UNIVERSITÀ DI PARMA

UNIVERSITÀ DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN FISICA

XXXIV CICLO

Towards a model independent approach to galaxy clustering

Coordinatore:
Chiar.mo Prof. Stefano Carretta
Tutore:
Chiar.mo Prof. Massimo Pietroni

Dottorando:
Marco Marinucci

Anni accademici 2018/2019-2020/2021



Sed, missis terrenis,
ad Coelestium speculationes
me contuli.

G. Galilei, *Sidereus Nuncius*

Acknowledgments

> A pleasure is fully grown only when it is remembered. You are speaking as if the pleasure were one thing and the memory another. It is all one thing. [...] When you and I met, the meeting was over very shortly, it was nothing. Now it is growing something as we remember it. But still we know very little about it. What it will be when I remember it as I lie down to die, what it makes in me all my days till then - that is the real meeting. The other is the beginning of it.
>
> C. S. Lewis, *Out of the silent planet*

First of all, I want to thank my supervisor Massimo Pietroni, a very good master and a dear friend, who have guided me through these years with patience and support. His ability to go straight to the heart of scientific questions together with his dedication and passion for his work, have been a source of inspirationn for me. With his suggestions and his corrections (sometimes a bit hard to digest for me), he always encouraged me to give my best; during all these years I have grown both as a man and as a researcher and I will always be grateful to him. I really hope that our friendly collaboration will last for a long time to come.

I am indebted to INFN and the University of Parma for the opportunity to carry out my PhD and to attend the many workshops and conferences to which I took part.

It has been a pleasure to work with many great and brilliant physicists during these years. I would like to thank Takahiro Nishimichi, my very first collaborator, with whom I have published my first two papers which have become a central part of my research. Thanks to Filippo Vernizzi that accepted to host me at the IPhT in Paris and for the time he dedicated to me: I have learnt a lot from our discussions about theoretical physics. I am grateful to Guido D'Amico,

from which I have learned the importance of details, which play a key role in good research work. Finally, thanks to Emiliano Sefusatti, that welcomed me in Trieste when I was still at the beginning of my career, for the time he spent with me on our shared projects.

I would also like to thank Patrick Valaegias and Zvonimir Vlah for having accepted to read and correct my thesis and for their useful comments that helped improving the quality of this work.

A grateful thanks goes to Vincent Desjaques and Adi Nusser, that welcomed me in the wonderful research group of the Technion Institute.

I wanna thank all the people that I have met during these years: Matteo Viel, Marco Peloso, Martin Crocce, Alkistis Pourtsidou, Maura Pavesi, Bill Wright, Lorenzo Piga, Ben Bose, Santiago Casas, Andrea Oddo, Matteo Biagetti, Davit Alkhanishvili, Eleonora Villa, Alessandro Carones, Viraj Nistane, Pratyush Pranav, Sarah Libanore, Alfredo Guerrera, Giovanni Verza.

I am grateful to my friends Giulio, Luca, Nicola, Emanuele, Matteo, Gianpaolo, Adriano, Sean, Cecilia, Flavia, Marina, Chiara, Giovanni, Andrea, Clarissa, Claudia, Claudio, Damiano, Miriam, Agnese, Saba, Luca, Anna, Pedro, Vanessa, Tommaso, Elisabetta, Sebastiano, Erica, Gloria, Tino, Sonia, Gianni, Elena, Hussam, Chiara, Gigi, Paolo, Carmine, and all those people that, during these years, supported me in the hardest times and shared with me the happiest ones.

Finally, I will always be grateful for my family, babbo e mamma, Chiara, Stefano and Maria, who always accompany me from afar with their love and their support.

A special thanks goes to my wife Chiara, my unwearied life partner. We faced all the difficulties together and we have shared the best moments of our lives. Her love is a special light in my life that helps me see the best part of me, giving me the strength to never give up. Without her, I would be nothing.

Contents

Abstract

Future galaxy redshift surveys will observe the large scale structure (LSS) of the universe with unprecedented precision. As an example, the Euclid satellite [1] planned launch is around the end of 2022. It will map around 1.5 billion of galaxies over a solid angle of 15000 deg^2 and at redshifts up to ~ 2. Euclid will be the one of the greatest galaxy survey and the major source of cosmological information for the next decade.

The analysis of the huge amount of data from these surveys will need theoretical modelling to be both accurate and flexible in order to constrain the cosmological parameters and, possibly, explore models beyond the standard ΛCDM model.

My PhD project focuses on the study of the role played by symmetries, in particular the (extended) galilean invariance, in providing cosmological tests that are as independent as possible from any cosmological model or computational scheme and its related approximation.

Among the results obtained in this program so far, we proposed a new method to measure galaxy bias, b_g, and the linear growth function, f, by combining the bispectrum (three-point correlation function) and power spectrum measurements, [2, 3]. The method is based on so-called *consistency relations* (CR) for the LSS [4, 5, 6, 7], which are non-perturbative statements based on symmetry principles, and therefore can potentially provide a theoretically clean and model-independent determination of b_g and f. We have tested this method against N-body simulations, validating the pipeline and providing accurate estimates of these crucial cosmological observables. We are currently using this CR-based approach within the Euclid mission, applying the procedure formulated in our previous works to the Euclid "flagship simulations", for which the results will be published soon.

More recently, we have developed the "LSS bootstrap" [8], a symmetry-based approach to constrain the kernels that appear in the perturbative expansion. It allows to determine to what extent the analytic structure of the kernels is dictated by symmetries, independently from the particular equations of motion considered. This approach provides a systematic way to construct the kernels at each perturbative order using only symmetry principles, and it can therefore naturally applied to non-standard cosmological models that satisfy the symmetries.

As a first application, I discuss an example of non standard gravity theory, the nDGP model [9], and show the constraining power of the approach.

We are currently working on the implementation of nDGP in the PyBird code [10], a fast code based on the FFTLog technology for the evaluation of the 1-loop integrals for biased tracers in redshift space. Some preliminary results on possible constraints on this model from BOSS-like survey [11] are shown.

The results presented here will manifest their practical value when applied to future generation survey, like the Euclid mission, of which I have been part for all the duration of my PhD studies.

Articles and copyright permissions

The results of this thesis have produced three papers:

1. M. Marinucci, T. Nischimici and M. Pietroni, *Measuring Bias via the Consistency Relations of the Large Scale Structure*, *Phys. Rev.* **D100** (2019), no. 12 123537, doi 10.1103/PhysRevD.100.123537;

2. M. Marinucci, T. Nishimichi and M. Pietroni, *Model independent measurement of the growth rate from the consistency relations of the LSS*, *JCAP* **07** (2020) 054, doi 10.1088/1475-7516/2020/07/054;

3. G. D'Amico, M. Marinucci, M. Pietroni and F. Vernizzi, *The Large Scale Structure Bootstrap: perturbation theory and bias expansion from symmetryes*, *JCAP* **10** (2021) 069, doi 10.1088/1475-7516/2021/10/069.

The first paper was published on the journal Physical Review D. As author of the mentioned article, I have the permission to use it or a portion of it in my thesis or dissertation, as specified by the APS Copyright Policies, that can be found here.

The second and the third paper were published on the Journal of Cosmology and Astrophysics (JCAP). As author of the article I have the permission to use them or a portion of them in my thesis or dissertation, as specified by the IOP's Author rights policy, that can be found here.

Chapter 1

Introduction

The next decade has the full potential of being a scientific golden age for cosmology, since the forthcoming generation of galaxy surveys will sample the large scale structure of the universe (LSS) over large volumes, making it possible to test cosmological models with unprecedented precision [1, 12, 13]. According to the current cosmological standard model, ΛCDM, at primordial times the universe expansion was driven by a scalar field, the inflaton, whose quantum perturbations are the seeds for the formation of primordial density perturbations. After the inflationary period, the inflaton decayed and filled the universe with relativistic particles.This event is usually called *reheating*, after which the standard *hot big bang* phase begun, in which the universe energy budget is dominated first by radiation and then by non relativistic matter.

During the matter-dominated epoch, the structures that we observe in the universe, such as the clusters of galaxies, developed through gravitational instability. At more recent times, around z$\sim O(0.5)$, the dark energy component (Λ) started dominating, causing the present accelerated expansion of the universe.

The ΛCDM model matches with extraordinary agreement the observations from the cosmic microwave background (CMB) [14, 15], the LSS of the late time universe [11] and measurements of the late time accelerated expansion [16, 17].

A LSS feature of particular relevance for precision cosmology is the *baryon acoustic oscillations* (BAO) peak in the two-point correlation function of galaxies, first detected in [18], see also [19, 20, 21]. The BAO peak is a direct signal coming from the interaction between baryonic matter and radiation in the early universe [22, 23, 24, 25]. Since the physical scale of

this sound horizon at recombination is a robust prediction, it can be used as a standard ruler to measure the expansion of the universe, and could give us insights on the nature of dark energy, responsible for the acceleration of the expansion of the universe [26].

All these cosmological observations and theoretical predictions have driven us to understand the composition of our universe at a very precise level, but some intriguing questions and issues remain to be faced. The unknown nature of dark matter and dark energy (which compose the 95% of the total matter-energy of the universe), possible non gaussianity of the primordial perturbations coming from the inflationary period, neutrino masses, and possible deviations from general relativity (GR) are only few examples of the open questions we need to address to complete the puzzle.

The main focus of this thesis will be the study of analytical models for the description of galaxy clustering from linear to mildly non-linear scales. The well established approach for the description of the dark matter field is the standard perturbation theory (SPT, for a review see [27]). This method consists in writing the non-linear fields as a perturbative expansion of initial linear fields, that are convolved, in momentum space, with some PT kernels. These convolution kernels are found by solving perturbatively the equation of motion of the density perturbations. This procedure, truncated at third order (1-loop), fails to reproduce the power spectrum (PS) on non linear scales, $k > k_{\rm NL} \simeq 0.2$ h/Mpc, where k is the wavenumber. Going to higher orders does not improve the situation [28]. Physically, the failure of PT is due to the perfect fluid approximation, which breaks down at small scales (higher k's) and late times (lower z's) due to multistreaming [29, 30, 31, 32, 33, 34, 35]. The effective field theory of LSS (EFTofLSS), see [36, 37, 38, 39], provides an accurate method for the calculation of the effect of non-linear scale physics on linear scales, adding appropriate UV counterterms to the perturbative expansion, incorporating physics beyond the perfect fluid approximation.

Moreover, when considering the galaxy field, which is what we actually observe, one has to take into account many other effects. First, one has to provide a relation between the dark matter perturbation and the galaxy number perturbation: this is usually realized by writing the latter as an expansion in derivatives of the gravitational potential [40]. The coefficients of this expansion are the *bias parameters*, which describe how galaxies trace the underlying potential, sourced by dark matter. Moreover one has to take into account the so-called *redshift*

space distortions. We infer the distance of a source that emits light by using the Hubble law, after having measured the redshift of the source. The peculiar velocity of galaxies add up to the Hubble velocity producing distortions in 3D maps of the sky obtained by redshift surveys. For a complete review of all these effects see [41].

The practical applicability of the EFTofLSS approach has been greatly enhanced by Fast Fourier Transform (FFT) methods to evaluate the momentum integrals appearing in the computation of the PS beyond linear order [42, 43, 44], which allow a fast and accurate implementation of Markov Chain Monte Carlo algorithms for cosmological parameter estimation. The EFTofLSS has recently been applied to the BOSS dataset [11] to extract cosmological parameters [45, 46], see also [47], leading to cosmological constraints independent from CMB observations. Moreover, within the BOSS collaboration, the bispectrum (BS) has been measured and analyzed, see [48, 49, 50].

All these recent results show, on one side, the great accuracy of the ΛCDM model in describing the history of the universe from primordial recent times. On the other side, these accurate predictions will need an increasing computational effort, given the accuracy required and the huge amount of data from future LSS observations.

In this thesis we explore the role played by symmetries in providing model independent predictions for the LSS.

A first step toward this direction was provided by [4, 5] were the Consistency Relations for the LSS (CR's) were derived. The content of the CR's is essentially kinematic: once the equivalence principle (EP) is assumed, they describe the effect of a long wavelength displacement field on short distance fluctuations. Therefore they are completely general and not based on any particular cosmological model or approximation scheme.

The simplest CR relates the squeezed limit of the BS, the configuration in which one momentum is taken much smaller that the other two, to the logarithmic derivative of the PS evaluated at the large momentum scale. The coefficient in front of the logarithmic derivative of the PS is protected against non-linearities by the equivalence principle. For a featureless PS, the protected term is of the same order of other, un-protected ones. However, as we showed in [2, 3], the wiggly BAO feature imprinted in LSS spectra provides a way to isolate

the CR-protected contribution to the BS from unprotected ones. The latter, although being parametrically of the same order in the squeezed limit, are either smooth or suppressed by factors $O(2\pi/(kr_{BAO}))$ with respect to the protected ones, where r_{BAO} is the BAO acoustic scale and k is the wavenumber of the short scale considered. Therefore, by comparing the BAO amplitudes in the BS and in the PS we were able to measure the CR coefficient, which depends on the linear bias parameter b_α, and, in redshift space, also on and the growth function, $f = d\ln D/d\ln a$. We discuss this procedure in detail, and present a pipeline for the real data analysis, in view of the Euclid mission.

Moving ahead in the direction of providing symmetry-based cosmological predictions, we have explored the role played by symmetries in dictating the structure of the PT kernels, independently on the details of the equations of motion: we called this approach the *bootstrap* of LSS [8], as it is similar in spirit to bootstrap approaches in field theory [51] and inflation [52, 53]. This is particularly relevant when considering models beyond ΛCDM, such as for dark energy and modified gravity. In particular, we will show that the structures of the kernels in all dark energy and modified gravity models with the same symmetries as ΛCDM are the same. Only a few time-dependent coefficients change, depending on the particular expansion history and possible time-dependence of the modified gravity parameters.

One symmetry plays a crucial role in fixing the structure of the kernels both for dark matter and for generic biased tracers: our results are based on the assumption that the dynamics is invariant under uniform spatial translations with arbitrary time dependence, which follows from diffeomorphism invariance in the non-relativistic limit [6]. We will refer to this as *extended* Galilean invariance (EGI) [54, 55] to distinguish it from the usual Galilean invariance relating inertial frames by a boost. This requirement translates in constraints on the analytic structure of the PT kernels.

We identify a hierarchy of EGI constraints, depending on the PT order of the induced displacement field. The order of the constraints is dictated by the number of momenta for which the sum vanishes. The *leading order* (LO) constraints, for example, fix the pole structure of the kernels as one of its external momenta goes to zero. The *next-to-leading* (NLO) constraints dictate the kernel structure as the sum of two external momenta vanishes, and so on.

In the case of biased tracers we end up with seven independent coefficients, equivalently

to the other bias expansions discussed in the literature for fixed cosmologies.

This approach provides a useful tool in two respects. On one hand, the deep role played by EGI provides a systematic way to derive the structure of the PT kernels for biased tracers at orders beyond the third, when NNLO, and so on, constraints start to play a role. On the other hand, the framework presented is optimal if one is interested in model-independent analyses of beyond ΛCDM scenarios, since, as we will see, we will be able to single out coefficients which depend only on cosmology and not on the tracer type.

We will investigate a particular modified gravity model, the nDGP model [9]. Here it is assumed that our universe is embedded in a 4-dimensional brane of a higher (5-)dimensional spacetime. The scale r_c at which the five-dimensional gravity start behaving as the four-dimensional one is called crossover scale. In this model a cosmological constant causing the accelerated expansion is provided by assuming this modification of gravity on large scales and it respects the same symmetries as ΛCDM.

The thesis is structured as follows. In chapter 2 we introduce SPT, with a detailed description of the perturbative scheme in the EdS approximation. A brief introduction to EFTofLSS follows. We will then illustrate the connection between the dark matter density field and the observed galaxy density field. Finally, the expression for the one-loop PS and the tree-level BS will be derived. Then, in chapter 3, we present a generalization of this approach to models with an exact time dependence beyond the EdS approximation. We will briefly review the Green's function approach and show some results from MCMC analysis of N-body data. We will introduce the nDGP model and present the novel implementation of this model within the public code PyBird [10], which will be used for MCMC analysis. In chapter 4 we will explore the use of CR's for model independent parameter estimation and show the results obtained against N-body simulations. We also provide a well-based pipeline for the extraction of the BS in the squeezed configuration and the measurement of the prefactor of the CR's. This pipeline will be applied to the Euclid flagship simulations, both in real and redshift space, and it will then be used for the future data analysis. Chapter 5 presents the full formalization of the LSS bootstrap and highlights the advantages of a new approach based only on symmetry. We conclude in chapter 6 with some remarks and possible future developments of this research.

Chapter 2

Perturbation theory and large scale structure

In this first chapter, we present and discuss the standard paradigm for the study of the large scale structure (LSS) of the universe. We will introduce the linear and non-linear perturbation theory of dark matter overdensity, and we will outline the standard procedure to construct actual observables. Then, we will discuss the connection between dark matter field and the galaxy density field, introducing observational effects such as the galaxy biasing, redshift space distortions, small scales effects and the correct modelling of the BAO feature. We will conclude with a short review of the state of the art.

2.1 Standard perturbation theory

The starting point is considering the dynamics of the dark matter particles within a FLRW background metric

$$ds^2 = -dt^2 + a^2(t)\delta_{ij}dx^i dx^j \,, \tag{2.1}$$

where t is the cosmic time as measured by a comoving observer. In the following we will make use of a new time variable, the conformal time τ defined as $a(t)dt = d\tau$.

Following [27] we consider a collection of dark matter particles of mass m in the non-relativistic limit and interacting only via gravitational interaction. The equation of motion for

a particle of velocity $\mathbf{v}$ at position $\mathbf{r}$ reads

$$\frac{d\mathbf{v}}{dt} = G\,m\,\sum_i \frac{\mathbf{r}_i - \mathbf{r}}{|\mathbf{r}_i - \mathbf{r}|^3}, \tag{2.2}$$

which can be rewritten, assuming a large number of particles, in terms of a smooth gravitational potential caused by the particle distribution,

$$\frac{d\mathbf{v}}{dt} = -\nabla_{\mathbf{r}}\phi, \qquad \phi(\mathbf{r}) = G\int d^3\mathbf{r}'\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \tag{2.3}$$

with ρ being the dark matter density distribution. From eq. (2.1) we can now write the Friedmann equations in terms of the conformal time τ

$$(\Omega_m + \Omega_\Lambda - 1)\mathcal{H}^2(\tau) = k, \tag{2.4}$$

$$\frac{\partial\mathcal{H}(\tau)}{\partial\tau} = -\frac{\Omega_m(\tau)}{2}\mathcal{H}^2(\tau) + \frac{\Lambda}{3}a^2(\tau) \equiv \left(\Omega_\Lambda - \frac{\Omega_m}{2}\right)\mathcal{H}^2, \tag{2.5}$$

where k is the space curvature, that we always consider to be null, and $\mathcal{H}(\tau) = d\ln a(\tau)/d\tau$ is the conformal Hubble parameter. The functions Ω_m and Ω_Λ are, respectively the matter and the cosmological constant energy density parameters. They are defined as the ratio between the matter energy density and the critical energy density, i.e. $\Omega_m(\tau) \equiv \rho_m(\tau)/\rho_{\rm cr}(\tau)$, and analogously for the cosmological constant. Notice that in eq. (2.4) and (2.5) we are not considering relativistic components of the universe, like radiation and neutrinos, because we will consider the matter perturbations in the matter domination epoch (also called Einstein-de Sitter model). We will discuss later the accuracy of this prescription, often adopted in LSS.

We define the dark matter overdensity $\delta(\mathbf{x})$ as

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

where $\bar{\rho}$ indicates the mean background density; we also define the peculiar velocity

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

Finally, we can split cosmological gravitational potential Φ splitting into the one generated by the background homogeneous matter distribution and the one sourced by the density fluctuations, ϕ,

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

where ϕ is normalized in such a way to satisfy the Poisson equation

$$\nabla^2\phi(\mathbf{x},\tau) = \frac{3}{2}\Omega_m(\tau)\mathcal{H}^2(\tau)\delta(\mathbf{x},\tau). \tag{2.9}$$

The evolution of the matter field is described by its phase space density function $f(\mathbf{x},\mathbf{p};\tau)$, where $\mathbf{p} = am\mathbf{u}$ is the comoving momentum perturbation. The equation governing the phase space density function evolution is the collisionless Boltzmann equation

$$\frac{df(\mathbf{x},\mathbf{p};\tau)}{d\tau} = \frac{\partial f}{\partial\tau} + \frac{\partial\mathbf{x}}{\partial\tau}\nabla_{\mathbf{x}}f + \frac{\partial\mathbf{p}}{\partial\tau}\nabla_{\mathbf{p}}f = 0\,, \tag{2.10}$$

which can be rewritten, using

$$\frac{d\mathbf{p}}{d\tau} = -am\nabla_{\mathbf{x}}\phi(\mathbf{x},\tau)\,, \tag{2.11}$$

as the well known Vlasov equation

$$\left(\frac{\partial}{\partial\tau} + \frac{p^i}{am}\frac{\partial}{\partial x^i} - am\nabla^i_{\mathbf{x}}\phi(\mathbf{x},\tau)\frac{\partial}{\partial p^i}\right) f = 0\,. \tag{2.12}$$

In writing eq. (2.12) we have assumed GR in the sub-horizon, or Newtonian, limit. Eq. (2.12) is in general hard to solve, even numerically. The general approach used in literature is to consider the evolution equations of the *moments* of the phase space function, defined as

$$\begin{aligned}
\rho(\mathbf{x},\tau) &= \int d^3\mathbf{p}\, f(\mathbf{x},\mathbf{p},\tau)\,,\\
\rho(\mathbf{x},\tau)u^i(\mathbf{x},\tau) &= \int d^3\mathbf{p}\,\frac{p^i}{am} f(\mathbf{x},\mathbf{p},\tau)\,,\\
\rho(\mathbf{x},\tau)\sigma^{ij}(\mathbf{x},\tau) &= \int d^3\mathbf{p}\,\frac{p^i}{am}\frac{p^j}{am} f(\mathbf{x},\mathbf{p},\tau) - u^i(\mathbf{x},\tau)u^j(\mathbf{x},\tau)\rho(\mathbf{x},\tau)\,.
\end{aligned} \tag{2.13}$$

These first three moments are, respectively, the matter density, the velocity and the velocity dispersion fields. Note that the number of possible moments is, in theory, infinite, as one can add to eq. (2.13) an infinite hierarchy of moments, see [39]. This becomes even more clear when taking different moments of eq. (2.12)

$$\begin{aligned}
&\frac{\partial\delta(\mathbf{x},\tau)}{\partial\tau} + \frac{\partial}{\partial x^i}\left[(1+\delta(\mathbf{x},\tau))\,u^i(\mathbf{x},\tau)\right] = 0\,,\\
&\frac{\partial u^i(\mathbf{x},\tau)}{\partial\tau} + \mathcal{H}(\tau)u^i(\mathbf{x},\tau) + u^j(\mathbf{x},\tau)\frac{\partial}{\partial x^j}u^i(\mathbf{x},\tau) =\\
&\qquad\qquad - \frac{\partial}{\partial x^i}\phi(\mathbf{x},\tau) - \frac{1}{\rho(\mathbf{x},\tau)}\frac{\partial}{\partial x^j}\left(\rho(\mathbf{x},\tau)\sigma^{ij}(\mathbf{x},\tau)\right)\,,\\
&\cdots\,,
\end{aligned} \tag{2.14}$$

where the dots indicate the equations for the higher moments. The first equation of (2.14) is the continuity equation and it couples the zeroth moment of the distribution function, i.e. ρ, to the first moment, the velocity. Moreover, the second equation is the Euler equation and couples the velocity to the velocity dispersion σ^{ij}, which is the second moment of the distribution function. In general, each of these equations describes the time evolution of the n-th moment, which is coupled to the $n+1$-th moment; this means that we need a prescription in order to truncate this hierarchy of equations. For instance, is we consider a negligible velocity dispersion $\sigma^{ij} = 0$ the continuity and Euler equations, together with the Poisson equation, form a closed system. Since the velocity dispersion quantifies the spread of the velocities of the particles at a certain position around the mean velocity $\mathbf{u}$, setting it to 0 means that the fluid is characterized by one single flow (or *stream*) and that at every position there is only one value for the particle velocity. This approximation, denoted as *single stream approximation*, is known to be valid for large scales and early times, while it breaks down after the particle trajectories cross each other, the so called *shell crossing*, and velocities start to virialize. Taking into account multi streaming phenomena is very difficult also for dark matter, since highly non-linear and non-perturbative physics starts playing an important role. Some attempts has been done in this direction, see for example [29, 30, 31, 32, 33, 34, 35].

Let us investigate a bit more eq. (2.12). We are implicitly assuming that eq. (2.12) is the evolution equation for the *microscopic* distribution function of the dark matter particles, that can be expressed as a sum of Dirac delta functions

$$f(\mathbf{x}, \mathbf{p}; \tau) = \sum_n \delta_D \left(\mathbf{x} - \mathbf{x}_n(\tau)\right) \delta_D \left(\mathbf{p} - \mathbf{p}_n(\tau)\right) \,, \tag{2.15}$$

where the sum runs over all the N particles. In cosmology we are actually interested in scales much larger than the mean inter-particle distance. Therefore, we can smooth the microscopic distribution function by averaging over a resolution scale [37]

$$\bar{f}(\mathbf{x}, \mathbf{p}; \tau, R) \equiv \frac{1}{V} \int d^3\mathbf{y}\, W\left(\left|\frac{\mathbf{y}}{R}\right|\right) f(\mathbf{x} + \mathbf{y}, \mathbf{p}; \tau) \,, \tag{2.16}$$

where $W(|\mathbf{x}|)$ is a smoothing window function normalized to unity

$$\frac{1}{V} \int d^3\mathbf{y}\, W\left(\left|\frac{\mathbf{y}}{R}\right|\right) = 1 \,. \tag{2.17}$$

Applying the filtering operation to the Vlasov equation, we obtain

$$\left(\frac{\partial}{\partial\tau}+\frac{p^i}{am}\frac{\partial}{\partial x^i}-am\nabla^i_{\mathbf{x}}\bar{\phi}(\mathbf{x},\tau;R)\frac{\partial}{\partial p^i}\right)\bar{f}(\mathbf{x},\mathbf{p};\tau,R)=C[f](\mathbf{x},\mathbf{p};\tau)\,, \tag{2.18}$$

where $C[f]$ is the (pseudo-)collisional term generated by the smoothing operation

$$C[f](\mathbf{x},\mathbf{p};\tau,R)=am\langle\nabla^i_{\mathbf{x}}\phi(\mathbf{x},\tau)\frac{\partial}{\partial p^i}f(\mathbf{x},\mathbf{p},\tilde{a})\rangle-am\nabla^i_{\mathbf{x}}\bar{\phi}(\mathbf{x},\tau)\frac{\partial}{\partial p^i}\bar{f}(\mathbf{x},\mathbf{p},\tau)\,, \tag{2.19}$$

where we have defined

$$\langle g\rangle(\mathbf{x})\equiv\int d^3\mathbf{y}\,W\left(\left|\frac{\mathbf{y}}{R}\right|\right)g(\mathbf{x}-\mathbf{y})\,. \tag{2.20}$$

We can define the first three moments of the smoothed distribution function

$$\begin{aligned}
\bar{\rho}(\mathbf{x},\tau)&=\int d^3\mathbf{p}\,\bar{f}(\mathbf{x},\mathbf{p},\tau)=\rho_0(\tau)(1+\bar{\delta}(\mathbf{x},\tau))\,,\\
\bar{\rho}(\mathbf{x},\tau)\bar{u}^i(\mathbf{x},\tau)&=\int d^3\mathbf{p}\,\frac{p^i}{am}\bar{f}(\mathbf{x},\mathbf{p},\tilde{a})\,,\\
\bar{\rho}(\mathbf{x},\tau)\bar{\sigma}^{ij}(\mathbf{x},\tau)&=\int d^3\mathbf{p}\,\frac{p^i}{am}\frac{p^j}{am}\bar{f}(\mathbf{x},\mathbf{p},\tau)-\bar{u}^i(\mathbf{x},\tau)\bar{u}^j(\mathbf{x},\tau)\bar{\rho}(\mathbf{x},\tau)\,,
\end{aligned} \tag{2.21}$$

where ρ_0 is the spatial average of $\rho(\mathbf{x},\tau)$ introduced in eq. (2.14). The (smoothed) density perturbation satisfy the Poisson equation

$$\nabla^2_{\mathbf{x}}\bar{\phi}(\mathbf{x},\tau)=\frac{3}{2}\mathcal{H}^2(\tau)\Omega_m(\tau)\bar{\delta}(\mathbf{x},\tau)\,. \tag{2.22}$$

The first two moments of the filtered Vlasov equation are

$$\frac{\partial}{\partial\tau}\bar{\delta}(\mathbf{x},\tau)+\frac{\partial}{\partial x^i}\left[\left(1+\bar{\delta}(\mathbf{x},\tau)\right)\bar{u}^i(\mathbf{x},\tau)\right]=0 \tag{2.23}$$

$$\frac{\partial}{\partial\tau}\bar{u}^i(\mathbf{x},\tau)+\mathcal{H}(\tau)\bar{u}^i(\mathbf{x},\tau)+\bar{u}^j(\mathbf{x},\tau)\frac{\partial}{\partial x^j}u^i(\mathbf{x},\tau)=-\frac{\partial}{\partial x^i}\bar{\phi}(\mathbf{x},\tau)-J^i\,, \tag{2.24}$$

where the new source in the Euler equation can be written as

$$J^i=J^i_\phi+J^i_\sigma\,, \tag{2.25}$$

with

$$\begin{aligned}
J^i_\phi(\mathbf{x},\tau)&=\frac{1}{\bar{\rho}}\langle(1+\delta)\nabla^i\phi\rangle(\mathbf{x},\tau)-\nabla^i\bar{\phi}(\mathbf{x},\tau)\,,\\
J^i_\sigma(\mathbf{x},\tau)&=\frac{1}{\bar{\rho}(\mathbf{x},\tau)}\frac{\partial}{\partial x^j}\left(\bar{\rho}(\mathbf{x},\tau)\bar{\sigma}^{ij}(\mathbf{x},\tau)\right)\,.
\end{aligned} \tag{2.26}$$

Eq. (2.26) suggests what is the nature of the different source terms. J^i_σ accounts for the effects due to the (smooth) velocity dispersion $\bar{\sigma}_{ij}$, which was present also in eq. (2.14). J^i_ϕ appears due to the filtering operation on the gravitational potential ϕ, and it is just the first moment of the collisional term in eq. (2.19), which disappears when the microscopic distribution function is considered. These two sources describe the effect of the smoothing operation on linear scales, i.e., when the small scale non-linearities are integrated. For the rest of this section we will not consider this effect and we will find solutions for the system in eq. (2.14). In section (2.2) we will briefly introduce the effective field theory of LSS (EFTofLSS), which provides a rigorous approach to take into account the contributions coming from small-scale perturbations.

We are now ready to solve perturbatively the system in eq. (2.14). From now on we will use the new time variable

$$d\eta = d\ln a = \mathcal{H} d\tau \,, \tag{2.27}$$

and we will rewrite the equations of motion in Fourier space using the notation

$$\delta(\mathbf{x},\eta) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{-i\mathbf{k}\cdot\mathbf{x}} \tilde{\delta}(\mathbf{k},\eta) \,. \tag{2.28}$$

In general, the velocity field can always be written as the sum of a divergence and a vorticity component

$$u^i(\mathbf{k},\eta) = -i\frac{k^i}{k^2} f_+(\eta)\mathcal{H}(\eta)\theta(\mathbf{k},\eta) + i\epsilon_{ijk}\frac{k^j}{k^2} w^k(\mathbf{k},\eta) \,, \tag{2.29}$$

where we have defined the (rescaled) velocity divergence as

$$\theta(\mathbf{x},\eta) = -\frac{\partial_i u^i}{f_+\mathcal{H}} \,, \qquad \theta(\mathbf{k},\eta) = i\frac{\mathbf{k}\cdot\mathbf{u}_k}{f_+\mathcal{H}} \,, \tag{2.30}$$

and the vorticity component as $w_k(\mathbf{k},\eta) = i\epsilon_{ijk}k^i u^j$. In particular, one can see that the latter is only sourced by the velocity dispersion σ_{ij}, meaning that we can neglect it in the following and consider it only to higher perturbative orders, see [37]. The definition of the velocity divergence, eq. (2.30), includes the (scale-independent) growth function, which is defined as the logarithmic derivative of the linear growing function

$$\delta_L(\mathbf{x},\eta) \equiv D_+(\eta)\delta_L(\mathbf{x},\eta_{\rm in}) \,, \quad \text{and} \quad f_+(\eta) \equiv \frac{d\ln D_+(\eta)}{d\eta} \,, \tag{2.31}$$

where $D_+(\eta)$ describes the time evolution of the *linear* density field δ_L, as we will show soon. The function D_+ describes the time evolution of the linear perturbations and it depends both on the cosmological model and the theory of gravity considered.

The system in eq. (2.14) reads, in Fourier space

$$\begin{aligned}
&\partial_\eta \delta_{\mathbf{k}} - f_+ \theta_{\mathbf{k}} = f_+ \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2} \alpha(\mathbf{q}_1,\mathbf{q}_2) \theta_{\mathbf{q}_1} \delta_{\mathbf{q}_2} \,, \\
&\partial_\eta \theta_{\mathbf{k}} - f_+ \theta_{\mathbf{k}} - \frac{3}{2}\frac{\Omega_m}{f_+}\left(\delta_{\mathbf{k}} - \theta_{\mathbf{k}}\right) = f_+ \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2} \beta(\mathbf{q}_1,\mathbf{q}_2) \theta_{\mathbf{q}_1}, \theta_{\mathbf{q}_2} \,,
\end{aligned} \tag{2.32}$$

where we have omitted the tilde and the time dependence to avoid clutter. In eq. (2.32) we have defined, using $\mathbf{q}_{1\ldots n} = \mathbf{q}_1 + \cdots + \mathbf{q}_n$

$$\mathcal{I}_{\mathbf{k},\mathbf{q}_1,\ldots,\mathbf{q}_n} \equiv \int \frac{d^3\mathbf{q}_1}{(2\pi)^3} \cdots \int \frac{d^3\mathbf{q}_n}{(2\pi)^3} \delta_D(\mathbf{k} - \mathbf{q}_{1\ldots\mathbf{q}_n}) \,. \tag{2.33}$$

Note that in eq. (2.32) the Poisson equation (2.9) has been used inside the Euler equation, and we have defined the two functions

$$\alpha(\mathbf{q}_1,\mathbf{q}_2) \equiv \frac{\mathbf{q}_1 \cdot (\mathbf{q}_1 + \mathbf{q}_2)}{q_1^2} \,, \qquad \beta(\mathbf{q}_1,\mathbf{q}_2) \equiv \frac{|\mathbf{q}_1 + \mathbf{q}_2|^2 \, \mathbf{q}_1 \cdot \mathbf{q}_2}{2\mathbf{q}_1^2 \mathbf{q}_2^2} \,. \tag{2.34}$$

These two functions encode the non-linear coupling between different modes in the evolution of the density and the velocity fields. Considering only the linear equations

$$\begin{aligned}
&\partial_\eta \delta_{\mathbf{k}} - f_+ \theta_{\mathbf{k}} = 0 \,, \\
&\partial_\eta \theta_{\mathbf{k}} - f_+ \theta_{\mathbf{k}} - \frac{3}{2}\frac{\Omega_m}{f_+}\left(\delta_{\mathbf{k}} - \theta_{\mathbf{k}}\right) = 0 \,,
\end{aligned} \tag{2.35}$$

and considering the time evolution for the (growing) linear density field as

$$\delta_L(\mathbf{k};\eta) = D_+(\eta) \delta_L(\mathbf{k};\eta_{\rm in}) \,, \tag{2.36}$$

we obtain the well-known relation between the linear density field and the linear velocity divergence: $\delta_L(\mathbf{k};\eta) = \theta_L(\mathbf{k};\eta)$. Then, we can look at the evolution equation for the linear growth function D

$$\frac{\partial^2 D}{\partial \eta^2} + \left(1 + \frac{\partial \ln \mathcal{H}}{d\eta}\right)\frac{\partial D}{\partial \eta} - \frac{3}{2}\Omega_m D = 0 \,. \tag{2.37}$$

If we consider a Einstein-de Sitter universe (EdS) completely filled with non-relativistic pressureless matter, i.e. $\Omega_m = 1$ and $\Omega_\Lambda = 0$, we find two solutions of the the type $D(\eta) \propto e^{s\eta}$,

with $s = 1$ representing the *growing mode*, since the growth factor grows with the scale factor $D_+ \sim a$, and, analogously, $s = -3/2$ is the *decaying mode*, with $D_- \sim a^{-3/2}$.

We are now ready to consider the nonlinear case. The usual prescription is to expand the density and velocity fields around their linear solutions. Here we present the solutions in the EdS case, the simplest one in which we can find exact perturbative solutions. We can expand as

$$\delta_{\mathbf{k}}(\eta) = \sum_{n=1}^{\infty} e^{n\eta} \delta_{\mathbf{k}}^{(n)} \,, \qquad \theta_{\mathbf{k}}(\eta) = -\mathcal{H}(\eta) \sum_{n=1}^{\infty} e^{n\eta} \theta_{\mathbf{k}}^{(n)}(\eta) \,, \tag{2.38}$$

where $n = 1$ represents the linear solution shown in eq. (2.31), $\delta^{(1)} = \theta^{(1)} = \delta_L$. We can now write a perturbative solution of the system in eq. (2.32)

$$\delta_{\mathbf{k}}^{(n)} = \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\ldots,\mathbf{q}_n} F_n(\mathbf{q}_1, \ldots, \mathbf{q}_n) \delta_L(\mathbf{q}_1) \ldots \delta_L(\mathbf{q}_n) \,, \tag{2.39}$$

$$\theta_{\mathbf{k}}^{(n)} = \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\ldots,\mathbf{q}_n} G_n(\mathbf{q}_1, \ldots, \mathbf{q}_n) \delta_L(\mathbf{q}_1) \ldots \delta_L(\mathbf{q}_n) \,. \tag{2.40}$$

F_n and G_n are homogeneous function of the momenta $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ and are constructed from the fundamental mode coupling functions $\alpha(\mathbf{q}_1, \mathbf{q}_2)$ and $\beta(\mathbf{q}_1, \mathbf{q}_2)$, defined in eq. (2.34), according to the recursion relation (valid for $n \geq 1$)

$$\begin{aligned} F_n(\mathbf{q}_1, \ldots, \mathbf{q}_n) = &\sum_{m=1}^{n-1} \frac{G_m(\mathbf{q}_1, \ldots, \mathbf{q}_m)}{(2n+3)(n-1)} [(2n+1)\alpha(\mathbf{q}_{1\ldots m}, \mathbf{q}_{m+1\ldots n}) F_{n-m}(\mathbf{q}_{m+1}, \ldots, \mathbf{q}_n) \\ &+ 2\beta(\mathbf{q}_{1\ldots m}, \mathbf{q}_{m+1\ldots n}) G_{n-m}(\mathbf{q}_{m+1}, \ldots, \mathbf{q}_n)] \,, \end{aligned} \tag{2.41}$$

$$\begin{aligned} G_n(\mathbf{q}_1, \ldots, \mathbf{q}_n) = &\sum_{m=1}^{n-1} \frac{G_m(\mathbf{q}_1, \ldots, \mathbf{q}_m)}{(2n+3)(n-1)} [3\alpha(\mathbf{q}_{1\ldots m}, \mathbf{q}_{m+1\ldots n}) F_{n-m}(\mathbf{q}_{m+1}, \ldots, \mathbf{q}_n) \\ &+ 2n\beta(\mathbf{q}_{1\ldots m}, \mathbf{q}_{m+1\ldots n}) G_{n-m}(\mathbf{q}_{m+1}, \ldots, \mathbf{q}_n)] \,. \end{aligned} \tag{2.42}$$

For $n = 2$ we have the following symmetrized expressions

$$F_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) = \frac{5}{7} + \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{2q_1 q_2} \left(\frac{q_1}{q_2} + \frac{q_2}{q_1} \right) + \frac{2}{7} \frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)^2}{q_1^2 q_2^2} \,, \tag{2.43}$$

$$G_2^{(s)}(\mathbf{q}_1, \mathbf{q}_2) = \frac{3}{7} + \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{2q_1 q_2} \left(\frac{q_1}{q_2} + \frac{q_2}{q_1} \right) + \frac{4}{7} \frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)^2}{q_1^2 q_2^2} \,. \tag{2.44}$$

It is possible to perform a fully analogous calculation for a general cosmology, considering the exact time dependence, as we will show in chapter 3.

The quantity we are interested in in cosmology is the power spectrum (PS) of the matter perturbation defined as

$$\langle \delta_{\mathbf{k}}(\eta)\delta_{\mathbf{k}'}(\eta)\rangle \equiv (2\pi)^3\delta_D(\mathbf{k}+\mathbf{k}')P(k,\eta)\,. \tag{2.45}$$

If we substitute the perturbative expansion of eq. (2.38) and eq. (2.39) in the expression for the PS, the expectation value becomes a perturbative expansion in the initial n-point correlation functions, $\langle\delta(\mathbf{q}_1,\eta_{in})\dots\delta(\mathbf{q}_n,\eta_{in})\rangle$. Assuming gaussianity for the initial conditions, we can apply the Wick's theorem, which ensures that the expectation values of an odd number of fields (with zero mean) vanish, while those of an even number of fields reduce to a sum of products of initial PS's, explicitly

$$\begin{aligned}
&\langle\varphi(\mathbf{q}_1)\dots\varphi(\mathbf{q}_{2p+1})\rangle = 0\,,\\
&\langle\varphi(\mathbf{q}_1)\dots\varphi(\mathbf{q}_{2p})\rangle = \sum_{\text{all contractions}}\ \prod_{\text{pairs (i,j)}}\langle\varphi(\mathbf{q}_i)\varphi(\mathbf{q}_j)\rangle.
\end{aligned} \tag{2.46}$$

This means that the first non linear contributions to the PS will contain four initial fields $\delta(\mathbf{k},\eta_{in})$, that is, two initial PS. Explicitly we have

$$P(k,\eta) = e^{2\eta}P_L(k) + e^{4\eta}P_{1-\text{loop}}(k) + \dots\,, \tag{2.47}$$

where we have defined the (initial) *linear power spectrum*

$$\langle\delta_L(\mathbf{k},\eta_{\text{in}})\delta_L(\mathbf{k}',\eta_{\text{in}})\rangle = (2\pi)^3\delta_D(\mathbf{k}+\mathbf{k}')P_L(k)\,, \tag{2.48}$$

and the 1-loop PS

$$P_{1-\text{loop}}(k) = P_{22}(k) + P_{13}(k)\,, \tag{2.49}$$

with

$$P_{22}(k) = 2\int\frac{d^3\mathbf{q}}{(2\pi)^3}\left[F_2^{(s)}(\mathbf{k}-\mathbf{q},\mathbf{q})\right]^2 P_L(q)P_L(|\mathbf{k}-\mathbf{q}|)\,, \tag{2.50}$$

and

$$P_{13}(k) = 6\int\frac{d^3\mathbf{q}}{(2\pi)^3}F_3^{(s)}(\mathbf{k},\mathbf{q},-\mathbf{q})\,P_L(q)P_L(k)\,. \tag{2.51}$$

It is possible to go beyond the 1-loop correction, and consider also 2-loop corrections and so on. This approach, standard perturbation theory (SPT), is know to perform well at large scales ($k \leq 0.15$ h/Mpc) and for early redshift ($z \geq 2$).

As you can see in eq. (2.47) in the EdS approximation the time dependence factorizes, and the n-th perturbative order evolves like $\sim a^n = e^{n\eta}$. The usual prescription adopted for the generalization to ΛCDM model is to substitute the scale factor with the growth function $a \to D_+$. This approximation has been tested to be accurate to better than 1% for mildly non-linear scales, see [56].

We will also consider the BS, defined as

$$\langle \delta_{\mathbf{k}_1}(\eta)\delta_{\mathbf{k}_2}(\eta)\delta_{\mathbf{k}_3}(\eta)\rangle \equiv (2\pi)^3\delta_D(\mathbf{k}_{123})B(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3;\eta)\,. \tag{2.52}$$

The interest in the BS is increasing, due to the possibility to measure it in future large scale survey [57, 58, 59, 60, 61, 62] (see [50, 63] for measurements of the BS with the BOSS observations, for other novel approaches to BS see [64, 65, 66, 67]). Substituting the perturbative solution for the matter density eq. (2.39), and assuming again gaussian initial condition for the linear field we obtain the leading order (LO) expression for the matter BS

$$B_{\rm LO}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3;\eta) = 2e^{2\eta}F_2^s(\mathbf{k}_1,\mathbf{k}_2)P_L(k_1)P_L(k_2) + 2 \text{ perms.}\,. \tag{2.53}$$

One is tempted to think that going to higher orders would progressively improve the accuracy of the SPT approach. Ref. [28] presents the calculation up to 3-loop order and showed that at redshift smaller than ~ 1.75 the agreement of the 3-loop correction with N-body simulations is even worse than the 2-loop, showing that the series is not convergent. At low redshift and small scales, the single stream approximations no longer holds, meaning that we have to take into account also the complicated multi streaming physics happening at these scales. In the next section we will describe a way to systematically include these unknown short distance (UV) effects on the computation of the PS at intermediate scales.

2.2 UV counterterms and Effective Field Theories

Our starting point is the smoothed Vlasov eq. (2.18) leading to the continuity and Euler equations (2.21). As we can see, in the Euler equation a source term appears, given in eq. (2.25) [37]. It can be interpreted as $-\partial_j\bar{\tau}^{ij}/\bar{\rho}$, where $\bar{\tau}^{ij}$ is the (smooth) effective stress tensor which carries the pressure and the viscosity forces induced by the small-scales fluctuations, see [36, 38]. The EFTofLSS is based on an expansion of $\bar{\tau}^{ij}$ over long wavelength field.

$$\langle \bar{\tau}_{ij} \rangle_{\bar{\delta}} = p_b \delta_{ij} + \rho_0 \left[c_s^2 \bar{\delta} \delta_{ij} - \frac{c_{bv}^2}{\mathcal{H}} \delta_{ij} \nabla_{\mathbf{x}}^k \bar{u}_k - \frac{3}{4} \frac{c_{sv}^2}{\mathcal{H}} \left(\nabla_{\mathbf{x}}^i \bar{u}_j + \nabla_{\mathbf{x}}^j \bar{u}_i - \frac{3}{2} \delta_{ij} \nabla_{\mathbf{x}}^k \bar{u}_k \right) \right] + \dots , \tag{2.54}$$

where the form of the expansion is dictated by rotational symmetry and the equivalence principle. In eq. (2.54), p_b, c_s, c_{bv} and c_{sv} are, respectively, the effective pressure, the adiabatic speed of sound, the bulk viscosity coefficient, and the shear viscosity coefficient, which depend on the smoothing scale R, see eq. (2.16). These function are not fixed by the SPT and they need to be measured from, for example, N-body simulations. In principle, one could calculate them numerically following a peak-background split argument, see [41]. The first EFT contribution to the density perturbation appears at third order

$$\delta^{\mathrm{ct},(3)}(\mathbf{k}) = -c(\eta, R) D^3(\eta) \frac{k^2}{k_{\mathrm{NL}}^2} W_R(\mathbf{k}) \delta^{(1)}(\mathbf{k}, \eta) \, , \tag{2.55}$$

where D is the linear growth function, $W_R(\mathbf{k})$ is the Fourier transform of the window function introduced in eq. (2.16) and k_{NL} is the non-linear scale. The effective coefficient c corresponds to an effective sound speed, with $\lim_{R \to 0} c = c_s^2$. The non-linear scale can be defined as the scale where the dimensionless linear matter power spectrum becomes unity [41]

$$\left. \frac{k^3 P_L(k)}{2\pi^2} \right|_{k = k_{\mathrm{NL}}} = 1 \, . \tag{2.56}$$

This is a rough estimate for the scale at which the non-linear corrections become of order one, and this scale corresponds to $k_{\mathrm{NL}}(z = 0) \simeq 0.25$ h/Mpc and $k_{\mathrm{NL}}(z = 0.5) \simeq 0.35$ h/Mpc, using a cosmology with the parameters of [68].

When we consider redshift space distortions (see section (2.3.2)) we have to take into account other two terms which depends on the orientation of the momentum $\mathbf{k}$ wrt the line-of-sight (see [69, 70])

$$\delta_{g,s}^{\mathrm{ct},(3)}(\mathbf{k}) = c_{\mathrm{ct},0} k^2 \delta^{(1)}(\mathbf{k}, \eta) + c_{\mathrm{ct},1} f \mu_k^2 \, k^2 \delta^{(1)}(\mathbf{k}, \eta) + c_{\mathrm{ct},2} f^2 \mu_k^4 \, k^2 \delta^{(1)}(\mathbf{k}, \eta) \, , \tag{2.57}$$

where $c_{\mathrm{ct},i}$ are new parameters that will enter in the PS. In eq. (2.57) we have included in the definition of $c_{\mathrm{ct},i}$ all the time dependencies appearing in eq. (2.55).

In the next section we will see how to connect the dark matter perturbations to the (observable) galaxy density perturbation, or, more generally, to a generic tracer overdensity. We

will make use again of the standard approach that consists of a perturbative expansion around the dark matter linear solution. We will also see how to take into account observable effects like the redshift space distortions (RSD) and discreteness effect.

2.3 From dark matter to galaxies

In the previous section we have showed the SPT results for the dark matter overdensity field, but what we actually observe are galaxies, HI emission lines, Lyman-α forests or any generic tracer which emits light[1]. What we usually do is to find a perturbative solution to (2.10) for the dark matter overdensity and then introduce some new parameters, the *bias parameters*, which describe the connection between the underlying gravitational potential generated by the dark matter field and the galaxy number density.

In order to create a 3D map of the sky we have to measure the distance of the particular tracer we are interested in, and we do it by inferring the distance by assuming the Hubble law. This procedure is limited by the fact that any tracer has in general a peculiar velocity, which adds to the Hubble one. This effect causes a distortion in the distribution of tracers which goes under the name of *redshift space distortions*. We will see how to take into account this effect and how to exploit it to measure cosmological parameters. The last ingredient we will consider is the discreteness effect.

2.3.1 Bias expansion

The galaxy density δ_g depends in general on the two quantities we have used so far to describe the dark matter field, the overdensity δ and the peculiar velocity $\mathbf{u}$, which is completely defined by its divergence θ at the order we are interested in. In order to consider also the non-locality introduced by gravity, we have to consider the dependence of the galaxy density on the local gravitational potential field, ϕ, which can be expressed in terms of δ using the Poisson equation (2.9). The equivalence principle ensures that a homogeneous shift in the gravitational potential can always be reabsorbed by a change of coordinates, which suggests that the galaxy density should only depend on the first derivative of the potential $\partial_{\mathbf{x},i}\phi$. Furthermore, terms like $\partial_{\mathbf{x},i}\phi$

[1]Weak lensing observations allow to directly measure the gravitational potential, for recent results see [71].

causes a uniform force which cannot be measured by any local observer as well. Only terms containing tidal effects, i.e., with at least two spatial derivatives of the gravitational potential $\partial_{\mathbf{x},i}\partial_{\mathbf{x},j}\phi$, are measurable and should enter in the perturbative expansion of the galaxy density. Therefore, we can define the quantity [40]

$$s_{ij} \equiv \frac{2}{3\Omega_m \mathcal{H}^2}\partial_{\mathbf{x},i}\partial_{\mathbf{x},j}\phi(\mathbf{x},\tau) - \frac{1}{3}\delta_{ij}\delta(\mathbf{x},\tau) = \left[\frac{\partial_{\mathbf{x},i}\partial_{\mathbf{x},j}}{\nabla^2} - \frac{1}{3}\delta_{ij}\right]\delta(\mathbf{x},\tau) \equiv \gamma_{ij}\delta(\mathbf{x},\tau)\,, \quad (2.58)$$

which is a trace-free part tensor. Since at linear order $\theta^{(1)} = \delta^{(1)}$, we can define two quantities which will appear at second order in the perturbative expansion

$$\eta(\mathbf{x},\tau) \equiv \theta(\mathbf{x},\tau) - \delta(\mathbf{x},\tau)\,, \quad (2.59)$$

and

$$t_{ij}(\mathbf{x},\tau) \equiv \left[\frac{\partial_{\mathbf{x},i}\partial_{\mathbf{x},j}}{\nabla^2} - \frac{1}{3}\delta_{ij}\right]\eta(\mathbf{x},\tau) = \gamma_{ij}\eta(\mathbf{x},\tau)\,. \quad (2.60)$$

It turns out that at second order, in the EdS approximation

$$\eta^{(2)}_{\mathbf{k}} = \frac{2}{7}s^{(1)\,2}_{\mathbf{k}} - \frac{4}{21}\delta^{(1)\,2}_{\mathbf{k}} = -\frac{2}{7}\mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2}\left(1 - \frac{(\mathbf{q}_1\cdot\mathbf{q}_2)^2}{q_1^2 q_2^2}\right)\delta^{(1)}_{\mathbf{q}_1}\delta^{(1)}_{\mathbf{q}_2}\,, \quad (2.61)$$

suggesting that a new variable that appears only at third order can be defined

$$\psi(\mathbf{x},\tau) \equiv \eta(\mathbf{x},\tau) - \frac{2}{7}s^2(\mathbf{x},\tau) + \frac{4}{21}\delta^2(\mathbf{x},\tau)\,, \quad (2.62)$$

where we have defined $s^2 \equiv s_{ij}s_{ij}$. Assuming homogeneity and isotropy, we can infer that the galaxy density should depend only on scalar quantities that can be constructed using δ, s_{ij}, t_{ij}, η and ψ. The combinations s_{ii} and t_{ii} are zero by construction and up to third order we can construct s^2, $st \equiv s_{ij}t_{ij}$ and $s^3 \equiv s_{ij}s_{jk}s_{ki}$. It can be shown that s^3 can be expressed in term of the other quantities, see [41], meaning that the galaxy density will depend on the following seven terms

$$\begin{aligned} &1^{\text{st}}\text{ order}: \quad \delta\,, \\ &2^{\text{nd}}\text{ order}: \quad \delta^2\,, s^2\,, \\ &3^{\text{rd}}\text{ order}: \quad \delta^3\,, \delta s^2, \psi, st\,, \end{aligned} \quad (2.63)$$

or, briefly,

$$\delta_g(\mathbf{x},\tau) = \sum_{\mathcal{O}} b_{\mathcal{O}}(\tau)\mathcal{O}(\mathbf{x},\tau)\,, \quad (2.64)$$

where $\mathcal{O}$ stands for each the operator listed in eq.(2.63). The bias parameters $b_{\mathcal{O}}$ depends only on time and not on space, as one could think in general. This is true as long as we consider only the impact of long-wavelength perturbations on the galaxy density: we have ignored so far the influence of small-scale perturbations on the formation of galaxies, which is stochastic, since we have implicitly assumed that the small-scale initial conditions are not correlated over long distance, i.e., we have assumed gaussian initial conditions. For the moment we will not consider the effect of stochastic terms, and they will be described in sec. (2.3.3).

We can write, in Fourier space

$$\begin{aligned}\delta_g(\mathbf{k}) = b_1\delta(\mathbf{k}) + b_2\delta^2(\mathbf{k}) + b_{s^2}s^2(\mathbf{k}) + \\ b_3\delta^3(\mathbf{k}) + b_{\delta s^2}\delta s^2(\mathbf{k}) + b_\psi\psi(\mathbf{k}) + b_{st}st(\mathbf{k})\,.\end{aligned} \tag{2.65}$$

The various b_i are the (deterministic) bias parameters. The terms appearing in eq.(2.65) are, explicitly

$$\delta^2(\mathbf{k}) = \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2}\delta_{\mathbf{q}_1}\delta_{\mathbf{q}_2}\,, \quad \delta^3(\mathbf{k}) = \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}\delta_{\mathbf{q}_1}\delta_{\mathbf{q}_2}\delta_{\mathbf{q}_3}\,, \tag{2.66}$$

$$s^2(\mathbf{k}) = \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2}S(\mathbf{q}_1,\mathbf{q}_2)\delta_{\mathbf{q}_1}\delta_{\mathbf{q}_2}\,, \tag{2.67}$$

$$\delta s^2(\mathbf{k}) = \frac{1}{3}\mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}\left[S(\mathbf{q}_1,\mathbf{q}_2) + S(\mathbf{q}_1,\mathbf{q}_3) + S(\mathbf{q}_2,\mathbf{q}_3)\right]\delta_{\mathbf{q}_1}\delta_{\mathbf{q}_2}\delta_{\mathbf{q}_3}\,, \tag{2.68}$$

$$st(\mathbf{k}) = \frac{1}{3}\mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}\left[S(\mathbf{q}_1,\mathbf{q}_{23})\left(G_2^s(\mathbf{q}_2,\mathbf{q}_3) - F_2^s(\mathbf{q}_2,\mathbf{q}_3)\right) + 2\text{ perms.}\right]\delta_{\mathbf{q}_1}\delta_{\mathbf{q}_2}\delta_{\mathbf{q}_3}\,, \tag{2.69}$$

$$\psi(\mathbf{k}) = \theta(\mathbf{k}) - \delta(\mathbf{k}) - \frac{2}{7}\mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2}\left[S(\mathbf{q}_1,\mathbf{q}_2) + \frac{2}{3}\right]\delta_{\mathbf{q}_1}\delta_{\mathbf{q}_2}\,, \tag{2.70}$$

where we have defined

$$S(\mathbf{q}_1,\mathbf{q}_2) \equiv \frac{(\mathbf{q}_1\cdot\mathbf{q}_2)^2}{q_1^2q_2^2} - \frac{1}{3}\,. \tag{2.71}$$

2.3.2 Redshift space distortions

Following [72] we can write the radial coordinate $\mathbf{s}$ of an observed galaxy as a sum of its Hubble flow, the position measured using the Hubble law, and the distortion caused by its peculiar velocity $\mathbf{u}$. The mapping between redshift space and real space can be written as

$$\mathbf{s} = \mathbf{x} + fu_z(\mathbf{x})\hat{z}\,, \tag{2.72}$$

where f is the growth function and u_z is the peculiar velocity component on the line-of-sight, here assumed to be the $\mathbf{z}$ axis, also called *plane-parallel*, or distant observer, approximation.

Assuming that the number density is conserved in the redshift-real space mapping, we can write

$$(1+\delta_{g,s})d^3\mathbf{s} = (1+\delta_g)d^3\mathbf{x}\,, \tag{2.73}$$

where $\delta_s(\mathbf{s})$ is the the overdensity field in redshift space. Using eq. (2.72) we have $d^3\mathbf{s} = J(\mathbf{x})d^3\mathbf{x}$, where

$$J(\mathbf{x}) \equiv 1 + f\partial_z u_z(\mathbf{x}) \tag{2.74}$$

is the exact Jacobian of the mapping, we obtain

$$\delta_{g,s}(\mathbf{s}) = \frac{\delta_g(\mathbf{x}) + 1 - J(\mathbf{x})}{J(\mathbf{x})} = \frac{\delta(\mathbf{x}) - f\partial_z u_z(\mathbf{x})}{1 + f\partial_z u_z(\mathbf{x})}. \tag{2.75}$$

We will work in Fourier space,

$$\delta_{g,s}(\mathbf{k}) = \int d^3\mathbf{s}\, e^{i\mathbf{k}\cdot\mathbf{s}}\delta_{g,s}(\mathbf{s}) = \int d^3\mathbf{x}\, e^{i\mathbf{k}\cdot\mathbf{x}} e^{-ifk_z u_z(\mathbf{x})}\left[\delta_g(\mathbf{x}) - f\partial_z u_z(\mathbf{x})\right]\,. \tag{2.76}$$

Eq. (2.76) describes the exact non-linear density field in redshift space as a function of the real space one in the plane-parallel approximation. Before we proceed with the perturbative approach, we can make some comments on eq. (2.76): first of all, the terms inside the square brackets express the squashing effect, the increase in the clustering amplitude due to linear falling. This effect was firstly studied in ref. [73] and goes under the name *Kaiser effect*. Moreover the exponential describes a non-linear effect, usually referred to as *Fingers-of-God* (FOG) effect, which erases power due to velocity dispersion along the line-of-sight, see for example ref. [74] and references therein. Our perturbative approach is based on the perturbative expansion of the exponential inside eq. (2.76) which gives

$$\delta_{g,s}(\mathbf{k}) = \sum_{n=1}^{\infty} \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\ldots,\mathbf{q}_n}\left[\delta_g(\mathbf{q}_1) + f\mu_1^2\theta(\mathbf{q}_1)\right]\frac{(f\mu_k k)^{n-1}}{(n-1)!}\frac{\mu_2}{q_2}\ldots\frac{\mu_n}{q_n}\theta(\mathbf{q}_2)\ldots\theta(\mathbf{q}_n)\,, \tag{2.77}$$

where we have defined $\mu_i \equiv \hat{\mathbf{q}}_i \cdot \hat{\mathbf{z}}$.

If we consider only the linear terms, i.e. $n = 1$, we recover the well known formula of the Kaiser effect for the matter density

$$\delta_s(\mathbf{k}) = (1 + f\mu^2)\delta(\mathbf{k})\,. \tag{2.78}$$

Similarly to what was presented in section (2.1) for the real space matter density, we can write the redshift space density as a perturbative convolution (in Fourier space) of the linear matter

density field, see eq. (2.39) and (2.40)

$$\delta^{(n)}_{g,s}(\mathbf{k}) = \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\ldots,\mathbf{q}_n} Z_n(\mathbf{q}_1,\ldots,\mathbf{q}_n)\delta_L(\mathbf{q}_1)\ldots\delta_L(\mathbf{q}_n)\,, \tag{2.79}$$

where $Z_n(\mathbf{q}_1,\ldots,\mathbf{q}_n)$ are the redshift space kernels. Using eq. (2.65) and expanding at each order eq. (2.77) we can define the redshift space kernels for the galaxy density: at leading order (LO) we have $\delta^{(1)}_g(\mathbf{k}) = b_1\,\delta_L(\mathbf{k})$ so

$$\delta_{g,s}(\mathbf{k}) = \delta_g(\mathbf{k}) + f\mu^2\delta(\mathbf{k}) = \left(b_1 + f\mu^2\right)\delta_L(\mathbf{k})\,, \tag{2.80}$$

which defines

$$Z_1(\mathbf{q}_1) = b_1 + f\mu_1^2\,. \tag{2.81}$$

Analogously, going to second order yields explicitly

$$\delta^{(2)}_g(\mathbf{k}) = \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2}\left[b_1 F_2(\mathbf{q}_1,\mathbf{q}_2) + b2 + b_{s^2}S(\mathbf{q}_1,\mathbf{q}_2)\right]\delta_L(\mathbf{q}_1)\delta_L(\mathbf{q}_2)\,, \tag{2.82}$$

and

$$\begin{aligned}\delta_{g,s}(\mathbf{k}) =& \mathcal{I}_{\mathbf{k},\mathbf{q}_1}\left[\delta^{(2)}_g(\mathbf{q}_1) + f\mu_1^2\theta^{(2)}(\mathbf{q}_1)\right] \\ &+ \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2}\left[\delta^{(1)}_g(\mathbf{q}_1) + f\mu_1^2\theta^{(1)}(\mathbf{q}_1)\right] f\mu_k k\frac{\mu_2}{q_2}\theta^{(1)}(\mathbf{q}_2)\,,\end{aligned} \tag{2.83}$$

which gives

$$\begin{aligned}Z_2(\mathbf{q}_1,\mathbf{q}_2) =& b_1 F_2^s(\mathbf{q}_1,\mathbf{q}_2) + f\mu_k^2\,G_2^s(\mathbf{q}_1,\mathbf{q}_2) + b_{s^2}\left(\frac{(\mathbf{q}_1\cdot\mathbf{q}_2)^2}{q_1^2q_2^2} - \frac{1}{3}\right) \\ &+ b_2 + \frac{f\mu_k k}{2}\left[(b_1 + f\mu_1^2)\frac{\mu_2}{q_2} + (b_1 + f\mu_2^2)\frac{\mu_1}{q_1}\right]\,.\end{aligned} \tag{2.84}$$

With analogous steps we can obtain the expression for the third order kernel

$$
\begin{aligned}
Z_3(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) =& b_1 F_3^s(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) + f\mu_k^2 G_3^s(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) \\
&+ b_\psi \left[G_3^s(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) - F_3(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3)\right] + b_3 \\
&\Bigg[2b_2 F_2^s(\mathbf{q}_1,\mathbf{q}_2) + 2b_{s^2} S(\mathbf{q}_1,\mathbf{q}_2) F_2^s(\mathbf{q}_2,\mathbf{q}_3) \\
&+ b_{st} S(\mathbf{q}_1,\mathbf{q}_{23})\Big(G_2^s(\mathbf{q}_2,\mathbf{q}_3) - F_2^s(\mathbf{q}_2,\mathbf{q}_3)\Big) \\
&- \frac{4b_\psi}{7} S(\mathbf{q}_1,\mathbf{q}_{23}) F_2^s(\mathbf{q}_2,\mathbf{q}_3) \\
&+ f\mu_k k \frac{\mu_1}{q_1}\left[b_1 F_2(\mathbf{q}_2,\mathbf{q}_3) + f\mu_{23}^2 G_2(\mathbf{q}_2,\mathbf{q}_3) + b_{s^2} S(\mathbf{q}_2,\mathbf{q}_3)\right] \\
&+ f\mu_k k \left[\left(b_1 + f\mu_1^2\right)\frac{\mu_{23}}{q_{23}} G_2(\mathbf{q}_2,\mathbf{q}_3)\right] \\
&+ \frac{(f\mu_k k)^2}{2}\left[\left(b_1 + f\mu_1^2\right)\frac{\mu_2}{q_2}\frac{\mu_3}{q_3}\right] + 2 \text{ perms.}\Bigg] .
\end{aligned}
\tag{2.85}
$$

2.3.3 Stochastic terms

The perturbative bias expansion described in Sec. (2.3.1) implicitly assumes that the influence of non-linear scales on galaxy formation is negligible, and we have not accounted for this effect so far. In order to take into account this effect in our galaxy density we have to introduce stochastic fields in the bias relation. We will denote the stochastic fields as ε, and they will be of order one in the perturbative expansion.

The set of stochastic contributions consists of all the terms of the deterministic bias expansion $\mathcal{O}$ in eq. (2.64), and we will indicate them with $\varepsilon_\mathcal{O}$. In general we can write [41]

$$
\delta_g(\mathbf{x},\tau) = \sum_\mathcal{O}\left[b_\mathcal{O}(\tau) + \varepsilon_\mathcal{O}(\mathbf{x},\tau)\right]\mathcal{O}(\mathbf{x},\tau) + \varepsilon(\mathbf{x},\tau) \tag{2.86}
$$

with

$$
\begin{aligned}
&1^{\text{st}} \text{ order}: \quad \varepsilon\,, \\
&2^{\text{nd}} \text{ order}: \quad \varepsilon_\delta \delta\,, \\
&3^{\text{rd}} \text{ order}: \quad \varepsilon_{\delta^2}\delta^2, \varepsilon_{s^2} s^2\,.
\end{aligned}
\tag{2.87}
$$

Note that since the $\varepsilon_\mathcal{O}$ are assumed to be uncorrelated with the large-scale perturbations, meaning that $\langle \varepsilon_\mathcal{O} \mathcal{O}' \rangle = 0$ for each $\mathcal{O}$ and $\mathcal{O}'$, they are completely described by their moments $\langle \varepsilon_\mathcal{O} \varepsilon_{\mathcal{O}'} \rangle$, $\langle \varepsilon_\mathcal{O} \varepsilon_{\mathcal{O}'} \varepsilon_{\mathcal{O}''} \rangle$, and so on.

In this work we will consider only the ε field as this is the main stochastic term considered in current LSS analysis, see [45, 46]. In particular, since the galaxy density contrast will depend non-locally from the dark matter perturbations, we have to consider also higher derivative terms for the stochastic field. In Fourier space this translates into an expansion in k^2 (for symmetry reason)

$$\langle \varepsilon(\mathbf{k})\varepsilon(\mathbf{k}')\rangle' = P_\varepsilon^{\{0\}} + P_\varepsilon^{\{2\}} k^2 + P_\varepsilon^{\{4\}} k^4 + \dots \,, \tag{2.88}$$

where $\langle \dots \rangle'$ indicates the correlator divided by $(2\pi)^3$ and the Dirac delta. We will explicitly calculate the $P^{\{0\}}$ term of this expansion. We will keep only the k^2 term, as it is the one relevant at the order we are interested in this work; at this order we have [75, 70]

$$\delta_{g,s}^{\mathrm{stoch},(3)}(k) = c_{\epsilon,1}\varepsilon + c_{\epsilon,2}k^2\varepsilon + c_{\epsilon,3} f\mu_k^2 k^2 \varepsilon \,. \tag{2.89}$$

As we mentioned above, the stochastic field ε is completely determined by its moment $\langle \varepsilon\varepsilon \rangle$ and in the following we will show the result for this correlator when considering only Poisson shot noise. The Poisson shot noise appears when we consider a discrete field (like, e.g., the number density of galaxies) and we approximate it with a continuous one. Indeed, considering the Poisson model, the number density of galaxies is

$$n(\mathbf{x}) = \sum_i \delta_D(\mathbf{x} - \mathbf{x}_i) \,, \tag{2.90}$$

while the (discrete) density contrast is given by

$$\delta_{\mathrm{d}}(\mathbf{x}) = \frac{n(\mathbf{x})}{\bar{n}} - 1 \,, \tag{2.91}$$

where we have defined the mean number density as $\bar{n} = \langle n(\mathbf{x})\rangle$. At this point, following [76] we can calculate the two-point correlation function of the discrete field

$$\xi_{\mathrm{d}}(\mathbf{x}_1, \mathbf{x}_2) = \langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\rangle = \frac{1}{\bar{n}^2}\langle n(\mathbf{x}_1) n(\mathbf{x}_2)\rangle - 1 \,. \tag{2.92}$$

The two-point of the number density field can be written as

$$\langle n(\mathbf{x}_1) n(\mathbf{x}_2)\rangle = \langle \sum_i \delta_D(\mathbf{x}_1 - \mathbf{x}_i)\delta_D(\mathbf{x}_2 - \mathbf{x}_i)\rangle + \langle \sum_{i,j} \delta_D(\mathbf{x}_1 - \mathbf{x}_i)\delta_D(\mathbf{x}_2 - \mathbf{x}_j)\rangle \,, \tag{2.93}$$

where we have separated the contributions coming from different points (the second term in the rhs of eq. (2.93)) and those coming from the same point, which will be responsible for the

shot noise term. In particular the former can be parametrized using the two-point correlation function of the smooth field

$$\langle \sum_{i,j} \delta_D(\mathbf{x}_1 - \mathbf{x}_i)\delta_D(\mathbf{x}_2 - \mathbf{x}_j)\rangle = \bar{n}^2 \left[1 + \xi(|\mathbf{x}_1 - \mathbf{x}_2|)\right] , \tag{2.94}$$

so that we obtain, for the discrete field,

$$\xi_{\rm d}(\mathbf{x}_1, \mathbf{x}_2) = \xi(|\mathbf{x}_1 - \mathbf{x}_2|) + \frac{1}{\bar{n}}\delta_D(\mathbf{x}_1 - \mathbf{x}_2) . \tag{2.95}$$

Going to Fourier space we obtain the well known result

$$P_{\rm d}(k) = P(k) + \frac{1}{\bar{n}} , \tag{2.96}$$

and, in the Poisson model, we can identify the two-point correlation function of the stochastic field as

$$\langle \varepsilon\varepsilon \rangle = \frac{1}{\bar{n}} . \tag{2.97}$$

Similarly, for the three-point function of the discrete field we have

$$\zeta_{\rm d}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \langle \delta_{\rm d}(\mathbf{x}_1)\delta_{\rm d}(\mathbf{x}_2)\delta_{\rm d}(\mathbf{x}_3)\rangle = \frac{\langle n(\mathbf{x}_1)n(\mathbf{x}_2)n(\mathbf{x}_3)\rangle}{\bar{n}^3} - \left[\frac{\langle n(\mathbf{x}_1)n(\mathbf{x}_2)\rangle}{\bar{n}^2} + 2\text{cyc.}\right] + 2 . \tag{2.98}$$

Analogously to eq. (2.93) we can rewrite the three-point correlator of the number density field as

$$\begin{aligned}
&\langle n(\mathbf{x}_1)n(\mathbf{x}_2)n(\mathbf{x}_3)\rangle \\
&= \langle \sum_{i} \delta_D(\mathbf{x}_1 - \mathbf{x}_i)\delta_D(\mathbf{x}_2 - \mathbf{x}_i)\delta_D(\mathbf{x}_3 - \mathbf{x}_i)\rangle \\
&+ \left[\langle \sum_{i,j} \delta_D(\mathbf{x}_1 - \mathbf{x}_i)\delta_D(\mathbf{x}_2 - \mathbf{x}_j)\delta_D(\mathbf{x}_3 - \mathbf{x}_j)\rangle + 2\text{cyc.}\right] \\
&+ \langle \sum_{i,j,k} \delta_D(\mathbf{x}_1 - \mathbf{x}_i)\delta_D(\mathbf{x}_2 - \mathbf{x}_j)\delta_D(\mathbf{x}_3 - \mathbf{x}_k)\rangle \\
&= \delta_D(\mathbf{x}_1 - \mathbf{x}_2)\delta_D(\mathbf{x}_1 - \mathbf{x}_3)\bar{n} + \left[\delta_D(\mathbf{x}_2 - \mathbf{x}_3)\bar{n}^2(1 + \xi_{12}) + 2\text{cyc.}\right] + \bar{n}^3\left(1 + \xi_{12} + \xi_{13} + \xi_{23} + \zeta\right) ,
\end{aligned} \tag{2.99}$$

where ζ is the three-point function of the continuous field. We use the notation $\xi_{ij} \equiv \xi(|\mathbf{x}_i - \mathbf{x}_j|)$. Using eq. (2.94) and (2.99) we obtain

$$\zeta_{\rm d}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \frac{1}{\bar{n}^2}\delta_D(\mathbf{x}_1 - \mathbf{x}_2)\delta_D(\mathbf{x}_1 - \mathbf{x}_3) + \left[\frac{1}{\bar{n}}\delta_D(\mathbf{x}_2 - \mathbf{x}_3)\xi_{12} + 2\text{cyc.}\right] + \zeta_{123} . \tag{2.100}$$

Going to Fourier space we get the discrete bispectrum

$$B_{\rm d}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) = B(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) + \frac{1}{\bar{n}}\left[P(k_1)+P_L(k_2)+P_L(k_3)\right] + \frac{1}{\bar{n}^2}\,, \tag{2.101}$$

where B is the continuous bispectrum. From eq. (2.101) we can write the shot noise contribution to the bispectrum as

$$B_\varepsilon(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) \equiv \frac{1}{\bar{n}^2} + \frac{1}{\bar{n}}\left[P_L(k_1)+P_L(k_2)+P_L(k_3)\right]\,. \tag{2.102}$$

Possible generalization of the stochastic contribution can be performed using the basis in eq. (2.87) presented in [41] or, for example, parametrizing possible deviations from the poissonian shot noise with additional parameters, see [61, 77].

2.4 1-loop power spectrum and tree-level bispectrum

From eqs. (2.79), (2.89) and (2.57) we can write the full galaxy density field at third order in redshift space

$$\begin{aligned}\delta^{(3)}_{g,s}(\mathbf{k},\eta) =& Z_1(k,\mu_k)\delta^{(1)}(\mathbf{k},\eta)\\ &+ \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2} Z_2(\mathbf{q}_1,\mathbf{q}_2)\delta^{(1)}(\mathbf{q}_1)\delta^{(1)}(\mathbf{q}_2)\\ &+ \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3} Z_3(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3)\delta^{(1)}(\mathbf{q}_1)\delta^{(1)}(\mathbf{q}_2)\delta^{(1)}(\mathbf{q}_3)\\ &+ \delta^{\rm stoch,(3)}_{g,s} + \delta^{\rm ct,(3)}_{g,s}\,,\end{aligned} \tag{2.103}$$

where the first three lines represent the deterministic and perturbative relation between galaxy and matter overdensities, while the last line include the UV and stochastic effects.

From eq. (2.103) we can calculate the 1-loop PS, which is given by

$$\begin{aligned}P_{g,s}(k,\mu_k) =& Z_1(k,\mu_k)^2 P_L(k) + 2\int\frac{d^3\mathbf{q}}{(2\pi)^3}\left[Z_2(\mathbf{k}-\mathbf{q},\mathbf{q})\right]^2 P_L(q)P_L(|\mathbf{k}-\mathbf{q}|)\\ &+ 6Z_1(k,\mu_k)P_L(k)\int\frac{d^3\mathbf{q}}{(2\pi)^3}Z_3(\mathbf{k},\mathbf{q},-\mathbf{q})P_L(q)\\ &+ c_{{\rm ct},0}k^2P_L(k) + c_{{\rm ct},1}f\mu_k^2k^2P_L(k) + c_{{\rm ct},2}f^2\mu_k^4k^2P_L(k)\\ &+ \frac{1}{\bar{n}}\left(c_{\varepsilon,1} + c_{\varepsilon,2}k^2 + c_{\varepsilon,3}f\mu_k^2\right)\,.\end{aligned} \tag{2.104}$$

In actual data analysis we consider angle-averaged multipoles of the PS, defined as

$$P^{(l)}_{g,s}(k) = \frac{2l+1}{2}\int_{-1}^{1} d\mu_k P_{g,s}(k,\mu_k)\mathcal{P}_l(\mu_k)\,, \tag{2.105}$$

where $\mathcal{P}_l(x)$ are the Legendre polynomials of order l.

Using eq. (2.103) we can also write the tree-level bispectrum. At this order we can neglect the effect of higher derivative terms, see [41], and write

$$\begin{aligned} B_{g,s}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) =& 2Z_1(\mathbf{k}_1)Z_1(\mathbf{k}_2)Z_2(\mathbf{k}_1,\mathbf{k}_2)P_L(k_1)P_L(k_2) + 2\,\text{cyc.} \\ &+ B_\varepsilon(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)\,, \end{aligned} \tag{2.106}$$

where B_ε is the shot noise contribution, given in eq. (2.102). Eqs. (2.104) and (2.106) are the fundamental predictions to be compared with observations. In recent data analysis, [45, 46, 47], the full one-loop PS has been used to constrain the cosmological parameters $\{A_s, n_s, \Omega_m, H\}$. The results of these analysis has shown the power of late-time observations, giving results in good accordance with those from the Planck mission [15].

2.5 IR resummation

At more recent times the non-linear structure formation affects the shape of the BAO peak, which is the imprint left by the interaction between the photons and baryons in the primordial fluid, see [18] and references therein. In particular, the main non linear effect are long wavelength displacements of order ~ 10 Mpc caused by bulk flows, which cause a damping of the BAO peak in the linear correlation function [78]. In Fourier space this effect translates into an exponential damping of the oscillating part of the power spectrum. This means that evolving the initial linear power spectrum considering only the linear effects of the growth of overdensity leads to larger (undamped) BAO at lower redshifts than those observed in the actual power spectrum. These bulk flow effects have to be taken into account in order to obtain unbiased estimates of the cosmological parameters.

The displacement due to bulk flows can always be written as an integral of the velocity perturbation

$$\mathbf{s}(\mathbf{x},\eta) = \int^\eta d\eta' \, \frac{\mathbf{u}(\mathbf{x},\eta')}{\mathcal{H}(\eta')}. \tag{2.107}$$

Since we are interested on the effect of large (IR) scales on mildly non linear scales, we can assume that, in Fourier space, the displacement $\mathbf{s}$ has support for modes $q < \Lambda$, where Λ is an infrared scale. Let us introduce now the new density field $\bar{\delta}$ which represents the density

perturbation where the effect of the IR displacement has been subtracted, explicitly

$$\delta(\mathbf{x},\tau) = \bar{\delta}(\mathbf{x} - \mathbf{s}(\mathbf{x},\tau),\tau)\,. \tag{2.108}$$

The field δ is the *true* field which contains the effect described above, while $\bar{\delta}$ is the field where the effect of IR displacements has not been taken into account. We are interested in resumming the effect of the IR displacement in the true field δ.Going to Fourier space, we obtain from eq. (2.108)

$$\delta(\mathbf{k}) = \int d^3\mathbf{x}\, e^{i\mathbf{k}\cdot\mathbf{x}}\bar{\delta}(\mathbf{x} - \mathbf{s}(\mathbf{x})) \simeq \int d^3\mathbf{x}\, e^{i\mathbf{k}\cdot(\mathbf{x}+\mathbf{s}(\mathbf{x}))}\bar{\delta}(\mathbf{x})\,, \tag{2.109}$$

where in the second equality we have neglected the $\mathbf{x}$-dependence of the displacement. If we calculate the power spectrum we obtain

$$\langle\delta_{\mathbf{k}}\delta_{\mathbf{k}'}\rangle = \int d^3\mathbf{x}d^3\mathbf{x}'\, e^{i(\mathbf{k}\cdot\mathbf{x}+\mathbf{k}'\cdot\mathbf{x}')}\langle e^{i(\mathbf{k}\cdot\mathbf{s}(\mathbf{x})+\mathbf{k}'\cdot\mathbf{s}(\mathbf{x}'))}\bar{\delta}(\mathbf{x})\bar{\delta}(\mathbf{x}')\rangle\,. \tag{2.110}$$

Translational invariance imposes that the argument of the angular brackets depends only on the modulus of the distance $r \equiv |\mathbf{x} - \mathbf{x}'|$ and a convenient change of coordinates

$$\begin{aligned} \mathbf{x} &= \frac{\mathbf{r}}{2} - \mathbf{R}\,, \\ \mathbf{x}' &= -\frac{\mathbf{r}}{2} - \mathbf{R}\,, \end{aligned} \tag{2.111}$$

gives

$$\langle\delta_{\mathbf{k}}\delta_{\mathbf{k}'}\rangle = (2\pi)^3\delta_D(\mathbf{k}+\mathbf{k}')\int d^3\mathbf{r}\, e^{i\mathbf{k}\cdot\mathbf{r}}\langle e^{i\mathbf{k}\cdot(\mathbf{s}(-\mathbf{r}/2)-\mathbf{s}(\mathbf{r}/2))}\bar{\delta}(-\mathbf{r}/2)\bar{\delta}(\mathbf{r}/2)\rangle\,. \tag{2.112}$$

We are interested in the effect of large scales ($< \Lambda$) on small scales ($> \Lambda$) and this separation of scales is reflected in eq. (2.112): we can break the correlator into his small and large scale part, since these two, at the order we are interested in, do not correlate. Explicitly this means that we can write

$$\langle e^{-i\mathbf{k}\cdot(\mathbf{s}(-\mathbf{r}/2)-\mathbf{s}(\mathbf{r}/2))}\bar{\delta}(-\mathbf{r}/2)\bar{\delta}(\mathbf{r}/2)\rangle = \langle e^{-i\mathbf{k}\cdot(\mathbf{s}(-\mathbf{r}/2)-\mathbf{s}(\mathbf{r}/2))}\rangle\langle\bar{\delta}(-\mathbf{r}/2)\bar{\delta}(\mathbf{r}/2)\rangle\,. \tag{2.113}$$

Given the range of scales on which the displacement $\mathbf{s}$ has support, we can treat it as a linear gaussian field (Zel'Dovich approximation, ZD, see [79]), from eq. (2.107)

$$\mathbf{s}(\mathbf{x}) = -i\int\frac{d^3\mathbf{p}}{(2\pi)^3}e^{-i\mathbf{p}\cdot\mathbf{x}}\frac{\mathbf{p}}{p^2}\delta_L(\mathbf{p}) \tag{2.114}$$

and from eq. (2.113)

$$\langle e^{-i\mathbf{k}\cdot\Delta\mathbf{s}(\mathbf{r})}\rangle = e^{-\frac{1}{2}\langle(\mathbf{k}\cdot\Delta(\mathbf{r})\mathbf{s})^2\rangle}\,, \tag{2.115}$$

where we have defined

$$\Delta\mathbf{s}(\mathbf{r}) \equiv \mathbf{s}(-\mathbf{r}/2) - \mathbf{s}(\mathbf{r}/2). \tag{2.116}$$

After some algebra, one finds

$$\Sigma^2(\mathbf{k},\mathbf{r},\Lambda) \equiv \frac{1}{2}\langle(\mathbf{k}\cdot\Delta\mathbf{s}(\mathbf{r}))^2\rangle = \int^{\Lambda}\frac{d^3\mathbf{p}}{(2\pi)^3}\frac{(\mathbf{k}\cdot\mathbf{p})^2}{p^4}\left(1-\cos\left(\mathbf{p}\cdot\mathbf{r}\right)\right)P_L(p) \tag{2.117}$$

$$= \frac{k^2}{6\pi^2}\int^{\Lambda} dp\, P_L(p)((1-j_0(pr)+2j_2(pr)))\,, \tag{2.118}$$

where $j_n(z)$ are the n-th order reduced Bessel functions, we can write

$$P(k) = \int d^3\mathbf{r}\, e^{i\mathbf{k}\cdot\mathbf{r}} e^{-\Sigma^2(\mathbf{k},\mathbf{r},\Lambda)}\bar{\xi}(\mathbf{r}), \tag{2.119}$$

where $\bar{\xi}$ is the two point correlation function of the density field without the IR displacements. Expanding the exponential factor at zeroth order we obtain

$$P(k) = \int \frac{d^3\mathbf{r}d^3\mathbf{q}}{(2\pi)^3} e^{i(\mathbf{k}-\mathbf{q})\cdot\mathbf{r}}\bar{P}(q) = \bar{P}(k)\,, \tag{2.120}$$

which tells us that, at the lowest order, the power spectrum of the density perturbations is the same of the one where the bulk flows effect was subtracted. At the first perturbative order we obtain the first IR correction

$$P(k)^{(1),\mathrm{IR}} = -\int\frac{d^3\mathbf{q}}{(2\pi)^3}\frac{(\mathbf{k}\cdot\mathbf{q})^2}{q^4}P_L(q)\left[\bar{P}(k) - \frac{1}{2}\left(\bar{P}(k_+)+\bar{P}(k_-)\right)\right]\,, \tag{2.121}$$

where we have defined $k_\pm \equiv |\mathbf{k}\pm\mathbf{q}|$. By expanding the difference inside the parentheses in eq. (2.121), the first non-vanishing contribution to the integral, after angular integration, is of order $\sim q^2\partial^2 P(k)/\partial k^2$, which, multiplied $((\mathbf{k}\cdot\mathbf{q})/q^2)^2$, is $O(q^0)$. Therefore, unless an enhancement factor comes out from the momentum dependence of the PS, then the contribution of eq. (2.121) is of order

$$P_L(k)\frac{\partial\ln P_L(k)}{\partial\ln k}\int^{\Lambda} d^3\mathbf{q}P_L(q) \sim P_L(k)\sigma_\delta^2(\Lambda)\,, \tag{2.122}$$

where we have defined

$$\sigma_\delta^2(\Lambda) \equiv \int^{\Lambda} d^3\mathbf{q}P_L(q)\,. \tag{2.123}$$

We can see that eq. (2.123) is parametrically of the same order of other terms that we are neglecting in this calculations. In other words, there is no clear criterium for resumming these contributions with respect to others.

On the other hand, if the linear power spectrum presents an oscillatory feature, the new scale introduces an enhancement factor at each PT order. Considering the BAO's, we can split the linear PS into an smooth and a oscillating component

$$P(k) = P_{\rm nw}(k) + P_{\rm w}(k)\,, \tag{2.124}$$

where the *wiggly* part is given by

$$P_{\rm w}(k) = A(k)\sin\left(kr_{BAO} + \phi(k)\right), \tag{2.125}$$

with $A(k)$ and $\phi(k)$ smooth functions of the momentum in the BAO range. If we consider the IR correction to the wiggle part of the power spectrum we obtain for the terms inside the parenthesis in eq. (2.121)

$$\left[\bar{P}_{\rm w}(k) - \frac{1}{2}\left(\bar{P}_{\rm w}(k_+) - \bar{P}_{\rm w}(k_-)\right)\right] = \bar{P}_{\rm w}(k)\left(1 - \cos\left(q\mu\, r_{BAO}\right)\right)\,, \tag{2.126}$$

where $\mu \equiv \hat{\mathbf{k}}\cdot\hat{\mathbf{q}}$. By expanding this result for small q's and multiplying by $((\mathbf{k}\cdot\mathbf{q})/q^2)^2$, we can see that the wiggly part of eq. (2.121) is of order $\sim k^2 r_{BAO}^2\sigma_\delta^2(\Lambda)$, that is, enhanced with respect to eq. (2.122) by a $k^2 r_{BAO}^2$ factor, that we want to resum at any order in PT. Using (2.126) in (2.121) we obtain

$$P_{\rm w}^{(1),\rm IR} = -\Sigma^2(k, r_{BAO}, \Lambda)P_{\rm w}(k)\,. \tag{2.127}$$

At the leading order we can calculate the effect of the IR displacement and resum it at all orders, at second order one finds

$$P_{\rm w}^{(2),\rm IR} = -\frac{1}{2}\Sigma^2(k, r_{BAO}, \Lambda)P_{\rm w}^{(1),\rm IR} = \frac{1}{2}\Sigma^4(k, r_{BAO}, \Lambda)P_{\rm w}(k)\,, \tag{2.128}$$

and we can easily generalize our results at order $n+1$

$$P_{\rm w}^{(n+1),\rm IR}(k) = -\frac{1}{n+1}\Sigma^2(k, r_{BAO}, \Lambda)P_{\rm w}^{(n),\rm IR}(k) = \frac{1}{(n+1)!}\Sigma^{2n}(k, r_{BAO}, \Lambda)P_{\rm w}(k)\,. \tag{2.129}$$

At leading order we can write

$$P_{LO}(k) = \bar{P}_L^{\rm nw}(k) + e^{-\Sigma^2(k, r_{BAO}, \Lambda)}\bar{P}_L^{\rm w}(k)\,. \tag{2.130}$$

Since we are interested to the 1-loop power spectrum, we have to go the next-to-leading order (NLO). We start by looking at the IR limit of the oscillating part of the 1-loop power spectrum in eq. (2.49), which gives

$$\begin{aligned} P^{\rm w}_{1-\rm loop}(k>\Lambda) &\simeq \frac{1}{2}\int^{\Lambda}\frac{d^3\mathbf{q}}{(2\pi)^3}\frac{(\mathbf{k}\cdot\mathbf{q})^2}{q^4}P_L(q)\left[P^{\rm w}_L(|\mathbf{k}+\mathbf{q}|)+P^{\rm w}_L(|\mathbf{k}-\mathbf{q}|)-2P^{\rm w}_L(k)\right] \\ &= -\Sigma^2(k,r_{BAO},\Lambda)P^{\rm w}_L(k)\,. \end{aligned} \tag{2.131}$$

Now we can apply directly the same procedure that lead us to eq. (2.130), obtaining

$$\begin{aligned} P(k) =& \bar{P}^{\rm nw}_L(k)+\bar{P}^{\rm nw}_{1-\rm loop} \\ &+ e^{-\Sigma^2(k,r_{BAO},\Lambda)}\left[1+\Sigma^2(k,r_{BAO},\Lambda)\right]\bar{P}^{\rm w}_L(k)+e^{-\Sigma^2(k,r_{BAO},\Lambda)}\bar{P}^{\rm w}_{1-\rm loop}(k)\,, \end{aligned} \tag{2.132}$$

where we have used the fact that when replacing $P^{\rm w}(k)$ with the 1-loop power spectrum already contains $\Sigma^2(k,r_{BAO},\Lambda)P^{\rm w}_L(k)$.

The results described above can be easily generalized to biased tracers, assuming that the equivalence principle holds, by substituting the matter power spectra with the biased tracers ones. On the other hand, going to redshift space requires a generalization of our results, meaning [80]

$$\begin{aligned} P^{\rm IR,res,LO+NLO}_{g,s}(k,\mu_k) =& Z_1^2(k,\mu_k)\left(P^{\rm nw}_L(k)+\left(1+\Sigma^2_{\rm tot}(\mu)\right)e^{-\Sigma^2_{\rm tot}(\mu_k)}P^{\rm w}_L(k)\right) \\ &+ P^{1-\rm loop}_{g,s}\left[P^{\rm nw}_L\right]+e^{-\Sigma^2_{\rm tot}(\mu_k)}P^{1-\rm loop,\,w}_{g,s}\,, \end{aligned} \tag{2.133}$$

where

$$\Sigma^2_{\rm tot}(\mu_k)\equiv(1+f\mu_k^2(2+f))\Sigma^2(k,r_{BAO},\Lambda)+f^2\mu_k^2(\mu_k^2-1)\delta\Sigma^2(k,r_{BAO},\Lambda)\,, \tag{2.134}$$

and

$$\delta\Sigma^2(k,r_{BAO},\Lambda)=\frac{1}{18\pi^2}\int^{\Lambda}dpP_L(p)\,j_2(pr_{BAO})\,. \tag{2.135}$$

In eq. (2.133), with $P^{1-\rm loop}_{g,s}[P^{\rm nw}_L]$ we mean that this term is calculated using the linear no-wiggle PS inside the 1-loop integrals, see the deterministic part of eq. (2.104). The wiggly part of the one loop power spectrum is given by

$$\begin{aligned} P^{1-\rm loop,\,w}_{g,s}(k,\mu) =& 4\int\frac{d^3\mathbf{p}}{(2\pi)^3}\left[Z_2(\mathbf{k}-\mathbf{p},\mathbf{p})\right]^2P^{\rm nw}_L(p)P^{\rm nw}_L(|\mathbf{k}-\mathbf{p}|)e^{-\Sigma^2_{\rm tot}(\mu_p)} \\ &+6Z_1(k,\mu_k)e^{-\Sigma^2_{\rm tot}(\mu_k)}P^{\rm w}_L(k)\int\frac{d^3\mathbf{p}}{(2\pi)^3}Z_3(\mathbf{k},\mathbf{p},-\mathbf{p})P^{\rm nw}_L(p) \\ &+6Z_1(k,\mu_k)P^{\rm nw}_L(k)\int\frac{d^3\mathbf{p}}{(2\pi)^3}Z_3(\mathbf{k},\mathbf{p},-\mathbf{p})P^{\rm w}_L(p)e^{-\Sigma^2_{\rm tot}(\mu_p)}\,. \end{aligned} \tag{2.136}$$

With similar arguments, for the tree-level bispectrum of galaxies in redshift space, we can write the leading-order prediction for the IR-resummation as [80]

$$B_{g,s}^{\mathrm{IR,res,LO}}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) = Z_1(\mathbf{k}_1)Z_1(\mathbf{k}_2)Z_2(\mathbf{k}_1,\mathbf{k}_2)\big[P_L^{\mathrm{nw}}(k_1)P_L^{\mathrm{nw}}(k_2) + P_L^{\mathrm{nw}}(k_1)e^{-\Sigma^2_{\mathrm{tot}}(\mu_{k_2})}P_L^{\mathrm{w}}(k_2) + P_L^{\mathrm{nw}}(k_2)e^{-\Sigma^2_{\mathrm{tot}}(\mu_{k_1})}P_L^{\mathrm{w}}(k_1)\big] + 2\,\mathrm{cyc.} \tag{2.137}$$

IR resummation is the last ingredient we need for the calculation of the 1-loop PS for biased tracers in redshift space: this will be the quantity that we will compare with observational data in order to extrapolate measurements of cosmological parameters. other approaches give analogous results, see [80, 81].

2.6 FFTLog method

The final goal of the study presented so far is to provide an observable to be compared with actual measurements. In our case the observable is the galaxy power spectrum in redshift space. In eq. (2.104) we have accounted for all the observational effects and IR-resummation, described in section (2.5). EFTofLSS was used to perform the analysis of the BOSS clustering dataset by [45, 46, 47]. In these works, the cosmological parameter space, composed by $\{H_0, A_s, n_s, \omega_b, \omega_c\}$, was sampled using the MCMC integration method [82, 83, 84, 85], which fully calculates the theoretical model at each step (varying the cosmological parameters). Standard numerical techniques takes some minutes to calculate the k-integrals in the one loop power spectrum for $\sim$30 points, running similar chains would be an extraordinary cost in terms of time and computational resources. In [43, 44] it was introduced a novel technique for the fast evaluation of the one loop integrals based on the FFTLog approach, that we will explain below. This technique was implemented in four main codes currently used for parameter estimation: `FAST-PT` [43, 86], PyBird [10], `CLASS-PT` [87] and `velocileptors` [88, 89].

The main idea behind the approach introduced in [44] is to approximate cosmologies with a sum of power law universes, k^α, with α complex. This is made possible by using the Fast Fourier Transform (FFT) of the PS in $\log k$ (Log). Given a range of scales, from k_{min} to k_{max},

the linear power spectrum with N k-points is given by

$$\bar{P}_L(k_n) = \sum_{m=-\frac{N}{2}}^{\frac{N}{2}} c_m k_n^{\nu+i\eta_m}, \tag{2.138}$$

where the coefficients c_m together with the frequencies η_m are defined by

$$c_m = \frac{1}{N}\sum_{l=0}^{N-1} P_L(k_l)^{-\nu} k_{\rm min}^{-i\eta_m} e^{-2\pi i m l/N}, \qquad \eta_m = \frac{2\pi m}{\log\left(k_{\rm max}/k_{\rm min}\right)}\,. \tag{2.139}$$

Here we denote the approximation for the linear power spectrum with $\bar{P}_L(k)$, while the coefficients c_m in eq. (2.139) are defined using the exact linear power spectrum $P_L(k)$. The parameter ν is a real number, while the powers in the power-law expansion are, in general, complex numbers. The interesting part of the decomposition eq. (2.138) is that it reduces the evaluation of a loop diagram for an arbitrary cosmology to evaluation of the same diagram for a set of different power-law universes with numerical (in general, time dependent) coefficients. Power-law momentum integral can be done analytically. In the simplest case of the one-loop power spectrum, the momentum integral for a power-law universe can be expressed entirely in terms of gamma functions.

The decomposition eq. (2.138) is useful because it separates the cosmology dependent portion, encoded entirely in the coefficients c_m, from the loop calculations which have been reduced to that of much more tractable cosmologies. For fixed value of bias ν, the momentum range $(k_{\rm min}, k_{\rm max})$ and the number of sampling points N the difficult part of the calculation which involves momentum integrals can be done only once and then used for any cosmology.

Recalling eqs. (2.50) and (2.51), we now express the explicit form of the kernels F_n in terms of integer powers of k^2, q^2 and $|\mathbf{k}-\mathbf{q}|^2$, obtaining

$$F_2(\mathbf{k}-\mathbf{q},\mathbf{q}) = \frac{5}{14} + \frac{3k^2}{28q^2} + \frac{3k^2}{28|\mathbf{k}-\mathbf{q}|^2} - \frac{5q^2}{28|\mathbf{k}-\mathbf{q}|^2} - \frac{5|\mathbf{k}-\mathbf{q}|^2}{28q^2} + \frac{k^4}{14|\mathbf{k}-\mathbf{q}|^2q^2}, \tag{2.140}$$

and (see also [90])

$$
\begin{aligned}
F_3(\mathbf{q},-\mathbf{q},\mathbf{k}) = \frac{1}{|\mathbf{k}-\mathbf{q}|^2}\Bigg[&\frac{5k^2}{126}-\frac{11\mathbf{k}\cdot\mathbf{q}}{108}+7\frac{(\mathbf{k}\cdot\mathbf{q})^2}{108k^2}-\frac{k^2(\mathbf{k}\cdot\mathbf{q})^2}{54q^4}+\frac{4(\mathbf{k}\cdot\mathbf{q})^3}{189q^4}\\
&-\frac{23k^2\mathbf{k}\cdot\mathbf{q}}{756q^2}+\frac{25(\mathbf{k}\cdot\mathbf{q})^2}{252q^2}-\frac{2(\mathbf{k}\cdot\mathbf{q})^3}{k^2q^2}\Bigg]\\
+\frac{1}{|\mathbf{k}+\mathbf{q}|^2}\Bigg[&\frac{5k^2}{126}+\frac{11\mathbf{k}\cdot\mathbf{q}}{108}-\frac{7(\mathbf{k}\cdot\mathbf{q})^2}{108k^2}-\frac{4k^2(\mathbf{k}\cdot\mathbf{q})^2}{27q^4}\\
&-\frac{53(\mathbf{k}\cdot\mathbf{q})^3}{189q^4}+\frac{23k^2\mathbf{k}\cdot\mathbf{q}}{756q^2}-\frac{121(\mathbf{k}\cdot\mathbf{q})^2}{756q^2}-\frac{5(\mathbf{k}\cdot\mathbf{q})^3}{27k^2q^2}\Bigg].
\end{aligned}
\tag{2.141}
$$

If we decompose $P_L(k)$ in self similar power law cosmologies using eq. (2.138), the one loop power spectrum becomes a sum of simple momentum integrals of the following form

$$
\int\frac{d^3q}{(2\pi)^3}\frac{1}{q^{2\nu_1}|\mathbf{k}-\mathbf{q}|^{2\nu_2}}\equiv k^{3-2\nu_{12}}I(\nu_1,\nu_2),
\tag{2.142}
$$

where ν_1 and ν_2 are in general complex numbers. The function $I(\nu_1,\nu_2)$ is dimensionless and can be easily calculated using the standard technique with Feynman parameters. The result is a well known expression see [90]

$$
I(\nu_1,\nu_2)=\frac{1}{8\pi^{3/2}}\frac{\Gamma\left(\frac{3}{2}-\nu_1\right)\Gamma\left(\frac{3}{2}-\nu_2\right)\Gamma\left(\nu_{12}-\frac{3}{2}\right)}{\Gamma(\nu_1)\Gamma(\nu_2)\Gamma(3-\nu_{12})},
\tag{2.143}
$$

where $\nu_{12}=\nu_1+\nu_2$. Using eq. (2.138) in the P_{22} term, eq. (2.50) we obtain

$$
\bar{P}_{22}(k)=2\sum_{m_1,m_2}c_{m_1}c_{m_2}\sum_{n_1,n_2}f_{22}(n_1,n_2)k^{-2(n_1+n_2)}\int\frac{d^3\mathbf{q}}{(2\pi)^3}\frac{1}{q^{2\nu_1-2n_1}|\mathbf{k}-\mathbf{q}|^{2\nu_2-2n_2}},
\tag{2.144}
$$

where c_{m_1} and c_{m_2} are the coefficients that contains the cosmological information coming from the FFT, while n_1 and n_2 are integer powers of q^2 and $|\mathbf{k}-\mathbf{q}|^2$ in the expanded $F_2^2(\mathbf{k}-\mathbf{q},\mathbf{q})$. The numerical coefficients are inside the function $f_{22}(n_1,n_2)$ and finally the complex exponent are $\nu_i=-1/2(\nu+i\eta_{m_i})$. Then using eq. (2.142) we have

$$
\bar{P}_{22}(k)=k^3\sum_{m_1,m_2}c_{m_1}k^{-2\nu_1}\,M_{22}(\nu_1,\nu_2)\,c_{m_2}k^{-2\nu_2},
\tag{2.145}
$$

where we have defined the matrix

$$
M_{22}(\nu_1,\nu_2)=\frac{\left(\frac{3}{2}-\nu_{12}\right)\left(\frac{1}{2}-\nu_{12}\right)\left[\nu_1\nu_2\left(98\nu_{12}^2-14\nu_{12}+36\right)-91\nu_{12}^2+3\nu_{12}+58\right]}{196\nu_1(1+\nu_1)\left(\frac{1}{2}-\nu_1\right)\nu_2(1+\nu_2)\left(\frac{1}{2}-\nu_2\right)}I(\nu_1,\nu_2).
\tag{2.146}
$$

Similar results hold for P_{13}, see ref. [44].

As we said, the cosmology dependence is all inside the c_m terms. For fixed number of sample points N, scale interval $[l_{\min}, k_{\max}]$ and bias ν, the matrix defined in eq. (2.146) is fixed and can be calculated once and for all. The calculation of the P_{22} integral then resides only in the calculation of one FFT in order to determine the c_m coefficients and then perform a matrix multiplication (2.145). Note that in eq. (2.140) and (2.141) we have shown the EdS kernels with fixed numerical coefficients, but we are interested in cosmologies with general time dependence, like ΛCDM and nDGP, that we will introduce in chapter (3). The same arguments leading to QFT-like integrals of the form (2.142) hold for these kind of cosmologies: as shown in [75] also the tracer density and velocity kernels can be decomposed in power of k^2, q^2 and $|\mathbf{k}-\mathbf{q}|^2$, allowing us to use FFTLog also for these models.

This procedure much faster than the direct computation and it has been shown to be quite accurate for a good choice of the bias parameter ν, [43, 44].

2.7 Conclusions

The state of the art measurements of cosmological parameters from LSS come from the analysis of the data from the Baryon Oscillation Spectroscopic Survey (BOSS, [91]). The full analysis of the power spectrum and the bispectrum was presented in [48, 49]. In [50] the authors claimed the detection of the BAO feature in the galaxy bispectrum, with consistent results with the power spectrum ones.

In [45, 46, 92, 47], the groups employ, for the first time, the fully consistent, PT based model for the anisotropic, redshift-space power spectrum, based on the EFTofLSS, including infrared resummation, described in eq. (2.104) and (2.133). The parameter space has been explored by using the common MCMC method, with a computation of the theoretical model at each step, while in most previous analysis the cosmology was fixed to a fiducial one and deviations from the fiducial cosmology were quantified using the Alcock-Paczynski parametrization [93]. In particular, the groups found that cosmological parameters constraints can be set independently from Planck results, i.e., without assuming any prior based on CMB experiments.

Chapter 3

SPT beyond ΛCDM

In chapter 2 we have presented the results for the one-loop PS and the tree-level BS using a perturbative approach and assuming an EDS background. For standard cosmologies, meaning within general relativity, we have $\Omega_m/f_+^2 \simeq 1$, which is exact in EdS, see [94]. Taking $\Omega_m/f_+^2 = 1$ exactly removes all the explicit time dependencies in the equations of motion of dark matter overdensity and velocity, which reflects in the time independence of the numerical coefficients of the kernels, see eqs. (2.43) and (2.44). After performing all the calculations in a EdS framework, one can substitute the scale factor with the linear growth function, i.e. $a(\tau) \to D_+(\tau)$, in order to have more accurate results, see [95]. The precision of this approximation was investigated, for example, by [56] who found that it performs to better than 1% at redshift $z = 0$ for $k \leq 0.2$ h/Mpc, and 0.1% at $z = 1$ on the same scales scales. For higher redshifts, the approximation performs better, since one enters the matter dominated era. Although for smaller scales and smaller redshifts it degrades, this approximation is used by many perturbative approaches [96, 45, 46], in order to simplify the calculations. In this chapter we will see how to include exact time dependence both in the equation of motion and in the perturbative solutions by using the Green's function approach. The importance of including these effects resides in the future possibility of distinguishing between beyond-ΛCDM models: these can include both time- and scale-dependent modifications, like the nDGP model [97], which will be the subject of study of the second part of this chapter, or the Hu-Sawicki $f(R)$ model [98]. These modified gravity toy models have been studied in detail with N-body simulations [99] and tested against observations [100, 101, 102].

Future surveys like the Euclid satellite will make possible to distinguish between different models, giving us more insight in the large scale structure of the universe.

3.1 Green's function approach

We recall that the Fourier space equations of motion for the dark matter overdensity, δ, and the rescaled velocity divergence, θ, are

$$\begin{aligned}
&\partial_\eta \delta_{\mathbf{k}} - f_+ \theta_{\mathbf{k}} = f_+ \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2} \alpha(\mathbf{q}_1,\mathbf{q}_2) \theta_{\mathbf{q}_1} \delta_{\mathbf{q}_2} \,, \\
&\partial_\eta \theta_{\mathbf{k}} - f_+ \theta_{\mathbf{k}} + \frac{3}{2}\frac{\Omega_m}{f_+}\mu\, \theta_{\mathbf{k}} + \frac{1}{f_+}\frac{k^2}{\mathcal{H}^2}\phi_{\mathbf{k}} = f_+ \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2} \beta(\mathbf{q}_1,\mathbf{q}_2) \theta_{\mathbf{q}_1} \theta_{\mathbf{q}_2} \,,
\end{aligned} \tag{3.1}$$

where α and β are the standard dark matter interaction vertices,

$$\alpha(\mathbf{q}_1,\mathbf{q}_2) = 1 + \frac{\mathbf{q}_1\cdot\mathbf{q}_2}{q_1^2} \qquad \text{and} \qquad \beta(\mathbf{q}_1,\mathbf{q}_2) = \frac{|\mathbf{q}_1+\mathbf{q}_2|^2 \mathbf{q}_1\cdot\mathbf{q}_2}{2q_1^2 q_2^2} \,. \tag{3.2}$$

The function $\mu = \mu(\eta)$ in the Euler equation carries possible linear modifications to the gravitational interaction. In GR one simply has $\mu(\eta) = 1$, and this is the case we will consider in this section. We will see in section (3.2) a specific model in which $\mu(\eta) \neq 1$. In general relativity, one closes these two equations with the Poisson equation (2.9). As we will see in more detail in sec. (3.2), the function μ modifies the Poisson equation at linear order, while other functions will be needed in order to take into account non-linear modifications.

We are now interested in the exact time dependence evolution within a ΛCDM background, where

$$H(\eta) = H_0 \sqrt{\Omega_{m,0} e^{-3\eta} + \Omega_{\Lambda,0}}, \quad \Omega_m(\eta) = \Omega_{m,0} \frac{H_0^2}{H^2(\eta)} e^{-3\eta} \,. \tag{3.3}$$

Note that also in modified gravity models where other forms of dark matter or dark energy are present one usually assumes the background evolution equations in eq. (3.3), and this is what is also assumed in the N-body simulations where these models are implemented (see section 3.2 and [99]). In the previous section we have considered the EdS approximation where $\Omega_m(\eta) = \Omega_{m,0} = 1$, $f_+ = 1$ and $\mathcal{H} = H_0 e^{-\frac{\eta}{2}}$. This assumption ensures that the numerical coefficients of the kernels that appear in the perturbative solutions of eq. (3.1) are fixed and time independent.

In the exact time dependence case we will make use of the Green's functions method to find the perturbative solutions to the equations of motion. We can find a n-th order solution of the form [103]

$$\begin{aligned}\delta^{(n)}_{\mathbf{k}}(\eta) &= \int_{-\infty}^{\eta} d\tilde{\eta}\left[G^{\delta}_{1}(\eta,\tilde{\eta})S^{\delta}_{n}(\mathbf{k},\eta) - G^{\delta}_{2}(\eta,\tilde{\eta})S^{\theta}_{n}(\mathbf{k},\eta)\right] ,\\ \theta^{(n)}_{\mathbf{k}}(\eta) &= \int_{-\infty}^{\eta} d\tilde{\eta}\left[G^{\theta}_{1}(\eta,\tilde{\eta})S^{\delta}_{n}(\mathbf{k},\eta) - G^{\theta}_{2}(\eta,\tilde{\eta})S^{\theta}_{n}(\mathbf{k},\eta)\right] ,\end{aligned} \tag{3.4}$$

where S^{σ}_{n}, with $\sigma = \{\delta, \theta\}$ are the n-th order source terms of the continuity, for $\sigma = \delta$, and the Euler equation, for $\sigma = \theta$. Substituting this solution into the equations of motion, we have

$$\begin{aligned}&\frac{dG^{\delta}_{\sigma}(\eta,\tilde{\eta})}{d\eta} - f_{+}G^{\theta}_{\sigma}(\eta,\tilde{\eta}) = \lambda_{\sigma}\delta_{D}(e^{\eta} - e^{\tilde{\eta}}) ,\\ &\frac{dG^{\theta}_{\sigma}(\eta,\tilde{\eta})}{d\eta} - f_{+}G^{\theta}_{\sigma}(\eta,\tilde{\eta}) + \frac{3}{2}\frac{\mu\Omega_{m}}{f_{+}}\left[G^{\theta}_{\sigma}(\eta,\tilde{\eta}) - G^{\delta}_{\sigma}(\eta,\tilde{\eta})\right] = (1+\lambda_{\sigma})\,\delta_{D}(e^{\eta} - e^{\tilde{\eta}})\end{aligned} \tag{3.5}$$

where $\lambda_1 = 1$ and $\lambda_2 = 0$. The Green's functions written above describe the response of the two fields δ and θ when the continuity equation and the Euler equation are perturbed. The system above can be solved analytically by using the linear growth and decay functions $D_{\pm}$, see ref. [75] for a full treatment of this method within wCDM cosmologies.

After having calculated all the Green's functions that account for the time evolution, the steps towards the 1-loop PS for biased tracers in redshift space are the same described in the previous chapter. The only main difference is that the convolution kernels introduced in eq. (2.39) and (2.40) will be time dependent. Schematically, for the density kernels, we have

$$F_{n}(\mathbf{q}_{1},\ldots,\mathbf{q}_{n}) \longrightarrow F_{n}(\mathbf{q}_{1},\ldots,\mathbf{q}_{n};\eta) .$$

For example, in ΛCDM the second order density kernel is

$$\begin{aligned}F_{2}(\mathbf{q}_{1},\mathbf{q}_{2};\eta) &= \beta(\mathbf{q}_{1},\mathbf{q}_{2}) + (\alpha_{s}-\beta)\,(\mathbf{q}_{1},\mathbf{q}_{2})\mathcal{G}^{\delta}_{1}(\eta)\\ &= \mathcal{G}^{\delta}_{1}(\eta) + \frac{\mathbf{q}_{1}\cdot\mathbf{q}_{2}}{2q_{1}q_{2}}\left(\frac{q_{1}}{q_{2}} + \frac{q_{2}}{q_{1}}\right) + \left(1-\mathcal{G}^{\delta}_{1}(\eta)\right)\frac{(\mathbf{q}_{1}\cdot\mathbf{q}_{2})^{2}}{q_{1}^{2}q_{2}^{2}} ,\end{aligned} \tag{3.6}$$

where $\mathcal{G}^{\delta}_{1}$ is a higher order Green's function, which, in ΛCDM, is expressed by

$$\mathcal{G}^{d}_{1}(\eta) = \int_{-\infty}^{\eta} d\tilde{\eta}\, G^{\delta}_{1}(\eta,\tilde{\eta})f_{+}(\tilde{\eta})\frac{D^{2}_{+}(\tilde{\eta})}{D^{2}_{+}(\eta)}e^{\tilde{\eta}} . \tag{3.7}$$

It is easy to verify that in the EdS limit $\mathcal{G}^{\delta}_{1} \to 5/7$ and the kernel in eq. (3.6) reproduces exactly the one in eq. (2.43).

The results for biased tracers in redshift space follow the same arguments described above. The bias expansion introduced in sect. (2.3.1) was derived using only symmetry principles of the equation of motion. For this reason, all those cosmological models that share the same symmetries will also need the same number of bias parameters. At third order in PT, the number of the deterministic biases, for models that satisfy the EP, is seven. For the exact time dependence, this was explicitly shown in [75]. In chapter (5) we will show how to reproduce this number of deterministic biases, only by considering symmetry principles without specifying the equations of motion.

The result for the one-loop power spectrum will simply be a generalization of eq. 2.104, for a full treatment see [75]. Notice that all these results are obtained for a generic time dependence for the equation of motion and assuming only a ΛCDM background. This means that all the results are completely general and are still valid for models where a time-dependent modification of gravity has been introduced and that satisfy the same symmetries of ΛCDM. The Green's function approach can be used to infer cosmological parameters for any model that present a time-dependent growth for the inhomogeneities. The difference between different models will reside only in the time dependent Green's function G_i^λ and the higher order functions such as $\mathcal{G}_i^\lambda$ and so on.

3.1.1 Comparison with EdS approximation

In figures (3.1) we show the result of an MCMC analysis against the N-body simulations presented in [104][1]. In particular we show the difference between the EdS case and the full exact-time ΛCDM. As you can see, the effect of considering exact-time does not affect the analysis, at least at this perturbative order (1-loop) and for the covariances used (that are, however, very small compared to the ones used for actual data analysis).

In fig. (3.2) we show the result of an MCMC analysis for the two redshifts of the BOSS dataset, CMASS2 ($z = 0.61$) and LOWZ ($z = 0.38$). The theoretical model is the one used in the EFTofLSS, assuming ΛCDM background. We have used the public code PyBird [10], where the one loop PS in redshift space for biased tracers is implemented, together with the

[1] All the details about the simulations and the fiducial values of the cosmological parameters can be found here.

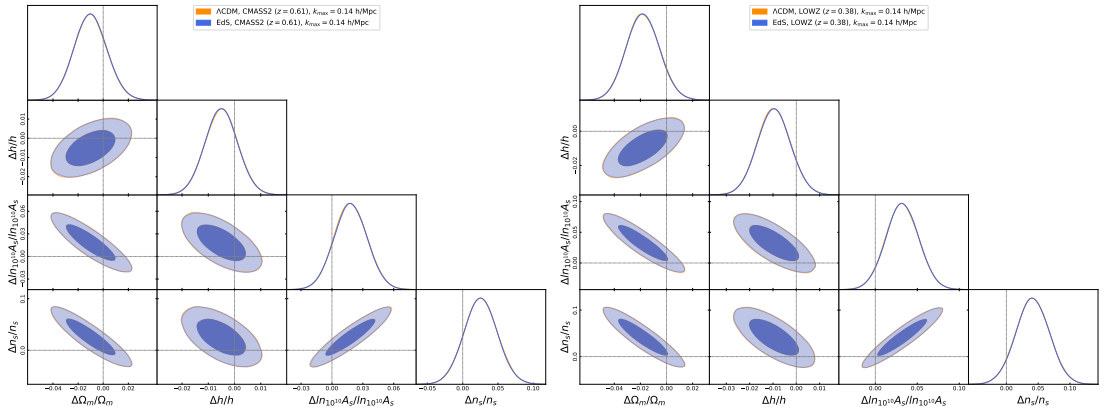

Figure 3.1: Comparison of the constraints on cosmological parameters using, in one case, the EdS approximation and, in the other, the exact time dependence of ΛCDM. The analysis was performed on the PT challenge simulations [104], for which the true values of the cosmological parameters are not public.

MCMC code Montepython [84]. The cutoff used is for both redshifts $k_{\rm max} = 0.14$ h/Mpc, for which the cosmological parameters measured are unbiased. The results of this analysis show compatible constraints between different redshifts, as one would expect. These analysis were performed without assuming any Planck prior. We only assumed a BBN prior on the baryon density, following [45, 46].

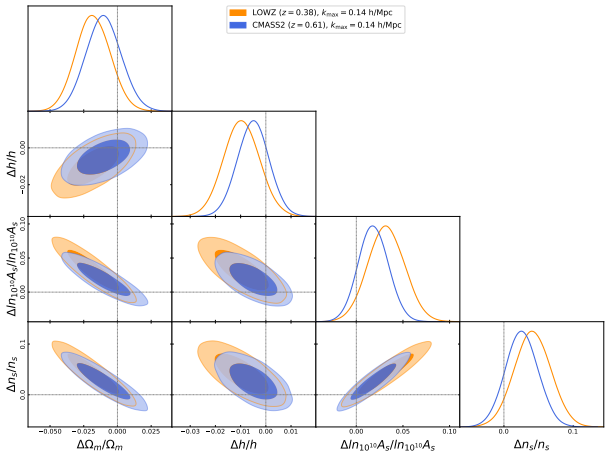

Figure 3.2: Posterior distribution for the cosmological parameters after the analysis of the PT challenge simulations using EFTofLSS in ΛCDM.

In the next section we will introduce the MG model nDGP. We will present a forecast of the

constraining power of BOSS-like surveys and introduce the future analysis we will perform on real data.

3.2 nDGP model

In theories beyond ΛCDM, where the linear growth is still scale-independent, like for instance in the normal branch of the Dvali, Gabadadze and Porrati (nDGP) braneworld model [9] or in the EFT of dark energy beyond linear order [105, 106] (see [103] for a comparison with simulations), new nonlinear couplings emerge. Their effect on the time-dependence of the coefficients can be derived by including them on the RHS of eq. (3.1), at the appropriate order.

In particular we will study the DGP model [9] where it is assumed that universe we live in is a four dimensional brane enclosed in a five-dimensional Minkowski spacetime. Thus, gravitational interaction is five-dimensional on large scales and reduces to a four-dimensional force, recovering GR, on small scales. The transition from five to four-dimensional happens at the so called crossover scale r_c, defined by the ratio between the five and the four-dimensional gravitational constants

$$r_c = \frac{G^{(5)}}{2G^{(4)}} \,. \tag{3.8}$$

The crossover scale represents the only free parameter of the theory. In this case the modified Friedman equation is given by

$$\varepsilon \frac{H}{r_c} = H^2 \left(1 - \Omega_{\rm m}(\eta)\right) \,, \tag{3.9}$$

where H is the Hubble parameter, Ω_m is fractional energy density of dark matter and $\varepsilon = \pm 1$ parametrizes the two DGP models: $\varepsilon = 1$ is the self-accelerated branch (*s*DGP), which was found to be unstable, see [107], while $\varepsilon = -1$ represents the normal branch (*n*DGP), which will be the subject of this study. In the latter model, accelerated expansion is achieved through a dynamical dark energy component [97], which reproduces an analogous behavior to the cosmological constant in ΛCDM. The general Hubble parameter is

$$H(\eta) = H_0 \left(\sqrt{\Omega_{m,0} e^{-3\eta} + \Omega_{DE}(\eta) + \Omega_{rc}} - \sqrt{\Omega_{rc}} \right) \,, \tag{3.10}$$

where Ω_{DE} is the density parameter of the dark energy component, and

$$\Omega_{rc} = \frac{1}{4r_c^2 H_0^2}\,. \tag{3.11}$$

Usually, in N-body simulations, the dark energy component is tuned so that the background history recovers the ΛCDM one, and only the perturbative corrections are studied. Matching eq. (3.10) with eq. (3.3) we obtain

$$\Omega_{DE}(\eta) = \Omega_\Lambda + 2\sqrt{\Omega_{rc}}\sqrt{\Omega_{m,0}e^{-3\eta} + \Omega_\Lambda}\,. \tag{3.12}$$

The theory also provides a screening mechanism for the small scales, for more details see [9, 97]. From now on, we will consider the ΛCDM background expressed in eq. 3.3.

The modified Poisson equation entering in the Euler equation (3.1) now reads, in Fourier space [103, 108]

$$-\frac{k^2}{\mathcal{H}^2}\phi_{\mathbf{k}}(\eta) = \frac{3\,\Omega_{\rm m}}{2}\,\mu(\eta)\delta_{\mathbf{k}} + S(\mathbf{k},\eta)\,, \tag{3.13}$$

where $\delta_{\mathbf{k}}$ is the dark matter overdensity. The function $\mu(\eta)$ can, in general, depend both on time and scale and characterize the linear modification to the gravitational interaction. Here we will study only time-dependent modifications to the clustering equations, but, in general, a scale-dependence could also be present. In the nDGP theory μ reads [106]

$$\mu(\eta) = 1 + \frac{1}{3\mathcal{B}(\eta)}\,, \qquad \mathcal{B}(\eta) = 1 + \frac{H(\eta)}{H_0}\frac{1}{\sqrt{\Omega_{rc}}}\left(\frac{2}{3} + \frac{1}{3}\partial_\eta \log\mathcal{H}\right)\,. \tag{3.14}$$

The source term $S(\mathbf{k},\eta)$ in eq. (3.13) is responsible for new mode couplings in modified gravity theories, including screening mechanism. In nDGP it reads

$$\begin{aligned} S(\mathbf{k},\eta) =& \mu_2(\eta)\left(\frac{3\,\Omega_{\rm m}}{2}\right)^2 \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2}\gamma(\mathbf{q}_1,\mathbf{q}_2)\delta_{\mathbf{q}_1}\delta_{\mathbf{q}_2} \\ &+ \mu_3(\eta)\left(\frac{3\,\Omega_{\rm m}}{2}\right)^3 \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}\gamma_3(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3)\delta_{\mathbf{q}_1}\delta_{\mathbf{q}_2}\delta_{\mathbf{q}_3} \\ &+ \mu_{22}(\eta)\left(\frac{3\,\Omega_{\rm m}}{2}\right)^3 \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}\gamma(\mathbf{q}_1,\mathbf{q}_2)\gamma(\mathbf{q}_{12},\mathbf{q}_3)\delta_{\mathbf{q}_1}\delta_{\mathbf{q}_2}\delta_{\mathbf{q}_3}\,, \end{aligned} \tag{3.15}$$

where the new kernels inside the integrals are given by

$$\begin{aligned} \gamma(\mathbf{k}_1,\mathbf{k}_2) &= 1 - \left(\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2\right)^2\,, \\ \gamma_3(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) &= 1 + 2\left(\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2\right)\left(\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_3\right)\left(\hat{\mathbf{k}}_2\cdot\hat{\mathbf{k}}_3\right) - \left(\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_3\right)^2 - \left(\hat{\mathbf{k}}_2\cdot\hat{\mathbf{k}}_3\right)^2 - \left(\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2\right)^2\,, \end{aligned} \tag{3.16}$$

where $\hat{\mathbf{k}} \equiv \mathbf{k}/k$. We anticipate that the cubic vertex proportional to μ_3 does not contribute to the power spectrum at one loop because it enters as $\gamma_3(\mathbf{k}, \mathbf{q}, -\mathbf{q}) = 0$ [106], but we have included it here for completeness. The new time dependent functions are given by

$$\mu_2(\eta) = -2H^2 r_c^2 \left(\frac{1}{3\mathcal{B}}\right)^3 , \qquad \mu_{22}(\eta) = 8H^4 r_c^4 \left(\frac{1}{3\mathcal{B}}\right)^5 . \tag{3.17}$$

As discussed above, we do not need to specify μ_3. General relativity is recovered by sending $r_c \to \infty$, where we have $\mu(\eta) \to 1$ and $\mu_2, \mu_{22} \to 0$.

Neglecting the term proportional to μ_3, the perturbation equations above then become

$$\partial_\eta \delta_{\mathbf{k}} - f_+ \theta_{\mathbf{k}} = S^\delta_{\mathbf{k}} , \tag{3.18}$$

$$\partial_\eta \theta_{\mathbf{k}} - f_+ \theta_{\mathbf{k}} + \frac{3}{2}\frac{\Omega_m}{f_+}\mu(\theta_{\mathbf{k}} - \delta_{\mathbf{k}}) = S^\theta_{\mathbf{k}} , \tag{3.19}$$

with

$$\begin{aligned} S^\delta_{\mathbf{k}} =& f_+ \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2} \alpha(\mathbf{q}_1, \mathbf{q}_2) \theta_{\mathbf{q}_1} \delta_{\mathbf{q}_2} , \\ S^\theta_{\mathbf{k}} =& f_+ \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2} \left[\beta(\mathbf{q}_1, \mathbf{q}_2) \theta_{\mathbf{q}_1} \theta_{\mathbf{q}_2} + \frac{\mu_2}{f_+^2} \left(\frac{3\Omega_{\rm m}}{2}\right)^2 \gamma(\mathbf{q}_1, \mathbf{q}_2) \delta_{\mathbf{q}_1} \delta_{\mathbf{q}_2} \right] \\ &+ \frac{\mu_{22}}{f_+} \left(\frac{3\Omega_{\rm m}}{2}\right)^3 \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3} \gamma(\mathbf{q}_1, \mathbf{q}_2) \gamma(\mathbf{q}_{12}, \mathbf{q}_3) \delta_{\mathbf{q}_1} \delta_{\mathbf{q}_2} \delta_{\mathbf{q}_3} . \end{aligned}$$

From eqs. (3.18) and (3.19) we can find the evolution equation for the growth function

$$\frac{d^2 D}{d^2 \eta} + \left(1 + \frac{d \ln \mathcal{H}}{d\eta}\right) \frac{dD}{d\eta} - \frac{3}{2}\Omega_m \mu = 0 , \tag{3.20}$$

where $\mu(\eta)$ is the time dependent function introduced in 3.14. In figure (3.3) we show its time behaviour for different values of Ω_{rc}, chosen following [97]. From now on we can follow the same steps described in the ΛCDM case: we look for a perturbative solution of the form (3.4), and we write the galaxy kernels in redshift space. The only difference now will be in the time dependent functions $\{\mathcal{G}_1^\lambda, \mathcal{G}_2^\lambda, \mathcal{U}_\sigma^\lambda, \mathcal{V}_{\sigma\tilde{\sigma}}^\lambda\}$, see appendix (A). One can easily verify that the degeneracies found for the ΛCDM model, are still valid, meaning that the tracers kernel in redshift space have the same form of those presented in [75].

3.3 Results

PyBird [10] is a Python code for the evaluation of the power spectrum of biased tracers in redshift space. The theoretical model used to calculate the one loop power spectrum is the

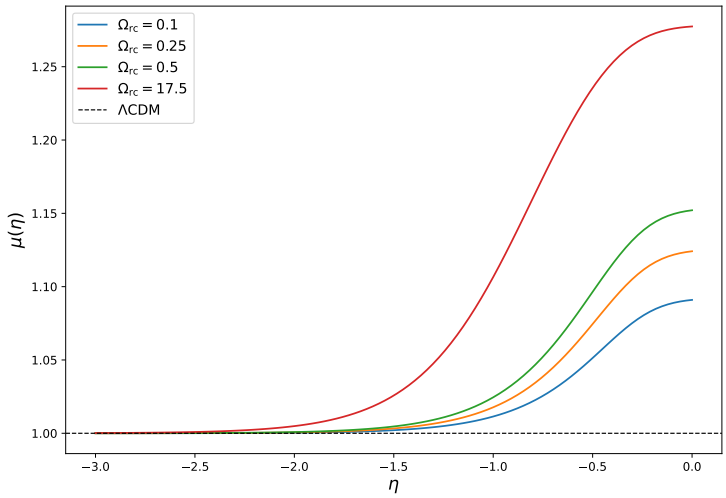

Figure 3.3: Comparison of μ as a function of the time variable η for different values of Ω_{rc}. The ΛCDM case is also shown for comparison.

EFTofLSS, the one described in the previous section and presented explicitly in [75]. The modular nature of this code[2] allows an easy implementation of the nDGP model we have described so far. The main modification we have implemented consists in solving numerically the differential equation for the growth function and the higher order Green's functions in the nDGP model, see app. (A). This procedure can easily be generalized to MG models with scale-independent growth, like, for example, clustering quintessence [109] (see [110] for the implementation of this model inside PyBird and constraints on it within the BOSS analysis).

At the linear level, nDGP modifies the linear growth function D_+ via the function μ in eq. (3.20). In figure (3.4) we show the results of the numerical solution for the linear growth function. As you can see, realistic values for Ω_{rc}, see [102], cause a percent deviation from ΛCDM for recent times.

This departure from ΛCDM is even more visible when the non-linear power spectrum, as shown in fig. (3.5).

The modification to D_+ due to nDGP is completely degenerate with the normalization of the power spectrum, parametrized by A_s. Therefore, using only one redshift observation

[2]See here for the public GitHub repository.

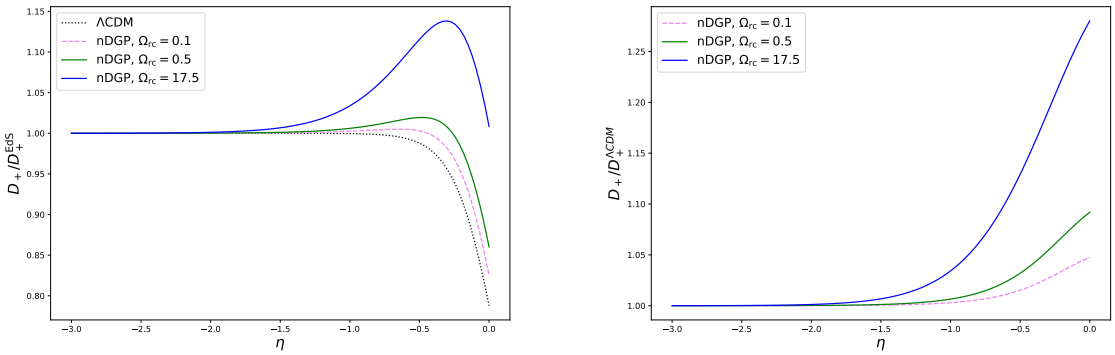

Figure 3.4: Left: comparison between growth functions in ΛCDM and nDGP (divided by the EdS). Right: ratio of growth function in ΛCDM and nDGP. We have used $\Omega_{m,0} = 0.315$.

.

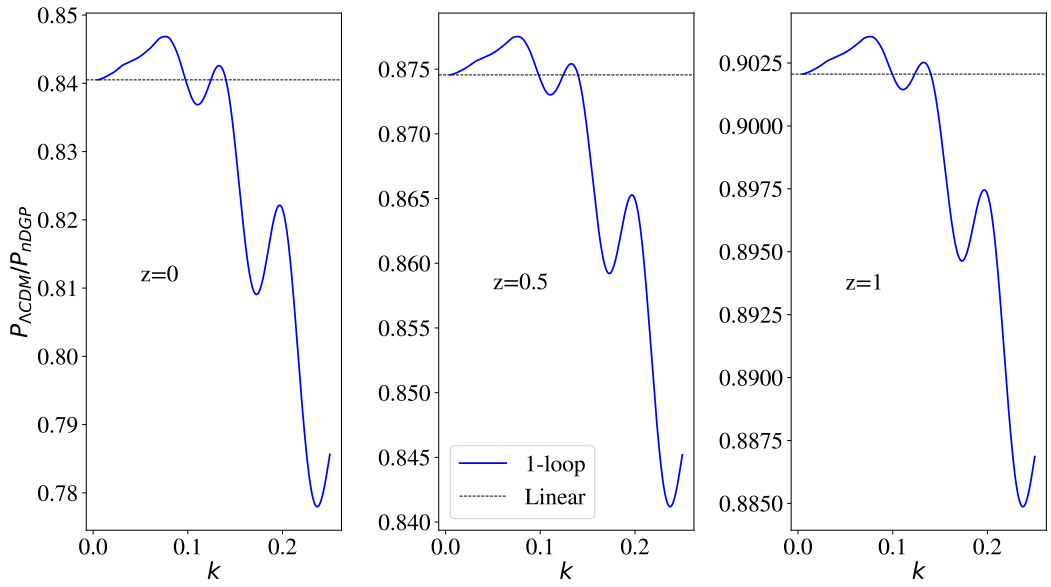

Figure 3.5: Comparison between the non-linear matter power spectra in ΛCDM cosmology and nDGP one at different redshifts. The cosmological parameters used are $A_s = 2.04 \times 10^{-9}$, $n_s = 0.966$, $h = 0.673$, $\omega_b = 0.022$, $\omega_c = 0.12$ and $\Omega_{rc} = 0.5$ for the nDGP cosmology.

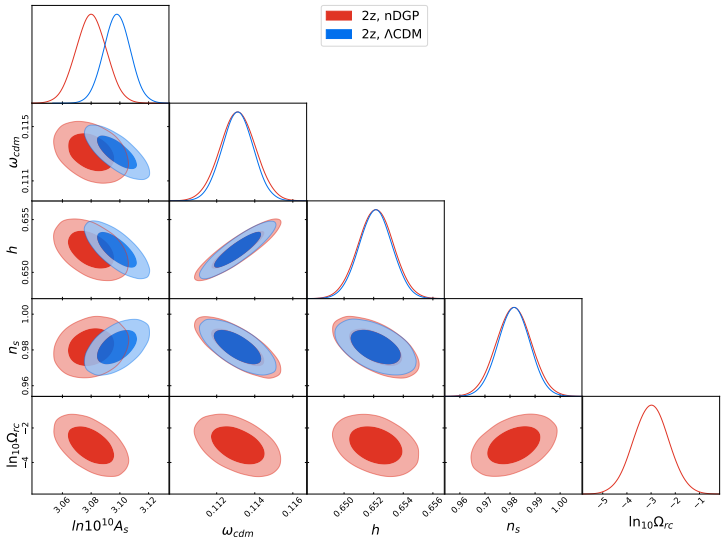

Figure 3.6: Fisher forecast using two redshift measurements of the PT challenge

for the PS will not give precise constraints on Ω_{rc} and A_s. Here we show, using a Fisher forecast [111], that using 2 redshift measurements will help in constraining Ω_{rc}. For the forecast we produce the PS monopole and quadrupole using PyBird with a fiducial cosmology $\{\ln_{10^{10}} A_s, \omega_{cdm}, h, n_s, \ln_{10}\Omega_{rc}\} = \{3.08, 0.113, 0.652, 0.9816, -3\}$ and for the covariances we use the public ones of the PT challenge simulations [104].

As you can see, a small value of Ω_{rc}, the fiducial value is $\Omega_{rc} = 10^{-3}$, causes a measurable effect on the scalar amplitude A_s. This promising result confirms our expectations on considering two redshifts for the analysis. Using different redshift bins, like, for example, in the Euclid mission, will increase even more the constraining power on these beyond-ΛCDM models.

The new version of the code is currently under testing and it will be used for the analysis of the BOSS dataset.

3.4 Future developments and conclusions

In this chapter we have briefly presented the Green's function approach for the exact-time dependence in models beyond the EdS approximation. The simplest model we considered is the ΛCDM model, and we have shown that the constraints on cosmological parameters

obtained in the two different models are the same, at least at this order, as shown in fig. (3.2). This is due to the accuracy of EdS in reproducing the ΛCDM theory, as also noted in ref. [75].

When beyond-ΛCDM models are considered, the exact-time dependence inclusion is crucial. Here we have considered a particular scale-independent MG model, nDGP. We have shown the results of the implementation inside PyBird, and we analyzed the results of the Fisher forecast on BOSS-like surveys. Although very promising, we expect to set only an upper limit to the new parameter Ω_{rc} when the BOSS dataset will be analyzed, which we leave for future studies. The new version of the code will be used for the analysis of the Euclid satellite dataset.

The natural prosecution of this work will be the inclusion of scale-dependent models, in order to account, for example, for massive neutrinos. We leave the exploration of these generalizations to future work.

Chapter 4

Consistency relations of the large scale structure

4.1 Equivalence Principle and BAO

The large scale structure of the Universe (LSS) is governed by nonlinear effects of different nature: the evolution of the dark matter (DM) field, redshift space distortions (RSD), and the bias of the field for the considered tracers (galaxies, halos...) with respect to the DM one. All these effects limit the application of analytical techniques to rather large scales, thus excluding large part of the data from actual analyses. It is therefore remarkable that fully nonlinear statements can be made, in the form of "consistency relations" (CR) [4, 5, 6, 7]. These are statements about the effect of perturbations at large scales on small scales ones, expressed in terms of relations between correlation functions of different order. As highlighted also in [6, 112], the equivalence principle affirms that one cannot measure a uniform gravitational field, ϕ_L, produced by a long-wavelength (linear) mode of a density perturbation with a local experiment. This translates into the fact that the effect of a uniform gravitational field can be reabsorbed performing a change of coordinates. Then, the n-points correlators of the density perturbations in presence of a long-mode gravitational potential, are the same as the n-points correlators of density perturbations in absence of the long mode but in transformed coordinates,

$$\langle \delta(\mathbf{x}_1, \eta_1) \dots \delta(\mathbf{x}_n, \eta_n) | \Phi_L \rangle = \langle \delta(\tilde{\mathbf{x}}_1, \tilde{\eta}_1) \dots \delta(\tilde{\mathbf{x}}_n, \tilde{\eta}_n) \rangle \,. \tag{4.1}$$

The content of the CR's is essentially kinematic: as explicited in eq. (4.1), it describes the effect of a long wavelength displacement field on short distance fluctuations. In linear perturbation theory, assuming adiabatic and gaussian initial conditions, the displacement field is directly related to a long wavelength matter density fluctuation, with a coefficient growing as $1/q$ as $q \to 0$, where q is the long wavelength wavenumber. Since a uniform displacement cannot have any consequence on equal-times correlation functions [54], CR's exhibit a characteristic $1/q$ growth only when considering unequal time correlators. On the other hand, for equal-time correlators the behavior of correlation functions as prescribed by CR's is in general 'obscured' by other terms, whose form is not dictated by the CR's and are formally of the same order in q. Therefore the practical use of CR's appears to be of a limited extent, as unequal-time correlation functions are not measurable (notice that "unequal-times" does not mean "unequal redshifts", as the latter involve density fields in regions inside the past lightcone, which is not the case for the former).

In particular we will consider the effect of a long-wavelength mode on the bispectrum (BS).

Assuming adiabatic and gaussian initial conditions, CR's single out a nonperturbative contribution to the squeezed limit BS, where by *nonperturbative* here it is meant that it does not rely on any approximation scheme (like, for instance Perturbation Theory (PT)) and that the result holds even beyond the perfect fluid approximation. By further assuming that the very 'long' modes at scale q can be described by linear PT (but with no approximation on the 'short' ones at scale k!) the CR's take the form of exact relations between the BS and the PS's evaluated at the scales k and q (see eq. (4.20) below). For the above reasons, the CR's hold not only for matter but for any tracer, both in real and in redshift space [113, 114]. The potential of CR's in constraining possible violations of gaussianity of the initial conditions has been investigated in [7, 115], while for violations of the EP see [113, 116].

At first sight, for theories respecting the EP and assuming gaussian initial conditions, the nontrivial content of the CR's might look empty, as in this case the contribution to the BS 'protected' by the CR's is parametrically of the same order, $O((k/q)^0)$, of other terms induced by different sources of nonlinearities at any PT order, for which theoretical control is limited by the reliability of PT in the considered range. However, as it was shown in [2, 3], the wiggly feature of Baryonic Acoustic Oscillations imprinted in the spectra of the tracers of LSS in

the late-time universe provide a way to isolate the CR-protected contribution to the PS from unprotected ones. The latter, although being parametrically of the same order in k/q in the squeezed limit, are either smooth or suppressed by factors $O(2\pi/(kr_s))$ with respect to the protected ones, where $r_s \simeq (100\,\mathrm{Mpc\,h^{-1}})$ is the BAO acoustic scale. Therefore, both in real and redshift space, by comparing the BAO amplitudes in the BS and in the PS we were able to measure the prefactor of the latter which is related to the linear bias parameter and the growth function, $f = d\ln D/d\ln a$. After fifteen years of the first clear detection of BAO in the galaxy correlation function by [18], the precision at which the BAO feature can be extracted from the distribution of galaxies has greatly been improved (see [11] for recent results from the PS and the correlation function). Furthermore, a high-significance detection of the BAO feature in the three-point functions has also been reported by [117]. Therefore, we expect that measurements of the CR's through the BAO feature in real data would be within the reach of large-scale experiments planned in the near future, such as LSST [118], Euclid [119], WFIRST [120].

In this chapter we follow the analysis first performed in [2, 3] both in real and redshift space, in which the CR's coefficients depend both on the large scale bias and the large scale growth function, $f = d\ln D/d\ln a$ (where D is the linear growth rate) thereby providing a way to break the degeneracy between the two. We will analyse large volume N-body simulations, confirming the validity of CR's. We will find that CR's alone are mostly sensitive to the bias b_α (where b_α is the linear bias of the tracer α), giving weak constraints on the parameter $\beta_\alpha = f/b_\alpha$. However, by combining the CR analysis with the independent extraction of the parameter β_α from the PS quadrupole to monopole ratio, the $f - \beta_\alpha$ degeneracy can be successfully broken. Moreover, we estimate the constraining potential of future surveys, in particular, Euclid [119], showing that it can reach better than 10% precision on the bias parameter, and therefore on f as well, in a manner completely free from assumptions on the biasing prescription as well as the underlying gravity theory.

The chapter is organised as follows. In sect. (4.2) we derive the CR's in the most general way for biased tracers in redshift space and show explicitly the relation between CR's and BAO physics; in sect. (4.3) we study the case when real space correlators are considered and measure the large scale galaxy bias against numerical simulations; in sect. (4.4) we generalize

the analysis to redhisft space, defining the multipoles approach and presenting also a forecast of the constraining power of this methodology when applied to future data from the Euclid survey. Finally in sect. (4.6) we summarize our conclusions and give our outlook on future developments.

The results of this chapter have been published in the two papers

- M. Marinucci, T. Nischimici and M. Pietroni, *Measuring Bias via the Consistency Relations of the Large Scale Structure*, *Phys. Rev.* **D100** (2019), no. 12 123537[1],
- M. Marinucci, T. Nishimichi and M. Pietroni, *Model independent measurement of the growth rate from the consistency relations of the LSS*, *JCAP* **07** (2020) 054[2],

and we have included these articles or part of them according to the journals copyright policies.

4.2 Consistency relations of the large scale structure

In this section, we derive the CR's in a way convenient for the purpose of this thesis, and, moreover, we specify to the equal-time limit, which was not treated in the original papers [4, 5].

4.2.1 General derivation

We consider the most general Boltzmann equation,

$$\left(\frac{\partial}{\partial\tau}+\frac{p_i}{am}\frac{\partial}{\partial x^i}-am\frac{\partial}{\partial x^i}\Phi(\mathbf{x},\tau)\frac{\partial}{\partial p_i}\right)f(\mathbf{x},\mathbf{p},\tau)=\mathcal{C}[f,\ldots](\mathbf{x},\mathbf{p},\tau)\,, \tag{4.2}$$

where $f(\mathbf{x},\mathbf{p},\tau)$ is the distribution function of a given species, not necessarily cold dark matter. The collision term at the RHS, takes into account possible non-gravitational interactions, and it involves f itself as well as the distribution functions of the other species taking part in the interactions. f could also represent the distribution of halos in a given mass range, or a

[1]Reprinted article with permission from M. Marinucci, T. Nischimici and M. Pietroni, *Measuring Bias via the Consistency Relations of the Large Scale Structure*, *Phys. Rev.* **D100** (2019), no. 12 123537 Copyright (2021) by the American Physical Society.

[2]

given type of galaxies, and in that case $\mathcal{C}$ would describe processes which change the comoving number density of these tracers, such as merging (see eq. (2.12) for the collisionless case). Eq. (4.2) is invariant under the time-dependent frame change

$$\mathbf{x} \to \bar{\mathbf{x}} = \mathbf{x} + \mathbf{d}(\tau)\,, \qquad \mathbf{p} \to \bar{\mathbf{p}} = \mathbf{p} + am\dot{\mathbf{d}}(\tau)\,, \tag{4.3}$$

provided we make the replacements

$$\begin{aligned} &f(\mathbf{x},\mathbf{p},\tau) \to \bar{f}(\mathbf{x},\mathbf{p},\tau) = f(\mathbf{x}-\mathbf{d}(\tau),\mathbf{p}-am\dot{\mathbf{d}}(\tau),\tau)\,,\\ &\frac{\partial}{\partial x^i}\Phi(\mathbf{x},\tau) \to \frac{\partial}{\partial x^i}\bar{\Phi}(\mathbf{x},\tau) = \frac{\partial}{\partial x^i}\Phi(\mathbf{x}-\mathbf{d}(\tau),\tau) - \mathcal{H}\dot{\mathbf{d}}(\tau) - \ddot{\mathbf{d}}(\tau)\,, \end{aligned} \tag{4.4}$$

and if the collisional term satisfies

$$\mathcal{C}[\bar{f},\ldots](\mathbf{x},\mathbf{p},\tau) = \mathcal{C}[f,\ldots](\mathbf{x}-\mathbf{d}(\tau),\mathbf{p}-am\dot{\mathbf{d}}(\tau),\tau)\,, \tag{4.5}$$

that is, the interaction rate is the same in the two frames. The transformation above is nothing but the equivalence Principle (EP), also called in this context the extended galilean invariance. Since it is an invariance of the Boltzmann equation, its consequences are not restricted to perturbation theory, but are valid at the fully nonlinear level, also including nonperturbative effects such as shell-crossing and multistreaming. Moreover, one should keep in mind that the symmetry holds for an *arbitrary* displacement $\mathbf{d}(\tau)$, independently on the identification of it as the infinite wavelength limit of a large scale cosmological perturbation. The last observation is crucial in order to disentangle the *dynamical* content of the consistency relation from the *statistical* one, related to the statistical properties of the cosmological perturbations such as adiabaticity and gaussianity.

The dynamical content is encoded in constraints on the mode-coupling vertices in the *soft* limit, that is, when one of the modes goes to zero. It is best analyzed in Fourier space, by replacing the homogeneous displacement $\mathbf{d}(\tau)$ with a scale dependent one,

$$\int d^3x\, e^{i\mathbf{x}\cdot\mathbf{q}}\mathbf{d}(\tau) = (2\pi)^3\delta_D(\mathbf{q})\mathbf{d}(\tau) \to \tilde{\mathbf{d}}(\mathbf{q},\tau)\,, \tag{4.6}$$

and then considering the $q \to 0$ limit. We will focus on the BS in redshift space

$$B^{(S)}_{\alpha\beta\gamma}(\mathbf{q},-\mathbf{k}_+,\mathbf{k}_-;\tau_\alpha,\tau_\beta,\tau_\gamma) \equiv \langle \delta^{(S)}_\alpha(\mathbf{q};\tau_\alpha)\delta^{(S)}_\beta(-\mathbf{k}_+;\tau_\beta)\delta^{(S)}_\gamma(\mathbf{k}_-;\tau_\gamma)\rangle'\,, \tag{4.7}$$

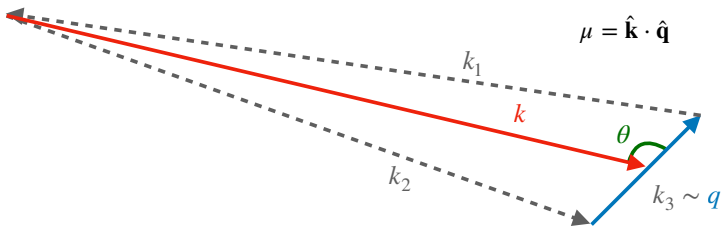

Figure 4.1: Graphic representation of the notation used for the triangle parametrization.

where the superscript (S) indicates that we are considering the redshift space quantities, $\mathbf{k}_\pm = \mathbf{k} \pm \frac{\mathbf{q}}{2}$, $q = |\mathbf{q}|$, $k_\pm = |\mathbf{k}_\pm|$, and the prime indicates that the expectation value has been divided by a $(2\pi)^3\delta_D(0)$ factor. $\delta^{(S)}_{\alpha,\beta,\gamma}$ indicate the density contrasts for different tracers (e.g. DM, baryons, a given galaxy type, ...), evaluated at times $\tau_{\alpha,\beta,\gamma}$, respectively.

By moving to another frame, the transformations (4.3)-(4.5) dictate the transformation of the BS. The density contrasts (obtained from the first moments of the distribution function) transform as

$$\delta^{(S)}(\mathbf{k},\tau) \to \bar{\delta}^{(S)}(\mathbf{k},\tau) = \delta^{(S)}(\mathbf{k},\tau) + \, iI_{\mathbf{k};\mathbf{q}',\mathbf{p}}\mathbf{p}\cdot\tilde{\mathbf{d}}(\mathbf{q}',\tau)\delta^{(S)}(\mathbf{p},\tau) + \cdots \,, \tag{4.8}$$

where

$$I_{\mathbf{k};\mathbf{p}_1,\mathbf{p}_2} \equiv \int \frac{d^3p_1}{(2\pi)^3}\frac{d^3p_2}{(2\pi)^3}(2\pi)\delta_D(\mathbf{k}-\mathbf{p}_1-\mathbf{p}_2)\,, \tag{4.9}$$

and the dots indicate higher orders in $\tilde{\mathbf{d}}$. Inserting it in (4.7) the additional contributions to the BS induced by the change of frame are obtained,

$$\begin{aligned} &iI_{-\mathbf{k}_+;\mathbf{q}',\mathbf{p}}\langle\mathbf{p}\cdot\tilde{\mathbf{d}}(\mathbf{q}',\tau_\beta)\delta^{(S)}_\alpha(\mathbf{q},\tau_\alpha)\delta^{(S)}_\beta(\mathbf{p},\tau_\beta)\delta^{(S)}_\gamma(\mathbf{k}_-;\tau_\gamma)\rangle' \\ &+(-\mathbf{k}_+\leftrightarrow\mathbf{k}_-\,,\;\beta\leftrightarrow\gamma)+(-\mathbf{k}_+\leftrightarrow\mathbf{q},\;\beta\leftrightarrow\alpha)\,, \end{aligned} \tag{4.10}$$

where the two parentheses at the second line stand for two contributions obtained from the one at the first line by performing the replacements indicated. When the uniform limit for the displacement (that is, the inverse of (4.6)) is taken, the sum of the three new contributions gives the BS itself multiplied by the coefficient

$$-\,i\left(\mathbf{k}_+\cdot\mathbf{d}(\tau_\beta)-\mathbf{k}_-\cdot\mathbf{d}(\tau_\gamma)-\mathbf{q}\cdot\mathbf{d}(\tau_\alpha)\right)\,, \tag{4.11}$$

which vanishes for $\tau_\alpha = \tau_\beta = \tau_\gamma$, as a consequence of the EP and translational invariance. This holds indeed at every order in $\mathbf{d}$, as it was shown in [54].

In order to obtain the CR's, one has to identify the displacement $\tilde{\mathbf{d}}(\mathbf{q},\tau)$ with the large scale displacements induced by the velocity perturbations, that is, one has to give a *statistical* content to it. Assuming that at large scales linear PT holds, we will then identify, in redshift space,

$$\tilde{\mathbf{d}}(\mathbf{q},\tau) = \frac{1}{\mathcal{H}f}\left(\mathbf{v}(\mathbf{q},\tau) + f\mathbf{v}(\mathbf{q},\tau)\cdot\hat{\mathbf{z}}\,\hat{\mathbf{z}}\right) = -i\frac{\delta_m(\mathbf{q},\tau)}{q^2}\left(\mathbf{q} + f\,\mathbf{q}\cdot\hat{\mathbf{z}}\,\hat{\mathbf{z}}\right), \tag{4.12}$$

with $\delta_m(\mathbf{q},\tau)$ the real space matter density field, which is related to the velocity field by the continuity equation, the first of eq. (2.32). Implicitly, we have assumed that all the different species fall, at large scales, with the same velocity field, which follows from the assumption of adiabatic initial conditions and, again, the EP. Inserting (4.12) in the first term in (4.10) we get, in the $q/k \to 0$ limit,

$$\simeq \frac{k}{q}\left(\mu + f\mu_q\mu_k\right)\langle\delta^{(S)}_\alpha(\mathbf{q},\tau_\alpha)\delta_m(-\mathbf{q},\tau_\beta)\rangle'\langle\delta^{(S)}_\beta(-\mathbf{k}_-,\tau_\beta)\delta^{(S)}_\gamma(\mathbf{k}_-,\tau_\gamma)\rangle', \tag{4.13}$$

where

$$\mu \equiv \frac{\mathbf{k}\cdot\mathbf{q}}{kq}, \qquad \mu_q \equiv \frac{\mathbf{q}\cdot\hat{\mathbf{z}}}{q}, \qquad \mu_k \equiv \frac{\mathbf{k}\cdot\hat{\mathbf{z}}}{k}. \tag{4.14}$$

A contribution proportional to k/q is obtained also from the second term in eq. (4.10), while the third one vanishes. Consistently with our assumption that linear PT holds at the scale q, we use the Kaiser relation to express the real space matter field in terms of the redshift space one for the tracer α,

$$\delta_m(\mathbf{q},\tau) = \frac{1}{b_\alpha + f\mu_q^2}\,\delta^{(S)}_\alpha(\mathbf{q},\tau), \tag{4.15}$$

where both f and b_α are evaluated at the time τ. So, combining with eq. (4.12), we have

$$\tilde{\mathbf{d}}(\mathbf{q},\tau) = -\frac{i}{q^2}\frac{\mathbf{q} + f\,\mathbf{q}\cdot\hat{\mathbf{z}}\,\hat{\mathbf{z}}}{b_\alpha + f\mu_q^2}\,\delta^{(S)}_\alpha(\mathbf{q},\tau). \tag{4.16}$$

Finally, we get

$$\begin{aligned}
&\lim_{q/k\to 0} B^{(S)}_{\alpha\beta\gamma}(\mathbf{q},-\mathbf{k}_+,\mathbf{k}_-;\tau_\alpha,\tau_\beta,\tau_\gamma) \\
&= -\frac{k}{q}\frac{\mu + f\mu_k\mu_q}{b_\alpha + f\mu_q^2}P^{(S)}_{\alpha\alpha}(\mathbf{q};\tau_\alpha,\tau_\alpha)\left[\frac{D(\tau_\gamma)}{D(\tau_\alpha)}P^{(S)}_{\beta\gamma}(\mathbf{k}_+;\tau_\beta,\tau_\gamma) - \frac{D(\tau_\beta)}{D(\tau_\alpha)}P^{(S)}_{\beta\gamma}(\mathbf{k}_-;\tau_\beta,\tau_\gamma)\right] \\
&+ O\left(\left(\frac{q}{k}\right)^0\right),
\end{aligned} \tag{4.17}$$

where the PS's are defined as

$$P^{(S)}_{\alpha\beta}(\mathbf{k};\tau_\alpha,\tau_\beta) \equiv \langle \delta^{(S)}_\alpha(\mathbf{k};\tau_\alpha)\delta^{(S)}_\beta(-\mathbf{k};\tau_\beta)\rangle' \,, \tag{4.18}$$

$D(\tau)$ is the linear matter growth factor and we have assumed the linear behavior of the PS time dependence at the *soft* scale q, $P^{(S)}_{\alpha\alpha}(\mathbf{q};\tau_\alpha,\tau_\beta) = P^{(S)}_{\alpha\alpha}(\mathbf{q};\tau_\alpha,\tau_\alpha)D(\tau_\beta)/D(\tau_\alpha)$. On the other hand, as we have already emphasized, the dynamics at the *hard* scale, k is completely nonlinear. The key point is that the structure of the first term at the RHS is protected against any kind of, perturbative and non-perturbative, nonlinear effect. By contrast, the form of the remaining terms, indicated as $O((q/k)^0)$, is not protected and will be modified in a less and less controllable way at increasing k values and decreasing redshifts.

When the *hard* scale PS is evaluated at different times, $\tau_\beta \neq \tau_\gamma$, the BS in the squeezed limit goes as $P_L(q)/q$, and in real space ($f = 0$) the contribution to the BS is a dipole, as it is proportional to μ/b_α. Physically, this contribution can be interpreted as the effect of the different large scale displacements, $\tilde{\mathbf{d}}(\mathbf{q},\tau_\beta)$ and $\tilde{\mathbf{d}}(\mathbf{q},\tau_\gamma)$ experienced by the two short-scale fields $\delta^{(S)}_\beta(\mathbf{k},\tau_\beta)$, and $\delta^{(S)}_\gamma(\mathbf{k},\tau_\gamma)$ at the two different times τ_β and τ_γ. This effect grows with the coherence length of the displacement, which explains the $1/q$ behavior, and moreover it depends on the orientation between the large scale and the short scale modes, which explains the dipole behavior.

The coefficient in front of the PS's goes (again, in real space) as $-\mu^2/b_\alpha \times O((q/k)^0)$. The μ^2 behavior can be understood as follows. As we have already recalled, see eq. (4.11), a perfectly uniform displacement field cannot give any contribution to the equal times BS. Therefore, the effect can depend only on the gradient of the large scale displacement/velocity field. More precisely, the $i-th$ spatial component of the displacement field can affect the clustering on short scales along the $j-th$ direction only via its $\partial_j d^i(\mathbf{x})$ component, leading to a contribution to the configuration space three point function proportional to

$$\begin{aligned}
&\langle \delta_\alpha(-\mathbf{R})\delta_\alpha\left(\frac{\mathbf{r}}{2}\right)\delta_\alpha\left(-\frac{\mathbf{r}}{2}\right)\rangle \\
&\quad = \langle \bar{\delta}_\alpha(-\mathbf{R}+\mathbf{d}(-\mathbf{R}))\bar{\delta}_\alpha\left(\frac{\mathbf{r}}{2}+\mathbf{d}\left(\frac{\mathbf{r}}{2}\right)\right)\bar{\delta}_\alpha\left(-\frac{\mathbf{r}}{2}+\mathbf{d}\left(-\frac{\mathbf{r}}{2}\right)\right)\rangle \\
&\quad = r^j\frac{\partial\xi_\alpha(r)}{\partial r^i}\,\langle\delta_\alpha(-\mathbf{R})\,\frac{\partial d^i(\mathbf{r})}{\partial r^j}\bigg|_{\mathbf{r}=0}\rangle + \cdots \,, \\
&\quad = \frac{2}{3\mathcal{H}^2}\frac{r^i r^j}{r^2}\frac{\partial\xi_\alpha(r)}{\partial\ln r}\,\langle\delta_\alpha(-\mathbf{R})\,\frac{\partial^2\Phi(\mathbf{r})}{\partial r^i\partial r^j}\bigg|_{\mathbf{r}=0}\rangle + \cdots
\end{aligned} \tag{4.19}$$

where $\xi_\alpha(r)$ is the correlation function and $\Phi(\mathbf{r})$ the gravitational potential, which has been related to the displacement $\mathbf{d}$ by means of linear PT. In Fourier space (see eq. (4.12)), $\partial_j d^i(\mathbf{r})$ gives $-q^i q^j/q^2 \times \delta_m(\mathbf{q})$ which, contracted to $k^i \partial P_\alpha(k)/\partial k^j = (k^i k^j)/k^2 \, dP_\alpha(k)/d\ln k$, gives the $-\mu^2$ dependence.

The reason of the $1/b_\alpha$ factor comes about because we want to trade the velocity field (which is responsible for the CR protected term) with the directly observable density field for the α tracer.

4.2.2 CR and BAO's

The equal-times CR for a single tracer, α, in redshift space, given eq. (4.17), reads,

$$\begin{aligned}\lim_{q/k\to 0} \frac{B_\alpha(\mathbf{q}, -\mathbf{k}_+, \mathbf{k}_-)}{P_\alpha(\mathbf{q})P_\alpha(\mathbf{k})} &= \lim_{q/k\to 0} -\frac{k}{q}\frac{\mu + f\mu_k\mu_q}{b_\alpha + f\mu_q^2}\frac{P_\alpha(\mathbf{k}_+) - P_\alpha(\mathbf{k}_-)}{P_\alpha(\mathbf{k})} + O\left(\left(\frac{q}{k}\right)^0\right) \\ &= -\frac{\mu^2 + f\mu\mu_k\mu_q}{b_\alpha + f\mu_q^2}\frac{\partial \ln P_\alpha(\mathbf{k})}{\partial \ln k} - \frac{\mu_k(\mu + f\mu_k\mu_q)(\mu_q - \mu\mu_k)}{b_\alpha + f\mu_q^2}\frac{\partial \ln P_\alpha(\mathbf{k})}{\partial \ln \mu_k} + O\left(\left(\frac{q}{k}\right)^0\right) ,\end{aligned} \tag{4.20}$$

where $k \equiv |\mathbf{k}|$, $q \equiv |\mathbf{q}|$, $\mathbf{k}_\pm \equiv \mathbf{k} \pm \mathbf{q}/2$, $\mu \equiv \mathbf{k}\cdot\mathbf{q}/(k\,q)$, $\mu_k \equiv \mathbf{k}\cdot\hat{z}/k$, and $\mu_q \equiv \mathbf{q}\cdot\hat{z}/k$, with $\hat{z}$ being the direction of the line of sight. We have omitted the time dependence and used the fact that, in the far observer approximation, the redshift space PS, $P_\alpha(\mathbf{k})$, depends only on k and μ_k. In the following, we will consider dark matter ($\alpha = m$), halos ($\alpha = h$), and galaxy ($\alpha = g$) tracers.

We stress that the 'linear bias' b_α appearing in the CR is not a parameter of a bias expansion, but is defined precisely as the limit between the real space PS for the tracer α and the α-matter cross-correlator (see [2] and the section (4.2.1)),

$$b_\alpha \equiv \lim_{q\to 0} \frac{P_{\alpha\alpha}(q)}{P_{\alpha m}(q)} , \tag{4.21}$$

the only assumptions entering this definition being the EP and adiabatic initial conditions, who ensure that all species move with the same velocity fields at large scales, and that linear PT holds at scales $q \to 0$.

Unlike the non-equal times CR's [113], the CR-protected contributions on the RHS of the equal times CR, namely the first term on the first line and the first two terms at the second one,

cannot be distinguished from the unprotected ones, the $O((q/k)^0)$ terms, by looking at a pole in q as the squeezed limit is approached. This is, at first sight, unfortunate, as the unequal-times BS is, differently from the equal-times one, not measurable from data. However, as it was discussed in [2] and will be elaborated on in the following, BAO oscillations provide a way to single out the CR protected terms.

Among the two terms at the last line of eq. (4.20), there is a hierarchy. The amplitudes of the oscillations in the logarithmic derivative of the PS with respect to k are enhanced with respect to the ones of the derivative with respect to μ_k. This can be understood by looking at models for the redshift space PS, such as [121], which can be cast in the form

$$P_\alpha(\mathbf{k}) \simeq F_{fog}(k\mu_k f\sigma_v)(b_\alpha + f\mu_k^2)^2 \left(P^0(k) + \Delta P_\alpha^{1-\text{loop}}(k,\mu_k)\right) , \tag{4.22}$$

where $F_{fog}(x)$ is a phenomenological smooth function, usually a gaussian or a lorentzian, which parametrizes the Fingers-of-God effect, $P^0(k)$ is the linear PS, and $\Delta P_\alpha^{1-\text{loop}}(k,\mu_k)$ are contributions of 1-loop order. In presence of a feature in the PS, like the BAO, we can write

$$P^0(k) = P^0_{\text{nw}}(k)\left(1 + A(k)\sin(kr_s)\right) , \tag{4.23}$$

where $P^0_{\text{nw}}(x)$ is the smooth component of the linear PS, and $r_s = O(100)\,\text{h}\,\text{Mpc}^{-1}$ is the BAO scale, we see that the oscillating part of the logarithmic derivative of the PS with respect to $\ln k$ is of order

$$\frac{\partial \ln P_\alpha(\mathbf{k})}{\partial \ln k} \sim A(k)\, kr_s\, \cos(kr_s) + \text{smooth/higher orders contributions}\,, \tag{4.24}$$

while

$$\frac{\partial \ln P_\alpha(\mathbf{k})}{\partial \ln \mu_k} \sim \frac{\partial \Delta A(\mathbf{k})}{\partial \ln \mu_k}\sin(kr_s) + \text{smooth/higher orders contributions}\,, \tag{4.25}$$

where $\Delta A(\mathbf{k})$ is the μ_k-dependent part of the 1-loop contribution to the amplitude of the oscillating part of the PS. Therefore, comparing the two oscillating contributions, we see that besides being of 1-loop order as opposed to linear, the latter is suppressed by an extra factor of order $1/kr_s$,

$$\frac{1}{kr_s}\frac{\Delta A(\mathbf{k})}{A(k)} = \frac{k_s}{2\pi k}\frac{\Delta A(\mathbf{k})}{A(k)} , \tag{4.26}$$

where we have defined $k_s = 2\pi r_s^{-1} \simeq 0.06\,\text{h}\,\text{Mpc}^{-1}$, and can then be safely neglected in the BAO range of scales for squeezed configurations.

Coming now to the oscillating part of the $O\left(\left(\frac{q}{k}\right)^0\right)$ terms in eq. (4.20), a perturbative analysis shows that they are of order

$$\frac{\Delta P_\alpha(k)}{P^0_{\rm nw}(k)} A(k)\sin(kr_s)\,, \tag{4.27}$$

where $\Delta P_\alpha(k)$ is a one-loop order contribution to the PS. Therefore, compared to the leading oscillatory contribution, (4.24), this one is parametrically suppressed as the ones in (4.26), and therefore will be neglected too. In section (4.3.2) and (4.4.2), we will verify from simulations, in real space and in redshift space respectively, that the difference between the LHS and the first term at the RHS of (4.20) is indeed smooth in the squeezed limit.

Summarising, in our analysis we will consider only the first term of eq. (4.20), that is,

$$\begin{aligned}\lim_{q/k\to 0} B_\alpha(\mathbf{q},\mathbf{k}_-,-\mathbf{k}_+) = -\frac{\mu^2+f\mu\mu_k\mu_q}{b_\alpha+f\mu_q^2}P_\alpha(\mathbf{q})\frac{\partial P_\alpha(\mathbf{k})}{\partial \ln k} \\ + \text{smooth/higher orders contributions.}\end{aligned} \tag{4.28}$$

4.3 Real space

Here we describe the analysis presented in ref. [2]. We will study the equal-time CR in real space both, which, from eq. (4.28), reads

$$\lim_{q/k\to 0}\frac{B_\alpha(\mathbf{q},\mathbf{k}_-,-\mathbf{k}_+)}{P_\alpha(q)} = -\frac{\mu^2}{b_\alpha}\frac{d\log P_\alpha(k)}{d\log k} + O\left(\left(\frac{q}{k}\right)^0\right)\,. \tag{4.29}$$

4.3.1 Check in Perturbation Theory

As a first test of the CR, we compute the matter BS in real space in the squeezed limit at lowest order in SPT. It is given by, see eq. (2.53)

$$\lim_{q/k\to 0}\frac{B^{\rm SPT}_{mmm}(q,k_-,k_+)}{P_{m,L}(q)P_{m,L}(k)} = -\mu^2\frac{d\log P_{m,L}(k)}{d\log k} + \frac{13+8\mu^2}{7} + O\left(\frac{q}{k}\right)\,, \tag{4.30}$$

where $P_{m,L}\equiv P_{mm,L}$ denotes the linear matter PS. Notice that the second term at the RHS is scale independent, as it is proportional to $(P_{m,L}(k_+)+P_{m,L}(k_-))/P_{m,L}(k)\to 2$. On the other hand, the first term, although subdominant, can be isolated from the rest thanks to its scale dependence, which is induced mainly by BAO oscillations. To see this, one writes the PS as

$$P_{m,L}(k) = P^{nw}_{m,L}(k)(1+A(k)\sin(kr_s)), \tag{4.31}$$

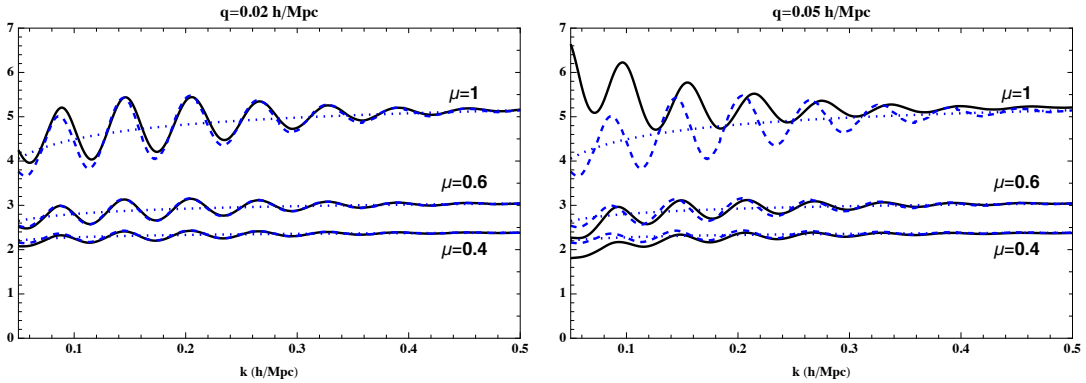

Figure 4.2: Comparison between the tree-level BS (divided by $P^0(q)P^0(k)$) (black-solid lines) and the RHS of eq. (4.30) (blue-dashed lines) as a function of k for two values for q. The blue-dotted lines are obtained by using the "no-wiggle" linear power spectrum, $P^{\rm nw}(k)$, in place of $P(k)$.

where r_s is the comoving sound horizon and both $P^{nw}_{m,L}(k)$ and $A(k)$ are featureless functions. Performing the logarithmic derivative, we get the first term at the RHS of (4.30) as

$$-\mu^2\Big[\left(\cos\left(kr_s\right)+\alpha(k)\sin\left(kr_s\right)\right)\,\frac{kr_sA(k)}{1+A(k)\sin\left(kr_s\right)}+\frac{d\log P^{nw}_m(k)}{d\log k}\Big]\,, \tag{4.32}$$

where we defined

$$\alpha(k)\equiv\frac{1}{kr_s}\frac{d\log A(k)}{d\log k}\,, \tag{4.33}$$

which takes values around $10\,\%$ in the k range of interest. So, the squeezed BS contains an oscillating component whose amplitude is enhanced by $kr_s\simeq 2\pi k/(0.05\,{\rm h\,Mpc^{-1}})$ with respect to eq. (4.31) and whose phase is shifted by $\sim\pi/2-\alpha(k)$, and a smooth component given by the second line of eq (4.32). The result is shown in fig. (4.2), for $q=0.02\,{\rm h\,Mpc^{-1}}$ and $q=0.05\,{\rm h\,Mpc^{-1}}$ and for three different values for μ. We see that the CR (eq. (4.30)) reproduces the BS as long as the squeezed limit ($q/k\ll 1$) holds. When q/k is not small, extra scale dependent terms come into play. However, the amplitudes of the oscillations in the RHS and in the LHS are still related by the CR, as we will show quantitatively in the next section.

4.3.2 Check in simulations

Once nonlinear effects are included, the CR (4.29) ensures that the first line in eq. (4.30) is modified just by changing the linear PS with the nonlinear one, while the second line is changed in

an uncontrolled way. However, due to parity invariance, the BS is symmetric under $k_+ \leftrightarrow k_-$, or, equivalently, $\mu \to -\mu$, which implies that the leading non-protected term in the squeezed limit should be proportional to $(P_m(k_+) + P_m(k_-))/P_m(k)$, that is, still featureless, although with an unknown coefficient.

To check this explicitly in fully nonlinear dynamics, we use 10 realizations of N-body simulations with $N = 2048^3$ mass elements performed in periodic cubes with the side length of $4\,h^{-1}\mathrm{Gpc}$. The error bars presented in this paper are obtained from the scattering among the 10 realizations.

The mass distribution evolved with the GADGET2 code [122] and the halo catalogs extracted by ROCKSTAR algorithm [123] at $z = 0$ will be presented in what follows (see the "fiducial cosmology" in [124] for other cosmological/numerical parameters). We measure the bispectrum using the quick FFT-based algorithm presented in [125] with the aliasing correction following [126]. We first store the data in bins of (k, q, μ) and check the results after summing up $B(q, k, \mu)/[P(q)P(k)]$ for different q bins up to some $q_{\rm max}$ weighting by the number of triangles.

We compare the oscillating part of the squeezed BS to that of the logarithmic derivative of the PS, and check if their amplitudes are related by the $-\mu^2/b_\alpha(q)$ factor of eq. (4.29). In order to extract the oscillating part, we will subtract smooth functions ("*poly*" in the figures) of the form

$$p(k, q_{\rm max}, \mu^2) = \sum_{i=-2}^{n} a_i(q_{\rm max}, \mu^2) k^i \,. \tag{4.34}$$

The negative powers in k, at each fixed $q_{\rm max}$, account for subleading terms in the squeezed limit, of order up to $q_{\rm max}^2/k^2$, while the positive powers account for the extra scale dependence induced for instance by the $d\log P_m^{nw}/d\log k$ contribution. The coefficients $a_i(q_{\rm max}, \mu^2)$ are fixed such that $p(k, q_{\rm max}, \mu^2)$ is the best fit to the chosen data (BS or logarithmic derivative); then the function is subtracted from the data to obtain only the oscillatory part, which is not captured by the fit for small enough n. We truncate our fit at $n = 2$, since with this value a satisfactory reduced χ^2 (see below) is obtained.

To quantify the goodness of the CR we measure the bias $b_\alpha(q_{max})$ in each μ bin, by mini-

mizing the χ^2 function

$$\chi^2 = \sum_{i=1}^{N} \frac{\left[\left(\frac{B}{PP}|_i - p_{B_i}\right) - \frac{1}{b_\alpha}\left(-\mu^2 \frac{d\ln P}{d\ln k}|_i - p_{P_i}\right)\right]^2}{\sigma_i^2}, \tag{4.35}$$

where all the q_{max}, k and μ dependencies are omitted. We denote with p_{B_i} and p_{P_i} the fitting curves relative to, respectively, the LHS and the RHS of the CR. σ_i are the errors for the BS in the $i-th$ k-bin, since the errors on the power spectra are much smaller. We use $N = 26$ linearly spaced bins from $k_{min} = 0.05$ h Mpc^{-1} to $k_{max} = 0.30$ h Mpc^{-1}. First, we performed

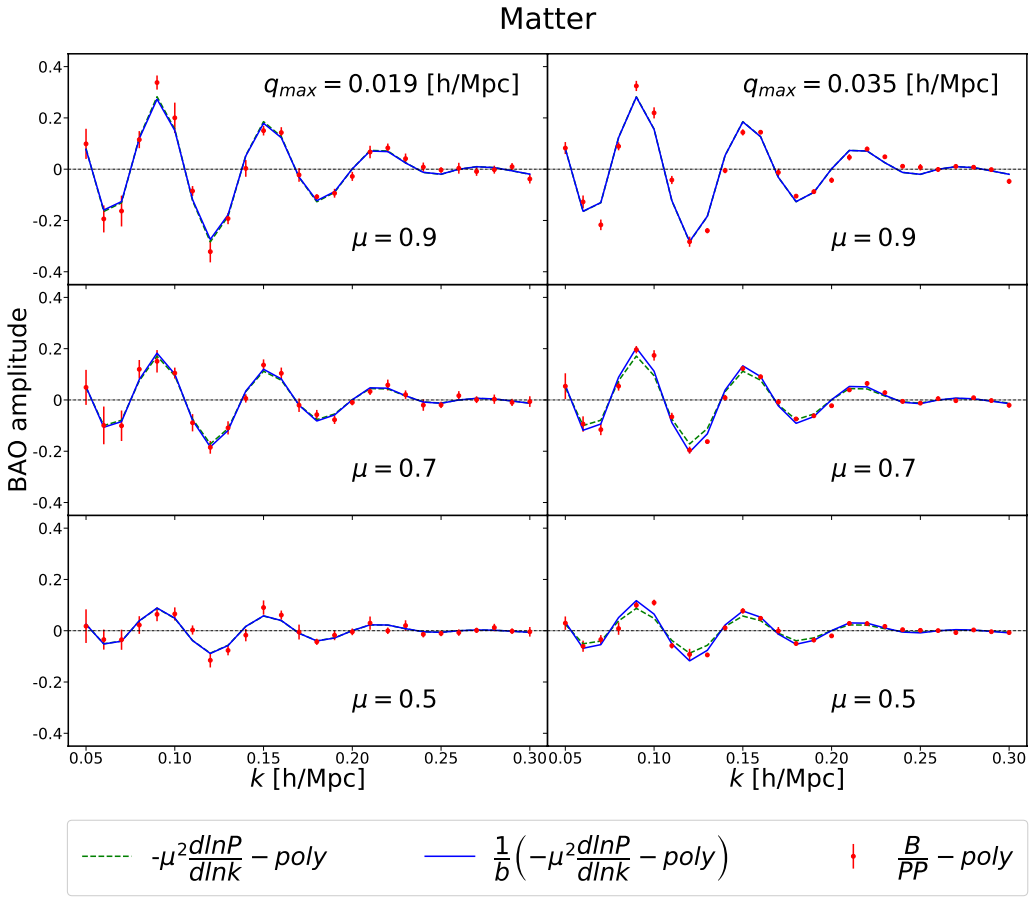

Figure 4.3: Comparison of the oscillating parts of the two sides of the CR, eq. (4.29), for different μ bins and different q_{max}, for matter at redshift $z = 0$. The red dots correspond to the BS, the blue lines are the oscillating part of the logarithmic derivative of the PS multiplied by $-\mu^2/b_m$, with b_m the best fit value of tab. (4.1). The green-dashed lines are computed assuming the expected value, $b_m = 1$.

the test on matter, for which the expected value is $b_m(q_{\rm max}) = 1$. In fig. (4.3) we show the

oscillating part of the BS (red points with error bars, estimated from the scatter among the 10 realizations) and of the logarithmic derivative of the PS multiplied by $-\mu^2/b_m$ (blue lines), for different μ-bins (rows) and for two different values for q_{max} (columns). In tab. (4.1) we give the best fit values for b_m in the different μ-bins and the relative $1-\sigma$ errors and reduced χ^2. As expected, going to higher q_{max} the quality of the fits gets worse, both because the squeezed limit is farther, and because the error bars are smaller since the number of BS configuration increases. The best fit values are always compatible with $b_m = 1$ for $q_{max} \lesssim 0.019$ h/Mpc.

$q = 0.019$ h/Mpc				$q = 0.035$ h/Mpc				$q = 0.061$ h/Mpc			
μ	b_m	σ_b	$\tilde{\chi}^2$	μ	b_m	σ_b	$\tilde{\chi}^2$	μ	b_m	σ_b	$\tilde{\chi}^2$
0.9	1.04	0.05	1.28	0.9	1.00	0.06	8.68	0.9	1.2	0.3	184
0.7	0.94	0.05	0.56	0.7	0.84	0.04	4.22	0.7	1.0	0.1	56
0.5	1.0	0.1	0.78	0.5	0.74	0.04	1.91	0.5	1.0	0.1	15
0.3	1.35	0.47	0.65	0.3	0.8	0.1	0.84	0.3	1.0	0.2	6

Table 4.1: Best fit values for b_m (expected value, $b_m = 1$) at $z = 0$ for different values of the maximum allowed q.

Having tested the CR on the matter field, we then used them to measure the bias of a given halo population, $b_h(q_{\rm max})$, via the $-\mu^2/b_h(q_{\rm max})$ modulation of the BAO oscillations, and compare the results with those obtained via the definition in eq. (4.21). We considered halos of masses $M > M_{min} = 10^{13} M_\odot$ at $z = 0$. The expected bias, as measured from eq. (4.21), is $b_h = 1.46 \pm 0.03$.

The results are shown in fig. (4.4), where, comparing the blue lines with the green-dashed ones, we see the effect of the halo bias in reducing the amplitude of the BAO oscillations with respect to the ones present in the logarithmic derivative of the PS. As shown in fig. (4.5) and Tab. (4.2), the extracted values for b_h are compatible with the expected one within the error bars, also for $q_{max} = 0.061$ h Mpc^{-1}, which gave bad fits in the DM case.

We have repeated the analysis for simulation data at redshift $z = 1$, finding qualitatively similar results (see fig. (4.6)). The only nonlinear effect limiting the effectiveness of our method is the damping of the BAO wiggles, which is stronger at lower redshifts.

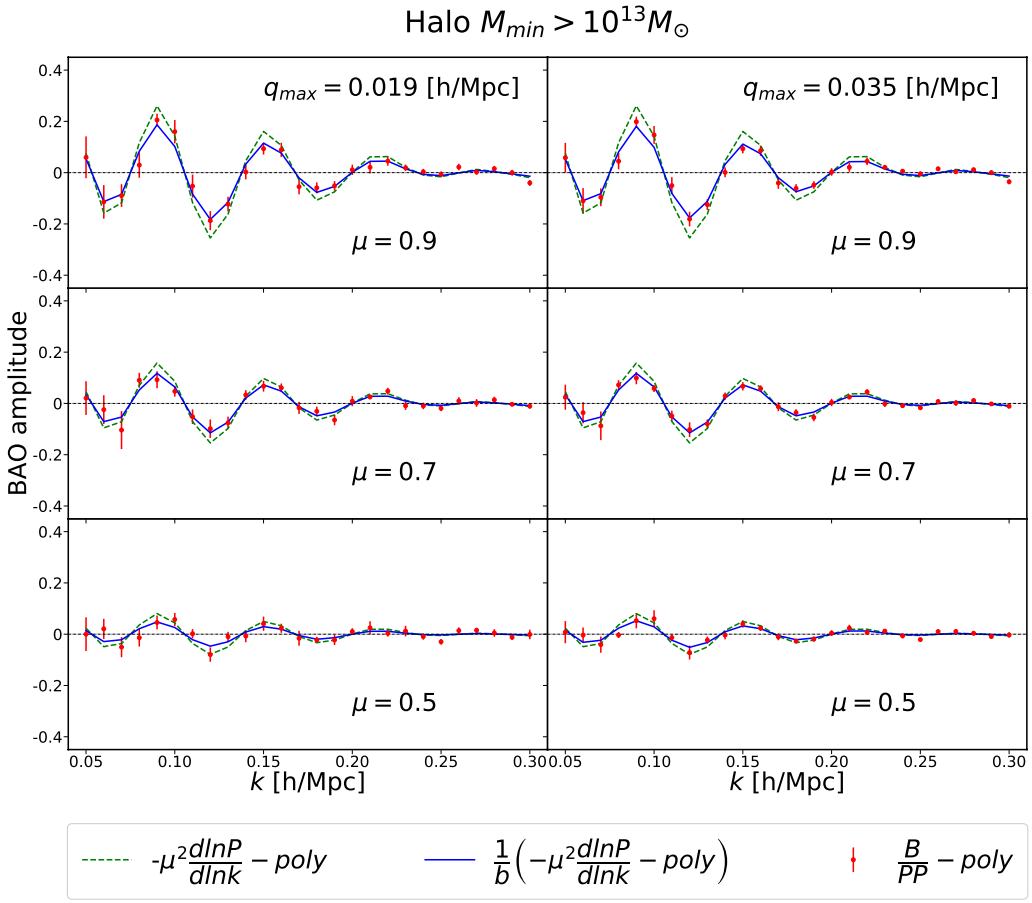

Figure 4.4: Same as fig. (4.3) for halos of mass $M > M_{min} = 10^{13} M_\odot$ at $z = 0$.

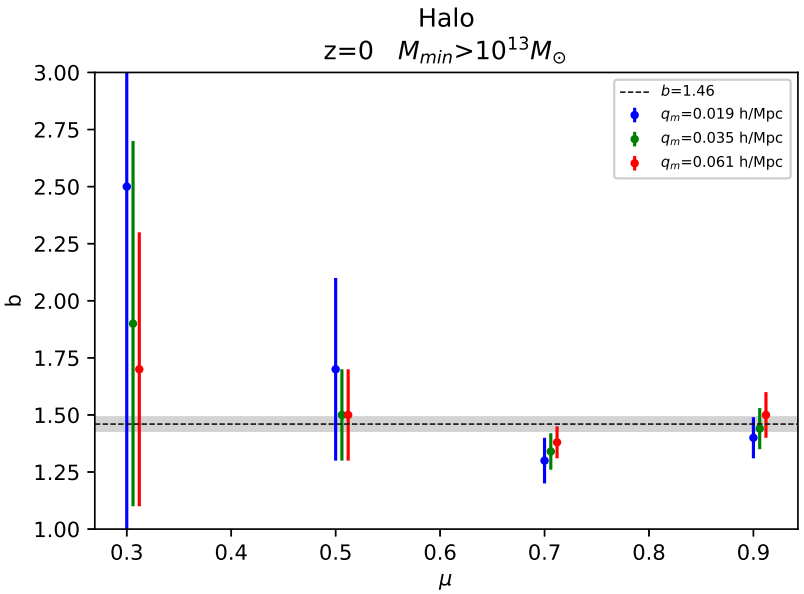

Figure 4.5: Values of the bias extracted for halos of mass $M > M_{min} = 10^{13}M_\odot$ at $z = 0$ for different μ-bins and for different values for $q_{\rm max}$. The shaded area represents the values extracted from the simulations using eq. (4.21) at 1-σ.

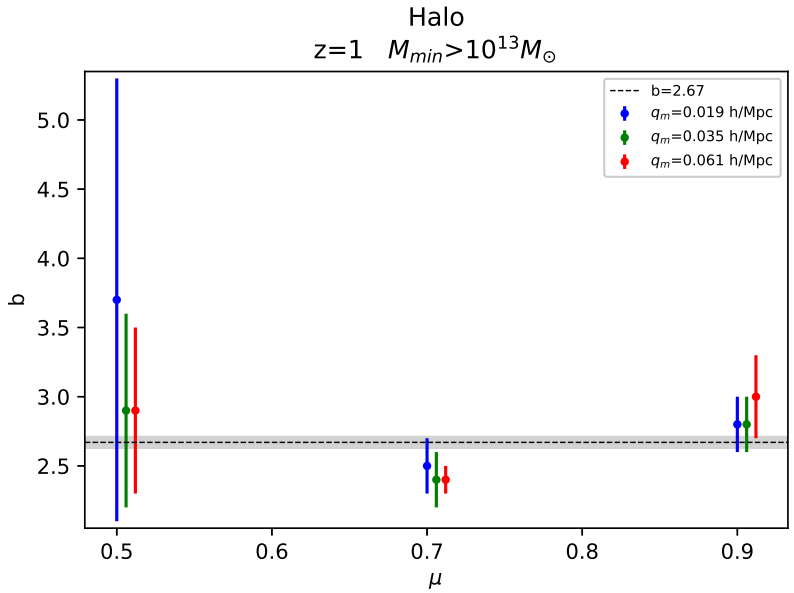

Figure 4.6: Same as fig. (4.5) at $z = 1$.

$q = 0.019$ h/Mpc				$q = 0.035$ h/Mpc				$q = 0.061$ h/Mpc			
μ	b_h	σ_b	$\tilde{\chi}^2$	μ	b_h	σ_b	$\tilde{\chi}^2$	μ	b_h	σ_b	$\tilde{\chi}^2$
0.9	1.40	0.09	0.8	0.9	1.44	0.09	1.1	0.9	1.5	0.1	1.6
0.7	1.3	0.1	0.6	0.7	1.34	0.2	0.6	0.7	1.38	0.07	0.6
0.5	1.7	0.4	1.1	0.5	1.5	0.2	1.0	0.5	1.5	0.2	1.1
0.3	2.5	1.9	0.6	0.3	1.9	0.8	0.6	0.3	1.7	0.6	0.7

Table 4.2: Bias values for halos with $M > M_{min} = 10^{13} M_\odot$ at $z = 0$. The expected value from eq. (4.21) is $b_h = 1.46 \pm 0.03$.

4.3.3 Discussion

Extracting bias from a PT-based approach differs substantially from the approach discussed here. The PT prediction (at tree level as well as at higher orders) requires assuming: 1) a cosmological model, namely, the linear matter power spectrum, its normalization σ_8 and the growth factor, and, 2) a bias model.

On the other hand, we are not assuming any model, neither for the cosmology nor for bias. Moreover, we are not relying on a computational scheme either, being it PT expansion or numerical simulations.

The crucial point is that what we call bias is not a model parameter, but a physical quantity. Thanks to the consistency relation, the cross correlator $P_{\alpha m}(q)$, and then $b_\alpha(q)$, can be measured by comparing two directly measurable quantities, namely, the amplitude of the BAO oscillations in the logarithmic derivative of the nonlinear power spectrum and that in the bispectrum, normalized as in eq. (4.29). The assumptions beyond our result are just the EP, adiabatic initial conditions, and the linear continuity equation at the very large scale q, which ensure that all tracers, and in particular matter and halos, fall with the same velocity field. This, from a theoretical point of view, singles out our proposal as being of a qualitatively different nature, rooting its robustness in being intrinsically nonperturbative and based on very general physical assumptions.

Moreover, in a PT approach, even at tree level, three bias parameters enter, namely, the linear bias b_1 and the second order bias parameters, b_2 and the "tidal bias" b_s (see e.g., [127]).

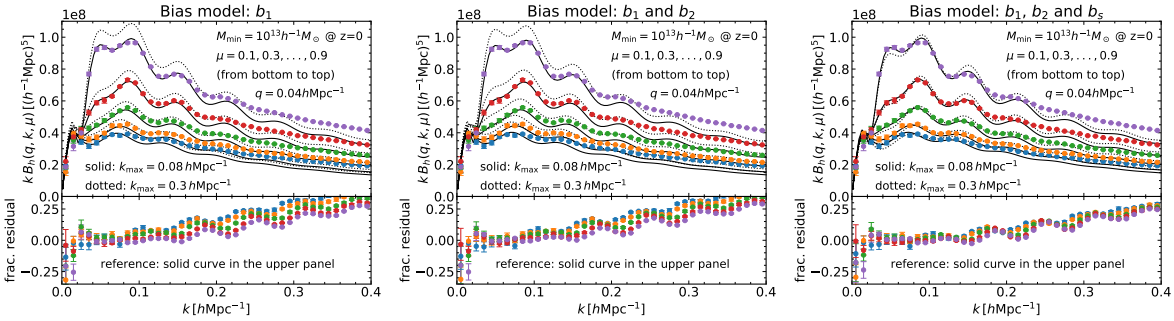

Figure 4.7: Fit to the full bispectrum in PT by using one, two, and three bias parameters

If we try to fit the measured bispectrum with this expression, these parameters are degenerate with each other, with the normalization of the power spectrum, and possibly with other cosmological parameters.

Together with the subtleties in the treatment of the scale-dependencies of the bias parameters, the end result of the fitting with the tree-level SPT template is clearly not as robust, from a theoretical point of view, as what we are proposing in this work.

To see this explicitly, we produced three plots, see fig. (4.7), in which we show the halo bispectrum at z=0 with different values of μ (in different colors) at a fixed q (=0.04 h Mpc^{-1}) as a function of k. The tree level calculation, but with different bias models (one for each figure, as indicated by the title) is fit to the data points, considering two different values of $k_{\rm max}$, below which the fitting was performed, $k_{\rm max}^{\rm low} = 0.08$ h Mpc^{-1} (solid lines), and $k_{\rm max}^{\rm high} = 0.3$ h Mpc^{-1} (dotted lines). We can see that, clearly, the tree-level PT does not provide a good fit over the BAO scale. We can see it quantitatively in fig. (4.8), which shows the reduced chi-squared for the three bias models as a function of $k_{\rm max}$. Indeed, it is close to unity only at very small $k_{\rm max}$.

For completeness, we summarize below the best fit bias parameters at the two values of $k_{\rm max}$ in Tab. (4.3). As we can see, the value of b_1, and, to a larger extent, those of b_2 and b_s, are affected by the choice of the bias model as well as the maximum wavenumber.

Moreover, we must keep in mind that these fits are performed by fixing the cosmological parameters of the PT computation to the true ones used in the simulation. In a real data analysis a scan over these parameters is needed, inducing an higher level of degeneracy.

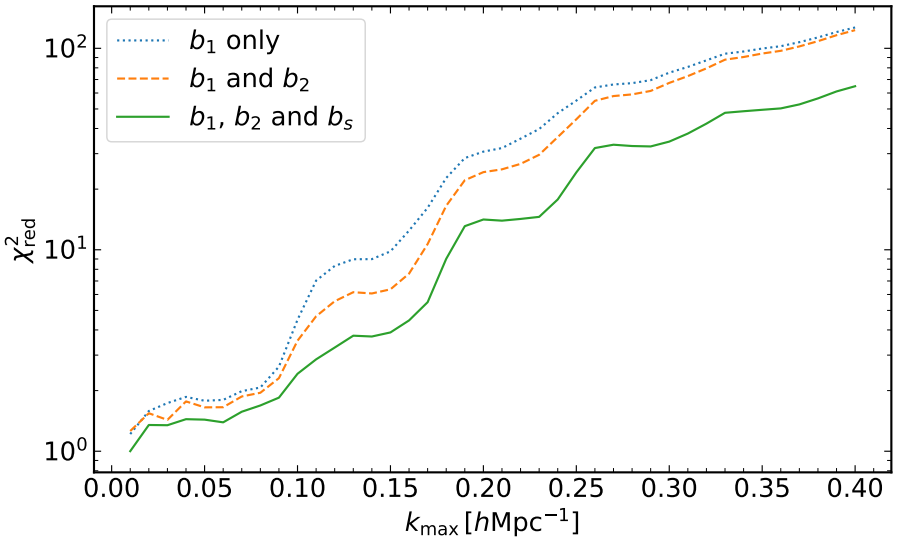

Figure 4.8: The reduced χ^2 from the three PT fits shown in fig. (4.7), as a function of $k_{\rm max}$.

bias model	$b_1(k^{\rm low}_{\rm max})$	$b_1(k^{\rm high}_{\rm max})$	$b_2(k^{\rm low}_{\rm max})$	$b_2(k^{\rm high}_{\rm max})$	$b_s(k^{\rm low}_{\rm max})$	$b_s(k^{\rm high}_{\rm max})$
b_1	1.347	1.389	-	-	-	-
b_1, b_2	1.336	1.357	0.0355	0.1332	-	-
b_1, b_2, b_s	1.393	1.497	-0.1021	-0.3468	-0.1371	-0.3447

Table 4.3: Bias values for halos with $M > M_{min} = 10^{13} M_\odot$ at $z = 0$, extracted from PT fits including one, two, and three bias parameters. Results for two different values for $k_{\rm max}$ are reported, with $k^{\rm low}_{\rm max} = 0.08$ h Mpc^{-1} and $k^{\rm high}_{\rm max} = 0.3$ h Mpc^{-1} . The expected value for b_h defined in eq. (4.21) is $b_h = 1.46 \pm 0.03$.

4.4 Redshift space

When we consider redshift space, the advantages of the method presented so far becomes even more clear. It is known that, besides for very large scales, redshift space distortions, in particular, Fingers of God effects are particularly different to model, requiring some measure of empirical parameterization. Our approach is free from this problem, as the redshift space version of the CR assumes the Kaiser relation only where it is reliable, namely, at the large scale q, while no modelling is required at the short scale k, see section (4.2.2) and eq. (4.28).

Moreover, in redshift space, the possibility of measuring, besides bias, the growth function

$f = d\log D/d\log a$ opens. Indeed, looking at eq. (4.28), one realizes that the μ^2/b_h prefactor of eq. (4.29) now becomes

$$\frac{\mu + \mu_k\,\mu_q\,f}{b_h + \mu_q^2 f}\,\mu\,, \tag{4.36}$$

where $\mu_k = \hat{\mathbf{k}}\cdot\hat{\mathbf{z}}$, $\mu_q = \hat{\mathbf{q}}\cdot\hat{\mathbf{z}}$ are the angles between the two modes and line of sight, here taken to be the z axis. Therefore, by considering different orientations of the triangles, or different multipoles, or again combining different tracers, f and b_h can be measured independently.

In this section we will perform this analysis for the redshift space version of the CR. In particular, we will investigate it using a multipoles decomposition, by integrating over the angular dependencies.

4.4.1 Multipoles

We will deal with the angular dependence of the BS by considering multipole expansions. In the following, we discuss the redshift space case, eq. (4.28), from which the real space results can be derived by taking $f \to 0$ and reinterpreting the PS's and the BS's as being the real space ones.

The BS in redshift space (in the distant observer approximation) depends on 5 coordinates: 3 of them (for instance q, k, and μ, see figure (4.1)) identify the triangular shape, while the remaining 2 are needed to define the orientation of the plane of the triangle with respect to the line of sight. Therefore, keeping q and k fixed, we are left with 3 angular coordinates, over which we will integrate with the measure

$$\int \frac{d^2\hat{k}}{4\pi}\int \frac{d^2\hat{q}}{4\pi}\,(2\pi)\delta_D\left(\phi_k + \phi_q\right) = \int_{-1}^{1}\frac{d\mu_k}{2}\int_{-1}^{1}\frac{d\mu_q}{2}\int_0^{2\pi}\frac{d\phi}{2\pi}\,, \tag{4.37}$$

where the delta-function in the first integral enforces rotation invariance around the z-axis, and we have defined $\phi = \phi_k - \phi_q$. The cosine μ is given in terms of the three independent variables as

$$\mu = \mu(\mu_k, \mu_q, \phi) = \sqrt{(1-\mu_k^2)(1-\mu_q^2)}\cos\phi + \mu_k\mu_q. \tag{4.38}$$

The PS's will be expanded in Legendre polynomials, $\mathcal{P}_l(\mu_k)$, as usual,

$$P_\alpha(\mathbf{k}) = \sum_{l=0}^{\infty} P_\alpha^{(l)}(k)\,\mathcal{P}_l(\mu_k)\,, \tag{4.39}$$

where

$$P_\alpha^{(l)}(k) \equiv \frac{2l+1}{2}\int_{-1}^{1} d\mu_k\, P_\alpha(\mathbf{k})\,\mathcal{P}_l(\mu_k)\,. \tag{4.40}$$

Concerning the PS at the large scale q, $P_\alpha(\mathbf{q})$, we will use the same expansion as above, with the additional assumption that linear PT holds at the scale q, which is consistent to what we have already assumed in deriving the CR. This implies that the Kaiser relation [73] can be used for the PS at this scale,

$$P_\alpha(\mathbf{q}) = (b_\alpha + f\mu_q^2)^2 P_L(q)\,, \tag{4.41}$$

leading to the well known expressions for the linear monopole and quadrupole,

$$P_\alpha^{(0)}(q) = b_\alpha^2\left(1+\frac{2}{3}\beta_\alpha+\frac{1}{5}\beta_\alpha^2\right)P_L(q)\,,\quad \frac{P_\alpha^{(2)}(q)}{P_\alpha^{(0)}(q)} = \frac{4\beta_\alpha}{21}\frac{7+3\beta_\alpha}{1+\frac{2}{3}\beta_\alpha+\frac{1}{5}\beta_\alpha^2}\,, \tag{4.42}$$

where $\beta_\alpha \equiv f/b_\alpha$.

Using the relations above, the RHS of (4.28) can be written as

$$-\frac{P_\alpha^{(0)}(q)}{\left(1+\frac{2}{3}\beta_\alpha+\frac{1}{5}\beta_\alpha^2\right)}\left(\frac{\mu^2}{b_\alpha}+\beta_\alpha\,\mu\mu_k\mu_q\right)\left(1+\beta_\alpha\,\mu_q^2\right)\frac{\partial P_\alpha(\mathbf{k})}{\partial\ln k}\,. \tag{4.43}$$

While the definition of the PS multipoles is unique, concerning the BS, different multipoles can be defined, as we can weight the angular integrations with Legendre polynomials in μ, μ_k, or μ_q. From eq. (4.38), we define BS multipoles with respect to μ as

$$B_\alpha^{(l)}(q,k) \equiv (2l+1)\int_{-1}^{1}\frac{d\mu_k}{2}\int_{-1}^{1}\frac{d\mu_q}{2}\int_0^{2\pi}\frac{d\phi}{2\pi}\,B_\alpha(\mathbf{q},\mathbf{k}_-,-\mathbf{k}_+)\,\mathcal{P}_l(\mu(\mu_k,\mu_q,\phi))\,. \tag{4.44}$$

The CR's for the monopole and the quadrupole in μ then read,

$$\begin{aligned}\lim_{q/k\to 0}\frac{B_\alpha^{(0)}(q,k)}{P_\alpha^{(0)}(q)P_\alpha^{(0)}(k)} =& -\left[\frac{1}{3b_\alpha}+\frac{b_\alpha-1}{9\,b_\alpha}\beta_\alpha\frac{1+\frac{3}{5}\beta_\alpha}{1+\frac{2}{3}\beta_\alpha+\frac{1}{5}\beta_\alpha^2}\right]\frac{d\ln P_\alpha^{(0)}(k)}{d\ln k}\\ &-\frac{2\beta_\alpha[2+b_\alpha(5+3\beta_\alpha)]}{225\,b_\alpha\left(1+\frac{2}{3}\beta_\alpha+\frac{1}{5}\beta_\alpha^2\right)}\frac{P_\alpha^{(2)}(k)}{P_\alpha^{(0)}(k)}\frac{d\ln P_\alpha^{(2)}(k)}{d\ln k}+\cdots\,,\end{aligned} \tag{4.45}$$

$$\begin{aligned}\lim_{q/k\to 0}\frac{B_\alpha^{(2)}(q,k)}{P_\alpha^{(0)}(q)P_\alpha^{(0)}(k)} =& -2\left[\frac{1}{3b_\alpha}+\frac{(b_\alpha-1)}{9\,b_\alpha}\beta_\alpha\frac{1+\frac{3}{5}\beta_\alpha}{1+\frac{2}{3}\beta_\alpha+\frac{1}{5}\beta_\alpha^2}\right]\frac{d\ln P_\alpha^{(0)}(k)}{d\ln k}\\ &-\frac{2\beta_\alpha[77+b_\alpha(98+75\beta_\alpha)]}{2205\,b_\alpha\left(1+\frac{2}{3}\beta_\alpha+\frac{1}{5}\beta_\alpha^2\right)}\frac{P_\alpha^{(2)}(k)}{P_\alpha^{(0)}(k)}\frac{d\ln P_\alpha^{(2)}(k)}{d\ln k}\\ &-\frac{8\beta_\alpha^2}{735\,\left(1+\frac{2}{3}\beta_\alpha+\frac{1}{5}\beta_\alpha^2\right)}\frac{P_\alpha^{(4)}(k)}{P_\alpha^{(0)}(k)}\frac{d\ln P_\alpha^{(4)}(k)}{d\ln k}+\cdots\,,\end{aligned} \tag{4.46}$$

where dots indicate smooth/subdominant contributions. Taking multipoles with respect to μ_k and μ_q, defined as

$$B_\alpha^{(l_{k,q})}(q,k) \equiv \frac{2l_{k,q}+1}{2}\int_{-1}^{1}\frac{d\mu_k}{2}\int_{-1}^{1}\frac{d\mu_q}{2}\int_0^{2\pi}\frac{d\phi}{2\pi}B_\alpha(\mathbf{q},\mathbf{k}_-.-\mathbf{k}_+)\mathcal{P}_{l_{k,q}}(\mu_{k,q}), \tag{4.47}$$

we get the same monopole equation as above, while, for the quadrupoles, we get,

$$\begin{aligned}\lim_{q/k\to 0}\frac{B_\alpha^{(l_k=2)}(q,k)}{P_\alpha^{(0)}(q)P_\alpha^{(0)}(k)} =& -\frac{2\beta_\alpha}{45\,b_\alpha}\frac{2+b_\alpha(5+3\beta_\alpha)}{1+\frac{2}{3}\beta_\alpha+\frac{1}{5}\beta_\alpha^2}\frac{d\ln P_\alpha^{(0)}(k)}{d\ln k}\\ &-\frac{105+43\beta_\alpha+55b_\alpha\beta_\alpha+33b_\alpha\beta_\alpha^2}{315\,b_\alpha\left(1+\frac{2}{3}\beta_\alpha+\frac{1}{5}\beta_\alpha^2\right)}\frac{P_\alpha^{(2)}(k)}{P_\alpha^{(0)}(k)}\frac{d\ln P_\alpha^{(2)}(k)}{d\ln k}\\ &-\frac{4\beta_\alpha}{315b_\alpha}\frac{2+b_\alpha(5+3\beta_\alpha)}{1+\frac{2}{3}\beta_\alpha+\frac{1}{5}\beta_\alpha^2}\frac{P_\alpha^{(4)}(k)}{P_\alpha^{(0)}(k)}\frac{d\ln P_\alpha^{(4)}(k)}{d\ln k}+\cdots,\end{aligned} \tag{4.48}$$

and

$$\begin{aligned}\lim_{q/k\to 0}\frac{B_\alpha^{(l_q=2)}(q,k)}{P_\alpha^{(0)}(q)P_\alpha^{(0)}(k)} =& -\frac{2\beta_\alpha}{63\,b_\alpha}\frac{7+b_\alpha(7+6\beta_\alpha)}{1+\frac{2}{3}\beta_\alpha+\frac{1}{5}\beta_\alpha^2}\frac{d\ln P_\alpha^{(0)}(k)}{d\ln k}\\ &-\frac{42+22\beta_\alpha+28b_\alpha\beta_\alpha+24b_\alpha\beta_\alpha^2}{315\,b_\alpha\left(1+\frac{2}{3}\beta_\alpha+\frac{1}{5}\beta_\alpha^2\right)}\frac{P_\alpha^{(2)}(k)}{P_\alpha^{(0)}(k)}\frac{d\ln P_\alpha^{(2)}(k)}{d\ln k}+\cdots.\end{aligned} \tag{4.49}$$

Notice that, unlike the quadrupole in μ, those in μ_k and μ_q are proportional to β_α, and therefore are non-vanishing only in redshift space.

4.4.2 Results in simulations

We analyse the same set of simulations already presented in ref. [2] and presented in section (4.3.2).

We measure the PS and the BS using fast Fourier transform. We first assign the particle mass or the halo number density on to 1024^3 grid points using Cloud-in-Cells (CIC) algorithm [128] in configuration space. After transforming to the Fourier space, we mitigate the aliasing effect [129] using the interlacing technique [126] and then divide the field by the CIC window function. We store the products of the resulting fields into bins to form the estimator of either the PS and the BS. In case of the PS, we prepare bins with the interval of $0.005\,\mathrm{h\,Mpc^{-1}}$. This is sufficient to resolve the BAO feature in detail. The product, $|\delta_\mathbf{k}|^2$, is averaged in the bins to obtain our estimator of the PS. In redshift space, we also consider $\mathcal{P}_\ell(\mu_\mathbf{k})|\delta_\mathbf{k}|^2$ to estimate

the multipole moments. In case of the halo PS, we subtract the standard Poissonian shot noise contribution, $V/N_{\rm h}$, where V is the simulation volume and $N_{\rm h}$ is the number of halos, from the monopole moment.

The estimator of the BS can be constructed in an analogous manner. We refined the binning scheme from that adopted in ref. [2] to better capture its configuration dependence. We consider a pair of wavevectors $(\mathbf{k}, \mathbf{q})$ and form a triangle $(\mathbf{q}, -\mathbf{k}_+, \mathbf{k}_-)$. We bin the triangles in q and k at every $0.01\,{\rm h\,Mpc^{-1}}$, and then we sum the ratio of the bispectrum to the linear PS in q up to a given $q_{\rm max}$, weighting the sum with the number of triangles in each q-bin,

$$\sum_{q\leq q_{\rm max}} N_{\rm tri}(q,k_i)\frac{B_\alpha^{(l)}(q,k_i)}{P_\alpha^{(0)}(q)} \equiv \frac{B_\alpha^{(l)}}{P_\alpha^{(0)}}(q_{\rm max},k_i)\,. \tag{4.50}$$

The remaining degree of freedom, the angle between the two wavevectors (and also the relative angle with respect to the line-of-sight direction in case of redshift space), is integrated to obtain the moment estimators: in ref. [2], we instead kept the angle dependence and estimated the BS in bins of q, k and μ. Since we know the expected angle dependence of the oscillatory feature, that is simply μ^2, we can fully express it with the first two even moments, monopole ($\ell = 0$) and the quadrupole ($\ell = 2$). This helps to obtain the BAO feature with smaller error bars. We subtracted the shot noise, $(V/N_{\rm h})^2 + (V/N_{\rm h})[P_{\rm h}(q) + P_{\rm h}(k_+) + P_{\rm h}(k_-)]$ from the halo monopole BS.

We are now ready to describe the procedure we have used to evaluate the bias parameter b_α and the parameter $\beta_\alpha = f/b_\alpha$ from the simulations, by using the CR's.

In order to do that, we have to fit, in k, the LHS's of the CR, binned up to a given $q_{\rm max}$, see eq. (4.50),

$$\frac{1}{P_\alpha^{(0)}(k_i)}\frac{B_\alpha^{(l)}}{P_\alpha^{(0)}}(q_{\rm max},k_i)\,, \tag{4.51}$$

with the RHS's, which we will model as

$$\begin{aligned} &-\sum_{l'=0,2} C^{(l,l')}(b_\alpha,\beta_\alpha)\left(\frac{d\ln P_\alpha^{(l')}(k_i)}{d\ln k} - \left.\frac{d\ln P_\alpha^{(l')}(k_i)}{d\ln k}\right|_{\rm smooth}\right) e^{-c(q_{max})k^2} \\ &\quad + p(\{a_i^{(l)}(q_{\rm max})\};k_i)\,, \end{aligned} \tag{4.52}$$

where the coefficients $C^{(l,l')}(b_\alpha,\beta_\alpha)$ can be read from eqs. (4.45) and (4.46), while the smooth functions $p(\{a_i^{(l)}(q)\};k)$ are going to fit the smooth contributions from the derivatives of the

PS multipoles together with the other smooth and or subdominant contributions discussed in sect. (4.2.2). The contribution of the PS hexadecapole ($l = 4$) to eqs. (4.46), is numerically negligible, and we do not include it in our analysis.

We have isolated the smooth contributions from the derivatives of the PS monopole and quadrupole by subtracting a spline fit. We tested alternative algorithms to extract the smooth contributions obtaining stable results for the extracted parameters. Moreover, we introduced a scale dependent BAO damping term, $e^{-c(q_{max})k^2}$, which models possible correlations between the long mode and the FoG damping beyond the squeezed limit.

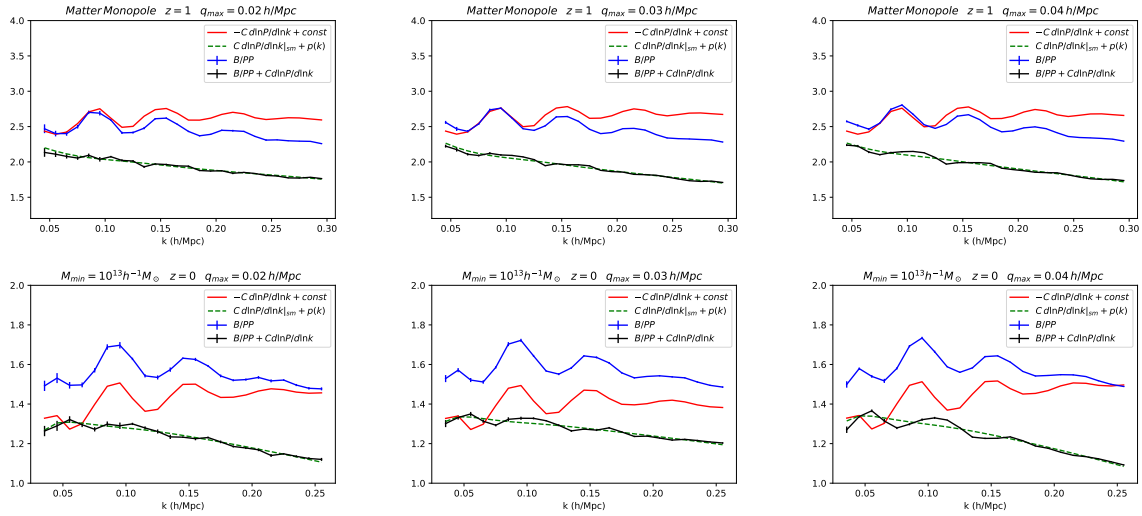

Figure 4.9: The BS to PS's ratios of eq. (4.51) (blue lines with error bars), and the oscillating part of the terms containing the logarithmic derivatives of the PS's in eq. (4.52) (red lines). The difference between the two is given by the black lines, together with the smooth fitting functions described in the text (green-dashed lines). The red lines include a constant offset for graphical purposes. The fiducial values for b_α, β_α and the best fitting values for the nuisance parameters have been used to produce these plots.

The form of the fitting functions $p(\{a_i^{(l)}(q)\}; k)$ is chosen in order to reproduce the leading expected contributions. It contains a term constant in k, as the lowest order PT result in the squeezed limit. Then, we include a negative contribution proportional to k^2, accounting for the leading contribution to the logarithmic derivative of the nonlinear PS from the Fingers of God effect. This can be understood by looking at the pre-factor in eq. (4.22). Finally, we

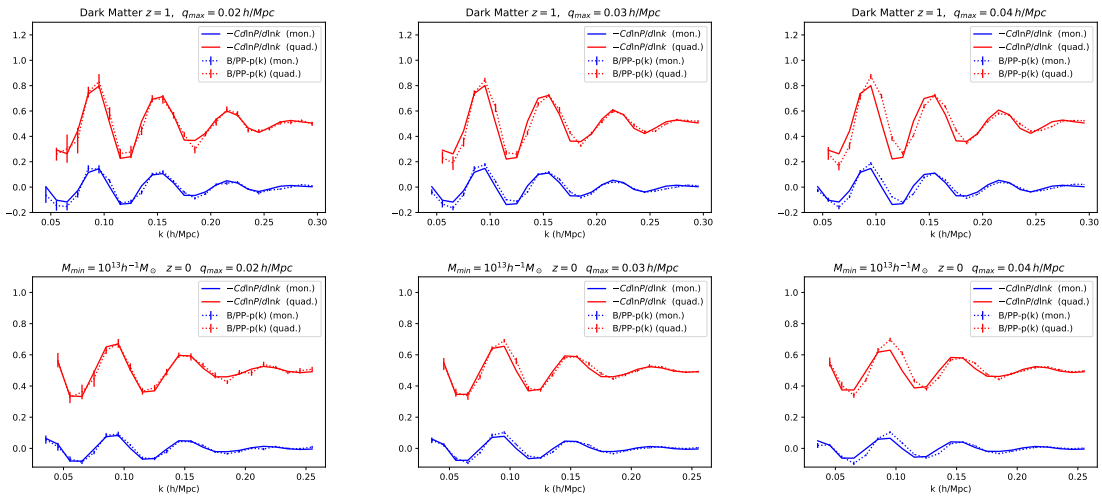

Figure 4.10: Comparison between the oscillating parts of the BS to PS ratios in eq. (4.51) (dotted lines) and that of eq. (4.52) (solid lines) both for the monopole (blue) and the quadrupole (red) components of the BS. The fiducial values for b_α, β_α and the best fitting values for the nuisance parameters have been used to produce these plots.

include also a k^{-2} term to control possible (small) deviations from the squeezed limit, which we expect to scale as $(q_{\rm max}/k)^2$.

Summarizing, the smooth function we will use in the CR's for the monopole and the quadrupole of the BS takes the form

$$p(\{a_i^{(l)}(q_{\rm max})\};k) = a_{-2}^{(l)}(q_{\rm max})\left(\frac{k}{q_{\rm max}}\right)^{-2} + a_0^{(l)}(q_{\rm max}) + a_2^{(l)}(q_{\rm max})\left(\frac{k}{\bar{k}}\right)^2, \tag{4.53}$$

where we have fixed the pivot scale $\bar{k} = 0.06\,{\rm h\,Mpc^{-1}}$. We have also considered extended polynomial fitting formulas, obtaining consistent results for the parameter estimations.

The smoothness of the difference between the LHS and the first term at the RHS of the CR can be verified from simulations, as we show in fig. (4.9), where we plot the ratio (4.51) of the BS to the PS's (blue lines), the sum of the logarithmic derivative of the PS multiplied by the appropriate coefficients, as in (4.52), without the subtraction of the smooth part (red lines), the difference between the two curves (black lines) and the smooth interpolation used to fit the latter, given by the sum of the terms containing the spline fits to the logarithmic derivatives of the PS's and the polynomial $p(\{a_i^{(l)}(q_{\rm max})\};k_i$ (green-dashed lines). The fiducial values of b_h and f have been used to evaluate the coefficients in these curves. The error bars in these plots

are dominated by those of the BS. As we can see, the difference between the BS to PS ratios in (4.51) and the terms in the logarithmic derivatives of the PS is smooth, with residual oscillations increasing as one moves away from the squeezed limit, by increasing $q_{\rm max}$. On the other hand, by increasing $q_{\rm max}$ the statistical errors are reduced, as more triangle configurations contribute to the BS, so a compromise has to be found between statistical power and the goodness of the squeezed limit approximation.

In fig. (4.10) we show the oscillating components of the BS to PS ratios, both for monopoles (dotted blue) and quadrupoles (dotted red) compared to the oscillating parts of eq. (4.52).

We introduce the following Log-likelihood function,

$$\chi^2_{CR}(b_\alpha, \beta_\alpha, \{a_i^{(l)}(q)\}, q_{\rm max}) \equiv \sum_i \left(\frac{r^2_{(0)}(k_i)}{\sigma^2_{(0),i}} + \frac{r^2_{(2)}(k_i)}{\sigma^2_{(2),i}} \right) , \tag{4.54}$$

where the $r_{(l)}(k_i)$'s are the differences between eq. (4.52) and eq. (4.51), and $\sigma_{(l),i}$'s are the corresponding errors on the BS measured from the simulations, evaluated in the i'th k-bin. We neglect the error on the PS as it is much smaller than that on the BS, and we assume diagonal covariances.

The CR's in eqs. (4.45) and (4.46) and (4.48), depend both on b_α and β_α, so, in principle, one can break the degeneracy between these two parameters by using the CR's alone. This is indeed the case, as we show in fig. (4.11). The BS monopole, as the $\mu-$ quadrupole (from eqs. (4.45) and (4.46)) are mostly sensitive to b_α but insensitive to β_α. On the other hand, when we combine the monopole with the μ_k or μ_q quadrupoles of eqs. (4.48) or (4.49) we can constrain also β_α, although only at the $\sim 40\,\%$ level. We do not combine different bispectrum quadrupoles, as they are not independent, and their cross-covariance would be non-trivial.

More effective constraints on β_α can be obtained by combining the CR's with measurements of the ratio between the PS quadrupole and monopole, in the linear regime, see the second of eqs. (4.42). This relation is valid in the Kaiser approximation [73], whose validity is limited to small $k's$. Therefore we will fit this ratio only up to $k^K_{max} \lesssim 0.02 - 0.03\ {\rm h\,Mpc^{-1}}$, where the ratio exhibits the plateau shown in fig. (4.12). Notice that in deriving the CR's we have assumed the validity of linear theory, and therefore of the Kaiser approximation, up to $q_{\rm max}$, so this procedure will be consistent as long as $k^K_{max} \lesssim q_{\rm max}$. Therefore, we will add to

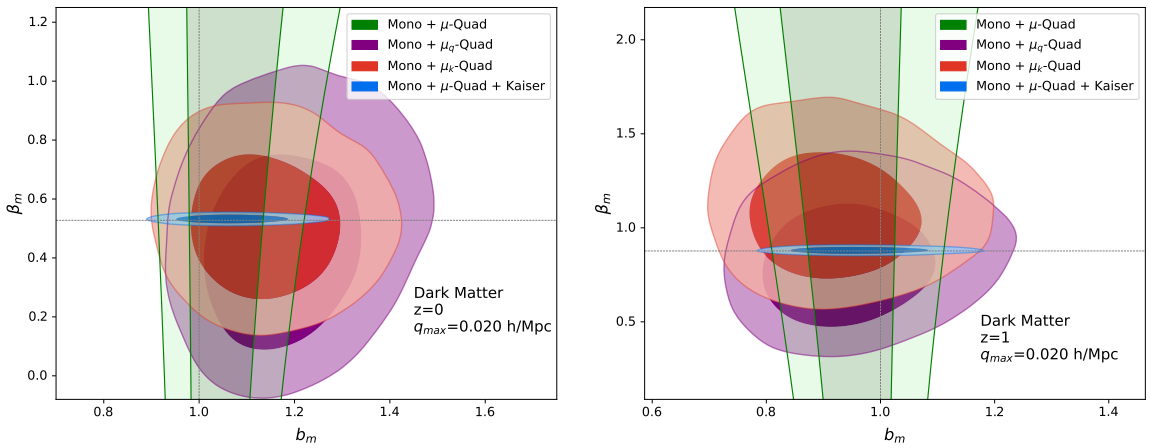

Figure 4.11: 1 and 2-σ constraints in the $b_m - \beta_m$ plane from CR's on the matter BS at redshift $z = 0$, (left) and $z = 1$ (right). The green contours are obtained combining the monopole and the μ-quadrupole of the BS (eq. (4.46)). Combining the monopole with the BS quadrupole in μ_k (eq. (4.48)), or the one in μ_q (eq. (4.49)), constrains also β_m as shown by the red and purple contours, respectively. Adding information on the PS quadrupole to monopole ratio, eq. (4.56), gives the blue contours. The dotted lines indicate the fiducial values for b_m and β_m.

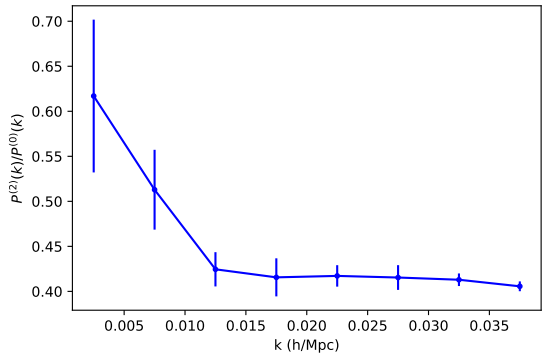

Figure 4.12: PS quadrupole to monopole ratio for halos of $M_{\rm min} = 10^{13} h^{-1}\, M_\odot$ at $z = 0$.

(4.54) the function

$$\chi^2_K(\beta_\alpha) = \sum_j \frac{r^2_K(\beta_\alpha; k_j)}{\sigma^2_{K,j}}\,, \tag{4.55}$$

where

$$r_K(\beta_\alpha; k_j) = \frac{P^{(2)}_\alpha(k_j)}{P^{(0)}_\alpha(k_j)} - \frac{4\beta_\alpha}{21}\frac{7+3\beta_\alpha}{1+\frac{2}{3}\beta_\alpha+\frac{1}{5}\beta^2_\alpha}\,, \tag{4.56}$$

and $\sigma_{K,j}$ is the error on the ratio between the PS quadrupole and monopole. We will show combined constraints obtained by minimizing the sum

$$\chi^2_{TOT} = \chi^2_{CR}(b_\alpha, \beta_\alpha, \{a^{(l)}_i(q)\}, q_{\rm max}) + \chi^2_K(\beta_\alpha)\,. \tag{4.57}$$

Summarizing, in our analysis we have 9 parameters, the physical ones $\{b_\alpha, \beta_\alpha\}$ and the fitting ones $\{a^{(l)}_{-2}, a^{(l)}_0, a^{(l)}_2, c\}$, with $l = 1, 2$, over which we will marginalize.

We first check our procedure for matter, for which we expect to extract values compatible with the fiducial ones, $b^{fid}_m = 1$ and $\beta^{fid}_m = f^{fid}$. We sample the log-likelihood function (4.57) using the MCMC Python library `emcee`[3][130]. The results of the analysis for matter are shown in Table (4.4) and in fig. (4.13), the plots are obtained using the plot library of Getdist[4][131].

As the constraining power of CR comes from the BAO's, we choose k values in the range in which they are present in the bispectra. For dark matter we take $k_{min} = 0.045\,{\rm h\,Mpc^{-1}}$ and $k_{max} = 0.30\,{\rm h\,Mpc^{-1}}$. Higher values of k_{max} do not improve our determinations of b_α and β_α. In fig. (4.13) we show the $68\,\%$ and $95\,\%$ confidence level regions in the $b_m - \beta_m$ plane for two different values of $q_{\rm max}$ ($= 0.02,\ 0.03\,{\rm h\,Mpc^{-1}}$), and, with dotted lines, the fiducial values. As we see, increasing $q_{\rm max}$ improves the constraints, due to the higher number of triangular configurations included in the BS measurement. Both values of $q_{\rm max}$ give unbiased values for the parameters. This is not the case by taking $q_{\rm max} = 0.04\,{\rm h\,Mpc^{-1}}$, which shows that this value is too far from the squeezed limit, as could have been anticipated also by looking at Figs. (4.9) and (4.10). Therefore, in our analysis on halos we will consider only $q_{\rm max} = 0.02,\ 0.03\,{\rm h\,Mpc^{-1}}$.

After having validated the procedure for dark matter we proceed in the analysis for halos with different masses at different redshifts. The results are presented in Table (4.5) and

[3] https://emcee.readthedocs.io/en/stable/

[4] https://getdist.readthedocs.io/en/latest/intro.html

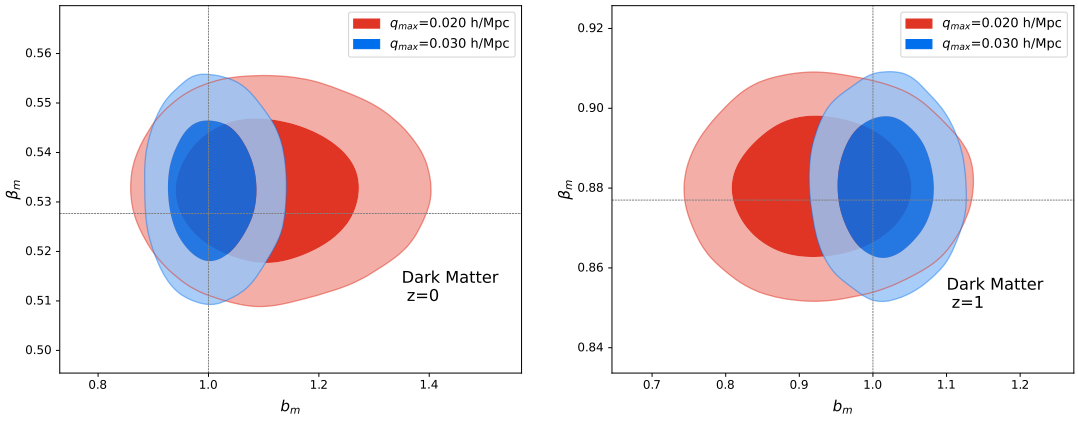

Figure 4.13: 1- and 2-σ constraints in the $b_m - \beta_m$ plane for dark matter at redshift $z = 0$ (left) and $z = 1$ (right), and for different values of q_{max}. The dotted lines are the expected values for b_m and β_m.

fig. (4.14) and are evaluated using $k_{min}(z = 0) = 0.035\,\mathrm{h\,Mpc^{-1}}$, $k_{max}(z = 0) = 0.26\,\mathrm{h\,Mpc^{-1}}$ and $k_{min}(z = 1) = 0.035\,\mathrm{h\,Mpc^{-1}}$, $k_{max}(z = 1) = 0.28\,\mathrm{h\,Mpc^{-1}}$. The results are compatible with the theoretical fiducial values at the $1 - \sigma$ level, with the expected values measured from the simulations using eq. (4.21) for b_h^{fid}, and eq. (4.42) for β_h^{fid}.

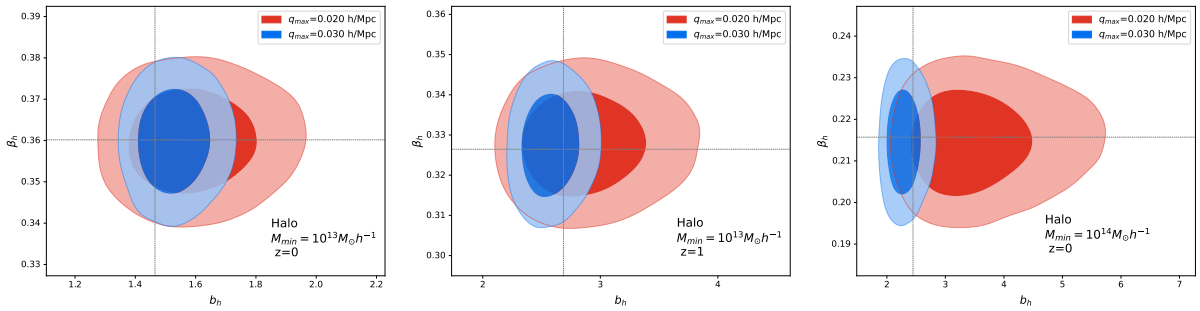

Figure 4.14: 1- and 2-σ constraints on b_h and β_h for halos of $M_{\rm min} = 10^{13} h^{-1}\,M_\odot$ at $z = 0$ (right) and $z = 1$ (middle) $M_{\rm min} = 10^{14} h^{-1}\,M_\odot$ at $z = 0$ (right) for two different values of q_{max} ($= 0.02,\ 0.03\,\mathrm{h\,Mpc^{-1}}$). The dotted lines are the expected values, obtained from direct measurements.

Our analysis shows that it is possible to break the degeneracy between the linear bias and the growth rate (or the β-parameter) with a good accuracy. In Table (4.4) and Table (4.5) we report the 68 % CL measurements we obtained for $q_{max} = 0.02, 0.03\,\mathrm{h\,Mpc^{-1}}$. We can see

$z=0$				
q_{max} (h/Mpc)	b_m	b_m^{fid}	$f=\beta_m$	f^{fid}
0.020	$1.11^{+0.12}_{-0.10}$	1	$0.532^{+0.010}_{-0.009}$	0.528
0.030	$1.01^{+0.05}_{-0.05}$	1	$0.533^{+0.009}_{-0.009}$	0.528
$z=1$				
q_{max} (h/Mpc)	b_m	b_m^{fid}	$f=\beta_m$	f^{fid}
0.020	$0.99^{+0.09}_{-0.08}$	1	$0.880^{+0.012}_{-0.012}$	0.877
0.030	$1.02^{+0.04}_{-0.04}$	1	$0.880^{+0.012}_{-0.012}$	0.877

Table 4.4: Best fit values for b_m and f at different redshifts for different values of the maximum allowed q. For each parameter, the quoted errors correspond to the 68% CL. of the one-dimensional probability distribution function.

that the results of the analysis for biased tracers in redshift space are fully consistent with the fiducial ΛCDM value for the growth function f.

Notice that, both for matter and halos, CR constrain mainly the bias parameter b_α, while the parameter β_α is constrained mostly by the independent measurement of the PS quadrupole to monopole ratio. As the latter is measured at better than $5\,\%$ accuracy for our simulations, the error on the derived growth function f is dominated by that on the bias parameter.

4.4.3 Estimating constraining power

We here present a forecast of the expected constraining power using the CR on the oscillatory part of the spectra alone, that is, not in conjunction with the PS quadrupole to monopole ratio. Since we have seen that the impact of the redshift space distortion is rather weak, a separate constraint on the tracer bias b_α and the growth rate parameter f from the consistency relation alone would be difficult. We thus focus on the constraint on b_α ignoring the redshift space distortions, that is, setting $f = 0$ in the CR's. As discussed earlier, we can then combine with the constraint on β_α from the redshift space distortion on the PS, to disentangle the degeneracy between the two parameters. Another simplification that we have made for the

$M_{\rm min} = 10^{13} h^{-1}\, M_\odot$ $\quad z = 0$				
q_{max} (h/Mpc)	b_h	b_h^{fid}	$f = \beta_h b_h$	f^{fid}
0.020	$1.58^{+0.15}_{-0.13}$	1.47	$0.57^{+0.06}_{-0.06}$	0.528
0.030	$1.53^{+0.08}_{-0.08}$	1.47	$0.55^{+0.04}_{-0.04}$	0.538
$M_{\rm min} = 10^{13} h^{-1}\, M_\odot$ $\quad z = 1$				
q_{max} (h/Mpc)	b_h	b_h^{fid}	$f = \beta_h b_h$	f^{fid}
0.020	$2.85^{+0.39}_{-0.32}$	2.686	$0.93^{+0.14}_{-0.14}$	0.877
0.030	$2.58^{+0.17}_{-0.16}$	2.686	$0.82^{+0.07}_{-0.07}$	0.877
$M_{\rm min} = 10^{14} h^{-1}\, M_\odot$ $\quad z = 0$				
q_{max} (h/Mpc)	b_h	b_h^{fid}	$f = \beta_h b_h$	f^{fid}
0.020	$3.40^{+0.83}_{-0.61}$	2.446	$0.73^{+0.18}_{-0.18}$	0.528
0.030	$2.29^{+0.21}_{-0.18}$	2.446	$0.49^{+0.06}_{-0.06}$	0.528

Table 4.5: Determination of b_h and β_h from the CR's for halos, for different values of $q_{\rm max}$. For each parameter, the quoted errors correspond to the 68% CL. of the one-dimensional probability distribution function.

forecast is to ignore nonlinear damping of BAOs. Including this effect would weaken the constraint especially from high wavenumbers, and thus the results presented here would give us the best-case scenario, but the purpose here is to give a rough idea on the statistical power brought by the consistency relations and the simplified treatment here must be fine.

We start with the construction of the BAO template based on the linear matter PS. As discussed in sect. (4.4.2), we take the logarithmic derivative, $d\ln P^0(k)/d\ln k$, and then subtract a B-spline fit to extract the oscillatory part. We use this as the template model after multiplying by $(1/3b_{\rm g})P_{\rm g}(q)P_{\rm g}(k)$ for the monopole and by $(2/3b_{\rm g})P_{\rm g}(q)P_{\rm g}(k)$ for the quadrupole of the galaxy BS, where $P_{\rm g}(k) = b_{\rm g}^2 P^0(k)$ is the linear galaxy PS with the bias parameter $b_{\rm g}$. We then estimate the covariance matrix of the galaxy BS, which is diagonal under the Gaussian

Table 4.6: Survey parameters considered in the forecast.

redshift	$V\,[h^{-3}\mathrm{Gpc}^3]$	$(n_\mathrm{g}/10^{-4})\,[h^3\mathrm{Mpc}^{-3}]$
$0.65 < z < 1.25$	22.64	15.86
$1.25 < z < 1.65$	20.66	8.86
$1.65 < z < 2.05$	22.69	2.61

assumption [132]:

$$[\Delta B_\mathrm{g}(k_1,k_2,k_3)]^2 = \frac{V}{N_\mathrm{tri}}\left[P_\mathrm{g}(k_1)+n_\mathrm{g}^{-1}\right]\left[P_\mathrm{g}(k_2)+n_\mathrm{g}^{-1}\right]\left[P_\mathrm{g}(k_3)+n_\mathrm{g}^{-1}\right], \tag{4.58}$$

$$\tag{4.59}$$

where N_tri is the number of Fourier triangles in a bin, which scales as V^2, and n_g is the galaxy number density specified later assuming a future survey setting. Since we specify the triangles by (q,k) and average over the angular dependence in our case, N_tri after taking this average can be expressed as

$$N_\mathrm{tri} = \left(\frac{2\pi}{3}\right)^2 \frac{(q^3_\mathrm{bin,max}-q^3_\mathrm{bin,min})(k^3_\mathrm{bin,max}-k^3_\mathrm{bin,min})}{k_\mathrm{f}^6}, \tag{4.60}$$

where $q_\mathrm{bin,min}$ and $q_\mathrm{bin,max}$ specify the minimum and the maximum wavenumber of the q bin and similarly for the k bin. We adopt the bin spacing of $0.005\,\mathrm{h\,Mpc}^{-1}$ for this forecast, and have confirmed that the results are virtually unchanged when we adopt a finer binning. In the above, we have excluded the contribution from redundant triangles (e.g., a triangle with negative q_z is equivalent to another with positive q_z) due to the reality condition, $\delta_{-\mathbf{q}} = \delta^*_{\mathbf{q}}$, and denote the fundamental wavenumber by $k_\mathrm{f} = 2\pi/V^{1/3}$. The error on the monopole moment of the BS is estimated using eq. (4.59) assuming that $P^0(k_1) = P^0(q)$ and $P^0(k_2) = P^0(k_3) = P^0(k)$ approximately hold over the triangles in a bin, and that of the quadrupole is obtained by further multiplying a factor 5 to eq. (4.59) to account for the weighting by the Legendre polynomial and our normalization convention. We ignore the error on the PS, which should be much smaller than that in the BS. We consider a Euclid-like survey and take the survey parameters from Table 3 in ref. [119]. Instead of considering the tomographic analysis with the 14 thin redshift bins over $0.65 < z < 2.05$ listed in that table, we consider three thick

redshift bins with similar volume as summarized in Table (4.6). We consider the survey area of $15,000\,\mathrm{deg}^2$ and take the values in "reference" case for the galaxy number density, averaged over the relevant fine redshift bins weighted by the volume. We propagate the error on the monopole and the quadrupole moment of the bispectra to the only parameter of the model template, b_g, to give the estimate of the statistical power of the consistency relation. We fix $b_\mathrm{g} = 1.5$ as the fiducial value for all the three bins. Changing this would give us a slight change in the relative contribution of the shot noise, but the final forecast is almost unchanged when we modify this to e.g., $b_\mathrm{g} = 1.6$.

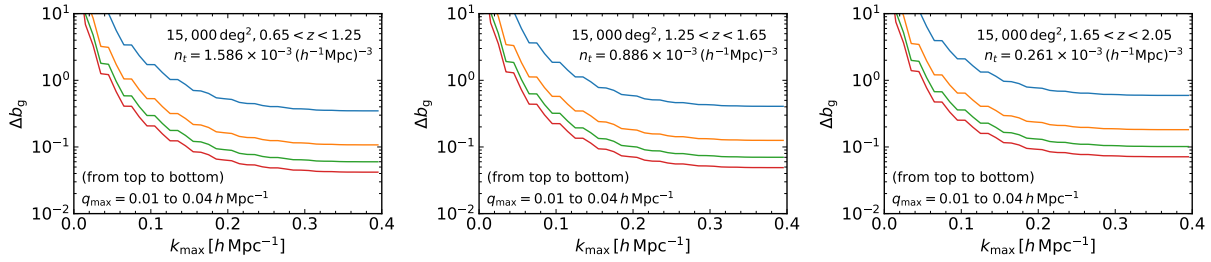

Figure 4.15: Forecast of the constraint on the galaxy bias parameter b_g from a Euclid-like survey in three tomographic redshift bins coming from the oscillatory part of the consistency relation alone. The results are shown as a function of the maximum hard wavenumber k_max for some values of the corresponding soft wavenumber limit, q_max.

We show in fig. (4.15) the expected 1-σ error on the bias parameter as a function of the maximum wavenumber included in the analysis. While the limit of the hard wavenumber, k_max, is indicated by the x-axis, we consider four values of q_max, the counterpart for the soft wavenumber, $0.01, 0.02, 0.03$ and $0.04\,\mathrm{h\,Mpc}^{-1}$. The smaller q_max is, we are restricting to more squeezed triangles and the resultant constraint is weaker. As we already see explicitly in previous sections, we can push to q_max to $0.03\,\mathrm{h\,Mpc}^{-1}$ quite safely without introducing a sizable bias in the consistency relation. While $q_\mathrm{max} = 0.04\,\mathrm{h\,Mpc}^{-1}$ might be slightly optimistic the improvement from $q_\mathrm{max} = 0.03\,\mathrm{h\,Mpc}^{-1}$ is smaller compared to that from $q_\mathrm{max} = 0.02\,\mathrm{h\,Mpc}^{-1}$ to $q_\mathrm{max} = 0.03\,\mathrm{h\,Mpc}^{-1}$.

It is clear from the figure that we can achieve a better than ten percent determination of the bias parameter for all the redshift bins, with the highest redshift bin slightly worse due to the larger shot noise error. Since the nonlinear damping of BAO is not very significant for scales

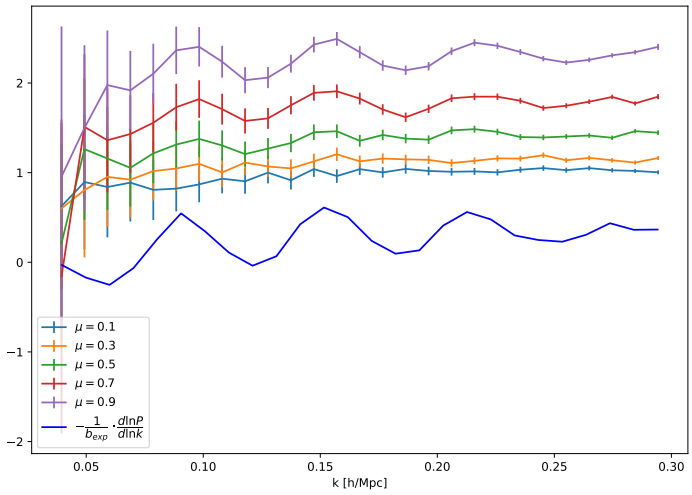

Figure 4.16: Squeezed limit of the BS in the BAO range for different values of the cosine μ, see fig. 4.1. The BAO feature is clearly visible and its amplitude get smaller for smaller μ values, as we expect from eq. (4.29). For comparison we show also the logarithmic derivative of the PS, divided by the expected bias value.

$k < 0.1\,\mathrm{h\,Mpc^{-1}}$ and the most of the constraining power is coming from the $k < 0.2\,\mathrm{h\,Mpc^{-1}}$, above which the shot noise error gets prominent, our estimate should be a good approximation even when the nonlinear effects are considered.

4.5 Future applications

After having presented a Fisher forecast on Euclid-like observations, we present here the application of the method presented so far to realistic simulations.

We are currently testing the accuracy of the CR-based method on estimating bias when applied to Euclid-like N-body simulations. These are simulations for the dark matter halos in real space with the specifics (volume, redshift range, number density of the sources) of the future Euclid satellite [1]. First, we provided a rigorous pipeline for the extraction of the squeezed limit of the BS. At this stage, we mapped the (k_1, k_2, k_3) triangular bins to (k, q, μ) and obtained the squeezed BS using the procedure described in sec. (4.4.2). In fig. (4.16) we show the

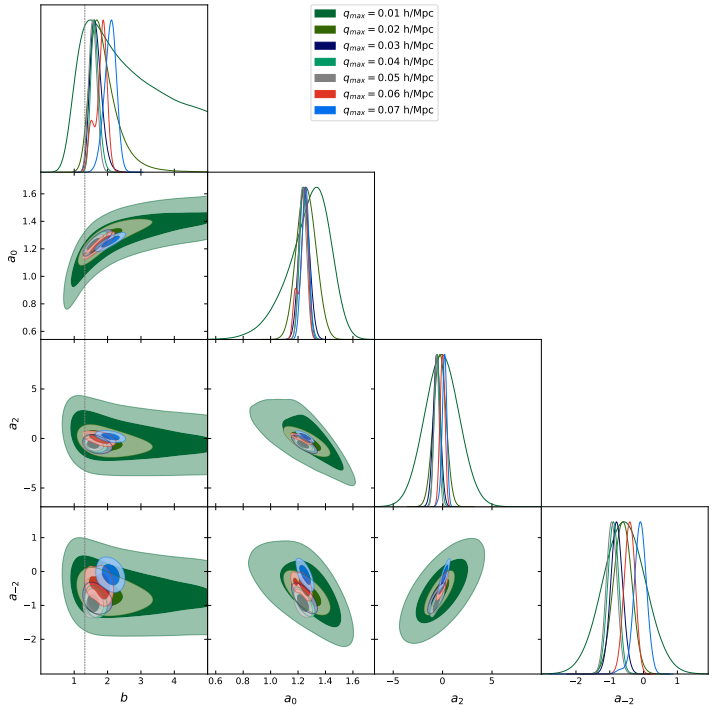

Figure 4.17: $q_{\rm max}$ dependence of the fitting parameters.

result of this procedure for different values of μ, verifying eq. (4.29). Here we will consider only the gaussian errors for the BS, since the PS errors are much smaller. The logarithmic derivative showed in fig. (4.16) is divide by the expected linear bias, $b_{\rm exp}$, measured using the SPT model for the BS of biased tracers, see [61]. We have then calculated the monopole of eq. (4.29) and for the analysis we will use the relation

$$\frac{B_t^{(0)}(k,q_{\rm max};z)}{P_t^{(0)}(q_{\rm max};z)P_t^{(0)}(k;z)} = -\frac{1}{3b_t(q_{max};z)}\frac{d\log P_t(k;z)}{d\log k} + a_0(q_{max};z) + a_2(q_{max};z)k^2 + \frac{a_{-2}(q_{max};z)}{k^2}\,. \tag{4.61}$$

Notice that in general the measured bias and coefficients of the smooth functions a_i's depend on q_{max} and on the redshift z. While the latter is a physical motivated dependence, the former is only due to our choice for $q_{\rm max}$. If the procedure we adopt is correct all these values will be independent on the $q_{\rm max}$ choice. On the other hand, we are considering the limit in which $q/k \to 0$, and considering higher values of $q_{\rm max}$ will make our CR much more unreliable. In figure (4.17) we show the MCMC analysis for the PS and BS monopoles at redshift $z = 0.90$.

Figure (4.17) shows that up to $q_{\rm max} \simeq 0.05$ h/Mpc the CR-based approach gives unbiased

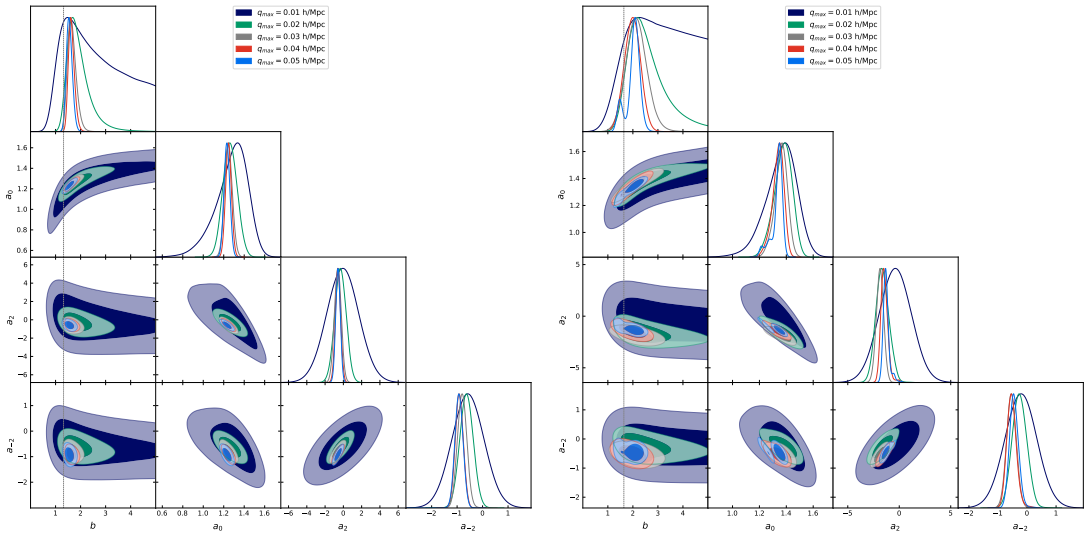

Figure 4.18: Results of the analysis for $z = 0.90, 1.19$.

and (almost) $q_{\max}$-independent values for b_t and a_i's. This results shows at what $q_{\max}$ our results are reliable and accurate, since higher values correspond to higher triangle numbers and a higher signal-to-noise ratio. The results of the analysis for the redshifts $z = 0.9$ and $z = 1.19$ are presented in fig. (4.18). These preliminary results show a model independent measurement of b_t with errors $\sim 8\%$ for $z = 0.90$ and $\sim 14\%$ for $z = 1.19$, using $k_{\max} = 0.26$ h/Mpc.

As mentioned above, a fundamental step would be the application of the CR-based method to actual observations. Despite the observation of the BAO features in the bispectrum of BOSS, [63], the width of the binning in the dataset[5] is too high to resolve the BAO in the squeezed limit of the BS. Here we have shown that Euclid observations will enable us to use the CR's to measure in a model-independent way the linear bias. The results of this analysis will be published soon within the Euclid collaboration.

4.6 Conclusions

In this chapter, we have investigated the CR's as a way to measure the large scale bias and the large scale growth rate in a model independent way. We have derived the relevant CR's in

[5]The PS monopole and quadtrupole and the BS monopole can be found here.

redshift space for the BS monopole and quadrupole and verified their validity on a set of large volume N-body simulations, both for DM and for halos of different mass, at different redshifts. As a first step we have verified the CR's for DM and halos in real space, and we were able to measure the large scale bias with a $\sim 10\%$ accuracy in a model independent way. This was one of the first application of the CR's, see ref. [2, 115].

Subsequently, we have performed an analogous analysis for the redshift space correlators. While the coefficients of the CR's depend on b_α and $\beta_\alpha = f/b_\alpha$ separately, the constraining power on β_α from CR's alone turns out to be very mild. However, when the CR measurements are combined with those on the PS quadrupole to monopole ratio, the $b_\alpha - \beta_\alpha$ degeneracy is completely broken.

When applied to a Euclid-like survey this approach would provide constraints on these parameters at better than 10% level. It is likely that this result can be further improved by modelling the leading contributions not protected by the CR's, and we think it will be very interesting to explore quantitatively this issue. In any case, while a ten percent error would not sound to be particularly good in modern cosmology, our constraints come completely free from model assumptions given as a bonus by just checking certain configurations of the BS. This, when combined with the redshift space distortion, which cannot break the degeneracy between the bias and the growth-rate parameter, would provide a unique way to constrain the gravitational growth. Moreover this procedure will be applied within the Euclid collaboration as a novel method for extracting the large scale galaxy bias. We are currently testing this method against the Euclid flagship simulations, finding results that are in agreement with our predictions presented in sect. (4.4.3) and that are compatible with the standard PT model, see for example [61, 77].

Considering different redshift bins, the extracted values for $f(z)$ would help constraining ΛCDM and modified scenarios as well. Having multiple tracers available would provide a way of testing the universality of the large scale growth rate, constraining possible velocity bias and violations of the EP. Moreover, a violation of the CR's can happen in presence of primordial non-gaussianities, see [115]. CR's could also be used in the context of those models that introduce primordial features in the PS, like [133, 134], in order to give better constraints on possible inflationary models. We leave the exploration of these applications to future work.

Chapter 5

The Large Scale Structure Bootstrap

In chapter (2) we have introduced SPT, the standard approach currently used for the calculation of the one-loop PS. This approach is based on the perturbative solution to the equations of motion of dark matter, eq. (2.32), subsequently used in the perturbative expansion of the galaxy density.

In this chapter we will describe a novel approach recently proposed in [8] which does not relies on the equations of motion but only on symmetry principles. The kernels of the perturbative solutions for the dark matter and the galaxies (or, in general, generic tracers) are built by using only terms that respect the symmetries considered. Then, symmetries are used to constrain the analytical form of the kernels, reducing the number of parameters needed. In the matter sector, we find that only few time dependent functions inside the kernels are not constrained by symmetries, but are defined by the cosmological model considered.

When galaxy density is considered, instead, we find that, at third perturbative order, the number of bias parameters is defined by the symmetries and does not depend on the particular cosmological model. At this order the number of bias parameters is seven, the same found in sect. (2.3.1), and we will show that this number is fixed in all the cosmologies that share the same symmetries. We will present a brief comparison between the ΛCDM and the nDGP (presented in section 3.2) models. We will also describe the UV effects within this approach. We will conclude with some possible future applications of the approach presented below.

The results of this chapter have been published in the paper

- G. D'Amico, M. Marinucci, M. Pietroni and F. Vernizzi, *The Large Scale Structure Bootstrap: perturbation theory and bias expansion from symmetryes*, *JCAP* **10** (2021) 069, doi 10.1088/1475-7516/2021/10/069[1],

and we have included this articles or part of it according to the journal (JCAP) copyright policies.

5.1 Constraints on PT kernels for general tracers

In this section we will see how to generalize the results obtained in PT (see Sec. 2.1) to a general cosmological model with only time-dependent growth factor.

We will make use of the time variable introduced in chapter 2, the logarithm of the scale factor defined in eq. 2.27 as $\eta \equiv \log a/a_0$. Beside the matter density contrast $\delta(\mathbf{x},\eta) \equiv \rho(\mathbf{x},\eta)/\bar{\rho}(\eta) - 1$, where ρ is the energy density and $\bar{\rho}$ its background value, we will also use the velocity divergence defined in eq. 2.30,

$$\theta(\mathbf{x},\eta) \equiv -\frac{\nabla \mathbf{u}(\mathbf{x},\eta)}{f_+(\eta)\mathcal{H}(\eta)} \,. \tag{5.1}$$

In the following, we will also consider the number density contrast of any tracer field $\delta_t(\mathbf{x},\eta)$, which can be galaxies, halos, etc. Using perturbation theory we can expand these quantities as

$$\delta(\mathbf{x},\eta) = \sum_{n=1}^{\infty} \delta^{(n)}(\mathbf{x},\eta), \qquad \theta(\mathbf{x},\eta) = \sum_{n=1}^{\infty} \theta^{(n)}(\mathbf{x},\eta), \qquad \delta_t(\mathbf{x},\eta) = \sum_{n=1}^{\infty} \delta_t^{(n)}(\mathbf{x},\eta) \,. \tag{5.2}$$

For the linear field it is convenient to introduce the *linear* doublet $\phi^\lambda_{\mathbf{k}}(\eta)$ ($\lambda = 1,2$) that includes the linear density contrast and the linear velocity divergence,

$$\phi^1(\mathbf{x},\eta) \equiv \delta^{(1)}(\mathbf{x},\eta) \,, \qquad \phi^2(\mathbf{x},\eta) \equiv \theta^{(1)}(\mathbf{x},\eta) \,. \tag{5.3}$$

After decomposing it in Fourier space,

$$\phi^\lambda(\mathbf{x},\eta) = \int \frac{d^3k}{(2\pi)^3} e^{-i\mathbf{k}\cdot\mathbf{x}}\, \phi^\lambda_{\mathbf{k}}(\eta) \,, \tag{5.4}$$

[1] © IOP Publishing Ltd and Sissa Medialab. Reproduced by permission of IOP Publishing. All rights reserved.

for each mode $\mathbf{k}$ the linear solution reads

$$\phi^{\lambda}_{\mathbf{k}}(\eta) = u^{\lambda}_{f}(\eta)\varphi_{\mathbf{k}}(\eta)\,, \qquad u^{\lambda}_{f}(\eta) \equiv \begin{pmatrix} 1 \\ \frac{f(\eta)}{f_+(\eta)} \end{pmatrix}\,, \tag{5.5}$$

where $\varphi_{\mathbf{k}}(\eta)$ is related to the initial field $\varphi_{\mathbf{k}}(0)$ by the linear growth, $\varphi_{\mathbf{k}}(\eta) = D(\eta)\varphi_{\mathbf{k}}(0)$, with $D(0) = 1$. For simplicity, and because this is what happens in most cosmological situations, in the following we will assume that the linear fields are in the growing mode, in which case $u^{\lambda}_{f_+}(\eta) \equiv (1,1)^T$. One can straightforwardly generalize our discussion to include the decaying mode as well.

5.1.1 Perturbation theory kernels

For the time being, we will consider only *deterministic evolution*, i.e., that the matter density contrast $\delta(\eta,\mathbf{x})$, the rescaled velocity divergence $\theta(\eta,\mathbf{x})$, and any tracer $\delta_t(\eta,\mathbf{x})$, are functionals of the linear matter density field $\varphi(\eta,\mathbf{x})$. Later, in sect. 5.4, we will discuss the stochastic contributions to the field evolution induced by small scale modes.

For a deterministic evolution, the non-linear fields can be written as

$$\delta(\mathbf{x},\eta) = \mathcal{F}[\varphi](\mathbf{x},\eta)\,, \qquad \theta(\mathbf{x},\eta) = \mathcal{G}[\varphi](\mathbf{x},\eta)\,, \qquad \delta_t(\mathbf{x},\eta) = \mathcal{K}[\varphi](\mathbf{x},\eta)\,. \tag{5.6}$$

In general, the non-linear velocity field has also a vorticity component, besides the divergence one. In standard perturbation theory, vorticity decays as $1/a$ at linear order [27] and, setting its initial condition to zero, it is not generated at any higher order. Therefore, as long as deterministic evolution is considered, it is consistent to consider a curl-free velocity field. On the other hand, small-scale stochasticity sources vorticity, as discussed in [135, 136], and symmetry arguments can be used to constrain these contributions as well. However, as concluded for instance in [137] the back-reaction of vorticity on the density and velocity divergence correlators is extremely suppressed on the scales reachable by PT methods, therefore we will not include these effects in this work.

Expanding the LHS of eq. (5.6) in perturbation theory using eq. (5.2) in Fourier space and

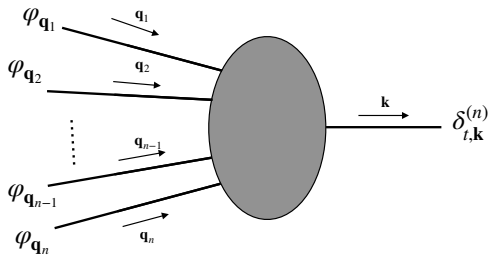

Figure 5.1: Symbolic representation of the $K_n(\mathbf{q}_1, \cdots, \mathbf{q}_n; a)$ halo amplitude.

expanding the RHS in the field φ, we obtain

$$\delta^{(n)}_{\mathbf{k}}(\eta) \equiv \mathcal{I}_{\mathbf{k};\mathbf{q}_1\cdots,\mathbf{q}_n} F_n(\mathbf{q}_1, \cdots, \mathbf{q}_n; \eta)\varphi_{\mathbf{q}_1}(\eta)\cdots\varphi_{\mathbf{q}_n}(\eta), \tag{5.7}$$

$$\theta^{(n)}_{\mathbf{k}}(\eta) \equiv \mathcal{I}_{\mathbf{k};\mathbf{q}_1\cdots,\mathbf{q}_n} G_n(\mathbf{q}_1, \cdots, \mathbf{q}_n; \eta)\varphi_{\mathbf{q}_1}(\eta)\cdots\varphi_{\mathbf{q}_n}(\eta), \tag{5.8}$$

$$\delta^{(n)}_{t,\mathbf{k}}(\eta) \equiv \mathcal{I}_{\mathbf{k};\mathbf{q}_1\cdots,\mathbf{q}_n} K_n(\mathbf{q}_1, \cdots, \mathbf{q}_n; \eta)\varphi_{\mathbf{q}_1}(\eta)\cdots\varphi_{\mathbf{q}_n}(\eta), \tag{5.9}$$

where we have redefined[2]

$$\mathcal{I}_{\mathbf{k};\mathbf{q}_1\cdots,\mathbf{q}_n} \equiv \frac{1}{n!}\int \frac{d^3q_1}{(2\pi)^3}\cdots\frac{d^3q_n}{(2\pi)^3}(2\pi)^3\delta_D\left(\mathbf{k} - \sum_{i=1}^{n}\mathbf{q}_i\right). \tag{5.10}$$

Notice that the functionals $\mathcal{F}$, $\mathcal{G}$, and $\mathcal{K}$ in (5.6) are in general non-local in space and in time. The space non-locality is due to the Poisson equation and to the fact that we consider only the divergence of the velocity fields. Moreover, the perturbative expansion of the kernels can be cast in a form that is local in time without any loss of generality, as in eqs. (5.7), (5.8), and (5.9). Notice also that the kernels F_n and G_n introduced in eqs. 5.7 and 5.8 are, in general, different from those defined in section 2.1 in eq. 2.39 and 2.40, as we will see soon.

The kernels F_n, G_n and K_n can be interpreted as *transition amplitudes* between n linear and one nonlinear fields, represented in Fig. 5.1. As usual, the delta function on the RHS of eq. (5.10) comes from assuming *translational invariance* of the field equations of motion. Moreover, *rotational invariance* imposes that the kernels depend on rotational invariant combinations of the momenta $\mathbf{q}_i$. In writing eqs. (5.7–5.9) we have assumed that the only scales entering the kernels are the external momenta $\mathbf{q}_i$, i.e. there are no other intrinsic scales in the problem (such as e.g. the neutrino masses, massive fields in modified gravity, etc.). Indeed, being obtained by

[2]Note that, differently from eq. (2.33), here we introduce the $(1/n!)$ factor.

a functional expansion around $\varphi_{\mathbf{q}_i}(\eta) = 0$, the kernels contain no information on the intrinsic scales of the initial power spectrum or of higher-order correlators of the initial conditions. We are also assuming no primordial non-gaussianity as, for biased tracers, it would induce a coupling between different scales which would affect the kernels (see, for instance [138]). Moreover, the perturbative expansion is well known to break down at short scales. This is usually taken into account by suitable counterterms. We ignore these terms for the time being and we discuss them in sect. 5.4.

The symmetry of the integration domain and of the multi-dimensional integration measure in (5.9) translates in the requirement that the amplitude is symmetric under exchange of any pair of external momenta, i.e.,

$$K_n(\mathbf{q}_1, \cdots, \mathbf{q}_i, \cdots, \mathbf{q}_j \cdots, \mathbf{q}_n; \eta) = K_n(\mathbf{q}_1, \cdots, \mathbf{q}_j, \cdots, \mathbf{q}_i \cdots, \mathbf{q}_n; \eta)\,. \tag{5.11}$$

Next, we will consider the two sets of symmetries to be imposed on our kernels.

5.1.2 Extended Galilean Invariance

We will use the invariance under time-dependent translations to derive relations between the kernels at different orders. In the non-relativistic limit, the dark matter fluid dynamics is invariant under the following change of coordinates [4, 5],

$$\eta \to \tilde{\eta} = \eta\,, \qquad \mathbf{x} \to \tilde{\mathbf{x}} = \mathbf{x} + \mathbf{d}(\eta)\,, \tag{5.12}$$

followed by an appropriate transformation of the density and velocity fields,

$$\delta(\mathbf{x}, \eta) \to \tilde{\delta}(\tilde{\mathbf{x}}, \tilde{\eta}) = \delta(\mathbf{x}, \eta) \qquad \mathbf{v}(\mathbf{x}, \eta) \to \tilde{\mathbf{v}}(\tilde{\mathbf{x}}, \tilde{\eta}) = \mathbf{v}(\mathbf{x}, \eta) + \mathcal{H}\,\partial_\eta \mathbf{d}(\eta)\,, \tag{5.13}$$

while the Newtonian potential transforms as

$$\Phi \to \Phi - \left[\mathcal{H}\partial_\eta(\mathcal{H}\partial_\eta \mathbf{d}) + \mathcal{H}^2 \partial_\eta \mathbf{d}\right] \cdot \mathbf{x}\,. \tag{5.14}$$

This is a symmetry regardless of the time dependence of $\mathbf{d}$. However, to derive relations between the kernels we will impose that the new solution generated by the time-dependent translation is also the long-wavelength limit of a physical mode satisfying the equations of

motion. As shown in Appendix B this assumption is not necessary but it simplifies a lot the derivation of the constraint (5.27).

We can thus consider $\mathbf{d}(\eta)$ as the zero momentum limit of a space-dependent field. Thus, at the linear level, in Fourier space we have the transformation

$$\delta^{(1)}_{\mathbf{k}} \to \tilde{\delta}^{(1)}_{\mathbf{k}} = \delta^{(1)}_{\mathbf{k}} + i(2\pi)^3\delta_D(\mathbf{k})\mathbf{k}\cdot\mathbf{d}^{(1)} \ , \qquad \mathbf{v}^{(1)}_{\mathbf{k}} \to \tilde{\mathbf{v}}^{(1)}_{\mathbf{k}} = \mathbf{v}^{(1)}_{\mathbf{k}} + (2\pi)^3\delta_D(\mathbf{k})\mathcal{H}\,\partial_\eta\mathbf{d}^{(1)} \ , \tag{5.15}$$

where $\mathbf{d}^{(1)}(\eta)$ is understood as the linear component of the full displacement field $\mathbf{d}(\eta)$. Using eqs. (5.3) and (5.5), this implies

$$\varphi_{\mathbf{k}} \to \tilde{\varphi}_{\mathbf{k}} = \varphi_{\mathbf{k}} + i(2\pi)^3\delta_D(\mathbf{k})\mathbf{k}\cdot\mathbf{d}^{(1)} \ . \tag{5.16}$$

The new fields generated by the time-dependent translation manifestly satisfy the linear continuity equation, $\partial_\eta\tilde{\delta}^{(1)}_{\mathbf{k}} = i\mathbf{k}\cdot\tilde{\mathbf{v}}^{(1)}_{\mathbf{k}}/\mathcal{H}$. Moreover, imposing that the new solution satisfies the Euler equation implies [139]

$$\mathcal{H}\partial_\eta(\mathcal{H}\partial_\eta\mathbf{d}^{(1)}) + \mathcal{H}^2\partial_\eta\mathbf{d}^{(1)} = 4\pi G\bar{\rho}a^2\mathbf{d}^{(1)} \ , \tag{5.17}$$

i.e., the displacement $\mathbf{d}^{(1)}(\eta)$ evolves in time following the same linear growth as the original field $\varphi_{\mathbf{k}}(\eta)$.

As shown by eq. (5.14), this coordinate transformation permits to remove the effect of the long-wavelength gravitational potential and settle in an inertial frame, as allowed by the Equivalence Principle [6]. This symmetry is not limited to the dark matter fluid dynamics. Assuming that all species fall in the same way in the gravitational field, it can be directly applied to any biased tracer such as the galaxy distribution [114, 139], replacing δ with δ_t in eq. (5.13),

$$\delta_t(\mathbf{x},\eta) \to \tilde{\delta}_t(\tilde{\mathbf{x}},\tilde{\eta}) = \delta_t(\mathbf{x},\eta) \ . \tag{5.18}$$

Moreover, the symmetry is non-perturbative, i.e. it remains valid even for very short wavelengths of δ_t, where the complex baryonic physics is difficult to model perturbatively.

Using eq. (5.18), we can write that under the coordinate transformation (5.12), the field transforms as,

$$\delta_t(\mathbf{x},\eta) \to \delta_t(\mathbf{x}-\mathbf{d},\eta) \ , \tag{5.19}$$

which, for each Fourier mode $\mathbf{k}$ can be written as

$$\delta_{t,\mathbf{k}}(\eta) \to e^{i\mathbf{k}\cdot\mathbf{d}(\eta)}\delta_{t,\mathbf{k}}(\eta) = \sum_{m=0}^{\infty} \frac{(i\mathbf{k}\cdot\mathbf{d})^m}{m!}\delta_{t,\mathbf{k}} \,. \tag{5.20}$$

Notice that both $\delta_{t,\mathbf{k}}(\eta)$ and $\mathbf{d}(\eta)$ in the equation above are nonlinear quantities, which we will expand perturbatively. In particular, we will first consider the effect of the linear contribution to $\mathbf{d}(\eta)$, $\mathbf{d}^{(1)}(\eta)$, and then the effect of the higher orders.

5.1.3 Leading Order

Expanding perturbatively the left- and RHS of eq. (5.20) using eq. (5.2) and treating $\mathbf{d}^{(1)}(\eta)$ as the zero momentum limit of a linear field gives, for any given order n,

$$\delta^{(n)}_{t,\mathbf{k}} \to \sum_{m=0}^{n} \frac{(i\mathbf{k}\cdot\mathbf{d}^{(1)})^m}{m!}\delta^{(n-m)}_{t,\mathbf{k}} \,, \tag{5.21}$$

where we have omitted the dependence on η to avoid cluttering. This equation shows how the LHS of eq. (5.9) transforms under (5.12). Let us now consider the RHS of eq. (5.9). If we now perform the shift (5.16) on this side of the equation we have, for the n-th order contribution and omitting again the η dependence,

$$\begin{aligned}
&\mathcal{I}_{\mathbf{k};\mathbf{q}_1\cdots,\mathbf{q}_n} K_n(\mathbf{q}_1,\cdots,\mathbf{q}_n)\varphi_{\mathbf{q}_1}\cdots\varphi_{\mathbf{q}_n} \\
&\to \mathcal{I}_{\mathbf{k};\mathbf{q}_1\cdots,\mathbf{q}_n} K_n(\mathbf{q}_1,\cdots,\mathbf{q}_n) \sum_{m=0}^{n} \begin{pmatrix} n \\ m \end{pmatrix} \epsilon_{\mathbf{q}_1}\cdots\epsilon_{\mathbf{q}_m}\varphi_{\mathbf{q}_{m+1}}\cdots\varphi_{\mathbf{q}_n} \,,
\end{aligned} \tag{5.22}$$

where for compactness we have defined $\epsilon_{\mathbf{k}}(\eta) \equiv i(2\pi)^3\delta_D(\mathbf{k})\mathbf{k}\cdot\mathbf{d}^{(1)}(\eta)$. Notice that the $m=0$ term in the second line coincides with $\delta^{(n)}_{t,\mathbf{k}}(\eta)$.

We can now equate the $\mathcal{O}((\mathbf{d}^{(1)})^m)$ contribution on the RHS of eq. (5.21), i.e.,

$$\frac{i^m}{m!}\left(\mathbf{k}\cdot\mathbf{d}^{(1)}\right)^m \delta^{(n-m)}_{t,\mathbf{k}} = \frac{i^m}{m!}\left(\mathbf{k}\cdot\mathbf{d}^{(1)}\right)^m \mathcal{I}_{\mathbf{k};\mathbf{p}_1\cdots,\mathbf{p}_{n-m}} K_{n-m}(\mathbf{p}_1,\cdots,\mathbf{p}_{n-m})\,\varphi_{\mathbf{p}_1}\cdots\varphi_{\mathbf{p}_{n-m}} \,, \tag{5.23}$$

with that on the RHS of eq. (5.22), and obtain a relation between the n-th order kernel and the $(n-m)$-th one,

$$\begin{aligned}
&\int d^3q_1\cdots d^3q_m(2\pi)^3\delta_D\left(\mathbf{k}-\mathbf{Q}_{n,0}\right)\, K_n(\mathbf{q}_1,\cdots,\mathbf{q}_m,\mathbf{q}_{m+1}\cdots\mathbf{q}_n)\mathbf{q}_1\cdot\mathbf{d}\,\delta_D(\mathbf{q}_1)\cdots\mathbf{q}_m\cdot\mathbf{d}\,\delta_D(\mathbf{q}_m) \\
&= (2\pi)^3\delta_D\left(\mathbf{k}-\mathbf{Q}_{n,m}\right) K_{n-m}(\mathbf{q}_{m+1},\cdots,\mathbf{q}_n)\,\left(\mathbf{k}\cdot\mathbf{d}\right)^m \,,
\end{aligned} \tag{5.24}$$

where we have defined

$$\mathbf{Q}_{n,m} \equiv \sum_{i=m+1}^{n} \mathbf{q}_i \,. \tag{5.25}$$

Due to the momentum delta functions in the first line and the arbitrariness of the shift $\mathbf{d}(\eta)$, the above relation fixes the soft limit of the n-th order kernel as

$$\lim_{\mathbf{q}_1,\cdots,\mathbf{q}_m\to 0} q_1^{i_1}\cdots q_m^{i_m}\, K_n(\mathbf{q}_1,\cdots,\mathbf{q}_m,\mathbf{q}_{m+1}\cdots\mathbf{q}_n) = Q_{n,m}^{i_1}\cdots Q_{n,m}^{i_m}\, K_{n-m}(\mathbf{q}_{m+1}\cdots\mathbf{q}_n) + \mathcal{O}(q)\,. \tag{5.26}$$

By further contracting by $q_1^{i_1}\cdots q_m^{i_m}$ we get a series of constraints on the pole structure of the kernel, which reads

$$\begin{aligned}\lim_{\mathbf{q}_1,\cdots,\mathbf{q}_m\to 0} & K_n(\mathbf{q}_1,\cdots,\mathbf{q}_m,\mathbf{q}_{m+1}\cdots\mathbf{q}_n) \\ &= \frac{\mathbf{q}_1\cdot\mathbf{Q}_{n,m}}{q_1^2}\cdots\frac{\mathbf{q}_m\cdot\mathbf{Q}_{n,m}}{q_m^2}\, K_{n-m}(\mathbf{q}_{m+1}\cdots\mathbf{q}_n) + \mathcal{O}((1/q)^{m-1})\,.\end{aligned} \tag{5.27}$$

These constraints enforce the symmetries of the equations of motions, namely EGI. They are closely related, but not completely equivalent, to the consistency relations of the LSS [4, 5]. Indeed, the latter involve fully non-perturbative correlators, as opposed to the intrinsically perturbative kernels appearing in (5.27). On the other hand, in order to derive the consistency relations, additional assumptions on the state of the system are needed, namely adiabaticity and gaussianity of the initial conditions. These are not necessary here, as least for unbiased tracers, as we implement the symmetries of the equations of motion, regardless of the state.

5.1.4 Next to Leading Order

Eq. (5.20) can also be read as follows. If the initial conditions, once evolved, produce a long mode $\mathbf{d}(\eta)$, then, in the limit of infinitely long wavelength, its effect on the nonlinear tracer field should factorize as

$$e^{i\mathbf{k}\cdot\mathbf{d}(\eta)}\delta_{t,\mathbf{k}}(\eta)\,, \tag{5.28}$$

where $\delta_{t,\mathbf{k}}(\eta)$ is $\mathbf{d}(\eta)$-independent. This statement holds beyond the assumption we made in the previous subsection, namely, that $\mathbf{d}(\eta)$ is the long wavelength limit of a linear field. Indeed,

the coupling of two linear modes, $\varphi_{\mathbf{q}_1}$ and $\varphi_{\mathbf{q}_2}$ gives a displacement field, in Fourier space,

$$\begin{aligned}\tilde{\mathbf{d}}^{(2)}_{\mathbf{q}}(\eta) &= \int^{\tau(\eta)} d\tau'\, \mathbf{v}^{(\mathbf{2})}_{\mathbf{q}}(\tau') \\ &= -i\frac{\mathbf{q}}{q^2}\int^{\eta} d\eta' f_+(\eta')\mathcal{I}_{\mathbf{q};\mathbf{q}_1,\mathbf{q}_2}\, G_2(\mathbf{q}_1,\mathbf{q}_2;\eta')\varphi_{\mathbf{q}_1}(\eta')\varphi_{\mathbf{q}_2}(\eta')\,, \\ &= -i\frac{\mathbf{q}}{q^2}\int^{\eta} d\eta' f_+(\eta')\frac{D_+(\eta')^2}{D_+(\eta)^2}\mathcal{I}_{\mathbf{q};\mathbf{q}_1,\mathbf{q}_2}\, G_2(\mathbf{q}_1,\mathbf{q}_2;\eta')\varphi_{\mathbf{q}_1}(\eta)\varphi_{\mathbf{q}_2}(\eta)\,. \end{aligned} \tag{5.29}$$

In the $\mathbf{q} = \mathbf{q}_1 + \mathbf{q}_2 \to 0$ limit, the coupling between these two modes gives a contribution to the nonlinear field of the form (5.28). The lowest contribution is of third order,

$$i\mathbf{k}\cdot\mathbf{d}^{(2)}(\eta)\delta^{(1)}_{t,\mathbf{k}}(\eta)\,, \tag{5.30}$$

where the 'zero mode' displacement is defined as

$$\mathbf{d}^{(2)}(\eta) \equiv \frac{1}{V}\int_V d^3x \int \frac{d^3q}{(2\pi)^3} e^{-i\mathbf{q}\cdot\mathbf{x}}\, \tilde{\mathbf{d}}^{(2)}_{\mathbf{q}}(\eta), \tag{5.31}$$

with $V = (2\pi)^3\delta_D(k=0)$. Using (5.29) in (5.30) then gives

$$\begin{aligned} &i\mathbf{k}\cdot\mathbf{d}^{(2)}(\eta)\delta^{(1)}_{t,\mathbf{k}}(\eta) \\ &\quad= \frac{1}{V}\mathcal{I}_{0;\mathbf{q}_1,\mathbf{q}_2}\frac{\mathbf{k}\cdot\mathbf{q}_{12}}{q_{12}^2}\int^{\eta} d\eta' f_+(\eta')\frac{D_+(\eta')^2}{D_+(\eta)^2}\, G_2(\mathbf{q}_1,\mathbf{q}_2;\eta')\varphi_{\mathbf{q}_1}(\eta)\varphi_{\mathbf{q}_2}(\eta)K_1(\mathbf{k};\eta)\varphi_{\mathbf{k}}(\eta)\,, \end{aligned} \tag{5.32}$$

where we have defined $\mathbf{q}_{ij} \equiv \mathbf{q}_i + \mathbf{q}_j$. NLO EGI requires that this contribution is contained in the expression for $\delta^{(3)}_{t,\mathbf{k}}(\eta)$ from eq. (5.9). In order to isolate it, we insert the identity

$$1 = \frac{(2\pi)^3}{V}\left(\delta_D(\mathbf{q}_{12}) + \delta_D(\mathbf{q}_{23}) + \delta_D(\mathbf{q}_{31})\right) + \left[1 - \frac{(2\pi)^3}{V}\left(\delta_D(\mathbf{q}_{12}) + \delta_D(\mathbf{q}_{23}) + \delta_D(\mathbf{q}_{31})\right)\right] \tag{5.33}$$

in the momentum integral. From the first term of the identity we get

$$\begin{aligned}\delta^{(3)}_{t,\mathbf{k}}(\eta) &\supset \mathcal{I}_{\mathbf{k};\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}\, K_3(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3;\eta)\varphi_{\mathbf{q}_1}(\eta)\varphi_{\mathbf{q}_2}(\eta)\varphi_{\mathbf{q}_3}(\eta)\frac{(2\pi)^3}{V}\left(\delta_D(\mathbf{q}_{12}) + \delta_D(\mathbf{q}_{23}) + \delta_D(\mathbf{q}_{31})\right) \\ &= \frac{1}{V}\,\mathcal{I}_{0;\mathbf{q}_1,\mathbf{q}_2}\, K_3(\mathbf{q}_1,\mathbf{q}_2,\mathbf{k};\eta)\varphi_{\mathbf{q}_1}(\eta)\varphi_{\mathbf{q}_2}(\eta)\varphi_{\mathbf{k}}(\eta), \end{aligned} \tag{5.34}$$

which should be identified with (5.32), therefore leading to the relation

$$\lim_{\mathbf{q}_{12}\to 0} K_3(\mathbf{q}_1,\mathbf{q}_2,\mathbf{k};\eta) \;\supset\; \frac{\mathbf{k}\cdot\mathbf{q}_{12}}{q_{12}^2}K_1(\mathbf{k};\eta)\int^{\eta} d\eta' f_+(\eta')\frac{D_+(\eta')^2}{D_+(\eta)^2}G_2(\mathbf{q}_1,\mathbf{q}_2;\eta')\,. \tag{5.35}$$

Notice that, as we will discuss in Sect. 5.1.6, mass and momentum conservation implies $G_2(\mathbf{q}_1, \mathbf{q}_2; \eta) \sim q_{12}^2$ as $q_{12} \to 0$. Therefore, unlike eq. (5.27), the kernels have no pole when the sum of two (or more, see next subsection) momenta vanish. Moreover, the term at the RHS of eq. (5.35) might also not be the leading one in this limit, as, for a generic tracer, constant contributions are also present. Nevertheless this equation fixes the structure of the terms containing the $\mathbf{k} \cdot \mathbf{q}_{12}/q_{12}^2$ combination in the $q_{12} \to 0$ limit which, as will see, provides meaningful constraints to the kernels.

5.1.5 $\mathbf{N^{l-1}LO}$

It is possible to generalize eq. (5.35) to a general order of the kernel and with the sum of an arbitrary number of momenta going to zero. Exploiting the coupling between l linear modes, one can expand the displacement field as

$$\tilde{\mathbf{d}}_{\mathbf{q}}^{(l)}(\eta) = -i\frac{\mathbf{q}}{q^2}\int^{\eta} d\eta' f_+(\eta') \left(\frac{D_+(\eta')}{D_+(\eta)}\right)^l \mathcal{I}_{\mathbf{q};\mathbf{q}_1,\ldots,\mathbf{q}_l} G_l(\mathbf{q}_1, \ldots, \mathbf{q}_l; \eta')\varphi_{\mathbf{q}_1}(\eta)\ldots\varphi_{\mathbf{q}_l}(\eta), \tag{5.36}$$

from which we can obtain the general relation for the n-th order kernel

$$\begin{aligned}\lim_{\mathbf{Q}_{l,0}\to 0} K_n(\mathbf{q}_1, \ldots, \mathbf{q}_l, \mathbf{q}_{l+1}, \ldots, \mathbf{q}_n) \supset \\ \frac{\mathbf{k}\cdot\mathbf{Q}_{l,0}}{Q_{l,0}^2}\int^{\eta} d\eta'\, f_+(\eta') \left(\frac{D_+(\eta')}{D_+(\eta)}\right)^l G_l(\mathbf{q}_1, \ldots, \mathbf{q}_l; \eta') K_{n-l}(\mathbf{q}_{l+1}, \ldots, \mathbf{q}_n; \eta)\,,\end{aligned} \tag{5.37}$$

where we recall that $\mathbf{Q}_{l,0} = \sum_{i=1}^{l} \mathbf{q}_i$ (see eq. (5.25)). The number of momenta l, for which the sum is set to zero in eq. (5.37), gives the order of the perturbative expansion for the long-wavelength displacement. Considering, for example, $l = 1$ and iterating m times eq. (5.37), gives the leading order extended Galilean invariance expressed in eq. (5.27). To obtain the NLO order one should choose $l = 2$, meaning that the sum of two internal momenta is set to zero, so that the first correction of this type arise for kernels K_n with $n \geq 3$, leading to eq. (5.35). The explicit calculation of the NNLO condition, which starts to play a role at the fourth PT order, is presented in Appendix C.

5.1.6 Mass and momentum conservation

In Sec. 5.1.2 we have discussed the behavior of the kernels $K_n(\mathbf{q}_1, \cdots \mathbf{q}_n; \eta)$ as one or more of the momenta of the linear fields, q_i, vanishes. In this section we inspect the opposite limit, in which the momentum of the nonlinear field, k, is $\ll q_i$ for all i's. We consider a tracer that satisfies mass and momentum conservation, such as for instance the dark matter density contrast.

Mass conservation imposes that the first term in this expansion vanishes at all times,

$$\delta_{\mathbf{k}=0}(\eta) = \int d^3x\, \delta(\mathbf{x}, \eta) = 0\,. \tag{5.38}$$

Momentum conservation, namely that the center of mass of the dark matter distribution is fixed, imposes that the second term vanishes,

$$\left.\frac{\partial \delta_{\mathbf{k}}(\eta)}{\partial k^i}\right|_{\mathbf{k}=0} = \int d^3x x^i\, \delta(\mathbf{x}, \eta) = 0\,. \tag{5.39}$$

These two conditions on (5.2) give the two independent constraints on the kernels,

$$\begin{aligned} &\lim_{\mathbf{Q}_{n,0}\to 0} F_n(\mathbf{q}_1, \cdots \mathbf{q}_n; \eta) = 0\,, \\ &\lim_{\mathbf{Q}_{n,0}\to 0} \frac{\partial}{\partial q_1^i} F_n(\mathbf{q}_1, \cdots \mathbf{q}_n; \eta) = 0\,, \end{aligned} \tag{5.40}$$

where the limit is taken by keeping all the individual q_i's non vanishing. The conditions above ensure that the density contrast decouples as

$$\delta_{\mathbf{k}}(\eta) = \mathcal{O}\left(k^2/q_i^2\right)\,, \tag{5.41}$$

when the external momentum is much smaller than the q_i's, i.e. $k \ll q_i$, as implied by general arguments on momentum conservation [79, 94, 135, 140].

The two conditions eq. (5.38) and (5.39) hold for the matter density contrast and velocity divergence [135], so that eq. (5.40) apply to the matter and velocity kernels , i.e. F_n and G_n. However, they do not hold for a generic tracer, such as the galaxy number density, which does not satisfy a conservation equation and for which eqs. (5.40) do not apply. Therefore, in the following we will denote as 'matter kernels' the ones satisfying the conditions in eq. (5.40). We turn now to derive their explicit forms by imposing the symmetries discussed above.

5.2 Matter Kernels

5.2.1 Imposing the constraints

Here we will implement all the constraints discussed in Sec. 5.1, including the last one, and derive the expressions of the matter kernels, i.e. for F_n and G_n. The details of these calculations are explicitly given in App. D. We will use isotropy to write the kernels in terms of rotationally invariant objects constructed from the external momenta. For $n = 1$ we have one invariant, i.e. $q_1^2 = k^2$. For $n \geq 2$ we need $3n - 3$ of them and we can take them from the $n(n+1)/2$ scalar products $\mathbf{q}_l \cdot \mathbf{q}_m$. For $n = 2$ these are q_1^2, q_2^2, $\mathbf{q}_1 \cdot \mathbf{q}_2$, and for $n = 3$ they are q_1^2, q_2^2, q_3^2, $\mathbf{q}_1 \cdot \mathbf{q}_2$, $\mathbf{q}_1 \cdot \mathbf{q}_3$ and $\mathbf{q}_2 \cdot \mathbf{q}_3$. We will then construct, up to third order, the most general dimensionless functions built out of the $\mathbf{q}_i$ momenta satisfying the above conditions. We restrict to rational functions, which is consistent with the perturbative nature of the present analysis, as expressed by eq. (5.2).

For $n = 1$, the kernel depends only on one momentum, $\mathbf{q}$, but isotropy implies that the dimensionless rotational invariant is just a numerical constant, implying

$$F_1(\mathbf{q}_1; \eta) = 1 \,. \tag{5.42}$$

For $n = 2$ there are two independent external momenta. The requirement discussed in sect. 5.1.2 above restricts the order of the poles, when one of the momenta vanishes, at most to the first. Therefore, the rotational invariants we can build out of $\mathbf{q}_1$ and $\mathbf{q}_2$ are limited to four:

$$1 \,, \qquad \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1^2}, \qquad \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_2^2}, \qquad \frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)^2}{q_1^2 q_2^2} \,. \tag{5.43}$$

We find it convenient, instead, to use as basis functions the following combinations

$$1, \qquad \gamma(\mathbf{q}, \mathbf{p}) = 1 - \frac{(\mathbf{q} \cdot \mathbf{p})^2}{q^2 p^2} \,, \qquad \beta(\mathbf{q}, \mathbf{p}) \equiv \frac{|\mathbf{q} + \mathbf{p}|^2 \, \mathbf{q} \cdot \mathbf{p}}{2 q^2 p^2} \,, \qquad \alpha_a(\mathbf{q}, \mathbf{p}) = \frac{\mathbf{q} \cdot \mathbf{p}}{q^2} - \frac{\mathbf{p} \cdot \mathbf{q}}{p^2} \,, \tag{5.44}$$

where the last combination appears only at $n \geq 3$. Indeed, requiring total symmetrization, eq. (5.11), restricts the invariants to three, leading to the most generic form for the second-order kernel,

$$F_2(\mathbf{q}_1, \mathbf{q}_2; \eta) = \, a_0^{(2)}(\eta) + a_1^{(2)}(\eta) \, \gamma(\mathbf{q}_1, \mathbf{q}_2) + a_2^{(2)}(\eta) \, \beta(\mathbf{q}_1, \mathbf{q}_2) \,. \tag{5.45}$$

Imposing the constraints discussed in sect. 5.1 gives $a_0^{(2)} = 0$ from mass conservation, and $a_2^{(2)} = 2$ from EGI (see App. D), so that the most general matter kernel at $n = 2$ is given by

$$F_2(\mathbf{q}_1, \mathbf{q}_2; \eta) = 2\beta(\mathbf{q}_1, \mathbf{q}_2) + a_1^{(2)}(\eta)\, \gamma(\mathbf{q}_1, \mathbf{q}_2)\,, \tag{5.46}$$

which contains only one independent function of time, $a_1^{(2)}(\eta)$.

For $n = 3$, the most general matter kernel can be obtained by taking the direct product between the independent structures built out of $\mathbf{q}_1$ and $\mathbf{q}_2$ and those built out of $\mathbf{q}_{12}$ and $\mathbf{q}_3$, and then symmetrizing over $\mathbf{q}_1$, $\mathbf{q}_2$, and $\mathbf{q}_3$,

$$(1, \gamma(\mathbf{q}_1, \mathbf{q}_2), \beta(\mathbf{q}_1, \mathbf{q}_2)) \otimes (1, \gamma(\mathbf{q}_{12}, \mathbf{q}_3), \alpha_a(\mathbf{q}_{12}, \mathbf{q}_3), \beta(\mathbf{q}_{12}, \mathbf{q}_3)) + \text{cyclic}\,. \tag{5.47}$$

The $\alpha_a(\mathbf{q}_1, \mathbf{q}_2)$ term in the first factor is absent due to symmetrization. The symmetrization process is completed by adding the remaining two cyclic permutations. Thus, the most general structure of the third-order kernel starts with 12 terms. Suppressing the dependence on η to avoid clutter, this reads

$$\begin{aligned} F_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) = {} & \frac{1}{3} a_0^{(3)} + a_1^{(3)} \gamma(\mathbf{q}_1, \mathbf{q}_2) + a_2^{(3)} \gamma(\mathbf{q}_{12}, \mathbf{q}_3) + a_3^{(3)} \beta(\mathbf{q}_1, \mathbf{q}_2) + a_4^{(3)} \beta(\mathbf{q}_{12}, \mathbf{q}_3) \\ & + a_5^{(3)} \gamma(\mathbf{q}_1, \mathbf{q}_2)\gamma(\mathbf{q}_{12}, \mathbf{q}_3) + a_6^{(3)} \beta(\mathbf{q}_1, \mathbf{q}_2)\beta(\mathbf{q}_{12}, \mathbf{q}_3) + a_7^{(3)} \gamma(\mathbf{q}_1, \mathbf{q}_2)\beta(\mathbf{q}_{12}, \mathbf{q}_3) \\ & + a_8^{(3)} \beta(\mathbf{q}_1, \mathbf{q}_2)\gamma(\mathbf{q}_{12}, \mathbf{q}_3) + \left(a_9^{(3)} + a_{10}^{(3)} \gamma(\mathbf{q}_1, \mathbf{q}_2) + a_{11}^{(3)} \beta(\mathbf{q}_1, \mathbf{q}_2) \right) \alpha_a(\mathbf{q}_{12}, \mathbf{q}_3) \\ & + \text{cyclic}\,. \end{aligned} \tag{5.48}$$

Imposing the symmetries discussed in sect. 5.1 provides 10 independent conditions on the 12 coefficients appearing in (5.48) (see again App. D for details), and thus restricts the total number of independent coefficients to three, i.e.

$$\left\{ a_1^{(2)},\ a_5^{(3)},\ a_{10}^{(3)} \right\}\,. \tag{5.49}$$

In this case the kernel reads

$$F_1(\mathbf{q}_1) = 1 \tag{5.50}$$

$$F_2(\mathbf{q}_1, \mathbf{q}_2) = 2\beta(\mathbf{q}_1, \mathbf{q}_2) + a_1^{(2)}\, \gamma(\mathbf{q}_1, \mathbf{q}_2)\,, \tag{5.51}$$

$$\begin{aligned} F_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) = {} & 2\, \beta(\mathbf{q}_1, \mathbf{q}_2)\beta(\mathbf{q}_{12}, \mathbf{q}_3) + a_5^{(3)} \gamma(\mathbf{q}_1, \mathbf{q}_2)\gamma(\mathbf{q}_{12}, \mathbf{q}_3) \\ & - 2 \left(a_{10}^{(3)} - h \right) \gamma(\mathbf{q}_1, \mathbf{q}_2)\beta(\mathbf{q}_{12}, \mathbf{q}_3) + 2(a_1^{(2)} + 2\, a_{10}^{(3)} - h)\beta(\mathbf{q}_1, \mathbf{q}_2)\gamma(\mathbf{q}_{12}, \mathbf{q}_3) \\ & + a_{10}^{(3)} \gamma(\mathbf{q}_1, \mathbf{q}_2)\alpha_a(\mathbf{q}_{12}, \mathbf{q}_3) + \text{cyclic}\,, \end{aligned} \tag{5.52}$$

where the time-dependent coefficient h is not independent and is defined below (eq. (5.56)).

All the kernel relations above hold analogously for the velocity divergence: G_n can be rewritten up to the third order as eq. (5.42), (5.46) and (5.52), i.e.,

$$G_1(\mathbf{q}_1) = 1 \tag{5.53}$$

$$G_2(\mathbf{q}_1, \mathbf{q}_2) = 2\beta(\mathbf{q}_1, \mathbf{q}_2) + d_1^{(2)}\,\gamma(\mathbf{q}_1, \mathbf{q}_2)\,, \tag{5.54}$$

$$\begin{aligned} G_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) = {} & 2\,\beta(\mathbf{q}_1, \mathbf{q}_2)\beta(\mathbf{q}_{12}, \mathbf{q}_3) + d_5^{(3)}\gamma(\mathbf{q}_1, \mathbf{q}_2)\gamma(\mathbf{q}_{12}, \mathbf{q}_3) \\ & - 2\,\left(d_{10}^{(3)} - h\right)\gamma(\mathbf{q}_1, \mathbf{q}_2)\beta(\mathbf{q}_{12}, \mathbf{q}_3) + 2(d_1^{(2)} + 2\,d_{10}^{(3)} - h)\beta(\mathbf{q}_1, \mathbf{q}_2)\gamma(\mathbf{q}_{12}, \mathbf{q}_3) \\ & + d_{10}^{(3)}\gamma(\mathbf{q}_1, \mathbf{q}_2)\alpha_a(\mathbf{q}_{12}, \mathbf{q}_3) + \text{cyclic}\,. \end{aligned} \tag{5.55}$$

The coefficient h is defined as

$$h(\eta) \equiv \int^{\eta} d\eta'\, f_+(\eta') \left[\frac{D_+(\eta')}{D_+(\eta)}\right]^2 d_1^{(2)}(\eta')\,, \tag{5.56}$$

with $d_1^{(2)}$ being the time-dependent coefficient of G_2 in eq. (5.54).

5.2.2 Time dependence

The discussion so far did not assume any specific cosmological model, namely a background evolution and a definite form of the system of Euler, continuity and Poisson equations. In this section we show how fixing a cosmology allows to derive the time dependence on the three parameters in eq. (5.49) and the corresponding ones for the velocity field. We will do this first in the standard ΛCDM case and then for a particular modified gravity model.

5.2.3 ΛCDM

Extending the notation of eq. (5.3) to nonlinear fields, and switching the time variable from η to

$$\chi \equiv \log \frac{D_+(\eta)}{D_+(0)}\,, \tag{5.57}$$

we can write the equation of motion for the n-th order ($n > 1$) one as (see e.g. [141]),

$$\left(\delta_{\lambda\lambda'}\partial_\chi + \Omega_{\lambda\lambda'}(\chi)\right)\Psi_{\mathbf{k}}^{\lambda',(n)}(\chi) = \mathcal{I}_{\mathbf{k};\mathbf{q}_1,\mathbf{q}_2}\gamma_{\lambda\lambda'\lambda''}(\mathbf{q}_1, \mathbf{q}_2)\sum_{m=1}^{n-1}\Psi_{\mathbf{q}_1}^{\lambda',(m)}(\chi)\Psi_{\mathbf{q}_2}^{\lambda'',(n-m)}(\chi)\,, \tag{5.58}$$

where Ψ^λ ($\lambda = 1, 2$) is the doublet

$$\Psi^1(\mathbf{x}, \eta) \equiv \delta(\mathbf{x}, \eta)\,, \qquad \Psi^2(\mathbf{x}, \eta) \equiv \theta(\mathbf{x}, \eta)\,. \tag{5.59}$$

Moreover, assuming ΛCDM, we have

$$\mathbf{\Omega}(\chi) = \begin{pmatrix} 0 & -1 \\ -\frac{3}{2}\frac{\Omega_m(\chi)}{f_+^2(\chi)} & \frac{3}{2}\frac{\Omega_m(\chi)}{f_+^2(\chi)} - 1 \end{pmatrix}\,, \tag{5.60}$$

$$\begin{aligned} \gamma_{121}(\mathbf{q}_1, \mathbf{q}_2) &= \gamma_{112}(\mathbf{q}_2, \mathbf{q}_1) = \frac{1}{2}\left(1 + \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1^2}\right)\,, \\ \gamma_{222}(\mathbf{q}_1, \mathbf{q}_2) &= \beta(\mathbf{q}_1, \mathbf{q}_2)\,, \end{aligned} \tag{5.61}$$

and zero otherwise.

Using eq. (5.2) we obtain the evolution equations of the kernels. At order $n = 2$, noting that $\gamma_{112}(\mathbf{q}_1, \mathbf{q}_2) + \gamma_{121}(\mathbf{q}_1, \mathbf{q}_2) = \gamma(\mathbf{q}_1, \mathbf{q}_2) + \beta(\mathbf{q}_1, \mathbf{q}_2)$ we get

$$\begin{aligned} &(\partial_\chi + 2)\, F_2(\mathbf{q}_1, \mathbf{q}_2; \chi) - G_2(\mathbf{q}_1, \mathbf{q}_2; \chi) = 2\big[\gamma(\mathbf{q}_1, \mathbf{q}_2) + \beta(\mathbf{q}_1, \mathbf{q}_2)\big]\,, \\ &\left(\partial_\chi + 1 + \frac{3}{2}\frac{\Omega_m(\chi)}{f_+^2(\chi)}\right) G_2(\mathbf{q}_1, \mathbf{q}_2; \chi) - \frac{3}{2}\frac{\Omega_m(\chi)}{f_+^2(\chi)} F_2(\mathbf{q}_1, \mathbf{q}_2; \chi) = 2\,\beta(\mathbf{q}_1, \mathbf{q}_2)\,. \end{aligned} \tag{5.62}$$

Moreover, using the form of the kernels in eq. (5.46) and selecting the growing mode initial condition, we get the system of two coupled equations for the kernel coefficients,

$$\begin{aligned} (\partial_\chi + 1) a_1^{(2)} &= 2 - a_1^{(2)} + d_1^{(2)}\,, \\ (\partial_\chi + 1) d_1^{(2)} &= \frac{3}{2}\frac{\Omega_m}{f_+^2}\left(a_1^{(2)} - d_1^{(2)}\right)\,. \end{aligned} \tag{5.63}$$

These equations are to be solved with the initial conditions

$$\lim_{\chi \to -\infty} \{a_1^{(2)},\, d_1^{(2)}\} = \{a_{1,\mathrm{EdS}}^{(2)},\, d_{1,\mathrm{EdS}}^{(2)}\}\,, \tag{5.64}$$

where the Einstein de Sitter values are the solutions corresponding to $\Omega_m / f_+^2 = 1$ which are finite in the $\chi \to -\infty$ limit,

$$a_{1,\mathrm{EdS}}^{(2)} = \frac{10}{7}\,, \qquad d_{1,\mathrm{EdS}}^{(2)} = \frac{6}{7}\,. \tag{5.65}$$

The analytic solutions are given by

$$\begin{aligned} a_1^{(2)}(\chi) &= 2 - e^{-\chi} \int_{-\infty}^{\chi} d\chi'\, e^{\chi'} \Delta a_1^{(2)}(\chi)\,, \\ d_1^{(2)}(\chi) &= a_1^{(2)}(\chi) - \Delta a_1^{(2)}(\chi)\,, \end{aligned} \tag{5.66}$$

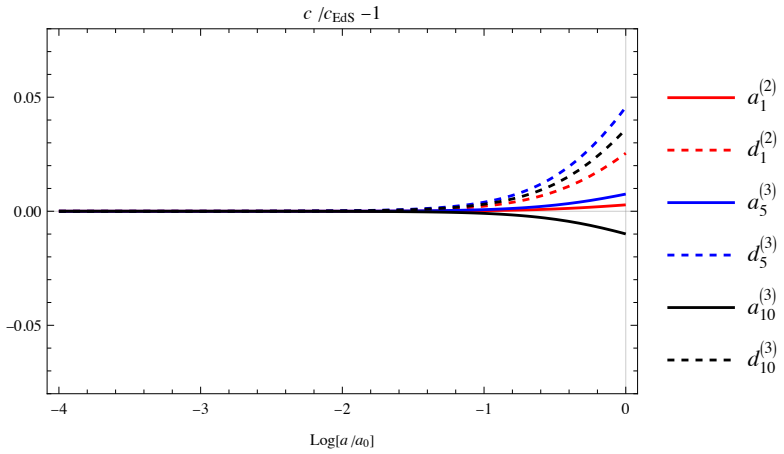

Figure 5.2: Evolution of the coefficients of the kernels in ΛCDM (for $\Omega_m^0 = 0.27$), normalized to the EdS limit.

where

$$\Delta a_1^{(2)}(\chi) = 2\, e^{-\chi} \int_{-\infty}^{\chi} d\chi'\; e^{\chi' - \int_{\chi'}^{\chi} d\chi'' \left(1 + \frac{3}{2}\frac{\Omega_m(\chi'')}{f_+^2(\chi'')}\right)} \,. \tag{5.67}$$

Going to the next order, at $n = 3$, and equating the coefficients of the different structures in γ, β and α_a, we get the coupled evolution equations for the 6 independent parameters appearing in the kernels for the density and velocity fields. These, as well as their analytic solutions, are written in App. E. An example of their behavior is given in Fig. 5.2. Notice that the deviations of the velocity coefficients (dashed lines) from the EdS is of order 4% at low redshifts, and is more pronounced than that of matter (solid lines).

Again, setting $\Omega_m / f_+^2 = 1$, the analytical solutions recover the results of the EdS limit for the third-order density-field kernels,

$$a^{(3)}_{5,\mathrm{EdS}} = \frac{8}{9}\,, \qquad a^{(3)}_{10,\mathrm{EdS}} = -\frac{1}{9}\,, \tag{5.68}$$

and the third-order velocity ones,

$$d^{(3)}_{5,\mathrm{EdS}} = \frac{8}{21}\,, \qquad d^{(3)}_{10,\mathrm{EdS}} = -\frac{1}{21}\,. \tag{5.69}$$

5.2.4 nDGP

We have introduced the nDGP model in section (3.2). In the notation adopted in this chapter, the analog for nDGP of the equations of motions eq. (5.58) for ΛCDM is

$$
\begin{aligned}
\left(\delta_{\lambda\lambda'}\partial_\chi + \Omega_{\lambda\lambda'}\right)\delta_{\mathbf{k}}^{\lambda',(n)} =& \mathcal{I}_{\mathbf{k};\mathbf{q}_1,\mathbf{q}_2}\gamma_{\lambda\lambda'\lambda''}(\mathbf{q}_1,\mathbf{q}_2)\sum_{m=1}^{n-1}\delta_{\mathbf{q}_1}^{\lambda',(m)}\delta_{\mathbf{q}_2}^{\lambda'',(n-m)} \\
&+ \frac{\mu_2}{f_+^2}\left(\frac{3}{2}\Omega_m\right)^2\mathcal{I}_{\mathbf{k};\mathbf{q}_1,\mathbf{q}_2}\gamma(\mathbf{q}_1,\mathbf{q}_2)\sum_{m=1}^{n-1}\delta_{\mathbf{q}_1}^{(m)}(\chi)\delta_{\mathbf{q}_2}^{(n-m)}(\chi) \\
&+ \frac{\mu_3}{f_+^2}\left(\frac{3}{2}\Omega_m\right)^3\mathcal{I}_{\mathbf{k};\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}\gamma(\mathbf{q}_1,\mathbf{q}_2)\sum_{m=1}^{n-2}\sum_{l=1}^{n-2}\delta_{\mathbf{q}_1}^{(m)}\delta_{\mathbf{q}_2}^{(l)}\delta_{\mathbf{q}_3}^{(n-m-l)}\,,
\end{aligned}
\tag{5.70}
$$

where the explicit dependence on χ has been removed to reduce clutter and

$$
\mathbf{\Omega} = \begin{pmatrix} 0 & -1 \\ -\frac{3}{2}\frac{\Omega_m}{f_+^2}\mu & \frac{3}{2}\frac{\Omega_m}{f_+^2}\mu - 1 \end{pmatrix}. \tag{5.71}
$$

Again, using eq. (5.2) and the one for the velocity field, we obtain the evolution equations for the kernels in this modified gravity scenario. At order $n = 2$ and using eq. (5.46) for the kernels we obtain

$$
\begin{aligned}
(\partial_\chi + 1)a_1^{(2)} &= 2 - a_1^{(2)} + d_1^{(2)} + 2\frac{\mu_2}{f_+^2}\left(\frac{3}{2}\Omega_m\right)^2, \\
(\partial_\chi + 1)d_1^{(2)} &= \frac{3}{2}\frac{\Omega_m}{f_+^2}\left(a_1^{(2)} - d_1^{(2)}\right) + 2\frac{\mu_2}{f_+^2}\left(\frac{3}{2}\Omega_m\right)^2,
\end{aligned}
\tag{5.72}
$$

to be solved with the same initial condition expressed in eq. (5.64). The analytic solution is

$$
\begin{aligned}
a_1^{(2)}(\chi) &= 2 - e^{-\chi}\int_{-\infty}^{\chi} d\chi'\, e^{\chi'}\left[\Delta a_1^{(2)}(\chi') - 2\frac{\mu_2}{f_+^2}\left(\frac{3}{2}\Omega_m\right)^2\right], \\
d_1^{(2)}(\chi) &= a_1^{(2)}(\chi) - \Delta a_1^{(2)}(\chi)\,,
\end{aligned}
\tag{5.73}
$$

where in nDGP

$$
\Delta a_1^{(2)}(\chi) = 2e^{-\chi}\int_{-\infty}^{\chi} d\chi'\, e^{\chi' - \int_{\chi'}^{\chi} d\chi''\left(1 + \frac{3}{2}\frac{\Omega_m(\chi'')}{f_+^2(\chi'')}\mu(\chi'')\right)}. \tag{5.74}
$$

Analogously to the ΛCDM case we can find the differential equations and the analytic solutions for the third order kernel and parameters: the result of these calculations are presented in App. E.2. An example of the behavior of the coefficients in a nDGP cosmology is presented

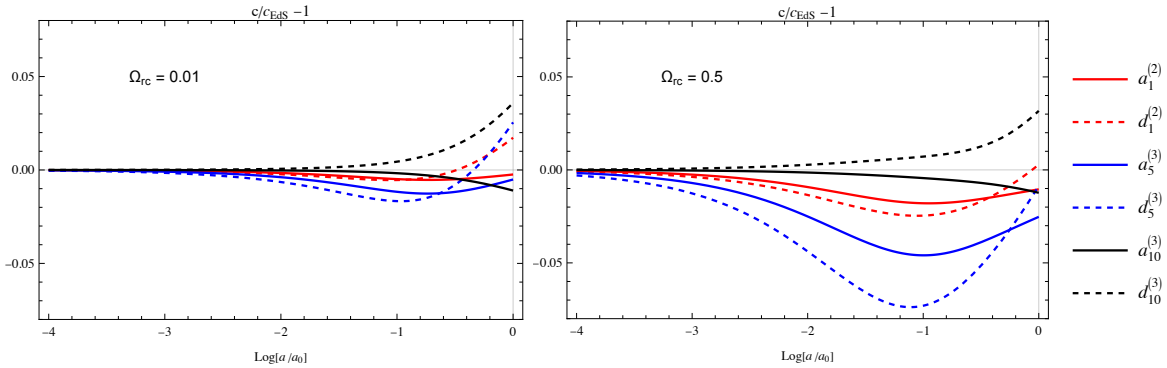

Figure 5.3: Evolution of the coefficients of the kernels in nDGP (for $\Omega_m^0 = 0.27$ and two possible values of the nDGP parameter $\Omega_{rc} = 0.01,\ 0.5$ in the left and right plot, respectively), normalized to the EdS limit.

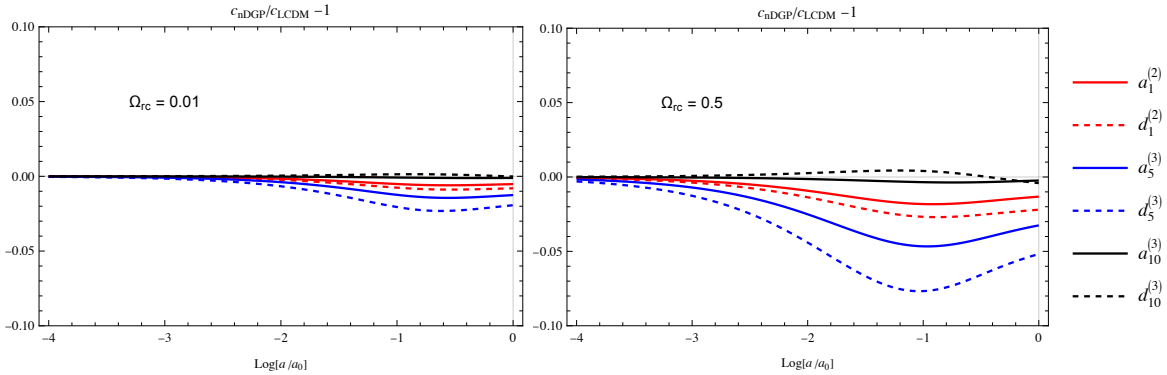

Figure 5.4: Comparison of the time evolution of the matter kernel coefficients in nDGP cosmologies compared to the ΛCDM ones ($\Omega_m^0 = 0.27$ and $\Omega_{rc} = 0.01,\ 0.5$ in the left and right plot, respectively).

in Fig. 5.3. Notice that, similarly to what happened in the ΛCDM case, the velocity coefficients (dashed lines) are more sensible to a change in the time behavior of the growing mode compared to the matter ones (solid lines). A comparison between the exact solutions in the ΛCDM case and the nDGP one is given in Fig. 5.4, for two representative values of the nDGP parameter Ω_{rc}.

5.3 General tracers

5.3.1 Kernels

The number density contrast for a generic tracer is not expected to satisfy a continuity equation, or to fulfill momentum conservation. Therefore, in order to obtain the most generic non-linear kernels for biased tracers, the condition from mass and momentum conservation given by eq. (5.40) should be lifted, while keeping the ones from EGI and the Equivalence Principle.

The explicit calculations of the kernels in this case can be also found in App. D. Up to third order, mass and momentum conservation give four independent constraints. So, compared to the matter case, which is described by three independent coefficients (one for $n = 2$ and two for $n = 3$), the kernels for general tracers have a total of 7 independent coefficients, which can be chosen to be

$$\left\{c_0^{(1)},\ c_0^{(2)},\ c_1^{(2)},\ c_0^{(3)},\ c_1^{(3)},\ c_5^{(3)},\ c_{10}^{(3)}\right\}\,. \tag{5.75}$$

The kernels are given by,

$$K_1(\mathbf{q}_1) = c_0^{(1)}\,, \tag{5.76}$$

$$K_2(\mathbf{q}_1,\mathbf{q}_2) = \ c_0^{(2)} + 2\,c_0^{(1)}\,\beta(\mathbf{q}_1,\mathbf{q}_2) + c_1^{(2)}\,\gamma(\mathbf{q}_1,\mathbf{q}_2)\,, \tag{5.77}$$

$$\begin{aligned} K_3(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) = {}& \frac{1}{3}\,c_0^{(3)} + c_1^{(3)}\gamma(\mathbf{q}_1,\mathbf{q}_2) + 2c_0^{(2)}\beta(\mathbf{q}_1,\mathbf{q}_2) \\ & + c_5^{(3)}\gamma(\mathbf{q}_1,\mathbf{q}_2)\gamma(\mathbf{q}_{12},\mathbf{q}_3) + 2\,c_0^{(1)}\beta(\mathbf{q}_1,\mathbf{q}_2)\beta(\mathbf{q}_{12},\mathbf{q}_3) \\ & + 2(h\,c_0^{(1)} - \ c_{10}^{(3)})\gamma(\mathbf{q}_1,\mathbf{q}_2)\beta(\mathbf{q}_{12},\mathbf{q}_3) + 2(c_1^{(2)} + 2\,c_{10}^{(3)} - h\,c_0^{(1)})\beta(\mathbf{q}_1,\mathbf{q}_2)\gamma(\mathbf{q}_{12},\mathbf{q}_3) \\ & + c_{10}^{(3)}\gamma(\mathbf{q}_1,\mathbf{q}_2)\alpha_a(\mathbf{q}_{12},\mathbf{q}_3) + \text{cyclic}\,, \end{aligned} \tag{5.78}$$

where the coefficient h has been defined in eq. (5.56). Notice that it enters the matter kernels and depends only on the underlying cosmology and not on the type of tracer. For instance, in the EdS case, it is given by (see eq. (5.65))

$$h_{\rm EdS}(\eta) = \frac{3}{7}\,. \tag{5.79}$$

The fact that it appears explicitly in the tracer kernels opens the possibility, at least in principle, to extract cosmological information in a model independent way.

5.3.2 Relation with other bias expansions

Here we compare the present approach to other bias expansions (see [142, 41] for a review on bias). For instance, we can compare our general kernel expansion up to third order in Eqs. (5.76), (5.77) and (5.78), with the bias expansion given in [143]. Up to third order in PT, the density contrast for a given biased tracer in configuration space is expressed as the sum of 7 independent operators,

$$\delta_t = b_1\,\delta + \frac{b_2}{2}\,\delta^2 + \frac{b_3}{3!}\,\delta^3 + b_{\mathcal{G}_2}\,\mathcal{G}_2(\Phi) + b_{\mathcal{G}_3}\,\mathcal{G}_3(\Phi) + b_{\delta\mathcal{G}_2}\,\delta\,\mathcal{G}_2(\Phi) + b_{\mathcal{G}_N}\,\mathcal{G}_N(\varphi_2,\varphi_1)\ , \quad (5.80)$$

where we have omitted the time dependence and we have defined the two Galilean-invariant combinations

$$\begin{aligned}\mathcal{G}_2(\Phi) &\equiv (\nabla_{ij}\Phi)^2 - \left(\nabla^2\Phi\right)^2\ ,\\ \mathcal{G}_3(\Phi) &\equiv \left(\nabla^2\Phi\right)^3 + 2\nabla_{ij}\Phi\nabla_{jk}\Phi\nabla_{ki}\Phi - 3\,(\nabla_{ij}\Phi)^2\,\nabla^2\Phi\,,\end{aligned} \quad (5.81)$$

where Φ is the Poisson potential normalized in such a way that $\nabla^2\Phi = \delta$, and the Galilean-invariant "non-local" combination

$$\mathcal{G}_N(\varphi_2,\varphi_1) \equiv \nabla_{ij}\varphi_2\nabla_{ij}\varphi_1 - \nabla^2\varphi_2\nabla^2\varphi_1\,, \quad (5.82)$$

with $\nabla^2\varphi_1 \equiv -\delta$ and $\nabla^2\varphi_2 \equiv -\mathcal{G}_2(\Phi)$.

The number of independent bias coefficients is the same as ours (cf. with eq. (5.75)):

Our basis: $1^{\rm st}$ order: $c_0^{(1)}$, $2^{\rm nd}$ order: $c_0^{(2)}$, $c_0^{(2)}$, $3^{\rm rd}$ order: $c_0^{(3)}$, $c_1^{(3)}$, $c_5^{(3)}$, $c_{10}^{(3)}$,

Ref. [143]: $1^{\rm st}$ order: b_1 , $2^{\rm nd}$ order: b_2, $b_{\mathcal{G}_2}$, $3^{\rm rd}$ order: b_3, $b_{\mathcal{G}_3}$, $b_{\delta\mathcal{G}_2}$, $b_{\mathcal{G}_N}$.

Indeed, by expanding the RHS of eq. (5.80) and equating the coefficients of the independent operators, we can relate them as

$$\begin{aligned}&c_0^{(1)} = b_1, \qquad c_0^{(2)} = b_2, \qquad c_0^{(3)} = b_3,\\ &c_1^{(2)} = b_1\,a_1^{(2)} - 2\,b_{\mathcal{G}_2}\,, \qquad c_1^{(3)} = b_2\,a_1^{(2)} - 2\,b_{\delta\mathcal{G}_2},\\ &c_5^{(3)} = b_1\,a_5^{(3)} - 2\,b_{\mathcal{G}_2}a_1^{(2)} + 2\,b_{\mathcal{G}_3} + 2\,b_{\mathcal{G}_N}, \qquad c_{10}^{(3)} = b_1\,a_{10}^{(3)} - \ b_{\mathcal{G}_3}\ ,\end{aligned} \quad (5.83)$$

or, inversely,

$$
\begin{aligned}
&b_1 = c_0^{(1)}, \qquad b_2 = c_0^{(2)}, \qquad b_3 = c_0^{(3)}, \\
&b_{\mathcal{G}_2} = \frac{1}{2}\left(a_1^{(2)} c_0^{(1)} - c_1^{(2)}\right), \qquad b_{\delta\mathcal{G}_2} = \frac{1}{2}\left(a_1^{(2)} c_0^{(2)} - c_1^{(3)}\right), \\
&b_{\mathcal{G}_3} = a_{10}^{(3)} c_0^{(1)} - c_{10}^{(3)}, \qquad b_{\mathcal{G}_N} = \frac{1}{2}\left(\left(a_1^{(2)}\right)^2 c_0^{(1)} - a_1^{(2)} c_1^{(2)} - a_5^{(3)} c_0^{(1)} + c_5^{(3)}\right) - a_{10}^{(3)} c_0^{(1)} + c_{10}^{(3)} .
\end{aligned}
\tag{5.84}
$$

Other basis expansions at this order have been given, for instance, in [144, 145, 146, 147, 148, 138, 149]. A comparison shows that, for a fixed cosmology, that is, for fixed $a_i^{(2,3)}$ coefficients, our basis can be related in a similar way as above to those in these references, (see also [150, 75] for explicit relations between coefficients).

For ΛCDM cosmology, our bias expansion can also be compared to the one presented in [75]. Notice that the time-dependent function Y defined in that reference, which carries information about the exact time dependence away from the EdS case, is related to our function h defined in eq. (5.56). More precisely we have: $Y = h/2 - 3/14$. However, our definition in eq. (5.56) is not restricted to ΛCDM. It applies to any cosmological model sharing the same symmetries as ΛCDM. Moreover, our derivation clarifies the physical origin of the tracer-independent function h, that is, EGI.

5.3.3 Relation with Fujita & Vlah

Before closing this section, we discuss in some detail the relation of the present approach to that of [150], which is the closest one to ours. The starting point, also in that paper, is to write down the most general kernels and then to reduce the number of independent coefficients by imposing symmetry-related constraints. In particular, the authors impose that the correlators satisfy the equal-time consistency relations [4, 5]. On top of that, they also impose extra conditions on the momentum dependence of the soft limit of the four-point function that do not derive directly from the equal time consistency relations. At third order and for a generic tracer they find 7 bias coefficients and 2 'universal' ones, which are independent on the tracer type and depend on the underlying cosmology only. We have shown that EGI at LO fixes one of these universal coefficients and at NLO relates the other to the $h(\eta)$ quantity defined in (5.56).

Let us see this in detail. Ref. [150] implements the requirement that the soft limit of the equal time mixed three point function

$$\lim_{q\to 0}\langle \delta_{\mathbf{q}}(\eta)\delta_{t_1,\mathbf{k}_+}(\eta)\delta_{t_2,\mathbf{k}_-}(\eta)\rangle\,, \tag{5.85}$$

where $\mathbf{k}_\pm \equiv \mathbf{k}\pm\mathbf{q}/2$, and '$t_1$' and '$t_2$' indicate two different tracers, has no poles in q. Eqs. (5.76) and (D.1) give,

$$\lim_{q\to 0}\langle \delta_{\mathbf{q}}(\eta)\delta_{t_1,\mathbf{k}_+}(\eta)\delta_{t_2,\mathbf{k}_-}(\eta)\rangle = -\frac{1}{2}\frac{\mathbf{k}\cdot\mathbf{q}}{q^2}\left(c^{(2)}_{2,t_1}(\eta)c^{(1)}_{0,t_2}(\eta) - c^{(2)}_{2,t_2}(\eta)c^{(1)}_{0,t_1}(\eta)\right) P_l(q;\eta)P_l(k;\eta)\,, \tag{5.86}$$

where $P_l(q;\eta)$ is the linear power spectrum. Requiring the absence of poles one finds

$$\frac{c^{(2)}_{2,t_1}(\eta)}{c^{(1)}_{0,t_1}(\eta)} = \frac{c^{(2)}_{2,t_2}(\eta)}{c^{(1)}_{0,t_2}(\eta)} \equiv 2\,\mathcal{C}_b(\eta)\,, \tag{5.87}$$

where the quantity $\mathcal{C}_b(\eta)$, the first 'universal' coefficient defined in ref. [150], does not depend on the tracer type. The authors also observe that in the EdS case $\mathcal{C}_b(\eta) = 1$. From eq. (D.4), this result is completely general, i.e., not limited to EdS. It is a consequence of the EGI of the system.

This can also be seen by extending the approach of [150] by using unequal-time consistency relation, which reads [4],

$$\lim_{q\to 0}\langle \delta_{\mathbf{q}}(\eta)\delta_{t_1,\mathbf{k}_+}(\eta_1)\delta_{t_2,\mathbf{k}_-}(\eta_2)\rangle = -\frac{\mathbf{k}\cdot\mathbf{q}}{q^2}\left(\frac{D_+(\eta_2)}{D_+(\eta)} - \frac{D_+(\eta_1)}{D_+(\eta)}\right) P_l(q;\eta)P_{t_1,t_2}(k;\eta_1,\eta_2)\,, \tag{5.88}$$

where $P_{t_1,t_2}(k;\eta_1,\eta_2)$ is the unequal time cross PS between the tracers t_1 and t_2. Using again eqs. (5.76) and (D.1) we get

$$\begin{aligned}\lim_{q\to 0}\langle \delta_{\mathbf{q}}(\eta)\delta_{t_1,\mathbf{k}_+}(\eta_1)\delta_{t_2,\mathbf{k}_-}(\eta_2)\rangle = \\ -\frac{1}{2}\frac{\mathbf{k}\cdot\mathbf{q}}{q^2}\left(\frac{D_+(\eta_2)}{D_+(\eta)}\frac{c^{(2)}_{2,t_2}(\eta_2)}{c^{(1)}_{0,t_2}(\eta_2)} - \frac{D_+(\eta_1)}{D_+(\eta)}\frac{c^{(2)}_{2,t_1}(\eta_1)}{c^{(1)}_{0,t_1}(\eta_1)}\right) P_l(q;\eta)P_{t_1,t_2}(k;\eta_1,\eta_2)\,,\end{aligned} \tag{5.89}$$

which reproduces the consistency relation (5.88) if

$$\frac{c^{(2)}_{2,t_1}(\eta)}{c^{(1)}_{0,t_1}(\eta)} = \frac{c^{(2)}_{2,t_2}(\eta)}{c^{(1)}_{0,t_2}(\eta)} = 2\,, \tag{5.90}$$

thus confirming $\mathcal{C}_b = 1$ (see eq. (5.87)) in any cosmology respecting the equivalence principle.

On the other hand, the other universal coefficient, $\mathcal{C}_d$, is indeed cosmology-dependent. By using the NLO condition (5.35), on the third order kernel in eq. (C.1) of [150], one can verify that

$$\mathcal{C}_d(\eta) = 2\, h(\eta)\,, \tag{5.91}$$

where the non-local in time, tracer-independent coefficient $h(\eta)$ is given in (5.56).

5.4 UV effects

Up to this point, we have implicitly assumed that PT is able to model the nonlinear behavior at all scales. As it is well known, this assumption fails at small scales, for wavenumbers larger than some value $k_{\rm NL}$. More precisely, the effect of "UV" modes with $q > k_{\rm NL}$ on the nonlinear field of a given tracer evaluated at $k < k_{\rm NL}$ is poorly reproduced by PT, see for instance [151]. In fact, the physics at small scales is unknown and its effects need to be taken into account as an expansion in $k/k_{\rm NL}$ [38].

In order to keep track of the UV corrections needed to correct PT up to a given order, we perform the shift

$$\varphi_{\mathbf{q}}(\eta) \to \varphi_{\mathbf{q}}(\eta) + \delta\varphi_{\mathbf{q}}(\eta)\,, \tag{5.92}$$

in the expansions (5.7), (5.8), or (5.9), where

$$\delta\varphi_{\mathbf{q}}(\eta) \equiv (\tilde{\varphi}_{\mathbf{q}}(\eta) - \varphi_{\mathbf{q}}(\eta))\,\theta(q-\Lambda) \equiv \tilde{\varphi}^{\rm uv}_{\mathbf{q}}(\eta) - \varphi^{\rm uv}_{\mathbf{q}}(\eta)\,. \tag{5.93}$$

$\tilde{\varphi}^{\rm uv}_{\mathbf{q}}(\eta)$ is the 'true' UV field which replaces the "wrong" one, $\varphi^{\rm uv}_{\mathbf{q}}(\eta)$ so that, once inserted in the loop corrections, the wrong PT behavior in the UV is replaced by the corrected one. In the following, we will assume the hierarchy $\Lambda \gg k_{\rm NL} \gg k$, where k is the momentum associated to the nonlinear field. In this limit, the UV effects can be expressed as powers of $k^2/k^2_{\rm NL}$, and the Λ dependence can be omitted[3], as we did in (5.92) and in the last of (5.93).

[3]The regularization scale Λ is non-physical. Therefore the fully renormalized PT observables should not depend on it.

Under the shift (5.92) the nonlinear expansion of the tracer field, eq. (5.9), gets extra contributions,

$$\begin{aligned}\delta^{(n)}_{t,\mathbf{k}}(\eta) &= \mathcal{I}_{\mathbf{k};\mathbf{q}_1\cdots,\mathbf{q}_n} K_n(\mathbf{q}_1,\cdots,\mathbf{q}_n;\eta)\varphi_{\mathbf{q}_1}(\eta)\cdots\varphi_{\mathbf{q}_n}(\eta),\\ &\to \sum_{m=0}^{n}\left(\begin{array}{c} n \\ m \end{array}\right)\mathcal{I}_{\mathbf{k};\mathbf{q}_1\cdots,\mathbf{q}_n} K_n(\mathbf{q}_1,\cdots,\mathbf{q}_n;\eta)\delta\varphi_{\mathbf{q}_1}(\eta)\cdots\delta\varphi_{\mathbf{q}_m}(\eta)\varphi_{\mathbf{q}_{m+1}}(\eta)\cdots\varphi_{\mathbf{q}_n}(\eta)\,,\end{aligned} \tag{5.94}$$

where the $m>0$ terms correct the wrong UV behavior of the $m=0$ one. We now discuss the effect of these corrections up to $n=3$. For $n=1$ we have no modification,

$$\delta^{(1)}_{t,\mathbf{k}}(\eta) \to K_1(\mathbf{k};\eta)\left(\varphi_{\mathbf{k}}(\eta)+\delta\varphi_{\mathbf{k}}(\eta)\right) = K_1(\mathbf{k};\eta)\varphi_{\mathbf{k}}(\eta)\,, \tag{5.95}$$

as, within the assumed hierarchy of scales, the UV field $\delta\varphi_{\mathbf{k}}(\eta)$ has no support on the IR momentum k. For $n=2$,

$$\begin{aligned}\delta^{(2)}_{t,\mathbf{k}}(\eta) \to&\mathcal{I}_{\mathbf{k};\mathbf{q}_1,\mathbf{q}_2} K_2(\mathbf{q}_1,\mathbf{q}_2;\eta)(\varphi_{\mathbf{q}_1}(\eta)+\delta\varphi_{\mathbf{q}_1}(\eta))(\varphi_{\mathbf{q}_2}(\eta)+\delta\varphi_{\mathbf{q}_2}(\eta))\\ =&\mathcal{I}_{\mathbf{k};\mathbf{q}_1,\mathbf{q}_2} K_2(\mathbf{q}_1,\mathbf{q}_2;\eta)\varphi_{\mathbf{q}_1}(\eta)\varphi_{\mathbf{q}_2}(\eta)\\ &+\mathcal{I}_{\mathbf{k};\mathbf{q}_1,\mathbf{q}_2} K_2(\mathbf{q}_1,\mathbf{q}_2;\eta)\left(\tilde{\varphi}^{\rm uv}_{\mathbf{q}_1}(\eta)\tilde{\varphi}^{\rm uv}_{\mathbf{q}_2}(\eta)-\varphi^{\rm uv}_{\mathbf{q}_1}(\eta)\varphi^{\rm uv}_{\mathbf{q}_2}(\eta)\right)\,,\end{aligned} \tag{5.96}$$

where in the last equation we have again made use of the hierarchy of scales. the new term contains only UV fields, and, for the matter and velocity fields, in the small k limit, it scales as $k^2/k^2_{\rm NL}$, due to momentum conservation. The full correction to the second order field will then have the form,

$$\mathcal{I}_{\mathbf{k};\mathbf{q}_1,\mathbf{q}_2} K_2(\mathbf{q}_1,\mathbf{q}_2;\eta)\left(\tilde{\varphi}^{\rm uv}_{\mathbf{q}_1}(\eta)\tilde{\varphi}^{\rm uv}_{\mathbf{q}_2}(\eta)-\varphi^{\rm uv}_{\mathbf{q}_1}(\eta)\varphi^{\rm uv}_{\mathbf{q}_2}(\eta)\right)\xrightarrow{k\to 0}\frac{k^2}{k^2_{\rm NL}}\epsilon^{(2)}_{\delta}(\eta)\,, \tag{5.97}$$

where the zero average stochastic field $\epsilon^{(2)}_{\delta}(\eta)$ will be assumed to be uncorrelated with $\varphi_{\mathbf{q}}(\eta)$.

On the other hand, for a general tracer there is no momentum constraint, and the new term behaves as a scale independent white noise for small k. Therefore, we may write

$$\begin{aligned}\delta^{(2)}_{\mathbf{k}}(\eta) &= \delta^{(2),PT}_{\mathbf{k}}(\eta)+\frac{k^2}{k^2_{\rm NL}}\epsilon^{(2)}_{\delta}(\eta)+O\left(\frac{k^4}{k^4_{\rm NL}}\right)\,,\\ \theta^{(2)}_{\mathbf{k}}(\eta) &= \theta^{(2),PT}_{\mathbf{k}}(\eta)+\frac{k^2}{k^2_{\rm NL}}\epsilon^{(2)}_{\theta}(\eta)+O\left(\frac{k^4}{k^4_{\rm NL}}\right)\,,\\ \delta^{(2)}_{t,\mathbf{k}}(\eta) &= \delta^{(2),PT}_{t,\mathbf{k}}(\eta)+\epsilon^{(2)}_{t}(\eta)+O\left(\frac{k^2}{k^2_{\rm NL}}\right)\,,\end{aligned} \tag{5.98}$$

where by $\delta^{(2),\,PT}_{\mathbf{k}}(\eta)$ and so on we indicate the uncorrected results of the previous sections, and we define the stochastic fields $\epsilon^{(2)}_{\delta,\theta,t}$ from the $k\to 0$ limit of the second line in eq. (5.96).

At $n=3$ we get

$$\begin{aligned}\delta^{(3)}_{t,\mathbf{k}}(\eta) \to & \delta^{(3),\,PT}_{t,\mathbf{k}}(\eta) \\ & + \mathcal{I}_{\mathbf{k};\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}\, K_3(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3;\eta)\Big[3\,\varphi^{\rm ir}_{\mathbf{q}_1}(\eta)\left(\tilde\varphi^{\rm uv}_{\mathbf{q}_2}(\eta)\tilde\varphi^{\rm uv}_{\mathbf{q}_3}(\eta)-\varphi^{\rm uv}_{\mathbf{q}_2}(\eta)\varphi^{\rm uv}_{\mathbf{q}_3}(\eta)\right) \\ & + \left(\tilde\varphi^{\rm uv}_{\mathbf{q}_1}(\eta)\tilde\varphi^{\rm uv}_{\mathbf{q}_2}(\eta)\tilde\varphi^{\rm uv}_{\mathbf{q}_3}(\eta)-\varphi^{\rm uv}_{\mathbf{q}_1}(\eta)\varphi^{\rm uv}_{\mathbf{q}_2}(\eta)\varphi^{\rm uv}_{\mathbf{q}_3}(\eta)\right)\Big]\,, \end{aligned} \tag{5.99}$$

where $\varphi^{\rm ir}_{\mathbf{q}}(\eta)\equiv\varphi_{\mathbf{q}}(\eta)\theta(\Lambda-q)$. The last line gives again noise terms uncorrelated with $\varphi_{\mathbf{q}}(\eta)$ and with the momentum dependences of eq. (5.98). On the other hand, the second line contains the IR component of the linear field. By taking the functional derivative of $\delta^{(3)}_{t,\mathbf{k}}(\eta)$ with respect to $\varphi^{\rm ir}_{\mathbf{k}}(\eta)$, this term gives

$$\frac{\delta\,\delta^{(3)}_{t,\mathbf{k}}(\eta)}{\delta\,\varphi^{\rm ir}_{\mathbf{k}}(\eta)} = \cdots + \frac{3}{(2\pi)^3}\int\frac{d^3q}{(2\pi)^3}K_3(\mathbf{k},\mathbf{q},-\mathbf{q};\eta)\left(\tilde\varphi^{\rm uv}_{\mathbf{q}}(\eta)\tilde\varphi^{\rm uv}_{-\mathbf{q}}(\eta)-\varphi^{\rm uv}_{\mathbf{q}}(\eta)\varphi^{\rm uv}_{-\mathbf{q}}(\eta)\right)\,, \tag{5.100}$$

which, again, vanishes as $k^2/k^2_{\rm NL}$ for matter density and velocity, and goes to a constant for a generic tracer as $k\to 0$. The expectation value of the RHS of (5.100) for matter and velocity corresponds to the 'sound speed' of the EfToLSS [38, 39, 152, 153],

$$c^2_{s,\delta}(\eta)=\lim_{k\to0}\frac{k^2_{\rm NL}}{k^2}\left\langle\frac{\delta\,\delta^{(3)}_{\mathbf{k}}(\eta)}{\delta\,\varphi^{\rm ir}_{\mathbf{k}}(\eta)}\right\rangle\,,\qquad c^2_{s,\theta}(\eta)=\lim_{k\to0}\frac{k^2_{\rm NL}}{k^2}\left\langle\frac{\delta\,\theta^{(3)}_{\mathbf{k}}(\eta)}{\delta\,\varphi^{\rm ir}_{\mathbf{k}}(\eta)}\right\rangle\,, \tag{5.101}$$

while, for a generic tracer it gives both a contribution degenerate with linear bias, at $O(k^0)$, and a sound speed one at $O(k^2)$. In summary, the $n=3$ result of eq. (5.99) can be expressed as

$$\begin{aligned}\delta^{(3)}_{\mathbf{k}}(\eta) &= \delta^{(3),\,PT}_{\mathbf{k}}(\eta)+\frac{k^2}{k^2_{\rm NL}}\left[\left(c^2_{s,\delta}(\eta)+\eta^{(3)}_{\delta}(\eta)\right)\varphi_{\mathbf{k}}(\eta)+\epsilon^{(3)}_{\delta}(\eta)\right]+O\left(\frac{k^4}{k^4_{\rm NL}}\right)\,,\\ \theta^{(3)}_{\mathbf{k}}(\eta) &= \theta^{(3),\,PT}_{\mathbf{k}}(\eta)+\frac{k^2}{k^2_{\rm NL}}\left[\left(c^2_{s,\theta}(\eta)+\eta^{(3)}_{\theta}(\eta)\right)\varphi_{\mathbf{k}}(\eta)+\epsilon^{(3)}_{\theta}(\eta)\right]+O\left(\frac{k^4}{k^4_{\rm NL}}\right)\,,\\ \delta^{(3)}_{t,\mathbf{k}}(\eta) &= \delta^{(3),\,PT}_{t,\mathbf{k}}(\eta)+\left[\tilde b_{0,t}(\eta)+\eta^{(3)}_{t}(\eta)+c^2_{s,t}(\eta)\frac{k^2}{k^2_{\rm NL}}\right]\varphi_{\mathbf{k}}(\eta)+\epsilon^{(3)}_{\theta}(\eta)+O\left(\frac{k^2}{k^2_{\rm NL}}\right)\,,\end{aligned} \tag{5.102}$$

where the stochastic fields $\eta^{(3)}_{\delta,\theta,t}(\eta)$ are given by the fluctuation of eq. (5.99) about its expectation value (multiplied by $k^2_{\rm NL}/k^2$ in the case of δ and θ).

5.5 Conclusions

In this chapter we have investigated the role of symmetries in determining the analytic structure of PT kernels, both for DM and for generic biased tracers. We have highlighted the prominent role played by EGI, and showed that the constraints imposed by this symmetry exhibit a rich structure. The role of EGI was already well appreciated in the literature [150], as it is at the basis of the consistency relations for the LSS [4, 5] (for recent applications see [2, 3]), which relate N-point correlators to $N-m$ ones as m external momenta vanish independently. This limit, and therefore the consistency relations, correspond to the EGI relations at LO in our language. On the other hand, we showed that the limits in which partial sums of the external momenta vanish are also governed by EGI, and provide extra constraints on the functional form of the kernels, as is summarized in eq. (5.37).

The number of independent EGI constraints increases with the PT order. At second order, the LO EGI gives one constraint, while, at third order, LO+NLO EGI provide eight independent conditions. Going to fourth order, one should include also NNLO constraints, and so on. We leave a systematic discussion of PT kernels beyond third order to future work.

Another noticeable feature of eq. (5.37) is its non-locality in time, which is made explicit at third order by the coefficient $h(\eta)$, defined in eq. (5.56). Time non-locality is a well known feature of the bias (and EFT) expansion (for a discussion, see for instance [147, 148]), where it is related to the convective time derivatives appearing in the equations of motion. It is not surprising that, in our discussion, it emerges as a consequence of EGI, which also provides a systematic way to take it into account at higher orders.

The symmetry-based framework defined in this paper allows for a general treatment of DM and biased tracers, as the two species differ only in mass and momentum conservation, which is enforced for the former but not for the latter. It also provides a useful language for using galaxy clustering data for model independent analyses of beyond ΛCDM scenarios, which could be classified in terms of symmetries rather than specific Lagrangians. Considering for instance the galaxy PS in redshift space at one loop, the cosmology-dependent coefficients are $h(\eta)$ and the ones appearing in the velocity kernels, $d_1^{(2)}$, $d_5^{(3)}$, and $d_{10}^{(3)}$. They can either be computed for a given model, as we did for ΛCDM and for nDGP in sect. 5.2.2, or fit from data

in a model independent analysis.

It was shown in [154] that in Horndeski theories EGI symmetry holds and the structure of the PT kernels is the same as that in ΛCDM. Beyond Horndeski (see e.g. [155, 156, 157]), EGI symmetry is broken and PT kernels and the bias expansion deviate from the standard ones, leading for instance to violations of the consistency relations (see [158] for a thorough study of the consequences of this on observables). It would be interesting to systematically study these deviations in a general setting and constrain them with data. A challenging direction is the extension of this approach to modified gravity models involving a scale-dependent growth, such as for instance $f(R)$ or massive neutrinos theories, which we leave for the future.

Chapter 6

Conclusions and outlook

The focus of this thesis was on the role played by symmetries in providing model independent statements about cosmological parameters and PT kernels.

Future surveys, such as Euclid [1, 159], will map the distribution of galaxies over volumes of considerable size and with unprecedent precision: it will represent one of the major sources of cosmological information for the next decade. The main observables that have been studied so far are the power spectrum (PS) and the bispectrum (BS), respectively the Fourier transform of the two- and three-point correlation function. The effective field theory of large scale structure (EFTofLSS, see [36, 37, 38, 39]) provides a rigorous method for the calculation of these correlation functions up to mildly non-linear scales, including corrections coming from small scale non-linearities. In particular, the EFT model for PS at one loop has been recently used for the analysis of the BOSS dataset [11], giving constraints on cosmological parameters that are competitive with CMB based observations, see [45, 46, 47]. The results of these and following analysis on current data (constraints on neutrino masses [160, 92], H_0 tension [161, 162], beyond-ΛCDM model [163, 10, 110], RSD [164, 165]) represent a great step forward in the study of LSS within the EFTofLSS. Moreover, in the BOSS collaboration the BS has been measured and analyses have been performed [48, 49, 50]. The first detection of the BAO features in the BS has been claimed in [63]. On the analytical side, the BS is experiencing growing interest as it can be used to break some degeneracies that are still present when only the PS is considered, for example the one between the linear growth rate f, the large scale bias b_1 and the power spectrum amplitude σ_8. Ref. [60] showed that by performing a joint analysis of the

PS and the BS one could reach a 10% accuracy on f, while [59] has shown that using only the bispectrum monopole reduces significantly the information content of the bispectrum, allowing only for a better estimation of the bias parameters. A first analysis using the EFT model for the tree-level BS has been recently performed on numerical simulations [166], which showed an improvement of 5-15% on the constraints on cosmological parameters.

All these recent results show, on one side, the accuracy of our theoretical modelling of the cosmological observables and, on the other side, the increasing computational power of the numerical techniques used for the analysis.

In the last decade, much interest has been paid to the role of symmetries in cosmology. Refs. [4, 5] derived the so called *consistency relations* (CR) for LSS, which are exact equalities between the N-point correlation function and the $(N-m)$-point one, when m momenta are sent to zero (soft limit). CR's are a direct consequence of the (extended) galilean invariance (EGI) of the system, which means that CR's represent general relations between observables that are model independent, non-perturbative and valid also on non-linear scales.

In [2, 3], we have shown how the CR's provide a reliable way to measure the large scale galaxy bias, b_g, and the linear growth rate, f. The equal-time CR for the BS takes the form, in redshift space,

$$\lim_{q/k\to 0}\frac{B_\alpha(\mathbf{q},-\mathbf{k}_+,\mathbf{k}_-)}{P_\alpha(\mathbf{q})P_\alpha(\mathbf{k})} = -\frac{\mu^2+f\mu\mu_k\mu_q}{b_\alpha+f\mu_q^2}\frac{\partial\ln P_\alpha(\mathbf{k})}{\partial\ln k} + \text{smooth/higher orders contributions}\,. \tag{6.1}$$

where $\mu\equiv\hat{\mathbf{k}}\cdot\hat{\mathbf{q}}$, $\mu_k\equiv\hat{\mathbf{k}}\cdot\hat{\mathbf{z}}$ and $\mu_q\equiv\hat{\mathbf{q}}\cdot\hat{\mathbf{z}}$. In eq. (6.1) the bias parameter b_α is not a parameter of a bias expansion, but it is defined as the large scale limit of the ratio between the real space power spectrum for the tracer α and the α-matter cross correlator

$$b_\alpha \equiv \lim_{q\to 0}\frac{P_{\alpha\alpha}(q)}{P_{\alpha m}(q)}\,. \tag{6.2}$$

Isolating the oscillating part of the two members of eq. (6.1), which is due to the presence of the baryon acoustic oscillations (BAO), it is possible to provide a model independent measurement of the two parameters b_α and f at $\sim 10\%$ precision, depending on the redshift and the tracer considered. These measurements have the same accuracy level of the constraints that comes

from joint analyses of PS and BS, see [60]. While the latter always assumes a cosmological model for the background evolution and a particular perturbative scheme, our method does not rely on these assumptions, as is the result of symmetry principles only. While current analysis with the BS reach very small $k_{\rm max}$ in order to have unbiased estimates of the cosmological parameters (in [77, 166] the authors used $k^{BS}_{\rm max} \simeq 0.08$-$0.09$ h/Mpc), we have shown that our approach can exploit information up to very non-linear scales $k_{\rm max} \simeq 0.25$ h/Mpc in a reliable and accurate way. This method is fast, since it uses only a limited number of triangles, and accurate, since it relies on well established physical motivations, the EGI and the BAO. The measurement of the (true) linear bias could be used as a prior for the MCMC analysis of LSS, making them faster and more reliable.

The analysis we report for the CR's are performed against N-body simulations. A crucial step would be the application of this method to actual observations, for example the BOSS data. Despite the observation of the BAO features in the bispectrum of BOSS, [63], the width of the bins in the publicly available dataset[1] is too high to resolve the BAO in the squeezed configurations of the BS, the one we are interested for the use of eq. (6.1).

This method will be of great use for the analysis of the Euclid future observation. By now, we are testing the use of CR's to measure the linear bias against the Euclid Flagship simulations. These are N-body simulations built with the specifics of the Euclid satellite, the observed number density of tracers, the volume and the redshift range. We have presented preliminary results of this promising analysis, for which the results will be published soon.

By investigating at a more deep level the role played by symmetries on cosmological scales, we formulated the first *LSS bootstrap* [8], presented also in this thesis. We found that symmetries constrain the analytical structure of the PT kernels. The symmetry that plays a prominent role is, again, the EGI, which ensures the invariance of cosmological observables under a general time dependent shift of the coordinates. The constraints imposed by this symmetry present a rich structure. The CR's used in our previous works and presented above, connect correlation functions of different orders when the sum of subset of external momenta are singularly sent to 0. We refer to the constraints that follow from taking single momenta in the soft limit

[1] The PS monopole and quadtrupole and the BS monopole can be found here.

as the leading order (LO). When the sum of two momenta is sent to zero, we have the next-to-leading order (NLO), and so on. Each order of the EGI gives independent constraint at the PT order considered. The number of these independent EGI constraints increases with the PT order. At second order, the leading order (LO) EGI gives one constraint, while, at third order, leading order and next-to-leading order (LO+NLO) EGI provide eight independent conditions. The $\mathrm{N}^{l-1}\mathrm{LO}$ constraint on the n-th order kernel is provided by the general relation

$$\lim_{\mathbf{q}_{1\dots l}\to 0} K_n(\mathbf{q}_1,\dots,\mathbf{q}_l,\mathbf{q}_{l+1},\dots,\mathbf{q}_l) \supset \\ \frac{\mathbf{k}\cdot\mathbf{q}_{1\dots l}}{q^2_{1\dots l}} \int^\eta d\eta' f_+(\eta') \left(\frac{D_+(\eta')}{D_+(\eta)}\right)^l G_l(\mathbf{q}_1,\dots,\mathbf{q}_l;\eta') K_{n-l}(\mathbf{q}_{l+1},\dots,\mathbf{q}_n;\eta)\,, \tag{6.3}$$

where K_n is the n-th order perturbative kernel for a generic tracer, while G_l is the l-th order velocity (divergence) kernel. Going to fourth order, one should include also NNLO constraints, and so on. The EGI constraints are valid for any model and for any tracer that satisfy the symmetries considered. This means that the kernels presented in the PT bootstrap can be used to test a wide variety of models beyond-ΛCDM by performing only one, model independent, MCMC analysis. The EGI is a gold mine that still needs to be deeply explored: one could go to the fourth perturbative order, which is required for the calculation of the BS at one-loop. Another step forward would be the inclusion of scale-dependent models in order to account, for example, for massive neutrinos. We leave the exploration of these generalizations to future work.

The theoretical modelling must always be accompanied by a sufficiently accurate numerical work. In this thesis we have also presented numerical tools based on the FFTLog technique [43, 44] for the fast evaluation of the integrals present in the loop corrections of the PS and the BS. In particular, we have studied a concrete example of modified gravity theory, the nDGP model, see [9, 97], and we have implemented it in PyBird [10], which is a public code currently used for the MCMC analysis of the LSS data[2]. We have shown preliminary results of the constraints on the nDGP parameter from a BOSS-like survey, using the simulation data from the PT challenge [104]. This code is being used in the preliminary tests within the Euclid collaboration and will be used for the future data analysis, that will give us more insight

[2]The public repository of the code can be found here.

on these models. As mentioned before, the inclusion of scale-dependent models would be of great interest for the future observations. A first step in this direction was reported in [167], which performs a PT calculations of the PS in presence of massive neutrinos and provides an accurate approximation for the FFTLog decomposition of the loop integrals. We are currently working on this issue, for which results will be published in the future.

In conclusion, the symmetry based methods for galaxy clustering represent a first step toward a more general approach to the LSS study. Through these methods we can provide model independent constraints on cosmological parameters (using the CR) and general perturbative structure (via the LSS bootstrap) that will help us with a reliable description of the cosmological observables, leading us to a better understanding of the universe we live in. The techniques presented in this thesis will help us in constraining possible deviations from the standard ΛCDM cosmological model. Moreover, the results presented here are easily implemented within fast codes used for the analysis of cosmological observations.

The programmed launch of the Euclid satellite together with the future data release will represent a significant test field for the methods presented in this thesis, and will reveal if the path we have undertaken is as accurate as it seems to promise.

Appendix A

Time-dependent functions for general time dependent cosmologies

The general form for the matter kernels in the approach used in chapter 3 up to third order is

$$K_\lambda^{(1)}(\mathbf{q}_1,\eta) = 1\,, \tag{A.1}$$

$$K_\lambda^{(2)}(\mathbf{q}_1,\mathbf{q}_2,a) = \alpha_s(\mathbf{q}_1,\mathbf{q}_2)\mathcal{G}_1^\lambda(\eta) + \beta(\mathbf{q}_1,\mathbf{q}_2)\mathcal{G}_2^\lambda(\eta)\,, \tag{A.2}$$

$$K_\lambda^{(3)}(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3,\eta) = \alpha^\sigma(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3)\mathcal{U}_\sigma^\lambda(\eta) + \beta^\sigma(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3)\mathcal{V}_{\sigma 2}^\lambda(\eta) + \gamma^\sigma(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3)\mathcal{V}_{\sigma 1}^\lambda(\eta)\,. \tag{A.3}$$

The six function that in eqs. A.1-A.3 are

$$\alpha^1(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) = \alpha(\mathbf{q}_3,\mathbf{q}_1+\mathbf{q}_2)\alpha_s(\mathbf{q}_1,\mathbf{q}_2), \quad \alpha^2(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) = \alpha(\mathbf{q}_3,\mathbf{q}_1+\mathbf{q}_2)\beta(\mathbf{q}_1,\mathbf{q}_2),$$
$$\beta^1(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) = 2\beta(\mathbf{q}_3,\mathbf{q}_1+\mathbf{q}_2)\alpha_s(\mathbf{q}_1,\mathbf{q}_2), \quad \beta^2(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) = 2\beta(\mathbf{q}_3,\mathbf{q}_1+\mathbf{q}_2)\beta(\mathbf{q}_1,\mathbf{q}_2),$$
$$\gamma^1(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) = \alpha(\mathbf{q}_1+\mathbf{q}_2,\mathbf{q}_3)\alpha_s(\mathbf{q}_1,\mathbf{q}_2), \quad \gamma^2(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) = \alpha(\mathbf{q}_1+\mathbf{q}_2,\mathbf{q}_3)\beta(\mathbf{q}_1,\mathbf{q}_2)\,,$$

while the time dependent functions are, in nDGP

$$\mathcal{G}_1^\lambda(\eta) = \int_{-\infty}^0 \left[G_1^\lambda(\eta,\tilde\eta)f_+(\tilde\eta) + G_2^\lambda(a,\tilde\eta)\frac{\mu_2(\tilde\eta)}{f_+(\tilde\eta)}\left(\frac{3\Omega_{\rm m}(\tilde\eta)}{2}\right)^2\right]\frac{D_+^2(\tilde\eta)}{D_+^2(\eta)}e^{-\tilde\eta}d\tilde\eta\,,$$
$$\mathcal{G}_2^\lambda(\eta) = \int_{-\infty}^0 G_2^\lambda(\eta,\tilde\eta)\left[f_+(\tilde\eta) - \frac{\mu_2(\tilde\eta)}{f_+(\tilde\eta)}\left(\frac{3\Omega_{\rm m}(\tilde\eta)}{2}\right)^2\right]\frac{D_+^2(\tilde\eta)}{D_+^2(\eta)}e^{-\tilde\eta}d\tilde\eta\,,$$

$$
\begin{aligned}
\mathcal{U}_1^\lambda(\eta) = \int_{-\infty}^{0} \Bigg\{ & G_1^\lambda(\eta,\tilde{\eta}) f_+(\tilde{\eta}) \mathcal{G}_1^\delta(\tilde{\eta}) \\
& + \frac{G_2^\lambda(a,\tilde{\eta})}{f_+(\tilde{\eta})} \left(\frac{3\Omega_{\rm m}(\tilde{\eta})}{2}\right)^2 \left[\mu_2(\tilde{\eta}) \mathcal{G}_1^\delta(\tilde{\eta}) + \frac{\mu_{22}(\tilde{\eta})}{2} \frac{3\Omega_{\rm m}(\tilde{\eta})}{2}\right] \Bigg\} \frac{D_+^3(\tilde{\eta})}{D_+^3(\eta)} e^{-\tilde{\eta}} d\tilde{\eta}\,, \\
\mathcal{U}_2^\lambda(\eta) = \int_{-\infty}^{0} \Bigg\{ & G_1^\lambda(\eta,\tilde{\eta}) f_+(\tilde{\eta}) \mathcal{G}_2^\delta(\tilde{\eta}) \\
& + \frac{G_2^\lambda(a,\tilde{\eta})}{f_+(\tilde{\eta})} \left(\frac{3\Omega_{\rm m}(\tilde{\eta})}{2}\right)^2 \left[\mu_2(\tilde{\eta}) \mathcal{G}_2^\delta(\tilde{\eta}) - \frac{\mu_{22}(\tilde{\eta})}{2} \frac{3\Omega_{\rm m}(\tilde{\eta})}{2}\right] \Bigg\} \frac{D_+^3(\tilde{\eta})}{D_+^3(\eta)} e^{-\tilde{\eta}} d\tilde{\eta}\,, \\
\mathcal{V}_{11}^\lambda(\eta) = \int_{-\infty}^{0} \Bigg\{ & G_1^\lambda(a,\tilde{\eta}) f_+(\tilde{\eta}) \mathcal{G}_1^\theta(\tilde{\eta}) \\
& + \frac{G_2^\lambda(\eta,\tilde{\eta})}{f_+(\tilde{\eta})} \left(\frac{3\Omega_{\rm m}(\tilde{\eta})}{2}\right)^2 \left[\mu_2(\tilde{\eta}) \mathcal{G}_1^\delta(\tilde{\eta}) + \frac{\mu_{22}(\tilde{\eta})}{2} \frac{3\Omega_{\rm m}(\tilde{\eta})}{2}\right] \Bigg\} \frac{D_+^3(\tilde{\eta})}{D_+^3(\eta)} e^{-\tilde{\eta}} d\tilde{\eta}\,, \\
\mathcal{V}_{21}^\lambda(\eta) = \int_{-\infty}^{0} \Bigg\{ & G_1^\lambda(\eta,\tilde{\eta}) f_+(\tilde{\eta}) \mathcal{G}_2^\theta(\tilde{\eta}) \\
& + \frac{G_2^\lambda(\eta,\tilde{\eta})}{f_+(\tilde{\eta})} \left(\frac{3\Omega_{\rm m}(\tilde{\eta})}{2}\right)^2 \left[\mu_2(\tilde{\eta}) \mathcal{G}_2^\delta(\tilde{\eta}) - \frac{\mu_{22}(\tilde{\eta})}{2} \frac{3\Omega_{\rm m}(\tilde{\eta})}{2}\right] \Bigg\} \frac{D_+^3(\tilde{\eta})}{D_+^3(\eta)} e^{-\tilde{\eta}} d\tilde{\eta}\,, \\
\mathcal{V}_{12}^\lambda(\eta) = \int_{-\infty}^{0} G_2^\lambda(\eta,\tilde{\eta}) \Bigg\{ & f_+(\tilde{\eta}) \mathcal{G}_1^\theta(\tilde{\eta}) \\
& - \frac{1}{f_+(\tilde{\eta})} \left(\frac{3\Omega_{\rm m}(\tilde{\eta})}{2}\right)^2 \left[\mu_2(\tilde{\eta}) \mathcal{G}_1^\delta(\tilde{\eta}) + \frac{\mu_{22}(\tilde{\eta})}{2} \frac{3\Omega_{\rm m}(\tilde{\eta})}{2}\right] \Bigg\} \frac{D_+^3(\tilde{\eta})}{D_+^3(\eta)} e^{-\tilde{\eta}} d\tilde{\eta}\,, \\
\mathcal{V}_{22}^\lambda(\eta) = \int_{-\infty}^{0} G_2^\lambda(a,\tilde{\eta}) \Bigg\{ & f_+(\tilde{\eta}) \mathcal{G}_2^\theta(\tilde{\eta}) \\
& - \frac{1}{f_+(\tilde{\eta})} \left(\frac{3\Omega_{\rm m}(\tilde{\eta})}{2}\right)^2 \left[\mu_2(\tilde{\eta}) \mathcal{G}_2^\delta(\tilde{\eta}) - \frac{\mu_{22}(\tilde{\eta})}{2} \frac{3\Omega_{\rm m}(\tilde{\eta})}{2}\right] \Bigg\} \frac{D_+^3(\tilde{\eta})}{D_+^3(\eta)} e^{-\tilde{\eta}} d\tilde{\eta}\,,
\end{aligned}
\tag{A.4}
$$

where G_σ^λ are the Green's functions, solutions of the equations (3.5). ΛCDM is recovered when $\mu_2 = \mu_{22} = 0$.

Appendix B

Galilean invariance for an arbitrary time dependence

In this Appendix we show that the constraints from EGI do not require to assume a particular time dependence for the displacement defining the velocity shift, as we did in eq. (5.15). Since we are going to have arbitrary time dependence, we have to allow different times for the linear fields in the expansion for $\delta_{t,\mathbf{k}}(\eta)$ in eq. (5.2), namely,

$$\begin{aligned}
\delta_{t,\mathbf{k}}(\eta) =_{p.t.} & \sum_{n=1}^{\infty} \frac{1}{n!} \int \frac{d^3q_1}{(2\pi)^3} \cdots \frac{d^3q_n}{(2\pi)^3} \\
& \times \int_{\eta_{in}}^{\eta} d\eta_1 \cdots \int_{\eta_{in}}^{\eta} d\eta_n \left. \frac{\delta^n \mathcal{F}_t[\varphi^{\lambda}_{\mathbf{q}}(\eta)](\mathbf{k},\eta)}{\delta\varphi^{\lambda_1}_{\mathbf{q}_1}(\eta_1)\cdots\delta\varphi^{\lambda_n}_{\mathbf{q}_n}(\eta_n)} \right|_{\varphi^{\lambda}_{\mathbf{q}}(\eta)=0} \varphi^{\lambda_1}_{\mathbf{q}_1}(\eta_1)\cdots\varphi^{\lambda_n}_{\mathbf{q}_n}(\eta_n) \\
= & \sum_{n=1}^{\infty} \mathcal{I}_{\mathbf{k};\mathbf{q}_1\cdots,\mathbf{q}_n} \int_{\eta_{in}}^{\eta} d\eta_1 \cdots \int_{\eta_{in}}^{\eta} d\eta_n \\
& \times \hat{K}_{n\lambda_1\cdots\lambda_n}(\mathbf{q}_1,\cdots,\mathbf{q}_n;\eta;\eta_1,\cdots,\eta_n) f^{+}_{\mathbf{q}_1}(\eta_1)\cdots f^{+}_{\mathbf{q}_n}(\eta_n)\varphi^{\lambda_1}_{\mathbf{q}_1}(\eta_1)\cdots\varphi^{\lambda_n}_{\mathbf{q}_n}(\eta_n)\,, \\
= & \sum_{n=1}^{\infty} \mathcal{I}_{\mathbf{k};\mathbf{q}_1\cdots,\mathbf{q}_n} K_n(\mathbf{q}_1,\cdots,\mathbf{q}_n;\eta)\varphi_{\mathbf{q}_1}(\eta)\cdots\varphi_{\mathbf{q}_n}(\eta)\,, \qquad \text{(B.1)}
\end{aligned}$$

where now

$$\begin{aligned}
& K_n(\mathbf{q}_1,\cdots,\mathbf{q}_n;\eta) \\
& = \int_{\eta_{in}}^{\eta} d\eta_1 \cdots \int_{\eta_{in}}^{\eta} d\eta_n \hat{K}_{n\lambda_1\cdots\lambda_n}(\mathbf{q}_1,\cdots,\mathbf{q}_n;\eta;\eta_1,\cdots,\eta_n) f^{+}_{\mathbf{q}_1}(\eta_1)\cdots f^{+}_{\mathbf{q}_n}(\eta_n) \\
& \qquad \times \frac{D_{\mathbf{q}_1}(\eta_1)}{D_{\mathbf{q}_1}(\eta)} \cdots \frac{D_{\mathbf{q}_n}(\eta_n)}{D_{\mathbf{q}_n}(\eta)}\, u^{\lambda_1}_{f,\mathbf{q}_1}(\eta_1)\cdots u^{\lambda_n}_{f,\mathbf{q}_n}(\eta_n)\,. \qquad \text{(B.2)}
\end{aligned}$$

Now, let us consider again a shift of the form

$$\varphi^{\lambda_i}_{\mathbf{q}_i}(\eta_i) \to \varphi^{\lambda_i}_{\mathbf{q}_i}(\eta_i) + i\delta_{\lambda 2}(2\pi)^3\delta_D(\mathbf{q}_i)\,\frac{\mathbf{q}_i\cdot\dot{\mathbf{d}}(\eta_i)}{f^+_{\mathbf{q}_i}(\eta_i)\mathcal{H}(\eta_i)}\,, \tag{B.3}$$

but with an arbitrary time dependence for $\mathbf{d}(\eta)$. Now, if the soft limit of the multi-time kernels satisfies

$$\begin{aligned}\lim_{\mathbf{q}_1\cdots\mathbf{q}_m\to 0} & q_1^{i_1}\cdots q_m^{i_m}\hat{K}_{n\lambda_1\cdots\lambda_n}(\mathbf{q}_1,\cdots,\mathbf{q}_n;\eta;\eta_1,\cdots,\eta_n)\\ &= Q^{i_1}_{n,m}\cdots Q^{i_m}_{n,m}\hat{K}_{n-m,\lambda_{m+1}\cdots\lambda_n}(\mathbf{q}_{m+1},\cdots,\mathbf{q}_n;\eta;\eta_{m+1},\cdots,\eta_n)\,,\end{aligned} \tag{B.4}$$

When inserted in eq. (B.1), each shift comes with a time integral of the form

$$\int_{\eta_{in}}^{\eta} d\eta_i\, f^+_{\mathbf{q}_i=0}(\eta_i)\,\frac{\dot{d}(\eta_i)}{f^+_{\mathbf{q}_i=0}(\eta_i)\mathcal{H}(\eta_i)} = d^i(\eta), \tag{B.5}$$

(where we have set $d^i(\eta_{in}) = 0$), which reproduces the $\mathbf{k}\cdot\mathbf{d}(\eta)$ factors in (5.21). Using (B.4) in (B.2), the EGI constraint of eq. (5.27) is recovered.

Appendix C

Next-to-Next-to-Leading order

In what follows we present a calculation of the Next-to-Next-to-Leading order (NNLO) contribution from the EGI. Going to this order implies that we consider a displacement field produced by the coupling of three linear modes, $\varphi_{\mathbf{q}_1}$, $\varphi_{\mathbf{q}_2}$ and $\varphi_{\mathbf{q}_3}$, given by

$$\tilde{\mathbf{d}}^{(3)}_{\mathbf{q}}(\eta) = -i\frac{\mathbf{q}}{q^2}\int^{\eta} d\eta'\, f_+(\eta')\left(\frac{D_+(\eta')}{D_+(\eta)}\right)^3 \mathcal{I}_{\mathbf{q},\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3} G_3(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3;\eta')\varphi_{\mathbf{q}_1}(\eta)\varphi_{\mathbf{q}_2}(\eta)\varphi_{\mathbf{q}_3}(\eta), \tag{C.1}$$

and the zero momentum displacement

$$\mathbf{d}^{(3)}(\eta) = \frac{1}{V}\int_V d^3\mathbf{x}\int \frac{d^3\mathbf{q}}{(2\pi)^3} e^{-i\mathbf{q}\cdot\mathbf{x}}\tilde{\mathbf{d}}^{(3)}_{\mathbf{q}}(\eta). \tag{C.2}$$

The NNLO gives its first contribution to the 4-th order kernel. By expanding eq. (5.20) at fourth order we have

$$\begin{aligned}\delta^{(4)}_{t,\mathbf{k}}(\eta) \to & \delta^{(4)}_{t,\mathbf{k}}(\eta) \\ & + i\mathbf{k}\cdot\mathbf{d}^{(1)}(\eta)\delta^{(3)}_{t,\mathbf{k}}(\eta) + i\left(\mathbf{k}\cdot\mathbf{d}^{(1)}(\eta)\right)^2\delta^{(2)}_{t,\mathbf{k}}(\eta) + i\left(\mathbf{k}\cdot\mathbf{d}^{(1)}(\eta)\right)^3\delta^{(1)}_{t,\mathbf{k}}(\eta) \\ & + i\mathbf{k}\cdot\mathbf{d}^{(2)}(\eta)\delta^{(2)}_{t,\mathbf{k}}(\eta) + i\mathbf{k}\cdot\mathbf{d}^{(1)}(\eta)\,\mathbf{k}\cdot\mathbf{d}^{(2)}(\eta)\delta^{(1)}_{t,\mathbf{k}}(\eta) \\ & + i\mathbf{k}\cdot\mathbf{d}^{(3)}(\eta)\delta^{(1)}_{t,\mathbf{k}}(\eta),\end{aligned} \tag{C.3}$$

where we can explicitly see that in the first line of the rhs eq. (C.3) we have the LO terms which give rise to the relation in eq. (5.27), while the NLO terms are collected in the second line, and give relations similar to eq. (5.35).

In the third line of eq. (C.3) it is shown the first NNLO contribution, which will give insights

on the behavior of the fourth order kernel when the sum of three momenta goes to zero. This contribution is generated by the couplings of three momenta and is given by

$$
\begin{aligned}
i\mathbf{k}\cdot\mathbf{d}^{(3)}(\eta)\delta^{(1)}_{t,\mathbf{k}}(\eta) =& \frac{1}{V}\mathcal{I}_{0,\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}\frac{\mathbf{k}\cdot\mathbf{q}_{123}}{q^2_{123}}\int^{\eta} d\eta'\, f_+(\eta')\left(\frac{D_+(\eta')}{D_+(\eta)}\right)^3 \\
&\times G_3(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3;\eta')K_1(\mathbf{k};\eta)\varphi_{\mathbf{q}_1}(\eta)\varphi_{\mathbf{q}_2}(\eta)\varphi_{\mathbf{q}_3}(\eta)\varphi_{\mathbf{k}}(\eta),
\end{aligned}
\tag{C.4}
$$

and this contribution should be contained in the expression for $\delta^{(4)}_{t,\mathbf{k}}$. To isolate it we adopt a analogous procedure to that presented in Sec (5.1.4), using the identity

$$
1 = \frac{(2\pi)^3}{V}\left[\delta_D(\mathbf{q}_{123}) + \text{cyclic}\right] + 1 - \frac{(2\pi)^3}{V}\left[\delta_D(\mathbf{q}_{123}) + \text{cyclic}\right], \tag{C.5}
$$

so that

$$
\begin{aligned}
\delta^{(4)}_{t,\mathbf{k}} &\supset \mathcal{I}_{\mathbf{k},\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3,\mathbf{q}_4}K_4(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3,\mathbf{q}_4;\eta)\varphi_{\mathbf{q}_1}(\eta)\varphi_{\mathbf{q}_2}(\eta)\varphi_{\mathbf{q}_3}(\eta)\varphi_{\mathbf{q}_4}(\eta)\frac{(2\pi)^3}{V}\delta_D(\mathbf{q}_{123}) \\
&= \frac{1}{V}\mathcal{I}_{0,\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3}K_4(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3,\mathbf{k})\varphi_{\mathbf{q}_1}(\eta)\varphi_{\mathbf{q}_2}(\eta)\varphi_{\mathbf{q}_3}(\eta)\varphi_{\mathbf{k}}(\eta),
\end{aligned}
\tag{C.6}
$$

which should be set equal to the contribution in eq. (C.4). We obtain

$$
\begin{aligned}
&\lim_{\mathbf{q}_{123}\to 0} K_4(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3,\mathbf{q}_4;\eta) \supset \\
&\frac{\mathbf{k}\cdot\mathbf{q}_{123}}{q^2_{123}}\int^{\eta} d\eta' f_+(\eta')\left(\frac{D_+(\eta')}{D_+(\eta)}\right)^3 G_3(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3;\eta')K_1(\mathbf{k};\eta) + O\left(\left(\frac{q_{123}}{k}\right)^2\right).
\end{aligned}
\tag{C.7}
$$

Appendix D

Detailed calculation of the bootstrap kernels

D.1 Case $n = 2$

We start from the most generic form of the kernel,

$$K_2(\mathbf{q}_1, \mathbf{q}_2; \eta) = c_0^{(2)}(\eta) + c_1^{(2)}(\eta)\, \gamma(\mathbf{q}_1, \mathbf{q}_2) + c_2^{(2)}(\eta)\, \beta(\mathbf{q}_1, \mathbf{q}_2)\,. \tag{D.1}$$

Unless it is necessary, to avoid clutter we suppress the explicit dependence on η. Now we impose the constraint eq. (5.27), which implies

$$\lim_{\mathbf{q}_2 \to 0} K_2(\mathbf{q}_1, \mathbf{q}_2) = \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_2^2} K_1(\mathbf{q}_1) = \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_2^2} c_0^{(1)} + \mathcal{O}(q_2^0)\,. \tag{D.2}$$

Taking the same limit in eq. (D.1) we obtain

$$\lim_{\mathbf{q}_2 \to 0} K_2(\mathbf{q}_1, \mathbf{q}_2) = \frac{1}{2} \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_2^2} c_2^{(2)}\,, \tag{D.3}$$

and comparing the two equations give, as anticipated, a relation between the coefficients of the kernels at different order,

$$c_2^{(2)} = 2\, c_0^{(1)}\,. \tag{D.4}$$

Therefore, the most generic kernel at $n = 2$ is given by eq. (5.77).

For conserved tracers, mass and momentum conservation, eq. (5.40), give

$$\begin{aligned} K_2(\mathbf{q}_1, -\mathbf{q}_1) &= c_0^{(2)} = 0\,, \\ \partial_{q_2^i} K_2(\mathbf{q}_1, \mathbf{q}_2)|_{\mathbf{q}_2=-\mathbf{q}_1} &= 0\,, \end{aligned} \tag{D.5}$$

where, due to the properties of the α_s and β functions, the second condition gives no constraint on the coefficients. The most generic matter kernel at $n = 2$ is thus given by eq. (5.46).

D.2 Case $n = 3$

The most general momentum structure for $K_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)$ is given by[1]

$$\begin{aligned} K_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) = \frac{1}{3}c_0^{(3)} &+ c_1^{(3)}\gamma(\mathbf{q}_1, \mathbf{q}_2) + c_2^{(3)}\gamma(\mathbf{q}_{12}, \mathbf{q}_3) + c_3^{(3)}\beta(\mathbf{q}_1, \mathbf{q}_2) + c_4^{(3)}\beta(\mathbf{q}_{12}, \mathbf{q}_3) \\ &+ c_5^{(3)}\gamma(\mathbf{q}_1, \mathbf{q}_2)\gamma(\mathbf{q}_{12}, \mathbf{q}_3) + c_6^{(3)}\beta(\mathbf{q}_1, \mathbf{q}_2)\beta(\mathbf{q}_{12}, \mathbf{q}_3) \\ &+ c_7^{(3)}\gamma(\mathbf{q}_1, \mathbf{q}_2)\beta(\mathbf{q}_{12}, \mathbf{q}_3) + c_8^{(3)}\beta(\mathbf{q}_1, \mathbf{q}_2)\gamma(\mathbf{q}_{12}, \mathbf{q}_3) \\ &+ \left(c_9^{(3)} + c_{10}^{(3)}\gamma(\mathbf{q}_1, \mathbf{q}_2) + c_{11}^{(3)}\beta(\mathbf{q}_1, \mathbf{q}_2)\right)\alpha_a(\mathbf{q}_{12}, \mathbf{q}_3) \\ &+ 2 \text{ permutations}\,. \end{aligned} \tag{D.6}$$

First we apply eq. (5.27), in the single vanishing momentum case,

$$\lim_{q_1\to 0} K_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) = \frac{(\mathbf{q}_2 + \mathbf{q}_3)\cdot\mathbf{q}_1}{q_1^2} K_2(\mathbf{q}_2, \mathbf{q}_3)\,, \tag{D.7}$$

while in the same limit eq. (D.6) gives

$$\begin{aligned} \lim_{q_1\to 0} K_3(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) = \frac{(\mathbf{q}_2 + \mathbf{q}_3)\cdot\mathbf{q}_1}{q_1^2}\Bigg[&\frac{1}{2}\left(c_3^{(3)} + c_4^{(3)}\right) - c_9^{(3)} + \gamma(\mathbf{q}_2, \mathbf{q}_3)\left(\frac{1}{2}\left(c_7^{(3)} + c_8^{(3)}\right) - c_{10}^{(3)}\right) \\ &+ \beta(\mathbf{q}_2, \mathbf{q}_3)\left(c_6^{(3)} - c_{11}^{(3)}\right)\Bigg] + \frac{(\mathbf{q}_2 - \mathbf{q}_3)\cdot\mathbf{q}_1}{q_1^2} c_{11}^{(3)}\,\alpha_a(\mathbf{q}_2, \mathbf{q}_3)\,. \end{aligned} \tag{D.8}$$

Comparing these two equations using the explicit form of K_2, eq. (5.77), gives the following constraints,

$$c_3^{(3)} + c_4^{(3)} - 2\,c_9^{(3)} = 2\,c_0^{(2)}, \qquad c_7^{(3)} + c_8^{(3)} - 2c_{10}^{(3)} = 2c_1^{(2)}\,, \qquad c_6^{(3)} - c_{11}^{(3)} = c_2^{(2)}\,, \qquad c_{11}^{(3)} = 0\,. \tag{D.9}$$

[1] A contribution of the form $(\mathbf{q}_1\cdot\mathbf{q}_2)(\mathbf{q}_2\cdot\mathbf{q}_3)(\mathbf{q}_3\cdot\mathbf{q}_1)/(q_1^2q_2^2q_3^2)$ can be shown to be linearly dependent from those generated by the direct product (5.47), and therefore needs not to be added.

The double limit gives

$$\lim_{q_1,q_2\to 0} K_3(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) = \frac{1}{2}\frac{\mathbf{q}_3\cdot\mathbf{q}_1}{q_1^2}\frac{\mathbf{q}_3\cdot\mathbf{q}_2}{q_2^2}\left(c_6^{(3)}-c_{11}^{(3)}\right) = \frac{\mathbf{q}_3\cdot\mathbf{q}_1}{q_1^2}\frac{\mathbf{q}_3\cdot\mathbf{q}_2}{q_2^2}K_1(\mathbf{q}_3)\,, \tag{D.10}$$

from which we get the (redundant) constraint

$$c_6^{(3)} - c_{11}^{(3)} = 2\,c_0^{(1)}\,. \tag{D.11}$$

Combined with the second and third relations in (D.9), this is equivalent to eq. (D.4).

Imposing eq. (5.35) we have

$$\lim_{q_{12}\to 0} K_3(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) \supset \frac{\mathbf{q}_{12}\cdot\mathbf{q}_3}{q_{12}^2}c_0^{(1)}(\eta)\int^\eta d\eta'\, f_+(\eta')\left[\frac{D_+(\eta')}{D_+(\eta)}\right]^2 G_2(\mathbf{q}_1,\mathbf{q}_2;\eta')\,, \tag{D.12}$$

while eq. (D.6) in the same limit gives

$$\begin{aligned}\lim_{q_{12}\to 0} K_3(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3) \supset & \frac{\mathbf{q}_{12}\cdot\mathbf{q}_3}{q_{12}^2}\left[\frac{1}{2}c_4^{(3)}+c_9^{(3)}+\gamma(\mathbf{q}_1,\mathbf{q}_2)\left(\frac{1}{2}c_7^{(3)}+c_{10}^{(3)}\right)+\beta(\mathbf{q}_1,\mathbf{q}_2)\left(\frac{1}{2}c_6^{(3)}+c_{11}^{(3)}\right)\right] \\ & + \frac{\mathbf{q}_{12}\cdot\mathbf{q}_3}{q_3^2}\left(\frac{1}{2}c_4^{(3)}-c_9^{(3)}\right)+\frac{(\mathbf{q}_{12}\cdot\mathbf{q}_3)^2}{q_{12}^2q_3^2}\left(c_4^{(3)}-c_2^{(3)}\right)\,,\end{aligned} \tag{D.13}$$

Using the structure of the velocity kernel, i.e.,

$$G_2(\mathbf{q}_1,\mathbf{q}_2) = 2\,\beta(\mathbf{q}_1,\mathbf{q}_2) + d_1^{(2)}\,\gamma(\mathbf{q}_1,\mathbf{q}_2)\,, \tag{D.14}$$

gives other five constraints,

$$\begin{aligned}&\frac{1}{2}c_4^{(3)}+c_9^{(3)}=0, \qquad \frac{1}{2}c_6^{(3)}+c_{11}^{(3)}=c_0^{(1)}, \quad \frac{1}{2}c_7^{(3)}+c_{10}^{(3)} = h\,c_0^{(1)}\,, \\ &c_4^{(3)}-c_2^{(3)}=0\,, \qquad \frac{1}{2}c_4^{(3)}-c_9^{(3)}=0\,,\end{aligned} \tag{D.15}$$

where we have defined

$$h(\eta) \equiv \int^\eta d\eta'\, f_+(\eta')\left[\frac{D_+(\eta')}{D_+(\eta)}\right]^2 d_1^{(2)}(\eta')\,. \tag{D.16}$$

Note that the last two conditions in eq. (D.15) come directly from the second line of eq. (D.13): the structure of these two terms is not present in the velocity kernel, see eq. (D.14), due to the momentum conservation for the velocity field, which constraints the kernel G_2 as in eq. (5.40).

Only 8 of the above equations are independent, leading to the following relations

$$c_2^{(3)} = c_4^{(3)} = c_9^{(3)} = c_{11}^{(3)} = 0\,, \qquad c_3^{(3)} = 2\,c_0^{(2)}\,, \qquad c_6^{(3)} = 2\,c_0^{(1)}\,, \tag{D.17}$$

and

$$c_7^{(3)} = -2\,c_{10}^{(3)} + 2\,h\,c_0^{(1)}\,, \qquad c_8^{(3)} = 4\,c_{10}^{(3)} + 2\,c_1^{(2)} - 2h\,c_0^{(1)}\,, \tag{D.18}$$

leaving with four (i.e. $\{c_0^{(3)}, c_1^{(3)}, c_5^{(3)}, c_{10}^{(3)}\}$) free coefficients out of 12. The form of the kernel after imposing these relations is given in eq. (5.78).

In the case of a conserved tracer, we can also impose the two conditions coming from mass and momentum conservation, eq. (5.40), which give respectively,

$$c_0^{(3)} + \frac{3}{2}c_1^{(3)} = 0, \qquad\qquad c_3^{(3)} = 0\,, \tag{D.19}$$

and

$$c_1^{(3)} = c_3^{(3)} = c_9^{(3)} = 0\,, \qquad c_{11}^{(3)} = 0\,, \tag{D.20}$$

allowing us to set

$$c_0^{(3)} = c_1^{(3)} = c_3^{(3)} = 0\,, \tag{D.21}$$

so that only two coefficients are left free. This leads to the matter kernel (5.52).

Appendix E

Time-dependence of the bootstrap kernels

E.1 ΛCDM

In this appendix we give the analytic solutions of the time evolution equations for the coefficients of the third order matter kernels, which are given by

$$
\begin{aligned}
(\partial_\chi + 2) a_5^{(3)}(\chi) &= \left(a_1^{(2)}(\chi) + d_1^{(2)}(\chi)\right) - a_5^{(3)}(\chi) + d_5^{(3)}(\chi)\,, \\
(\partial_\chi + 2) d_5^{(3)}(\chi) &= \frac{3}{2}\frac{\Omega_m(\chi)}{f_+^2(\chi)}\left(a_5^{(3)}(\chi) - d_5^{(3)}(\chi)\right)\,, \\
(\partial_\chi + 2) a_{10}^{(3)}(\chi) &= -\frac{1}{2}\left(a_1^{(2)}(\chi) - d_1^{(2)}(\chi)\right) - a_{10}^{(3)}(\chi) + d_{10}^{(3)}(\chi)\,, \\
(\partial_\chi + 2) d_{10}^{(3)}(\chi) &= \frac{3}{2}\frac{\Omega_m(\chi)}{f_+^2(\chi)}\left(a_{10}^{(3)}(\chi) - a_{10}^{(3)}(\chi)\right)\,.
\end{aligned}
\tag{E.1}
$$

The solutions are given by

$$
\begin{aligned}
a_5^{(3)}(\chi) &= \Delta C_5^{(3)}(\chi) + d_5^{(3)}(\chi) = \Delta a_5^{(3)}(\chi) + d_5^{(3)}(\chi)\,, \\
d_5^{(3)}(\chi) &= e^{-2\chi}\int_{-\infty}^{\chi} d\chi'\, e^{2\chi'}\frac{3}{2}\frac{\Omega_m(\chi')}{f_+^2(\chi')}\Delta a_5^{(3)}(\chi')\,, \\
\Delta a_5^{(3)}(\chi) &= e^{-2\chi}\int_{-\infty}^{\chi} d\chi'\, e^{2\chi' - \int_{\chi'}^{\chi} d\chi''\left(1+\frac{3}{2}\frac{\Omega_m(\chi'')}{f_+^2(\chi'')}\right)} a_+^{(2)}(\chi')\,,
\end{aligned}
\tag{E.2}
$$

where $a_+^{(2)}(\chi) \equiv a_1^{(2)}(\chi) + d_1^{(2)}(\chi)$, can be read from eqs. (5.66), (5.67), and, analogously, by,

$$
\begin{aligned}
a_{10}^{(3)}(\chi) &= \Delta C_{10}^{(3)}(\chi) + d_{10}^{(3)}(\chi) = \Delta a_{10}^{(3)}(\chi) + d_{10}^{(3)}(\chi)\,, \\
d_{10}^{(3)}(\chi) &= e^{-2\chi} \int_{-\infty}^{\chi} d\chi'\, e^{2\chi'}\, \frac{3}{2}\frac{\Omega_m(\chi')}{f_+^2(\chi')} \Delta a_{10}^{(3)}(\chi')\,, \\
\Delta a_{10}^{(3)}(\chi) &= -\frac{e^{-2\chi}}{2} \int_{-\infty}^{\chi} d\chi'\, e^{2\chi' - \int_{\chi'}^{\chi} d\chi'' \left(1+\frac{3}{2}\frac{\Omega_m(\chi'')}{f_+^2(\chi'')}\right)} a_-^{(2)}(\chi')\,,
\end{aligned}
\tag{E.3}
$$

where $a_-^{(2)}(\chi) \equiv a_1^{(2)}(\chi) - d_1^{(2)}(\chi)$.

E.2 nDGP

Here we give the analytic solutions of the time evolution equations for the coefficients of the third order matter kernels in a nDGP cosmology, which are given by

$$
\begin{aligned}
(\partial_\chi + 2) a_5^{(3)}(\chi) &= \left(a_1^{(2)}(\chi) + d_1^{(2)}(\chi)\right) - a_5^{(3)}(\chi) + d_5^{(3)}(\chi) + 2\left(\frac{3}{2}\Omega_m\right)^2 \left(\frac{3}{2}\frac{\Omega_m}{f_+^2}\mu_3 + \frac{\mu_2}{f_+^2} a_1^{(2)}\right)\,, \\
(\partial_\chi + 2) d_5^{(3)}(\chi) &= \frac{3}{2}\frac{\Omega_m(\chi)}{f_+^2(\chi)}\mu(\chi)\left(a_5^{(3)}(\chi) - d_5^{(3)}(\chi)\right) + 2\left(\frac{3}{2}\Omega_m\right)^2 \left(\frac{3}{2}\frac{\Omega_m}{f_+^2}\mu_3 + \frac{\mu_2}{f_+^2} a_1^{(2)}\right)\,, \\
(\partial_\chi + 2) a_{10}^{(3)}(\chi) &= -\frac{1}{2}\left(a_1^{(2)}(\chi) - d_1^{(2)}(\chi)\right) - a_{10}^{(3)}(\chi) + d_{10}^{(3)}(\chi)\,, \\
(\partial_\chi + 2) d_{10}^{(3)}(\chi) &= \frac{3}{2}\frac{\Omega_m(\chi)}{f_+^2(\chi)}\mu(\chi)\left(a_{10}^{(3)}(\chi) - a_{10}^{(3)}(\chi)\right)\,.
\end{aligned}
\tag{E.4}
$$

The solutions are given by

$$
\begin{aligned}
a_5^{(3)}(\chi) &= \Delta a_5^{(3)}(\chi) + d_5^{(3)}(\chi)\,, \\
d_5^{(3)}(\chi) &= e^{-2\chi} \int_{-\infty}^{\chi} d\chi'\, e^{2\chi'} \times \\
&\quad \times \left[\frac{3}{2}\frac{\Omega_m(\chi')}{f_+^2(\chi')}\mu(\chi')\Delta a_5^{(3)}(\chi') + 2\left(\frac{3}{2}\Omega_m(\chi')\right)^2 \left(\frac{3}{2}\frac{\Omega_m(\chi')}{f_+^2(\chi')}\mu_3 + \frac{\mu_2(\chi')}{f_+^2(\chi')} a_1^{(2)}(\chi')\right)\right]\,, \\
\Delta a_5^{(3)}(\chi) &= e^{-2\chi} \int_{-\infty}^{\chi} d\chi'\, e^{2\chi' - \int_{\chi'}^{\chi} d\chi'' \left(1+\frac{3}{2}\frac{\Omega_m(\chi'')}{f_+^2(\chi'')}\mu(\chi'')\right)}\, a_+^{(2)}(\chi')\,,
\end{aligned}
\tag{E.5}
$$

and, analogously, by,

$$
\begin{aligned}
a_{10}^{(3)}(\chi) &= \Delta a_{10}^{(3)}(\chi) + d_{10}^{(3)}(\chi)\,, \\
d_{10}^{(3)}(\chi) &= e^{-2\chi} \int_{-\infty}^{\chi} d\chi'\, e^{2\chi'}\, \frac{3}{2}\frac{\Omega_m(\chi')}{f_+^2(\chi')}\mu(\chi') \Delta a_{10}^{(3)}(\chi')\,, \\
\Delta a_{10}^{(3)}(\chi) &= -\frac{e^{-2\chi}}{2} \int_{-\infty}^{\chi} d\chi'\, e^{2\chi' - \int_{\chi'}^{\chi} d\chi'' \left(1+\frac{3}{2}\frac{\Omega_m(\chi'')}{f_+^2(\chi'')}\mu(\chi'')\right)} a_-^{(2)}(\chi')\,.
\end{aligned}
\tag{E.6}
$$

Bibliography

[1] **EUCLID** Collaboration, R. Laureijs *et. al.*, *Euclid Definition Study Report*, 1110.3193.

[2] M. Marinucci, T. Nishimichi and M. Pietroni, *Measuring Bias via the Consistency Relations of the Large Scale Structure*, *Phys. Rev.* **D100** (2019), no. 12 123537 [1907.09866].

[3] M. Marinucci, T. Nishimichi and M. Pietroni, *Model independent measurement of the growth rate from the consistency relations of the LSS*, *JCAP* **07** (2020) 054 [2005.09574].

[4] M. Peloso and M. Pietroni, *Galilean invariance and the consistency relation for the nonlinear squeezed bispectrum of large scale structure*, *JCAP* **05** (2013) 031 [1302.0223].

[5] A. Kehagias and A. Riotto, *Symmetries and Consistency Relations in the Large Scale Structure of the Universe*, *Nucl.Phys.* **B873** (2013) 514–529 [1302.0130].

[6] P. Creminelli, J. Noreña, M. Simonović and F. Vernizzi, *Single-Field Consistency Relations of Large Scale Structure*, *JCAP* **1312** (2013) 025 [1309.3557].

[7] P. Valageas, A. Taruya and T. Nishimichi, *Consistency relations for large scale structures with primordial non-Gaussianities*, *Phys. Rev. D* **95** (2017), no. 2 023504 [1610.00993].

[8] G. D'Amico, M. Marinucci, M. Pietroni and F. Vernizzi, *The large scale structure bootstrap: perturbation theory and bias expansion from symmetries*, *JCAP* **10** (2021) 069 [2109.09573].

[9] G. R. Dvali and G. Gabadadze, *Gravity on a brane in infinite volume extra space*, *Phys. Rev. D* **63** (2001) 065007 [hep-th/0008054].

[10] G. D'Amico, L. Senatore and P. Zhang, *Limits on wCDM from the EFTofLSS with the PyBird code*, *JCAP* **01** (2021) 006 [2003.07956].

[11] **BOSS** Collaboration, S. Alam *et. al.*, *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample*, *Mon. Not. Roy. Astron. Soc.* **470** (2017), no. 3 2617–2652 [1607.03155].

[12] **LSST** Collaboration, v. Ivezić *et. al.*, *LSST: from Science Drivers to Reference Design and Anticipated Data Products*, *Astrophys. J.* **873** (2019), no. 2 111 [0805.2366].

[13] **DESI** Collaboration, A. Aghamousa *et. al.*, *The DESI Experiment Part I: Science,Targeting, and Survey Design*, 1611.00036.

[14] **Planck** Collaboration, N. Aghanim *et. al.*, *Planck 2018 results. I. Overview and the cosmological legacy of Planck*, *Astron. Astrophys.* **641** (2020) A1 [1807.06205].

[15] **Planck** Collaboration, N. Aghanim *et. al.*, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [1807.06209]. [Erratum: Astron.Astrophys. 652, C4 (2021)].

[16] **Supernova Cosmology Project** Collaboration, S. Perlmutter *et. al.*, *Measurements of Ω and Λ from 42 high redshift supernovae*, *Astrophys. J.* **517** (1999) 565–586 [astro-ph/9812133].

[17] **Supernova Search Team** Collaboration, A. G. Riess *et. al.*, *Observational evidence from supernovae for an accelerating universe and a cosmological constant*, *Astron. J.* **116** (1998) 1009–1038 [astro-ph/9805201].

[18] **SDSS** Collaboration, D. J. Eisenstein *et. al.*, *Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies*, *Astrophys. J.* **633** (2005) 560–574 [astro-ph/0501171].

[19] E. Gaztanaga, A. Cabre and L. Hui, *Clustering of Luminous Red Galaxies IV: Baryon Acoustic Peak in the Line-of-Sight Direction and a Direct Measurement of H(z)*, *Mon. Not. Roy. Astron. Soc.* **399** (2009) 1663–1680 [0807.3551].

[20] W. J. Percival *et. al.*, *Measuring the matter density using baryon oscillations in the SDSS*, *Astrophys. J.* **657** (2007) 51–55 [astro-ph/0608635].

[21] **SDSS** Collaboration, M. Tegmark *et. al.*, *Cosmological Constraints from the SDSS Luminous Red Galaxies*, *Phys. Rev. D* **74** (2006) 123507 [astro-ph/0608632].

[22] P. J. E. Peebles and J. T. Yu, *Primeval adiabatic perturbation in an expanding universe*, *Astrophys. J.* **162** (1970) 815–836.

[23] R. A. Sunyaev and Y. B. Zeldovich, *Small scale fluctuations of relic radiation*, *Astrophys. Space Sci.* **7** (1970) 3–19.

[24] J. R. Bond and G. Efstathiou, *Cosmic background radiation anisotropies in universes dominated by nonbaryonic dark matter*, *Astrophys. J. Lett.* **285** (1984) L45–L48.

[25] J. A. Holtzman, *Microwave background anisotropies and large scale structure in universes with cold dark matter, baryons, radiation, and massive and massless neutrinos*, *Astrophys. J. Suppl.* **71** (1989) 1–24.

[26] W. J. Percival, S. Cole, D. J. Eisenstein, R. C. Nichol, J. A. Peacock, A. C. Pope and A. S. Szalay, *Measuring the Baryon Acoustic Oscillation scale using the SDSS and 2dFGRS*, *Mon. Not. Roy. Astron. Soc.* **381** (2007) 1053–1066 [0705.3323].

[27] F. Bernardeau, S. Colombi, E. Gaztanaga and R. Scoccimarro, *Large scale structure of the universe and cosmological perturbation theory*, *Phys. Rept.* **367** (2002) 1–248 [astro-ph/0112551].

[28] D. Blas, M. Garny and T. Konstandin, *Cosmological perturbation theory at three-loop order*, *JCAP* **1401** (2014), no. 01 010 [1309.3308].

[29] A. Taruya and S. Colombi, *Post-collapse perturbation theory in 1D cosmology – beyond shell-crossing*, *Mon. Not. Roy. Astron. Soc.* **470** (2017), no. 4 4858–4884 [1701.09088].

[30] C. Rampf and U. Frisch, *Shell-crossing in quasi-one-dimensional flow*, *Mon. Not. Roy. Astron. Soc.* **471** (2017), no. 1 671–679 [1705.08456].

[31] P. McDonald and Z. Vlah, *Large-scale structure perturbation theory without losing stream crossing*, *Phys. Rev.* **D97** (2018), no. 2 023508 [1709.02834].

[32] S. Saga, A. Taruya and S. Colombi, *Lagrangian cosmological perturbation theory at shell-crossing*, *Phys. Rev. Lett.* **121** (2018), no. 24 241302 [1805.08787].

[33] A. Halle, T. Nishimichi, A. Taruya, S. Colombi and F. Bernardeau, *Power spectrum response of large-scale structure in 1D and in 3D: tests of prescriptions for post-collapse dynamics*, 2001.10417.

[34] M. Pietroni, *Structure formation beyond shell-crossing: nonperturbative expansions and late-time attractors*, *JCAP* **1806** (2018), no. 06 028 [1804.09140].

[35] S.-F. Chen and M. Pietroni, *Asymptotic expansions for Large Scale Structure*, *JCAP* **06** (2020) 033 [2002.11357].

[36] D. Baumann, A. Nicolis, L. Senatore and M. Zaldarriaga, *Cosmological Non-Linearities as an Effective Fluid*, *JCAP* **1207** (2012) 051 [1004.2488].

[37] M. Pietroni, G. Mangano, N. Saviano and M. Viel, *Coarse-Grained Cosmological Perturbation Theory*, *JCAP* **1201** (2012) 019 [1108.5203].

[38] J. J. M. Carrasco, M. P. Hertzberg and L. Senatore, *The Effective Field Theory of Cosmological Large Scale Structures*, *JHEP* **1209** (2012) 082 [1206.2926].

[39] A. Manzotti, M. Peloso, M. Pietroni, M. Viel and F. Villaescusa-Navarro, *A coarse grained perturbation theory for the Large Scale Structure, with cosmology and time independence in the UV*, *JCAP* **1409** (2014), no. 09 047 [1407.1342].

[40] P. McDonald and A. Roy, *Clustering of dark matter tracers: generalizing bias for the coming era of precision LSS*, *JCAP* **08** (2009) 020 [0902.0991].

[41] V. Desjacques, D. Jeong and F. Schmidt, *Large-Scale Galaxy Bias*, *Phys. Rept.* **733** (2018) 1–193 [1611.09787].

[42] M. Schmittfull, Z. Vlah and P. McDonald, *Fast large scale structure perturbation theory using one-dimensional fast Fourier transforms*, *Phys. Rev. D* **93** (2016), no. 10 103528 [1603.04405].

[43] J. E. McEwen, X. Fang, C. M. Hirata and J. A. Blazek, *FAST-PT: a novel algorithm to calculate convolution integrals in cosmological perturbation theory*, *JCAP* **09** (2016) 015 [1603.04826].

[44] 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* **1804** (2018), no. 04 030 [1708.08130].

[45] 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** (2020) 005 [1909.05271].

[46] M. M. Ivanov, M. Simonović and M. Zaldarriaga, *Cosmological Parameters from the BOSS Galaxy Power Spectrum*, *JCAP* **05** (2020) 042 [1909.05277].

[47] S.-F. Chen, Z. Vlah and M. White, *A new analysis of the BOSS survey, including full-shape information and post-reconstruction BAO*, 2110.05530.

[48] H. Gil-Marín, J. Noreña, L. Verde, W. J. Percival, C. Wagner, M. Manera and D. P. Schneider, *The power spectrum and bispectrum of SDSS DR11 BOSS galaxies – I. Bias and gravity*, *Mon. Not. Roy. Astron. Soc.* **451** (2015), no. 1 539–580 [1407.5668].

[49] H. Gil-Marín, L. Verde, J. Noreña, A. J. Cuesta, L. Samushia, W. J. Percival, C. Wagner, M. Manera and D. P. Schneider, *The power spectrum and bispectrum of SDSS DR11 BOSS*

galaxies – II. Cosmological interpretation, *Mon. Not. Roy. Astron. Soc.* **452** (2015), no. 2 1914–1921 [1408.0027].

[50] H. Gil-Marín, W. J. Percival, L. Verde, J. R. Brownstein, C.-H. Chuang, F.-S. Kitaura, S. A. Rodríguez-Torres and M. D. Olmstead, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: RSD measurement from the power spectrum and bispectrum of the DR12 BOSS galaxies*, *Mon. Not. Roy. Astron. Soc.* **465** (2017), no. 2 1757–1788 [1606.00439].

[51] D. Simmons-Duffin, *The Conformal Bootstrap*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in Fields and Strings*, 2, 2016. 1602.07982.

[52] N. Arkani-Hamed, D. Baumann, H. Lee and G. L. Pimentel, *The Cosmological Bootstrap: Inflationary Correlators from Symmetries and Singularities*, *JHEP* **04** (2020) 105 [1811.00024].

[53] E. Pajer, D. Stefanyszyn and J. Supeł, *The Boostless Bootstrap: Amplitudes without Lorentz boosts*, *JHEP* **12** (2020) 198 [2007.00027].

[54] R. Scoccimarro and J. Frieman, *Loop corrections in nonlinear cosmological perturbation theory*, *Astrophys.J.Suppl.* **105** (1996) 37 [astro-ph/9509047].

[55] B. Jain and E. Bertschinger, *Selfsimilar evolution of cosmological density fluctuations*, *Astrophys.J.* **456** (1996) 43 [astro-ph/9503025].

[56] M. Pietroni, *Flowing with Time: a New Approach to Nonlinear Cosmological Perturbations*, *JCAP* **0810** (2008) 036 [0806.0971].

[57] V. Desjacques, D. Jeong and F. Schmidt, *The Galaxy Power Spectrum and Bispectrum in Redshift Space*, *JCAP* **12** (2018) 035 [1806.04015].

[58] D. Gualdi and L. Verde, *Galaxy redshift-space bispectrum: the Importance of Being Anisotropic*, *JCAP* **06** (2020) 041 [2003.12075].

[59] V. Yankelevich and C. Porciani, *Cosmological information in the redshift-space bispectrum*, *Mon. Not. Roy. Astron. Soc.* **483** (2019), no. 2 2078–2099 [1807.07076].

[60] N. Agarwal, V. Desjacques, D. Jeong and F. Schmidt, *Information content in the redshift-space galaxy power spectrum and bispectrum*, *JCAP* **03** (2021) 021 [2007.04340].

[61] 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** (2020) 056 [1908.01774].

[62] P. Gagrani and L. Samushia, *Information Content of the Angular Multipoles of Redshift-Space Galaxy Bispectrum*, *Mon. Not. Roy. Astron. Soc.* **467** (2017), no. 1 928–935 [1610.03488].

[63] D. W. Pearson and L. Samushia, *A Detection of the Baryon Acoustic Oscillation features in the SDSS BOSS DR12 Galaxy Bispectrum*, *Mon. Not. Roy. Astron. Soc.* **478** (2018), no. 4 4500–4512 [1712.04970].

[64] D. Gualdi, H. Gil-Marín, R. L. Schuhmann, M. Manera, B. Joachimi and O. Lahav, *Enhancing BOSS bispectrum cosmological constraints with maximal compression*, *Mon. Not. Roy. Astron. Soc.* **484** (2019), no. 3 3713–3730 [1806.02853].

[65] H. L. Child, M. Takada, T. Nishimichi, T. Sunayama, Z. Slepian, S. Habib and K. Heitmann, *Bispectrum as Baryon Acoustic Oscillation Interferometer*, *Phys. Rev.* **D98** (2018) 123521 [1806.11147].

[66] H. L. Child, Z. Slepian and M. Takada, *A Physical Picture of Bispectrum Baryon Acoustic Oscillations in the Interferometric Basis*, 1811.12396.

[67] J. Byun, A. Oddo, C. Porciani and E. Sefusatti, *Towards cosmological constraints from the compressed modal bispectrum: a robust comparison of real-space bispectrum estimators*, *JCAP* **03** (2021) 105 [2010.09579].

[68] **Planck** Collaboration, P. Ade *et. al.*, *Planck 2015 results. XIII. Cosmological parameters*, 1502.01589.

[69] L. Senatore and M. Zaldarriaga, *Redshift Space Distortions in the Effective Field Theory of Large Scale Structures*, 1409.1225.

[70] A. Perko, L. Senatore, E. Jennings and R. H. Wechsler, *Biased Tracers in Redshift Space in the EFT of Large-Scale Structure*, 1610.09321.

[71] B. Giblin *et. al.*, *KiDS-1000 catalogue: Weak gravitational lensing shear measurements*, *Astron. Astrophys.* **645** (2021) A105 [2007.01845].

[72] R. Scoccimarro, H. M. P. Couchman and J. A. Frieman, *The Bispectrum as a Signature of Gravitational Instability in Redshift-Space*, *Astrophys. J.* **517** (1999) 531–540 [astro-ph/9808305].

[73] N. Kaiser, *Clustering in real space and in redshift space*, *Mon.Not.Roy.Astron.Soc.* **227** (1987) 1–27.

[74] M. U. SubbaRao, M. A. Aragón-Calvo, H. W. Chen, J. M. Quashnock, A. S. Szalay and D. G. York, *Visualization of large scale structure from the sloan digital sky survey*, .

[75] Y. Donath and L. Senatore, *Biased Tracers in Redshift Space in the EFTofLSS with exact time dependence*, *JCAP* **10** (2020) 039 [2005.04805].

[76] K. C. Chan and L. Blot, *Assessment of the Information Content of the Power Spectrum and Bispectrum*, *Phys. Rev. D* **96** (2017), no. 2 023528 [1610.06585].

[77] 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*, 2108.03204.

[78] D. J. Eisenstein, H.-j. Seo and . White, Martin J., *On the Robustness of the Acoustic Scale in the Low-Redshift Clustering of Matter*, *Astrophys.J.* **664** (2007) 660–674 [astro-ph/0604361].

[79] Y. B. Zel'dovich, *Survey of modern cosmology*, *Adv. Astron. Astrophys.* **3** (1965), no. 241.

[80] M. M. Ivanov and S. Sibiryakov, *Infrared Resummation for Biased Tracers in Redshift Space*, *JCAP* **07** (2018) 053 [1804.05080].

[81] L. Senatore and M. Zaldarriaga, *The IR-resummed Effective Field Theory of Large Scale Structures*, *JCAP* **1502** (2015), no. 02 013 [1404.5954].

[82] D. Foreman-Mackey, D. W. Hogg, D. Lang and J. Goodman, *emcee: The MCMC Hammer*, *Publ. Astron. Soc. Pac.* **125** (2013) 306–312 [1202.3665].

[83] T. Brinckmann and J. Lesgourgues, *MontePython 3: boosted MCMC sampler and other features*, 1804.07261.

[84] B. Audren, J. Lesgourgues, K. Benabed and S. Prunet, *Conservative Constraints on Early Cosmology: an illustration of the Monte Python cosmological parameter inference code*, *JCAP* **1302** (2013) 001 [1210.7183].

[85] J. Torrado and A. Lewis, *Cobaya: Code for Bayesian Analysis of hierarchical physical models*, *JCAP* **05** (2021) 057 [2005.05290].

[86] X. Fang, J. A. Blazek, J. E. McEwen and C. M. Hirata, *FAST-PT II: an algorithm to calculate convolution integrals of general tensor quantities in cosmological perturbation theory*, *JCAP* **02** (2017) 030 [1609.05978].

[87] 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** (2020), no. 6 063533 [2004.10607].

[88] S.-F. Chen, Z. Vlah and M. White, *Consistent Modeling of Velocity Statistics and Redshift-Space Distortions in One-Loop Perturbation Theory*, *JCAP* **07** (2020) 062 [2005.00523].

[89] S.-F. Chen, Z. Vlah, E. Castorina and M. White, *Redshift-Space Distortions in Lagrangian Perturbation Theory*, *JCAP* **03** (2021) 100 [2012.04636].

[90] E. Pajer and M. Zaldarriaga, *On the Renormalization of the Effective Field Theory of Large Scale Structures*, *JCAP* **1308** (2013) 037 [1301.7182].

[91] **BOSS** Collaboration, K. S. Dawson *et. al.*, *The Baryon Oscillation Spectroscopic Survey of SDSS-III*, *Astron. J.* **145** (2013) 10 [1208.0022].

[92] T. Colas, G. D'amico, L. Senatore, P. Zhang and F. Beutler, *Efficient Cosmological Analysis of the SDSS/BOSS data from the Effective Field Theory of Large-Scale Structure*, *JCAP* **06** (2020) 001 [1909.07951].

[93] C. Alcock and B. Paczynski, *An evolution free test for non-zero cosmological constant*, *Nature* **281** (1979) 358–359.

[94] P. J. E. Peebles, *The Large-Scale Structure of the Universe*. Princeton University Press, 1980.

[95] A. Nusser and J. M. Colberg, *The omega dependence in the equations of motion*, *Mon. Not. Roy. Astron. Soc.* **294** (1998) 457 [astro-ph/9705121].

[96] F. Bernardeau, M. Crocce and R. Scoccimarro, *Constructing Regularized Cosmic Propagators*, *Phys.Rev.* **D85** (2012) 123519 [1112.3895].

[97] F. Schmidt, *Cosmological Simulations of Normal-Branch Braneworld Gravity*, *Phys. Rev. D* **80** (2009) 123003 [0910.0235].

[98] W. Hu and I. Sawicki, *Models of f(R) Cosmic Acceleration that Evade Solar-System Tests*, *Phys. Rev. D* **76** (2007) 064004 [0705.1158].

[99] B. Li, *Simulating Large-Scale Structure for Models of Cosmic Acceleration*. 2514-3433. IOP Publishing, 2018.

[100] K. Koyama, *Gravity beyond general relativity*, *Int. J. Mod. Phys. D* **27** (2018), no. 15 1848001.

[101] L. Lombriser, *Constraining chameleon models with cosmology*, *Annalen Phys.* **526** (2014) 259–282 [1403.4268].

[102] L. Lombriser, W. Hu, W. Fang and U. Seljak, *Cosmological Constraints on DGP Braneworld Gravity with Brane Tension*, *Phys. Rev. D* **80** (2009) 063536 [0905.1112].

[103] B. Bose, K. Koyama, M. Lewandowski, F. Vernizzi and H. A. Winther, *Towards Precision Constraints on Gravity with the Effective Field Theory of Large-Scale Structure*, *JCAP* **04** (2018) 063 [1802.01566].

[104] 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** (2020), no. 12 123541 [2003.08277].

[105] G. Cusin, M. Lewandowski and F. Vernizzi, *Nonlinear Effective Theory of Dark Energy*, *JCAP* **04** (2018) 061 [1712.02782].

[106] G. Cusin, M. Lewandowski and F. Vernizzi, *Dark Energy and Modified Gravity in the Effective Field Theory of Large-Scale Structure*, *JCAP* **1804** (2018), no. 04 005 [1712.02783].

[107] M. A. Luty, M. Porrati and R. Rattazzi, *Strong interactions and stability in the DGP model*, *JHEP* **09** (2003) 029 [hep-th/0303116].

[108] K. Koyama, A. Taruya and T. Hiramatsu, *Non-linear Evolution of Matter Power Spectrum in Modified Theory of Gravity*, *Phys.Rev.* **D79** (2009) 123512 [0902.0618].

[109] E. Sefusatti and F. Vernizzi, *Cosmological structure formation with clustering quintessence*, *JCAP* **1103** (2011) 047 [1101.1026].

[110] G. D'Amico, Y. Donath, L. Senatore and P. Zhang, *Limits on Clustering and Smooth Quintessence from the EFTofLSS*, 2012.07554.

[111] M. Tegmark, *Measuring cosmological parameters with galaxy surveys*, *Phys. Rev. Lett.* **79** (1997) 3806–3809 [astro-ph/9706198].

[112] T. Baldauf, M. Mirbabayi, M. Simonović and M. Zaldarriaga, *Equivalence Principle and the Baryon Acoustic Peak*, *Phys. Rev.* **D92** (2015), no. 4 043514 [1504.04366].

[113] M. Peloso and M. Pietroni, *Ward identities and consistency relations for the large scale structure with multiple species*, *JCAP* **1404** (2014) 011 [1310.7915].

[114] P. Creminelli, J. Gleyzes, M. Simonović and F. Vernizzi, *Single-Field Consistency Relations of Large Scale Structure. Part II: Resummation and Redshift Space*, *JCAP* **1402** (2014) 051 [1311.0290].

[115] A. Esposito, L. Hui and R. Scoccimarro, *Nonperturbative test of consistency relations and their violation*, *Phys. Rev. D* **100** (2019), no. 4 043536 [1905.11423].

[116] P. Creminelli, J. Gleyzes, L. Hui, M. Simonović and F. Vernizzi, *Single-Field Consistency Relations of Large Scale Structure. Part III: Test of the Equivalence Principle*, *JCAP* **1406** (2014) 009 [1312.6074].

[117] Z. Slepian *et. al.*, *Detection of baryon acoustic oscillation features in the large-scale three-point correlation function of SDSS BOSS DR12 CMASS galaxies*, *Mon. Not. Roy. Astron. Soc.* **469** (2017), no. 2 1738–1751 [1607.06097].

[118] J. Tyson, D. Wittman, J. Hennawi and D. Spergel, *LSST: A Complementary probe of dark energy*, *Nucl. Phys. B Proc. Suppl.* **124** (2003) 21–29 [astro-ph/0209632].

[119] L. Amendola *et. al.*, *Cosmology and fundamental physics with the Euclid satellite*, *Living Rev. Rel.* **21** (2018), no. 1 2 [1606.00180].

[120] O. Doré *et. al.*, *WFIRST: The Essential Cosmology Space Observatory for the Coming Decade*, 1904.01174.

[121] A. Taruya, T. Nishimichi and S. Saito, *Baryon Acoustic Oscillations in 2D: Modeling Redshift- space Power Spectrum from Perturbation Theory*, *Phys. Rev.* **D82** (2010) 063522 [1006.0699].

[122] V. Springel, *The Cosmological simulation code GADGET-2*, *Mon.Not.Roy.Astron.Soc.* **364** (2005) 1105–1134 [astro-ph/0505010].

[123] P. S. Behroozi, R. H. Wechsler and H.-Y. Wu, *The Rockstar Phase-Space Temporal Halo Finder and the Velocity Offsets of Cluster Cores*, *Astrophys. J.* **762** (2013) 109 [1110.4372].

[124] T. Nishimichi *et. al.*, *Dark Quest. I. Fast and Accurate Emulation of Halo Clustering Statistics and Its Application to Galaxy Clustering*, *Astrophys. J.* **884** (2019) 29 [1811.09504].

[125] T. Baldauf, L. Mercolli, M. Mirbabayi and E. Pajer, *The Bispectrum in the Effective Field Theory of Large Scale Structure*, *JCAP* **1505** (2015), no. 05 007 [1406.4135].

[126] E. Sefusatti, M. Crocce, R. Scoccimarro and H. Couchman, *Accurate Estimators of Correlation Functions in Fourier Space*, *Mon. Not. Roy. Astron. Soc.* **460** (2016), no. 4 3624–3636 [1512.07295].

[127] T. Baldauf, U. Seljak, V. Desjacques and P. McDonald, *Evidence for Quadratic Tidal Tensor Bias from the Halo Bispectrum*, *Phys. Rev.* **D86** (2012) 083540 [1201.4827].

[128] R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles*. Taylor and Francis, 1981.

[129] Y. P. Jing, *Correcting for the alias effect when measuring the power spectrum using FFT*, *Astrophys. J.* **620** (2005) 559–563 [astro-ph/0409240].

[130] D. Foreman-Mackey, D. W. Hogg, D. Lang and J. Goodman, *emcee: The MCMC Hammer*, *Publ. Astron. Soc. Pac.* **125** (2013) 306–312 [1202.3665].

[131] A. Lewis, *GetDist: a Python package for analysing Monte Carlo samples*, 1910.13970.

[132] R. Scoccimarro, S. Colombi, J. N. Fry, J. A. Frieman, E. Hivon and A. Melott, *Nonlinear evolution of the bispectrum of cosmological perturbations*, *The Astrophysical Journal* **496** (Apr, 1998) 586–604.

[133] F. Beutler, M. Biagetti, D. Green, A. Slosar and B. Wallisch, *Primordial Features from Linear to Nonlinear Scales*, *Phys. Rev. Res.* **1** (2019), no. 3 033209 [1906.08758].

[134] A. Vasudevan, M. M. Ivanov, S. Sibiryakov and J. Lesgourgues, *Time-sliced perturbation theory with primordial non-Gaussianity and effects of large bulk flows on inflationary oscillating features*, *JCAP* **09** (2019) 037 [1906.08697].

[135] L. Mercolli and E. Pajer, *On the velocity in the Effective Field Theory of Large Scale Structures*, *JCAP* **1403** (2014) 006 [1307.3220].

[136] J. J. M. Carrasco, S. Foreman, D. Green and L. Senatore, *The Effective Field Theory of Large Scale Structures at Two Loops*, *JCAP* **1407** (2014) 057 [1310.0464].

[137] S. Pueblas and R. Scoccimarro, *Generation of Vorticity and Velocity Dispersion by Orbit Crossing*, *Phys.Rev.* **D80** (2009) 043504 [0809.4606].

[138] 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** (2015) 029 [1503.08826].

[139] B. Horn, L. Hui and X. Xiao, *Soft-Pion Theorems for Large Scale Structure*, *JCAP* **09** (2014) 044 [1406.0842].

[140] A. A. Abolhasani, M. Mirbabayi and E. Pajer, *Systematic Renormalization of the Effective Theory of Large Scale Structure*, *JCAP* **05** (2016) 063 [1509.07886].

[141] M. Crocce and R. Scoccimarro, *Renormalized cosmological perturbation theory*, *Phys. Rev. D* **73** (2006) 063519 [astro-ph/0509418].

[142] V. Desjacques, M. Crocce, R. Scoccimarro and R. K. Sheth, *Modeling scale-dependent bias on the baryonic acoustic scale with the statistics of peaks of Gaussian random fields*, *Phys.Rev.* **D82** (2010) 103529 [1009.3449].

[143] A. Eggemeier, R. Scoccimarro and R. E. Smith, *Bias Loop Corrections to the Galaxy Bispectrum*, *Phys. Rev. D* **99** (2019), no. 12 123514 [1812.03208].

[144] K. C. Chan, R. Scoccimarro and R. K. Sheth, *Gravity and Large-Scale Non-local Bias*, *Phys. Rev. D* **85** (2012) 083509 [1201.3614].

[145] S. Saito, T. Baldauf, Z. Vlah, U. Seljak, T. Okumura and P. McDonald, *Understanding higher-order nonlocal halo bias at large scales by combining the power spectrum with the bispectrum*, *Phys. Rev. D* **90** (2014), no. 12 123522 [1405.1447].

[146] V. Assassi, D. Baumann, D. Green and M. Zaldarriaga, *Renormalized Halo Bias*, *JCAP* **08** (2014) 056 [1402.5916].

[147] M. Mirbabayi, F. Schmidt and M. Zaldarriaga, *Biased Tracers and Time Evolution*, *JCAP* **07** (2015) 030 [1412.5169].

[148] L. Senatore, *Bias in the Effective Field Theory of Large Scale Structures*, *JCAP* **11** (2015) 007 [1406.7843].

[149] T. Fujita, V. Mauerhofer, L. Senatore, Z. Vlah and R. Angulo, *Very Massive Tracers and Higher Derivative Biases*, *JCAP* **01** (2020) 009 [1609.00717].

[150] T. Fujita and Z. Vlah, *Perturbative description of biased tracers using consistency relations of LSS*, *JCAP* **10** (2020) 059 [2003.10114].

[151] T. Nishimichi, F. Bernardeau and A. Taruya, *Response function of the large-scale structure of the universe to the small scale inhomogeneities*, *Phys. Lett.* **B762** (2016) 247–252 [1411.2970].

[152] T. Baldauf, L. Mercolli and M. Zaldarriaga, *Effective field theory of large scale structure at two loops: The apparent scale dependence of the speed of sound*, *Phys. Rev. D* **92** (2015), no. 12 123007 [1507.02256].

[153] E. Noda, M. Peloso and M. Pietroni, *A Robust BAO Extractor*, *JCAP* **1708** (2017), no. 08 007 [1705.01475].

[154] M. Crisostomi, M. Lewandowski and F. Vernizzi, *Consistency relations for large-scale structure in modified gravity and the matter bispectrum*, *Phys. Rev. D* **101** (2020), no. 12 123501 [1909.07366].

[155] M. Zumalacárregui and J. García-Bellido, *Transforming gravity: from derivative couplings to matter to second-order scalar-tensor theories beyond the Horndeski Lagrangian*, *Phys. Rev. D* **89** (2014) 064046 [1308.4685].

[156] J. Gleyzes, D. Langlois, F. Piazza and F. Vernizzi, *Healthy theories beyond Horndeski*, *Phys. Rev. Lett.* **114** (2015), no. 21 211101 [1404.6495].

[157] D. Langlois and K. Noui, *Degenerate higher derivative theories beyond Horndeski: evading the Ostrogradski instability*, *JCAP* **02** (2016) 034 [1510.06930].

[158] M. Lewandowski, *Violation of the consistency relations for large-scale structure with dark energy*, *JCAP* **08** (2020) 044 [1912.12292].

[159] **Euclid Theory Working Group** Collaboration, L. Amendola *et. al.*, *Cosmology and fundamental physics with the Euclid satellite*, 1206.1225.

[160] 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** (2020), no. 8 083504 [1912.08208].

[161] 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** (2020) 032 [2002.04035].

[162] G. D'Amico, L. Senatore, P. Zhang and H. Zheng, *The Hubble Tension in Light of the Full-Shape Analysis of Large-Scale Structure Data*, *JCAP* **05** (2021) 072 [2006.12420].

[163] M. M. Ivanov, E. McDonough, J. C. Hill, M. Simonović, M. W. Toomey, S. Alexander and M. Zaldarriaga, *Constraining Early Dark Energy with Large-Scale Structure*, *Phys. Rev. D* **102** (2020), no. 10 103502 [2006.11235].

[164] M. M. Ivanov, O. H. E. Philcox, M. Simonović, M. Zaldarriaga, T. Nishimichi and M. Takada, *Cosmological constraints without fingers of God*, 2110.00006.

[165] G. D'Amico, L. Senatore, P. Zhang and T. Nishimichi, *Taming redshift-space distortion effects in the EFTofLSS and its application to data*, 2110.00016.

[166] M. M. Ivanov, O. H. E. Philcox, T. Nishimichi, M. Simonović, M. Takada and M. Zaldarriaga, *Precision analysis of the redshift-space galaxy bispectrum*, 2110.10161.

[167] A. Aviles, A. Banerjee, G. Niz and Z. Slepian, *Clustering in Massive Neutrino Cosmologies via Eulerian Perturbation Theory*, 2106.13771.