can see that the model fits the data well and that the residuals for $k > k_{\rm max}$ are within the estimate of the one-loop theoretical error.

It is important to emphasize once again that the statistical errors in these two figures are roughly 8 times smaller than the case of interest corresponding to the volume $V = 8\ ({\rm Gpc}/h)^3$. This implies that for our main analysis, the theoretical model presented above is perfectly adequate to capture all relevant nonlinearities and it should lead to unbiased answers. This can be verified using the covariance matrix rescaled to the volume of the baseline analysis, still fitting the PT Challenge data. The corresponding triangle plots for $\alpha, b_1, b_1\alpha$ in this analysis for two different choices of $k_{\rm max}$ are shown in Fig. 28. One can see that the inference of the amplitude and b_1 is unbiased, indicating that the simple Eulerian model guarantees a sufficient level of accuracy in describing the galaxy overdensity field on large scales and for the volume considered.

It is also interesting to look at the behavior of errors in Fig. 28 and Tab. I as a function of $k_{\rm max}$. We can see the improvement on the combination $b_1\alpha$ is smaller than the mode counting expectation. This is due to the fact that marginalization over higher order biases effectively introduces an error which makes the constraints weaker. On the other hand, the errors on b_1 and α improve significantly with the increasing $k_{\rm max}$. This is in line with the expectation from the information content of the bispectrum,

[21] Equivalent approaches for large scale structure correlators have been implemented in other publicly available codes, such as `CLASS-PT` [227], `velocileptors` [228] and `CLASS-1loop` [229].

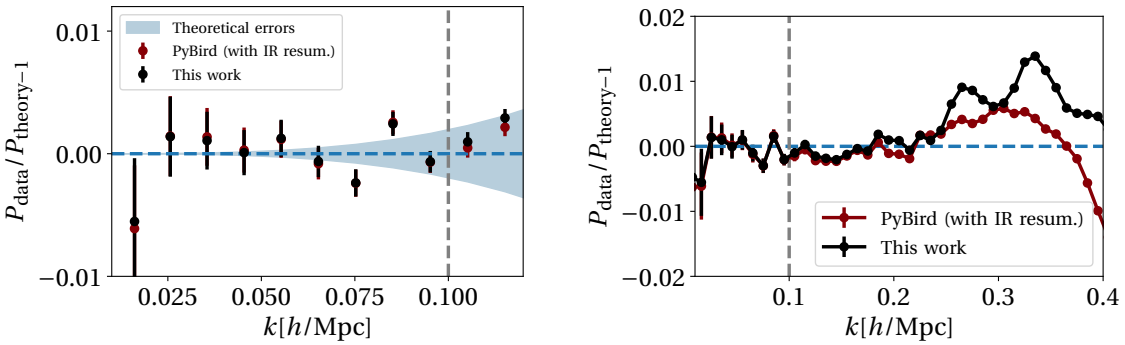

Figure 26: Comparison between the data and the best-fit model in a joint power spectrum and bispectrum analysis up to $k_{\rm max} = 0.1\ h/{\rm Mpc}$ at $z = 0.6$ for the galaxy power spectrum. In the left panel the shaded region represents the 2-loop theoretical error of Equation (285). An independent fit obtained using the `PyBird` code which accounts for the IR resummation is also shown. Statistical error bars on data correspond to the full PT Challenge volume.

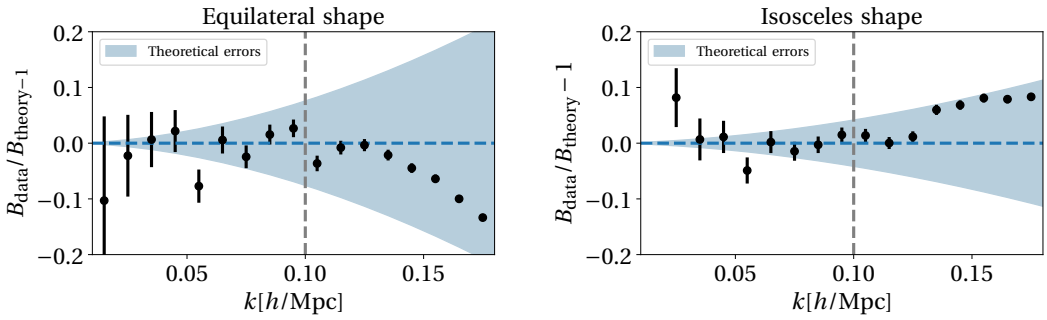

Figure 27: Comparison between the data and the best-fit model in a joint power spectrum and bispectrum analysis up to $k_{\rm max} = 0.1\ h/{\rm Mpc}$ at $z = 0.6$ for the equilateral and isosceles (i.e., $k_1 = 0.015\ h/{\rm Mpc}$, $k_2 = k_3$) shapes of galaxy bispectrum. Statistical error bars on data correspond to the full PT Challenge volume. The shaded region represents the 1-loop theoretical error of Equation (286).

where relative errors improve faster than $k_{\rm max}^{-3/2}$ in agreement with Equation (294). In conclusion, the joint P+B analysis of the PT Challenge simulation boxes shows that the theoretical model presented in the previous section fits the data well and leads to errors in agreement with basic arguments based on perturbation theory.

It is interesting to ask how this picture changes when the theoretical error is included in the analysis. As we discussed, the theoretical error is an additional correlated covariance whose purpose is to mimic marginalizing over neglected higher order corrections in the perturbation theory calculations. We repeat the same analysis of the PT Challenge but with the 2-loop power spectrum and 1-loop bispectrum theoretical errors included in the covariance. The results are shown in Fig. 29. The triangle plots for $\alpha, b_1, b_1\alpha$ are shown with and without theoretical errors in the covariance, for two different values of $k_{\rm max}$. The inclusion of the theoretical error makes the constraints on α worse by 10% and 30% at $k_{\rm max} = 0.1\ h/{\rm Mpc}$ and $k_{\rm max} = 0.12\ h/{\rm Mpc}$ respectively. Such results are expected. On larger scales, the impact of theoretical error is much smaller, given that the model is more precise. Nevertheless, the inclusion of the theoretical error does not change the error bars dramatically. In particular, they cannot explain the difference of roughly factor of 3 between the result we find here and the SBI results for the joint P+B analysis in [171]. We will come back to this point in Sec. VI H.

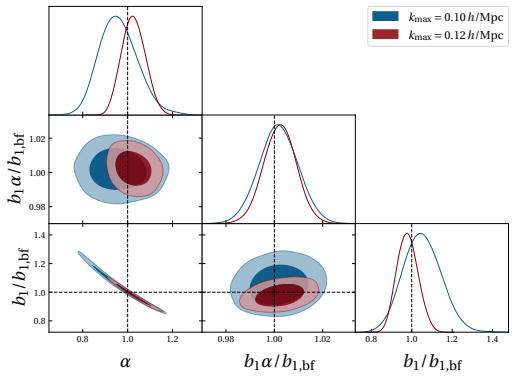

Figure 28: P+B analysis of PT challenge data. The covariance used in the analysis is computed using the reduced volume $V = (2\ {\rm Gpc}/h)^3$.

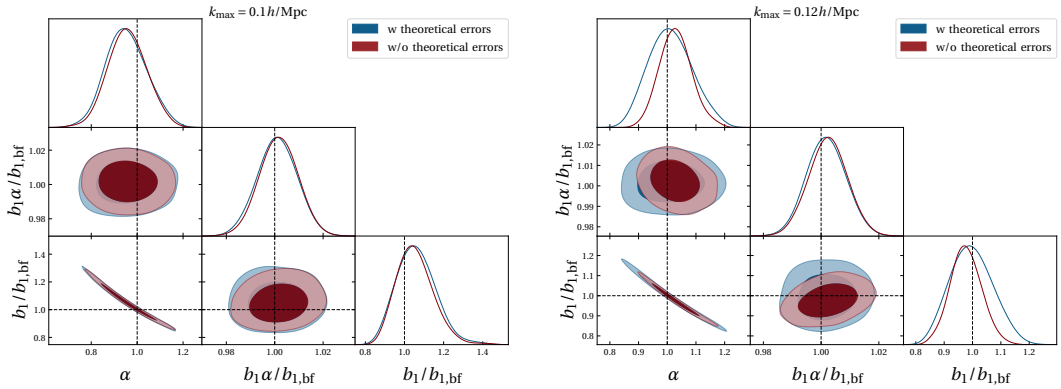

Figure 29: P+B analysis of the PT challenge data with and without theoretical errors for (*Left*) $k_{\rm max} = 0.1\ h/{\rm Mpc}$ and (*Right*) $k_{\rm max} = 0.12\ h/{\rm Mpc}$. The covariance used in the analysis is computed using the reduced volume $V = (2\ {\rm Gpc}/h)^3$.

F P+B and P+B+T forecasts

After extensive tests of the simple EPT model on large scales using large-volume PT Challenge simulations and verifying that it is adequate for the joint P+B analysis, we are finally ready to present the main results. Our goals are to estimate the error bars of the conventional P+B analysis in a box of the volume $V = 8\ ({\rm Gpc}/h)^3$ for tracers similar to the one studied in [171] and [212] and estimate the possible improvements coming from including the galaxy trispectrum in the analysis.

We will not analyze the real simulation data due to unavailability of the trispectrum that is an important part of our results. Instead, using the baseline model (for which we know to match the reality on large scales) we generate the mock data vector containing P, B and T and the wiggly part of the power spectrum for $k > k_{\rm max}$, and perform the analysis of this data vector. In order to do that, for simplicity, we use the same baseline cosmology as the PT Challenge which is very close to standard ΛCDM universe. We also have to specify some fiducial values for the following set of parameters:

$$\theta = \{\alpha, b_1, b_2, b_{\mathcal{G}_2}, b_3, b_{\mathcal{G}_3}, b_{\mathcal{G}_2\delta}, b_{\Gamma_3}, c_s^2, c_{\rm w}^2, c_1, c_2, c_3\}\ . \tag{308}$$

Obviously $\alpha_{\rm fid} = 1$. For other parameters we use two different sets of fiducial values depending on the tracer.

First, we chose the linear bias $b_{1,\rm fid} = 1.7$. We also set the noise to be given by the number density of the dark matter halos in real space equal to $\bar{n} = 1.3 \times 10^{-3}\ ({\rm Mpc}/h)^3$ at fixed redshift $z = 0.5$. The linear

bias, the noise and redshift are chosen to be similar to the SNG sample studied in [171]. The fiducial values of other parameters are far less important, and we verify that explicitly in our analysis. Since the sample is dense, the nonlinear bias parameters are small. We choose the following values: $b_{2,\text{fid}} = b_{\mathcal{G}_{2,\text{fid}}} = b_{\Gamma_3,\text{fid}} = b_3 = b_{\mathcal{G}_3,\text{fid}} = b_{\delta\mathcal{G}_2,\text{fid}} = 0.1$. The fiducial counterterms are given by $c^2_{s,\text{fid}} = 1$, $c^2_{\text{w,fid}} = 0$ and we choose that all fiducial noise free parameters are equal to 1, $c_{1,\text{fid}} = c_{2,\text{fid}} = c_{3,\text{fid}} = 1$. When running the MCMC chains we use the following priors:

$$\begin{gathered}
\mathcal{P}(\alpha) = \mathcal{U}(0.5, 1.5), \\
\mathcal{P}(b_1) = \mathcal{N}(1.0, 5.0), \\
\mathcal{P}(b_2) = \mathcal{P}(b_{\mathcal{G}_2}) = \mathcal{P}(b_{\Gamma_3}) = \mathcal{P}(b_3) = \mathcal{P}(b_{\mathcal{G}_3}) = \mathcal{P}(b_{\delta\mathcal{G}_2}) = \mathcal{N}(0., 1.0), \\
\mathcal{P}(c_s^2) = \mathcal{N}(0, 5.0), \\
\mathcal{P}(c_{\text{w}}^2) = \mathcal{N}(0, 1.0), \\
\mathcal{P}(c_1) = \mathcal{P}(c_2) = \mathcal{P}(c_3) = \mathcal{N}(1, 1)\ .
\end{gathered} \tag{309}$$

The variance of each parameter is chosen to match the choice in [171].

The second set of fiducial nuisance parameters is chosen to describe the sample at $z = 1$ used in the beyond 2-point function challenge paper [212]. From [212] we estimate that $b_1 \approx 2.5$ and $\bar{n} \sim 4.5 \times 10^{-4} (h/\text{Mpc})^3$. For the fiducial values of the higher order bias parameters we use the measurements for a similar linear bias in [230, 231]. This is not exact, but it provides a good estimate of the nonlinear biases which for less dense samples can be large. Labeling with a tilde the biases in [230, 231], the rotation to our basis is

$$b_1 = \tilde{b}_1\ , \quad b_{\mathcal{G}_2} = \tilde{b}_{K^2} \quad b_2 = \tilde{b}_2 + \frac{4}{3}\tilde{b}_{K^2}\ , \tag{310}$$

for linear and quadratic bias and

$$b_{\Gamma_3} = \tilde{b}_{\text{td}}\ , \quad b_{\mathcal{G}_3} = -\tilde{b}_{K^3}\ , \quad b_{\mathcal{G}_2\delta} = \tilde{b}_{\delta K^2} + \frac{1}{2}\tilde{b}_{K^3} - \frac{8}{21}\tilde{b}_{\text{td}}\ , \quad b_3 = \tilde{b}_3 + 4\tilde{b}_{\delta K^2} + \frac{4}{3}\tilde{b}_{K^3}\ , \tag{311}$$

for cubic biases. Relaying on the measurements in [230, 231] we estimate

$$b_1 \approx 2.5\ , \quad b_2 \approx 0.2\ , \quad b_{\mathcal{G}_2} \approx -0.6\ , \quad b_{\Gamma_3} \approx 1.5\ , \quad b_{\mathcal{G}_2\delta} \approx 0.5\ , \quad b_{\mathcal{G}_3} \approx -1\ , \quad b_3 \approx -2\ . \tag{312}$$

The other parameters are the same as before. In particular, $c^2_{s,\text{fid}} = c_{1,\text{fid}} = c_{2,\text{fid}} = c_{3,\text{fid}} = 1$. We also use the same priors as in Equation (309).

We checked that a different choice of fiducial values of the nonlinear bias parameters in both cases does not significantly impact our results, beyond small 10% corrections. More generally, all our results should be correct only up to $10 - 20\%$, given several approximations that we use in the likelihood and the uncertainty in the fiducial parameters. This is acceptable, since we do not aim to have correct errors to better precision. We want to answer the question whether the error bars in FBI differ from P+B very significantly (by more than a factor of 2). Any mistake of the order of 10% is not very important for answering this question.

We begin by presenting the analysis performed on the mock data for a sample in [171] generated as described above, using the power spectrum and the bispectrum. The results are shown in Fig. 30 and Table II. There are three important points to make regarding the joint P+B analysis:

- First, the error bars are in rough agreement with the mode counting estimate in Equation (292). This is the confirmation that the main degeneracy $\alpha - b_1$ is broken by the bispectrum and that the remaining degeneracies of the amplitude α with the nonlinear biases b_2 and $b_{\mathcal{G}_2}$ are much milder (even though they degrade the errors by a factor of a few compared to the ideal case where all

	P+B	P+B+T
$k_{\rm max} = 0.1\ h/{\rm Mpc}$		
α	0.968±0.064	0.966±0.054
b_1	1.764±0.118	1.767±0.100
$k_{\rm max} = 0.12\ h/{\rm Mpc}$		
α	0.981±0.046	0.979±0.035
b_1	1.736±0.086	1.741±0.065

Table II: 1σ constraints and best fit parameter for P+B and P+B+T analyses using mock data.

biases are known).

- Second, the error bars from a joint P+B analysis are much closer to the FBI results of [171] and roughly a factor of 3 smaller than the SBI errors in the same paper. This is a very interesting result and we will discuss it in more detail in the next section.

- Finally, the error bars follow the estimate for the scaling with $k_{\rm max}$ given in Equation (294), providing another confirmation for our estimates. In particular, going from $k_{\rm max} = 0.1\ h/{\rm Mpc}$ to $k_{\rm max} = 0.12\ h/{\rm Mpc}$, we expect roughly a 30% improvement, in agreement with our analysis on the mock data.

Having obtained the error bars from the joint analysis of power spectrum and bispectrum, we proceed to investigate additional possible sources of information on top of this baseline analysis.

The three-level four-point function is the first higher order statistic that can be consistently constructed when the galaxy overdensity field is evolved up to third order in perturbation theory. In fact, the trispectrum completes the set of correlation functions that can be reliably predicted with the cubic model. In this section we present the joint P+B+T analysis and investigate the impact of the trispectrum on constraining the amplitude α.

The results of the joint P+B+T analysis are shown in Fig. 30 and Table II. The two main findings are:

- Adding the trispectrum to the P+B analysis reduces the error bars on the amplitude α by $\sim 20\%$ and $\sim 30\%$ at $k_{\rm max} = 0.1\,h/{\rm Mpc}$ and $k_{\rm max} = 0.12\ h/{\rm Mpc}$ respectively. This is an improvement in agreement with our SNR estimates. It indicates that there are no further major degeneracy breaking, which is expected given that there are several new free parameters in the trispectrum model. Nevertheless, even a $\sim 20\%$ reduction of error bars is interesting and it is important to see if the impact of the trispectrum is the same in a more realistic setup (e.g. including the full nonlinear covariance or doing the analysis in redshift space).

- The error bars on α at redshift $z = 0.5$ are 0.054 and 0.035 for $k_{\rm max} = 0.10\ h/{\rm Mpc}$ and $k_{\rm max} = 0.12\ h/{\rm Mpc}$ respectively. This is in a good agreement with the FBI analysis of [171] (given our uncertainty in the results of $10-20\%$). Such result is not obvious, and it is rather nontrivial. The two approaches are using different perturbation theory models and measure α in two different ways. Therefore, assuming the absence of accidental cancellations, we interpret this agreement as a strong indication that the full complexity of the field level analysis on large scales can be captured by the leading n-point functions computed in the simple Eulerian model, in agreement with expectations based on perturbation theory.

.

We now apply the same methodology to a sample which resembles the one analyzed in [212]. Note that the simulation volume is the same $V = (2\,{\rm Gpc}/h)^3$ but at slightly higher redshift $z = 1$. Using the exact same pipeline, we perform four different analyses. The results are shown in Fig.33, Fig.32 and Tab. IV.

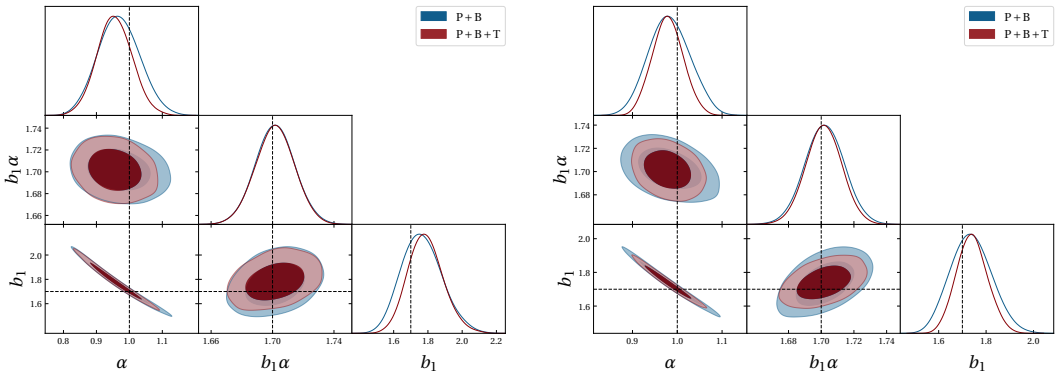

Figure 30: Comparison between P+B analysis and P+B+T using mock data for $k_{\rm max} = 0.1\ h/{\rm Mpc}$ (*Left*) and $k_{\rm max} = 0.12\ h/{\rm Mpc}$ (*Right*).

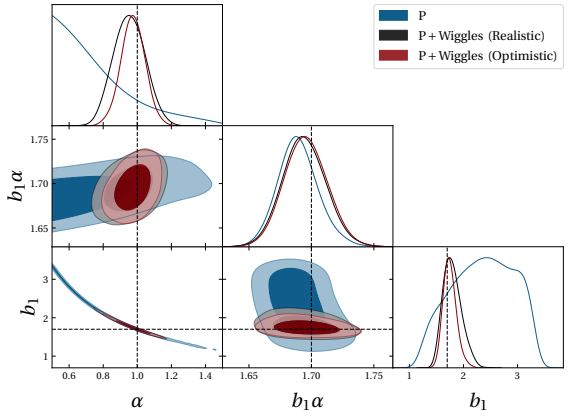

Figure 31: Comparison between constraints obtained using only the broadband fitted until $k_{\rm max} = 0.1\ h/{\rm Mpc}$ and when adding also wiggles (with $k_{\rm max,w} = 0.3\ h/{\rm Mpc}$).

- First, we perform a joint P+B analysis with $k_{\rm max,P} = 0.3\,h/{\rm Mpc}$ and $k_{\rm max,B} = 0.15\,h/{\rm Mpc}$. We do two versions of this analysis. In one of them we fix b_{Γ_3}, c_2, c_3 to their fiducial values. This setup is exactly the same as the EFT P+B "restricted" analysis in [212] (see Fig. 22), with the same choice of prior. We recover very similar error bar on α of 3%. In the other version we also vary b_{Γ_3}, c_2, c_3 within their usual priors. This corresponds to the baseline EFT P+B analysis in [212]. In this case we find a $\sim 4\%$ error, which is smaller than what is reported in [212] . However, given that our results are uncertain at the $O(20\%)$ level, we achieve a relatively good agreement with the EFT P+B analyses in [212]. This serves as a cross check that our pipeline works well and that our choice of fiducial parameters is reasonable.

- Second, we perform the joint P+B analysis with $k_{\rm max} = 0.12\,h/{\rm Mpc}$ using priors given in Equation (309). This scale cut corresponds to $k_{\rm max} = 0.1\,h/{\rm Mpc}$ of [212] where the cubic filter in k-space is used. Since we use the spherical momentum shells, the two choices have a comparable number of Fourier modes (for more details see [212]). The error on α in this case is 5.5%, approximately 30% larger than in the FBI "extended" analysis of [212]. This is similar to the results we found in other cases and indicates the broad agreement between the FBI and P+B analyses.

- Third, we add the trispectrum and perform the joint P+B+T analysis with the same cut $k_{\rm max} = 0.12h/{\rm Mpc}$. Similarly to what happens for other tracers, also here we find an improvement of roughly 20% compared to the P+B analysis. The joint P+B+T result is in rough agreement with the

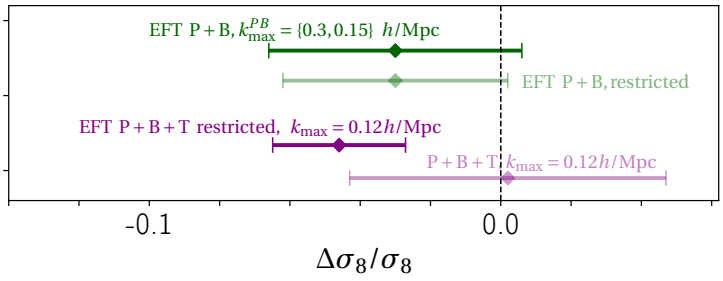

Figure 32: Error bars on the amplitude of the liner density field obtained in different analyses described in the main text. This figure is supposed to be compared with Fig. 22 in [212]. The central values are chosen to match the Fig. 22 in [212]. While our error bars are different at $O(20\%)$ level, the general trend for all example is the same.

FBI error of 4% obtained in [212].

- Finally, we run a P+B+T analysis at $k_{\rm max} = 0.12\ h/$Mpc but this time fixing all cubic biases and the bispectrum noise terms. In other words, the only free parameters remaining are the linear and quadratic biases, the dark matter counterterm and the power spectrum shot noise. We choose this setup to be as close as possible to the FBI "restricted" analysis in [212]. In this case we find a significantly smaller error on α of 2%. This matches nicely the FBI "restricted" analysis in [212].

The last example is particularly interesting. If we compare the "restricted" P+B+T and P+B analyses at the same $k_{\rm max} = 0.12\ h/$Mpc, we find that the addition of the trispectrum when cubic biases are fixed makes the constraints better by more than a factor of 2. Note that the impact of fixing cubic biases in the P+B analysis is minimal. This is a very significant improvement. Does this mean that our intuition for the relevance of the trispectrum is wrong? We think not. It is still true that the trispectrum has much lower SNR than the power spectrum and the bispectrum. However, since the trispectrum in this particular setup comes with no new free parameters by design, even this small signal is very useful to break the residual degeneracy between α and b_2 and $b_{\mathcal{G}_2}$ in the bispectrum. This degeneracy breaking is what leads to much smaller error bar on α. Clearly, this example is somewhat artificial and designed to maximize the impact of n-point functions beyond the bispectrum. For a real tracer where the cubic and higher order biases are not zero, such approach is inappropriate, as evident from the biased result of the FBI "restricted" analysis in [212]. Pushing this to the extreme, one can always design a setup or find a cosmological parameter such that the relevant information is only in the higher order correlation functions. For example, if we wanted to measure the primordial non-Gaussianity parameter $\tau_{\rm NL}$, the FBI would be much much better than the P+B analysis. However, these special situations are not what we have in mind when we discuss the comparison of the field-level inference and the standard analyses. In a generic setup, our arguments are still valid.

In order to summarize our results and visually compare them to analyses in [212], we produce Fig. 32 which is supposed to be compared to the Fig. 22 in [212]. We choose the central values of the intervals to match the results in [212] and we plot our forecast for the error bars. While all our results are uncertain at the $O(20\%)$ level, we find an overall good agreement with results of [212] and we recover all features and trends. These results are reassuring and provide additional evidence for the simple relation between the n-point functions and the field level inference on large scales.

In conclusion, all our results are in agreement with expectations from perturbation theory. For dark matter halos in real space, where the only cosmological parameter is the amplitude of the linear field α, most of the information comes from the bispectrum, except in some artificial examples where the higher order correlation functions can dominate the SNR. Even though the power spectrum has the highest signal-to-noise, due to the exact degeneracy between b_1 and α in linear theory, the amplitude can be measured from one-loop power spectrum only with a modest precision of roughly 20% in a box

of volume $V = 8\,(\mathrm{Gpc}/h)^3$ and with $k_{\rm max} = 0.12\ h/\mathrm{Mpc}$. On the other hand, adding the bispectrum improves these errors to approximately $5-6\%$ (the exact number slightly depends on the sample). Addition of the trispectrum further reduces the error to roughly $4-5\%$. These errors are in good agreement with the measurement of α using FBI [171, 212].

For simplicity, in this section we focused on only one of the two samples studied in [171]. However, we also checked that for the other sample at $z = 1$ and with even larger number density, we find very similar results as the FBI analysis and all our conclusions remain the same in that case as well.

G BAO wiggles

In addition to the baseline analysis of the correlation functions up to a given $k_{\rm max}$, we will also consider information on the amplitude of the linear density field in the broadening of the BAO wiggles. While one can easily extend the full analysis of the correlation functions beyond $k_{\rm max}0.1h/\mathrm{Mpc}$, we choose to focus on the question of how much information is in the wiggles only. This is important, since it gives an estimate of how much additional signal on the amplitude α one may ideally recover from the displacement field on smaller scales. In fact, one nontrivial way in which the amplitude α affects the BAO is through the broadening of the BAO peak. This nonlinear effect is very well understood and can be described to very high precision in perturbation theory [150–153]. First, one can decompose the linear power spectrum into the smooth and wiggly parts [227, 232–234];

$$P_{\rm lin}(k,z) = P^{\rm nw}_{\rm lin}(k,z) + P^{\rm w}_{\rm lin}(k,z)\ . \tag{313}$$

The bulk flows induced by the long wavelength fields affect the features such as BAO, leading to the damping of the oscillations. The full expression for the infrared-resummed galaxy power spectrum is schematically given by

$$P^{\rm IR}_g(k,z) = P^{\rm nw}_g(k,z) + e^{-\alpha^2\Sigma^2_\Lambda(z)k^2}(1+\alpha^2\Sigma^2_\Lambda(z)k^2)\alpha^2 b_1^2 P^{\rm w}_{\rm lin}(k,z) + e^{-\alpha^2\Sigma^2_\Lambda(z)k^2}P^{\rm w}_{g,1-\rm loop}(k,z)\ , \tag{314}$$

where $P^{\rm nw}_g$ is the standard EPT power spectrum up to one loop computed using $P^{\rm nw}_{\rm lin}$ and $P^{\rm w}_{g,1-\rm loop}$ is the wiggly part of the EPT one-loop power spectrum. The BAO damping factor is given by [150]

$$\Sigma^2_\Lambda(z) \equiv \frac{1}{6\pi^2}\int_0^\Lambda dq\, P_{\rm lin}(q,z)\left[1 - j_0(q\ell_{\rm BAO}) + 2j_2(q\ell_{\rm BAO})\right]\ , \tag{315}$$

where the BAO scale is $\ell_{\rm BAO} \approx 110\,h/\mathrm{Mpc}$, $j_n(x)$ are spherical Bessel functions of order n, and Λ is the scale up to which the long-wavelength displacements are taken into account. The typical value used in practice is $\Lambda = 0.2\,h/\mathrm{Mpc}$, but the result is almost insensitive to this choice [150, 227, 235]. Given the numerical values of $\Sigma^2_\Lambda(z)$ for relevant redshifts, the damping is important only for high k. Note that the formula for the wiggly part of the galaxy power spectrum is *non-perturbative* and almost perfectly describes the BAO. Therefore, unlike the broadband, it can be safely used at all scales.

It is clear from Equation (314) how the parameter α can be measured from the BAO wiggles. Apart from the small loop corrections controlled by the perturbation theory parameter $\Delta^2(k)$, α multiplies the large parameter Σ^2_Λ which controls the broadening of the BAO peak. In other words, the amplitude of the linear density field can be simply estimated from how much the nonlinear evolution damps the BAO wiggles. This is easy to intuitively understand in real space. On large scales, one can approximate the two-point correlation function using only the linear theory. From the smooth part one can measure the combination $b_1\alpha$. Then, focusing on the BAO peak only, the broadening is controlled by $\alpha^2\Sigma^2_\Lambda$ which has only one free parameter. This is the essence of the statement that large displacements can be used to measure α without strong sensitivity on nonlinear galaxy bias if a feature is present in the initial conditions. Crucially, this is true in Fourier space only for high-k part of the BAO wiggles,

beyond $k_{\rm max} \approx 0.2\ h/{\rm Mpc}$.

It is possible to test how much information on α comes from the BAO wiggles using the formulas above. While one can use Equation (314) directly, just to keep the discussion as simple as possible, we will use in our forecast the following simplified model for the IR-resummed wiggly part of the galaxy power spectrum

$$P_g^{\rm w}(k,z) = e^{-\alpha^2\Sigma_\Lambda^2(z)k^2}\alpha^2 b_1^2 P_{\rm lin}^{\rm w}(k,z) + c_{\rm w}^2\alpha^2 b_1^2 k^2 P_{\rm lin}^{\rm w}(k,z)\ . \tag{316}$$

This equation phenomenologically captures all features of the exact result and in a Fisher forecast it is simpler to implement and interpret. We introduce a single parameter $c_{\rm w}^2$ that encompasses all nonlinear bias and possible other nonlinear corrections and it is of order $O(1)$. The wiggly part of the galaxy power spectrum at high k can then be analyzed separately from the broadband (see for example [236–239]).

We will consider two different setups.

- Optimistic: In this setup we set $c_{\rm w} = 0$ and $\Lambda \to \infty$. This is the best case scenario, since we neglect all nonlinear corrections with bias parameters and we replace Σ_Λ^2 with Σ_∞^2, to approximately take into account the total broadening from the displacement *and* loop corrections. This makes the effect as large as possible and the constraints as good as possible.

- Realistic: In this setup we set $\Lambda = 0.2\ h/{\rm Mpc}$ and leave $c_{\rm w}^2$ as a free parameter with a Gaussian prior of width equal to 1. In this way the information on α comes only from the part of the broadening induced with large displacements and it is to some extent degenerate with other nonlinearities.

The likelihood in these cases has the same form as for the power spectrum, except that the signal and the model contain the wiggly part of the galaxy power spectrum only

$$\ln\mathcal{L}_{\rm w} = -\frac{1}{2}\sum_{k=k_{\rm max}}^{\bar{k}}\frac{(P_g^{\rm w}(k,\theta) - \hat{P}_g^{\rm w}(k))^2}{{\rm Cov}_P(k)}\ . \tag{317}$$

In our implementation we choose $\bar{k} = 0.3\ h/{\rm Mpc}$ since the wiggles are essentially completely damped beyond this scale. Note that the covariance matrix is the same as for the power spectrum. Clearly, the SNR on the amplitude α is suppressed by the amplitude of the wiggly part of the power spectrum which is smaller than $O(10\%)$.

In Fig. 31 and III we show the constraints on b_1 and α obtained for the power spectrum up to $k_{\rm max} = 0.1\ h/{\rm Mpc}$ with and without adding the information from the wiggles with $k_{\rm max} < k \leq \bar{k} = 0.3\ h/{\rm Mpc}$. We present the constraints for the realistic and optimistic scenario. The main findings are:

- The information on the amplitude α in the wiggles is significantly higher than in the power spectrum at $k < k_{\rm max}$. Even though the wiggles have small amplitude ($\sim 10\%$ of the power spectrum), given the robustness of the BAO broadening, the constraints from the large displacements are stronger than the one-loop power spectrum.

- Even though adding the wiggles improves the results, they are still not as good as adding the bispectrum on large scales. In principle, it is not obvious that this had to be the case. We conclude that in a ΛCDM-like cosmology, in order to get optimal constraints, one has to include the higher-order n-point functions.

- The results do not change dramatically going from optimistic to realistic scenario. This implies that the constraints from the BAO wiggles are rather robust. Note that this information is automatically included in the full-shape analyses of the power spectrum that use $k_{\rm max} \sim 0.3\ h/{\rm Mpc}$.

	P	P+Wiggles (realistic)	P+Wiggles (optimistic)
α	0.743±0.185	0.95±0.099	0.975±0.072
b_1	2.402±0.535	1.801±0.191	1.748±0.124

Table III: 1σ constraints and best fit parameter for broadband analysis of the power spectrum and when including also the wiggles in the realistic and in the optimistic scenario.

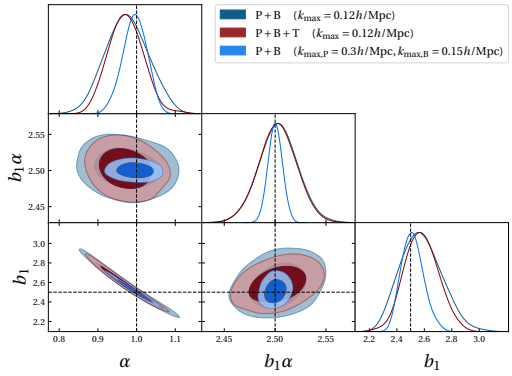

Figure 33: P+B and P+B+T analyses of the sample of halos similar to [212]. P+B results for $k_{\rm max,P} = 0.3h/{\rm Mpc}$, $k_{\rm max,B} = 0.15h/{\rm Mpc}$ fixing b_{Γ_3}, c_2, c_3 to their fiducial values are also shown.

H *Comparison with SBI*

The results presented in the previous section indicate that the conventional analyses based on correlation functions and the FBI analysis lead to the same results in the perturbative regime, as expected based on theoretical arguments [213, 214]. This seems to be in contradiction with the findings in [171] where the FBI leads to up to 5 times better error bars than the SBI P+B analysis. The new puzzle is why the SBI and the standard P+B analyses differ so much. In this section we try to provide some possible explanations, even though the conclusive arguments are still lacking and more investigation is needed to fully clarify this issue.

The key to understanding the different results is to understand the differences in the settings of the SBI analysis of [171] and the settings in this section. There are three major differences. One is the theoretical model used. SBI analysis of [171] is based on LPT, while we use the Eulerian theory. As we argued in detail in this work and as has been checked in the literature before, the two models are practically indistinguishable on the relevant scales and we do not think that this can be the source of the disagreement. The second major difference is the likelihood used in the analyses. SBI does not assume any likelihood (even though it has been shown that a Gaussian likelihood is an excellent approximation [182]) and uses a large number of realizations to estimate the relevant scatter in the bins for the data vector. On the other hand, we use a simple Gaussian likelihood with the Gaussian covariance, neglecting the cross-covariance between P, B and T. We have also justified this choice in the previous sections and there is numerous evidence in the literature that this is a good approximation. Finally, the third major difference is the choice of bias parameters in the analyses and the treatment of priors on

	P+B	P+B+T	P+B ($k_{\rm max,P} = 0.3h/{\rm Mpc}$, $k_{\rm max,B} = 0.15h/{\rm Mpc}$)
α	0.975±0.055	0.973±0.045	0.995±0.031
b_1	2.575±0.151	2.578±0.125	2.514±0.0810

Table IV: 1σ constraints and best fit parameters for P+B and P+B+T with $k_{\rm max} = 0.12\,h/{\rm Mpc}$ (first and second columns) and P+B with $k_{\rm max,P} = 0.3\,h/{\rm Mpc}$ and $k_{\rm max,B} = 0.15\,h/{\rm Mpc}$ (third column).

these nuisance parameters. We believe that this can be the main source of disagreement.

In this work we follow the usual logic used in perturbation theory and analyses based on the n-point functions. First, given the statistical error bars of the data, we estimate to what perturbative order we have to compute observables in order to be unbiased. In our examples, such estimate indicates that the one-loop power spectrum and the tree-level bispectrum are sufficient. Once the order in PT is fixed, the bias parameters that have to be used in the analysis are also determined. It is important to emphasize that here we do not truncate the bias expansion, we are only truncating the loop expansion of the correlation functions, given the total SNR that we expect for a given data set. For example, in the case where we compute power spectrum up to one-loop and the bispectrum at tree level, even though we use the full cubic bias model, only one of the cubic biases appears explicitly in the one-loop power spectrum (b_{Γ_3}). The other cubic biases either give zero contribution at this order in PT or just renormalize b_1. The complexity of the nonlinear field at this order in PT and for the simplest observables such as P and B is just not sufficient for other bias parameters to appear explicitly. Obviously, cubic biases would appear if we had gone to higher order in the loop expansion or higher n-point functions. However, since we estimate that they are not necessary to fit the data, we can safely drop these refinements of the basic perturbative prediction. To conclude, even though our model is based on the full bias expansion up to the cubic order, 3 out of 4 cubic bias parameters do not appear in the analysis. This is a consequence of bias renormalization and the loop expansion and not restricting the number of bias parameters that are used in the analysis.

On the other hand, the practical implementation of SBI in [171] uses the nonlinear field computed at cubic order with all relevant bias parameters to predict P and B. This is similar in spirit to the field-level inference, where one can only truncate the forward model to some order in the bias expansion, without the possibility to explicitly restrict the likelihood to certain correlation functions or number of loops. Therefore, the estimated power spectrum and bispectrum contain higher order contributions beyond what we use in our model. For example, the power spectrum has a part of the 2-loop terms, such as various P_{33} contributions. The bispectrum has a part of the one-loop contributions (diagrams usually labeled as B_{123} and B_{222}) and even some two-loop contributions such as (B_{233}). It is interesting to explore whether this difference leads to a large difference in the error bars in the two P+B analyses.

Naively, one may think that this difference is irrelevant, since the higher order terms are small in PT. However, this is not obvious. While the variance of the dark matter field is small for $k_{\rm max}$ used in the analysis, the fact that there are many contributions from the bias expansion means that there are combinations of bias parameters that might produce a large effect. With all bias parameters uncorrelated and of order one, such combinations are naturally explored in the MCMC chains. Marginalizing over such parameters can then make the error bars on α significantly higher than the standard expectation. We do not have a clear evidence that this is the main origin of the disagreement, but it is certainly an effect to be carefully taken into account in any comparison.

In order to illustrate this point better, let us imagine that we keep going to higher and higher orders in PT hoping to improve the precision of our inference. If we keep $k_{\rm max}$ fixed, this means that we fit the same data with a theory that has a rapidly increasing number of free parameters. Without careful choice of the priors, even though the individual loop diagrams may be small, summing a large number of such terms would inevitably lead to large effects. In such case the *total* loop contribution stops being small and the perturbativity in this sense is lost. The details of when this happens for a given $k_{\rm max}$ and the given observable depend on the sample and the priors, but one always has to check if this is the case. The hints of this behavior were already noticed in the analyses of the one-loop bispectrum [132, 240, 241]. On the other hand, for the consistency of PT, adding higher order terms should at worst leave the constraints unchanged, it should never make them worse.

What are the usual ways to mitigate this problem? One possible procedure is to add informative and physically motivated priors on bias parameters which makes them correlated. These correlations can be estimated from some external analysis (see for example [242–246] for HOD informed priors

obtained from a simulation-based approach and from analytical predictions respectively). Another possible solution is to use the so-called perturbativity priors [132], which we already mentioned in the introduction. These theoretically motivated priors do not rely on any galaxy evolution physics and are imposed in order to control the size of the loop corrections predicted by PT (see [132] for more details). It is important to emphasize that these two procedures lead to similar results, indicating that the real tracers indeed never produce large fluctuations due to bias nonlinearities. Perturbativity priors are also one implicit assumption that we always make for perturbation theory to make sense. Note that requiring that fluctuations in the galaxy density field are small on large scales still allows that all biases are of order 1, but imposes correlations among them, similar to the ones found in HOD or hydrodynamical simulations.

If one accepts the perturbativity prior, what are the consequences for the SBI analysis? In principle, one would have to restrict the sampling only to those values of biases that do not produce large loop corrections. In this case, the expectation is that the cubic terms in the bias expansion would be much less relevant and the SBI will be similar to the standard analysis of the tree-level bispectrum presented in this work. This intuition provides a simple way to test the relevance of cubic terms. For example, one can set all cubic parameters to zero in the SBI analysis and try to recover the standard results. The other easy check is to plot the one-loop bispectrum amplitudes as a function of scale for various shapes (similar to Fig. 27) for all samples in the chains in the SBI and see if the envelope they create coincides with the estimate of the theoretical error.

If we assume that the wide priors are the origin of the disagreement between the SBI and the standard analysis, one may wonder why the FBI still gives good results with presumably large fluctuations in the galaxy density field. The answer is that the FBI is always optimal if the model is close enough to the data. For the combination of biases that do not respect perturbativity the model would contain a large fluctuation in the galaxy density field with the particular template given by PT. FBI compares such model with the whole field and assigns a small likelihood to such set of bias parameters. Therefore, there is no contradiction in having small errors in the FBI approach and large errors from P+B if the nuisance parameters are degenerate enough in the P+B combination.

In conclusion, one possible origin of disagreement can be that including higher-order loop contributions on very large scales where they are supposed to be negligible can lead to artificial increase in the variance of the inferred cosmological parameters. In order to avoid this, one can either do the lower order calculation as we did in this work, or impose some perturbativity priors in the inference. If one does either of these, the results obtained using n-point functions at fixed loop order and the SBI analysis should lead to similar constraints. Luckily, this can be relatively easily tested and we leave this for future work.

VII A different summary statistics: the marked power spectrum

A *Theory model*

In the previous section we have computed the bispectrum and trispectrum for a biased tracer in real space. This task seemed easy as we were working in an ideal setup with a simulation box in real space. However, in real world scenarios, when we analyze galaxy statistics on the light-cone the situation dramatically changes. The inclusion of observational effects such as RSD, the window function, and fiber collisions gets more and more complicated as we go to higher order statistics beyond the power spectrum.

For this reason, there has been increasing interest in the study of alternative summary statistics that can efficiently compress information even in the strongly non-linear regime while remaining computationally manageable. A partial list of these summaries includes methods such as the Wavelet Scattering Transform [247–252], k-nearest neighbors [253, 254], skew-spectra [255, 256], clipping [257] and Minkowski functionals [258, 259]. So far, these summaries have been either explored through Fisher analysis or used as inputs for likelihood-free inference.

In this section, we focus on another promising statistic that has received significant attention in the literature: the marked power spectrum. This is defined as the two-point correlator of a weighted version of the density field, where the weight ("mark") can represent galaxy properties, halo merger history or - as considered in this work – the average density of the surrounding environment.

The marked power spectrum was first introduced in [260]. In the context of Cosmology, it was recently pointed out in [261] that a density-weighted mark could be a powerful probe of Modified Gravity. Its effectiveness at constraining many cosmological parameters – and notably neutrino masses – was shown by the authors of [262, 263] in a simulation-based study. Its sensitivity to primordial non-Gaussianity was instead recently pointed out in [264].

Previous works [261, 263, 265–267] emphasized two main factors that explain why the marked power spectrum is so efficient at extracting cosmological information, compared to the standard power spectrum. The former is that, by correlating the field with the environment, the marked power spectrum is effectively incorporating information from higher-order correlation functions. The latter is that the mark can be chosen to up-weight low-density regions. These carry significant information but, at the same time, give very little contribution to the conventional power spectrum.

The primary aim of this chapter is to shed further insight into these issues by significantly extending previous analytical studies of the marked power spectrum in several directions. First, we extend the EFTofLSS-based formalism introduced by [265], developing for the first time a theoretical framework that incorporates primordial non-Gaussianity (PNG) into the perturbative description of the marked power spectrum. Additionally, we conduct a Fisher matrix analysis to give an analytical quantitative assessment of the marked power spectrum's sensitivity to various parameters, including PNG and neutrino masses; in this assessment, we consider different choices of the density-weight parameters of the mark in order to verify the impact of under/over-weighting low-density regions. Finally, we make a detailed comparison between the constraining power of the marked power spectrum and that of a combination of the conventional power spectrum and bispectrum ($P + B$). In previous works, the marked 2-point statistic was typically compared to the standard power spectrum, but $P + B$ looks like a more natural benchmark, considering that the marked field inherently includes higher-order information, as noted earlier. This will be made more transparent in our analytical model, which displays new terms in the one-loop power spectrum of the marked fields that are very similar to the tree-level bispectrum and to the convolution of linear power spectra.

In the current study, we focus specifically on the form of the mark introduced in [261], which is

widely used in cosmological studies on this topic. Recently, [268] highlighted that this ansatz can be optimized in a parameter-dependent manner to enhance estimation efficiency. We leave this optimization — especially in the context of PNG analysis, which is a primary focus of our study — for future work. Additionally, we restrict our analysis to biased tracers in real space, with an extension to redshift space also planned for a future work.

The standard statistic employed in Large Scale Structure analysis is the power spectrum of the galaxy number density perturbation [22]

$$n_g(\boldsymbol{x}) \equiv \bar{n}_g(1 + \delta_g(\boldsymbol{x}))\,, \tag{318}$$

where $\bar{n}_g \equiv \langle n_g(\boldsymbol{x})\rangle$ is the mean galaxy density and $\delta_g(\boldsymbol{x})$ is the usual overdensity field. In this work, we are interested in studying an alternative summary statistic to extract more information from the galaxy clustering, the *Marked Power Spectrum* (see refs. [212, 264] for other examples of summary statistics), in which the galaxy number density of Equation 318 is weighted by a (space-dependent) mark

$$\rho_{\rm M}(\boldsymbol{x}) = m(\boldsymbol{x}) n_g(\boldsymbol{x}) = m(\boldsymbol{x})\bar{n}[1 + \delta_g(\boldsymbol{x})]\,. \tag{319}$$

The mark function usually adopted in literature is [262, 265, 266, 269, 270]

$$m(\boldsymbol{x}) = \left(\frac{1+\delta_s}{1+\delta_s+\delta_{g,R}(\boldsymbol{x})}\right)^p\,, \tag{320}$$

which is a function of the smoothed galaxy overdensity field, $\delta_{g,R}(\boldsymbol{x})$[23], and R, δ_s, and p are hyper-parameters of the model that can be used to enhance underdense ($p > 0$) and overdense ($p < 0$) regions. Marking the overdensity field with this non-linear function of the smoothed field will make the marked density field more non-Gaussian, allowing us to access information on the large-scale structure distribution that is usually included in higher-order correlation functions.

By defining the mean mark $\bar{m}$ as

$$\langle \rho_{\rm M}(\boldsymbol{x})\rangle = \langle m(\boldsymbol{x}) n_g(\boldsymbol{x})\rangle \equiv \bar{m}\bar{n} \tag{322}$$

we can consider the marked overdensity field

$$\delta_M(\boldsymbol{x}) \equiv \frac{\rho_M(\boldsymbol{x}) - \langle\rho_M\rangle}{\langle\rho_M\rangle} = \frac{m(\boldsymbol{x})}{\bar{m}}[1 + \delta_g(\boldsymbol{x})] - 1\,, \tag{323}$$

and expand it perturbatively in the non-linear fields δ_g and $\delta_{g,R}$. Expanding the mark, we first obtain

$$\begin{aligned} m(\boldsymbol{x}) &= \left(\frac{1+\delta_s}{1+\delta_s+\delta_{g,R}}\right)^p = \sum_{n=0}^{\infty}(-1)^n \frac{p(p+1)\dots(p+n-1)}{n!}\frac{{\delta_{g,R}}^n}{(1+\delta_s)^n} \\ &\equiv \sum_{n=0}^{\infty}(-1)^n C_n {\delta_{g,R}}^n(\boldsymbol{x})\,, \end{aligned} \tag{324}$$

with

$$C_n \equiv \frac{p(p+1)\dots(p+n-1)}{n!}\frac{1}{(1+\delta_s)^n}\,. \tag{325}$$

[22] Throughout this work, we will always refer to galaxies, but the present approach refers to any biased tracer of the dark matter density field.

[23] We will use a Gaussian smoothing function (in Fourier space) which reads

$$W_R(k) = e^{-\frac{k^2R^2}{2}}\,, \quad \delta_{g,R}(\boldsymbol{k}) = W_R(k)\delta_g(\boldsymbol{k})\,. \tag{321}$$

For biased tracers, such as galaxies, we need to be careful about the convergence of the series in (324). Taking a large enough R, the time evolution of the smoothed field is $\delta_{g,R}(z) \simeq b_1(z)\, D(z)\, \delta_0$, with b_1 and D being the linear bias and the growth factor; the convergence of the series is guaranteed simply by $|\delta_{g,R}(\boldsymbol{x})/(1+\delta_s)| < 1$, and, noting that the fluctuation of the smoothed field is $\sigma^2_{RR}(z) \equiv \langle \delta^2_{m,R}(\boldsymbol{x})\rangle$, we expect convergence if

$$\sigma_{RR}(0) < \frac{1+\delta_s}{b_1(z)D(z)}\,. \tag{326}$$

We can proceed by expanding perturbatively as

$$\delta_g(\boldsymbol{x}) = \sum_{n=0}^{\infty} \delta_g^{(n)}(\boldsymbol{x})\,, \qquad \delta_{g,R}(\boldsymbol{x}) = \sum_{n=0}^{\infty} {\delta_{g,R}}^{(n)}(\boldsymbol{x})\,, \tag{327}$$

where, in Fourier space, we have ${\delta_{g,R}}^{(n)}(\boldsymbol{k}) = W_R(k)\delta_g^{(n)}(\boldsymbol{k})$. Now, inserting (327) inside the marked density field definition, Equation (323), we obtain

$$\begin{aligned}
\bar{m}(1+\delta_{\rm M}(\boldsymbol{x})) &= m(\boldsymbol{x})(1+\delta_g(\boldsymbol{x})) \\
&= \sum_{n=0}^{\infty}(-1)^n C_n \left(\sum_{l=0}^{\infty} {\delta_{g,R}}^{(l)}(\boldsymbol{x})\right)^n \left[1+\sum_{j=1}^{\infty}\delta_g^{(j)}(\boldsymbol{x})\right] \\
&\equiv 1+\sum_{n=0}^{\infty} {\delta_{\rm M}}^{(n)}(\boldsymbol{x})\,,
\end{aligned} \tag{328}$$

from which we obtain

$$ {\delta_{\rm M}}^{(1)}(\boldsymbol{x}) = C_0\delta_g^{(1)}(\boldsymbol{x}) - C_1{\delta_{g,R}}^{(1)}(\boldsymbol{x})\,, \tag{329}$$

$$ {\delta_{\rm M}}^{(2)}(\boldsymbol{x}) = C_0\delta_g^{(2)}(\boldsymbol{x}) - C_1{\delta_{g,R}}^{(1)}\delta_g^{(1)}(\boldsymbol{x}) - C_1{\delta_{g,R}}^{(2)}(\boldsymbol{x}) + C_2{\delta_{g,R}}^{(1)}{\delta_{g,R}}^{(1)}(\boldsymbol{x})\,, \tag{330}$$

$$\begin{aligned}
{\delta_{\rm M}}^{(3)}(\boldsymbol{x}) &= C_0\delta_g^{(3)}(\boldsymbol{x}) - C_1{\delta_{g,R}}^{(1)}\delta_g^{(2)}(\boldsymbol{x}) - C_1{\delta_{g,R}}^{(2)}\delta_g^{(1)}(\boldsymbol{x}) - C_1{\delta_{g,R}}^{(3)}(\boldsymbol{x}) \\
&\quad C_2{\delta_{g,R}}^{(1)}{\delta_{g,R}}^{(1)}\delta_g^{(1)}(\boldsymbol{x}) + 2C_2{\delta_{g,R}}^{(1)}{\delta_{g,R}}^{(2)}(\boldsymbol{x}) - C_3{\delta_{g,R}}^{(1)}{\delta_{g,R}}^{(1)}{\delta_{g,R}}^{(1)}(\boldsymbol{x})\,.
\end{aligned} \tag{331}$$

To proceed, we expand the various fields involved into powers of the linear matter perturbation field $\delta^{(1)}(\boldsymbol{k})$. For biased tracers in real space, we have

$$\delta_g^{(n)}(\boldsymbol{k},z) = \mathcal{I}_{\boldsymbol{k},\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n} K_n(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n;z)\delta^{(1)}(\boldsymbol{q}_1)\ldots\delta^{(1)}(\boldsymbol{q}_n)\,, \tag{332}$$

where we have defined

$$\mathcal{I}_{\boldsymbol{k},\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n} \equiv \int \frac{d^3\boldsymbol{q}_1}{(2\pi)^3}\cdots\int \frac{d^3\boldsymbol{q}_n}{(2\pi)^3}(2\pi)^3\delta_D(\boldsymbol{k}-\boldsymbol{q}_{1\ldots n})\,, \tag{333}$$

and K_n are the well known real space kernels for a generic biased tracer. Analogously, we can define new kernels for the marked field as

$$ {\delta_{\rm M}}^{(n)}(\boldsymbol{k}) = \mathcal{I}_{\boldsymbol{k},\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n} H_n(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n;z)\delta^{(1)}(\boldsymbol{q}_1)\ldots\delta^{(1)}(\boldsymbol{q}_n)\,, \tag{334}$$

with

$$\begin{aligned} H_1(\boldsymbol{k}) &= C_{\delta_{\rm M}}(k)K_1(\boldsymbol{k})\,, \\ H_2(\boldsymbol{k}_1,\boldsymbol{k}_2) &= C_{\delta_{\rm M}}(k)K_2(\boldsymbol{k}_1,\boldsymbol{k}_2) + C_{\delta_{\rm M}{}^2}(k_1,k_2)K_1(\boldsymbol{k}_1)K_1(\boldsymbol{k}_2)\,, \\ H_3(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) &= C_{\delta_{\rm M}}(k)K_3(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) + 2C_{\delta_{\rm M}{}^2}(k_1,k_{23})K_1(\boldsymbol{k}_1)K_2(\boldsymbol{k}_2,\boldsymbol{k}_3) \\ &\quad + C_{\delta_{\rm M}{}^3}(k_1,k_2,k_3)K_1(\boldsymbol{k}_1)K_1(\boldsymbol{k}_2)K_1(\boldsymbol{k}_3)\,, \end{aligned} \tag{335}$$

where the H_3 still needs to be symmetrized of the three momenta. We have defined the functions

$$C_{\delta_{\rm M}}(k) = C_0 - C_1 W_R(k)\,, \tag{336}$$

$$C_{\delta_{\rm M}{}^2}(k) = C_2 W_R(k_1)W_R(k_2) - \frac{C_1}{2}[W_R(k_1) + W_R(k_2)]\,, \tag{337}$$

$$\begin{aligned} C_{\delta_{\rm M}{}^3}(k_1,k_2,k_3) = &- C_3 W_R(k_1)W_R(k_2)W_R(k_3) \\ &+ \frac{C_2}{3}[W_R(k_2)W_R(k_3) + W_R(k_1)W_R(k_3) + W_R(k_1)W_R(k_2)]\,. \end{aligned} \tag{338}$$

Notice that these functions implicitly contain information about the choice of the adopted mark through the coefficients C_i. Finally, from this expansion, we can compute the power spectrum of the marked density field up to third perturbative order, i.e. the one loop marked power spectrum. We have

$$\bar{m}^2 M(\boldsymbol{k}) = \bar{m}^2 |\delta_{\rm M}(\boldsymbol{k})|^2 \equiv M_{11}(\boldsymbol{k}) + 2M_{13}(\boldsymbol{k}) + M_{22}(\boldsymbol{k})\,, \tag{339}$$

with

$$M_{11}(\boldsymbol{k}) = H_1^2(\boldsymbol{k})P_L(k)\,, \tag{340}$$

$$M_{13}(\boldsymbol{k}) = 3H_1(\boldsymbol{k})P_L(k)\int \frac{d^3\boldsymbol{q}}{(2\pi)^3} H_3(\boldsymbol{k},\boldsymbol{q},-\boldsymbol{q})P_L(q)\,, \tag{341}$$

$$M_{22}(\boldsymbol{k}) = 2\int \frac{d^3\boldsymbol{q}}{(2\pi)^3} |H_2(\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q})|^2 P_L(q)P_L(|\boldsymbol{k}-\boldsymbol{q}|)\,, \tag{342}$$

where P_L is the linear power spectrum. By writing explicitly the expression for the 22 term in Equation 342 we obtain

$$\begin{aligned} M_{22}(k) = &\, 2C^2_{\delta_{\rm M}}(k)\int_{\boldsymbol{q}} K_2^2(\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q})P_{\rm lin}(|\boldsymbol{k}-\boldsymbol{q}|)P_{\rm lin}(q) \\ &+ 2\int_{\boldsymbol{q}} C^2_{\delta_{\rm M}{}^2}(|\boldsymbol{k}-\boldsymbol{q}|,q)K_1^2(\boldsymbol{k}-\boldsymbol{q})K_1^2(\boldsymbol{q})P_{\rm lin}(|\boldsymbol{k}-\boldsymbol{q}|)P_{\rm lin}(q) \qquad (343) \\ &+ 4C_{\delta_{\rm M}}(k)\int_{\boldsymbol{q}} C_{\delta_{\rm M}{}^2}(|\boldsymbol{k}-\boldsymbol{q}|,q)K_1(\boldsymbol{k}-\boldsymbol{q})K_1(\boldsymbol{q})K_2(\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q})P_{\rm lin}(|\boldsymbol{k}-\boldsymbol{q}|)P_{\rm lin}(q)\,. \qquad (344) \end{aligned}$$

The first line is a simple rescaling of the usual P_{22} term of the galaxy one-loop power spectrum. This piece should contain, at most, the same amount of information of the P_{22} term, modulo the dependence on the scale of the $C_{\delta_{\rm M}}$ function. The additional terms in the second and third line of Equation 344, which are generated by the presence of the smoothed field in the mark function, are convolutions of bispectrum- and (disconnected) trispectrum-like terms with some scale-dependent kernels given by the marking procedure. These terms arise because the marked field is by definition more non-Gaussian, and already at the two-point function level, it contains information from higher order correlation functions of the unmarked field.

As pointed out in [265, 266], the marked power spectrum requires more caution compared to the standard case as the inclusion of renormalized operators in the bias expansion leads to the addition of a

zero-lag contribution, which, in Fourier space, manifests as

$$\delta_g(\boldsymbol{k}) \to \delta_g(\boldsymbol{k}) - \frac{b_2}{2}\sigma^2(2\pi)^3\delta_D(\boldsymbol{k})\,, \tag{345}$$

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

$$\delta_{\rm M}{}^{(2)}(\boldsymbol{k}) \to \delta_M^{(2)} - \frac{b_2}{2}\sigma^2 C_{\delta_{\rm M}}(k)(2\pi)^3\delta_D(\boldsymbol{k})\,, \tag{346}$$

$$\delta_{\rm M}{}^{(3)}(\boldsymbol{k}) - b_2\sigma^2 C_{\delta_{\rm M}{}^2}(k,0)\delta^{(1)}(\boldsymbol{k})\,. \tag{347}$$

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

In this work, we will compare the information content of a standard analysis that involves galaxy power spectrum and bispectrum against an analysis performed by adopting only two-point correlation functions, namely the one-loop power spectrum and marked power spectrum. The model for the galaxy bispectrum in real space is given by

$$B(k_1,k_2,k_3) = 2K_1(\boldsymbol{k}_1)K_1(\boldsymbol{k}_2)K_2(\boldsymbol{k}_1,\boldsymbol{k}_2)P_L(k_1)P_L(k_2) + 2 \text{ perms.}\,, \tag{348}$$

where we stopped at the tree level. As for the one-loop contribution, one would need to go to the fourth perturbative order. The tree-level approximation has been shown to work properly up to $k^B_{max} \simeq 0.08/0.10\,h/\text{Mpc}$ [183, 271], while the one-loop is needed to extend the maximum k and get unbiased results [240, 241, 272].

The perturbative results presented in the previous section need to be corrected in order to achieve better agreement with smaller, more non-linear scales. The Effective Field Theory of Large Scale Structure (EFTofLSS), originally presented in [139, 140, 143], proposes itself as a completion treatment of standard perturbation theory, addressing its main issues: UV divergence and the treatment of imperfections in the cosmological fluid. For the standard one-loop power spectrum in real space, this is achieved with the introduction of a third-order counterterm, which, in Fourier space, reads $\delta^{(ct)}(\boldsymbol{k}) = -c_s^2k^2\delta^{(1)}(\boldsymbol{k})$. This term accounts for the backreaction of small-scale physics on long-wavelength modes. The speed of sound c_s^2 is a free parameter of the theory and will be fitted with observational data. For the marked power spectrum, it was shown that this term is enough to capture the UV divergences in real space; all other terms are manifestly convergent for hard loop momenta $p \gg k$ due to the presence of smoothing windows that depend on the physical scale R, used in the definition of the mark. Ref. [265, 266] showed that the marked theory does not require additional counterterms, and the relevant counterterm in $M(\boldsymbol{k})$ is

$$M_{ct}(\boldsymbol{k}) = -2c_s^2k^2C^2_{\delta_{\rm M}}(k)P_L(k)\,, \tag{349}$$

with c_s^2 being redshift dependent. In the case of biased tracers, additional UV divergences arise

$$M_{13}(\boldsymbol{k}) \supset 2C_{\delta_{\rm M}}(k)C_{\delta_{\rm M}{}^2}(k,0)K_1^2(\boldsymbol{k})P_L(k)\int_{\boldsymbol{q}}\frac{b_2}{2}P_L(q)\,, \tag{350}$$

$$M_{22}(\boldsymbol{k}) \supset \frac{b_2^2}{2}C^2_{\delta_{\rm M}}(k)\int_{\boldsymbol{q}}P_L(q)P_L(|\boldsymbol{k}-\boldsymbol{q}|) = \frac{b_2^2}{2}C^2_{\delta_{\rm M}}(k)\int_{\boldsymbol{q}}[P_L(q)P_L(|\boldsymbol{k}-\boldsymbol{q}|) - P_L^2(q)] \tag{351}$$

$$+ \frac{b_2^2}{2}C^2_{\delta_{\rm M}}(k)\int_{\boldsymbol{q}}P_L^2(q)\,. \tag{352}$$

The M_{13} contribution vanishes when including the proper renormalized bias operator as in Equation 347, while the second term contains a divergent part in the last line that can be completely absorbed when

considering the shot noise term. Indeed, the stochastic part of the marked power spectrum

$$M_{\rm stoch}(k) = C^2_{\delta_{\rm M}}(k) P_{\rm shot}\,, \tag{353}$$

for constant $P_{\rm shot}$ reabsorbs the UV divergence present in the M_{22} term. For this reason, following [266], we will assume this shape for the noise contribution henceforth.

For the infrared resummation, that accounts for the correct treatment of long-wavelength (IR) displacements on the baryon acoustic peak [150, 152], we will follow the approximation of [265, 266]: for the terms in the marked spectrum involving $C^2_{\delta_{\rm M}}(k)P_{NL}(k)$ we replace P_{NL} with its IR-resummed form, while for the other terms that involve the linear power spectrum $P_L(k)$, we replace it with the IR-resummed one. The IR-resummed power spectrum for matter has the form

$$P^{\rm IR-res}_{\rm LO}(k) = P^{nw}_{\rm L}(k) + e^{-k^2\Sigma^2} P^{w}_{\rm L} \tag{354}$$

$$P^{\rm IR-res}_{\rm NLO}(k) = P^{nw}_{\rm L}(k) + P^{nw}_{\rm 1-loop}(k) + e^{-k^2\Sigma^2}\left[(1+k^2\Sigma^2)P^{w}_{\rm L} + P^{w}_{\rm 1-loop}\right]\,, \tag{355}$$

where nw and w parts are the broadband (no-wiggle) and the oscillatory parts of the power spectrum, and Σ is the velocity dispersion. Similar relations hold for biased tracers. The effectiveness of this approximation has been shown to fit the simulated data without any residual wiggles appropriately.

Summarizing, the one-loop EFT model for the marked power spectrum of biased tracers in real space has the following form

$$\bar{m}^2 M(k) = M_{11}(k) + M_{22}(k) + 2M_{13}(k) + M_{ct}(k) + M_{\rm stoch}(k)\,, \tag{356}$$

which contains the following five parameters[24]

$$\{b_1, b_2, b_{\mathcal{G}_2}, c_s^2, P_{\rm shot}\}\,. \tag{357}$$

At the perturbative order that we are considering, the galaxy bispectrum presents only the following stochastic terms

$$B_{\rm stoch}(k_1,k_2,k_3) = P_{\rm shot}\left[P_L(k_1) + P_L(k_2) + P_L(k_3)\right] + B_{\rm shot}\,, \tag{358}$$

adding the parameter $B_{\rm shot}$ to the list of varied parameters in Equation 357.

Ref. [266] showed that, differently from the standard perturbative treatment of the power spectrum, when considering marked statistics, the perturbative series converges more slowly, and loop terms can give non-zero contributions also to linear scales. The authors came up with a low-k correction for the marked power spectrum that, in real space, is

$$M|_{\rm low-}k(k) = C^2_{\delta_{\rm M}}(k)\left[a_0 P_L(k) + c_0\right]\,, \tag{359}$$

with the addition of two more parameters. We can already exclude the second one since it is completely degenerate with the shot noise term, leaving only the free parameter a_0. One should also vary this parameter, but we will fix it for the following reasons. In [266], they showed that for biased tracers in real space this should not matter much, which also confirmed by our fit against simulations shown in section VII C; in this work we want to explore what is the actual gain of the marked power spectrum compared to standard perturbative analyses and fixing $a_0 = 0$ represents the optimistic scenario where we can completely trust our theory model and exploit all of its information content.

---

[24] Notice that we are fixing $b_{\Gamma_3} = 0$, as the power spectrum alone is poorly sensitive to this parameter [273, 274].

B Primordial non-gaussianities of the non-local type

The main goal of this chapter is to investigate the constraining power of marked statistics on primordial non-gaussianities (PNG). In particular, we are interested in primordial non-gaussianities of the non-local type, meaning the non-gaussianities that are suppressed in the squeezed limit and that can arise from quadratic or higher order interaction in the inflationary field [275].

In this work, we focus on non-local PNG because the potential gain from analyzing the bispectrum is maximized in this scenario, particularly due to the absence of the scale-dependent bias peaked at large scales that appear in local PNGs [276–278]. Some works have recently explored non-local PNG in the context of machine-learning base approaches to enhance signal extraction [279, 280]. Additionally, within the framework of the Cosmological Collider [281], recently adopted on the BOSS survey [282], several theoretical models involving massive particles during inflation can be effectively approximated by equilateral and orthogonal shapes in the non-local non-Gaussianity regime. This approximation allows for a more accurate representation of interactions in the early universe and provides a pathway for distinguishing between different inflationary models.

It is useful to represent non-local PNG using a linear combination of two shapes, equilateral and orthogonal [283], depending on the triangular configuration over which the bispectrum is peaked. Concretely, the two parameters for the amplitudes of different shapes, $f_{\rm NL}^{\rm eq}$ and $f_{\rm NL}^{\rm or}$, can be directly linked to physical observables of the inflaton field. The Lagrangian of general single-field models in the context of the effective field theory of inflation is given by [275, 283]

$$S = \int d^4x\sqrt{-g}\Bigg[-\frac{M_{\rm Pl}^2\dot{H}}{c_{\pi,s}^2}\left(\dot{\pi}^2 - c_{\pi,s}^2\frac{(\partial_i\pi)^2}{a^2}\right) - M_{\rm Pl}^2\dot{H}(1-c_{\pi,s}^{-2})\dot{\pi}\frac{(\partial_i\pi)^2}{a^2} + \left(M_{\rm Pl}^2\dot{H}(1-c_{\pi,s}^{-2}) - \frac{4}{3}M_3^4\right)\dot{\pi}^3\Bigg], \tag{360}$$

with the scalar perturbation π being related to the curvature perturbation $\zeta = -H\pi$. The two leading interactions in Equation 360 are given by the third order terms $\dot{\pi}(\partial_i\pi)^2$ and $\dot{\pi}^3$, which produce specific bispectra with amplitudes [284]

$$f_{\rm NL}^{(\dot{\pi}(\partial_i\pi)^2)} = -(85/324)(c_{\pi,s}^{-2}-1)\,, \quad \text{and} \quad f_{\rm NL}^{\dot{\pi}^3} = -(10/243)(c_{\pi,s}^{-2}-1)\left[\tilde{c}_3 + (3/2)c_{\pi,s}^2\right], \tag{361}$$

where we have defined the dimensionless parameter $\tilde{c}_3$ through $\tilde{c}_3(c_{\pi,s}^{-2}-1) = 2M_3^4c_{\pi,s}^2/(\dot{H}M_{\rm Pl}^2)$. The two EFT shapes can be projected onto the equilateral and orthogonal shapes as

$$f_{\rm NL}^{\rm equil} = \frac{1-c_{\pi,s}^2}{c_{\pi,s}^2}\left[-0.275 - 0.0780c_s^2 - 0.53\tilde{c}_3\right], \tag{362}$$

$$f_{\rm NL}^{\rm ortho} = \frac{1-c_{\pi,s}^2}{c_{\pi,s}^2}\left[0.0159 - 0.0167c_{\pi,s}^2 - 0.01113\tilde{c}_3\right]. \tag{363}$$

$$B_\zeta(k_1,k_2,k_3) = \frac{18}{5}f_{\rm NL}\Delta_\zeta^4\frac{\mathcal{S}(k_1,k_2,k_3)}{k_1^2k_2^2k_3^2}\,, \tag{364}$$

with the equilateral and orthogonal templates defined as [39, 283]

$$\mathcal{S}_{\rm eq}(k_1,k_2,k_3) = \left(\frac{k_1}{k_2} + 5\text{ perms.}\right) - \left(\frac{k_1^2}{k_2k_3} + 2\text{ perms.}\right) - 2\,, \tag{365}$$

$$\mathcal{S}_{\rm ort}(k_1,k_2,k_3) = (1+p)\frac{\Delta}{e_3} - p\frac{\Gamma^3}{e_3^2}\,, \tag{366}$$

where $p = 8.52587$, $\Delta = (k_T - 2k_1)(k_T - 2k_2)(k_T - 2k_3)$

$$k_T = k_1 + k_2 + k_3, \quad e_2 = k_1k_2 + k_1k_3 + k_2k_3, \quad e_3 = k_1k_2k_3, \quad \Gamma = \frac{2}{3}e_2 - \frac{1}{3}(k_1^2 + k_2^2 + k_3^2)\,. \tag{367}$$

Here we are making use of the so-called *orthogonal-LSS* shape, that, contrary to the *orthogonal-CMB* shape, does not present the usual scale dependent bias effect on large scales [283]. This choice is motivated by many inflationary models that can physically produce such signal [283, 285].

The PNG affects the late-time dark matter distribution, inducing an additional bispectrum term given by

$$\langle \delta^{(1)}\delta^{(1)}\delta^{(1)}\rangle = f_{\rm NL}B_{111}(k_1,k_2,k_3)(2\pi)^3\delta_D(\boldsymbol{k}_{123}) \tag{368}$$

with

$$f_{\rm NL}B_{111}(k_1,k_2,k_3) = \mathcal{T}(k_1)\mathcal{T}(k_2)\mathcal{T}(k_3)B_\zeta(k_1,k_2,k_3)\,, \tag{369}$$

where we have introduced the transfer function $\mathcal{T}(k) \equiv \delta^{(1)}(\boldsymbol{k})/\zeta(\boldsymbol{k}) = (P_{11}(k)/P_\zeta(k))^{1/2}$. Notice that this can be easily calculated via the definition of the power spectrum of initial fluctuations

$$P_\zeta(k) = \Delta_\zeta^2 k^{-3}\left(\frac{k}{k_*}\right)^{n_s-1}\,. \tag{370}$$

Analysis of *Planck* data finds $\Delta_\zeta^2 \simeq 4.1\times10^{-8}$, $n_s \simeq 0.96$ for the pivot scale $k_* = 0.05\ {\rm Mpc}^{-1}$ [286]. The initial bispectrum also enters the 1-loop galaxy power spectrum through the 1-2 term

$$f_{\rm NL}P_{12}(\boldsymbol{k}) = 2f_{\rm NL}K_1(\boldsymbol{k})\int_{\boldsymbol{q}} K_2(\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q})B_{111}(k,q,|\boldsymbol{k}-\boldsymbol{q}|)\,. \tag{371}$$

In addition, non-local PNG modulates galaxy formation, which is captured by the scale-dependent galaxy bias [46**?**]

$$\delta_g = b_1\delta + f_{\rm NL}b_\zeta\left(\frac{k}{k_{\rm NL}}\right)^2\zeta + \text{nonlinear}\,, \tag{372}$$

so that the final model for the galaxy power spectra and bispectra in real space is

$$P_g(\boldsymbol{k}) = P_{g,\rm G}(k) + f_{\rm NL}\left(P_{12}(\boldsymbol{k}) + 2b_\zeta K_1(k)\frac{k^2}{k_{\rm NL}^2}\frac{P_L(k)}{\mathcal{T}(k)}\right)\,, \tag{373}$$

$$B_g(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) = B_{g,\rm G}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) + f_{\rm NL}K_1(k_1)K_1(k_2)K_1(k_3)B_{111}(k_1,k_2,k_3)\,, \tag{374}$$

where $P_{g,\rm G}$ and $B_{g,\rm G}$ are the standard Gaussian power spectrum and bispectrum models. Following [287], we neglect all higher order correction terms, such as $O(f_{\rm NL}^2)$ PNG terms, non-Gaussian two-loop terms and contributions generated by higher order terms such as $\delta\nabla^2\zeta$, as they are subdominant for the scales considered in this work. Notice that for PNG of the local type some of these terms should be included as in that case PNG bias terms would dominate on large scales.

For the marked field, we have the additional term coming from the initial bispectrum contribution

$$f_{\rm NL}M_{12}(\boldsymbol{k}) = 2f_{\rm NL}H_1(\boldsymbol{k})\int_{\boldsymbol{q}} H_2(\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q})B_{111}(k,q,|\boldsymbol{k}-\boldsymbol{q}|)\,, \tag{375}$$

so that the marked power spectrum and bispectrum for galaxies in real space are

$$M_g(\boldsymbol{k}) = M_{g,\rm G}(\boldsymbol{k}) + f_{\rm NL}\left(M_{12}(\boldsymbol{k}) + 2b_\zeta K_1(k)\frac{k^2}{k_{\rm NL}^2}C^2_{\delta_{\rm M}}(k)\frac{P_L(k)}{\mathcal{T}(k)}\right) \tag{376}$$

$$BM_g(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) = BM_{g,\rm G}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) + H_1(\boldsymbol{k}_1)H_1(\boldsymbol{k}_2)H_1(\boldsymbol{k}_3)f_{\rm NL}B_{111}(k_1,k_2,k_3)\,, \tag{377}$$

with

$$BM_{g,\rm G}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) = 2H_1(\boldsymbol{k}_1)H_1(\boldsymbol{k}_2)H_2(\boldsymbol{k}_1,\boldsymbol{k}_2)P_L(k_1)P_L(k_2) + 2\text{ perms.}\,. \tag{378}$$

Writing down $M_{12}(\boldsymbol{k})$ explicitly:

$$\begin{aligned} M_{12}(\boldsymbol{k}) = 2C_{\delta_M}(k)^2 b_1\int_{\boldsymbol{q}}\left[b_1F_2(\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q}) + \frac{b_2}{2} + b_{\mathcal{G}_2}\sigma^2(\boldsymbol{k}-b\boldsymbol{q},\boldsymbol{q})\right]B_{111}(k,|\boldsymbol{k}-\boldsymbol{q}|,q)) \\ +2C_{\delta_M}(k)b_1^3\int_{\boldsymbol{q}} C^2_{\delta_M}(|\boldsymbol{k}-\boldsymbol{q}|,q)B_{111}(k,|\boldsymbol{k}-\boldsymbol{q}|,q)\,. \end{aligned} \tag{379}$$

Through this work, we will assume that the following relation holds for the primordial non-Gaussian bias and the linear bias

$$b_\zeta = \frac{18}{5}\delta_c(b_1-1)\,, \tag{380}$$

motivated by the peak-background split argument [46], with δ_c being the critical overdensity for the spherical collapse.

It is worth noticing that, at the power spectrum level, the PNG contributions are entered only at the one-loop level. Moreover, as shown in [287], the effect of equilateral and orthogonal non-Gaussianities is very degenerate with the second order biases b_2 and $b_{\mathcal{G}_2}$. These two effects conspire to make power spectrum only analyses very inefficient in measuring non-local PNG, making the inclusion of the bispectrum essential in order to to constrain them. The marked power spectrum seems very promising in this direction by bringing information from higher-order functions back to a two-point function, as also shown in [267]. This could naively explain the results of ref. [264], where the authors find that the marked power statistic helps in constraining non-local PNG much more than that of the local type.

C Comparison to simulations

Before proceeding with the Fisher forecast, we compare our predictions with the Quijote [288] and Quijote-PNG [285] simulations. These are N-body simulations of volume 1 $(\text{Gpc}/h)^3$, containing 512^3 particles each, and run using the TreePM code GADGET-III from initial conditions generated at $z = 127$ be the codes 2LPTIC [289] and 2LPTPNG [285, 290], for the simulations with and without PNG respectively. We will compare our PT predictions with real space dark matter halos identified in each simulation by the standard Friends-of-friends algorithm [291] by setting the linking length parameter to $b = 0.2$ and considering halos with more the 20 dark matter particles. We compute the marked power spectrum for halos with minimum mass $M_{\rm min} = 3.2\times10^{13}\,M_\odot/h$, following the procedure described in [292, 293]. The fiducial cosmology is $\{\sigma_8, n_s, \Omega_b, \Omega_m, h\} = \{0.834, 0.9624, 0.049, 0.3175, 0.6711\}$. We compare with the

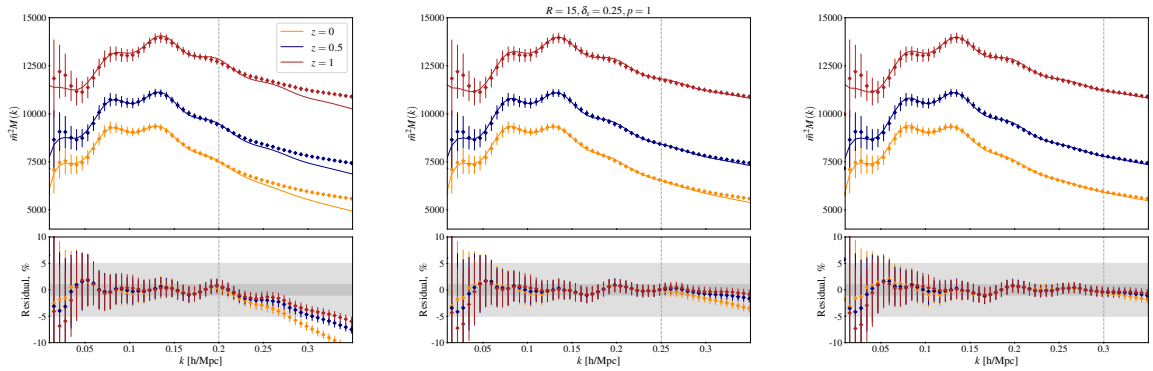

Figure 34: Results of the fitting of the marked power spectrum from the Quijote-PNG simulations with equilateral PNG, $f_{\rm NL}^{\rm eq} = 100$ for different redshifts. The left plot is obtained fitting the simulations up to $k^M_{\rm max} = 0.2\,h/{\rm Mpc}$, the central one with $k^M_{\rm max} = 0.25\,h/{\rm Mpc}$ and the right one with $k^M_{\rm max} = 0.3\,h/{\rm Mpc}$. The diamonds represent the mean marked power spectrum from the simulations, while the solid lines are the results obtained from the EFT fit.

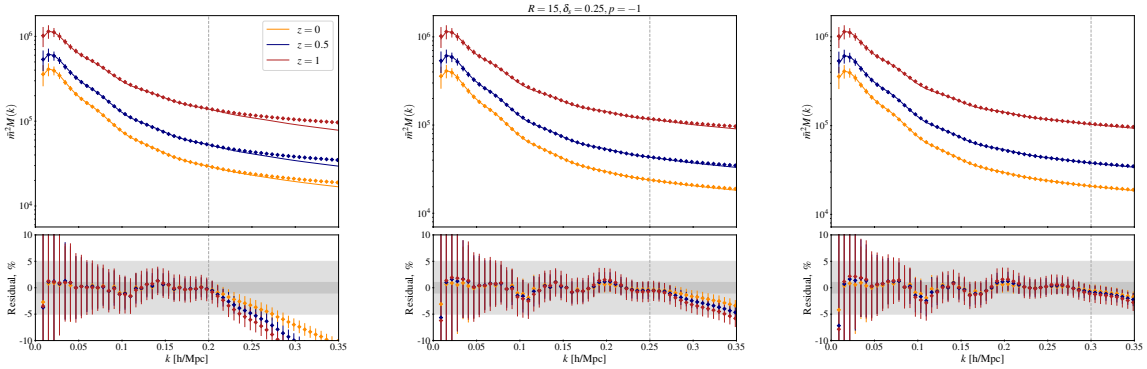

Figure 35: Same as figure 34, with $p = -1$.

simulations at redshifts $z = \{0, 0.5, 1\}$ and mark parameters $\{R = 15\,h/{\rm Mpc}, \delta_s = 0.25\}$ and $p = \{-1, 1, 2\}$ to test our PT model and numerical implementation by fixing the cosmological parameters to their fiducial values and fitting the bias, small-scales and stochastic parameters $\{b_1, b_2, b_{\mathcal{G}_2}, c_s^2, P_{\rm shot}\}$ to the simulation power spectrum. The results of this fits are shown in figures 34, 35, 36, where we display the mean marked power spectrum obtained from the PNG simulations with equilateral shape and $f_{\rm NL} = +100$ fitted up to $k_{\rm max} = \{0.20, 0.25, 0.30\}\ h/{\rm Mpc}$. To perform this comparison, we estimate the data covariance numerically using the scattering among the different realizations of the simulations with the same cosmology.

We observe that the theoretical predictions with the EFT of LSS are in excellent agreement with the simulations for the cases with $p = 1$ and $p = -1$, within 1 %, up to the maximum $k_{\rm max}$ adopted for the fitting and for different redshifts. These results confirm again the additional terms introduced in [266] to control the low-k behavior of the marked power spectrum are not needed for biased tracers in real space, also in the presence of non-local PNG. Moreover, we do not see any residual wiggles, corroborating the validity of the approximation adopted for the IR-resummation. For the case $p = 2$ more caution is needed. Looking at the expression for the PT kernels for the marked power spectrum, Equation (335) and Equation (338) we see that the linear contribution can go to zero for some values of k in some $\{R, \delta_s, p\}$

configuration. The condition for these terms to go to zero is given by

$$W_R(k) = \frac{1+\delta_s}{p}\,, \tag{381}$$

where we used $C_0 = 1$ and $C_1 = p/(1+\delta_s)$. For $\{R, \delta_s, p\} = \{15, 0.25, 2\}$, this condition is satisfied for $k \simeq 0.06\,h/\mathrm{Mpc}$, implying that around that scales some contribution of the marked power spectrum go to zero. For the halo marked spectrum, the agreement between data and out model decreases to $10-20\%$, as shown in figure 36. We can see that, especially for low redshifts, the scales around $k \simeq 0.06\,h/\mathrm{Mpc}$ are poorly fitted, and the disagreement is more pronounced when the k_{max} is higher. We still observe an agreement of $\sim 5\%$ for the redshift and the scales considered for this analysis, and hence we will use this configuration to perform our forecast.

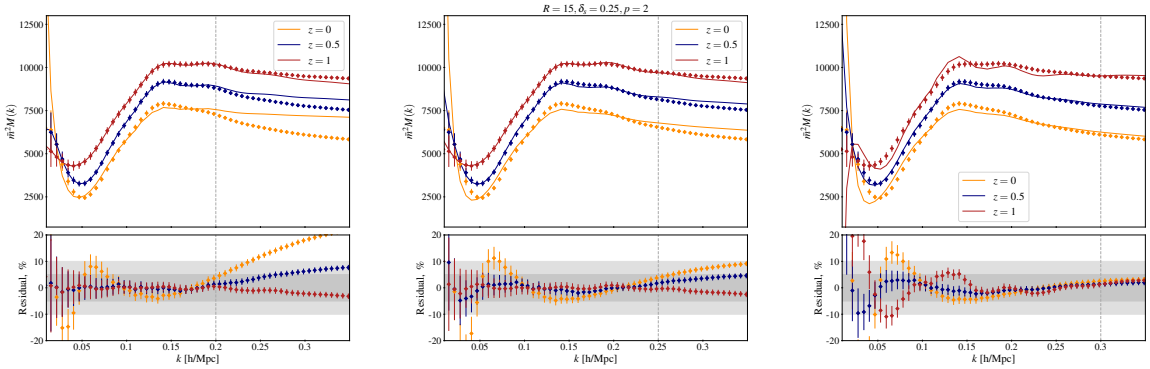

Figure 36: Same as figure 34 and 35, with $p = 2$.

D *Effect of different mark parameters*

In this section we investigate how the marked power spectrum $M(k)$ responds to variations in the mark parameters. We also assess the effectiveness of pertrubation theory in reproducing the simulation data for various combinations of these parameters. We recall that we are adopting the usual prescription for the mark, defined as:

$$m(\mathbf{x}) = \left(1 + \left(\frac{\delta_{g,R}(\mathbf{x})}{1+\delta_s}\right)\right)^{-p} \tag{382}$$

where $\delta_{g,R}$ is the local smoothed density contrast at scale R, δ_s is the density scaling parameter that sets the effective normalization, also referred as bias, and p is the hyper-parameter controlling the strength of the down-weighting in high-density regions.

As fully explained in the main text, by tuning R, δ_s, and p, the mark selectively enhances the contribution from underdense or overdense regions, where nonlinear effects could be weaker and perturbation theory is expected to be more reliable. We explore separately the effect of each parameter, as they play different roles in the non-linear modeling adopted in this work. We measure the marked power spectrum from the Quijote simulations using the mark parameters combinations showed in table V.

For each configuration, we compute the marked power spectrum $M(k)$, and its covariance $C_M(k)$ from the scattering among different realizations of the same simulations. We show here only the redshift $z = 0.5$ as it is of main interest for the results presented in this thesis, but we stress that the same arguments are general for each redshift considered.

We first investigate the problematic case of $p = 2$ that can present peculiar features that could be

R	15	15	15	15	15	15	15	15	15	5	10	30
δ_s	0.25	1	2	3	4	0.25	0.25	0.25	0.25	0.25	0.25	0.25
p	1	2	2	2	2	2	-1	1.25	1.5	1	1	1

Table V: Mark parameters used in this section.

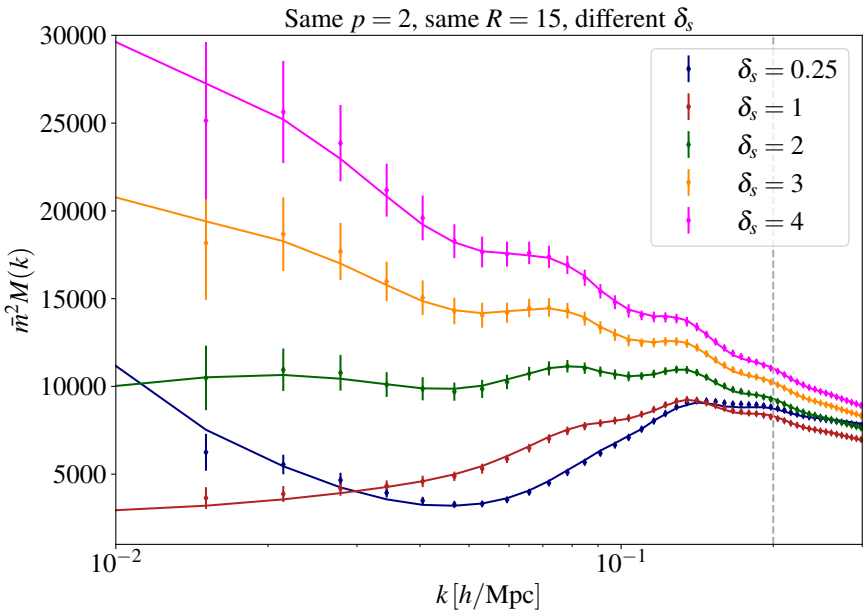

Figure 37: Effect of varying the bias δ_s, fixing the other mark parameters p and R.

hard to capture using an analytical prescription, as already discussed in the main text. Equations. (338) show that for some combination of the mark parameters, there are scales for which some contributions of the marked power spectrum get canceled exactly: in particular this is the case when the function $C_{\delta_M}(k)$ goes to zero. This condition translates to

$$1 - C_1 W_R(k) = 0\,, \tag{383}$$

with $C_1 = p/(1 + \delta_s)$. This equation is always satisfied for some value of k when the combination of parameters $\{p, \delta_s\}$ satisfies $p \geq 1 + \delta_s$ and $\delta_s \geq -1$. For negative values of p and positive for δ_s the marked power spectrum is always positive and never goes to zero. On the other hand, when $p = 2$, if $-1 \leq \delta_s \leq 1$, then some contributions of the marked power spectrum go to zero and the theoretical modeling becomes more challenging and less accurate. This is shown explicitly in figure 37, where $p = 2$ and $R = 15$ are kept fixed and δ_s is varied. For $\delta_s = 0.25$, C_{δ_M} goes zero for $k \simeq 0.06\,h/\mathrm{Mpc}$, which are the scales that are poorly fitted by our theoretical model. For values $\delta_s \geq 1$, we can see that the perturbative modeling fits perfectly with the simulation data, up to mildly non-linear scales.

We also consider the effect of varying the exponent p, keeping $\delta_s = 0.25$ fixed. The condition on p to not have unwanted zero-crossing of the marked power spectrum is $p \leq 1.25$. In fact, we see that the cases with $p = 2$ and $p = 1.5$ are problematic and the theoretical model is not accurate enough to fit the larger scales. For smaller values of p, the theoretical predictions sits almost perfectly on top of the simulations data.

It is also worth investigating the role of the smoothing scale R. The condition to have a well-behaved perturbative series up to the scales we are considering, is given by Equation (326), which is strongly affected by R. Smaller R values make the fluctuations larger and, consequently, give a larger value

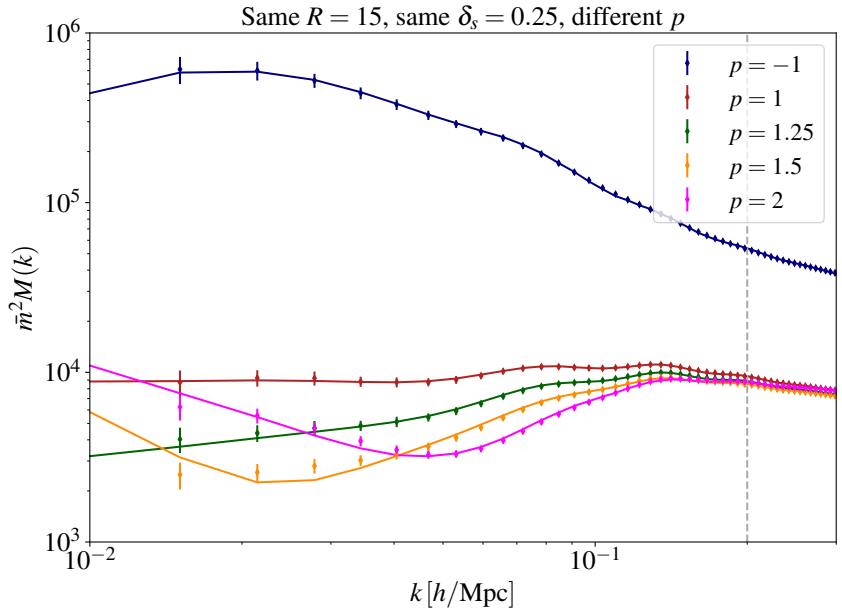

Figure 38: Effect of varying the exponent scale p, fixing the other mark parameters R and δ_s. We note that for $p \geq 1.5$ the theoretical model is less accurate due to the marked power spectrum features that appear in this mark combinations.

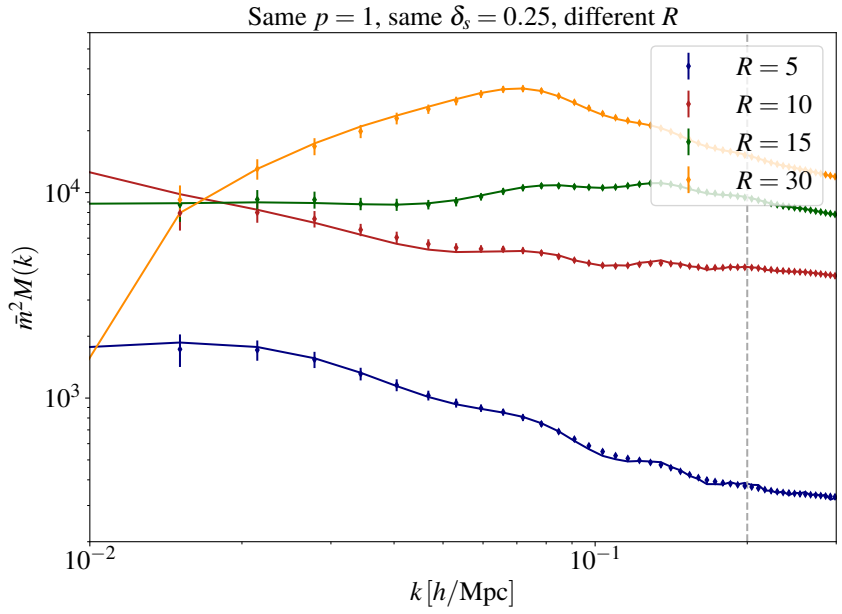

Figure 39: Effect of varying the smoothing scale R, fixing the other mark parameters p and δ_s. s

for δ_{RR}. This implies that smaller R cause the perturbative series to converge more slowly. Despite this, we observe that the agreement within theoretical model and data is in exquisite agreement for the scales analyzed, as shown in figure 39.

E Forecasts

We are now ready to study the amount of information accounted for by the marked power spectrum and compare it with a standard power spectrum plus bispectrum analysis. This can be readily addressed by assuming, at leading order, that the statistics considered follow a multivariate Gaussian distribution and compute the Fisher matrix. The statistics that we will consider are the power spectrum, the marked power spectrum and the bispectrum for galaxies in real space. We will compute these statistics by varying the value of $k_{\rm max}$ for P and for M, while we will keep fixed $k^B_{\rm max} = 0.10h/{\rm Mpc}$ for the bispectrum. This has been shown to be a mildly conservative choice as the tree-level model for the bispectrum is valid up to these scales for the volume that we will consider in this work [183, 271].

Previous analyses showed that the marked power spectrum always outperforms the standard power spectrum because it is able to keep track of the cosmic web through the non-linear mark in Equation 320: different values of the exponent p can enhance voids ($p > 0$) or nodes ($p < 0$). [262] finds that positive values of p are optimal to constrain the total neutrino mass and modified gravity theories, given that voids are more sensitive to these two ingredients. In this work, we will perform a Fisher forecast up to mildly non-linear scales using $p = 2, 1$ to enhance underdense regions and $p = -1$ to mark the overdense ones to constrain massive neutrinos and non-local-primordial non-gaussianities and compare with the standard analysis of power spectrum combined with bispectrum. The fiducial cosmology adopted here is one of the Quijote simulations for the standard cosmological parameters, with the additional parameters $\{M_\nu, f^{\rm eq}_{\rm NL}, f^{\rm or}_{\rm NL}\} = \{0.06, 0., 0.\}$. We will run the analysis separately for massive neutrinos and primordial non-gaussianities.

For the fiducial of the galaxy biases, we adopt a semi-analytic model of galaxy formation to fix them; specifically, for the linear bias we adopt the following model[25]

$$b_1(z) = 0.9 + 0.4z\,. \tag{384}$$

For the higher order biases we adopt the following fitting formulae [29, 294], obtained from a combination of N-body simulations and halo occupation modeling

$$b_2(z) = -0.704 - 0.208z + 0.183z^2 - 0.00771z^3\,. \tag{385}$$

For $b_{\mathcal{G}_2}$ and b_{Γ_3} we use instead the co-evolution model [29, 295], which gives

$$b_{\mathcal{G}_2}(z) = -\frac{2}{7}(b_1(z) - 1), \quad b_{\Gamma_3}(z) = \frac{23}{42}(b_1(z) - 1)\,. \tag{386}$$

For the higher derivative term, we adjusted the prescription of [273] in order to match with the BOSS data

$$c_s = -D^2(z)[{\rm Mpc}/h]^2\,, \tag{387}$$

with $D(z)$ being the growth factor for linear perturbations. Finally, the fiducial values for the stochastic terms are set to the Poisson sampling prediction

$$P_{\rm shot} = \bar{n}_g^{-1} \quad B_{\rm shot} = \bar{n}_g^{-2}\,. \tag{388}$$

To compute the one-loop prediction of the standard and the marked power spectrum we implement the public FFTLog-based code CLASS-PT [227], an extension of the Boltzmann solver CLASS [296] that includes the EFTofLSS as non-linear model. For massive neutrino cosmologies we just modify the

[25] Notice that this model should describe Euclid-like Hα targets [273], but we will employ this model for BOSS-like volumes as well since we are only interested in the relative change between different statistics.

z	b_1	b_2	$b_{\mathcal{G}_2}$	b_{Γ_3}	$c_s^2[\mathrm{Mpc}/h]^2$	$P_{\rm shot}[\mathrm{Mpc}/h]^3$	$B_{\rm shot}[\mathrm{Mpc}/h]^6 \times 10^{-3}$
0.61	1.14	-0.38	-0.041	0.079	0.53	3333	11111

Table VI: Table of fiducial values for the redshift, bias parameters, counterterm and shot-noise parameters adopted for the forecast.

linear power spectrum and input it inside the one-loop integrals, neglecting the non-linear effects in the perturbative kernels and the scale dependent growth function as this approximation has been shown to perform well for the current constraints on neutrino masses [273, 297].

We perform a forecast for BOSS CMASS2 redshift $z = 0.61$ and total volume $V_{\rm BOSS} = 3.83\ (\mathrm{Gpc}/h)^3$, with number density $\bar{n}_g^{\rm BOSS} = 3\times10^{-4}\ (h/\mathrm{Mpc})^3$. We consider diagonal Gaussian covariances, after having tested that this approximation correctly reproduces the ones obtained with the Quijote simulations. They are defined as

$$C_P(k) = \frac{2(2\pi)^3}{VV_s}P^2(k), \quad C_M(k) = \frac{2(2\pi)^3}{VV_s}M^2(k), \tag{389}$$

$$C_B(k_1, k2, k_3) = \frac{(2\pi)^6 s_{123}}{VV_{123}}P(k_1)P(k_2)P(k_3) \tag{390}$$

where $V_s = 4k^2\delta k$ and $V_{123} = 8\pi^2 k_1 k_2 k_3 \delta k^3$ are the shell volumes for the power spectrum and the bispectrum, $\delta k = 0.01\ \ h/\mathrm{Mpc}$, s_{123} is the triangular symmetry factor and V is the volume of the survey. To avoid double counting of the information content in the different spectra, we also include the cross covariance between the marked and the standard power spectrum

$$C_{PM}(k) = \frac{2(2\pi)^3}{VV_s}P(k)M(k)\,. \tag{391}$$

In ref. [262] it has been shown that the Gaussian approximation for the covariance matrix is particularly good for the marked power spectrum, while for the power spectrum and the bispectrum one should, in principle, also include non-Gaussian terms. These are expected to give a negligible contribution up to the scales considered, although for the bispectrum one should include them to properly estimate the forecasted errorbars, especially for primordial non-Gaussainities [298].

In standard LSS analyses, the abundance of baryons is usually constrained using BBN priors [299, 300], and hence we decide to fix it to its fiducial value. We vary the cosmological parameters $\{\ln(10^{10}A_s), n_s, \omega_{\rm cdm}, h, M_\nu, f_{\rm NL}^{\rm Equil}, f_{\rm NL}^{\rm Ortho}\}$, with the total neutrino mass and the two PNG parameters varied separately. For galaxies, we fix the third order bias parameter b_{Γ_3} to its theoretical prediction, as it is poorly constrained [273], and we vary $\{b_1, b_2, b_{\mathcal{G}_2}, c_s, P_{\rm shot}, B_{\rm shot}\}$.

The list of fiducial parameters is listed in tab. VI

In the following we will present our results and discuss various effects that contribute to the cosmological parameter measurement from the standard and marked power spectrum and their combination with the bispectrum. We first discuss the results obtained analyzing the dark matter only case and then we will go through the results for galaxies.

LSS observation always involve luminous sources (galaxies, quasars, . . .), hence studying the dark matter field is a merely academic exercise; nevertheless, it can still be worthy as it can give useful insights on where the physical information is. It also represents an optimal case, since the unknown bias parameters will degrade the constraints and the discrete nature of galaxies will diminish the signal-to-noise ratio due to the shot-noise. The results of our analysis are shown in fig. 40 for the neutrino mass M_ν and fig. 41 for non-local PNG. We decide not to show the comparison between the standard power spectrum and the marked one as it has been shown that, with the same volumes and sky-cuts, the

latter gives better constraints on almost all the cosmological parameters. We compare here two different

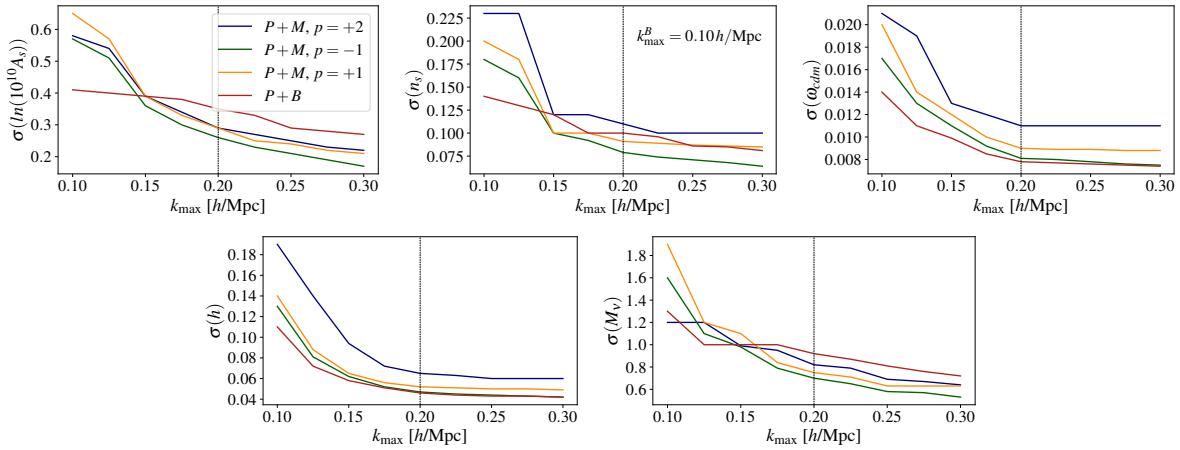

Figure 40: 2σ forecasted errorbars for dark matter only in presence of massive neutrinos with BOSS CMASS volume and redshift (see main text). The dashed vertical line represents the maximum wave vector adopted in standard analyses [274].

mark configuration to understand if the information is coming from different structures of LSS web (voids as underdense regions or filaments and nodes for overdense ones) or it is just information taken form higher order correlation functions. The effects of massive neutrinos on LSS are mainly a rescaling of the growth of linear perturbation and a erasing of small scale clustering below the free streaming scale [156]. The former is the most visible and detectable, while the latter is slightly degenerate with the small scales parameter, the counterterm c_s^2 that already appears at level of matter, and the higher order biases that are present for biased tracers. For all the parameters, in the analysis where the total neutrino mass is varied, we can see that different marks, $p = 2, 1$ for voids and $p = -1$ for nodes, give constraints that are almost comparable, with slightly better results for positive p. Most importantly, we can see that these constraints are comparable with the $P + B$ analysis, which, we remind, has a fixed $k^B_{\rm max} = 0.10\,h/{\rm Mpc}$: this confirms that the additional information of the marked power spectrum is mostly a bispectrum-like contribution. In particular, for the neutrino mass M_ν we can see that for $k_{\rm max} \gtrsim 0.17\,h/{\rm Mpc}$, the constraints obtained using the marked spectra are slightly better compared to the $P + B$ analysis, most probably due to the smaller scales probed with M.

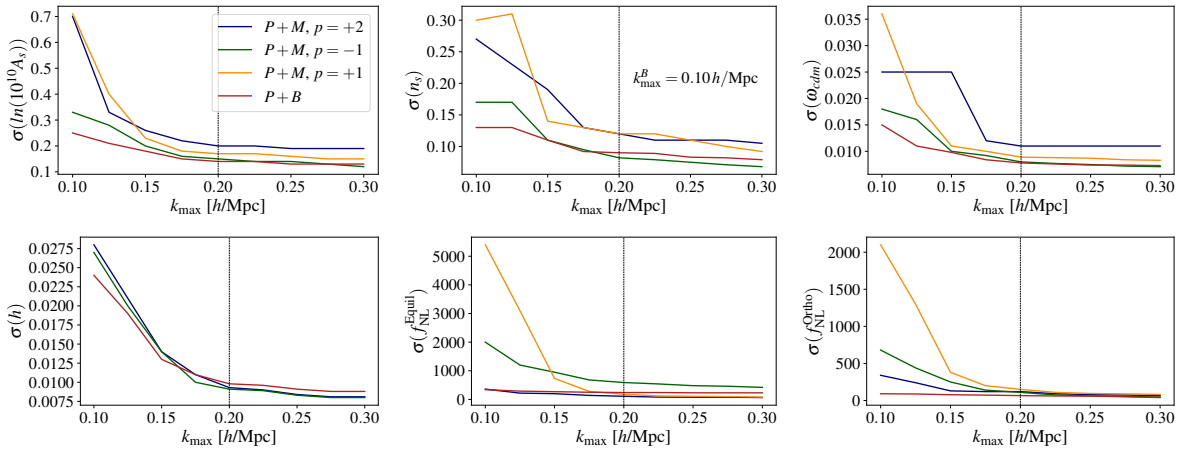

Figure 41: Same as figure 40 but for cosmologies with primordial non-gaussianities.

For the primordial non-gaussianities, the outcomes are slightly different. Non-local non-gaussianities affect the standard power spectrum only through one-loop contribution (differently from the well-known

large scales effect for non-gaussianities of the local type [276, 277]) and hence their effect is either tiny and degenerate with higher order biases or counterterms. This means that the bispectrum plays a crucial role in constraining this kind of non-gaussianities as it is the main source of signal. Figure 41 shows that the errors $\sigma(f_{\rm NL}^{\rm Equil})$, $\sigma(f_{\rm NL}^{\rm Ortho})$ from the $P+M$ analysis are always bigger than the $P+B$ one, at least on the scales at which EFT of LSS is still reliable for BOSS volumes. We see that the case with $p=2$ slightly outperforms the $P+B$ analysis for PNG of equilateral shape. As we already discussed in the previous section, the theoretical results for this mark are less accurate compared to the other two and hence these results might be slightly affected by this. We see that the differences between a positive and a negative p are negligible and that the results we obtain are very similar to those obtained with the bispectrum: this again confirms that the information content of the marked power spectrum is in fact a subset of the information contained in higher order correlation functions. While for the total neutrino mass this was not obvious as the contribution of the bispectrum to its constraints are not essential, in the case of non-local primordial non-gaussianities this results to be evident.

At this stage, it is worth comparing our findings with the results of [262], where the authors investigated the constraining power of the marked power spectrum on massive neutrinos using a simulation-based approach. Using the cb (cold dark matter + baryons) component, they find that using the marked power spectrum produces an improvement of ~ 3 over the standard one at $k_{\rm max}=0.5\,h/{\rm Mpc}$. If consider a direct comparison between M and P, we find a similar improvement in our Fisher forecast – at scales ($k_{\rm max}=0.20-0.30\,h/{\rm Mpc}$ – when we consider the same volume, redshift and mark parameters. This again confirms the agreement with previous results, even though here we focus on the comparison with the standard $P+B$ analysis, which seems a more natural one, considering that M does incorporate higher-order information.

The previous analysis in [262] also emphasizes a void-enhancing mark is the optimal choice for constraining neutrino masses. This result is again qualitatively consistent with our findings for the matter field analysis, see figure 40. Moreover, we find that a similar conclusion holds for equilateral PNG's: as pointed out in previous works, see e.g. [301], the signal from equilateral PNG is maximized for non-linear filamentary structure and this could explain why a "void-enhanced" marked power spectrum could be effective in constraining this shape.

Biased tracers, such as galaxies, introduce a certain degree of freedom through the biases, that parametrize our incomplete knowledge about the connection between dark matter and galaxies. They are the main source of degradation of cosmological information and indeed much effort has been recently made to constrain them using hydro-dynamical simulations [243–245]. For what concerns our work, we know that, in real space, specific biases are almost perfectly degenerate with cosmological parameters. When the power spectrum alone is considered, the linear bias is completely degenerate with the primordial amplitude A_s, and this degeneracy can be slightly broken by using the information from the IR-resummed BAO. This degeneracy can be broken by combining the power spectrum and the bispectrum, at the price of introducing a new slight degeneracy among the scalar amplitude and the second order bias b_2. Including modes from smaller scales could indeed alleviate these degeneracies, allowing to constrain higher order biases through the one-loop terms. This eventually results in worse signal-to-noise for the cosmological parameters. For massive neutrinos we observe in figure 42 that the marked power spectra with $p=\pm1$ perform poorly for small $k_{\rm max}$'s, but they reach the level of constraints of $p=2$ when the loop terms are more important around $k_{\rm max}\simeq 0.2\,h/{\rm Mpc}$.

Interestingly, we see that $\sigma(M_\nu)$ for the $P+M$ case, for maximum k around the non-linear scale, can reach the level of the $P+B$ one, confirming what was already found in redshift space in [263] and [302]. It is also interesting to notice that the main effect of adding the bispectrum is usually a better measurement of the second order biases b_2 and $b_{\mathcal{G}_2}$, which are usually less constrained from P alone analyses. We see that the inclusion of the marked power spectrum allows to measure these biases at the same level of $P+B$ analyses, or even better in the case of $b_{\mathcal{G}_2}$.

Neutrino mass affects the linear power spectrum through a modification of the linear growth and

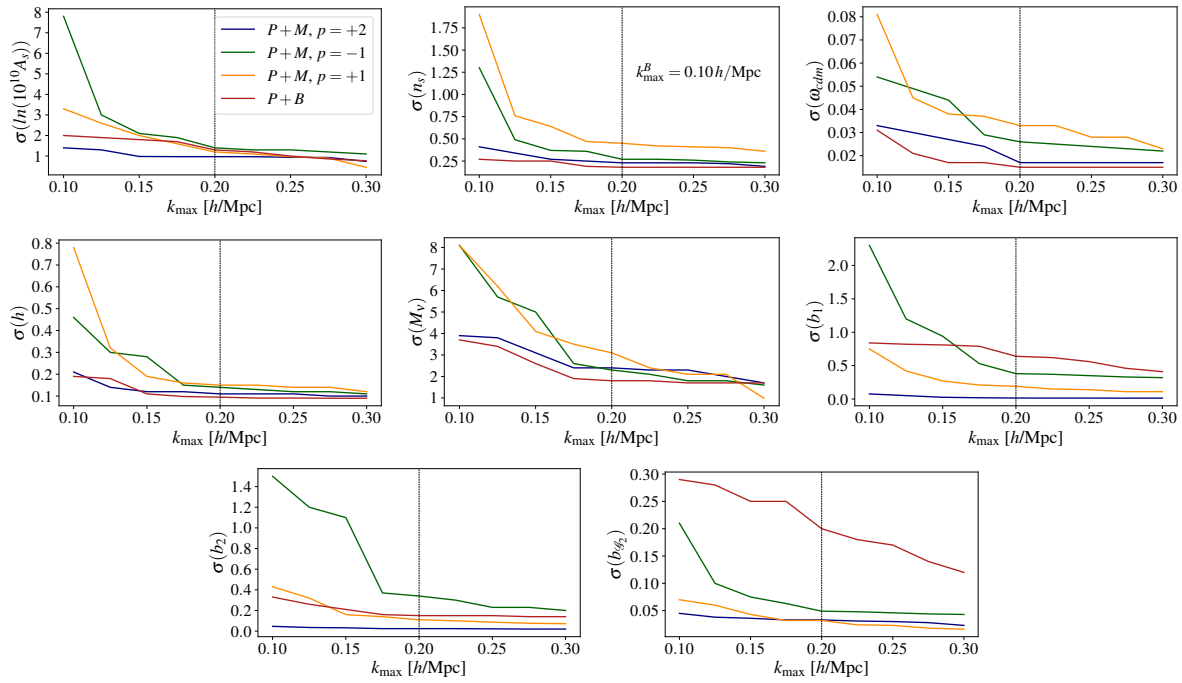

Figure 42: 2σ constraints from galaxies for massive neutrinos cosmologies and BOSS CMASS volume, number density and redshift. The dashed vertical line represent the maximum wave vector adopted in standard analysis [274].

a damping around and below the free streaming scale, which is a nonlinear effect. The galaxy power spectrum alone is mostly sensitive to the former effects, while the latter is degenerate with higher order biases and counterterms. As shown in figure 43, the marked power spectrum, especially for marks with positive p, breaks some of the degeneracies among the total neutrino mass M_ν and the biased tracers, improving the constraints on the bias parameters. This should not change dramatically when we will consider redshift space distortions since, at first order, these will allow to have a better measurement of the linear bias b_1, which does not appear to be degenerate with the neutrino total mass.

For non-local primordial non-gaussianities, we see a similar behavior to what we found with the dark matter field. We already mentioned that for the shapes studied here, the bispectrum is essential in order to constrain primordial non-gaussianities, differently from the case of massive neutrinos cosmologies, where most of the information is taken from linear scales. The results of our analysis are shown in figure 44. We notice that in both cases where $p = \pm 1$, the constraints on $f_{\rm NL}^{\rm Equil}$ and $f_{\rm NL}^{\rm Ortho}$ obtained combining P and M never reach those obtained with $P + B$, even for very non-linear $k_{\rm max}$ and the same is true for almost all the other parameters varied in this analysis, with the exception of the higher order biases. The most promising results are obtained using $p = 2$, where still some caution is required since the theoretical modeling is less accurate due to the particular features of the marked power spectrum for this case, see fig. 36. This is again a confirmation of what we mentioned earlier for massive neutrinos: if the information about a cosmological parameter within the galaxy power spectrum comes both from linear and non-linear effects, then the marked power spectrum can improve the constraints obtained when combining with the bispectrum, as it is for M_ν;[26] on the other hand, non-local PNG affect the galaxy power spectrum only through a one-loop effect, and basically all the information is contained in the bispectrum, hence the marked power spectrum can only do as well as the bispectrum does, for the scales considered. It is important however to stress that our results are obtained in real space and

[26] We stress again that usual analyses as in ref. [264] always compare P, M, B and various combinations of them at the same $k_{\rm max} = 0.5\,h/{\rm Mpc}$. Here we are using the maximum scale at which the EFTofLSS is valid, since going to higher $k_{\rm max}$ would give biased results in realistic analyses due to exclusion of higher order loops entering in the analytical model. We are, however, in agreement with their findings.

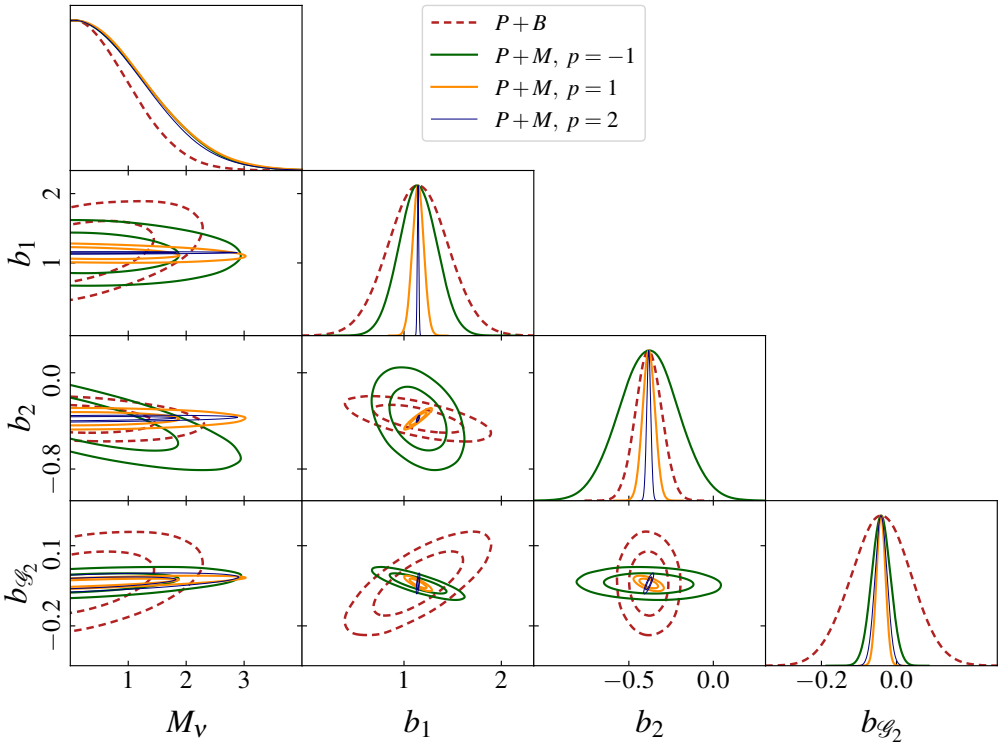

Figure 43: Triangle plot for massive neutrinos cosmology and biased tracers. We compare $P + M$ with $k^{P,M}_{\rm max} = 0.20\,h/{\rm Mpc}$ and $P + B$ with $k^{B}_{\rm max} = 0.10\,h/{\rm Mpc}$.

without accounting for redshift space distortions, survey mask effects, binning, which are expected to degrade even more our results.

For biased tracers we also see, as shown in figure 45, a small degeneracy among the PNG parameters and the linear bias b_1: this is a consequence of the universality relation assumed for the non-Gaussian bias in Equation 380.

The corner plots presented in this work suggest that combining the power spectrum and the bispectrum with different marks would give even better results either for massive neutrinos and primordial non-gaussianities. This kind of combination is beyond the scope of this work and we will investigate it in future work accounting also for redshift space distortions and other observational effects.

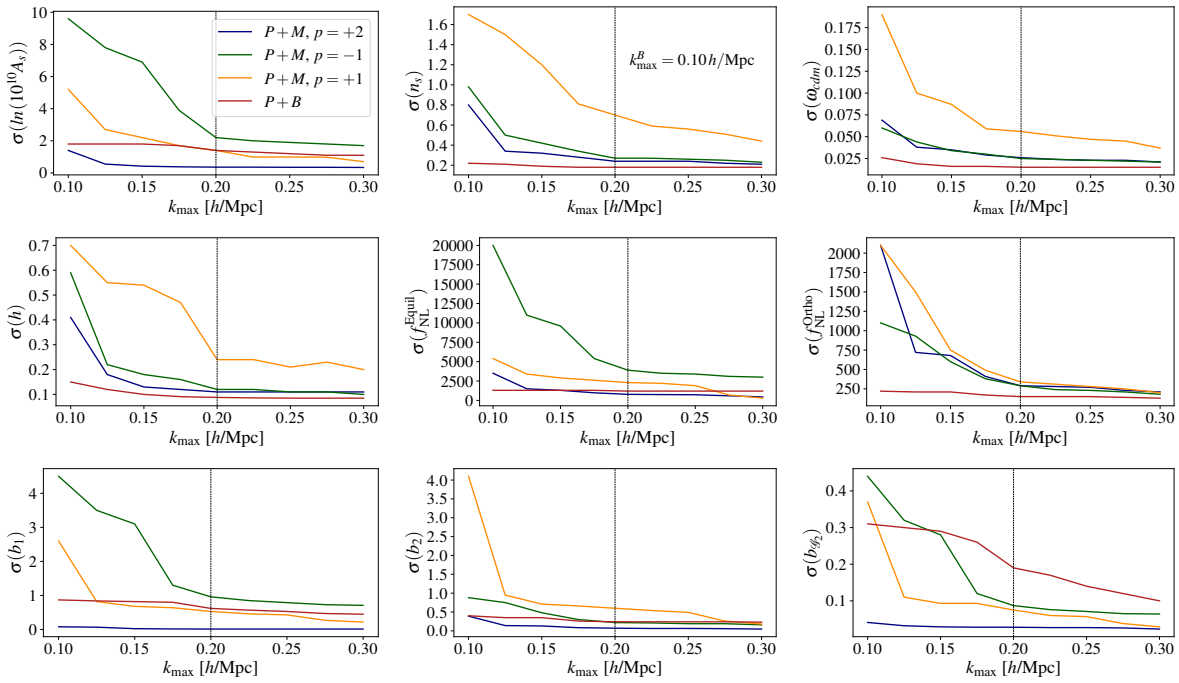

Figure 44: Same as figure 42 for cosmologies with primordial non-gaussianities.

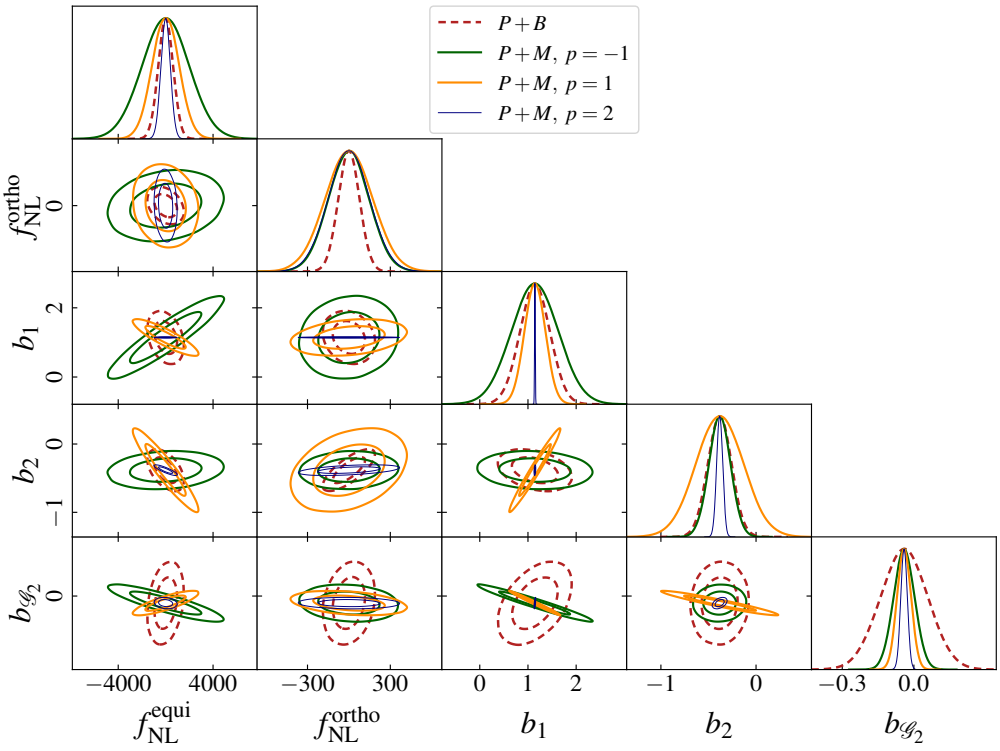

Figure 45: Same as figure 43 but for primordial non-gaussianities.

VIII Cosmological parameters from realistic Stage IV mocks: the case of *Nancy Grace Roman Space Telescope*

To give an example of an application of the theoretical model for the 1-loop power spectrum developed in this second part of the thesis, in this chapter we analyze realistic galaxy mocks on the light-cone that mimic the spectroscopic data that will be obtained by the Nancy Grace Roman Space Telescope. This realistic forecast can be useful in order to assess the capabilities of the Full Shape analysis of galaxy datasets obtained with the Roman experiment to constrain late-time cosmological parameters such as H_0, Ω_m, σ_8. Before delving into the data analysis part, let us introduce the observational effects that must be added to our power spectrum model when dealing with real-world applications.

A *Theory model for the observed power spectrum*

So far we have obtained the model for the galaxy power spectrum up to one-loop order in real space. On the other hand, in real spectroscopic surveys we measure the position of galaxies via their redshift so that RSD must be unavoidably taken into account. In the first part of the thesis we have derived the mapping from real to redshift space at linear order. However, this is not enough for the 1-loop power spectrum and non-linear PT kernels in redshift space must be derived. These non-linear kernels are well known in literature and we report them here for completeness. Up to third order we have [25]:

$$Z_1(\boldsymbol{k}) = b_1 + f\mu^2\,, \tag{392}$$

$$Z_2(\boldsymbol{k}_1,\boldsymbol{k}_2) = \frac{b_2}{2} + b_{\mathcal{G}_2}\left(\frac{(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)^2}{k_1^2k_2^2}-1\right) + b_1F_2(\boldsymbol{k}_1,\boldsymbol{k}_2) + f\mu^2G_2(\boldsymbol{k}_1,\boldsymbol{k}_2) + \frac{f\mu k}{2}\left(\frac{\mu_1}{k_1}(b_1+f\mu_2^2)+\frac{\mu_2}{k_2}(b_1+f\mu_1^2)\right)\,, \tag{393}$$

$$\begin{aligned} Z_3(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) &= 2b_{\Gamma_3}\left[\frac{(\boldsymbol{k}_1\cdot(\boldsymbol{k}_2+\boldsymbol{k}_3))^2}{k_1^2(\boldsymbol{k}_2+\boldsymbol{k}_3)^2}-1\right]\left[F_2(\boldsymbol{k}_2,\boldsymbol{k}_3)-G_2(\boldsymbol{k}_2,\boldsymbol{k}_3)\right] \\ &+ b_1F_3(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) + f\mu^2G_3(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) + \frac{(f\mu k)^2}{2}(b_1+f\mu_1^2)\frac{\mu_2}{k_2}\frac{\mu_3}{k_3} \\ &+ f\mu k\frac{\mu_3}{k_3}\left[b_1F_2(\boldsymbol{k}_1,\boldsymbol{k}_2)+f\mu_{12}^2G_2(\boldsymbol{k}_1,\boldsymbol{k}_2)\right] + f\mu k(b_1+f\mu_1^2)\frac{\mu_{23}}{k_{23}}G_2(\boldsymbol{k}_2,\boldsymbol{k}_3) \\ &+ b_2F_2(\boldsymbol{k}_1,\boldsymbol{k}_2) + 2b_{\mathcal{G}_2}\left[\frac{(\boldsymbol{k}_1\cdot(\boldsymbol{k}_2+\boldsymbol{k}_3))^2}{k_1^2(\boldsymbol{k}_2+\boldsymbol{k}_3)^2}-1\right]F_2(\boldsymbol{k}_2,\boldsymbol{k}_3) + \frac{b_2f\mu k}{2}\frac{\mu_1}{k_1} \\ &+ b_{\mathcal{G}_2}f\mu k\frac{\mu_1}{k_1}\left[\frac{(\boldsymbol{k}_2\cdot\boldsymbol{k}_3)^2}{k_2^2k_3^2}-1\right]\,, \end{aligned} \tag{394}$$

where $\boldsymbol{k} = \boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3$ and the kernel Z_3 has to be symmetrized over its arguments. Note that we omit the time dependence for simplicity but recall that all biases depend on time. Thus, we have that the

model for the 1-loop power spectrum in redshift space reads

$$P_{g,RSD}(\boldsymbol{k},z) = Z_1^2(\boldsymbol{k},z)P_{11}(k,z) + 2\int \frac{d\boldsymbol{q}}{(2\pi)^3} Z_2^2(|\boldsymbol{k}-\boldsymbol{q}|,\boldsymbol{q},z)P_{11}(|\boldsymbol{k}-\boldsymbol{q}|,z)P_{11}(q,z) \tag{395}$$

$$+ 6P_{11}(k,z)\int \frac{d\boldsymbol{q}}{(2\pi)^3} Z_3(\boldsymbol{q},-\boldsymbol{q},\boldsymbol{k},z)P_{11}(q,z)\,. \tag{396}$$

In redshift space in order to isolate the different geometrical contributions, it is customary to expand the power spectrum in Legendre polynomials as

$$P_\ell(k,z) = \sum_\ell \frac{2\ell+1}{2} L_\ell(\mu)P_{g,RSD}(k,\mu,z)\,. \tag{397}$$

In all that follows we will consider only the monopole $\ell = 0$ and quadrupole $\ell = 2$ of the power spectrum, as for the level of precision of the Roman mocks ,the hexadipole ($\ell = 4$) does not add significant information. Let us point out some crucial aspects of introducing nonlinear RSD. In the non-linear mapping from real to redshift space there are products of fields evaluated at the same location. These are of the type $\delta_g \boldsymbol{u}_\parallel, \boldsymbol{u}_\parallel^2, \delta_g \boldsymbol{u}_\parallel^2, \boldsymbol{u}_\parallel^3$. As we have seen, such contact terms lead to uncontrolled UV behavior of the theory and thus counterterms must be introduced. Crucially, compared to what happened for the higher derivative term in the bias expansion, some RSD counterterms are not degenerate with the dark matter counterterms as they have a different momentum structure. This can be understood intuitively since the mapping to redshift space breaks spatial isotropy (with respect to the $\hat{z}$ axis) identifying a preferred direction. Thus, the redshift space mapping will introduce terms that cannot be proportional to k^2 but rather to $(\boldsymbol{k}\hat{n})^2 = (k\mu)^2$. In particular, the lowest order that contributes to the quadrupole of the power spectrum has to be of the form [187, 303, 304]:

$$c_2 f\mu^2 k^2 \tag{398}$$

Crucially, non-linear effects such as the Finger-of-God (FoG), i.e., the apparent radial elongation of galaxy clusters in redshift space caused by the line-of-sight velocity dispersion of galaxies within virialized structures, reduce the range of scales over which we can trust our perturbative model. To ameliorate this issue, one then has several possibilities. FoG can be phenomenologically approximated by a Gaussian or Lorentzian damping factor of the form

$$e^{-(k\mu\sigma_v^2)^2/2}, \qquad \frac{1}{1+\frac{1}{2}(k\mu\sigma_v^2)^2}\,, \tag{399}$$

that multiplies the power spectrum model or one can introduce a next-to-leading-order counterterm of the form $c_4\mu^4k^4f^2(b_1+f\mu^2)P_{11}$ to the 1-loop model. Note that the RSD counterterm in Equation(398) can be seen as the leading-order contribution of the FoG damping factor.

However, in what follows, for the precision of the Roman analyses, and for the range of scales considered, we will neglect higher-order contributions from FoG. We will do this for the following reason.The impact of FoG effects is strictly correlated to the zero-crossing of the quadrupole of the power spectrum [305]. We noticed that the quadrupole of the Roman mocks remains well above zero for the range of scales considered in the analysis; this seems to suggest that non-perturbative FoG effects are not so strong for the galaxy sample analyzed and a perturbative approach can be trusted.

Another ingredient that must be included in our power spectrum model in redshift space is the contributions from large scale bulk flows. In previous sections, we have derived the IR resummation formula that accounts for this effect in real space. In redshift space, the formula for the damping factor of

the wiggly power spectrum gets slightly modified [235, 306]:

$$\Sigma^2_{RSD} = (1 + f\mu^2(2 + f))^2\Sigma^2 + f^2\mu^2(\mu^2 - 1)\delta\Sigma^2\,, \tag{400}$$

where Σ^2 is the usual real space damping factor and $\delta\Sigma^2$ is given by

$$\delta\Sigma^2 = \int^{\Lambda} \frac{1}{2\pi^2} P_{11}(q) j_2(q\ell_{BAO})\,. \tag{401}$$

In the actual measurement of the galaxy power spectrum, we need to assume a cosmological model to convert galaxy's position from angular and redshift coordinates into 3D comoving cartesian coordinates. However, in general it is possible that the fiducial model is different than the true cosmology. This can lead to a distortion in the observed galaxy power spectrum known as the Alcock-Paczynski (AP) effect[307]. To account for this issue we have to rewrite the true Fourier modes (denoted with ′) in terms of the fiducial one. This can be easily obtained from the Jacobian of the coordinate transformation between the real parallel and perpendicular distances $d_\parallel, d_\perp$ and the fiducial ones. We have

$$k'_\perp = k_\perp/\alpha_\perp\,, \qquad k'_\parallel = k_\parallel/\alpha_\parallel\,, \tag{402}$$

where

$$\alpha_\perp = \frac{D_A(z)H_0}{D_{A,\rm fid}(z)H_{0,\rm fid}}\,, \qquad \alpha_\parallel = \frac{H_{\rm fid}(z)H_0}{H(z)H_{0,\rm fid}}\,, \tag{403}$$

since distances are measured in h/Mpc. So that, defining $F \equiv \alpha_\parallel/\alpha_\perp$, we obtain

$$\begin{aligned} k' &= \frac{k}{\alpha_\perp}\sqrt{1 + \mu^2\left(\frac{1}{F} - 1\right)}\,, \\ \mu' &= \frac{\mu}{F}\frac{1}{\sqrt{1 + \mu^2\left(\frac{1}{F} - 1\right)}}\,. \end{aligned} \tag{404}$$

The change of coordinates also rescales the overall multipoles by the determinant of the Jacobian i.e., $\frac{1}{\alpha_\perp^2\alpha_\parallel}$. Note that, while in principle, since we are analyzing galaxy mock catalogs, the true underlying cosmology is known, the AP effect must be included in any case to consistently account for the different values of the cosmological parameters explored during the MCMC. The final power spectrum multipoles are then given by

$$P_{\ell,\rm model}(k,z) = \frac{2\ell+1}{2\alpha_\perp^2\alpha_\parallel}\int d\mu L_\ell(\mu)[P_{g,RSD}(k'(k),\mu'(\mu),z) + c^\ell f^{\ell/2}\mu^\ell k^2 P_{11}(k'(k),z) + P_{\rm shot}]\,, \tag{405}$$

where $P_{\rm shot}$ is the shot noise term. Note that even if $\frac{1}{\bar{n}}$ is subtracted from the monopole during the measurement, it is important to keep the shot noise term in the model to keep track of deviations from the Poissonian limit, which are expected on general grounds. In what follows we divide the total light-cone in redshift bins, evaluating our model at the effective redshift of that specific slice $z_{\rm eff} = \frac{z_{\rm min}+z_{\rm max}}{2}$.[27] As discussed extensively in the first part of the thesis, this is a simplification in the modeling as each galaxy is located in principle at a different peculiar redshift. In fact, the effective redshift approximation compresses the full 3D clustering into a simpler and potentially less informative 2D one. However, we will

[27] In real survey data analysis, the effective redshift of the bin is usually obtained by a weighted sum of each peculiar redshift, for example $z_{\rm eff} = \frac{\sum_i^{Ngal.} w_{FKP,i}w_{c,i}z_i}{\sum_i^{Ngal.} w_{FKP,i}w_{c,i}}$ (see e.g.,[308]) where w_{FKP} and w_c are respectively FKP and compensated weights, which will be discussed later in the main text.

employ the effective redshift approximation for the following reasons. The full 3D clustering is impossible to be correctly analyzed in Fourier space since, as we have seen, if we drop the flat sky assumption the mapping between real to redshift space introduces mode coupling effects that make the power spectrum itself ill-defined. One way to approximately account for this effect is the so-called unequal time formalism introduced in the first part of this thesis. However, as we have seen, at leading order unequal time effects are completely negligible when we are analyzing a single tracer of the underlying dark matter field. Moreover, from the 3D analysis in configuration space, we showed that radial effects are proportional to the width of the bin. In what follows we adopt redshift slices of width $\Delta z = 0.2$, so that radial effects will be mitigated.

The final observational effect that we must include in our analysis is the window function, which accounts for the angular and radial mask of the survey. This is done directly in Fourier space in order to avoid multiple Fourier transforms of the power spectrum model, which can be numerically unstable for very high k due to the presence of the counterterms. In configuration space, the windowed multipoles of the correlation function are simply obtained by $\xi_{\ell,W}(s) = Q_{\ell,\ell'}\xi_{\ell'}(s)$, where $Q_{\ell,\ell'} = C_{\ell,\ell',\ell''}Q_{\ell''}(s)$, $Q_\ell(s)$ are the window function multipoles and:

$$C_{0,\ell',\ell''} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{5} \end{bmatrix}_{\ell',\ell''}, \quad C_{2,\ell',\ell''} = \begin{bmatrix} 0 & 1 \\ 1 & \frac{2}{7} \end{bmatrix}_{\ell',\ell''}, \tag{406}$$

where repeated indices are summed over. Going to Fourier space, we have that:

$$\begin{aligned} P_{\ell,W}(k) &= 4\pi(-i)^\ell \int ds s^2 j_\ell(ks)\xi_{\ell,W}(s) \\ &= \frac{2}{\pi}(-i)^\ell i^\ell \int ds s^2 j_\ell(ks) \int dk' k'^2 j_{\ell'}(k's) Q_{\ell,\ell'}(s) P_{\ell'}(k') \,, \end{aligned} \tag{407}$$

now defining

$$W_{\ell,\ell'}(k,k') = \frac{2}{\pi}(-i)^\ell i^\ell k'^2 \int ds s^2 j_\ell(ks) Q_{\ell,\ell'}(s) j_{\ell'}(k's) \,, \tag{408}$$

we have the convolution.

$$P_{\ell,W}(k) = \int dk W_{\ell,\ell'}(k,k') P'_\ell(k') \,. \tag{409}$$

Now let us see how to estimate $Q_\ell(s)$ directly in Fourier space, using the same estimator as the power spectrum [309, 310]. To do so, firstly, let us introduce the estimator for the power spectrum multipoles often referred to in the literature as the "Yamamoto estimator" [311]:

$$P_\ell = \frac{2\ell+1}{N_P} \int \frac{d\Omega_k}{4\pi} \left[\int d^3\mathbf{r}_1 \, F(\mathbf{r}_1) e^{i\mathbf{k}\cdot\mathbf{r}_1} \int d^3\mathbf{r}_2 \, F(\mathbf{r}_2) e^{-i\mathbf{k}\cdot\mathbf{r}_2} L_\ell(\hat{\mathbf{k}}\cdot\hat{\mathbf{r}}_2) - P_\ell^{\rm noise}(\mathbf{k}) \right] , \tag{410}$$

where Ω_k represents the solid angle in Fourier space, and L_ℓ is the Legendre polynomial of order ℓ. The weighted density field $F(\mathbf{r})$ is defined as

$$F(\mathbf{r}) = w_{\rm fkp}(\mathbf{r}) \left[n_g(\mathbf{r}) - \alpha n_s(\mathbf{r}) \right] , \tag{411}$$

where n_g and n_s are the number densities for the galaxy catalog and random catalog respectively, and α' is the ratio of the number of real galaxies to random galaxies $\alpha = N_{\rm galaxies}/N_{\rm random}$.[28] The normalization

[28] To mitigate shot-noise contributions, the random catalog should contain way more points than the galaxy catalog, with $\alpha = 10$ a

is given by

$$N_P = \int d^3\mathbf{r}\,[n_g(\mathbf{r})w_{\rm fkp}(\mathbf{r})]^2\,. \tag{412}$$

The shot noise $P_\ell^{\rm noise}$ is

$$P_\ell^{\rm noise}(\mathbf{k}) = (1+\alpha)\int d^3\mathbf{r}\,n_g(\mathbf{r})w_{\rm fkp}^2(\mathbf{r})L_\ell(\hat{\mathbf{k}}\cdot\hat{\mathbf{r}})\,. \tag{413}$$

The FKP (Feldman-Kaiser-Peacock) weights, first derived in [312], minimize the variance of the estimator at a desired power spectrum value. Denoted as $w_{\rm FKP}$, these weights are given by

$$w_{\rm FKP}(\mathbf{r}) = \frac{1}{1+n_g(\mathbf{r})P_0}\,, \tag{414}$$

where P_0 is the power spectrum amplitude in units of $h^{-3}{\rm Mpc}^3$ where the estimator is optimized. For typical galaxy survey analyses, a value of order $P_0 = 10^4 h^{-3}{\rm Mpc}^3$ is usually assumed.[29] Then ,the expectation value of the power spectrum FKP estimator can be rewritten as

$$\left\langle\hat{P}_\ell(k)\right\rangle = \frac{(2\ell+1)}{N_P}(-i)^\ell\int ds\,s^2\,j_\ell(ks)\int d\Omega_s\int d^3r_1\,\Theta(r_1)\Theta(r_1+s)\,\bar{n}_w(r_1)\bar{n}_w(r_1+s)\,\mathcal{L}_\ell(\mu)\,\xi(\mathbf{s},\mathbf{r}_1)\,, \tag{415}$$

where Θ is the Heaviside step function which is 1 if the galaxy is inside the survey mask and 0 otherwise. $\bar{n}_w(\boldsymbol{r}) = w_{\rm FKP}(\boldsymbol{r})\bar{n}(\boldsymbol{r})$ is the weighted mean galaxy density. Let us assume that we are calculating the power spectrum of a random field whose distribution is approaching Poisson i.e., $\xi(\mathbf{s},\mathbf{r}_1) = 1$ and which has the same mask as the survey. We can write

$$\left\langle\hat{P}_{{\rm rand},\ell}(k)\right\rangle = 4\pi(-i)^\ell\int_0^\infty ds\,s^2\,j_\ell(ks)\,Q_\ell(s)\,, \tag{416}$$

where the multipoles of the window function are

$$Q_\ell(s) \equiv \frac{(2\ell+1)}{N_P}\int\frac{d\Omega_s}{4\pi}\int d^3r_1\,\Theta(r_1)\Theta(r_1+s)\,\bar{n}_w(r_1)\bar{n}_w(r_1+s)\,\mathcal{L}_\ell(\mu)\,. \tag{417}$$

We now invert Equation (416) using the spherical Bessel closure relation:

$$\int_0^\infty dk\,\frac{(ks)^2}{2\pi^2}j_\ell(ks)j_\ell(ks') = \frac{1}{4\pi}\delta_D(s-s')\,. \tag{418}$$

We have:

$$\begin{aligned}\int_0^\infty dk\,\frac{k^2}{2\pi^2}\,i^\ell\,j_\ell ks')\left\langle\hat{P}_{{\rm rand},\ell}(k)\right\rangle &= 4\pi\int_0^\infty ds\,s^2\,Q_\ell(s)\left[\int_0^\infty dk\,\frac{k^2}{2\pi^2}\,i^\ell(-i)^\ell\,j_\ell(ks)\,j_\ell(ks')\right]\\ &= 4\pi\int_0^\infty ds\;Q_\ell(s)\left[\frac{1}{4\pi}\delta_D(s-s')\right]\\ &= Q_\ell(s')\,.\end{aligned} \tag{419}$$

---

reasonable choice in terms of precision and efficiency.

[29] In real survey analysis, on top of FKP weights, additional completeness weights (w_c) are often included in the measurements. These weights include the systematics and observational effects which are specifics of a particular survey and how they are included in the measurement depends on the specific experiment. In our case, since we are analyzing galaxy mocks we will neglect those weights.

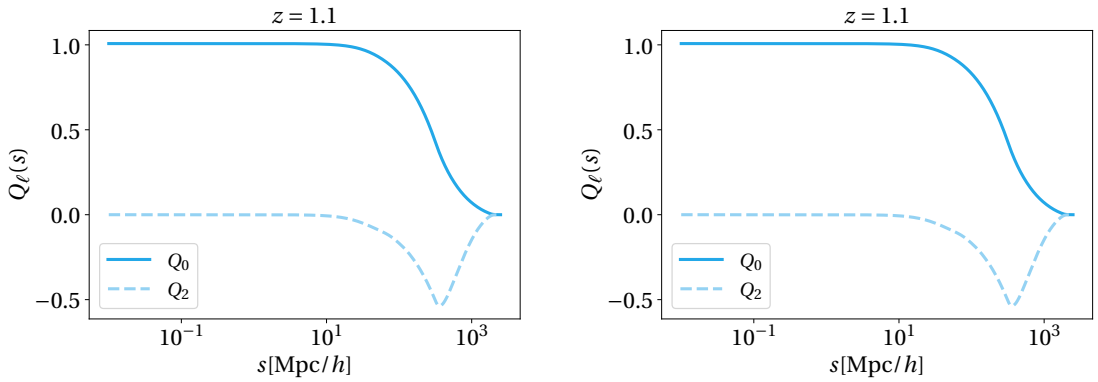

Figure 46: Q_0, Q_2 for the first two slices of the Roman galaxy mocks

Relabeling $s' \to s$, we obtain:

$$Q_\ell(s) = i^\ell \int_0^\infty \frac{dk}{2\pi^2}\, k^2 \left\langle P_{\mathrm{rand},\ell}(k) \right\rangle j_\ell(ks)\,. \tag{420}$$

Therefore $Q_\ell(s)$ can be obtained by Fourier transforming the power spectrum of a random distribution with the same angular and radial distributions as the data. Such a catalog of random objects is also required for the estimation of the power spectrum. Note that the window function multipoles in Fourier space must have the same normalization as the power spectrum so that:

$$Q_\ell(k) = \alpha \left\langle P_{\mathrm{rand},\ell}(k) \right\rangle \tag{421}$$

As an example, the $Q_\ell(s)$ obtained following the outlined procedure is shown in Fig.46 for the first redshift slice. Lastly, when constructing the galaxy overdensity field for power spectrum estimation, one must specify the mean galaxy density. In practice, this is taken to be the survey's mean density, which ensures that the measured power spectrum vanishes at the survey scale—a correction known as the integral constraint. To prevent biases in cosmological parameter inference, the same constraint must also be imposed on the theoretical power spectrum model. While the integral constraint formally affects only the $k = 0$ mode, the survey window function couples neighboring modes, causing its effect to extend to finite k. The integral constraints-corrected power spectrum model reads [308, 309, 313, 314]:

$$P_\ell(k) = \int dk' k'^2 W_{\ell,\ell'}(k,k') P_{\ell'}(k') - \frac{Q_\ell(k)}{Q_0(0)} \int dk' k'^2 W_{\ell,\ell'}(0,k') P_{\ell'}(k') \tag{422}$$

However, in the analysis presented here, we used $k_{\mathrm{min}} = 0.01h/\mathrm{Mpc}$ and we have checked that for this range of scales, the effect of the integral constraint is completely negligible.

B Data analysis

Our dataset is represented by the low redshift bins of the Roman High Latitude Spectroscopic Survey Hα mocks on the light-cone generated by applying Galacticus[315] to the simulated dark matter halos of the UNIT simulations [316]. The fiducial cosmology of the simulation is $\{\omega_b, \omega_{\mathrm{cdm}}, h, \ln(10^{10}A_s), n_s\} = \{0.0212, 0.1205, 0.6774, 3.0253, 0.9667\}$. See [317, 318] for a more detailed description of the galaxy mocks. We select all the galaxies with dust-attenuated Hα flux higher than 10^{-16} erg s^{-1} cm^{-2} which corresponds to the 6.5σ nominal depth for the Roman GRS. We then randomly sub-select 55% of the total number of galaxies to account for the purity of the measurements and selection effects. The galaxies are selected between $1 < z < 1.4$ and divided into two redshift bins of width $\Delta z = 0.2$. The final total number of

observed galaxies is $\sim 4.5 \times 10^6$ and they cover a sky area of ~ 2000 deg^2. For the measurement of the power spectrum multipoles, we use the public code `nbodykit`[319]. We used 512^3 as the mesh grid size, and the Triangular Shaped Cloud (TSC) algorithm for the mass assignment scheme. We corrected for aliasing effects with the interlacing technique [320, 321] and divided the obtained density field (in Fourier space) by the window function introduced during the mass assignment procedure [320]. We measured the power spectra in k-bins of width $\Delta k = 0.01h/\text{Mpc}$ from $0.01h/\text{Mpc} < k < 0.3h/\text{Mpc}$, but in the analysis we consider only bins up to $k_{\text{max}} = 0.2h/\text{Mpc}$ for both redshift slices. While in real survey data analysis a more rigorous procedure must be used to determine the maximum wavenumber over which our perturbative model breaks down, in the simple analysis presented here, we adopt this choice in order to be similar to the one in [322, 323].[30]

Finally, to assess the uncertainty of the measurements we need to estimate the covariance matrix. To do so without relying on the Gaussian approximation, for each redshift slice we used a set of 1000 EZMocks calibrated to resemble the measurements of the Roman mocks. However, since those EZmocks were firstly produced in [317], they were calibrated to mimic an earlier version of the Galacticus Roman mocks. To overcome this, we do the following. In each redshift slice we measure the multipoles of the power spectrum from all the EZmocks, find their mean $\bar{P}_\ell$ and the 1σ error. We then multiply the fractional error $\bar{P}_\ell/\sigma_\ell$ by the measurements from the Galacticus mock obtaining the rescaled uncertainty $f_\ell = \bar{P}_\ell/\sigma_\ell$. Then we estimate the covariance matrix in the usual way:

$$\tilde{C}_{ij,\ell,\ell'} = \frac{1}{N_s - 1} \sum_n^{N_s} \left[P^n_{i,\ell} - \bar{P}_{i,\ell} \right] \left[P^n_{j,\ell'} - \bar{P}_{j,\ell'} \right] , \tag{423}$$

from this covariance we calculate the correlation matrix

$$R_{ij,\ell,\ell'} = \frac{C_{ij,\ell,\ell'}}{\sqrt{C_{ii,\ell\ell} C_{jj,\ell'\ell'}}} , \tag{424}$$

and multiply the normalized correlation matrix for the fractional difference to obtain the final covariance matrix

$$C_{ij,\ell,\ell'} = f_{i,\ell} f_{j,\ell} R_{ij,\ell\ell'} . \tag{425}$$

This procedure can preserve the structure in the correlation matrix from the EZmocks, and have the correct or scaled uncertainty in the diagonal terms for the updated Galacticus measurements.

The measurement of the multipoles of the power spectrum and the associated error-bars extracted from the diagonal of the covariance matrix of Equation (425) is shown in Fig. 47 for both redshift slices. Equipped with the theory model, the data, and the covariance matrix we can run an MCMC analysis varying the following parameters:

$$\{\omega_b, \omega_{\text{cdm}}, h, \ln(10^{10} A_s), b_{1,i}, b_{2,i}, b_{\mathcal{G}_2,i}, b_{\Gamma_3,i}, c_{0,i}, c_{2,i}, P_{\text{shot}}\} , \tag{426}$$

where $i = 1, 2$ is the number of redshift slices. We decided to fix the spectral index of the primordial power spectrum n_S to its fiducial value since, while CMB data can measure this parameter with exquisite precision, LSS data are not so sensitive to this early-Universe parameter. In previous full-shape analyses, n_s has been either fixed or a prior from Planck data has been put on it (see e.g.,[299, 300, 323, 324]). Furthermore, as customary in full shape analyses, we put a BBN prior on the baryon density parameter

[30] Since the dominant next to leading order (NLO) effects are NLO FoG contributions, one possibility to determine the maximum wavenumber is to pick k_{max} as the maximum wavenumber below which one gets unbiased result without including the NLO FoG counterterm in the model.

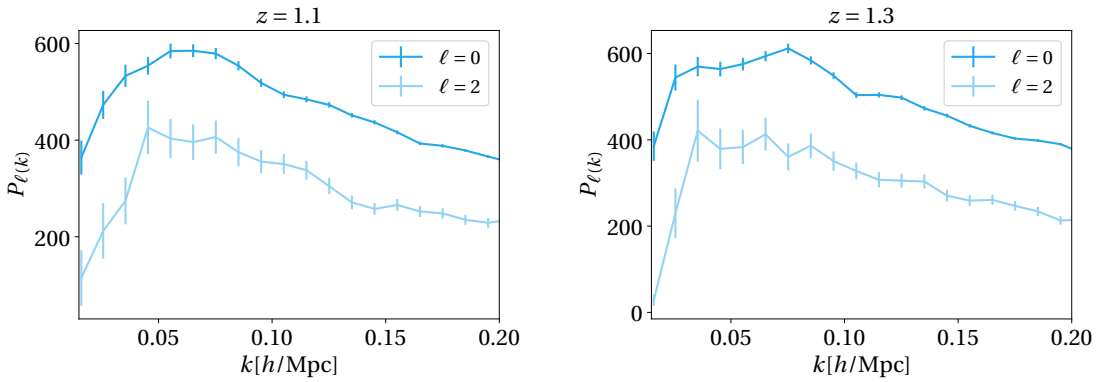

Figure 47: Measured monopole and quadrupole for the range of scales considered in the analysis in the two low-redshift bins of the Roman mocks. Error bars from the diagonal of Equation425 are also showed.

based on [325]

$$\mathcal{P}(\omega_{\rm b}) = \mathcal{N}(\omega_{\rm fid}, 0.00055)\,. \tag{427}$$

Previous analyses have also included additional parameters to the power spectrum noise model. Terms such as $\epsilon_1 k^2$ and $\epsilon_2 \mu^2 k^2$ have been introduced to model the first order higher-derivative contributions to the noise model from the monopole and quadrupole respectively [227, 300]. However, we find that, for the two slices of the Roman mocks analyzed here, the inclusion of such terms does not change sensibly the results. We used flat priors for the cosmological parameters. We will adopt the following choice of priors for the nuisance parameters.

$$\mathcal{P}(b_{1,i}) = \mathcal{N}(1.5, 2), \quad \mathcal{P}(b_{2,i}) = \mathcal{P}(b_{\mathcal{G}_2,i}) = \mathcal{P}(b_{\Gamma_r,i}) = \mathcal{N}(0, 3), \quad \mathcal{P}(c_{0,i}) = \mathcal{P}(c_{2,i}) = \mathcal{N}(0, 30)\,, \tag{428}$$

where $\mathcal{N}(a, b)$ is a Gaussian of mean a and $\sigma = b$. To run the MCMC chains, we make use of the public `PyBird` [226] code associated with the `MontePython` sampler. `PyBird` is a Python-based implementation of the 1-loop power spectrum model in redshift space outlined in this thesis. The only difference is the implementation of the IR resummation scheme, which is done following the approach of [152, 153] which is slightly different with respect to the one explained here. However, the two approaches have extensively been checked in the literature to give the same results (see e.g.,[225, 326]). In Fig.48 we present the posteriors of the cosmological parameters combining the two redshift slices. We see that the posteriors are unbiased and that the late time parameters H_0, σ_8 and $\Omega_{\rm m}$ can be measured with $\sim 2.5\%, 4\%, 7\%$ precision. While this is a first forecast of the capabilities of the Full Shape analysis of galaxy datasets obtained with the Roman experiment, a more detailed and updated analysis of Roman galaxy mocks is the scope of ongoing work.

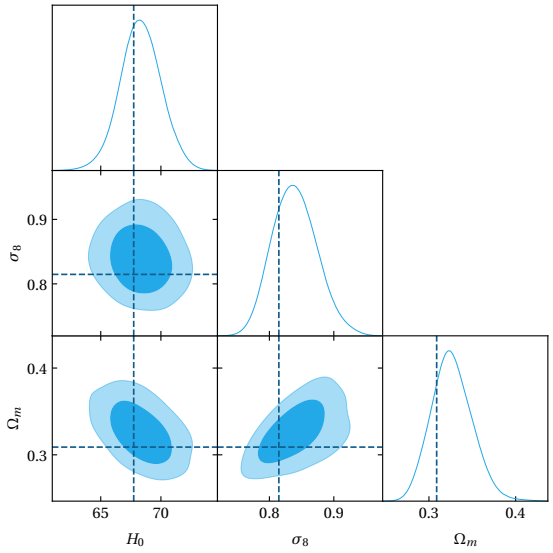

Figure 48: Posteriors for the cosmological parameters of the Roman galaxy mocks analysis with $k_{\rm max} = 0.2h/{\rm Mpc}$

IX Final remarks

The current decade is the golden age of galaxy redshift surveys. Experiments like DESI, Euclid, SPHEREx are currently taking data from millions of galaxies across huge portions of the sky, while surveys like the LSST and the Nancy Grace Roman will be operating in the next couple of years. This experimental effort will deliver the most precise map of the cosmic web in the history of humankind. However, such an observational cornerstone will need a commensurate theoretical effort on the modeling side in order to interpret this huge amount of data in the best possible way, shedding new light on open questions in cosmology. With the level of precision in the measurements guaranteed by modern galaxy surveys, we must question the validity of every approximation that was valid for past experiments.

The first part of this thesis is focused on some of these approximations: the plane parallel and the flat sky approximation. In a nutshell, these simplifications in the modeling of galaxy statistics compress the 3-dimensional clustering into a simpler 2D observable. Considering only very large scales when the galaxy overdensity field can be modeled linearly, we tackle this issue both in configuration space and in Fourier space. In configuration space, wide angle models which drop both the plane parallel and the flat sky approximation are well established in literature so that we can compare directly the accuracy of the 3D model and the flat sky/plane parallel one. Based on theoretical forecasts, we find that these simplifications can sensibly bias measurements of the growth of structure and of the level of local Primordial non-Gaussianity of current and upcoming surveys. Interestingly, we find that the majority of the accuracy loss comes from the fact that the flat sky model fails to model the radial part of the correlators. Namely, the flat sky assumes that all galaxies in a given redshift slice are located at the same radial position, while in reality each galaxy is located at a different peculiar redshift. Thus, we introduced the Radial 3D model, which, while being way simpler than the full wide angle one, can preserve the radial information and the 3-dimensionality of the clustering. We verified that this simpler model can achieve accuracy levels comparable to those of the wide-angle model. In Fourier space, a way to model such radial correlators has been introduced very recently. Such "unequal time" correlations arise as simple additive contributions to the standard "equal time" power spectrum and are proportional to powers of the radial distance between the galaxies in the pairs, in a perturbative way. In this thesis, we presented the extension of the unequal time formalism to redshift space and

with the inclusion of local Primordial non-Gaussianity contributions. We find that, at first order, such contributions are negligible when considering correlators of the same dark matter tracer. On the other hand, when cross-correlating two different biased tracers, such unequal time contributions are sizable and can be, in principle, degenerate with $f_{\rm NL}$ local, due to the similar scale-dependence.

In the second part of this thesis, we focused on the mildly non-linear regime of the galaxy overdensity field. Firstly, we review schematically the formalism that we adopted to model non-linearities: the so-called Effective Field Theory of Large Scale Structure. Then we investigated the relations between cosmological analyses based on conventional summary statistics and field level inference methods. Perturbation theory arguments predict that if the galaxy field is forward modeled up to m-order then the information content in the field level posterior (once marginalized over initial conditions) must be comparable to the information of the summary statistics that can be constructed at that given order in PT. We checked this prediction explicitly, going to third order in PT. In particular, we studied a simple setup with dark matter halos in real space, with the amplitude of the linear density field σ_8 as the only free cosmological parameter. We showed that Eulerian perturbation theory is adequate for describing this system on large scales, computed the leading n-point functions and performed a joint power spectrum, bispectrum and trispectrum analysis using synthetic data. Beyond the bispectrum which is crucial for breaking the degeneracy between A and the linear bias, we find that the addition of the trispectrum reduces the error on the amplitude by only $20-30\%$. Our results for the joint analysis are in good agreement with recent field-level analyses in the same setup. This implies that the field-level inference on large scales does not get significant information from large displacements beyond those in Eulerian kernels or higher-order n-point functions beyond the trispectrum. We provide further evidence for this, showing that the dependence of the error bars on the maximum wavenumbers used in the analysis is the same in the two approaches.

Then, we investigated another alternative to the standard statistics: the marked power spectrum. This unconventional statistic, which is essentially the 2-point function of two galaxy overdensities in Fourier space that have been weighted by some non-linear function, has emerged as a promising tool for extracting cosmological information from the large-scale structure of the Universe. Here, we presented the first comprehensive analytical study of the marked power spectrum's sensitivity to primordial non-Gaussianity (PNG) of the non-local type. We extend previous effective field theory frameworks to incorporate PNG, developing a complete theoretical model that we have validated against the Quijote simulation suite. With Fisher forecasts, we compared the constraining power of the marked power spectrum against traditional approaches combining the power spectrum and bispectrum. We explored different choices of mark parameters to evaluate their impact on parameter constraints, particularly focusing on equilateral and orthogonal PNG as well as neutrino masses. Our analysis shows that while marking up underdense regions yields optimal constraints in the low shot-noise regime, the marked power spectrum's performance for discrete tracers with BOSS-like number densities does not surpass that of standard P+B analysis at mildly non-linear scales ($k \lesssim 0.25h/\mathrm{Mpc}$). However, the marked approach offers several practical advantages, including simpler estimation procedures and potentially more manageable systematic effects. Our theoretical framework reveals how the marked power spectrum incorporates higher-order correlation information through terms resembling tree-level bispectra and power spectrum convolutions.

Lastly, as a simple application of the 1-loop power spectrum model, we analyze realistic galaxy mocks on the light-cone that mimic the sample of galaxies that will be observed by the *Nancy Grace Roman Space Telescope*. After measuring the power spectrum multipoles, including observational systematics such as the Alcock-Paczyński effect and window function, we showed that the EFT model can recover the fiducial cosmology at $k_{\rm max} = 0.2h/\mathrm{Mpc}$ with a minimal choice of free parameters. Furthermore, we forecasted the error bars for the late time parameters $H_0, \sigma_8, \Omega_{\rm m}$ from the full shape analyses of galaxy datasets obtained with the Roman Telescope, obtaining promising results.

In conclusion, at its core, this thesis has addressed two closely related aspects of large-scale structure

modeling. In the first part, we investigated the validity of several commonly adopted approximations in the literature, showing that some of them can bias cosmological analyses in certain scenarios, and we developed new models to overcome these limitations. In the second part, we examined the effectiveness of recently proposed methods that aim to extract more precise cosmological information from galaxy survey data compared to conventional analyses, and we argued that, at least in the cases studied here, the gain in precision is not substantial. A straightforward continuation of the work presented here is to test the validity of our conclusions—mostly obtained through theoretical forecasts or N-body simulations—against the analysis of forthcoming high-precision galaxy survey data. As extensively discussed, such datasets will soon become available, making the timing of the results of this thesis particularly appropriate.

---

[1] N. Aghanim *et al.* (Planck), Planck 2018 results. VI. Cosmological parameters, Astron. Astrophys. **641**, A6 (2020), [Erratum: Astron.Astrophys. 652, C4 (2021)], arXiv:1807.06209 [astro-ph.CO].

[2] A. G. Riess *et al.* (Supernova Search Team), Observational evidence from supernovae for an accelerating universe and a cosmological constant, Astron. J. **116**, 1009 (1998), arXiv:astro-ph/9805201.

[3] S. Perlmutter *et al.* (Supernova Cosmology Project), Measurements of Ω and Λ from 42 High Redshift Supernovae, Astrophys. J. **517**, 565 (1999), arXiv:astro-ph/9812133.

[4] A. A. Starobinsky, Spectrum of relict gravitational radiation and the early state of the universe, JETP Lett. **30**, 682 (1979).

[5] A. A. Starobinsky, A New Type of Isotropic Cosmological Models Without Singularity, Phys. Lett. B **91**, 99 (1980).

[6] A. H. Guth, The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems, Phys. Rev. D **23**, 347 (1981).

[7] A. D. Linde, A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems, Phys. Lett. B **108**, 389 (1982).

[8] G. F. Smoot *et al.* (COBE), Structure in the COBE differential microwave radiometer first year maps, Astrophys. J. Lett. **396**, L1 (1992).

[9] Y. Akrami *et al.* (Planck), Planck 2018 results. IX. Constraints on primordial non-Gaussianity, Astron. Astrophys. **641**, A9 (2020), arXiv:1905.05697 [astro-ph.CO].

[10] S. Galli, M. Martinelli, A. Melchiorri, L. Pagano, B. D. Sherwin, and D. N. Spergel, Constraining Fundamental Physics with Future CMB Experiments, Phys. Rev. D **82**, 123504 (2010), arXiv:1005.3808 [astro-ph.CO].

[11] U. Seljak and M. Zaldarriaga, Signature of gravity waves in polarization of the microwave background, Phys. Rev. Lett. **78**, 2054 (1997), arXiv:astro-ph/9609169.

[12] M. Kamionkowski, A. Kosowsky, and A. Stebbins, A Probe of primordial gravity waves and vorticity, Phys. Rev. Lett. **78**, 2058 (1997), arXiv:astro-ph/9609132.

[13] E. Allys *et al.* (LiteBIRD), Probing Cosmic Inflation with the LiteBIRD Cosmic Microwave Background Polarization Survey, PTEP **2023**, 042F01 (2023), arXiv:2202.02773 [astro-ph.IM].

[14] K. Abazajian *et al.* (CMB-S4), CMB-S4: Forecasting Constraints on Primordial Gravitational Waves, Astrophys. J. **926**, 54 (2022), arXiv:2008.12619 [astro-ph.CO].

[15] K. S. Dawson *et al.* (BOSS), The Baryon Oscillation Spectroscopic Survey of SDSS-III, Astron. J. **145**, 10 (2013), arXiv:1208.0022 [astro-ph.CO].

[16] A. Aghamousa *et al.* (DESI), The DESI Experiment Part I: Science,Targeting, and Survey Design, (2016), arXiv:1611.00036 [astro-ph.IM].

[17] Y. Mellier *et al.* (Euclid), Euclid. I. Overview of the Euclid mission, (2024), arXiv:2405.13491 [astro-ph.CO].

[18] O. Doré *et al.*, Cosmology with the SPHEREX All-Sky Spectral Survey, (2014), arXiv:1412.4872 [astro-ph.CO].

[19] A. Abate *et al.* (LSST Dark Energy Science), Large Synoptic Survey Telescope: Dark Energy Science Collaboration, (2012), arXiv:1211.0310 [astro-ph.CO].

[20] Y. Wang *et al.*, The High Latitude Spectroscopic Survey on the Nancy Grace Roman Space Telescope, Astrophys. J. **928**, 1 (2022), arXiv:2110.01829 [astro-ph.CO].

[21] D. J. Schlegel *et al.*, The MegaMapper: A Stage-5 Spectroscopic Instrument Concept for the Study of Inflation

and Dark Energy, (2022), arXiv:2209.04322 [astro-ph.IM].
[22] S. Dodelson and F. Schmidt, *Modern Cosmology* (Academic Press, 2020).
[23] P. J. Peebles, *The Large-Scale Structure of the Universe* (Princeton University Press, 1980).
[24] A. J. S. Hamilton, Linear redshift distortions: A Review, in *Ringberg Workshop on Large Scale Structure* (1997) arXiv:astro-ph/9708102.
[25] F. Bernardeau, S. Colombi, E. Gaztanaga, and R. Scoccimarro, Large scale structure of the universe and cosmological perturbation theory, Phys. Rept. **367**, 1 (2002), arXiv:astro-ph/0112551.
[26] N. Kaiser, On the Spatial correlations of Abell clusters, Astrophys. J. Lett. **284**, L9 (1984).
[27] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, The Statistics of Peaks of Gaussian Random Fields, Astrophys. J. **304**, 15 (1986).
[28] H. J. Mo and S. D. M. White, An Analytic model for the spatial clustering of dark matter halos, Mon. Not. Roy. Astron. Soc. **282**, 347 (1996), arXiv:astro-ph/9512127.
[29] V. Desjacques, D. Jeong, and F. Schmidt, Large-Scale Galaxy Bias, Phys. Rept. **733**, 1 (2018), arXiv:1611.09787 [astro-ph.CO].
[30] N. Kaiser, Clustering in real space and in redshift space, Mon. Not. Roy. Astron. Soc. **227**, 1 (1987).
[31] D. S. Salopek and J. R. Bond, Stochastic inflation and nonlinear gravity, Phys. Rev. D **43**, 1005 (1991).
[32] E. Komatsu and D. N. Spergel, Acoustic signatures in the primary microwave background bispectrum, Phys. Rev. D **63**, 063002 (2001), arXiv:astro-ph/0005036.
[33] L. Verde, R. Jimenez, M. Kamionkowski, and S. Matarrese, Tests for primordial nonGaussianity, Mon. Not. Roy. Astron. Soc. **325**, 412 (2001), arXiv:astro-ph/0011180.
[34] N. Bartolo, E. Komatsu, S. Matarrese, and A. Riotto, Non-Gaussianity from inflation: Theory and observations, Phys. Rept. **402**, 103 (2004), arXiv:astro-ph/0406398.
[35] A. Gangui, F. Lucchin, S. Matarrese, and S. Mollerach, The Three point correlation function of the cosmic microwave background in inflationary models, Astrophys. J. **430**, 447 (1994), arXiv:astro-ph/9312033.
[36] P. Creminelli and M. Zaldarriaga, Single field consistency relation for the 3-point function, JCAP **10**, 006, arXiv:astro-ph/0407059.
[37] V. Acquaviva, N. Bartolo, S. Matarrese, and A. Riotto, Second order cosmological perturbations from inflation, Nucl. Phys. B **667**, 119 (2003), arXiv:astro-ph/0209156.
[38] J. M. Maldacena, Non-Gaussian features of primordial fluctuations in single field inflationary models, JHEP **05**, 013, arXiv:astro-ph/0210603.
[39] D. Babich, P. Creminelli, and M. Zaldarriaga, The Shape of non-Gaussianities, JCAP **08**, 009, arXiv:astro-ph/0405356.
[40] N. Arkani-Hamed, P. Creminelli, S. Mukohyama, and M. Zaldarriaga, Ghost inflation, JCAP **04**, 001, arXiv:hep-th/0312100.
[41] P. Creminelli, On non-Gaussianities in single-field inflation, JCAP **10**, 003, arXiv:astro-ph/0306122.
[42] M. Alishahiha, E. Silverstein, and D. Tong, DBI in the sky, Phys. Rev. D **70**, 123505 (2004), arXiv:hep-th/0404084.
[43] S. Matarrese and L. Verde, The effect of primordial non-Gaussianity on halo bias, Astrophys. J. Lett. **677**, L77 (2008), arXiv:0801.4826 [astro-ph].
[44] S. Matarrese, L. Verde, and R. Jimenez, The Abundance of high-redshift objects as a probe of non-Gaussian initial conditions, Astrophys. J. **541**, 10 (2000), arXiv:astro-ph/0001366.
[45] N. Dalal, O. Dore, D. Huterer, and A. Shirokov, The imprints of primordial non-gaussianities on large-scale structure: scale dependent bias and abundance of virialized objects, Phys. Rev. D **77**, 123514 (2008), arXiv:0710.4560 [astro-ph].
[46] F. Schmidt and M. Kamionkowski, Halo Clustering with Non-Local Non-Gaussianity, Phys. Rev. D **82**, 103002 (2010), arXiv:1008.0638 [astro-ph.CO].
[47] T. Giannantonio, C. Porciani, J. Carron, A. Amara, and A. Pillepich, Constraining primordial non-Gaussianity with future galaxy surveys, Mon. Not. Roy. Astron. Soc. **422**, 2854 (2012), arXiv:1109.0958 [astro-ph.CO].
[48] A. Slosar, C. Hirata, U. Seljak, S. Ho, and N. Padmanabhan, Constraints on local primordial non-Gaussianity from large scale structure, JCAP **08**, 031, arXiv:0805.3580 [astro-ph].
[49] A. Barreira, On the impact of galaxy bias uncertainties on primordial non-Gaussianity constraints, JCAP **12**, 031, arXiv:2009.06622 [astro-ph.CO].
[50] A. Barreira, Can we actually constrain f_{NL} using the scale-dependent bias effect? An illustration of the impact of galaxy bias uncertainties using the BOSS DR12 galaxy power spectrum, JCAP **11**, 013, arXiv:2205.05673

[astro-ph.CO].
[51] E. Fondi, L. Verde, F. Villaescusa-Navarro, M. Baldi, W. R. Coulton, G. Jung, D. Karagiannis, M. Liguori, A. Ravenni, and B. D. Wandelt, Taming assembly bias for primordial non-Gaussianity, (2023), arXiv:2311.10088 [astro-ph.CO].
[52] N. Dalal and W. J. Percival, Estimating non-gaussian bias using counts of tracers, (2025), arXiv:2503.21024 [astro-ph.CO].
[53] M. Marinucci, V. Desjacques, and A. Benson, Non-Gaussian assembly bias from a semi-analytic galaxy formation model, Mon. Not. Roy. Astron. Soc. **524**, 325 (2023), arXiv:2303.10337 [astro-ph.CO].
[54] T. Lazeyras, A. Barreira, F. Schmidt, and V. Desjacques, Assembly bias in the local PNG halo bias and its implication for f $_{NL}$ constraints, JCAP **01**, 023, arXiv:2209.07251 [astro-ph.CO].
[55] B. A. Reid, L. Verde, K. Dolag, S. Matarrese, and L. Moscardini, Non-gaussian halo assembly bias, Journal of Cosmology and Astroparticle Physics **2010** (07), 013.
[56] R. K. Sheth, H. J. Mo, and G. Tormen, Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes, Mon. Not. Roy. Astron. Soc. **323**, 1 (2001), arXiv:astro-ph/9907024.
[57] S. Zaroubi and Y. Hoffman, Clustering in redshift space: Power spectrum, (1993), arXiv:astro-ph/9311013.
[58] A. Raccanelli and Z. Vlah, Power spectrum in the cave, (2023), arXiv:2305.16278 [astro-ph.CO].
[59] A. Raccanelli and Z. Vlah, Observed power spectrum and frequency-angular power spectrum, Phys. Rev. D **108**, 043537 (2023), arXiv:2306.00808 [astro-ph.CO].
[60] Z. Gao, A. Raccanelli, and Z. Vlah, Asymptotic connection between full- and flat-sky angular correlators, Phys. Rev. D **108**, 043503 (2023), arXiv:2306.02993 [astro-ph.CO].
[61] R. Laureijs *et al.* (EUCLID), Euclid Definition Study Report, (2011), arXiv:1110.3193 [astro-ph.CO].
[62] A. Blanchard *et al.* (Euclid), Euclid preparation: VII. Forecast validation for Euclid cosmological probes, Astron. Astrophys. **642**, A191 (2020), arXiv:1910.09273 [astro-ph.CO].
[63] A. G. Adame *et al.* (DESI), DESI 2024 VI: Cosmological Constraints from the Measurements of Baryon Acoustic Oscillations, (2024), arXiv:2404.03002 [astro-ph.CO].
[64] Y.Wang *et al.*, *in prep.*
[65] E. V. Linder, Cosmic growth history and expansion history, Phys. Rev. D **72**, 043529 (2005), arXiv:astro-ph/0507263.
[66] N. Kaiser, Clustering in real space and in redshift space, Mon. Not. Roy. Astron. Soc. **227**, 1 (1987).
[67] A. J. S. Hamilton, Linear redshift distortions: A Review, in *Ringberg Workshop on Large Scale Structure* (1997) arXiv:astro-ph/9708102.
[68] A. S. Szalay, T. Matsubara, and S. D. Landy, Redshift space distortions of the correlation function in wide angle galaxy surveys, Astrophys. J. Lett. **498**, L1 (1998), arXiv:astro-ph/9712007.
[69] P. Papai and I. Szapudi, Non-Perturbative Effects of Geometry in Wide-Angle Redshift Distortions, Mon. Not. Roy. Astron. Soc. **389**, 292 (2008), arXiv:0802.2940 [astro-ph].
[70] A. Raccanelli, L. Samushia, and W. J. Percival, Simulating Redshift-Space Distortions for Galaxy Pairs with Wide Angular Separation, Mon. Not. Roy. Astron. Soc. **409**, 1525 (2010), arXiv:1006.1652 [astro-ph.CO].
[71] A. Raccanelli, D. Bertacca, D. Jeong, M. C. Neyrinck, and A. S. Szalay, Doppler term in the galaxy two-point correlation function: wide-angle, velocity, Doppler lensing and cosmic acceleration effects, Phys. Dark Univ. **19**, 109 (2018), arXiv:1602.03186 [astro-ph.CO].
[72] L. R. Abramo and D. Bertacca, Disentangling the effects of Doppler velocity and primordial non-Gaussianity in galaxy power spectra, Phys. Rev. D **96**, 123535 (2017), arXiv:1706.01834 [astro-ph.CO].
[73] S. Andrianomena, C. Bonvin, D. Bacon, P. Bull, C. Clarkson, R. Maartens, and T. Moloi, Testing General Relativity with the Doppler magnification effect, Mon. Not. Roy. Astron. Soc. **488**, 3759 (2019), arXiv:1810.12793 [astro-ph.CO].
[74] D. Jeong, F. Schmidt, and C. M. Hirata, Large-scale clustering of galaxies in general relativity, Phys. Rev. D **85**, 023504 (2012), arXiv:1107.5427 [astro-ph.CO].
[75] D. Bertacca, R. Maartens, A. Raccanelli, and C. Clarkson, Beyond the plane-parallel and Newtonian approach: Wide-angle redshift distortions and convergence in general relativity, JCAP **10**, 025, arXiv:1205.5221 [astro-ph.CO].
[76] F. Semenzato, D. Bertacca, and A. Raccanelli, The full-sky Spherical Fourier-Bessel power spectrum in general relativity, (2024), arXiv:2406.09545 [astro-ph.CO].
[77] J. E. Bautista *et al.*, The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: measurement

of the BAO and growth rate of structure of the luminous red galaxy sample from the anisotropic correlation function between redshifts 0.6 and 1, Mon. Not. Roy. Astron. Soc. **500**, 736 (2020), arXiv:2007.08993 [astro-ph.CO].

[78] A. G. Sanchez *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. Roy. Astron. Soc. **433**, 1202 (2013), arXiv:1303.4396 [astro-ph.CO].

[79] L. Samushia, W. J. Percival, and A. Raccanelli, Interpreting large-scale redshift-space distortion measurements, Mon. Not. Roy. Astron. Soc. **420**, 2102 (2012), arXiv:1102.1014 [astro-ph.CO].

[80] M. Icaza-Lizaola *et al.* (eBOSS), The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 LRG sample: structure growth rate measurement from the anisotropic LRG correlation function in the redshift range $0.6 < z < 1.0$, Mon. Not. Roy. Astron. Soc. **492**, 4189 (2020), arXiv:1909.07742 [astro-ph.CO].

[81] A. Tamone *et al.* (eBOSS), The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Growth rate of structure measurement from anisotropic clustering analysis in configuration space between redshift 0.6 and 1.1 for the Emission Line Galaxy sample, Mon. Not. Roy. Astron. Soc. **499**, 5527 (2020), arXiv:2007.09009 [astro-ph.CO].

[82] I. Szapudi, Wide angle redshift distortions revisited, Astrophys. J. **614**, 51 (2004), arXiv:astro-ph/0404477.

[83] A. Raccanelli, D. Bertacca, R. Maartens, C. Clarkson, and O. Doré, Lensing and time-delay contributions to galaxy correlations, Gen. Rel. Grav. **48**, 84 (2016), arXiv:1311.6813 [astro-ph.CO].

[84] A. Raccanelli, D. Bertacca, O. Doré, and R. Maartens, Large-scale 3D galaxy correlation function and non-Gaussianity, JCAP **08**, 022, arXiv:1306.6646 [astro-ph.CO].

[85] J. Yoo and U. Seljak, Wide Angle Effects in Future Galaxy Surveys, Mon. Not. Roy. Astron. Soc. **447**, 1789 (2015), arXiv:1308.1093 [astro-ph.CO].

[86] M. S. Vogeley and A. S. Szalay, Eigenmode analysis of galaxy redshift surveys I. theory and methods, Astrophys. J. **465**, 34 (1996), arXiv:astro-ph/9601185.

[87] M. Tegmark, Measuring cosmological parameters with galaxy surveys, Phys. Rev. Lett. **79**, 3806 (1997), arXiv:astro-ph/9706198.

[88] J. L. Bernal, N. Bellomo, A. Raccanelli, and L. Verde, Beware of commonly used approximations. Part II. Estimating systematic biases in the best-fit parameters, JCAP **10**, 017, arXiv:2005.09666 [astro-ph.CO].

[89] A. Raccanelli, L. Verde, and F. Villaescusa-Navarro, Biases from neutrino bias: to worry or not to worry?, Mon. Not. Roy. Astron. Soc. **483**, 734 (2019), arXiv:1704.07837 [astro-ph.CO].

[90] C. Bonvin, L. Hui, and E. Gaztanaga, Optimising the measurement of relativistic distortions in large-scale structure, JCAP **08**, 021, arXiv:1512.03566 [astro-ph.CO].

[91] V. Tansella, G. Jelic-Cizmek, C. Bonvin, and R. Durrer, COFFE: a code for the full-sky relativistic galaxy correlation function, JCAP **10**, 032, arXiv:1806.11090 [astro-ph.CO].

[92] M. Alvarez *et al.*, Testing Inflation with Large Scale Structure: Connecting Hopes with Reality, (2014), arXiv:1412.4671 [astro-ph.CO].

[93] G. R. Dvali, G. Gabadadze, and M. Porrati, 4-D gravity on a brane in 5-D Minkowski space, Phys. Lett. B **485**, 208 (2000), arXiv:hep-th/0005016.

[94] F. Schmidt, Cosmological Simulations of Normal-Branch Braneworld Gravity, Phys. Rev. D **80**, 123003 (2009), arXiv:0910.0235 [astro-ph.CO].

[95] M. A. Luty, M. Porrati, and R. Rattazzi, Strong interactions and stability in the DGP model, JHEP **09**, 029, arXiv:hep-th/0303116.

[96] K. Koyama, Ghosts in the self-accelerating universe, Class. Quant. Grav. **24**, R231 (2007), arXiv:0709.2399 [hep-th].

[97] A. Raccanelli, D. Bertacca, D. Pietrobon, F. Schmidt, L. Samushia, N. Bartolo, O. Dore, S. Matarrese, and W. J. Percival, Testing Gravity Using Large-Scale Redshift-Space Distortions, Mon. Not. Roy. Astron. Soc. **436**, 89 (2013), arXiv:1207.0500 [astro-ph.CO].

[98] R. Liu, G. Valogiannis, N. Battaglia, and R. Bean, Constraints on f(R) and normal-branch Dvali-Gabadadze-Porrati modified gravity model parameters with cluster abundances and galaxy clustering, Phys. Rev. D **104**, 103519 (2021), arXiv:2101.08728 [astro-ph.CO].

[99] M. Bosi, N. Bellomo, and A. Raccanelli, Constraining extended cosmologies with GW×LSS cross-correlations, JCAP **11**, 086, arXiv:2306.03031 [astro-ph.CO].

[100] L. Piga, M. Marinucci, G. D'Amico, M. Pietroni, F. Vernizzi, and B. S. Wright, Constraints on modified gravity from the BOSS galaxy survey, JCAP **04**, 038, arXiv:2211.12523 [astro-ph.CO].

[101] A. Raccanelli and Z. Vlah, Power spectrum in the cave, (2023), arXiv:2305.16278 [astro-ph.CO].
[102] A. Raccanelli and Z. Vlah, Observed power spectrum and frequency-angular power spectrum, Phys. Rev. D **108**, 043537 (2023), arXiv:2306.00808 [astro-ph.CO].
[103] Z. Gao, A. Raccanelli, and Z. Vlah, Asymptotic connection between full- and flat-sky angular correlators, Phys. Rev. D **108**, 043503 (2023), arXiv:2306.02993 [astro-ph.CO].
[104] A. F. Heavens and A. N. Taylor, A Spherical Harmonic Analysis of Redshift Space, Mon. Not. Roy. Astron. Soc. **275**, 483 (1995), arXiv:astro-ph/9409027.
[105] W. J. Percival *et al.* (2dFGRS), The 2dF Galaxy Redshift Survey: Spherical harmonics analysis of fluctuations in the final catalogue, Mon. Not. Roy. Astron. Soc. **353**, 1201 (2004), arXiv:astro-ph/0406513.
[106] J. Yoo and V. Desjacques, All-Sky Analysis of the General Relativistic Galaxy Power Spectrum, Phys. Rev. D **88**, 023502 (2013), arXiv:1301.4501 [astro-ph.CO].
[107] F. Semenzato, D. Bertacca, and A. Raccanelli, The full-sky Spherical Fourier-Bessel power spectrum in general relativity, (2024), arXiv:2406.09545 [astro-ph.CO].
[108] J. Yoo, General Relativistic Description of the Observed Galaxy Power Spectrum: Do We Understand What We Measure?, Phys. Rev. D **82**, 083508 (2010), arXiv:1009.3021 [astro-ph.CO].
[109] C. Bonvin and R. Durrer, What galaxy surveys really measure, Phys. Rev. D **84**, 063505 (2011), arXiv:1105.5280 [astro-ph.CO].
[110] A. Challinor and A. Lewis, The linear power spectrum of observed source number counts, Phys. Rev. D **84**, 043516 (2011), arXiv:1105.5292 [astro-ph.CO].
[111] D. Bertacca, R. Maartens, A. Raccanelli, and C. Clarkson, Beyond the plane-parallel and Newtonian approach: Wide-angle redshift distortions and convergence in general relativity, JCAP **10**, 025, arXiv:1205.5221 [astro-ph.CO].
[112] A. Raccanelli, D. Bertacca, D. Jeong, M. C. Neyrinck, and A. S. Szalay, Doppler term in the galaxy two-point correlation function: wide-angle, velocity, Doppler lensing and cosmic acceleration effects, Phys. Dark Univ. **19**, 109 (2018), arXiv:1602.03186 [astro-ph.CO].
[113] M. Borzyszkowski, D. Bertacca, and C. Porciani, LIGER: mock relativistic light-cones from Newtonian simulations, Mon. Not. Roy. Astron. Soc. **471**, 3899 (2017), arXiv:1703.03407 [astro-ph.CO].
[114] D. Bertacca, Generalisation of the Kaiser Rocket effect in general relativity in the wide-angle galaxy 2-point correlation function, Int. J. Mod. Phys. D **29**, 2050085 (2020), arXiv:1912.06887 [gr-qc].
[115] M. Y. Elkhashab, C. Porciani, and D. Bertacca, The large-scale monopole of the power spectrum in a Euclid-like survey: wide-angle effects, lensing, and the 'finger of the observer', Mon. Not. Roy. Astron. Soc. **509**, 1626 (2021), arXiv:2108.13424 [astro-ph.CO].
[116] M. Y. Elkhashab *et al.* (Euclid), Euclid preparation. The impact of relativistic redshift-space distortions on two-point clustering statistics from the Euclid wide spectroscopic survey, (2024), arXiv:2410.00956 [astro-ph.CO].
[117] L. Verde and S. Matarrese, Detectability of the effect of Inflationary non-Gaussianity on halo bias, Astrophys. J. Lett. **706**, L91 (2009), arXiv:0909.3224 [astro-ph.CO].
[118] V. Desjacques and U. Seljak, Primordial non-Gaussianity from the large scale structure, Class. Quant. Grav. **27**, 124011 (2010), arXiv:1003.5020 [astro-ph.CO].
[119] Z. Gao, Z. Vlah, and A. Challinor, Flat-sky angular power spectra revisited, JCAP **02**, 003, arXiv:2307.13768 [astro-ph.CO].
[120] J. Yoo, A. L. Fitzpatrick, and M. Zaldarriaga, A New Perspective on Galaxy Clustering as a Cosmological Probe: General Relativistic Effects, Phys. Rev. D **80**, 083514 (2009), arXiv:0907.0707 [astro-ph.CO].
[121] J. Yoo, General Relativistic Description of the Observed Galaxy Power Spectrum: Do We Understand What We Measure?, Phys. Rev. D **82**, 083508 (2010), arXiv:1009.3021 [astro-ph.CO].
[122] T. Steele, A. Raccanelli, and Z. Vlah, Projecting Unequal Time Fields and Correlators of Large Scale Structure, (2025), arXiv:2502.09518 [astro-ph.CO].
[123] D. Jeong, F. Schmidt, and C. M. Hirata, Large-scale clustering of galaxies in general relativity, Phys. Rev. D **85**, 023504 (2012), arXiv:1107.5427 [astro-ph.CO].
[124] F. Spezzati, E. Vanzan, A. Raccanelli, Z. Vlah, and D. Bertacca, Observed unequal-time power spectrum, Phys. Rev. D **112**, 023530 (2025), arXiv:2503.09759 [astro-ph.CO].
[125] R. Maartens, G.-B. Zhao, D. Bacon, K. Koyama, and A. Raccanelli, Relativistic corrections and non-Gaussianity in radio continuum surveys, JCAP **02**, 044, arXiv:1206.0732 [astro-ph.CO].
[126] A. Pillepich, C. Porciani, and O. Hahn, Universal halo mass function and scale-dependent bias from N-body

simulations with non-Gaussian initial conditions, Mon. Not. Roy. Astron. Soc. **402**, 191 (2010), arXiv:0811.4176 [astro-ph].

[127] M. Grossi, L. Verde, C. Carbone, K. Dolag, E. Branchini, F. Iannuzzi, S. Matarrese, and L. Moscardini, Large-scale non-Gaussian mass function and halo bias: tests on N-body simulations, Mon. Not. Roy. Astron. Soc. **398**, 321 (2009), arXiv:0902.2013 [astro-ph.CO].

[128] P. McDonald and U. Seljak, How to measure redshift-space distortions without sample variance, JCAP **10**, 007, arXiv:0810.0323 [astro-ph].

[129] U. Seljak, Extracting primordial non-gaussianity without cosmic variance, Phys. Rev. Lett. **102**, 021302 (2009), arXiv:0807.1770 [astro-ph].

[130] E. Sefusatti, M. Crocce, S. Pueblas, and R. Scoccimarro, Cosmology and the Bispectrum, Phys. Rev. D **74**, 023522 (2006), arXiv:astro-ph/0604505.

[131] D. Wadekar and R. Scoccimarro, Galaxy power spectrum multipoles covariance in perturbation theory, Phys. Rev. D **102**, 123517 (2020), arXiv:1910.02914 [astro-ph.CO].

[132] D. Braganca, Y. Donath, L. Senatore, and H. Zheng, Peeking into the next decade in Large-Scale Structure Cosmology with its Effective Field Theory, (2023), arXiv:2307.04992 [astro-ph.CO].

[133] R. Takahashi, N. Yoshida, M. Takada, T. Matsubara, N. Sugiyama, I. Kayo, A. J. Nishizawa, T. Nishimichi, S. Saito, and A. Taruya, Simulations of Baryon Acoustic Oscillations II: Covariance matrix of the matter power spectrum, Astrophys. J. **700**, 479 (2009), arXiv:0902.0371 [astro-ph.CO].

[134] R. Takahashi, N. Yoshida, M. Takada, T. Matsubara, N. Sugiyama, I. Kayo, T. Nishimichi, S. Saito, and A. Taruya, Non-Gaussian Error Contribution to Likelihood Analysis of the Matter Power Spectrum, Astrophys. J. **726**, 7 (2011), arXiv:0912.1381 [astro-ph.CO].

[135] Y. B. Zeldovich, Gravitational instability: An Approximate theory for large density perturbations, Astron. Astrophys. **5**, 84 (1970).

[136] T. Matsubara, Resumming Cosmological Perturbations via the Lagrangian Picture: One-loop Results in Real Space and in Redshift Space, Phys. Rev. D **77**, 063530 (2008), arXiv:0711.2521 [astro-ph].

[137] R. A. Porto, L. Senatore, and M. Zaldarriaga, The Lagrangian-space Effective Field Theory of Large Scale Structures, JCAP **05**, 022, arXiv:1311.2168 [astro-ph.CO].

[138] Z. Vlah, M. White, and A. Aviles, A Lagrangian effective field theory, JCAP **09**, 014, arXiv:1506.05264 [astro-ph.CO].

[139] D. Baumann, A. Nicolis, L. Senatore, and M. Zaldarriaga, Cosmological Non-Linearities as an Effective Fluid, JCAP **07**, 051, arXiv:1004.2488 [astro-ph.CO].

[140] J. J. M. Carrasco, M. P. Hertzberg, and L. Senatore, The Effective Field Theory of Cosmological Large Scale Structures, JHEP **09**, 082, arXiv:1206.2926 [astro-ph.CO].

[141] M. M. Ivanov, Effective Field Theory for Large-Scale Structure (2023) arXiv:2212.08488 [astro-ph.CO].

[142] S. Matarrese and M. Pietroni, Resumming Cosmic Perturbations, JCAP **06**, 026, arXiv:astro-ph/0703563.

[143] M. Pietroni, G. Mangano, N. Saviano, and M. Viel, Coarse-Grained Cosmological Perturbation Theory, JCAP **01**, 019, arXiv:1108.5203 [astro-ph.CO].

[144] M. Pietroni, Flowing with Time: a New Approach to Nonlinear Cosmological Perturbations, JCAP **10**, 036, arXiv:0806.0971 [astro-ph].

[145] S. Anselmi and M. Pietroni, Nonlinear Power Spectrum from Resummed Perturbation Theory: a Leap Beyond the BAO Scale, JCAP **12**, 013, arXiv:1205.2235 [astro-ph.CO].

[146] M. Crocce and R. Scoccimarro, Renormalized cosmological perturbation theory, Phys. Rev. D **73**, 063519 (2006), arXiv:astro-ph/0509418.

[147] P. McDonald, Dark matter clustering: a simple renormalization group approach, Phys. Rev. D **75**, 043514 (2007), arXiv:astro-ph/0606028.

[148] W. Wang *et al.*, The boundary of cosmic filaments, Mon. Not. Roy. Astron. Soc. **532**, 4604 (2024), arXiv:2402.11678 [astro-ph.CO].

[149] D. Galárraga-Espinosa *et al.*, Evolution of cosmic filaments in the MTNG simulation, Astron. Astrophys. **684**, A63 (2024), arXiv:2309.08659 [astro-ph.CO].

[150] T. Baldauf, M. Mirbabayi, M. Simonović, and M. Zaldarriaga, Equivalence Principle and the Baryon Acoustic Peak, Phys. Rev. D **92**, 043514 (2015), arXiv:1504.04366 [astro-ph.CO].

[151] M. Mirbabayi, M. Simonović, and M. Zaldarriaga, Baryon Acoustic Peak and the Squeezed Limit Bispectrum, (2014), arXiv:1412.3796 [astro-ph.CO].

[152] L. Senatore and M. Zaldarriaga, The IR-resummed Effective Field Theory of Large Scale Structures, JCAP **02**, 013, arXiv:1404.5954 [astro-ph.CO].
[153] L. Senatore and G. Trevisan, On the IR-Resummation in the EFTofLSS, JCAP **05**, 019, arXiv:1710.02178 [astro-ph.CO].
[154] L. Senatore, Bias in the Effective Field Theory of Large Scale Structures, JCAP **11**, 007, arXiv:1406.7843 [astro-ph.CO].
[155] V. Assassi, D. Baumann, D. Green, and M. Zaldarriaga, Renormalized Halo Bias, JCAP **08**, 056, arXiv:1402.5916 [astro-ph.CO].
[156] J. Lesgourgues and S. Pastor, Massive neutrinos and cosmology, Phys. Rept. **429**, 307 (2006), arXiv:astro-ph/0603494.
[157] F. Villaescusa-Navarro, F. Marulli, M. Viel, E. Branchini, E. Castorina, E. Sefusatti, and S. Saito, Cosmology with massive neutrinos I: towards a realistic modeling of the relation between matter, haloes and galaxies, JCAP **03**, 011, arXiv:1311.0866 [astro-ph.CO].
[158] E. Castorina, E. Sefusatti, R. K. Sheth, F. Villaescusa-Navarro, and M. Viel, Cosmology with massive neutrinos II: on the universality of the halo mass function and bias, JCAP **02**, 049, arXiv:1311.1212 [astro-ph.CO].
[159] L. Hui and K. P. Parfrey, The Evolution of Bias: Generalized, Phys. Rev. D **77**, 043527 (2008), arXiv:0712.1162 [astro-ph].
[160] M. LoVerde, Halo bias in mixed dark matter cosmologies, Phys. Rev. D **90**, 083530 (2014), arXiv:1405.4855 [astro-ph.CO].
[161] E. Castorina, C. Carbone, J. Bel, E. Sefusatti, and K. Dolag, DEMNUni: The clustering of large-scale structures in the presence of massive neutrinos, JCAP **07**, 043, arXiv:1505.07148 [astro-ph.CO].
[162] F. Villaescusa-Navarro, A. Banerjee, N. Dalal, E. Castorina, R. Scoccimarro, R. Angulo, and D. N. Spergel, The imprint of neutrinos on clustering in redshift-space, Astrophys. J. **861**, 53 (2018), arXiv:1708.01154 [astro-ph.CO].
[163] C.-T. Chiang, W. Hu, Y. Li, and M. Loverde, Scale-dependent bias and bispectrum in neutrino separate universe simulations, Phys. Rev. D **97**, 123526 (2018), arXiv:1710.01310 [astro-ph.CO].
[164] F. Schmidt and F. Beutler, Imprints of Reionization in Galaxy Clustering, Phys. Rev. D **96**, 083533 (2017), arXiv:1705.07843 [astro-ph.CO].
[165] A. Lamberts, P. Chang, C. Pfrommer, E. Puchwein, A. E. Broderick, and M. Shalaby, Patchy blazar heating: diversifying the thermal history of the intergalactic medium, Astrophys. J. **811**, 19 (2015), arXiv:1502.07980 [astro-ph.CO].
[166] D. Tseliakhovich and C. Hirata, Relative velocity of dark matter and baryonic fluids and the formation of the first structures, Phys. Rev. D **82**, 083520 (2010), arXiv:1005.2416 [astro-ph.CO].
[167] V. Assassi, D. Baumann, D. Green, and M. Zaldarriaga, Renormalized Halo Bias, JCAP **08**, 056, arXiv:1402.5916 [astro-ph.CO].
[168] R. Angulo, M. Fasiello, L. Senatore, and Z. Vlah, On the Statistics of Biased Tracers in the Effective Field Theory of Large Scale Structures, JCAP **09**, 029, arXiv:1503.08826 [astro-ph.CO].
[169] G. D'Amico, M. Marinucci, M. Pietroni, and F. Vernizzi, The large scale structure bootstrap: perturbation theory and bias expansion from symmetries, JCAP **10**, 069, arXiv:2109.09573 [astro-ph.CO].
[170] 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, JCAP **04**, 030, arXiv:1708.08130 [astro-ph.CO].
[171] N.-M. Nguyen, F. Schmidt, B. Tucci, M. Reinecke, and A. Kostić, How much information can be extracted from galaxy clustering at the field level?, (2024), arXiv:2403.03220 [astro-ph.CO].
[172] D. Gualdi, S. Novell-Masot, H. Gil-Marín, and L. Verde, Matter trispectrum: theoretical modelling and comparison to N-body simulations, JCAP **01**, 015, arXiv:2009.02290 [astro-ph.CO].
[173] D. Gualdi and L. Verde, Integrated trispectrum detection from BOSS DR12 NGC CMASS, JCAP **09**, 050, arXiv:2201.06932 [astro-ph.CO].
[174] A. Chudaykin, M. M. Ivanov, and M. Simonović, Optimizing large-scale structure data analysis with the theoretical error likelihood, Phys. Rev. D **103**, 043525 (2021), arXiv:2009.10724 [astro-ph.CO].
[175] D. Wadekar, M. M. Ivanov, and R. Scoccimarro, Cosmological constraints from BOSS with analytic covariance matrices, Phys. Rev. D **102**, 123521 (2020), arXiv:2009.00622 [astro-ph.CO].
[176] T. Baldauf, M. Mirbabayi, M. Simonović, and M. Zaldarriaga, LSS constraints with controlled theoretical uncertainties, (2016), arXiv:1602.00674 [astro-ph.CO].

[177] M. G. Moreira, F. Andrade-Oliveira, X. Fang, H.-J. Huang, E. Krause, V. Miranda, R. Rosenfeld, and M. Simonović, Mitigating baryonic effects with a theoretical error covariance, Mon. Not. Roy. Astron. Soc. **507**, 5592 (2021), arXiv:2104.01397 [astro-ph.CO].

[178] A. Aires, N. Kokron, R. Rosenfeld, F. Andrade-Oliveira, and V. Miranda, Mitigation of nonlinear galaxy bias with a theoretical-error likelihood, JCAP **04**, 070, arXiv:2410.08930 [astro-ph.CO].

[179] A. Maraio, A. Hall, and A. Taylor, Mitigating baryon feedback bias in cosmic shear through a theoretical error covariance in the matter power spectrum, Mon. Not. Roy. Astron. Soc. **537**, 1749 (2025), arXiv:2410.12500 [astro-ph.CO].

[180] 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, JCAP **03**, 056, arXiv:1908.01774 [astro-ph.CO].

[181] S. Martin, P. Schneider, and P. Simon, The bispectrum covariance beyond Gaussianity: A log-normal approach, Astron. Astrophys. **540**, A9 (2012), arXiv:1109.0944 [astro-ph.CO].

[182] B. Tucci and F. Schmidt, EFTofLSS meets simulation-based inference: σ_8 from biased tracers, JCAP **05**, 063, arXiv:2310.03741 [astro-ph.CO].

[183] M. M. Ivanov, O. H. E. Philcox, T. Nishimichi, M. Simonović, M. Takada, and M. Zaldarriaga, Precision analysis of the redshift-space galaxy bispectrum, Phys. Rev. D **105**, 063512 (2022), arXiv:2110.10161 [astro-ph.CO].

[184] K. C. Chan and L. Blot, Assessment of the Information Content of the Power Spectrum and Bispectrum, Phys. Rev. D **96**, 023528 (2017), arXiv:1610.06585 [astro-ph.CO].

[185] S. Novell-Masot, H. Gil-Marín, and L. Verde, On approximations of the redshift-space bispectrum and power spectrum multipoles covariance matrix, JCAP **06**, 048, arXiv:2306.03137 [astro-ph.CO].

[186] M. Mirbabayi, F. Schmidt, and M. Zaldarriaga, Biased Tracers and Time Evolution, JCAP **07**, 030, arXiv:1412.5169 [astro-ph.CO].

[187] L. Senatore and M. Zaldarriaga, Redshift Space Distortions in the Effective Field Theory of Large Scale Structures, (2014), arXiv:1409.1225 [astro-ph.CO].

[188] G. Cabass, M. M. Ivanov, M. Lewandowski, M. Mirbabayi, and M. Simonović, Snowmass white paper: Effective field theories in cosmology, Phys. Dark Univ. **40**, 101193 (2023), arXiv:2203.08232 [astro-ph.CO].

[189] T. Baldauf, E. Schaan, and M. Zaldarriaga, On the reach of perturbative descriptions for dark matter displacement fields, JCAP **03**, 017, arXiv:1505.07098 [astro-ph.CO].

[190] T. Baldauf, E. Schaan, and M. Zaldarriaga, On the reach of perturbative methods for dark matter density fields, JCAP **03**, 007, arXiv:1507.02255 [astro-ph.CO].

[191] M. Schmittfull, M. Simonović, V. Assassi, and M. Zaldarriaga, Modeling Biased Tracers at the Field Level, Phys. Rev. D **100**, 043514 (2019), arXiv:1811.10640 [astro-ph.CO].

[192] F. Schmidt, F. Elsner, J. Jasche, N. M. Nguyen, and G. Lavaux, A rigorous EFT-based forward model for large-scale structure, JCAP **01**, 042, arXiv:1808.02002 [astro-ph.CO].

[193] C. Modi, M. White, A. Slosar, and E. Castorina, Reconstructing large-scale structure with neutral hydrogen surveys, JCAP **11**, 023, arXiv:1907.02330 [astro-ph.CO].

[194] M. Schmittfull, M. Simonović, M. M. Ivanov, O. H. E. Philcox, and M. Zaldarriaga, Modeling Galaxies in Redshift Space at the Field Level, JCAP **05**, 059, arXiv:2012.03334 [astro-ph.CO].

[195] F. Schmidt, An n-th order Lagrangian Forward Model for Large-Scale Structure, JCAP **04**, 033, arXiv:2012.09837 [astro-ph.CO].

[196] A. Obuljen, M. Simonović, A. Schneider, and R. Feldmann, Modeling HI at the field level, Phys. Rev. D **108**, 083528 (2023), arXiv:2207.12398 [astro-ph.CO].

[197] J. Stadler, F. Schmidt, and M. Reinecke, Fast, accurate and perturbative forward modeling of galaxy clustering. Part I. Galaxies in the restframe, JCAP **04**, 089, arXiv:2409.10937 [astro-ph.CO].

[198] J. Stadler, F. Schmidt, M. Reinecke, and M. Esposito, Fast, Accurate and Perturbative Forward Modeling of Galaxy Clustering Part II: Redshift Space, (2024), arXiv:2411.04513 [astro-ph.CO].

[199] G. Cabass and F. Schmidt, The EFT Likelihood for Large-Scale Structure, JCAP **04**, 042, arXiv:1909.04022 [astro-ph.CO].

[200] F. Elsner, F. Schmidt, J. Jasche, G. Lavaux, and N.-M. Nguyen, Cosmology inference from a biased density field using the EFT-based likelihood, JCAP **01**, 029, arXiv:1906.07143 [astro-ph.CO].

[201] F. Schmidt, G. Cabass, J. Jasche, and G. Lavaux, Unbiased Cosmology Inference from Biased Tracers using the EFT Likelihood, JCAP **11**, 008, arXiv:2004.06707 [astro-ph.CO].

[202] G. Cabass and F. Schmidt, The Likelihood for LSS: Stochasticity of Bias Coefficients at All Orders, JCAP **07**, 051,

arXiv:2004.00617 [astro-ph.CO].
[203] F. Schmidt, Sigma-Eight at the Percent Level: The EFT Likelihood in Real Space, JCAP **04**, 032, arXiv:2009.14176 [astro-ph.CO].
[204] A. Kostić, N.-M. Nguyen, F. Schmidt, and M. Reinecke, Consistency tests of field level inference with the EFT likelihood, JCAP **07**, 063, arXiv:2212.07875 [astro-ph.CO].
[205] I. Babić, F. Schmidt, and B. Tucci, BAO scale inference from biased tracers using the EFT likelihood, JCAP **08** (08), 007, arXiv:2203.06177 [astro-ph.CO].
[206] J. Stadler, F. Schmidt, and M. Reinecke, Cosmology inference at the field level from biased tracers in redshift-space, JCAP **10**, 069, arXiv:2303.09876 [astro-ph.CO].
[207] I. Babić, F. Schmidt, and B. Tucci, Straightening the Ruler: Field-Level Inference of the BAO Scale with LEFTfield, (2024), arXiv:2407.01524 [astro-ph.CO].
[208] M. Peron, T. Nishimichi, M. Pietroni, and A. Taruya, Renormalized Perturbation Theory at Field-level: the LSS bootstrap in GridSPT, (2025), arXiv:2506.07105 [astro-ph.CO].
[209] L. Amendola, M. Marinucci, M. Pietroni, and M. Quartin, Improving precision and accuracy in cosmology with model-independent spectrum and bispectrum, JCAP **01**, 001, arXiv:2307.02117 [astro-ph.CO].
[210] M. Marinucci, K. Pardede, and M. Pietroni, Bootstrapping Lagrangian perturbation theory for the large scale structure, JCAP **10**, 051, arXiv:2405.08413 [astro-ph.CO].
[211] A. Ansari, A. Banerjee, S. Jain, and S. Lalsodagar, Bootstrapping LSS perturbation theory beyond third order, (2025), arXiv:2504.01078 [astro-ph.CO].
[212] E. Krause *et al.* (Beyond-2pt), A Parameter-Masked Mock Data Challenge for Beyond-Two-Point Galaxy Clustering Statistics, (2024), arXiv:2405.02252 [astro-ph.CO].
[213] G. Cabass, M. Simonović, and M. Zaldarriaga, Cosmological information in perturbative forward modeling, Phys. Rev. D **109**, 043526 (2024), arXiv:2307.04706 [astro-ph.CO].
[214] F. Schmidt, On the Connection between Field-Level Inference and n-point Correlation Functions, (2025), arXiv:2504.15351 [astro-ph.CO].
[215] A. Taruya, T. Nishimichi, and D. Jeong, Grid-based calculation for perturbation theory of large-scale structure, Phys. Rev. D **98**, 103532 (2018), arXiv:1807.04215 [astro-ph.CO].
[216] M. Peloso and M. Pietroni, Galilean invariance and the consistency relation for the nonlinear squeezed bispectrum of large scale structure, JCAP **05**, 031, arXiv:1302.0223 [astro-ph.CO].
[217] A. Kehagias and A. Riotto, Symmetries and Consistency Relations in the Large Scale Structure of the Universe, Nucl. Phys. B **873**, 514 (2013), arXiv:1302.0130 [astro-ph.CO].
[218] P. Creminelli, J. Gleyzes, M. Simonović, and F. Vernizzi, Single-Field Consistency Relations of Large Scale Structure. Part II: Resummation and Redshift Space, JCAP **02**, 051, arXiv:1311.0290 [astro-ph.CO].
[219] P. Creminelli, J. Noreña, M. Simonović, and F. Vernizzi, Single-Field Consistency Relations of Large Scale Structure, JCAP **12**, 025, arXiv:1309.3557 [astro-ph.CO].
[220] E. Sefusatti, M. Crocce, and V. Desjacques, The Matter Bispectrum in N-body Simulations with non-Gaussian Initial Conditions, Mon. Not. Roy. Astron. Soc. **406**, 1014 (2010), arXiv:1003.0007 [astro-ph.CO].
[221] A. Moradinezhad Dizgah, M. Biagetti, E. Sefusatti, V. Desjacques, and J. Noreña, Primordial Non-Gaussianity from Biased Tracers: Likelihood Analysis of Real-Space Power Spectrum and Bispectrum, JCAP **05**, 015, arXiv:2010.14523 [astro-ph.CO].
[222] D. Alkhanishvili, C. Porciani, E. Sefusatti, M. Biagetti, A. Lazanu, A. Oddo, and V. Yankelevich, The reach of next-to-leading-order perturbation theory for the matter bispectrum, Mon. Not. Roy. Astron. Soc. **512**, 4961 (2022), arXiv:2107.08054 [astro-ph.CO].
[223] A. Eggemeier, R. Scoccimarro, R. E. Smith, M. Crocce, A. Pezzotta, and A. G. Sánchez, Testing one-loop galaxy bias: Joint analysis of power spectrum and bispectrum, Phys. Rev. D **103**, 123550 (2021), arXiv:2102.06902 [astro-ph.CO].
[224] A. Oddo, F. Rizzo, E. Sefusatti, C. Porciani, and P. Monaco, Cosmological parameters from the likelihood analysis of the galaxy power spectrum and bispectrum in real space, JCAP **11**, 038, arXiv:2108.03204 [astro-ph.CO].
[225] T. Nishimichi, G. D'Amico, M. M. Ivanov, L. Senatore, M. Simonović, M. Takada, M. Zaldarriaga, and P. Zhang, Blinded challenge for precision cosmology with large-scale structure: results from effective field theory for the redshift-space galaxy power spectrum, Phys. Rev. D **102**, 123541 (2020), arXiv:2003.08277 [astro-ph.CO].
[226] G. D'Amico, L. Senatore, and P. Zhang, Limits on wCDM from the EFTofLSS with the PyBird code, JCAP **01**, 006, arXiv:2003.07956 [astro-ph.CO].

[227] A. Chudaykin, M. M. Ivanov, O. H. E. Philcox, and M. Simonović, Nonlinear perturbation theory extension of the Boltzmann code CLASS, Phys. Rev. D **102**, 063533 (2020), arXiv:2004.10607 [astro-ph.CO].
[228] S.-F. Chen, Z. Vlah, E. Castorina, and M. White, Redshift-Space Distortions in Lagrangian Perturbation Theory, JCAP **03**, 100, arXiv:2012.04636 [astro-ph.CO].
[229] D. Linde, A. Moradinezhad Dizgah, C. Radermacher, S. Casas, and J. Lesgourgues, CLASS-OneLoop: accurate and unbiased inference from spectroscopic galaxy surveys, JCAP **07**, 068, arXiv:2402.09778 [astro-ph.CO].
[230] T. Lazeyras, C. Wagner, T. Baldauf, and F. Schmidt, Precision measurement of the local bias of dark matter halos, JCAP **02**, 018, arXiv:1511.01096 [astro-ph.CO].
[231] T. Lazeyras and F. Schmidt, Beyond LIMD bias: a measurement of the complete set of third-order halo bias parameters, JCAP **09**, 008, arXiv:1712.07531 [astro-ph.CO].
[232] D. Baumann, D. Green, and B. Wallisch, Searching for light relics with large-scale structure, JCAP **08**, 029, arXiv:1712.08067 [astro-ph.CO].
[233] J. Hamann, S. Hannestad, J. Lesgourgues, C. Rampf, and Y. Y. Y. Wong, Cosmological parameters from large scale structure - geometric versus shape information, JCAP **07**, 022, arXiv:1003.3999 [astro-ph.CO].
[234] Z. Vlah, U. Seljak, M. Y. Chu, and Y. Feng, Perturbation theory, effective field theory, and oscillations in the power spectrum, JCAP **03**, 057, arXiv:1509.02120 [astro-ph.CO].
[235] D. Blas, M. Garny, M. M. Ivanov, and S. Sibiryakov, Time-Sliced Perturbation Theory II: Baryon Acoustic Oscillations and Infrared Resummation, JCAP **07**, 028, arXiv:1605.02149 [astro-ph.CO].
[236] O. H. E. Philcox, M. M. Ivanov, M. Simonović, and M. Zaldarriaga, Combining Full-Shape and BAO Analyses of Galaxy Power Spectra: A 1.6\% CMB-independent constraint on H_0, JCAP **05**, 032, arXiv:2002.04035 [astro-ph.CO].
[237] P. Zhang, G. D'Amico, L. Senatore, C. Zhao, and Y. Cai, BOSS Correlation Function analysis from the Effective Field Theory of Large-Scale Structure, JCAP **02** (02), 036, arXiv:2110.07539 [astro-ph.CO].
[238] S.-F. Chen, Z. Vlah, and M. White, A new analysis of galaxy 2-point functions in the BOSS survey, including full-shape information and post-reconstruction BAO, JCAP **02** (02), 008, arXiv:2110.05530 [astro-ph.CO].
[239] H. Gil-Marín, How to optimally combine pre-reconstruction full shape and post-reconstruction BAO signals, JCAP **05** (05), 040, arXiv:2203.05581 [astro-ph.CO].
[240] O. H. E. Philcox, M. M. Ivanov, G. Cabass, M. Simonović, M. Zaldarriaga, and T. Nishimichi, Cosmology with the redshift-space galaxy bispectrum monopole at one-loop order, Phys. Rev. D **106**, 043530 (2022), arXiv:2206.02800 [astro-ph.CO].
[241] G. D'Amico, Y. Donath, M. Lewandowski, L. Senatore, and P. Zhang, The BOSS bispectrum analysis at one loop from the Effective Field Theory of Large-Scale Structure, JCAP **05**, 059, arXiv:2206.08327 [astro-ph.CO].
[242] K. Akitsu, Mapping the galaxy-halo connection to the galaxy bias: implication to the HOD-informed prior, (2024), arXiv:2410.08998 [astro-ph.CO].
[243] M. M. Ivanov, C. Cuesta-Lazaro, S. Mishra-Sharma, A. Obuljen, and M. W. Toomey, Full-shape analysis with simulation-based priors: Constraints on single field inflation from BOSS, Phys. Rev. D **110**, 063538 (2024), arXiv:2402.13310 [astro-ph.CO].
[244] M. M. Ivanov, A. Obuljen, C. Cuesta-Lazaro, and M. W. Toomey, Full-shape analysis with simulation-based priors: Cosmological parameters and the structure growth anomaly, Phys. Rev. D **111**, 063548 (2025), arXiv:2409.10609 [astro-ph.CO].
[245] H. Zhang, M. Bonici, G. D'Amico, S. Paradiso, and W. J. Percival, HOD-informed prior for EFT-based full-shape analyses of LSS, JCAP **04**, 041, arXiv:2409.12937 [astro-ph.CO].
[246] M. M. Ivanov, Simulation-Based Priors without Simulations: an Analytic Perspective on EFT Parameters of Galaxies, (2025), arXiv:2503.07270 [astro-ph.CO].
[247] B. R.-S. Blancard *et al.*, SIMBIG: Galaxy Clustering Analysis with the Wavelet Scattering Transform, (2023), arXiv:2310.15250 [astro-ph.CO].
[248] G. Valogiannis and C. Dvorkin, Towards an optimal estimation of cosmological parameters with the wavelet scattering transform, PRD **105**, 103534 (2022), arXiv:2108.07821 [astro-ph.CO].
[249] G. Valogiannis, F. Villaescusa-Navarro, and M. Baldi, Towards unveiling the large-scale nature of gravity with the wavelet scattering transform, arXiv e-prints , arXiv:2407.18647 (2024), arXiv:2407.18647 [astro-ph.CO].
[250] G. Valogiannis, S. Yuan, and C. Dvorkin, Precise cosmological constraints from BOSS galaxy clustering with a simulation-based emulator of the wavelet scattering transform, PRD **109**, 103503 (2024), arXiv:2310.16116 [astro-ph.CO].

[251] M. Eickenberg, E. Allys, A. Moradinezhad Dizgah, P. Lemos, E. Massara, M. Abidi, C. Hahn, S. Hassan, B. Regaldo-Saint Blancard, S. Ho, S. Mallat, J. Andén, and F. Villaescusa-Navarro, Wavelet Moments for Cosmological Parameter Estimation, arXiv e-prints , arXiv:2204.07646 (2022), arXiv:2204.07646 [astro-ph.CO].

[252] M. Peron, G. Jung, M. Liguori, and M. Pietroni, Constraining primordial non-Gaussianity from large scale structure with the wavelet scattering transform, JCAP **2024** (7), 021, arXiv:2403.17657 [astro-ph.CO].

[253] A. Banerjee and T. Abel, Tracer-field cross-correlations with k-nearest neighbour distributions, MNRAS **519**, 4856 (2023), arXiv:2210.05140 [astro-ph.CO].

[254] W. R. Coulton, T. Abel, and A. Banerjee, Small-scale signatures of primordial non-Gaussianity in k-nearest neighbour cumulative distribution functions, MNRAS **534**, 1621 (2024), arXiv:2309.15151 [astro-ph.CO].

[255] M. Schmittfull and A. M. Dizgah, Galaxy skew-spectra in redshift-space, JCAP **2021** (3), 020, arXiv:2010.14267 [astro-ph.CO].

[256] J. Hou, A. M. Dizgah, C. Hahn, M. Eickenberg, S. Ho, P. Lemos, E. Massara, C. Modi, L. Parker, and B. R.-S. Blancard, Cosmological constraints from the redshift-space galaxy skew spectra, PRD **109**, 103528 (2024), arXiv:2401.15074 [astro-ph.CO].

[257] F. Simpson, A. F. Heavens, and C. Heymans, Clipping the cosmos. II. Cosmological information from nonlinear scales, PRD **88**, 083510 (2013), arXiv:1306.6349 [astro-ph.CO].

[258] M. Lippich and A. G. Sánchez, MEDUSA: Minkowski functionals estimated from Delaunay tessellations of the three-dimensional large-scale structure, MNRAS **508**, 3771 (2021), arXiv:2012.08529 [astro-ph.CO].

[259] W. Liu, A. Jiang, and W. Fang, Probing massive neutrinos with the Minkowski functionals of the galaxy distribution, JCAP **2023** (9), 037, arXiv:2302.08162 [astro-ph.CO].

[260] S. Dietrich, On correlation of marked point processes, Math.Nacr **116**, 197 (1984).

[261] M. White, A marked correlation function for constraining modified gravity models, JCAP **11**, 057, arXiv:1609.08632 [astro-ph.CO].

[262] E. Massara, F. Villaescusa-Navarro, S. Ho, N. Dalal, and D. N. Spergel, Using the Marked Power Spectrum to Detect the Signature of Neutrinos in Large-Scale Structure, Phys. Rev. Lett. **126**, 011301 (2021), arXiv:2001.11024 [astro-ph.CO].

[263] E. Massara, F. Villaescusa-Navarro, C. Hahn, M. M. Abidi, M. Eickenberg, S. Ho, P. Lemos, A. Moradinezhad Dizgah, and B. R.-S. Blancard, Cosmological Information in the Marked Power Spectrum of the Galaxy Field, Astrophys. J. **951**, 70 (2023), arXiv:2206.01709 [astro-ph.CO].

[264] G. Jung, A. Ravenni, M. Liguori, M. Baldi, W. R. Coulton, F. Villaescusa-Navarro, and B. D. Wandelt, Quijote-PNG: Optimizing the summary statistics to measure Primordial non-Gaussianity, Astrophys. J. **976**, 109 (2024), arXiv:2403.00490 [astro-ph.CO].

[265] O. H. E. Philcox, E. Massara, and D. N. Spergel, What does the marked power spectrum measure? Insights from perturbation theory, Phys. Rev. D **102**, 043516 (2020), arXiv:2006.10055 [astro-ph.CO].

[266] O. H. E. Philcox, A. Aviles, and E. Massara, Modeling the Marked Spectrum of Matter and Biased Tracers in Real- and Redshift-Space, JCAP **03**, 038, arXiv:2010.05914 [astro-ph.CO].

[267] H. Ebina and M. White, An Analytically Tractable Marked Power Spectrum, arXiv e-prints , arXiv:2409.17133 (2024), arXiv:2409.17133 [astro-ph.CO].

[268] J. A. Cowell, D. Alonso, and J. Liu, Hitting the mark: Optimising Marked Power Spectra for Cosmology, arXiv e-prints , arXiv:2409.05695 (2024), arXiv:2409.05695 [astro-ph.CO].

[269] M. White, A marked correlation function for constraining modified gravity models, JCAP **11**, 057, arXiv:1609.08632 [astro-ph.CO].

[270] A. Aviles, K. Koyama, J. L. Cervantes-Cota, H. A. Winther, and B. Li, Marked correlation functions in perturbation theory, JCAP **01**, 006, arXiv:1911.06362 [astro-ph.CO].

[271] M. M. Ivanov, O. H. E. Philcox, G. Cabass, T. Nishimichi, M. Simonović, and M. Zaldarriaga, Cosmology with the galaxy bispectrum multipoles: Optimal estimation and application to BOSS data, Phys. Rev. D **107**, 083515 (2023), arXiv:2302.04414 [astro-ph.CO].

[272] G. D'Amico, Y. Donath, M. Lewandowski, L. Senatore, and P. Zhang, The one-loop bispectrum of galaxies in redshift space from the Effective Field Theory of Large-Scale Structure, JCAP **07**, 041, arXiv:2211.17130 [astro-ph.CO].

[273] A. Chudaykin and M. M. Ivanov, Measuring neutrino masses with large-scale structure: Euclid forecast with controlled theoretical error, JCAP **11**, 034, arXiv:1907.06666 [astro-ph.CO].

[274] M. M. Ivanov, M. Simonović, and M. Zaldarriaga, Cosmological Parameters from the BOSS Galaxy Power

Spectrum, JCAP **05**, 042, arXiv:1909.05277 [astro-ph.CO].
[275] C. Cheung, P. Creminelli, A. L. Fitzpatrick, J. Kaplan, and L. Senatore, The Effective Field Theory of Inflation, JHEP **03**, 014, arXiv:0709.0293 [hep-th].
[276] S. Matarrese and L. Verde, The effect of primordial non-Gaussianity on halo bias, Astrophys. J. Lett. **677**, L77 (2008), arXiv:0801.4826 [astro-ph].
[277] V. Desjacques, U. Seljak, and I. Iliev, Scale-dependent bias induced by local non-Gaussianity: A comparison to N-body simulations, Mon. Not. Roy. Astron. Soc. **396**, 85 (2009), arXiv:0811.2748 [astro-ph].
[278] N. Dalal, O. Dore, D. Huterer, and A. Shirokov, The imprints of primordial non-gaussianities on large-scale structure: scale dependent bias and abundance of virialized objects, Phys. Rev. D **77**, 123514 (2008), arXiv:0710.4560 [astro-ph].
[279] U. Giri, M. Münchmeyer, and K. M. Smith, Robust neural network-enhanced estimation of local primordial non-Gaussianity, Phys. Rev. D **107**, L061301 (2023), arXiv:2205.12964 [astro-ph.CO].
[280] Y. Kvasiuk, M. Münchmeyer, and K. Smith, A Tale of Two Fields: Neural Network-Enhanced non-Gaussianity Search with Halos, (2024), arXiv:2410.01007 [astro-ph.CO].
[281] N. Arkani-Hamed and J. Maldacena, Cosmological Collider Physics, (2015), arXiv:1503.08043 [hep-th].
[282] G. Cabass, O. H. E. Philcox, M. M. Ivanov, K. Akitsu, S.-F. Chen, M. Simonović, and M. Zaldarriaga, BOSS Constraints on Massive Particles during Inflation: The Cosmological Collider in Action, (2024), arXiv:2404.01894 [astro-ph.CO].
[283] L. Senatore, K. M. Smith, and M. Zaldarriaga, Non-Gaussianities in Single Field Inflation and their Optimal Limits from the WMAP 5-year Data, JCAP **01**, 028, arXiv:0905.3746 [astro-ph.CO].
[284] Y. Akrami *et al.* (Planck), Planck 2018 results. IX. Constraints on primordial non-Gaussianity, Astron. Astrophys. **641**, A9 (2020), arXiv:1905.05697 [astro-ph.CO].
[285] W. R. Coulton, F. Villaescusa-Navarro, D. Jamieson, M. Baldi, G. Jung, D. Karagiannis, M. Liguori, L. Verde, and B. D. Wandelt, Quijote-PNG: Simulations of Primordial Non-Gaussianity and the Information Content of the Matter Field Power Spectrum and Bispectrum, Astrophys. J. **943**, 64 (2023), arXiv:2206.01619 [astro-ph.CO].
[286] Y. Akrami *et al.* (Planck), Planck 2018 results. X. Constraints on inflation, Astron. Astrophys. **641**, A10 (2020), arXiv:1807.06211 [astro-ph.CO].
[287] G. Cabass, M. M. Ivanov, O. H. E. Philcox, M. Simonović, and M. Zaldarriaga, Constraints on Single-Field Inflation from the BOSS Galaxy Survey, Phys. Rev. Lett. **129**, 021301 (2022), arXiv:2201.07238 [astro-ph.CO].
[288] F. Villaescusa-Navarro *et al.*, The Quijote simulations, Astrophys. J. Suppl. **250**, 2 (2020), arXiv:1909.05273 [astro-ph.CO].
[289] M. Crocce, S. Pueblas, and R. Scoccimarro, Transients from Initial Conditions in Cosmological Simulations, Mon. Not. Roy. Astron. Soc. **373**, 369 (2006), arXiv:astro-ph/0606505.
[290] R. Scoccimarro, L. Hui, M. Manera, and K. C. Chan, Large-scale Bias and Efficient Generation of Initial Conditions for Non-Local Primordial Non-Gaussianity, Phys. Rev. D **85**, 083002 (2012), arXiv:1108.5512 [astro-ph.CO].
[291] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White, The Evolution of Large Scale Structure in a Universe Dominated by Cold Dark Matter, Astrophys. J. **292**, 371 (1985).
[292] G. Jung, D. Karagiannis, M. Liguori, M. Baldi, W. R. Coulton, D. Jamieson, L. Verde, F. Villaescusa-Navarro, and B. D. Wandelt, Quijote-PNG: Quasi-maximum Likelihood Estimation of Primordial Non-Gaussianity in the Nonlinear Halo Density Field, ApJ **948**, 135 (2023), arXiv:2211.07565 [astro-ph.CO].
[293] G. Jung *et al.*, Quijote-PNG: The Information Content of the Halo Mass Function, Astrophys. J. **957**, 50 (2023), arXiv:2305.10597 [astro-ph.CO].
[294] V. Yankelevich and C. Porciani, Cosmological information in the redshift-space bispectrum, Mon. Not. Roy. Astron. Soc. **483**, 2078 (2019), arXiv:1807.07076 [astro-ph.CO].
[295] M. M. Abidi and T. Baldauf, Cubic Halo Bias in Eulerian and Lagrangian Space, JCAP **07**, 029, arXiv:1802.07622 [astro-ph.CO].
[296] D. Blas, J. Lesgourgues, and T. Tram, The Cosmic Linear Anisotropy Solving System (CLASS) II: Approximation schemes, JCAP **07**, 034, arXiv:1104.2933 [astro-ph.CO].
[297] M. M. Ivanov, M. Simonović, and M. Zaldarriaga, Cosmological Parameters and Neutrino Masses from the Final Planck and Full-Shape BOSS Data, Phys. Rev. D **101**, 083504 (2020), arXiv:1912.08208 [astro-ph.CO].
[298] T. Flöss, M. Biagetti, and P. D. Meerburg, Primordial non-Gaussianity and non-Gaussian covariance, Phys. Rev. D **107**, 023528 (2023), arXiv:2206.10458 [astro-ph.CO].

[299] M. M. Ivanov, M. Simonović, and M. Zaldarriaga, Cosmological Parameters from the BOSS Galaxy Power Spectrum, JCAP **05**, 042, arXiv:1909.05277 [astro-ph.CO].
[300] G. D'Amico, J. Gleyzes, N. Kokron, K. Markovic, L. Senatore, P. Zhang, F. Beutler, and H. Gil-Marín, The Cosmological Analysis of the SDSS/BOSS data from the Effective Field Theory of Large-Scale Structure, JCAP **05**, 005, arXiv:1909.05271 [astro-ph.CO].
[301] A. Lewis, The real shape of non-Gaussianities, JCAP **10**, 026, arXiv:1107.5431 [astro-ph.CO].
[302] C. Hahn and F. Villaescusa-Navarro, Constraining M_ν with the bispectrum. Part II. The information content of the galaxy bispectrum monopole, JCAP **04**, 029, arXiv:2012.02200 [astro-ph.CO].
[303] A. Perko, L. Senatore, E. Jennings, and R. H. Wechsler, Biased Tracers in Redshift Space in the EFT of Large-Scale Structure, (2016), arXiv:1610.09321 [astro-ph.CO].
[304] M. Lewandowski, L. Senatore, F. Prada, C. Zhao, and C.-H. Chuang, EFT of large scale structures in redshift space, Phys. Rev. D **97**, 063526 (2018), arXiv:1512.06831 [astro-ph.CO].
[305] A. Baleato Lizancos, U. Seljak, M. Karamanis, M. Bonici, and S. Ferraro, Selecting samples of galaxies with fewer Fingers-of-God, JCAP **07**, 014, arXiv:2501.10587 [astro-ph.CO].
[306] M. M. Ivanov and S. Sibiryakov, Infrared Resummation for Biased Tracers in Redshift Space, JCAP **07**, 053, arXiv:1804.05080 [astro-ph.CO].
[307] C. Alcock and B. Paczynski, An evolution free test for non-zero cosmological constant, Nature **281**, 358 (1979).
[308] F. Beutler *et al.* (BOSS), The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: Anisotropic galaxy clustering in Fourier-space, Mon. Not. Roy. Astron. Soc. **466**, 2242 (2017), arXiv:1607.03150 [astro-ph.CO].
[309] F. Beutler, E. Castorina, and P. Zhang, Interpreting measurements of the anisotropic galaxy power spectrum, JCAP **03**, 040, arXiv:1810.05051 [astro-ph.CO].
[310] T. Simon, P. Zhang, V. Poulin, and T. L. Smith, Updated constraints from the effective field theory analysis of the BOSS power spectrum on early dark energy, Phys. Rev. D **107**, 063505 (2023), arXiv:2208.05930 [astro-ph.CO].
[311] K. Yamamoto, M. Nakamichi, A. Kamino, B. A. Bassett, and H. Nishioka, A Measurement of the quadrupole power spectrum in the clustering of the 2dF QSO Survey, Publ. Astron. Soc. Jap. **58**, 93 (2006), arXiv:astro-ph/0505115.
[312] H. A. Feldman, N. Kaiser, and J. A. Peacock, Power spectrum analysis of three-dimensional redshift surveys, Astrophys. J. **426**, 23 (1994), arXiv:astro-ph/9304022.
[313] F. Beutler and P. McDonald, Unified galaxy power spectrum measurements from 6dFGS, BOSS, and eBOSS, JCAP **11**, 031, arXiv:2106.06324 [astro-ph.CO].
[314] F. Beutler *et al.* (BOSS), The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Testing gravity with redshift-space distortions using the power spectrum multipoles, Mon. Not. Roy. Astron. Soc. **443**, 1065 (2014), arXiv:1312.4611 [astro-ph.CO].
[315] A. J. Benson, Galacticus: A Semi-Analytic Model of Galaxy Formation, New Astron. **17**, 175 (2012), arXiv:1008.1786 [astro-ph.CO].
[316] C.-H. Chuang *et al.*, UNIT project: Universe N-body simulations for the Investigation of Theoretical models from galaxy surveys, Mon. Not. Roy. Astron. Soc. **487**, 48 (2019), arXiv:1811.02111 [astro-ph.CO].
[317] Z. Zhai, C.-H. Chuang, Y. Wang, A. Benson, and G. Yepes, Clustering in the simulated H α galaxy redshift survey from Nancy Grace Roman Space Telescope, Mon. Not. Roy. Astron. Soc. **501**, 3490 (2021), arXiv:2008.09746 [astro-ph.CO].
[318] Z. Zhai, A. Benson, Y. Wang, G. Yepes, and C.-H. Chuang, Prediction of Hα and [OIII] emission line galaxy number counts for future galaxy redshift surveys, Mon. Not. Roy. Astron. Soc. **490**, 3667 (2019), arXiv:1907.09680 [astro-ph.GA].
[319] N. Hand, Y. Feng, F. Beutler, Y. Li, C. Modi, U. Seljak, and Z. Slepian, nbodykit: an open-source, massively parallel toolkit for large-scale structure, Astron. J. **156**, 160 (2018), arXiv:1712.05834 [astro-ph.IM].
[320] Y. P. Jing, Correcting for the alias effect when measuring the power spectrum using FFT, Astrophys. J. **620**, 559 (2005), arXiv:astro-ph/0409240.
[321] E. Sefusatti, M. Crocce, R. Scoccimarro, and H. Couchman, Accurate Estimators of Correlation Functions in Fourier Space, Mon. Not. Roy. Astron. Soc. **460**, 3624 (2016), arXiv:1512.07295 [astro-ph.CO].
[322] A. G. Adame *et al.* (DESI), DESI 2024 VII: cosmological constraints from the full-shape modeling of clustering measurements, JCAP **07**, 028, arXiv:2411.12022 [astro-ph.CO].
[323] A. G. Adame *et al.* (DESI), DESI 2024 V: Full-Shape Galaxy Clustering from Galaxies and Quasars, (2024),

arXiv:2411.12021 [astro-ph.CO].

[324] A. Chudaykin, M. M. Ivanov, and O. H. E. Philcox, Reanalyzing DESI DR1: 1. CDM Constraints from the Power Spectrum and Bispectrum, (2025), arXiv:2507.13433 [astro-ph.CO].

[325] N. Schöneberg, The 2024 BBN baryon abundance update, JCAP **06**, 006, arXiv:2401.15054 [astro-ph.CO].

[326] M. Maus *et al.*, A comparison of effective field theory models of redshift space galaxy power spectra for DESI 2024 and future surveys, JCAP **01**, 134, arXiv:2404.07272 [astro-ph.CO].