Primordial Non-Gaussianity

in the

Effective Field Theory of Large Scale Structure

About the cover: The back cover is Janet Towbin's 2016 drawing *Cosmic Web*. The front is a detailed version of it.

ISBN:
Geprint door: ProefschriftMaken || www.proefschriftmaken.nl

Primordial Non-Gaussianity in the Effective Field Theory of Large Scale Structure

Niet-Gaussische Begincondities in de Effectieve Veldentheorie van de Groteschaalstructuur

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. H. R. B. M. Kummeling, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op maandag 4 juni 2018 des middags te 12.45 uur

door

Drian Christoph van der Woude

geboren op 13 oktober 1989
te 's-Gravenhage

Promotor:	Prof. dr. S.J.G. Vandoren
Copromotor:	Dr. E. Pajer

Contents

List of Publications

The original content described in this thesis was first published in the following scientific works.

- V. Assassi, D. Baumann, E. Pajer, Y. Welling, and D. van der Woude, "Effective Theory of Large-Scale Structure with Primordial Non-Gaussianity," JCAP **1511** (2015) 024, arXiv:1505.06668 [astro-ph.CO] [1] (Chapter 4).
- Y. Welling, D. van der Woude, and E. Pajer, "Lifting Primordial Non-Gaussianity Above the Noise," JCAP **1608** no. 08, (2016) 044, arXiv:1605.06426 [astro-ph.CO] [2] (Chapter 5).
- E. Pajer and D. van der Woude, "Divergence of Perturbation Theory in Large Scale Structures," Manuscript under review at JCAP, arXiv:1710.01736 [astro-ph.CO] [3] (Chapter 6).

Other publications to which the author contributed during his PhD (not included in this thesis):

- G. Cabass, E. Pajer, and D. van der Woude, "A Robust Window on Primordial Non-Gaussianity," Manuscript in preparation.
- S. Jazayeri, E. Pajer, and D. van der Woude, "Adiabatic Modes for a Solid Universe," Manuscript in preparation.

Preface

"Je gaat het pas zien als je het doorhebt."

– Johan Cruijff

If I had to summarize the past four years as a beginning researcher in theoretical physics, it would be: the pursuit of understanding. I feel extremely fortunate to have been able to make this my sole job for so long, to have been able to take full advantage of the richness of opportunities at the Institute for Theoretical Physics here in Utrecht, to have met so many fun and kind people from all over the world sharing the same passion for understanding, and to have been able to learn so much first hand from my daily supervisor Enrico Pajer.

Despite the fact that I am surely no exception to the standard lore that one leaves the PhD with many more questions than one enters it with, I am very happy to report here that there are at least three modest research questions to which we have found a novel understanding. This thesis is my attempt to share them with you.

I sincerely hope to be able to communicate at least part of the joy I experienced learning these things for the first time. Enjoy the read!

Drian van der Woude
May 2018, Utrecht

1 Introduction

The purpose of this thesis is to improve our description and understanding of the evolution of the distribution of matter on the largest possible scales. To engage the less specialized reader and to argue for the relevance of the topic, this introduction provides a brief personal invitation to the subject matter. It is therefore not aimed to be scientifically complete or unbiased. At the end we outline the structure of the scientific content of the thesis.

1.1 Scales in physics

My favorite artistic expression of the relevance of scales is Georges Seurat's 1884 pointillist painting *A Sunday Afternoon on the Island of La Grande Jatte*, see Figure 1.1. If one zooms in on the painting, one discovers that it is entirely made up of little dots of paint. Obviously, none of these dots individually represent the millions of dollars that the painting is worth as a whole: it is only when one considers the large collection of them, that the aesthetic value becomes apparent. If one thinks about this a bit longer, the same principle is actually true of any image printed anywhere: there is always a fundamental scale at which the image is pixelated. Most of the times we do not notice it, since our eyes also have a finite resolution and if the pixel size is much smaller than this resolution, we interpret the image as continuous. A good example are the black letters you are reading now. In detail, they consist of a large collection of black dots, but also include some very small blank spots in between. When we see them, our visual mechanism goes through some sort of averaging procedure: regions consisting of mostly black dots are interpreted as a single black 'pixel'. In that sense, Seurat's painting is not that special at all. The reason it's special is that the fundamental pixel size is precisely such that at a regular distance from the image, one is able to appreciate both the pixelated nature *and* the emergent aesthetic beauty simultaneously. To me, this is one of the appealing aspects of impressionism: it allows one to see in one shot the true

Figure 1.1: *A Sunday Afternoon on the Island of La Grande Jatte* by Georges Seurat.

nature of how a collection of meaningless ingredients can conspire to form a beautiful emergent world.

In theoretical physics, this concept of the world changing upon zooming in and out has become one of the cornerstones of modern research[1]. A condensed matter theorist is always aware of a fundamental scale that limits his or her field theoretic description and the standard model of elementary particle physics is presently merely considered as an effective theory at the scales we have probed so far. What sets a theoretical physicist apart from an art aficionado is that in physics we make this observation mathematically precise: how exactly is our description of the world on certain scales sensitive to its constituents on much smaller scales? The Effective Field Theory of Large Scale Structure (EFT of LSS) is a theory designed specifically to an-

[1] A mathematically more precise definition of these statements will be given in the next chapter, but for the impatient reader: zooming out refers to coarse-graining or smoothing operations on fields.

swer that question for the distribution of matter on the largest cosmological scales.

1.2 Cosmology and the large scale structure of the universe

As you can imagine, today's definition of 'large scale structure' is not what it used to be. For instance, when in 1914 the British astronomer and physicist Sir Arthur Eddington wrote the book *Stellar Movements and the Structure of the Universe*, he was merely referring to the Milky Way: a microscopic system from the point of view of this thesis [2].

Despite this apparent subjectivity of the meaning of 'large scales', the scales this thesis deals with are very well defined. More appropriately, they can be referred to as cosmologically large scales. Let me take the opportunity to clarify a common confusion between cosmology and astronomy at this point. Even though there is plenty of cross-contamination between the fields (as there ought to be in any mature science), and it is not always clear to which field a particular study should be assigned, the rough idea is that astronomy deals with the objects that make up the universe, whereas cosmology deals with the universe as a whole. In terms of scales, cosmology is the study of scales at which the universe can be described as 'almost' homogeneous and isotropic. More mathematically speaking, cosmology deals with the evolution of the universe on scales that are perturbatively close to the homogeneous and isotropic background. The Milky Way is an example of a complicated object that belongs to the realm of astronomy. The two are compared in Figure 1.2.

In many ways, it is remarkable that the homogeneous universe is an interesting field of study at all. On the one hand, it could very well have been extremely boring to study a homogeneous world. Einstein, for one, for a long time believed the homogeneous universe had to be static: it just stays as it is forever.[3] On the other hand, it could have been highly irregular, way too complicated for us to describe. Fortunately, it turns out to be

[2] The smallest scales described in this thesis are more than a factor hundred larger than the Milky Way.

[3] For a more general historical discussion, see for instance Peebles's classic work on large scale structure [4].

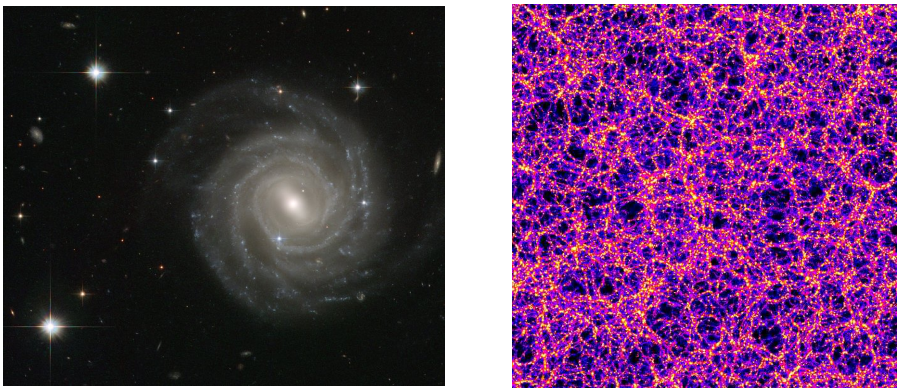


Figure 1.2: The left Figure shows a photograph by the Hubble space telescope of galaxy UGC 12158 (credit: ESA/Hubble and NASA), which is thought to resemble the Milky Way. The right Figure is a simulated image of cosmological large scale structures (credit: HPCC group of the University of Washington). The scale of the picture on the right is approximately a thousand times larger. The large scale distribution of matter indeed seems pretty homogeneous.

right in between these extremes, and it is often on that boundary between complexity and simplicity where the most interesting scientific questions lie, as beautifully argued by Timothy Gowers at the Millennium Celebration of the Clay Mathematics Institute.[4] The key reasons are really the pillars of cosmological research:

- The cosmological principle: the universe is statistically homogeneous and isotropic on large enough scales.
- General relativity: homogeneous and isotropic spacetimes are, generically, dynamical.

By statistical homogeneity and isotropy we mean that the distribution of galaxies is as if they have been placed at random, with no preferred position or orientation.[5] It is this large scale simplicity of the universe that allows for exact approximate solutions for its dynamics, as one of the few examples

[4] I highly recommend his lecture *On the Importance of Mathematics*. It is on Youtube.

[5] Mathematically, statistical homogeneity and isotropy are implemented at the level of correlation functions.

of exact solutions to Einstein's equations ever found [6]. Originally, it was Einstein himself who, in lack of firm observational confirmation, was mostly inspired by Mach's principle to consider closed, homogeneous and spherical universes. His belief in a static universe famously led him to introduce the cosmological constant [5]. Eventually, it took the likes of Friedmann, Lemaitre, Robertson and Walker to establish the unique solution to spatially homogeneous and isotropic spacetimes, the FLRW solution [7]:

$$ds^2 = -dt^2 + a^2(t)\left[\frac{dr^2}{1-\kappa r^2} + r^2 d\Omega^2\right], \tag{1.2.1}$$

for some function of time $a(t)$ multiplying the spatial metric in spherical coordinates, with $\kappa \in \{-1, 0, 1\}$ (normalized such that a is the physical radius of curvature), for spatially open, flat or closed universes respectively. Around that time, observational evidence began to point in the same direction: Hubble [6] and Lemaitre [7] showed that the distribution of galaxies known at the time was indeed consistent with the large scale homogeneity. As such, the FLRW model became the working model for the evolution for the universe and the field of cosmology was born.

After the Second World War, Gamow realized[8] that the expansion of the universe implies an early very hot and dense state, during which the energy density of the universe was dominated by radiation. The hot big bang scenario was born. As its main feat, it was able to explain the abundance of the simplest chemical elements. Moreover, it predicted an omnipresent blackbody distribution of photons, as a remnant of the recombination of electrons and protons to form neutral hydrogen in the early universe, which today we know as the Cosmic Microwave Background Radiation (CMB). It is often referred to as 'the first light', and it yields a wonderful snapshot of the way the universe looked at the age of 380.000 years old. The most beautiful and scientifically useful image was provided by the Planck collaboration and is depicted in Figure 1.3. The relation between this map of the CMB and the

[6] Other famous examples are the nonlinear solution in vaccuum (black holes), and the perturbative, linear solution for gravitational waves.

[7] Friedmann was the first to find this solution, but it was largely neglected up untill his death of typhoid in 1925. Lemaitre later discovered the same solution independently. In 1935 Robertson and Walker rigorously proved it to be the only spatially homogeneous and isotropic solution to Einstein's equations.

[8] See for instance the discussion in [8]

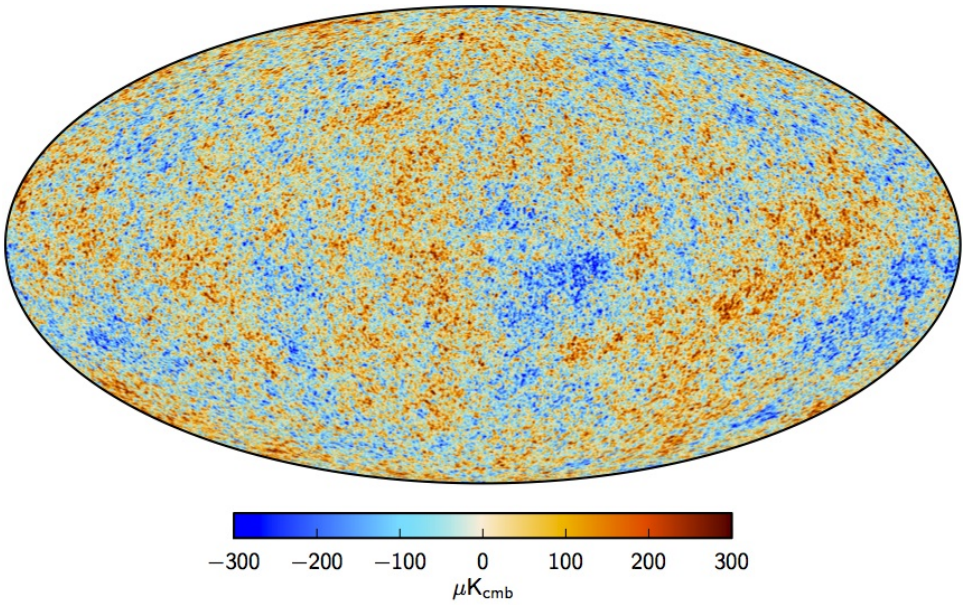


Figure 1.3: The Planck collaboration's 2015 image of the universe at the time of recombination. It is the projection of a two-dimensional sphere on the sky on this flat piece of paper. The image shows that the universe was homogeneous to a very high degree: fluctuations are rarely more than a 10^{-5} fraction of the average CMB temperature of 2.7K.

LSS we see today is simple: since the energy density in radiation dilutes faster than Dark Matter (DM), at some point the gravitational potential (which sources the dark matter perturbations) was no longer influenced by oscillations in the photon-baryon fluid, and dark matter perturbations became free to collapse under the influence of gravity in the remaining 14 billion years.[9] In this sense, we have a very complete picture of the history of the universe, including its perturbations.

The question we have not addressed so far is the question that directly leads to the work of this thesis: to what extent does the description of the universe as purely homogeneous and isotropic match the evolution of the real universe, which is in fact highly irregular on the scales of the Milky

[9] To be precise, dark matter perturbation grow logarithmically during radiation domination, and linearly during matter domination.

Way. This question recently led to some discussion in the context of Dark Energy (see, for instance [9, 10]). Interestingly, the most convincing paper on this topic in my opinion [11], which addresses the question from the point of view of effective field theory, is in fact the paper that first proposed the EFT of LSS by asking a similar question for the evolution of perturbations: how do the evolution equations for density perturbations in the universe on cosmological scales compare to similar evolution equations in a universe without any small scale irregularities? In chapter 3 we review how the EFT of LSS addresses this question. The original work of this thesis studies a particular extension and the reach of the EFT of LSS. In the next section we outline in more detail the content of each Chapter.

1.3 Outline

The thesis is divided into two preliminary chapters, introducing the topic for non-cosmologists, and three core chapters, which are (slight modifications of) the the following three papers: [1–3]. Let us here summarize the content and results of each chapter.

- **Preliminaries I: The Cosmological Collider** 2. The aim of this Chapter is to introduce the concept of Primordial Non-Gaussianity (PNG) and show how density perturbations in the late universe can be used as a tool to learn about the physical content of the high energy very early universe.

- **Preliminaries II: Large Scale Structure as an Effective Fluid** 3. In this Chapter we extend the discussion of perturbations in the late universe to density perturbations in Large Scale Structure (LSS) today. Since by now, these scales have become sub-Hubble (a concept we explain later) and nonlinear, we need to extend the theory beyond linear perturbation theory. We argue that the only consistent theory in this respect is the EFT of LSS, which we introduce from a bottom-up and top-down perspective.

- **The EFT of LSS with Primordial Non-Gaussianity** 4. We argue that the EFT of LSS has to be extended to properly account for the presence of non-Gaussian correlations in the initial conditions. We

show that a well defined new set of (non-local) operators has to be added to the expansion of the stress tensor, which makes the new theory consistent and renormalizable. Finally, we take some first steps towards a numerical analysis of the relevance of various contributions to matter bispectrum, and argue that a solid control of theoretical uncertainties is relevant.

- **Lifting Primordial Non-Gaussianity Above the Noise** 5. We forecast how well (near) future LSS surveys will be able to constrain the level of non-Gaussianity in the initial conditions, assuming an idealized survey that yields us the dark matter distribution directly. We find that the EFT is good enough (and an improvement compared to SPT) not to limit improving the current upper bounds from CMB surveys. On the other hand, we show that it is unlikely that at the order we are working at, we will get to often quoted theoretically interesting benchmarks. We thoroughly analyze the relevance of theoretical uncertainties in our statistical analysis.

- **Divergence of Perturbation Theory in Large Scale Structures** 6. To study the maximal reach of any perturbative approach to LSS, we study the convergence of SPT (as a proxy for any similar perturbation theory, including the EFT of LSS) for gravitational collapse in one spatial dimension, for which resummations are possible. We find that for Fourier space observables, the convergence of SPT depends on the initial conditions, but converges for realistic ones. We show that it diverges for the real space correlation function, as well as real space count-in-cell statistics. We argue that the origin of this divergence is related to the observation that SPT fails to capture correctly 'rare events' of high densities and velocities. We provide a naive estimate of the lower bound of non-perturbative corrections to LSS.

The core Chapters contain a conclusion and discussion themselves, but are also briefly recapped in section 7. A Dutch summary and information about the author can be found at the very end of the thesis.

2 Preliminaries I: The Cosmological Collider

The most amazing aspect of modern cosmology, if not science in general, is that there are very good reasons to believe the large scale structures that make up our universe are a direct consequence of quantum fluctuations on top of an extremely rapidly expanding early universe. The physical mechanism behind this early phase of so called inflation is still unknown and remains one of outstanding major questions in cosmology. As in particle physics, one way to learn more about the theory of inflation, is to study the interactions of its perturbations. This is precisely what we do here. We argue that, in some sense, inflation acts as a cosmological collider, whose interactions are imprinted in the statistics of the large scale structure of the universe, much like the interactions in a particle collider are imprinted on its final states.

The content of this chapter is largely based on [12, 13], but can be found in most standard cosmology textbooks. In the following sections, we try to argue as comprehensively as possible how the statements above come into practice. For this reason, we often stick to the simplest example of a scalar field and highlight essential features. We do not go into the detailed production mechanism for Primordial non-Gaussianity (PNG), as in the end this thesis only deals with the final outcome that serves as an initial condition for the large scale structure (LSS) computations of the next section. Nonetheless, we do hope to convince the reader of the relevance of PNG for high energy physics.

2.1 The evolution of physical scales in an expanding universe

The dynamics of the universe is given by the Einstein equations

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}. \tag{2.1.1}$$

As a consequence of homogeneity and isotropy, the energy-momentum tensor has to be of the perfect fluid form

$$T_{\mu\nu} = (\rho + p)U_\mu U_\nu + p g_{\mu\nu}, \tag{2.1.2}$$

where $g_{\mu\nu}$ is the FLRW metric, and $U^\mu = (1, 0, 0, 0)$ for an observer comoving with the fluid. This automatically yields a conservation equation for the energy-momentum of the universe

$$\dot{\rho} + 3H(\rho + p) = 0, \tag{2.1.3}$$

for the FLRW metric (1.2.1). Here $H = \dot{a}/a$. The remaining information in Einstein's equation is the so called Friedmann equation,

$$H^2 = \frac{8\pi G}{3}\rho - \frac{\kappa}{a^2}. \tag{2.1.4}$$

Together, these equations dictate expansion rate of the universe. By now, we have established the history summarized in Table 2.1. In this thesis, the

Energy density:	?	? – $(\leq 1\text{MeV})^4$	$(\leq 1\text{MeV})^4$ – $(1\text{eV})^4$	$(1\text{eV})^4$ – $(10^{-3}\text{eV})^4$	$(10^{-3}\text{eV})^4$
Era:	Inflation	Pre- and reheating	Radiation	Matter	Dark energy
Energy source:	Inflaton?	Inflaton/Radiation?	Radiation	Matter	Dark energy
Expansion:	$H \approx$ constant	$H \sim a^{-3}$	$H \sim a^{-4}$	$H \sim a^{-3}$	$H \sim$ constant

Table 2.1: History of the universe in chronological order, where we use energy density as time variable.

main relevance of the history of the universe is its effect on the evolution equations for perturbations. The detailed evolution equations depend on the physical content of the universe and are introduced below. Before getting there, we highlight a particular effect the expansion of the universe has on the evolution of physical scales.

Linear perturbation theory is most easily done by Fourier transforming a field of perturbations. Our Fourier conventions for perturbations $\varphi(x, t)$ are

$$\varphi(\boldsymbol{x}, t) = \int \frac{d^3k}{(2\pi)^3}\varphi(k, t)e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \equiv \int_k \varphi(k, t)e^{i\boldsymbol{k}\cdot\boldsymbol{x}}, \tag{2.1.5}$$

where $k = |\boldsymbol{k}|$. We can then study the evolution of one its Fourier modes $\varphi_k(t)$, which contributes $2\varphi_k(t)\cos(kx)$ in real space. More precisely, we will

study the evolution of the distance between two of its peaks as a function of time. Dropping order one numbers, the physical distance between them is

$$d = \int \sqrt{ds^2} = a\Delta x \sim \frac{a}{k}. \tag{2.1.6}$$

In other words, it simply increases with the scale factor. The point is here to compare this to the other scale in the universe: the Hubble scale (which we sometimes refer to as Hubble horizon). One way to understand the relevance of the Hubble scale is to ask how far two points have to be separated such that the physical distance between them increases faster than the speed of light. In units where the speed of light is unity, we then find

$$c = 1 = \dot{d}_{crit} = H d_{crit} \quad \Longrightarrow \quad d_{crit} = H^{-1}. \tag{2.1.7}$$

The next question is whether the distance between the peaks ever gets so far apart that they recede faster than the speed of light. Defining the conformal Hubble rate $\mathcal{H} = aH$, we find that the condition $d > d_{crit}$ implies $k > \mathcal{H}^{-1}$. Since k is a mere constant, the question whether a certain physical scale grows to get out of contact, or shrinks to so called sub-Hubble distances, is determined by the condition

$$\dot{\mathcal{H}} = \ddot{a} > 0. \tag{2.1.8}$$

In other words it depends on whether the universe expands in an accelerated or decelerated fashion. In terms of the history of the universe, this means that during inflation, modes become super-Hubble, whereas during matter and radiation domination the Hubble radius catches up again until eventually the modes become sub-Hubble. As we show below, this observation is the crucial reason today's perturbations teach us about quantum fluctuations during inflation. The short story is that tiny sub-Hubble perturbations during inflation are rapidly stretched to super-Hubble scales. Since at that point, not even light or gravity can communicate between peaks, the evolution becomes much simpler to compute in most cases[1]. After the universe settles into the later history, we just wait until modes re-enter the Hubble horizon, such that they become observable. In the following subsections, we make this more precise.

[1] As we comment on below, this holds for single field-attractor models of inflation in particular.

2.2 Inflation: background

Inflation is a hypothesized phase of near exponential expansion in the early universe. Equivalently, it is defined by the condition $\dot{H} \approx 0$. Going back to the Friedmann and conservation equation, this implies a near constant energy density ρ, and negative pressure s

$$\rho + p \approx 0. \tag{2.2.1}$$

If these conditions are satisfied, indeed $a(t) \sim \exp\{Ht\}$, and the curvature term in the Friedmann equation rapidly becomes irrelevant. In fact, the latter observation was one of the three main early motivations for inflation.

Namely, during radiation and matter domination, such a curvature term grows compared to the energy density. This is surprising, as its observed current contribution to the Friedmann equation is zero up to corrections of order 10^{-5}, meaning at early times it had to be even smaller. This is called the curvature problem of cosmology, and it is solved by a long enough initial phase of accelerated expansion. Similarly, the dilution of potential magnetic monopoles during inflation solves the monopole problem. Finally, the observed homogeneity of the CMB constituted the horizon problem: at the time of the CMB, many of the patches we observe today were outside of the Hubble horizon at the time, which, as shown in the previous section, means they were not able to communicate. Again, a period of accelerated expansion solves the issue as it allows distant regions to have been able to communicate earlier on.

Assuming nothing spectacular happens between the end of inflation and the beginning of the radiation era (see Table 2.1), one can obtain a rough estimation of minimum number of e-folds (the number of times the universe expands by a factor e). We assume that the energy density at the end of inflation was comparable to the energy density at the beginning if the radiation era, $\rho_{I,f} \approx \rho_{rad,i}$. From the horizon and curvature problem, one finds the following condition on the number of e-folds [12],

$$e^N > \frac{a_{I,f} H_{I,f}}{a_0 H_0}, \tag{2.2.2}$$

where the zero indicates today. The success of nucleosynthesis, the theory of the cosmological formation of light elements in the early hot universe,

suggests that radiation domination went back to at least as early as 1MeV, the time of neutron-proton conversion. On the other hand, in order for the classical background during inflation to make sense at all, we require the energy density at the end of inflation to be much less than the Planck energy density $\left(10^{19}\text{GeV}\right)^4$.[2] These two extremes suggest the minimal number of e-folds N to be between 17 and 68, assuming a standard history. Typically, these requirements are sufficient to solve the monopole problem as well.

Nonetheless, the strongest argument in favor of inflation comes from the theory of perturbations. In particular, detailed analysis of the (polarization) spectrum perturbations in the photon distribution in the CMB strongly suggest these perturbations to be of super-Hubble origin, with coherent initial phases. Inflation is a natural mechanism to produce such super-Hubble perturbations, as the discussion in the previous section suggests. For a more elaborate discussion of the latter argument, see [13, 14]. On top of this, the near scale independence of the amplitude of perturbations provides evidence for the near exponential expansion. We go into the theory of perturbations in more detail in the next section.

The physical mechanism behind inflation is still largely unknown. As a toy model, one often has in mind a scalar field slowly rolling down a potential,

$$S_\phi = \int d^4x\sqrt{-g}\left[-\frac{1}{2}\left(\partial\phi\right)^2 - V(\phi)\right] \tag{2.2.3}$$

One can easily check that for this field to be the dominant source of energy during the (homogeneous) inflationary phase of the universe, one requires

$$\dot{\phi}^2 \ll |V(\phi)|. \tag{2.2.4}$$

In order to sustain this condition for an extended period of time in simple models, one usually also assumes

$$|\ddot{\phi}| \ll H|\dot{\phi}|. \tag{2.2.5}$$

The homogeneous equation of motion for the scalar field then reduces to a single time-derivative (attractor) equation

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0 \quad \rightarrow \quad 3H\dot{\phi} + V'(\phi) = 0. \tag{2.2.6}$$

[2] The current upper bound on the tensor to scalar ratio (from the Planck collaboration) immediately yields an upper bound on the scale of inflation at the order of $10^{16}GeV$.

The slow roll conditions can be defined as

$$\varepsilon = -\frac{\dot{H}}{H^2} \approx \frac{M_{\rm pl}^2}{2}\left(\frac{V'(\phi)}{V(\phi)}\right)^2 \ll 1 \tag{2.2.7}$$

$$\eta = \frac{\dot{\varepsilon}}{H\varepsilon} \approx 4\varepsilon - 2M_{\rm pl}^2\frac{V''}{V} \ll 1, \tag{2.2.8}$$

where $M_{\rm pl}^{-2} = 8\pi G$ and the latter approximate equalities are slow roll approximations. The size of these parameters is related tot the number of e-folds of inflation, but the precise relation depends on the model (see e.g. [15]).

2.3 Inflation: perturbations

Most of the information about the physical nature of inflation is obtained from its perturbations. For that reason, the literature on inflationary models and its observational consequences for perturbations is vast. In the interest of continuity, we will skip many of the interesting extensions and subtleties of inflationary perturbations and stick to the simple scalar field example.

As always, the starting point is simply expanding the fields in the action in background and perturbations. In our case, this means expansion in homogeneous and space dependent functions. In our case there are in principle 'two' fields: the scalar field and the metric $g_{\mu\nu}$. We expand them as

$$\phi(x,t) = \bar{\phi}(t) + \varphi(x,t) \tag{2.3.1}$$

$$g_{\mu\nu}(x,t) = \bar{g}_{\mu\nu}(t) + h_{\mu\nu}(x,t). \tag{2.3.2}$$

Many of the subtle, and interesting aspects of cosmological perturbation theory can be ascribed to the fact that gravity is a gauge theory. For one, not all perturbations above correspond to physical degrees of freedom. Moreover, the homogeneity of the background refers to a specific (FLRW) reference frame, which is not always explicit. One of the consequences of this is a very powerful statement due to Weinberg about the late-time behavior of perturbations. We comment on this theorem below. For now, we forget about gravitational perturbations for a moment and focus on the scalar perturbations. [3]

[3] This can be thought of as the decoupling limit in this case.

Perturbation theory during inflation is an example of semi-classical gravity: the homogeneous background is treated classically, but the fluctuations are considered as quantum fields on a time dependent background. The classic work on this is [16]. A more pedagogical introduction is given in [17]. Sticking to cosmology, canonical quantization is actually a relatively straightforward generalization of the quantization of the harmonic oscillator. The equation of motion for the free operator φ is simply

$$\ddot{\hat{\varphi}} + 3H\dot{\hat{\varphi}} + V''(\bar{\phi})\hat{\varphi} - \left(\frac{\nabla^2}{a^2}\right)\hat{\varphi} = 0. \tag{2.3.3}$$

One can define its conjugate momentum $\hat{\pi} = \partial L_\phi/\partial\dot{\varphi} = a^3\dot{\hat{\varphi}}$, to write down the canonical commutation relations,

$$[\hat{\varphi}(x,t), \hat{\varphi}(y,t)] = 0\ , \quad \left[\hat{\varphi}(x,t), \hat{\dot{\varphi}}(y,t)\right] = \frac{i}{a^3}\delta_D^3(x-y), \tag{2.3.4}$$

where δ_D^3 is the three dimensional Dirac delta function and the momentum commutator vanishes as well.

Finally, one needs to specify a vacuum state. In Minkowski space, for which time translations are a good symmetry, there is a well defined Hamiltonian, which allows us to find the lowest energy state of a free scalar field. On time-dependent backgrounds this is more tricky. However, for the modes that are relevant for cosmology today (roughly the ones observed in the CMB), there is a natural candidate. Namely, in the inflationary paradigm, these modes are believed to originate on sub-Hubble scales before being stretched to super-Hubble scales. On scales much smaller than the Hubble scale, we expect the expansion of the universe to be negligible. This can be seen by looking for solutions of the form $\hat{\varphi}(x,t) = \hat{\varphi}_q(t)e^{iqx}$ to 2.3.3. If $q/a \gg H$, this reduces to

$$\ddot{\hat{\varphi}}_q + \left(\frac{q^2}{a^2}\right)\hat{\varphi}_q = 0, \tag{2.3.5}$$

which is essentially the equation of motion for the Fourier components of a free field on Minkowski, with (approximately time-independent) frequency $k = q/a$. In that case, one expands the operator as

$$\hat{\varphi}(\boldsymbol{x},t) = \int_k \left[\varphi_k(t)a_k e^{ia\boldsymbol{k}\cdot\boldsymbol{x}} + \varphi_k^\star(t)a_k^\dagger e^{-ia\boldsymbol{k}\cdot\boldsymbol{x}}\right], \tag{2.3.6}$$

where a and $a^\dagger$ are creation and annihilation operators, which satisfy

$$[a_k, a_{k'}] = 0\ , \quad \left[a_k, a^\dagger_{k'}\right] = (2\pi)^3\delta_D(k-k') = (2\pi)^3 a^3\delta_D(q-q'). \tag{2.3.7}$$

Then, the properly normalized solution for the time dependence in the sub-Hubble limit is

$$\varphi_k(t) = \frac{1}{\sqrt{2k}} e^{-i\int k dt}, \tag{2.3.8}$$

and the vacuum state $|0\rangle$ is selected by demanding that

$$a_k|0\rangle = 0\ \forall \boldsymbol{k}. \tag{2.3.9}$$

Through the identification $\alpha_q = a^{-3/2}a_q$, we can extend the Minkowski example to the cosmological setting and expand

$$\hat{\varphi}(\boldsymbol{x}, t) = \int_q \left[\varphi_q(t)\alpha_q e^{i\boldsymbol{q}\cdot\boldsymbol{x}} + \varphi^\star_q(t)\alpha^\dagger_q e^{-i\boldsymbol{q}\cdot\boldsymbol{x}}\right], \tag{2.3.10}$$

such that $\alpha_q, \alpha^\dagger_q$ satisfy the standard time-independent commutation relations, and in that case

$$\varphi_q(t) = a^{-3/2}\varphi_{k\to q/a}(t). \tag{2.3.11}$$

The full solution for the mode function in the decoupling limit is then given by a solution to

$$\ddot{\varphi}_q + 3H\dot{\varphi}_q + V''(\bar{\phi})\varphi_q - \left(\frac{q^2}{a^2}\right)\varphi_k = 0, \tag{2.3.12}$$

such that the solution reduces to the 'positive frequency' mode functions 2.3.8 in the early time, or large q limit. This choice of mode functions and creation and annihilation operators is known as the Bunch-Davies vaccuum.

This concludes this section. We showed how the natural quantum oscillations of the vacuum are treated on an inflationary background. In the next section we study the full evolution of the perturbations including gravity.

2.4 Connecting inflation to observations

Before making the connection, we have to understand how gauge transformations act on perturbations. Standard gauge transformations act on both coordinates and fields, for instance

$$\tilde{g}_{\mu\nu}(\tilde{x}) = g_{\alpha\beta}(x)\frac{\partial x^\alpha}{\partial \tilde{x}^\mu}\frac{\partial x^\beta}{\partial \tilde{x}^\nu}. \tag{2.4.1}$$

In cosmology, we prefer to leave the coordinates fixed, and interpret the entire change as a change of the perturbed field. For the metric perturbation, this means

$$\Delta h_{\mu\nu}(x) = \tilde{g}_{\mu\nu}(x) - g_{\mu\nu}(x), \tag{2.4.2}$$

leaving everything else fixed. The same logic is applied to all other tensors. Since the background is isotropic, at linear order in perturbations we can distinguish spatial scalars, vectors and tensors, which do not couple - the SVT decomposition. Counting degrees of freedom teaches us that there is a single gauge invariant scalar degree of freedom, one vector and one tensor. We focus on the scalar degree of freedom. A common choice of gauge for the metric is the Newtonian gauge, for which the scalar part of the metric reads

$$g_{00} = -1 - 2\Phi \quad g_{0i} = 0\,, \quad g_{ij} = a^2\left(1 - 2\Psi\right). \tag{2.4.3}$$

In this gauge, one can define the gauge invariant comoving curvature perturbation $\mathcal{R}$,[4]

$$\mathcal{R} = -\Psi - \frac{H}{\dot{\bar{\phi}}}\varphi, \tag{2.4.4}$$

which can once again be expanded as

$$\hat{\mathcal{R}}(\boldsymbol{x}, t) = \int_q \left[\mathcal{R}_q(t)\alpha_q e^{i\boldsymbol{q}\cdot\boldsymbol{x}} + \mathcal{R}_q^\star(t)\alpha_q^\dagger e^{-i\boldsymbol{q}\cdot\boldsymbol{x}}\right], \tag{2.4.5}$$

[4] For a definition of $\mathcal{R}$ before fixing the gauge, we can write $\mathcal{R} = A/2 + H\delta u$, where A and δu are particular perturbations of the metric and energy-momentum tensor, see section 5.1 of [12].

The components then satisfy the Mukhanov-Sasaki equation

$$\mathcal{R}_q'' + \frac{2}{z} z' \mathcal{R}_q' + q^2 R_q = 0, \tag{2.4.6}$$

where a prime denotes differentiation with respect to conformal time τ, defined through

$$d\tau = \frac{dt}{a(t)}, \tag{2.4.7}$$

and we have defined

$$z = \frac{a\dot{\phi}}{H}. \tag{2.4.8}$$

We can study two limits to this equation: $q/a \gg H$ and $q/a \ll H$. The first limit is effectively the one studied in the previous section. In this case, the constraint equation tells us Ψ_q is negligible, and $\mathcal{R}_q$ is related to the scalar field perturbations by a (time-dependent) multiplicative factor. Moreover, the vacuum discussion is unaltered. On super-Hubble scales, we can drop the q^2 term and find the solution

$$\mathcal{R}_{q\to 0} = c_1 + c_2 \int d\tau \frac{1}{z^2}. \tag{2.4.9}$$

The second solution decays as $\mathcal{R}_q \sim a^{-2}$ up to slow-roll corrections. Note however that the terms we have neglected, of order $q^2/(aH)^2$ decay similarly, so it is not always obvious how to relate these subleading terms. For a more elaborate discussion on this, see [18, 19].

The Weinberg theorem

The main result above is that the modes generically freeze to a constant. This interesting fact goes back to a 'trick' Weinberg presented in 2003 to find the long wavelength solution to the linear system of perturbation equations on FLRW backgrounds [18]. This is much more general than the explicit scalar field example we worked out here: for fluctuations on an FLRW background, there are always two physical solutions for $\mathcal{R}$, one of which decays

on super-Hubble scales, and one that approaches a constant. If the fluctuations are sourced by a single scalar degree of freedom, this leaves no other option for the solution than to match the Weinberg solutions.[5]

The reason is that these modes on super-Hubble scales become equivalent to a (large) coordinate change on the FLRW background. In other words, it has no locally observable physical effects (in this limit). Since this holds on all FLRW backgrounds, this means we do not have to solve any equations as long as the modes are outside of the Hubble horizon! This is the true power of this theorem: it allows us to relate inflationary perturbations to late-times (such as the CMB), when the modes re-enter. There is no need to solve the complicated reheating era. Moreover, they extend to symmetries of the full quantum theory, and can be used to derive Ward identities, see [23–25].

Finally, we note that in the post big bang universe, one can simply test whether the perturbations in our universe satisfy this particular solution, which are also referred to as adiabatic modes. Both the name and the observability come from the observation that for the adiabatic solution described above, all scalars have the property that,

$$\frac{\delta s}{\dot{\bar{s}}} = \frac{\delta \rho}{\dot{\bar{\rho}}}, \tag{2.4.10}$$

for any scalar s. This can be tested for in the CMB (other scalars are for instance the cold dark matter density ρ_m by itself) and everything is consistent with adiabaticity within a couple of percent [26]. Note however that even if multiple scalar fields were active during inflation, producing non-adiabatic perturbations, at late times, a period of local thermal equilibrium with no non-zero conserved quantities always drives the perturbations towards the adiabatic solution [27].

Classicalization

The only thing we have skipped so far is how these inherently quantum mechanical perturbations during inflation translate into classical stochastic

[5] By now, there have been a number of exceptions. Solid inflation is an example of a single field model with a different symmetry breaking pattern that evades the theorem [20]. Non-attractor models also have a slightly different story [21], althouh it is still in the realm of general arguments based on symmetries of the theory [22].

variables at late times. This can only happen if some sort of decoherence sets in as modes exit the Hubble horizon, such that operators can be interpreted as classical stochastic variables - correlation functions are expectation values with respect to some classical ensemble. This identification is only possible if all operators commute, as classical variables do. We can check if this is the case by defining the classicality condition

$$\mathcal{C} = \frac{\langle [\mathcal{R}, \dot{\mathcal{R}}] \rangle^2}{\langle \dot{\mathcal{R}}^2 \rangle \langle \mathcal{R}^2 \rangle}, \tag{2.4.11}$$

which assumes near Gaussianity for the estimate to be reliable. Going back to 2.3.4, we find that the commutator of $\mathcal{R}$ and $\dot{\mathcal{R}}$ vanishes as a^{-3}, up to slow roll corrections. Since we saw that $\mathcal{R}$ itself goes to a constant at late times, whereas $\dot{\mathcal{R}}$ goes as a^{-2} up to slow roll corrections, we conclude that indeed a classical stochastic interpretation is possible. This argument makes the case for decoherence at least reasonable.

Ergodic theorem

Finally, we assume that in today's universe, every mode we observe is a particular realization of the ensemble defined by the averages above (explicit ensemble averages are studied below in the form of correlators). In order to obtain more information about this ensemble, we need multiple measurements. For this, we employ the ergodic theorem, which tells us that under reasonable assumptions, ensemble averages are equal to spatial averages. More explicitly, Appendix D of [12] shows that if the correlations between the field values of φ at distantly separated points goes to zero rapidly enough, the difference between the ensemble average and spatial average around some point z_0 goes to zero. That is, let,

$$\Delta_R^2(x_1, x_2, \ldots) \equiv \Big\langle \left(\int_z W_R(z) \varphi(x_1 + z) \varphi(x_2 + z) \cdots - \langle \varphi(x_1) \varphi(x_2) \cdots \rangle \right)^2 \Big\rangle, \tag{2.4.12}$$

where $W_R \equiv \sqrt{\pi R^2}^{-3} e^{-|z-z_0|^2/R^2}$. Then $R \to \infty$ implies $\Delta_R \sim \mathcal{O}\left(R^{-3/2}\right)$.

This finally allows us to truly relate cosmological observables to early universe quantum correlators.

The power spectrum

At the level of linear perturbation theory, we are essentially dealing with the theory of a free, single scalar variable, which is completely determined by its two point correlation function. More explicitly, the decomposition in terms of creation and annihilation operators implies Wick's theorem for correlators: everything can be expressed in terms of the two-point correlator. Here we compute the two-point function, known as the power spectrum, from inflation, and explicitly show how it relates to cosmological observables.

From the quantum mechanical point of view, we simply use 2.4.5 to compute the equal time two point correlator,

$$\langle 0|\mathcal{R}(\boldsymbol{x},t)\mathcal{R}(\boldsymbol{y},t)|0\rangle = \int_q |\mathcal{R}_q|^2 e^{i\boldsymbol{q}(\boldsymbol{x}-\boldsymbol{y})}, \tag{2.4.13}$$

which is a function of the norm of the distance only as a consequence of translation and rotation invariance.

From the late-time observational point of view, let us suppose that we have a prescription for relating the field $\mathcal{R}(x,t)$ to some observable, say the local density of photons. We come back to this prescription later. Then, we can do the following: take a fixed distance r, compute the product $\mathcal{R}(x)\mathcal{R}(x+r)$ for a large collection of points x that are nicely distributed over the universe, and take the average. This defines the correlation function $\xi_{\mathcal{R}}(r)$,

$$\xi_{\mathcal{R}}(r) = \langle \mathcal{R}(x,t)\mathcal{R}(x+r,t)\rangle_{x\in\text{Universe}}. \tag{2.4.14}$$

The statement of this section now simply comes down to the statement that for $x-y=r$,

$$\xi_{\mathcal{R}}(r) = \langle 0|\mathcal{R}(x,t)\mathcal{R}(y,t)|0\rangle. \tag{2.4.15}$$

In other words, taking its Fourier transform, we find that the power spectrum $P(k)$, which tells us the variance of the amplitude of fluctuations with wavelength k, is given by

$$P_{\mathcal{R}}(q) = \int d^3r e^{i\boldsymbol{q}\cdot\mathbf{r}}\xi_{\mathcal{R}}(r) = |\mathcal{R}_q|^2. \tag{2.4.16}$$

The power spectrum and inflation

For the scalar field, slow roll example above, we can in principle compute the prediction for the power spectrum given the inflationary potential. The answer for the dimensionless power spectrum, which is a measure of the relative amplitude of perturbations, is given by (e.g. in [13])

$$\Delta^2_{\mathcal{R}}(q) \equiv \frac{q^3}{2\pi^2} P_{\mathcal{R}}(q) = \frac{1}{8\pi^2} \frac{H^2}{\varepsilon M^2_{\rm pl}}\bigg|_{q=aH} , \tag{2.4.17}$$

where the latter expression means we evaluate the time dependent functions at a time given by q, making the time independence of this quantity manifest. Due to the near scale invariance of the inflationary dynamics, this function is mostly captured by its amplitude and slope,

$$\Delta^2_{\mathcal{R}}(q^\star) \sim \frac{1}{\varepsilon} \frac{H_I^2}{M^2_{\rm pl}} , \quad \frac{d\log \Delta^2_{\mathcal{R}}(q)}{d\log k} \equiv n_s - 1 = -2\varepsilon - \eta , \tag{2.4.18}$$

for some pivot scale $q^\star$. At present, these values are the only nonzero measured numbers teaching us about the early universe! A third number, the minimal number of e-folds is derived from the success of the theory of nucleosynthesis. At present, this only teaches us about the combined unknowns H, ε, η. A measurement of the tensor power spectrum would break this degeneracy. However, assuming ε and η are not very different, the observed dimensionless power spectrum of order 10^{-9} combined with $n_s - 1 \approx 0.04$[6], suggests the Hubble rate during inflation to be of order 10^{14}GeV, which means the energy density was of order 10^{16}GeV. This is what makes inflation so interesting from a high energy physics point of view. However, we stress that this relies on the assumption that ε and η are of the same order. In fact, the current upper bound on tensor modes already falsifies this assumption: $\varepsilon < \eta$. For a discussion of the implications of the latter, see for instance [28].

The discussion so far was limited to linear equations of motion for the perturbations. In other words, the two-point statistics are all we have in a

[6] The actual Planck constraints [26] are $10^9 A_s e^{-2\tau} = 1.881 \pm 0.014$, $n_s = 0.9652 \pm 0.0062$, which are defined through $\Delta^2_R(k) = A_s (k/k_p)^{n_s - 1}$ (assuming small scale dependence of n_s), where $k_p = 0.05{\rm Mpc}^{-1}$, and τ is the reionization optical depth, which is smaller than 0.1.

Gaussian universe, but if we include interactions in the theory there is much more information in higher order correlators. This brings us to the topic of non-Gaussianity, which, if observable, can teach us much more about the physics of the early universe.

2.5 Primordial non-Gaussianity

The most obvious consequence of the Gaussian nature of the free theory is that any three point correlator vanishes. However, if we go beyond the free theory, interactions induce so called non-Gaussianity, most directly seen in a non-vanishing three point function. As a schematic example, consider classical nonlinear evolution of the variable $\mathcal{R}$ described above,

$$D_\tau \mathcal{R}(x,\tau) \supset f(\tau)\mathcal{R}^2(x,\tau), \tag{2.5.1}$$

for some second order differential time operator D. Here the $\supset$ sign indicates this is only one of the terms that make up the left hand side. Then the solution for $\mathcal{R}$ can be found perturbatively in the linear solution. It leads to a second order contribution of the form

$$\mathcal{R}^{(2)}(x,\tau) \supset \int d\tilde{\tau} G_D(\tau,\tilde{\tau}) f(\tilde{\tau}) \mathcal{R}^{(1)}(x,\tilde{\tau})\mathcal{R}^{(1)}(x,\tilde{\tau}) = \tilde{f}(\tau)\left(\mathcal{R}^{(1)}(x)\right)^2, \tag{2.5.2}$$

where G_D is the Green's function associated to the differential operator, and using that $\mathcal{R}$ is constant at linear order. Putting this into a three point correlator, and forgetting about the function f for the moment, we find

$$\langle \mathcal{R}(x)\mathcal{R}(y)\mathcal{R}(z)\rangle_c \supset \langle \left(\mathcal{R}^{(1)}(x)\right)^2 \mathcal{R}^{(1)}(y)\mathcal{R}^{(1)}(z)\rangle_c = 2\xi_{\mathcal{R}}(x-y)\xi_{\mathcal{R}}(x-z), \tag{2.5.3}$$

which is obviously nonzero. The subscript 'c' in this expression refers to 'connected', meaning we subtract from the correlator all contributions that can be written as products of lower order (connected) correlators. For a precise definition, see [29]. Most of the correlators in this thesis refer to connected correlators.

More commonly, one studies the three point function in Fourier space. Statistical homogeneity of the above correlator then forces the three point

function in Fourier space to be proportional to an overall Dirac-delta function,

$$\langle \mathcal{R}_{k_1} \mathcal{R}_{k_2} \mathcal{R}_{k_3} \rangle_c \equiv (2\pi)^3 B_{\mathcal{R}}(k_1, k_2, k_3) \delta_D(k_1 + k_2 + k_3). \tag{2.5.4}$$

The function B is called the bispectrum. In the example above, this leads to a contribution to the bispectrum of the form,

$$B_{\mathcal{R}}(k_1, k_2, k_3) \supset 2 P_{\mathcal{R}}(k_1) P_{\mathcal{R}}(k_2). \tag{2.5.5}$$

The permutation symmetry between the momenta is restored by considering all insertions of the quadratic term in 2.5.3. We comment on more general parameterizations of the bispectrum below.

The full computation of the three point function in a particular universe involves much more that the simple example above. Most importantly, we need to go beyond the free theory computation to properly incorporate the nonlinearities in the action for inflationary perturbations. Maldacena was the first to compute them in the slow roll example, see [30]. At present, there are many computations in various models of the primordial bispectrum (i.e. the three-point function for $\mathcal{R}$ before the modes re-enter the Hubble horizon). We will not go into all these models. Instead, we focus on the outcome: the bispectrum for $\mathcal{R}$, as defined in 6.3.28. In the next Chapter we show how the primordial bispectrum in $\mathcal{R}$ is related to the bispectrum of for instance the matter density field. In sections 4.2.1 we provide a generic parametrization of the primordial bispectrum and comment on some models that it captures.

Before concluding, we would like to finish with a brief discussion of the relevance of PNG (meaning the primordial bispectrum) for the physics of inflation. This is nicely summarized in, for instance, [31]. Most prominently, Creminelli and Zaldarriaga [32] proved that in single field, attractor models of inflation, the so called squeezed limit of the bispectrum, in which one of the momenta becomes much smaller than the others, must be given by

$$B_{\mathcal{R}}(q, k, -k - q)_{q/k \to 0} = (n_s - 1) P_{\mathcal{R}}(q) P_{\mathcal{R}}(k). \tag{2.5.6}$$

A detection above this level would therefore rule out the simplest models of inflation. On the other hand, there is a very interesting theoretical bound on the value of the bispectrum when all of the momenta are of the same order. In the squeezed limit, this type of non-Gaussianity (called equilateral non-Gaussianity) decays as $(q/k)^2$, but for other configurations one in principle

needs a specific model to work with. Conveniently, however, for generic single field slow roll models, an effective field theory approach was designed to compute its perturbations [33]. This can be used to show on general grounds that the slow roll theory becomes strongly coupled if the parameter $f_{\rm NL}^{loc}$ is larger than one, see [34, 35]. Here $f_{\rm NL}$ is defined as the dimensionless parameter in front of a template shape (as a function of momenta) for the bispectrum.

Statistical homogeneity and isotropy restrict the bispectrum to be a function of three numbers only. Pictorially, it is a function of the shape of the momentum triangle only. Often, one parametrizes the bispectrum as a function of the norm of two momenta and the angle in between, as we do in 4.2.1. Another way to parametrize the functions is as a function of the norm of all three momenta. This is used in 4.5.1 to define the three shapes we consider for our numerical analysis. For instance, local non-Gaussianity is defined by the following contribution to the bispectrum,

$$B_{\mathcal{R}}(k_1, k_2, k_3) = 2 f_{\rm NL}^{loc} \left(P_{\mathcal{R}}(k_1) P_{\mathcal{R}}(k_2) + \text{perms} \right) . \tag{2.5.7}$$

Observationally, the full bispectrum is then typically matched against a collection of such template functions.[7]

In summary, this chapter shows how PNG can in principle teach us about the single field as well as the slow roll nature of inflation. In the next section we dive into large scale structure. We discuss how $\mathcal{R}$ is related to the density field of matter in LSS, and how we describe the evolution of perturbations as they re-enter the Hubble horizon.

[7] We remark that this may not be the optimal search for non-Gaussianities, see for instance [36].

3 Preliminaries II: Large Scale Structure as an Effective Fluid

In the previous chapter we showed how to relate the gauge invariant perturbation $\mathcal{R}$ to the physics of inflation, how $\mathcal{R}_q$ can be interpreted as a classical stochastic variable after Hubble crossing during inflation, and how it is conserved along the history of the post big bang universe as long as $q \ll \mathcal{H}$. On the other hand, our telescopes observe modes that are inside the horizon by definition. More precisely, we observe density fluctuations inside our observable universe. Therefore, we require a description for what happens to the modes as they re-enter the Hubble horizon. How are they related to density fluctuations, and what is their dynamics?

It turns out that sub-Hubble evolution in the radiation era is quite different from sub-Hubble evolution during matter domination. During radiation domination dark matter perturbations only grow logarithmically, due to the oscillations in the photon-baryon fluid, but as dark matter becomes the dominant energy source, they begin to grow linearly with a.[1] To summarize this history, a transfer function is defined that relates the density perturbations at various scales at the beginning of the matter era to the super-Hubble initial values for $\mathcal{R}_q$. Since the super-Hubble $\mathcal{R}_q$ is time-independent, this definition of the transfer function is independent of initial time.

In this chapter we first briefly derive the transfer function to set up the initial conditions for the subsequent evolution of perturbations during the matter era, which teaches us that linear perturbation theory breaks down at late times and short scales. This serves as the motivation to introduce the EFT of LSS as the only perturbation theory for LSS that consistently takes into account our inability to describe the evolution of perturbations on short scales, by effectively parameterizing its effect on large scales. The introduction of the EFT is the main purpose of this Chapter, and we derive it both from a top-down and a bottom-up perspective.

[1] In flat space they would have grown exponentially with time, but the expansion of the universe causes it to be a power law instead.

3.1 The transfer function: setting the initial conditions for LSS

For the purpose of this thesis, we assume that during matter domination, the density perturbations are purely in the form of cold dark matter, [2]

$$\delta_m(x,\tau) = \frac{\delta\rho(x,\tau)}{\bar{\rho}(\tau)} \equiv \delta(x,\tau). \tag{3.1.1}$$

To connect this to $\mathcal{R}$, we first observe that in Newtonian gauge (see [12]), in the absence of anisotropic stress (implying $\Phi = \Psi$), Einstein's equations relate the density field to the Newtonian potential through the relativistic Poisson equation (see [37])

$$\delta = -\frac{3}{2}\frac{k^2}{\mathcal{H}^2}\Phi - \frac{2}{\mathcal{H}}\Phi' - 2\Phi. \tag{3.1.2}$$

Using the definition of $\mathcal{R}$, and another Einstein equation, the Newtonian potential is in turn related to the comoving curvature perturbation as

$$\mathcal{R} = -\Phi + \frac{\mathcal{H}(\Phi' + \mathcal{H}\Phi)}{4\pi G a^2(\bar{\rho} + \bar{p})}. \tag{3.1.3}$$

To simplify the treatment a bit, let us focus on universes with a fixed equation of state $\bar{p} = w\bar{\rho}$. The equation of motion for Φ is then given by ([37]),

$$\Phi'' + 3(1+w)\mathcal{H}\Phi' + wk^2\Phi = 0, \tag{3.1.4}$$

and the super-Hubble relation between the curvature perturbation and Φ is,

$$\mathcal{R}_i = -\frac{5+3w}{3+3w}\Phi_i\ , \quad \mathcal{R}_i' = \Phi_i' = 0. \tag{3.1.5}$$

This sets the initial condition for both radiation and matter domination. The constancy of $\mathcal{R}$ holds throughout the transition as well, but for Φ, one has to set the initial condition for modes that enter around matter-radiation equality a bit earlier and solve the full equation for Φ during the transition.

[2] The relation of dark matter perturbations to actual observables is nontrivial, but outside of the scope of this thesis.

Using that $\mathcal{H} = \tau^{-1}$ during radiation domination, and $\mathcal{H} = 2/\tau$ during matter domination, one can find the solution for Fourier mode k during radiation domination to be related to the spherical Bessel function $j_1(x)$ (see [37]),

$$\Phi_k = -2\mathcal{R}_{k,i}\left(\frac{\sin x - x\cos x}{x^3}\right) = -2\mathcal{R}_{k,i}\frac{j_1(x)}{x}, \tag{3.1.6}$$

where $x = k\tau/\sqrt{3}$, which indeed reduces to a constant on super-Hubble scales. On sub-Hubble scales, it describes oscillations with decaying amplitude as $\tau^{-2} \sim a^{-2}$.

During matter domination, the dominant mode is simply a constant, even on sub-Hubble scales. Using 3.1.5, one can show that the transition from radiation to matter domination produces a factor 9/10 drop in the potential. This evolution of various modes was nicely summarized in Figure 3.1, taken from [37]. Taking a snapshot at the beginning of matter domination, the

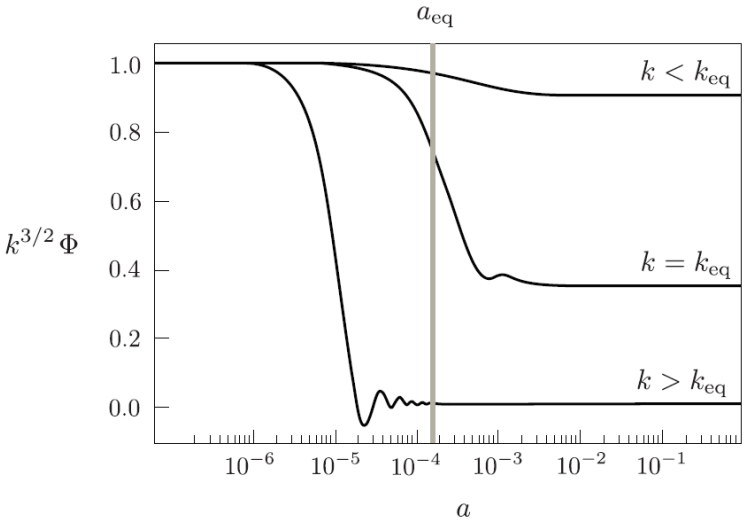


Figure 3.1: The evolution of Fourier modes of the Newtonian potential with the same initial condition, as a function of time. Modes for which $k < k_{eq}$ re-enter the Hubble horizon during the matter era, modes with $k > k_{eq}$ during radiation domination. Image taken from [37].

above allows us to define a transfer function for the matter era initial condi-

tion for the density field. This matter transfer function is defined through,[3]

$$\delta(k, \tau_i) \equiv T_m(k, \tau_i)\mathcal{R}_i(k)\,, \tag{3.1.7}$$

where $\mathcal{R}_i$ is evaluated when the mode was still super-Hubble, and τ_i is some time early in the matter era. The subsequent evolution of the matter density is given by its equations of motion and the growing mode initial condition. The dimensionless initial power spectrum for density fluctuations is then given by

$$\Delta^2(k, \tau_i) = \frac{k^3}{2\pi^2} P(k) = T_m^2(k, \tau_i)\Delta_R^2(k), \tag{3.1.8}$$

which we plot in Figure 3.2.

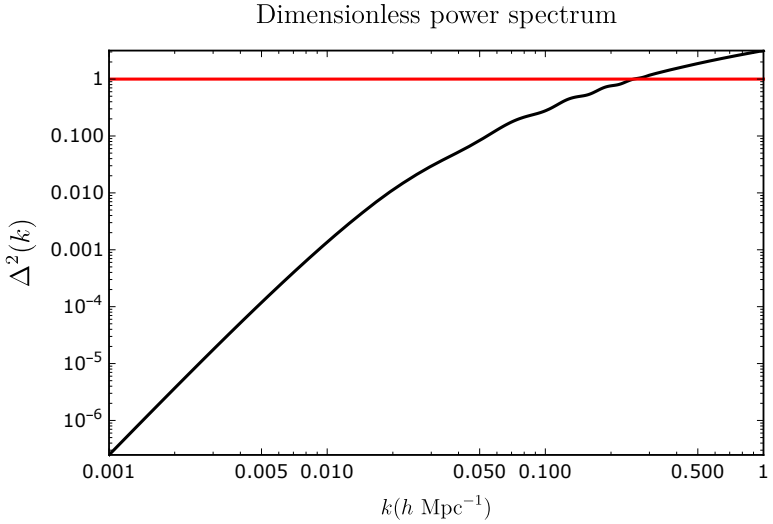


Figure 3.2: Dimensionless power spectrum from CAMB, for a flat ΛCDM cosmology with the following parameters: $\Omega_m^0 = 0.27$, $\Omega_\Lambda^0 = 0.73$, $h = 0.70$. The amplitude of the primordial potential is $\Delta_\varphi^2(k_0) \simeq 8.7 \times 10^{-10}$, defined at the pivot scale $k_0 = 0.002\,h\,\mathrm{Mpc}^{-1}$. The nonlinear regime begins around $k = 0.2\,h\,\mathrm{Mpc}^{-1}$.

[3] Note that we use a slightly different definition in the next Chapters. The transfer function is here implicitly defined by solving the evolution equation all the way to the matter era and comparing the density contrast to the initial perturbation $\mathcal{R}$.

Finally, this finishes our discussion of linear perturbation theory. Next, we study what happens at higher orders in perturbations.

3.2 Beyond linear theory

The effective field theory of large scale structure (EFT of LSS) is designed to accurately describe the evolution of matter density perturbations (which we assume to be cold) beyond the linear level. Let us spend some words on why we are interested in this in the first place.

Linear theory for the evolution of the density contrast is simply the combination of the Poisson equation (3.1.2) and evolution of the potential 3.1.4, which were summarized in the transfer function 3.1.7. Now, if linear theory were in fact correct all the way to the present, the dimensionless power spectrum today would be given by 3.1.8, upon replacing τ_i with τ_0, plotted in Figure 3.2.

However, one expects the dimensionless power spectrum to be a good measure of the expansion parameter in LSS: when it becomes of order one, nonlinear corrections are expected to start dominating over the linear solution. The 'nonlinear' $k_{\rm NL}$ at which this happens is estimated as,

$$\Delta^2(k_{\rm NL}) \equiv 1. \tag{3.2.1}$$

Moreover, in the quasi-linear regime, nonlinear corrections are expected to modify the linear results at order, [4]

$$\text{Error} \sim \frac{\text{Non-linear}}{\text{Linear}} \sim \Delta^2(k). \tag{3.2.2}$$

This is the reason we need to go beyond linear theory. More concretely, in 4 we show that in order to learn about primordial non-Gaussianity (PNG), we really have to control our errors to a very high precision.

From Figure 3.2, it is obvious that these errors become more and more significant as we go to higher and higher wavenumbers. One might therefore

[4] A more accurate estimation of the size of nonlinear corrections in our universe was performed in [38]. Namely, the presence of the baryon acoustic oscillation (BAO) scale in the initial conditions, renders an IR resummation necessary. Since this in independent of the UV sensitivity that the EFT deals with, we do not comment on it here. The IR resummed theory can be found in [39, 40].

wonder why we do not just stick to very large scales, at which linear theory is a very good approximation. The answer is that we only learn about cosmology from statistical averages: the observation of a single mode does not teach us anything. It is only upon averaging over space (or modes) that we find robust patterns. The more modes we can describe, the more modes we can include in our analysis, and the better our constraints on cosmological parameters (such as PNG) becomes. Roughly, the statistical error is expected to go down as $N^{-1/2}$, where for three dimensional LSS surveys the number of modes N scales as $(k_{max}/k_{min})^3$. The full analysis in the context of PNG is the topic of Chapter 5, in which we show that it is indeed challenging to include enough modes in our theory. Before getting there, we introduce the means to describe the mildly nonlinear regime: the EFT of LSS.

3.3 The Effective Field Theory of Large Scale Structure

There are various approaches to the evolution of density perturbations during the matter era beyond linearity. Most of them are some sort of extension of what is called Standard Perturbation Theory (SPT). The purpose of this section is to argue that, regardless of its eventual practical usefulness, the EFT of LSS is the *only* description of LSS that consistently parameterizes our ignorance of the dynamics beyond the nonlinear scale.

As for most effective theories in physics, there are two perspectives on it: a so called top-down or bottom-up perspective. The top-down perspective is useful when the short scale ('UV') theory is known, but complicated to solve, in which case one 'integrates out' degrees of freedom that are irrelevant for the description on much larger scales. The bottom-up perspective is useful to write down a large scale theory that consistently takes into account our ignorance of what is happening on shorter scales. In the case of LSS we have access to both perspectives, and they unambiguously lead to the EFT of LSS as the only consistent perturbation theory for LSS.

We present both perspectives here, starting with the bottom-up approach.

3.3.1 Bottom up: smoothing nonlinear equations

From a bottom up perspective, the starting point are the perturbations on top of the homogeneous and isotropic FLRW background. Going beyond linearity is complicated for at least two reasons: a) GR is nonlinear in nature, which in principle induces an infinite set of interactions for the perturbations, and b) beyond linearity different Fourier modes couple, which makes it an infinite set of coupled, highly nonlinear equations. In order to deal with them, a physicist often resorts to perturbation theory. Our case is no exception.

Newtonian approximation

The first approximation applicable to the evolution of cold dark matter (which can be considered a proxy for non-interacting, pressureless, classical matter moving at non-relativistic speeds) in the matter era is the expansion in the smallness of velocities. We follow the arguments of [11]. Crucial to this expansion is the observation that, whereas the density contrast in our universe is much larger than unity on small scales, the Newtonian potential (working in Newtonian gauge) is always of order 10^{-5}, except in the vicinity of black holes [5].

We now argue that an expansion in the peculiar velocity v^i is valid as well. First, one can prove the self consistency in perturbation theory, yielding $v^2 \sim \phi\delta$ on all scales, i.e. v^2 is of order 10^{-10} on large scales, growing to order 10^{-5} around the nonlinear scale. Moreover, for virialized systems we also find $v^2 \sim \phi$. Alternatively, one can simply measure galactic velocities. Altogether, we conclude that $v^2 \lesssim \phi$ on all scales. Thus, we perturb in velocities as well.

Since vectors and tensors produced during inflation are typically small and decay on sub-Hubble scales in linear theory, we neglect them in this treatment. Finally, we restrict ourselves to perturbations whose wavelength is much shorter than the Hubble horizon. Assuming the matter content of the universe is non-interacting, non-relativistic (cold dark) matter, these approximations allow us to write down the conservation and Poisson equations

[5] For a discussion on gauge dependence of this discussion, or on how to deal with order one metric fluctuations near black holes, we refer to the paper [11].

in the Newtonian approximation[6],

$$(\partial_\tau + \boldsymbol{v}\cdot\boldsymbol{\nabla})\,\delta = -(1+\delta)\boldsymbol{\nabla}\cdot\boldsymbol{v}\ , \tag{3.3.1}$$

$$(\partial_\tau + \boldsymbol{v}\cdot\boldsymbol{\nabla})\,\boldsymbol{v} = -\mathcal{H}\boldsymbol{v} - \boldsymbol{\nabla}\phi\ , \tag{3.3.2}$$

where ϕ is the gravitational potential, which satisfies the Poisson equation

$$\Delta\Phi \equiv \Delta\phi = \frac{3}{2}\mathcal{H}^2\Omega_m\delta\ . \tag{3.3.3}$$

Here $\Omega_m = \rho_m(\tau)/\bar{\rho}_{tot}(\tau)$ is the time dependent fractional energy density of matter, which is equal to unity until dark energy starts to set in. Also note the discontinuity in the use of capital letters for the Newtonian potential compared to the previous chapter (which was not to confuse notation with the scalar field $\phi(t)$). These equations form the starting point of most perturbative treatments of LSS.

Standard Perturbation Theory

Since the perfect fluid equations of motion are nonlinear partial differential equations, generic closed form solutions are not known. Standard perturbation theory (SPT) is a first attempt to deal with them. We introduce this method first, as the EFT is only a slight modification of it. We comment on its shortcomings afterwards.

The first observation to make is that vorticity, $w = \boldsymbol{\nabla}\times\boldsymbol{v}$, decays in linear theory. Moreover, in the absence of a stress tensor, as in 3.3.1, one can show that if it starts out as zero, it remains zero. Even though the EFT eventually does introduce stress, which indeed sources vorticity, we can safely neglect it for questions about the density as we are after in this thesis. Thus we can write the equations of motion in terms of the density δ and the velocity divergence $\theta \equiv \boldsymbol{\nabla}\cdot\boldsymbol{v}$,

$$\partial_\tau\delta + \theta = -\delta\theta - \frac{\partial_i\theta}{\partial^2}\partial_i\delta, \tag{3.3.4}$$

$$(\partial_\tau + \mathcal{H})\theta + \frac{3}{2}\Omega_m\mathcal{H}^2\delta = -\left(\frac{\partial_i\partial_j\theta}{\partial^2}\right)^2 - \frac{\partial_i\theta}{\partial^2}\partial_i\theta, \tag{3.3.5}$$

[6] See [4] for a rigorous definition of the Newtonian limit. It is for instance shown how the fields defined here are related to general relativistic perturbations in Newtonian gauge. See also [12] for the more general discussion of (gauge invariant) perturbations.

where the nonlocal operations are clear once we specify initial conditions and move to Fourier space. Together, these equations form a coupled set of first order equations with some nontrivial derivative structure. The full perturbative treatment can be found in 4.C. For the purpose of this introduction, let us qualitatively study the behavior of a theory like this. Combining the two equations, one obtains a second order equation for the density field, which schematically reads,

$$\mathcal{D}_\tau \delta(x,\tau) \simeq \delta(x,\tau)^2, \tag{3.3.6}$$

where $\mathcal{D}_\tau$ is some second order differential time operator, and we have assumed the simplest possible nonlinear interaction for convenience.

Now, SPT solves these equations perturbatively by assuming $\delta(x,\tau) \ll 1$. Using a Green's function G, for which

$$\mathcal{D}_\tau G(\tau,\tau') = \delta_D(\tau-\tau'), \tag{3.3.7}$$

the equation of motion can then be rewritten in integral form as

$$\delta(x,\tau) = \delta_L(x,\tau) + \int_{\tau'} G(\tau,\tau')\delta(x,\tau)^2, \tag{3.3.8}$$

where the linear solution satisfies the equations $\mathcal{D}_\tau \delta_L = 0$. Generically, there are two such solutions, depending on the initial conditions. We have selected the growing mode, which matches the adiabatic mode initial condition. Notice that the contact interaction translates to a convolution in Fourier space,

$$\delta(k,\tau) = \delta_L(k,\tau) + \int_{\tau'} G(\tau,\tau') \int_q \delta(k-q,\tau)\delta(q,\tau). \tag{3.3.9}$$

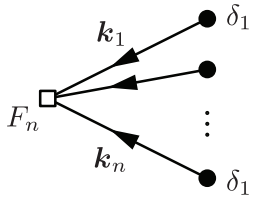
This shows explicitly how a contact interaction leads to contributions of arbitrarily short scales (large q) to the evolution of any mode k, even in a classical theory. In a matter dominated universe, the time dependences are nice enough that the perturbative solution becomes

$$\delta(\boldsymbol{k},\tau) = \sum_{n=1}^{\infty} a^n(\tau)\delta_n(k), \tag{3.3.10}$$

where the generic expression is,

$$\delta_n(\boldsymbol{k}) = \int_{q_1,\ldots,q_{n-1}} F_n(k-q_s, q_1, \ldots, q_{n-1})\delta_1(k-q_s) \times \ldots \times \delta_1(q_{n-1}), \tag{3.3.11}$$

where $\delta_1(k) \equiv \delta_L(k,\tau)/a(\tau)$ and $q_s \equiv \sum_i q_i$. In the simplified example the F_n are just numbers, but generically they are some complicated function of momenta, see Appendix 4.C. Physically, this can be thought of as the interaction of several linear modes δ_1 as they propagate in time and space. Diagrammatically, we can represent this as



$$\equiv \quad F_n(\boldsymbol{k}_1, \cdots, \boldsymbol{k}_n)\,(2\pi)^3\delta_D\big(\boldsymbol{k} - \boldsymbol{k}_{1\ldots n}\big)\ . \tag{3.3.12}$$

Since in cosmology we are interested statistical averages, as elaborated on in the previous chapter, we mainly use this theory to compute correlators: averages over the ensemble of initial conditions. In a perturbative expansion, and assuming Gaussian initial conditions (such that Wick's theorem applies) the power spectrum can be written as

$$P_\delta(k) \equiv \langle\delta(\boldsymbol{k},\tau)\delta(-\boldsymbol{k},\tau)\rangle' = \underbrace{P_{11}(k)}_{\text{tree}} + \underbrace{P_{13}(k) + P_{22}(k)}_{\text{one-loop}} + \cdots\ , \tag{3.3.13}$$

where $P_{mn}(k) \equiv (\langle\delta_{(m)}(\boldsymbol{k},\tau)\delta_{(n)}(-\boldsymbol{k},\tau)\rangle' + \text{perm.})$. The prime denotes the suppression of the overall momentum conserving Dirac-delta function times $(2\pi)^3$, which is guaranteed to multiply any correlator as a consequence of statistical homogeneity. Notice also that in order to keep later expressions compact we have dropped the explicit time arguments in $P_\delta(k)$ and $P_{mn}(k)$. Unless stated otherwise, density correlators should always be understood as being evaluated at time τ. As an example, the explicit formula for P_{13} is given by,

$$P_{13} = 6P_{11}(k)\int_q F_3(k,q,-q)P_{11}(q). \tag{3.3.14}$$

Notice the appearance of the 'loop' integral. We will also be interested in the late-time matter bispectrum which, for Gaussian initial conditions, has the following loop expansion

$$B_\delta(k_1,k_2,k_3) \equiv \langle \delta(\boldsymbol{k}_1,\tau)\delta(\boldsymbol{k}_2,\tau)\delta(\boldsymbol{k}_3,\tau)\rangle' = \\ = \underbrace{B_{112}}_{\text{tree}} + \underbrace{B_{222} + B^{(\text{I})}_{123} + B^{(\text{II})}_{123} + B_{114}}_{\text{one-loop}} + \cdots , \tag{3.3.15}$$

where $B_{lmn}(k_1,k_2,k_3) \equiv (\langle \delta_{(l)}(\boldsymbol{k}_1,\tau)\delta_{(m)}(\boldsymbol{k}_2,\tau)\delta_{(n)}(\boldsymbol{k}_3,\tau)\rangle' + \text{perms})$. The diagrammatic representations of the one-loop contributions to the power spectrum and bispectrum are shown in fig. 3.3. (See [29] for a more complete description of the Feynman rules of SPT.) In these diagrams, every square represents a density mode at late times, formed out of several initial density modes (the arrows). The dots indicate which modes in the initial conditions are Wick contracted to form the statistical average. Interestingly, we find that loops naturally arise in this statistical field theory. We stress that this does not have anything to do with the quantum origin of perturbations. It is merely a manifestation of the fact that in a nonlinear field theory, modes of different momenta couple, and for our statistical initial conditions, we expect all of them to contribute to the evolution of a particular mode k.

The trouble with SPT

The fact that SPT has to be modified for theoretical consistency can be seen from (at least) the following arguments, all of which are related in some sense:

- The density contrast in not small on all scales. As we showed, in a nonlinear theory, modes of all wavelengths couple to a particular long mode k. The perturbative expansion of perfect fluid equations is therefore not well motivated.

- To elaborate on the previous point: the theory is sensitive to small scale physics, whose physics we do not control analytically. On the one hand this can be seen from the loops running over arbitrarily high momenta. On the other hand, in real space this can be seen from the contact interactions.

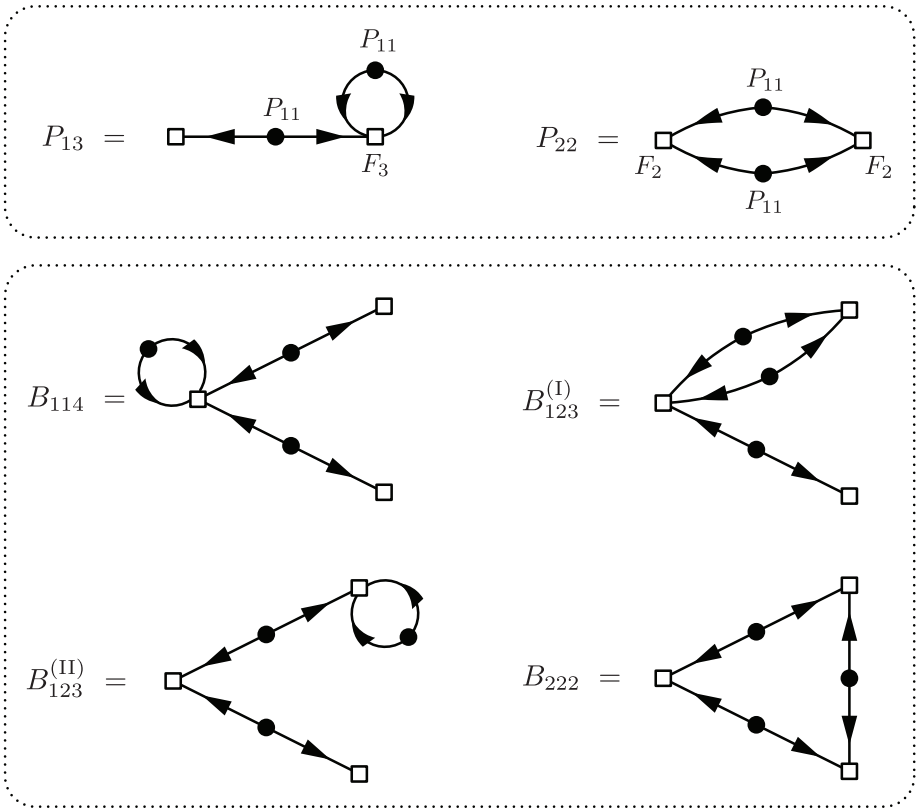


Figure 3.3: Diagrammatic representation of the one-loop power spectrum and the one-loop bispectrum for Gaussian initial conditions.

- Moreover, for generic initial conditions the loops are divergent, rendering the theory unpredictive. This happens not to be the case for our universe, but it signals the theoretical inconsistency of the theory. More physically, it tells us that we are applying the theory beyond its scope.

- The theory does not take into account deviations from a perfect fluid. This was an assumption that works well in linear theory and on large scales, but we know that on short (cluster) scales, the motion of matter cannot be described as a perfect fluid. This fact should be incorpo-

rated in the theory: we expect the perfect fluid description to become worse as we approach short scales. This is not apparent from the SPT equations.

- Finally, the theory is not 'closed under renormalization', to use common high energy terminology. In this case, it means that if we try to fix some of the problems above by introducing a cutoff, the theory becomes cutoff dependent. In particular, some cutoff-dependent terms appear in the solution that do not find their counterpart in the equations of motion.

The first step to address these issues is to realize that from a bottom-up perspective, the density and velocity only make sense as smoothed quantities.

Smoothing the equations: appearance of the stress tensor

The first, rather intuitive, step is to smooth the equations in real space in order to define spatially averaged values of the density and velocity. A smoothing function $W_\Lambda(|\boldsymbol{x}-\boldsymbol{x}'|)$ is a normalized (to unity) function that is equal to unity when $|\boldsymbol{x}-\boldsymbol{x}'| \ll \Lambda^{-1}$ and quickly goes to zero for $|\boldsymbol{x}-\boldsymbol{x}'| > \Lambda^{-1}$. A common choice is the Gaussian window function

$$W_{\Lambda,G}(\boldsymbol{x}) \equiv \frac{1}{(2\pi\Lambda^{-2})^{3/2}} e^{-\frac{1}{2}\boldsymbol{x}^2\Lambda^2}. \tag{3.3.16}$$

Smoothed variables are then defined as

$$X_\ell(\boldsymbol{x}) \equiv [X]_\Lambda(\boldsymbol{x}) \equiv \int \mathrm{d}^3\boldsymbol{x}'\; W_\Lambda(|\boldsymbol{x}-\boldsymbol{x}'|)\, X(\boldsymbol{x}')\ . \tag{3.3.17}$$

This automatically fixes the first SPT problem: there is a well-defined expansion parameter if we choose $\Lambda < k_{\rm NL}$. One way to see this is that for the Gaussian smoothing kernel, the power spectrum of the density is exponentially suppressed for $k > \Lambda$,

$$P_{\delta_l}(k) = P_\delta(k) e^{-\frac{k^2}{\Lambda^2}}. \tag{3.3.18}$$

In other words, the dimensionless power spectrum is always smaller than unity and is therefore a good expansion parameter. Smoothing does introduce two new issues though: the theory becomes artificially dependent on

the cutoff Λ, and it turns out that a product of fields smoothed is not equal to the product of smoothed fields (i.e. smoothing and multiplication do not commute). We first comment on the latter. Let us define 'short fields' through $f_s \equiv f - f_l$, $g_s \equiv g - g_l$ for some fields f, g. The smoothed product then becomes

$$[fg]_\Lambda = [f_l g_l]_\Lambda + [f_l g_s]_\Lambda + [f_s g_l]_\Lambda + [f_s g_s]_\Lambda \,. \tag{3.3.19}$$

As shown in [11], the first three terms can be rewritten as

$$[f_l g_l]_\Lambda + [f_l g_s]_\Lambda + [f_s g_l]_\Lambda = f_l g_l + \frac{1}{\Lambda^2}\left[\nabla f_l \nabla g_l + \ldots\right] + \ldots \tag{3.3.20}$$

where the terms in brackets consist of the different distributions of the two derivatives on the fields, and the latter dots denote higher orders in Λ^{-1}. Restricting our theory to modes with $k < \Lambda$, these derivative corrections are suppressed. The short-short contribution cannot be simplified. This is once again a manifestation of the fact that products of short scale fields can have long wave support. Let us apply the smoothing to the Euler and continuity equation 3.3.1. For convenience, we define the smoothing of the velocity field in terms of the momentum density field

$$\rho_l \boldsymbol{v}_l \equiv \boldsymbol{\pi}_l = [\boldsymbol{\pi}]_\Lambda = [\rho \boldsymbol{v}]_\Lambda \,. \tag{3.3.21}$$

Subtleties of this approach were analyzed in [41], but have no further implications for this work. This definition is convenient if we rewrite the continuity and Euler equation as

$$\begin{aligned} \partial_\tau \tilde{\rho} + \boldsymbol{\nabla}\left(\tilde{\rho}\boldsymbol{v}\right) &= 0 \\ \partial_\tau\left(\tilde{\rho}\boldsymbol{v}^i\right) + \mathcal{H}\left(\tilde{\rho}\boldsymbol{v}^i\right) + \boldsymbol{\nabla}_j(\tilde{\rho}\boldsymbol{v}^i\boldsymbol{v}^j) &= -\tilde{\rho}\boldsymbol{\nabla}_i\phi, \end{aligned} \tag{3.3.22}$$

where $\tilde{\rho} \equiv 1 + \delta$. To conveniently smooth these equations, observe that by applying partial integration and the relation $\partial_x W_\Lambda(x - x') = -\partial_{x'} W_\Lambda(x - x')$, a total derivative can always be taken out of the smoothing operation. The only nontrivial term to smooth is the force term $\tilde{\rho}\boldsymbol{\nabla}_i\phi$, but as shown in [11], using the Poisson equation and some integrations by parts, we find

$$\begin{aligned} [\tilde{\rho}\boldsymbol{\nabla}_i\phi]_\Lambda =& \tilde{\rho}_l \boldsymbol{\nabla}_i\phi_l + \boldsymbol{\nabla}_j\left[\frac{1}{8\pi G a^2 \bar{\rho}}\left(2\partial^i\phi_s\partial^j\phi_s - \delta^{ij}(\partial_k\phi_s)^2\right)\right]_\Lambda \\ &+ \mathcal{O}\left(\frac{\boldsymbol{\nabla}^2}{\Lambda^2}\tilde{\rho}_l\boldsymbol{\nabla}\phi_l\right), \end{aligned} \tag{3.3.23}$$

where ϕ_l and ϕ_s are straightforwardly related to the density field through smoothing the linear Poisson equation. Altogether, we find that the continuity equation is unchanged, whereas the the Euler equation gets modified,

$$\begin{aligned} \partial_\tau \tilde{\rho}_l + \boldsymbol{\nabla}\left(\tilde{\rho}_l \boldsymbol{v}_l\right) &= 0 \\ \partial_\tau\left(\tilde{\rho}_l \boldsymbol{v}_l^i\right) + \mathcal{H}\left(\tilde{\rho}_l \boldsymbol{v}_l^i\right) + \boldsymbol{\nabla}_j(\tilde{\rho}_l \boldsymbol{v}_l^i \boldsymbol{v}_l^j) &= -\tilde{\rho}_l \boldsymbol{\nabla}_i \phi_l - \boldsymbol{\nabla}_j \tau_\Lambda^{ij}, \end{aligned} \tag{3.3.24}$$

where the new stress tensor

$$\begin{aligned} \tau_\Lambda^{ij} = &\left[\frac{1}{8\pi G a^2 \bar{\rho}}\left(2\partial^i\phi_s\partial^j\phi_s - \delta^{ij}(\partial_k\phi_s)^2\right)\right]_\Lambda + \left[\tilde{\rho}_s \boldsymbol{v}_s^i \boldsymbol{v}_s^j\right]_\Lambda \\ &+ \mathcal{O}\left(\frac{1}{\Lambda^2}\frac{\boldsymbol{\nabla}^2\phi_l \boldsymbol{\nabla}^2\phi_l}{8\pi G a^2\bar{\rho}}, \frac{\tilde{\rho}_l \boldsymbol{\nabla}\boldsymbol{v}_l^i \boldsymbol{\nabla}\boldsymbol{v}_l^j}{\Lambda^2}\right). \end{aligned} \tag{3.3.25}$$

As such, we find that smoothing introduces a new term in the equations of motion, which is cutoff dependent and cannot be computed in the long wavelength theory as it explicitly depends on the behavior of the fields on short scales. The upside is that we have obtained a theory that manifestly consists of fields only fluctuating on large scales. Since we observe small deviations from a perfect fluid on large scales, we can also assume the tensor is small on large scales. This means that we do have perturbative control over this theory. The next step is to deal with the cutoff dependence and interpret this new stress tensor physically. This is where the EFT comes in.

The EFT: local expansion of the stress tensor

Let us forget about the higher derivative contributions to the stress tensor for the moment. Later, we show how these terms are naturally retrieved from the way we deal with the remaining part of the stress tensor later. Naively, the stress tensor is some unknown function of space and time that influences our theory in an unpredictable manner. Here we show that there is still a way to make progress. The crucial observation is that, in the absence of PNG, the stress tensor is local in the following sense.

Let us first suppose there are no interactions between the long wavelength fields in our theory and the short fields that make up the stress tensor. In that case, we declare that whatever the stress tensor is, its self-correlations are local, i.e. its self-correlations are proportional to a spatial Dirac delta

function or derivatives theoreof,

$$\langle \tau_\Lambda^{ij}(\boldsymbol{x})\tau_\Lambda^{kl}(\boldsymbol{y})\rangle = c^{ijkl}(\Lambda)\delta_D(\boldsymbol{x}-\boldsymbol{y}) + d^{ijklm}(\Lambda)\frac{\boldsymbol{\nabla}^m_{(\boldsymbol{x}-\boldsymbol{y})}}{\Lambda}\delta_D(\boldsymbol{x}-\boldsymbol{y}) + \dots \tag{3.3.26}$$

In the next section, where we derive the effective theory from the top down, we show that this locality is physically motivated by the fact that dark matter particles typically travel less than the nonlinear distance, such that the nonlinear scale acts as an effective mean free path. From the bottom-up perspective, we simply assume by construction that there are no new degrees of freedom that have long distance correlations. This is the analogue of integrating out heavy degrees of freedom in more common effective quantum field theories (see for instance [42]). The local character of these self-correlations is the reason the backreaction of the stress tensor on the equations of motion can be neglected for long mode correlators of the density field. Moreover, as a consequence of mass and momentum conservation, the stress tensor appearance in the equation of motion for the density field always appears with two additional spatial derivatives, which makes it even more subleading on large scales.[7]

The true relevance of the stress tensor comes from the observation that the theory is interacting, which means the short scales are influenced by long modes dynamically. We stress however, that this means they are influenced by *locally observable* long modes. For instance, (generalized) Galilian invariance and the equivalence priciple guarantee that long velocity or acceleration modes are not locally observable. This allows for an expansion of the stress tensor in terms of the locally observable long modes.

Before explaining this expansion, though, we highlight one more unusual aspect of our theory: it is non-local in time. The reason is that the timescale for short-scale evolution (e.g. gravitational collapse) is also Hubble - there is no hierarchy in time scales. This means that the stress tensor at position $\boldsymbol{x}$ and time τ can not only be influenced by the density field at that position and time, but also by the density field at some earlier time at the earlier position of the 'fluid' element. This fluid position is calculated by simply tracing the evolution of a hypothetical particle at this position back in time

[7] For a more elaborate discussion and analytical analysis of these points for particular initial conditions, see [41, 43].

along the long mode velocity field. It is implicitly defined as follows,

$$\boldsymbol{x}_{fl}[\boldsymbol{x},\tau;\tau'] = \boldsymbol{x} - \int_{\tau'}^{\tau} \boldsymbol{v}_l(x_{fl}[\boldsymbol{x},\tau;\tau''],\tau'')d\tau'' \; ; \quad \boldsymbol{x}_{fl}[\boldsymbol{x},\tau;\tau] = \boldsymbol{x}. \quad (3.3.27)$$

For future reference, we also define the Lagrangian coordinate $\boldsymbol{q} \equiv \boldsymbol{x}_{fl}(\tau_i)$. We now claim that the stress tensor generically is some functional of contractions and powers of the gravitational tidal tensor,

$$\tau^{ij}_{\Lambda}(\boldsymbol{x},\tau) = \mathcal{F}\left[\boldsymbol{\nabla}_i\boldsymbol{\nabla}_j\phi(\boldsymbol{x}_{fl},\tau')\right]. \quad (3.3.28)$$

A full treatment of the non-locality in time can be found in [44], as the stress tensor can be interpreted as a biased tracer as well. At the order we will be interested in, it was shown that by using the equations of motion, the theory can effectively be written in a local in time form as well (see also [45, 46]). Keeping in mind that eventually all we care about are long mode correlators, let us decompose

$$\tau^{ij}_{\Lambda}(\boldsymbol{x},\tau) = \langle\tau^{ij}_{\Lambda}(\boldsymbol{x},\tau)\rangle_s + \Delta\tau^{ij}_{\Lambda}(\boldsymbol{x},\tau), \quad (3.3.29)$$

where the index short refers to a statistical average over the short mode initial conditions, which is spatially independent. The difference, which we refer to as stochastic noise, precisely has the self-correlation properties discussed above. Using the equations of motion, we can then expand the averaged stress tensor in derivatives and powers of fluctuations as,

$$\begin{aligned}\langle\tau^{ij}_{\Lambda}\rangle_s(\boldsymbol{x},\tau) =& \frac{p(\tau)}{\bar{\rho}}\delta^{ij} + c_s^2(\tau,\Lambda)\delta_l - \frac{c^2_{bv}(\tau,\Lambda)}{\mathcal{H}}\boldsymbol{\nabla}\cdot\boldsymbol{v}_l\delta^{ij} \\ &- \frac{3}{4}\frac{c^2_{sv}(\tau,\Lambda)}{\mathcal{H}}\left(\boldsymbol{\nabla}^i\boldsymbol{v}_l^j + \boldsymbol{\nabla}^j v_l^i - \frac{2}{3}\delta ij\boldsymbol{\nabla}\cdot\boldsymbol{v}_l\right) + \dots \,, \quad (3.3.30)\end{aligned}$$

where the dots stand for higher derivatives and powers of the fields. Here, the coefficients are functions of time only as a consequence of the short scale average. Here we have written the expansion in terms of the long mode fluid variables to give a fluid interpretation, but they are all related to the tidal field through the equations of motion. Thus we see that the EFT perspective naturally introduces corrections to the perfect fluid in the form a speed of sound and viscosity coefficients. Moreover, we have added additional, cutoff dependent parameters to the theory, which have the right structure to cancel cutoff dependent terms from SPT loops (after smoothing the initial power spectrum). This is what we show next.

Renormalization

The renormalization of the EFT of LSS is most pedagogically done for scaling universe initial conditions ($P_{11}(k) \sim k^n$), for which, in an Einstein-de Sitter universe the non-linear scale is the only physical scale that appears. This was extensively treated in [43]. Here we just show the simplest example of how it works.

Let us consider the contribution $P_{13}(k)$ to the power spectrum for $k \ll \Lambda$ in the smoothed theory,

$$P(k) \supset P_{13,\Lambda} = 6P_{11,\Lambda}(k) \int_q F_3(k, q, -q) P_{11,\Lambda}. \tag{3.3.31}$$

In the new theory, we can plug the stress tensor in the equations of motion to find its contribution to the density in Fourier space, $k^2 c^2(\tau, \Lambda)\delta_l(k, \tau)$, where c is some combination of the speed of sound and viscosity terms above. This leads to a new contribution to the power spectrum of the form

$$P(k) \supset k^2 c^2(\Lambda) P_{11,\Lambda}. \tag{3.3.32}$$

Combining them, we find a contribution to the power spectrum of the form

$$P(k) \supset 6P_{11,\Lambda}(k) \int_q F_3(k, q, -q) P_{11,\Lambda} + k^2 c^2(\Lambda) P_{11,\Lambda}. \tag{3.3.33}$$

First note that to leading order in an expansion in k/Λ, $P_{11,\Lambda} = P_{11}(k)$. Then, the only remaining cutoff dependence is in the coefficient c, and some nontrivial cutoff dependence coming from the loop. The key observation is that the momentum dependence of this new term turns out to be exactly the same as the momentum dependence of the loop. This is again a consequence of mass and momentum conservation,

$$\lim_{q \to \infty} F_3(k, -q, q) \propto \frac{k^2}{q^2}, \tag{3.3.34}$$

which guarantees that for $q \gg k$,

$$\int_{q \gg k} F_3(k, q, -q) P_{11,\Lambda} \sim k^2 \int_{q \gg k} \frac{1}{q^2} P_{11,\Lambda} = k^2 f(\Lambda), \tag{3.3.35}$$

for some function f of dimension Λ^{-2}. In a scaling universe, the separation of the integral into a finite and (cutoff dependent) divergent part can be done more formally using for instance dimensional regularization. Here we have shown how the new terms from the stress tensor cancel the cutoff dependence from the loop. As always, we should also allow for a finite contribution to the free coefficient c. On dimensional grounds, this should go as length squared. The physical scale that sets this length has to be $k_{\rm NL}^{-1}$ - the distance of influence of the short scales that we have integrated out. We stress however that the only measurable quantity is the sum of the loop and 'counterterm'. Matching measurement with theory at some scale Λ then allows us to extract $c(\Lambda)$, after which the theory is fixed and independent of the cutoff. For details, see [47]. We comment on the time dependence of the counterterms more elaborately in Chapter 4.

Altogether, we have argued how the bottom-up approach led us to include an expanded stress tensor into the theory for LSS perturbations, that accounts for our ignorance of non-perturbatuive short scale physics, and leads to dynamics beyond the perfect fluid approximation. Moreover, it leads to a well-defined, consistent perturbation theory that successfully addresses all issues with SPT addressed above. Importantly, it signals its own breakdown as we try to apply the theory beyond the nonlinear scale, as at that point an infinite tower of operators (products of fields and derivatives) is expected to become relevant. In the next section, we explain from a top-down perspective how this theory is understood as a gravitationally induced fluid from taking moments of the Boltzmann equation.

3.3.2 Top-down: the gravitational fluid

The 'UV' theory for LSS is a large collection of dark matter particles moving about due to their mutual gravitational interaction, for which we neglect (general) relativistic effects. The full statistical information in such a theory is contained in the collection of multi-particle distribution functions, $f_N[(\boldsymbol{x}_1, \boldsymbol{p}_1), \ldots, (\boldsymbol{x}_N, \boldsymbol{p}_N)]$, whose ensemble averages yield the joint probability of finding N particles with those coordinates, see [4]. Each of them satisfy a Liouville equation that describes the conservation of particle number in phase space. Ensemble averaging those equations finally results in the BBGKY equations for the joint probability densities. Conveniently, Inagaki ([48]) showed that if we assume that the gravitational acceleration of a

particular particle is substantially influenced by many others, the BBGKY system reduces to the fluid limit: a single Liouville equation (the Vlasov equation) for a continuous single particle distribution function $f(\boldsymbol{x}, \boldsymbol{p}, \tau)$, whose correlations are related to probabilities as,

$$dP_{12} = \langle f(1)f(2)\rangle d^3\boldsymbol{x}_1 d^3\boldsymbol{x}_2 d^3\boldsymbol{p}_1 d^3\boldsymbol{p}_2, \tag{3.3.36}$$

where the average is over the statistical ensemble, and dP_{12} is the joint probability of finding a particle at 1 and 2 simultaneously. The Vlasov equation is given by,

$$\frac{df}{d\tau} = \frac{\partial f}{\partial \tau} + \frac{\boldsymbol{p}}{am} \cdot \boldsymbol{\nabla} f - am \boldsymbol{\nabla}\phi \cdot \frac{\partial f}{\partial \boldsymbol{p}} = 0, \tag{3.3.37}$$

where m is the mass of a particle, $\boldsymbol{p} = am\boldsymbol{v}$,

$$1 + \delta(\boldsymbol{x}, \tau) = \frac{m}{\bar{\rho}_m a^3} \int f(\boldsymbol{x}, \boldsymbol{p}, \tau) d^3p, \tag{3.3.38}$$

where $\bar{\rho}_m = \Omega_m \bar{\rho}$, and the potential is given by the Poisson equation 3.3.3 as usual. Real space equations of motion are the moments of the Vlasov equation

$$\int d^3p \, \frac{p_{i_1}}{am} \ldots \frac{p_{i_n}}{am} \frac{df}{d\tau}(\boldsymbol{x}, \boldsymbol{p}, \tau) = 0. \tag{3.3.39}$$

In principle, this leads to an infinite set of equations for moments of the distribution function. The zeroth moment is the relation to the density above, and the first and second are given by

$$(1 + \delta)\boldsymbol{v} = \frac{m}{\bar{\rho}_m a^3} \int f(\boldsymbol{x}, \boldsymbol{p}, \tau) \frac{\boldsymbol{p}}{am} d^3p \, , \tag{3.3.40}$$

$$(1 + \delta)\sigma_{ij} = \frac{m}{\bar{\rho}_m a^3} \int f(\boldsymbol{x}, \boldsymbol{p}, \tau) \frac{p_i}{am} \frac{p_j}{am} d^3p - (1 + \delta) v_i v_j \, . \tag{3.3.41}$$

The first two moments give rise to the conservation equations in the form of 3.3.24, but here they are unsmoothed quantities still. The higher moments give continuity equations for the stress tensor in terms of higher moments of the distribution function.

Before continuing, we make two remarks. First, it turns out that setting the stress tensor to zero is a self-consistent ansatz: the generation of the

stress tensor is proportional to the tensor itself. Second, the zero stress tensor case is equivalent to the assumption that there is only a single velocity associated with every position (see [49]),

$$f(\boldsymbol{x},\boldsymbol{p},\tau)=\frac{\bar{\rho}_m a^3}{m}[1+\delta(\boldsymbol{x},\tau)]\delta_D[\boldsymbol{p}-am\boldsymbol{v}(\boldsymbol{x},\tau)]. \tag{3.3.42}$$

A distribution that possesses this property is called a single streaming. Deviations from a perfect fluid are therefore directly related to multi-streaming. Despite this apparent simplicity, in order to define a perturbative large scale theory from the Vlasov-Poisson system, we have to smooth the distribution function

$$f_l(\boldsymbol{x},\boldsymbol{p},\tau)=\int d^3x' W_\Lambda(|\boldsymbol{x}-\boldsymbol{x}'|)f(\boldsymbol{x},\boldsymbol{p},\tau). \tag{3.3.43}$$

Since such a smoothing region contains many particles, this automatically induces multi-streaming, which in turn excites all higher moments. Next, we show how this system of equations can be truncated consistently to form the effective fluid on large scales.

Boltzmann hierarchy truncation: appearance of the effective fluid

The system of equations for the moments, or rather cumulants, can be derived as follows. Let us define the moment generating function[8]

$$M(\boldsymbol{l})=\frac{m}{\bar{\rho}_m a^3}\int d^3p\, e^{\boldsymbol{l}\cdot\frac{\boldsymbol{p}}{am}} f(\boldsymbol{x},\boldsymbol{p},\tau) \tag{3.3.44}$$

The cumulant generating function, which generates the part of the moments that is independent of lower moments, is defined as

$$C(\boldsymbol{l})=\log M(\boldsymbol{l}). \tag{3.3.45}$$

The zeroth, first and second moment are $(1+\delta)$, $\boldsymbol{v}_l$ and σ_{ij} respectively. The smoothed versions are defined as

$$C_l(\boldsymbol{l})=\log M_l(\boldsymbol{l})\ ;\quad M_l(\boldsymbol{l})=\frac{m}{\bar{\rho}_m a^3}\int d^3p\, e^{\boldsymbol{l}\cdot\frac{\boldsymbol{p}}{am}} f_l(\boldsymbol{x},\boldsymbol{p},\tau). \tag{3.3.46}$$

[8] We need to assume as physical input this exists. A thermal distribution would do the job for small enough $|\boldsymbol{l}|$, for instance. More generally, we expect the number of particles at high energies to be strongly suppressed.

The equation of motion for the cumulant generating function can be derived from the Vlasov equation. However, as we encountered before, the smoothing of a product of spatially varying fields is not equal to product of the smoothed fields. This applies to the last term in the Vlasov equation. The smoothed equations of motion therefore encounter higher derivative corrections and contributions of the form $[f_s\phi_s]_\Lambda$. We neglect the derivative corrections here as we assume them to be subdominant for most questions, in which case the equation of motion for the smoothed cumulant generating function is ([49])

$$\frac{\partial C_l}{\partial \tau} + \mathcal{H}\, \boldsymbol{l}\cdot\boldsymbol{\nabla}_{\boldsymbol{l}} C_l + \boldsymbol{\nabla} C_l \cdot \boldsymbol{\nabla}_{\boldsymbol{l}} C_l + \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}_{\boldsymbol{l}} C_l + \boldsymbol{l}\cdot\boldsymbol{\nabla}_{\boldsymbol{l}}\phi_l + \frac{1}{M_l}\left[M_s\, \boldsymbol{l}\cdot\boldsymbol{\nabla}\phi_s\right]_\Lambda = 0, \tag{3.3.47}$$

where $M_s = M - M_l$ and $\boldsymbol{\nabla}_{\boldsymbol{l}}$ denotes a derivative with respect to the vector $\boldsymbol{l}$. The last term is the new term that we referred to. Note that if the fundamental theory consist of separate particles, M_s comes from a single stream distribution function. The cumulant equations of motion are obtained from derivatives of this equations with respect to $\boldsymbol{l}$ at $\boldsymbol{l} = 0$. For instance, the first derivative yields the Euler equation with stress

$$\frac{\partial \boldsymbol{v}_l}{\partial \tau} + \mathcal{H}\, \boldsymbol{v}_l + (\boldsymbol{v}_l\cdot\boldsymbol{\nabla})\,\boldsymbol{v}_l = -\boldsymbol{\nabla}\phi_l - \frac{1}{\tilde{\rho}_l}\boldsymbol{\nabla}_j\left(\tilde{\rho}_l\sigma^l_{ij}\right) - \frac{1}{\tilde{\rho}_l}\left[\tilde{\rho}_s\boldsymbol{\nabla}\phi_s\right]_\Lambda. \tag{3.3.48}$$

Combining the last two terms yields exactly the Euler equation with stress tensor we found form the bottom-up approach 3.3.24. What remains to be shown is that we can neglect the effect of higher order cumulants on the stress tensor: do the equations of motion generate a growth in the stress tensor due to higher order cumulants and short scale effects we have integrated out?

The structure of the cumulant equations is as follows

$$\partial_\tau c_j \sim \sum_{j'=0}^{j} c_{j+1-j'}\boldsymbol{\nabla} c_{l'} + \boldsymbol{\nabla} c_{j+1} + \left(\frac{\partial}{\partial \boldsymbol{l}}\right)^j \left(\frac{1}{M_l(\boldsymbol{l})}\left[M_s(\boldsymbol{l})\,\boldsymbol{l}\cdot\boldsymbol{\nabla}\phi_s\right]_\Lambda\right)\Bigg|_{l=0}, \tag{3.3.49}$$

where c_j is the j-th cumulant. Let us forget about the last term for a second. Then, assuming the time dependence is of order Hubble, the contribution of a cumulant of order $j+1$ to a cumulant of order j is always of order

$$c_j \supset \frac{k_l}{\mathcal{H}} c_{j+1}, \tag{3.3.50}$$

where k_l is some long momentum. Naively, one might worry that $k/\mathcal{H} \gg 1$. However, we have to keep in mind that from its definition, we can estimate $c_{j+1} \sim v_p c_j$. Thus, the effect of higher order cumulants is of order

$$\frac{k}{\mathcal{H}} c_{j+1} \sim \frac{k v_p}{\mathcal{H}} c_j. \tag{3.3.51}$$

We estimate the particle velocity from linear theory as the velocity at the nonlinear scale. The variance of the dimensionless velocity is then given by (again from [11])

$$\Delta_v^2(k_{\rm NL}) \sim \frac{\mathcal{H}^2}{k^2} \Delta_\delta^2(k_{\rm NL}) \sim \frac{\mathcal{H}^2}{k^2}, \tag{3.3.52}$$

which implies $v_p \sim \mathcal{H}/k_{\rm NL}$. Thus,

$$\frac{k}{\mathcal{H}} c_{j+1} \sim \frac{k}{k_{\rm NL}} c_j. \tag{3.3.53}$$

In an ordinary fluid, the effect of higher moments would be suppressed by $k l_{\rm mfp}$, where $l_{\rm mfp}$ is the mean free path. Interestingly, for ordinary fluids the mean free path is small due to interactions, whereas interactions are absent in our gravitational fluid. The effective small mean free path is a direct consequence of the finite lifetime of the universe and its smooth initial conditions: particles did not have the time to move very far since the beginning of the matter era. This shows that higher cumulants feed back into the equations as

$$\sigma_{ij} \supset \left(\frac{\boldsymbol{\nabla}}{\mathcal{H}}\right)^{l-2} c_l \sim \left(\frac{k}{k_{\rm NL}}\right)^{l-2} \sigma_{ij}. \tag{3.3.54}$$

These effects are nicely captured by the derivative expansion of the stress tensor advocated in the bottom-up picture. The smallness of the stress tensor itself was studied numerically and analytically in [49], and indeed verified. Finally, we comment on the effect on the last term in 3.3.49. From 3.3.48 and rewritten in 3.3.25, we recall that its contribution to the velocity is of order $(k/\mathcal{H})\sigma_{ij}$. Going to higher cumulants, we find that its contribution to c_j is of order $v_l^n v_p^{j-1-n} (k/\mathcal{H})\sigma_{ij}$, where $0 \leq n \leq j-1$, which is indeed also under perturbative control.

Conclusion

Finally, we conclude that from the particle picture (top-down) we derived the same effective fluid equations as we did from the bottom-up construction. Moreover, we argued that the smallness of $(k/\mathcal{H})v_p$, where $(1/\mathcal{H})v_p$ acts as an effective mean free path, guarantees the consistency of the Boltzmann truncation, where the effects of higher cumulants backreact as higher derivative corrections, or at higher orders in perturbations. In the bottom-up construction, the expansion in fields and derivatives of the effective stress tensor effectively captures this effect of higher order cumulants on the density and velocity field. It is in this sense that these degrees of freedom are integrated out to arrive at an effective theory for long wavelength density and velocity perturbations.

Thus we arrived at the EFT of LSS for Gaussian initial conditions. In the next chapter, we extend our theory to include long-short correlations in the initial conditions.

4 The EFT of Large Scale Structure with Primordial Non-Gaussianity

4.1 Introduction

Large-scale structure (LSS) surveys will play an increasingly important role in probing the initial conditions and subsequent evolution of our universe [31]. However, extracting primordial information from the observations will be challenging, and understanding the many sources of late-time nonlinearities will be essential for realizing the full potential of the future data. On sufficiently large scales, fluctuations in the dark matter density are small and therefore amenable to a perturbative treatment [29, 50]. In contrast, dark matter is strongly clustered on small scales and perturbation theory is insufficient to describe its dynamics. Moreover, gravitational nonlinearities couple short and long modes, so even the perturbative regime isn't immune to our uncertainties about the evolution of LSS on small scales.

Failing to account for the backreaction of the short-scale nonlinearities on the long-wavelength universe can bias the theoretical interpretation of future observations. Recently, this problem has been addressed using the methods of effective field theory (EFT) [11, 47]. In this approach, dark matter fluctuations are separated into long and short modes. While the short modes are nonperturbative and can only be modelled through numerical simulations, their effects at long distances can be captured systematically by adding corrections to the evolution equations for the long-wavelength perturbations. These corrections can be organized in a double expansion in powers of the long-wavelength fluctuations and spatial derivatives. The allowed terms are constrained by symmetries, and only a finite number of terms is required to describe observations at a finite level of precision.

The EFT description of the long-wavelength fluctuations is complete if (and only if) the set of operators correcting the fluid equations is closed under renormalization. By this we mean that, at a given order, loops don't generate new operators, but only mix the existing operators. The renormalization of the power spectrum was treated in [43, 47] (at one-loop) and in [51] (at two-loop), while the one-loop renormalization of the bispectrum was presented in [45, 52] (for Gaussian initial conditions). Extensions to Lagrangian perturbation theory have appeared in [38], halo statistics were treated in [46, 53–57], redshift space distortions have been included in [58], and baryonic effects were discussed in [59]. Since probing primordial non-Gaussianity is a key motivation underlying these developments, it is essential

to systematically develop the EFT approach for these more general initial conditions. In this Chapter, we work out the extension of the effective theory of large-scale structure (EFT-of-LSS) to non-Gaussian initial conditions. We will present explicitly the renormalization of the one-loop dark matter power spectrum and bispectrum for a wide class of primordial non-Gaussianities (PNG).

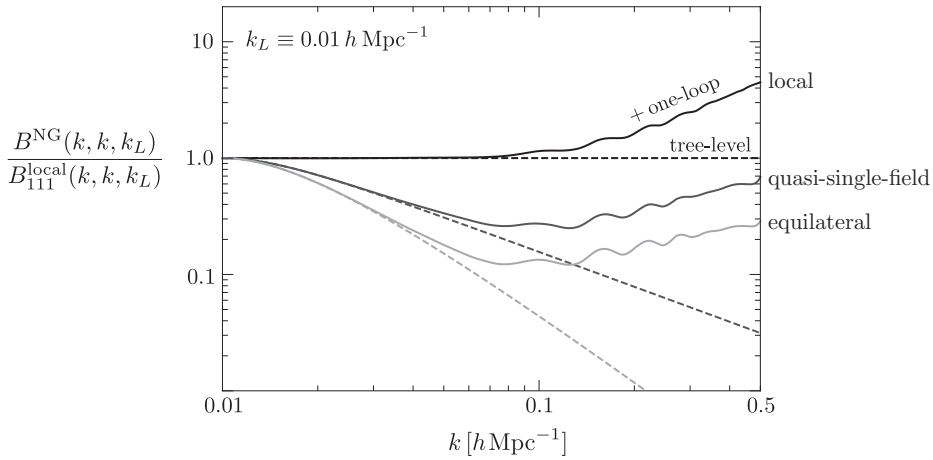


Figure 4.1: Solid (dashed) lines show the full (tree-level) non-Gaussian SPT contribution at $z = 0$ for three representative models of primordial non-Gaussianity (see §4.5.3 for more details).

The importance of an accurate treatment of nonlinear corrections to the dark matter bispectrum is illustrated in fig. 4.1. Shown is the bispectrum computed in standard perturbation theory (SPT) for PNG of local, equilateral and quasi-single-field type (see §4.5.3 for more details). Dashed lines refer to the tree-level signal, while solid lines contain the leading-order effects of gravitational evolution. We see that the loop corrections become relevant on relatively large scales and tend to decrease the difference among the three primordial signals. Even for an idealized observable such as the dark matter bispectrum, linear evolution is therefore insufficient to fully exploit the potential of current and future data sets. In this work, we will show how nonlinear gravitational corrections can be computed in a systematic and self-consistent way.

The outline of the Chapter is as follows. In Section 4.2, we review the shortcomings of standard perturbation theory and explain how they are addressed in the EFT-of-LSS. We then define the types of non-Gaussian initial conditions studied in this work (see also Appendix 4.A). In Section 4.3, we show how PNG leads to new terms in the stress tensor of the EFT-of-LSS. We derive these terms both from the 'top-down', by coarse graining the equations of motion (cf. Appendix 4.B), and from the 'bottom-up', by constructing the most general stress tensor consistent with the symmetries. In Section 4.4, we show that our set of EFT terms is closed under renormalization. We explicitly derive the one-loop counterterms required by the renormalization procedure (with details given in Appendix 4.C). In Section 4.5, we compute the renormalized one-loop dark matter bispectrum numerically (see also Appendix 4.D). We provide a preliminary analysis of the shapes of the distinct EFT contributions. A more detailed treatment will appear in [60]. Our conclusions are summarized in Section 4.6.

Notation and Conventions

The most important variables used in this work are collected in Appendix 4.E. Both conformal time τ and the scale factor a are employed as measures of time evolution. Three-dimensional vectors will be denoted in boldface ($\boldsymbol{x}$, $\boldsymbol{k}$, etc.) or with Latin subscripts (x_i, k_i, etc.). The magnitude of vectors is defined as $k \equiv |\boldsymbol{k}|$ and unit vectors are written as $\hat{\boldsymbol{k}} \equiv \boldsymbol{k}/k$. We sometimes write the sum of n vectors as $\boldsymbol{k}_{1\ldots n} \equiv \boldsymbol{k}_1 + \ldots + \boldsymbol{k}_n$. We will often use the following shorthand for three-dimensional momentum integrals

$$\int_{\boldsymbol{p}} (\ldots) \equiv \int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3} (\ldots) \,.$$

We denote the three-dimensional Laplacian by $\triangle \equiv \delta^{ij}\partial_i\partial_j$. A prime on correlation functions, $\langle\cdots\rangle'$, indicates that an overall momentum-conserving delta function is being dropped. Our convention for the dimensionless power spectrum is

$$\Delta^2(k) \equiv \frac{k^3}{2\pi^2} P(k) \,.$$

We also define a dimensionless bispectrum as

$$\mathcal{B}(k_1, k_2, k_3) \equiv \left(\frac{k_1^3}{2\pi^2}\right)^2 B(k_1, k_2, k_3) \,.$$

We will typically use $\boldsymbol{p}$ for short-scale fluctuations and reserve $\boldsymbol{k}$ for long-wavelength modes. We use X_ℓ and X_s for the long-wavelength and short-wavelength parts of a quantity X. The Gaussian and non-Gaussian parts of X are $X^{\rm G}$ and $X^{\rm NG}$. Finally, $X^{[L]}$ denotes the spin-L part of X.

When we present numerical results, the linear power spectrum is computed with the Boltzmann code CAMB [61], using a flat ΛCDM cosmology with $\Omega_m^0 = 0.27$, $\Omega_\Lambda^0 = 0.73$ $h = 0.70$. The initial power spectrum of the gravitational potential is taken to be of a power law form with amplitude $\Delta^2_\varphi = 8.7 \times 10^{-10}$ and spectral index $n_s = 0.96$, defined at the pivot scale $k_0 = 0.002\,h\mathrm{Mpc}^{-1}$.

4.2 Preliminaries: Non-Gaussian Initial Conditions

Standard perturbation theory (SPT) and the Effective Field Theory were introduced in 3. For a much more detailed exposition, we refer the reader to the classic review [29]. Further details may also be found in Appendix 4.C. Primordial non-Gaussianity (PNG) was introduced in 2. Here we define our way of parameterizing them conveniently.

4.2.1 Non-Gaussian Initial Conditions

We choose to define our "initial" conditions at some time τ_{in} after matter-radiation equality, but early enough that nonlinearities in the prior evolution can still be ignored. The linearly-evolved potential $\phi_{(1)}(\boldsymbol{k},\tau)$ and the linearly-evolved dark matter density contrast $\delta_{(1)}(\boldsymbol{k},\tau)$ can then be written in terms of the primordial potential $\varphi(\boldsymbol{k})$:

$$\phi_{(1)}(\boldsymbol{k},\tau) = T(k,\tau)\varphi(\boldsymbol{k}) \ , \tag{4.2.1}$$

$$\delta_{(1)}(\boldsymbol{k},\tau) = -\frac{2}{3}\frac{k^2}{\mathcal{H}^2\Omega_m}\phi_{(1)}(\boldsymbol{k},\tau) \equiv M(k,\tau)\varphi(\boldsymbol{k}) \ , \tag{4.2.2}$$

where the transfer function $T(k,\tau)$ captures the evolution of the gravitational potential in the radiation era.[1] To simplify the notation, we will typically drop the time arguments, but the functions $T(k)$ and $M(k)$ should

[1] As in Chapter 3, the transfer function is implicitly defined to capture the effect of the evolution of perturbations.

always be evaluated at time τ. We define the power spectrum of the primordial potential as

$$P_\varphi(k) \equiv \langle \varphi(\boldsymbol{k})\varphi(-\boldsymbol{k})\rangle' \equiv \frac{2\pi^2}{k^3}\Delta_\varphi^2 \left(\frac{k}{k_0}\right)^{n_s-1} , \tag{4.2.3}$$

where we have assumed a nearly scale-invariant power law ansatz motivated by inflation. The power spectrum of the linearly-evolved density contrast $\delta_{(1)}(\boldsymbol{k},\tau)$ then is

$$P_{11}(k) = [M(k)]^2 P_\varphi(k) . \tag{4.2.4}$$

Recall that both $M(k)$ and $P_{11}(k)$ are to be evaluated at time τ, while $P_\varphi(k)$ is time independent.

Primordial non-Gaussianity (PNG) leads to higher-order correlations beyond the power spectrum. The leading diagnostic for PNG is the bispectrum

$$B_\varphi(k_1,k_2,k_3) \equiv \langle \varphi(\boldsymbol{k}_1)\varphi(\boldsymbol{k}_2)\varphi(\boldsymbol{k}_3)\rangle' , \tag{4.2.5}$$

$$B_{111}(k_1,k_2,k_3) = [M(k_1)M(k_2)M(k_3)]B_\varphi(k_1,k_2,k_3) . \tag{4.2.6}$$

For perturbative PNG, the potential φ can be expanded around a Gaussian field φ_g. For instance, at lowest order, we may write [62]

$$\begin{aligned}\varphi(\boldsymbol{k}) =& \varphi_g(\boldsymbol{k}) + f_{\rm NL}\int_{\boldsymbol{p}} K_{\rm NL}(\boldsymbol{p},\boldsymbol{k}-\boldsymbol{p})\left[\varphi_g(\boldsymbol{p})\varphi_g(\boldsymbol{k}-\boldsymbol{p}) - P_g(p)\,(2\pi)^3\delta_D(\boldsymbol{k})\right] \\ &+ \dots ,\end{aligned} \tag{4.2.7}$$

where we have subtracted the power spectrum of the Gaussian field, $P_g(p) \equiv \langle \varphi_g(\boldsymbol{p})\varphi_g(-\boldsymbol{p})\rangle'$, to ensure that $\langle\varphi\rangle = 0$. Eq. (4.2.7) is the most general quadratic expansion satisfying the requirements of statistical homogeneity and isotropy. The momentum-dependent kernel function $K_{\rm NL}(\boldsymbol{k}_1,\boldsymbol{k}_2)$ parametrizes the shape of the non-Gaussianity. Substituting (4.2.7) into (4.2.5), and keeping only the leading term in an expansion in $f_{\rm NL}\varphi_g$, we get

$$B_\varphi(k_1,k_2,k_3) = 2f_{\rm NL}\,K_{\rm NL}(\boldsymbol{k}_1,\boldsymbol{k}_2)\,P_\varphi(k_1)P_\varphi(k_2) + 2 \text{ perms} . \tag{4.2.8}$$

The permutations in (4.2.8) imply that the bispectrum alone does not uniquely determine the kernel function $K_{\rm NL}(\boldsymbol{k}_1,\boldsymbol{k}_2)$ (i.e. different kernel functions

can give rise to the same bispectum, but different trispectrum). However, the ambiguity in the choice of kernel disappears in the squeezed limit, $q \equiv k_1/k_2 \ll 1$, where we get

$$K_{\rm NL}(\boldsymbol{k}_1, \boldsymbol{k}_2) \xrightarrow{q\to 0} \frac{B_\varphi(k_1, k_2, k_3)}{4 f_{\rm NL} P_\varphi(k_1) P_\varphi(k_2)} \Big(1 + \mathcal{O}(q)\Big) \,. \tag{4.2.9}$$

In this limit, the kernel and its leading scaling with q are uniquely defined in terms of the bispectrum. As we will see, the squeezed limit will be particularly relevant for our investigation.

Statistical homogeneity requires that $K_{\rm NL}(\boldsymbol{k}_1, \boldsymbol{k}_2)$ is only a function of k_1, k_2 and $\hat{\boldsymbol{k}}_1 \cdot \hat{\boldsymbol{k}}_2$. It will be convenient to express the angular dependence as an expansion in terms of Legendre polynomials $P_L(\hat{\boldsymbol{k}}_1 \cdot \hat{\boldsymbol{k}}_2)$, and write the squeezed limit as

$$\boxed{K_{\rm NL}(\boldsymbol{k}_1, \boldsymbol{k}_2) \xrightarrow{q\to 0} \sum_{L,i} a_{L,i} \left(\frac{k_1}{k_2}\right)^{\Delta_i} P_L(\hat{\boldsymbol{k}}_1 \cdot \hat{\boldsymbol{k}}_2)} \,, \tag{4.2.10}$$

where, by symmetry, L has to be an even integer (see Appendix 4.A for more details). The ansatz (4.2.10) captures many physically relevant cases. (Non-Gaussianity in feature models [63] may fall outside of this parameterization.) We can organize the different contributions by the order L (the "spin") of the Legendre polynomial and treat each scaling Δ_i separately:

- *Scalar contributions*

  First, we consider the scalar contributions; i.e. we set $a_{L\geq 2} \equiv 0$. The different squeezed limits are then distinguished by their scaling dimensions Δ:

  - For $\Delta = 0$, the kernel is a momentum-independent constant, $K_{\rm NL}(\boldsymbol{k}_1, \boldsymbol{k}_2) = 1$, and the ansatz (4.2.7) corresponds to a *local* expression for φ in real space [64], i.e. $\varphi(\boldsymbol{x}) = \varphi_g(\boldsymbol{x}) + f_{\rm NL}(\varphi_g^2(\boldsymbol{x}) - \langle\varphi_g^2\rangle)$. This is the case of *local non-Gaussianity*.
  - For $\Delta = 2$, the squeezed limit is suppressed by two additional powers of the low-momentum mode k_1. This is characteristic of higher-derivative interactions in single-field inflation, which produce *equilateral non-Gaussianity* [33, 65].

- A squeezed limit with intermediate momentum scaling, $\Delta \in [0, 1.5]$, can arise if the inflaton interacts with massive scalar particles during inflation. In these models of *quasi-single-field inflation* [66], nonlinear interactions of the additional scalars can be mediated to the inflaton sector, creating observable non-Gaussianity with a characteristic signature in the squeezed limit: the scaling dimensions Δ are functions of the masses of the extra particles. Coupling the inflaton to operators in a conformal field theory [67] allows to extend the intermediate momentum scaling to the regime $\Delta \in [0, 2]$.

- *Higher-spin contributions*

  Various physical mechanisms can lead to an angular dependence in the squeezed limit [68]:

  - The inflaton may couple to massive higher-spin particles. In this case, the angular dependence is given by the Legendre polynomial of order the spin of the particle [69]. At tree level, only particles with even spin contribute to the bispectrum in the squeezed limit.
  - At loop level, the angular dependence induced by the interaction with higher-spin particles can be different from just a simple Legendre polynomial. For example, coupling the inflaton to a $U(1)$ gauge field via the interaction $I(\phi)F^2$ [70] leads to a shape of the form (4.2.10) with $a_2 = a_0/2$.
  - Curvature perturbations sourced by large-scale primordial magnetic fields [71, 72] also can lead to non-zero a_0, a_1 and a_2.
  - The bispectrum produced in solid inflation [20] corresponds to $a_2 \gg a_0$.

4.3 Coarse Graining

In the EFT-of-LSS the short-scale fluctuations have to be integrated out both in the equations of motion and in the initial conditions. The presence of non-Gaussianity in the initial conditions yields some non-trivial features which have to be taken into account in a consistent renormalization of the loop expansion. In contrast to a usual fluid (such as air or water), for dark

matter there isn't a hierarchy of time scales between the evolution of short-wavelength and long-wavelength fluctuations. This implies that the short scales keep memory of their initial conditions. For Gaussian initial conditions, this isn't very important since all scales are statistically independent. As a result, there are no initial correlations between the long-wavelength fluctuations of the EFT and the short scales which are being integrated out. However, in the presence of primordial non-Gaussianity, the initial statistics of the short scales depends on the long-wavelength fluctuations. This initial dependence will affect the dynamics of the short modes, which in turn will then backreact on the evolution of the long-wavelength fluctuations. The objective of this section is to understand how this memory effect can be incorporated in the EFT-of-LSS.

In §4.3.1, we coarse grain the dark matter equations of motion to derive the stress tensor of the effective theory. As we explained in §3.3, this stress tensor can be expanded in terms of the long-wavelength fields and their derivatives. In the presence of PNG, the coefficients of this expansion will be spatially modulated. In §4.3.2, we derive this dependence by coarse-graining the initial condition (4.2.7). This introduces new, non-dynamical fields in the EFT. We collect the leading-order non-Gaussian corrections in §4.3.3.

4.3.1 Smoothing the Equations of Motion

To set up a consistent expansion of the stress tensor in the presence of PNG, we rewrite the expansion in this section in a slightly different form, such that the potential dependence on initial conditions is more explicit.

At first order in the long-wavelength fluctuations and to leading order in derivatives, we can write (3.3.30) as

$$\begin{aligned}\tau^{ij}(\boldsymbol{x},\tau) \;&=\; c^{ij}(\boldsymbol{x},\tau) \;+\; c^{ij}{}_{kl}(\boldsymbol{x},\tau)\,\partial^k\partial^l\phi_\ell(\boldsymbol{x},\tau) \\ &\quad+\; \hat{c}^{ij}{}_{kl}(\boldsymbol{x},\tau)\,\partial^k v_\ell^l(\boldsymbol{x},\tau) \;+\; \cdots\;. \end{aligned} \tag{4.3.1}$$

We will sometimes suppress the indices on the coefficients in (4.3.1) and collectively refer to them as $c(\boldsymbol{x},\tau)$. Notice that the coefficients $c(\boldsymbol{x},\tau)$ depend on position through their dependence on the initial short-scale fluctuations

[2]

$$c(\boldsymbol{x},\tau) \equiv c[\phi_s(\boldsymbol{q},\tau_{in}),\tau] \ . \tag{4.3.2}$$

Since the short scales have been removed by coarse graining, the relation in (4.3.2) is non-local. More precisely, the coefficients $c(\boldsymbol{x},\tau)$ will depend on the value of ϕ_s over a patch of size Λ^{-1} centred around the Lagrangian coordinate $\boldsymbol{q}(\boldsymbol{x},\tau)$. We will come back to this point in the next section.

When the small-scale fluctuations are replaced by their statistical ensemble averages, the stress tensor becomes

$$\tau^{ij} = \left\langle \tau^{ij} \right\rangle_s + \Delta\tau^{ij} \ , \tag{4.3.3}$$

where $\langle\cdots\rangle_s$ denotes an average over many realizations of the short modes in a fixed long-wavelength background and $\Delta\tau^{ij}$ is a stochastic term which accounts for the statistical deviation from the average. We will sometimes refer to $\langle\tau^{ij}\rangle_s$ as the "viscosity" part of the stress tensor and to $\Delta\tau^{ij}$ as the "noise" part. We get the viscosity contribution to the stress tensor by replacing the coefficients $c(\boldsymbol{x},\tau)$ in (4.3.1) by their statistical averages $\langle c(\boldsymbol{x},\tau)\rangle_s$. For Gaussian initial conditions, the average over short-scale fluctuations is independent of the long-wavelength fluctuations. As a result, the averaged coefficients $\langle c(\boldsymbol{x},\tau)\rangle_s$ become simple cutoff-dependent parameters of the EFT. However, primordial non-Gaussianity gives rise to correlations between different scales and the averaged coefficients will depend (non-locally) on the initial long-wavelength fluctuations. This dependence can be determined by integrating out the short-scale fluctuations in the initial conditions, which we shall do next.

4.3.2 Smoothing the Initial Conditions

We apply the filtering procedure (3.3.17) to the primordial potential (4.2.7). Isolating the terms which contribute to the coupling between the long and

[2] In (4.3.2), we have assumed that the initial short-scale fluctuations are fully determined by the short-scale potential $\phi_s(\boldsymbol{q},\tau_{in})$. We could also have added a dependence on the short-scale velocity $v_s^i(\boldsymbol{q},\tau_{in})$ and its dispersion. However, at early times, every scale is perturbative so that one can use the linear equations of motion to express the velocity in terms of ϕ_s. In particular, assuming that these scales are in their growing mode, one can show that the velocity is determined by the gradient of the potential, $v_s^i \propto \partial^i\phi_s$.

short modes, we get

$$\varphi_\ell(\boldsymbol{k}) \supset f_{\rm NL}\int_{\tilde{\boldsymbol{p}}} K_{\rm NL}(\boldsymbol{k}-\tilde{\boldsymbol{p}},\tilde{\boldsymbol{p}})\left[\varphi_g^s(\boldsymbol{k}-\tilde{\boldsymbol{p}})\varphi_g^s(\tilde{\boldsymbol{p}}) - P_g(\tilde{p})\,(2\pi)^3\delta_D(\boldsymbol{k})\right]$$
$$\equiv f_{\rm NL}\psi_J(\boldsymbol{k})\ , \tag{4.3.4}$$

$$\varphi_s(\boldsymbol{p}) \supset 2f_{\rm NL}\int_{\tilde{\boldsymbol{p}}} K_{\rm NL}(\boldsymbol{p}-\tilde{\boldsymbol{p}},\tilde{\boldsymbol{p}})\,\varphi_g^s(\boldsymbol{p}-\tilde{\boldsymbol{p}})\,\varphi_g^\ell(\tilde{\boldsymbol{p}})\ . \tag{4.3.5}$$

The field ψ_J in (4.3.4) comes about because the split between short and long-wavelength fluctuations is done at the level of the Gaussian field φ_g. For non-Gaussian initial conditions, the relation between the physical primordial potential φ and the Gaussian one φ_g is nonlinear, so that two short modes of the Gaussian field can combine to produce a long-wavelength fluctuation. The field ψ_J is a noise term which precisely captures this effect. This field does not correlate with the long-wavelength fluctuations, but it has non-trivial correlations with the noise terms of the stress tensor, $\Delta\tau^{ij}$. We will explain this in more detail in the next section. The right-hand side of (4.3.5) encodes the dependence of the short modes on the long-wavelength fluctuations. Writing $\phi_g^s(\boldsymbol{p},\tau_{in}) \equiv T(p,\tau_{in})\varphi_g^s(\boldsymbol{p})$ and going to position space, we find

$$\phi_s(\boldsymbol{x},\tau_{in}) \simeq \phi_g^s(\boldsymbol{x},\tau_{in}) + 2f_{\rm NL}\int_{\boldsymbol{k}}\int_{\boldsymbol{p}} K_{\rm NL}(\boldsymbol{k},\boldsymbol{p})\,\phi_g^s(\boldsymbol{p},\tau_{in})\,\varphi_g^\ell(\boldsymbol{k})\,e^{i(\boldsymbol{p}+\boldsymbol{k})\cdot\boldsymbol{x}}\ . \tag{4.3.6}$$

Since $k \ll p$, the result in (4.3.6) only depends on the squeezed limit of the kernel function $K_{\rm NL}$ (and the corresponding bispectrum). In §4.2.1, we introduced the following ansatz

$$K_{\rm NL}(\boldsymbol{k},\boldsymbol{p}) \xrightarrow{k\ll p} \sum_{L,i} a_{L,i}\left(\frac{k}{p}\right)^{\Delta_i} P_L(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{p}})\left[1+\mathcal{O}\left(k^2/p^2\right)\right]\ . \tag{4.3.7}$$

For each scaling Δ in this sum, we can treat the different orders ("spins") in the Legendre expansion separately:

- *Spin-0*

  For $L=0$, the initial short-scale fluctuations can be written as

$$\phi_s(\boldsymbol{x},\tau_{in}) = \phi_g^s(\boldsymbol{x},\tau_{in}) + f_{\rm NL}\,\alpha^{(s)}(\boldsymbol{x})\,\psi(\boldsymbol{x})\ , \tag{4.3.8}$$

where

$$\boxed{\psi(\boldsymbol{k}) \equiv \left(\frac{k}{\mu}\right)^{\Delta} \varphi_g^{\ell}(\boldsymbol{k})}\,, \tag{4.3.9}$$

and $\alpha^{(s)}(\boldsymbol{p}) \equiv 2a_0\,(\mu/p)^{\Delta}\,\phi_g^s(\boldsymbol{p},\tau_{in})$. The scale μ is an arbitrary momentum scale introduced to make $\psi(\boldsymbol{x})$ dimensionless. As we shall see, the field ψ will play an important role in the rest of this Chapter. It describes how the statistics of the small scales is modulated by the presence of the long modes. For local non-Gaussianity, we have $\Delta = 0$ and therefore $\psi(\boldsymbol{x}) = \varphi_g(\boldsymbol{x})$. Eq. (4.3.8) then reduces to a more familiar expression

$$\phi_s(\boldsymbol{x},\tau_{in}) = \left(1 + 2f_{\rm NL}^{\rm local}\varphi_g^{\ell}(\boldsymbol{x})\right)\phi_g^s(\boldsymbol{x},\tau_{in})\ . \tag{4.3.10}$$

We see that local non-Gaussianity simply modulates the amplitude of the small-scale fluctuations.

To compute the dependence of the coefficients $\langle c(\boldsymbol{x},\tau)\rangle_s$ on ψ, we substitute (4.3.8) into (4.3.2). At first order in an expansion in ψ, we find

$$\begin{aligned}\langle c(\boldsymbol{x},\tau)\rangle_s \ &= \langle c[\phi_g^s(\boldsymbol{q},\tau_{in}),\tau]\rangle_s \\ &\quad + f_{\rm NL}\int_{\tilde{\boldsymbol{q}}}\left\langle \frac{\delta c(\boldsymbol{x},\tau)}{\delta\phi_s(\tilde{\boldsymbol{q}},\tau_{in})}\,\alpha^{(s)}(\tilde{\boldsymbol{q}},\tau_{in})\right\rangle_s \psi(\tilde{\boldsymbol{q}})\ ,\end{aligned} \tag{4.3.11}$$

where the derivative is evaluated for the Gaussian field configuration ϕ_g^s. Furthermore, since the integral only has support for $|\tilde{\boldsymbol{q}} - \boldsymbol{q}(\boldsymbol{x},\tau)| < \Lambda^{-1}$, we can pull the long-wavelength field $\psi(\tilde{\boldsymbol{q}}) \approx \psi(\boldsymbol{q})$ out of the integral (up to corrections that are higher order in ∇^2/Λ^2). Translation invariance then guarantees that the remaining integral is independent of the position $\boldsymbol{x}$, and eq. (4.3.11) can be written as

$$\langle c(\boldsymbol{x},\tau)\rangle_s \ \equiv\ \mathfrak{c}(\Lambda,\tau) + f_{\rm NL}\,\mathfrak{c}_\psi(\Lambda,\tau)\,\Psi(\boldsymbol{x},\tau)\ , \tag{4.3.12}$$

where we have defined

$$\Psi(\boldsymbol{x},\tau) \equiv \psi(\boldsymbol{q}(\boldsymbol{x},\tau))\ . \tag{4.3.13}$$

In Section 4.4, we will see why it is important that field ψ is evaluated at the Lagrangian position $\boldsymbol{q}$. Restoring the indices, but dropping the arguments of the coefficients and the fields, we get

$$\langle c^{ij}\rangle_s^{[0]} = \left[\mathfrak{c}^{[0]} + f_{\rm NL}\,\mathfrak{c}_\psi^{[0]}\Psi\right]\delta^{ij} \ , \tag{4.3.14}$$

$$\begin{aligned}\langle c^{ijkl}\rangle_s^{[0]} &= \left[\mathfrak{c}_0^{[0]} + f_{\rm NL}\,\mathfrak{c}_{\psi,0}^{[0]}\Psi\right]\delta^{ij}\delta^{kl} \\ &\quad + \left[\mathfrak{c}_1^{[0]} + f_{\rm NL}\,\mathfrak{c}_{\psi,1}^{[0]}\Psi\right]\left[\delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk}\right] \ ,\end{aligned} \tag{4.3.15}$$

and similarly for the coefficient $\langle \hat{c}^{ijkl}\rangle_s^{[0]}$. The tensor structure of (4.3.14) and (4.3.15) is fixed by statistical isotropy.

- *Spin-2*

  For $L = 2$, equation (4.3.6) becomes

$$\phi_s(\boldsymbol{x},\tau_{in}) = \phi_g^s(\boldsymbol{x},\tau_{in}) + f_{\rm NL}\,\alpha_{ij}^{(s)}(\boldsymbol{x})\,\psi^{ij}(\boldsymbol{x}) \ , \tag{4.3.16}$$

  where

$$\boxed{\psi^{ij}(\boldsymbol{k}) \equiv \frac{3}{2}\mathcal{P}^{ij}(\hat{\boldsymbol{k}})\left(\frac{k}{\mu}\right)^{\Delta}\varphi_g^\ell(\boldsymbol{k})} \ , \tag{4.3.17}$$

  and $\alpha_{ij}^{(s)}(\boldsymbol{p}) \equiv a_2\,(\mu/p)^\Delta\,(\hat{p}_i\hat{p}_j)\phi_g^s(\boldsymbol{p},\tau_{in})$. In (4.3.17), we have defined the projection operator $\mathcal{P}^{ij}(\hat{\boldsymbol{k}}) \equiv \hat{k}^i\hat{k}^j - \frac{1}{3}\delta^{ij}$. Note that the tensor ψ^{ij} is traceless and hence contains no scalar contribution. Substituting (4.3.16) into (4.3.2), we find

$$\langle c^{ij}\rangle_s^{[2]} = \mathfrak{c}^{[2]}\,\delta^{ij} + f_{\rm NL}\,\mathfrak{c}_\psi^{[2]}\,\Psi^{ij} \ , \tag{4.3.18}$$

$$\begin{aligned}\langle c^{ijkl}\rangle_s^{[2]} &= \mathfrak{c}_0^{[2]}\,\delta^{ij}\delta^{kl} + \mathfrak{c}_1^{[2]}\left[\delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk}\right] \\ &\quad + f_{\rm NL}\left[\mathfrak{c}_{\psi,0}^{[2]}\Psi^{ij}\delta^{kl} + \mathfrak{c}_{\psi,1}^{[2]}\Psi^{kl}\delta^{ij} + \right. \\ &\qquad \left. \mathfrak{c}_{\psi,2}^{[2]}\left[\Psi^{ik}\delta^{jl} + \Psi^{jk}\delta^{il} + \Psi^{il}\delta^{jk} + \Psi^{jl}\delta^{ik}\right]\right]\end{aligned} \tag{4.3.19}$$

  where we have defined $\Psi^{ij}(\boldsymbol{x},\tau) \equiv \psi^{ij}(\boldsymbol{q}(\boldsymbol{x},\tau))$. Although we have dropped the arguments, the coefficients in (4.3.18) and (4.3.19) of course still depend on the cutoff Λ and time τ.

- *Spin-4*

  For $L = 4$, equation (4.3.6) becomes

$$\phi_s(\boldsymbol{x}, \tau_{in}) = \phi_g^s(\boldsymbol{x}, \tau_{in}) + f_{\rm NL}\, \alpha^{(s)}_{ijkl}(\boldsymbol{x}) \psi^{ijkl}(\boldsymbol{x}) \ , \tag{4.3.20}$$

  where

$$\boxed{\psi^{ijkl}(\boldsymbol{k}) \equiv \frac{35}{8} \mathcal{P}^{ijkl}(\hat{\boldsymbol{k}}) \left(\frac{k}{\mu}\right)^{\Delta} \varphi_g^{\ell}(\boldsymbol{k})} \ , \tag{4.3.21}$$

  and $\alpha^{(s)}_{ijkl}(\boldsymbol{p}) \equiv a_4 \, (\mu/p)^{\Delta} \, (\hat{p}_i \hat{p}_j \hat{p}_k \hat{p}_l) \phi_g^s(\boldsymbol{p}, \tau_{in})$. The projection tensor $\mathcal{P}^{ijkl}(\hat{\boldsymbol{k}})$ ensures that ψ^{ijkl} is symmetric and traceless in any two of its indices

$$\begin{aligned} \mathcal{P}^{ijkl}(\hat{\boldsymbol{k}}) \equiv & \hat{k}^i \hat{k}^j \hat{k}^k \hat{k}^l - \frac{1}{7} \Big(\delta^{ij} \hat{k}^k \hat{k}^l + 5 \text{ perms} \Big) \\ & + \frac{1}{35} \Big(\delta^{ij} \delta^{kl} + 2 \text{ perms} \Big) \ . \end{aligned} \tag{4.3.22}$$

  Let us note that, at leading order in derivatives, the coefficients with only two indices, c^{ij}, cannot depend on ψ^{ijkl} since we would need to contract two of the indices. As a result, the field ψ^{ijkl} only contributes to the coefficients with four indices

$$\langle c^{ijkl} \rangle_s^{[4]} = \mathfrak{c}_0^{[4]} \delta^{ij} + \mathfrak{c}_1^{[4]} [\delta^{ik} \delta^{jl} + \delta^{il} \delta^{jk}] + f_{\rm NL}\, \mathfrak{c}_{\psi}^{[4]} \Psi^{ijkl} \ , \tag{4.3.23}$$

  where, as before, we have defined $\Psi^{ijkl}(\boldsymbol{x}, \tau) \equiv \psi^{ijkl}(\boldsymbol{q}(\boldsymbol{x}, \tau))$.

- *Higher-spin*

  It should be clear that higher-spin contributions ($L > 4$) will be parametrized by higher-order symmetric traceless tensors. However, at the order we are working at, these higher-order tensors will need to be contracted with one or several Kronecker deltas. Hence, they do not contribute to the stress tensor and will therefore not be considered in this work.

4.3.3 Effective Stress Tensor

Putting the results of the previous two sections together, the stress tensor becomes a functional of the long-wavelength fields[3] $\{v^i, \Phi \equiv 2\phi/(3\mathcal{H}^2\Omega_m)\}$, the non-dynamical fields $\{\Psi, \Psi^{ij}, \Psi^{ijkl}\}$, and their derivatives. In this section, we present the most general expression for the viscosity and noise parts of the stress tensor.

Viscosity part

We write the viscosity part of the stress tensor as

$$\langle \tau^{ij} \rangle_s = \mathcal{F}[\partial^i\partial^j\Phi, \partial^i v^j, \Psi, \Psi^{ij}, \Psi^{ijkl}, \cdots] \,, \tag{4.3.24}$$

where the ellipses refer to terms with higher derivatives. Since the long-wavelength fluctuations are perturbative, we can expand the functional $\mathcal{F}$ in (4.3.24) in powers of the fluctuations and their derivatives. We will focus on terms which involve single powers of the new long-wavelength fields $\{\Psi, \Psi^{ij}, \Psi^{ijkl}\}$, since the Gaussian terms have been discussed in [45, 47, 52, 73]. The stress tensor can then be written as the sum of three terms organized by spin:

$$\langle \tau^{ij} \rangle_s^{\rm NG} = f_{\rm NL} \left[\langle \tau^{ij} \rangle_s^{[0]} + \langle \tau^{ij} \rangle_s^{[2]} + \langle \tau^{ij} \rangle_s^{[4]} \right] \,. \tag{4.3.25}$$

We look at each of these contributions in turn.

- *Scalar contributions*

  The leading spin-0 contributions to the stress tensor are

  $$\begin{aligned} \frac{1}{\bar\rho}\langle \tau^{ij} \rangle_s^{[0]} &= \left[g\,\Psi + c_1\,\Psi\delta + \frac{c_2}{\mathcal{H}}\,\Psi\theta \right]\delta^{ij} + c_3\,\Psi\partial^i\partial^j\Phi \\ &\quad + \frac{c_4}{2\mathcal{H}}\,\Psi\left(\partial^i v^j + \partial^j v^i\right) \,, \end{aligned} \tag{4.3.26}$$

  where the coefficients $g(\tau)$ and $c_i(\tau)$ are dimensionless and time dependent. At second order in perturbation theory, the expression in (4.3.26)

[3] To avoid clutter in the expressions, we will drop the subscripts on the long-wavelength fields from now on: i.e. we set $\{v^i_\ell, \Phi_\ell\} \to \{v^i, \Phi\}$.

simplifies, since some terms are related by the equations of motion. In particular, using $v^i_{(1)} = -[\mathcal{H}f]\partial^i\Phi_{(1)}$, with $f \equiv d\ln D_1/d\ln a$, we get

$$\boxed{\frac{1}{\bar{\rho}}\langle\tau^{ij}\rangle^{[0]}_s \;=\; \big[g\Psi + g_1\Psi\delta\big]\delta^{ij} + g_2\Psi s^{ij}} \;, \tag{4.3.27}$$

where we have defined $g_1 \equiv c_1 + \frac{1}{3}c_3 - f\,c_2$, $g_2 \equiv c_3 - f\,c_4$ and introduced the tidal tensor $s_{ij} \equiv \partial_i\partial_j\Phi - \frac{1}{3}\delta_{ij}\triangle\Phi$.

- *Higher-spin contributions*

  Similarly, the leading spin-2 and spin-4 contributions to the stress tensor are

$$\frac{1}{\bar{\rho}}\langle\tau^{ij}\rangle^{[2]}_s \;=\; \tilde{g}\,\Psi^{ij} + \tilde{g}_1\,\Psi^{ij}\,\delta + \tilde{g}_2(\Psi^{ik}s_k{}^j + \Psi^{jk}s_k{}^i) + \tilde{g}_3\,\Psi^{kl}s_{kl}\,\delta^{ij} \;, \tag{4.3.28}$$

$$\frac{1}{\bar{\rho}}\langle\tau^{ij}\rangle^{[4]}_s \;=\; \hat{g}\,\Psi^{ijkl}s_{kl} \;. \tag{4.3.29}$$

We see that a bispectrum with a squeezed limit of the form (4.3.7), generally yields eight additional parameters in the stress tensor (for each value of Δ_i). When the squeezed limit is isotropic, the number of additional parameters reduces to three.

Noise part

The noise term in (4.3.3) arises because, for any specific realization, the short-scale modes fluctuate away from their ensemble averages. This stochastic contribution to the effective stress tensor can also be expanded in powers of the long-wavelength fields $\{\Phi, \Psi, \Psi^{ij}, \Psi^{ijkl}\}$. Just like the viscosity term in (4.3.25), the non-Gaussian contributions to the noise term can be organized in terms of spin:

$$(\Delta\tau^{ij})^{\rm NG} = f_{\rm NL}\left[(\Delta\tau^{ij})^{[0]} + (\Delta\tau^{ij})^{[2]} + (\Delta\tau^{ij})^{[4]}\right] \;. \tag{4.3.30}$$

We look at each of these contributions in turn.

- *Scalar contributions*

  The scalar contribution is given by

$$\boxed{\frac{1}{\bar{\rho}}(\Delta\tau^{ij})^{[0]} \;=\; J_{\psi}^{ij}\,\Psi}\;, \tag{4.3.31}$$

  where J_{ψ}^{ij} is a random variable representing stochastic noise, i.e. contributions which are uncorrelated with the long-wavelength fluctuations. This noise term also correlates with the leading-order Gaussian noise term, $(\Delta\tau^{ij})^{\mathrm{G}} \supset J_0^{ij}$. As we shall see in Section 4.4, the correlation of these two noise terms will be important in the renormalization of non-Gaussian loop diagrams. Furthermore, noise terms are uncorrelated on large scales, i.e. their correlation functions in position space are proportional to delta functions. This means that, in Fourier space, the correlator of J_{ψ}^{ij} with J_0^{ij} is

$$\langle J_0^{ij}(\boldsymbol{k},\tau)J_{\psi}^{kl}(\boldsymbol{k}',\tau)\rangle = \mathcal{J}^{ijkl}(\boldsymbol{k},\tau)(2\pi)^3\delta_D(\boldsymbol{k}+\boldsymbol{k}')\;, \tag{4.3.32}$$

  where $\mathcal{J}^{ijkl}(\boldsymbol{k},\tau)$ is an analytic function, which around $\boldsymbol{k}=\boldsymbol{0}$ can be expanded as

$$\mathcal{J}^{ijkl}(\boldsymbol{k},\tau) = \mathcal{J}(\Lambda,\tau)\delta^{ij}\delta^{kl} + \hat{\mathcal{J}}(\Lambda,\tau)\left[\delta^{ik}\delta^{jl}+\delta^{il}\delta^{jk}\right] + \mathcal{O}(k^2)\;. \tag{4.3.33}$$

  Furthermore, as advertised earlier, the noise terms of the stress tensor, $J_{0,\psi}^{ij}$, also correlate with the noise term in the initial conditions, ψ_J:

$$\begin{aligned}\langle\psi_J(\boldsymbol{k})J_a^{ij}(\boldsymbol{k}',\tau)\rangle' &= \int_{\boldsymbol{p}} K_{\mathrm{NL}}(\boldsymbol{k}-\boldsymbol{p},\boldsymbol{p})\langle\varphi_g^s(\boldsymbol{k}-\boldsymbol{p})\varphi_g^s(\boldsymbol{p})J_a^{ij}(\boldsymbol{k}',\tau)\rangle' \\ &= \tilde{\mathcal{J}}_a(\Lambda,\tau)\delta^{ij}+\mathcal{O}(k^2)\;,\end{aligned} \tag{4.3.34}$$

  where $a\in\{0,\psi\}$ and we have assumed that the integral is analytic in $\boldsymbol{k}$.

- *Higher-spin contributions*

  The leading spin-2 and spin-4 contributions to the noise part of the

stress tensor are

$$\frac{1}{\bar{\rho}}(\Delta\tau^{ij})^{[2]} = J_\psi^{ijkl}\Psi_{kl} \ , \tag{4.3.35}$$

$$\frac{1}{\bar{\rho}}(\Delta\tau^{ij})^{[4]} = J_\psi^{ijklmn}\Psi_{klmn} \ , \tag{4.3.36}$$

where the tensors J_ψ^{ijkl} and J_ψ^{ijklmn} are uncorrelated with long-wavelength fluctuations, but can be correlated with ψ_J and J_0^{ij}.

Both the viscosity and noise parts of the stress tensor are crucial in a consistent renormalization of matter correlation functions. As we will see in the next section, they generate new solutions which have precisely the correct momentum dependence to absorb the divergences arising from the loop diagrams of standard perturbation theory.

4.4 Renormalization

In this section, we describe the one-loop renormalization of the power spectrum and the bispectrum in the presence of primordial non-Gaussianity. We will focus on loop diagrams arising from non-Gaussian initial conditions, since loops coming from Gaussian initial conditions have already been studied in [45, 51, 52]. Moreover, we will restrict the presentation to the case of isotropic (spin-0) initial conditions. The generalization to PNG with non-trivial angular dependence is straightforward, and left as an exercise to the reader.

In §4.4.1, we classify the divergences arising in the non-Gaussian loops of standard perturbation theory. We then show, in §4.4.2, that the solutions generated by the stress tensor (4.3.27) act as the appropriate counterterms. When we compute correlators of the field Ψ it will be convenient to expand the right-hand side of (4.3.13) around the Eulerian position $\boldsymbol{x}$:

$$\begin{aligned}\Psi(\boldsymbol{x},\tau) &= \psi(\boldsymbol{q}(\boldsymbol{x},\tau)) \\ &= \psi(\boldsymbol{x}) + \boldsymbol{\nabla}\psi(\boldsymbol{x})\cdot\boldsymbol{\nabla}\Phi(\boldsymbol{x},\tau) + \cdots \ ,\end{aligned} \tag{4.4.1}$$

where, in the last line, we have used linear perturbation theory to write the velocity v^i in terms of the gradient of the gravitational potential $\partial^i\Phi$. The

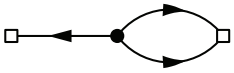
correlation between δ and ψ is

$$P_{1\psi}(k) \equiv \langle \delta_{(1)}(\boldsymbol{k},\tau)\psi(-\boldsymbol{k})\rangle' = \frac{(k/\mu)^{\Delta}}{M(k)} P_{11}(k) \ , \tag{4.4.2}$$

where Δ and $M(k)$ were defined in (4.3.9) and (4.2.2), respectively. Finally, in §4.4.3, we discuss the time dependence of the renormalized EFT parameters.

4.4.1 Loops in Standard Perturbation Theory

There is only one non-Gaussian contribution to the one-loop power spectrum

$P_{12}(k)$ =  $= \int_{\boldsymbol{p}} F_2(\boldsymbol{p}, \boldsymbol{k}-\boldsymbol{p}) B_{111}(k, p, |\boldsymbol{k}-\boldsymbol{p}|)$

$+$ perm . (4.4.3)

Extracting the UV behavior of this integral, we find the following divergence

$$P_{12}(k)\Big|_{\Lambda\to\infty} = -\frac{1}{21} f_{\rm NL} \left[2\sigma^2(\Lambda)\, k^2\, P_{1\psi}(k) + \sigma^2_{\psi}(\Lambda)\, k^2 M(k)\right] + \text{perm} \ , \tag{4.4.4}$$

where we have defined

$$\sigma^2(\Lambda) \equiv \int^{\Lambda} \frac{\mathrm{d}p}{2\pi^2} \frac{a_0}{(p/\mu)^{\Delta}} P_{11}(p) \ , \tag{4.4.5}$$

$$\sigma^2_{\psi}(\Lambda) \equiv \int^{\Lambda} \frac{\mathrm{d}p}{2\pi^2} \frac{K_{\rm NL}(\boldsymbol{p},-\boldsymbol{p})}{[M(p)]^2} P^2_{11}(p) \ . \tag{4.4.6}$$

Recall that $M(p)$ and $P_{11}(p)$ are evaluated at time τ (or scale factor a). The coefficients $\sigma^2(\Lambda)$ and $\sigma^2_{\psi}(\Lambda)$ are therefore also time dependent. In (4.4.6), we assumed that $K_{\rm NL}(\boldsymbol{p},-\boldsymbol{p})$ is independent of the direction of the loop momentum $\boldsymbol{p}$. We see that the first term in (4.4.4) is proportional to $P_{1\psi}(k)$, and will therefore be renormalized by a counterterm proportional to ψ. On the other hand, the second divergence in (4.4.4) is proportional to $M(k)$ and is renormalized by the noise contribution to the initial conditions ψ_J.

There are four non-Gaussian contributions to the bispectrum (see fig. 4.2):

$$B_{113}^{(\mathrm{I})} = 3P_{11}(k_2)\int_{\boldsymbol{p}} F_3(\boldsymbol{k}_1+\boldsymbol{p},-\boldsymbol{p},\boldsymbol{k}_2)B_{111}(k_1,p,|\boldsymbol{k}_1+\boldsymbol{p}|) + 5 \text{ perms} , \tag{4.4.7}$$

$$B_{113}^{(\mathrm{II})} = 3B_{111}(k_1,k_2,k_3)\int_{\boldsymbol{p}} F_3(\boldsymbol{k}_1,\boldsymbol{p},-\boldsymbol{p})P_{11}(p) + 2 \text{ perms} , \tag{4.4.8}$$

$$B_{122}^{(\mathrm{I})} = 4\int_{\boldsymbol{p}} F_2(\boldsymbol{k}_3+\boldsymbol{p},-\boldsymbol{p})F_2(\boldsymbol{p},\boldsymbol{k}_2-\boldsymbol{p})B_{111}(k_1,|\boldsymbol{k}_3+\boldsymbol{p}|,|\boldsymbol{k}_2-\boldsymbol{p}|)P_{11}(p) + 2 \text{ perms} , \tag{4.4.9}$$

$$B_{122}^{(\mathrm{II})} = F_2(\boldsymbol{k}_1,\boldsymbol{k}_2)P_{11}(k_2)P_{12}(k_1) + 5 \text{ perms} . \tag{4.4.10}$$

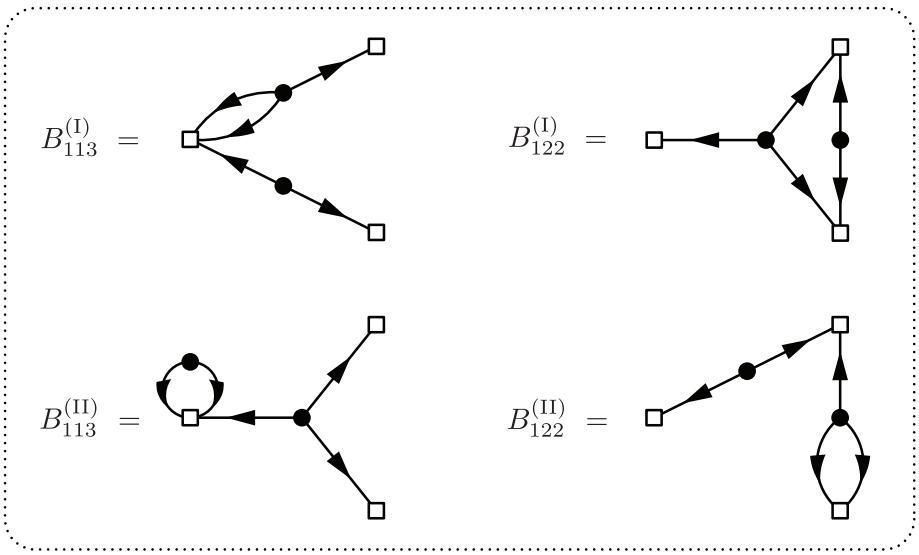


Figure 4.2: Diagrammatic representation of the non-Gaussian contributions to the one-loop bispectrum. The diagrams of type II are renormalized by the same counterterms that renormalize the one-loop power spectrum.

We note that the divergences appearing in $B_{122}^{(\mathrm{II})}$ already arise in P_{12} and are therefore renormalized by the same counterterms. Hence, we turn our attention to the diagrams $B_{113}^{(\mathrm{I})}$, $B_{113}^{(\mathrm{II})}$ and $B_{122}^{(\mathrm{I})}$:

- The UV limit of $B_{113}^{(\mathrm{I})}$ is

$$B_{113}^{(\mathrm{I})}\Big|_{p\to\infty} = -\left[\frac{4}{105}k_3^2 + \frac{1}{21}\frac{\boldsymbol{k}_1\cdot\boldsymbol{k}_2}{k_2^2}\,k_3^2 + \frac{7}{45}\frac{(\boldsymbol{k}_3\cdot\boldsymbol{k}_2)^2}{k_2^2}\right] \times f_{\mathrm{NL}}\left[2\sigma^2(\Lambda)P_{1\psi}(k_1) + \sigma_\psi^2(\Lambda)M(k_1)\right]P_{11}(k_2) \;+\; 5 \text{ perms} \,, \tag{4.4.11}$$

  where $\sigma^2(\Lambda)$ and $\sigma_\psi^2(\Lambda)$ were defined in (4.4.5) and (4.4.6), respectively. Let us focus on the term proportional to $\boldsymbol{k}_1\cdot\boldsymbol{k}_2$. It is easy to see that this contribution is renormalized by the counterterm $\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\Phi$. However, in order for this term to transform correctly under boosts, it has to appear in the combination

$$\psi(\boldsymbol{x}) + \boldsymbol{\nabla}\psi(\boldsymbol{x})\cdot\boldsymbol{\nabla}\Phi(\boldsymbol{x},\tau) \,, \tag{4.4.12}$$

  i.e. the field ψ has to be evaluated at the Lagrangian position $\boldsymbol{q}(\boldsymbol{x},\tau)$; cf. eq. (4.4.1). For this to happen, the divergences in (4.4.4) and (4.4.11) have to have the same numerical coefficient (highlighted in red). It is reassuring that our computation reproduces this fact.

- The UV limit of $B_{113}^{(\mathrm{II})}$ is

$$B_{113}^{(\mathrm{II})}\Big|_{p\to\infty} = -\frac{61}{630}\sigma_v^2(\Lambda)\left(k_1^2+k_2^2+k_3^2\right)B_{111}(k_1,k_2,k_3) \,, \tag{4.4.13}$$

  where we have defined

$$\sigma_v^2(\Lambda) \equiv \int^\Lambda \frac{\mathrm{d}p}{2\pi^2}\,P_{11}(p) \,. \tag{4.4.14}$$

  Since the divergence in (4.4.13) is the same as that appearing in P_{13} and $B_{123}^{(\mathrm{II})}$ (cf. fig. 3.3), it will be renormalized by the same (Gaussian) counterterm.

- The UV limit of $B_{122}^{(\mathrm{I})}$ is

$$B_{122}^{(\mathrm{I})}\Big|_{p\to\infty} = -\frac{26}{147}f_{\mathrm{NL}}\left[k_2^2k_3^2 - \frac{40}{13}(\boldsymbol{k}_2\cdot\boldsymbol{k}_3)^2\right] \times\left[2\hat{\sigma}^2(\Lambda)P_{1\psi}(k_1) + \hat{\sigma}_\psi^2(\Lambda)\,M(k_1)\right] + 2 \text{ perms} \,, \tag{4.4.15}$$

where we have defined

$$\hat{\sigma}^2(\Lambda) \equiv \int^{\Lambda} \frac{\mathrm{d}p}{2\pi^2} \frac{a_0}{p^2(p/\mu)^{\Delta}} P_{11}^2(p) \ , \tag{4.4.16}$$

$$\hat{\sigma}_{\psi}^2(\Lambda) \equiv \int^{\Lambda} \frac{\mathrm{d}p}{2\pi^2} \frac{K_{\rm NL}(\boldsymbol{p}, -\boldsymbol{p})}{p^2[M(p)]^2} P_{11}^3(p) \ . \tag{4.4.17}$$

We notice that all terms in (4.4.15) are *analytic* in two of the external momenta. For instance, the first permutation is analytic in k_2 and k_3. In position space, these divergences therefore are proportional to derivatives of delta functions. As we will explain below, these terms are renormalized by the noise term in the effective stress tensor (4.3.31).

4.4.2 Renormalization in the EFT-of-LSS

Next, we demonstrate that the stress tensor (4.3.27) generates new solutions which precisely cancel the cutoff dependence arising from the loop diagrams in SPT. Many of the computational details will be relegated to Appendix 4.C.

Including the stress tensor, the equations of motion for δ and θ are

$$\mathcal{D}_{\delta}\delta \equiv \mathcal{H}^2 \left[-a^2\partial_a^2 + \left(\frac{3}{2}\Omega_m - 3 \right) a\partial_a + \frac{3}{2}\Omega_m \right] \delta = \mathcal{S}_{\beta} - \mathcal{H}\partial_a(a\mathcal{S}_{\alpha}) + \tau_{\theta} \ , \tag{4.4.18}$$

$$\begin{aligned} \mathcal{D}_{\theta}\theta \equiv \mathcal{H}^2 & \left[+a^2\partial_a^2 + \left(4 - \frac{3}{2}\Omega_m \right) a\partial_a + (2 - 3\Omega_m) \right] \theta \\ & = \partial_a(a\mathcal{S}_{\beta}) - \frac{3}{2}\Omega_m \mathcal{H}\mathcal{S}_{\alpha} + \partial_a(a\tau_{\theta}) \ , \end{aligned} \tag{4.4.19}$$

where ∂_a denotes a derivative with respect to the scale factor, $\mathcal{S}_{\alpha}$ and $\mathcal{S}_{\beta}$ are nonlinear source terms defined in Appendix 4.C, and

$$\tau_{\theta} \equiv -\partial_i \left[\frac{1}{\rho} \partial_j \tau^{ij} \right] \ . \tag{4.4.20}$$

Substituting (4.3.27) and (4.3.31), we find that τ_{θ} is the sum of a contribution from $\langle \tau^{ij} \rangle_s$ and one from the stochastic term $\Delta\tau^{ij}$:

$$\tau_{\theta} = \tau_v + \tau_n \ , \tag{4.4.21}$$

where, to second order in the fluctuations and including the Gaussian contributions (see [45, 52]), we have

$$\tau_v = -d^2 \triangle\delta - e_1 \triangle(\delta^2) - e_2 \triangle(s^2) - e_3 \partial_i(s^{ij}\partial_j\delta) \\ - f_{\rm NL}\big[g\big(\triangle\Psi - \partial_i(\delta\partial^i\Psi)\big) + g_1 \triangle(\Psi\delta) + g_2\partial_i\partial_j(\Psi s^{ij})\big] \ , \tag{4.4.22}$$

$$\tau_n = -\partial_i\big[\partial_j J_0^{ij} - (\delta\partial_j J_0^{ij})\big] - \partial_i\partial_j(J_1^{ij}\delta) - \partial_i\partial_j(J_2^{ij}{}_{kl}s^{kl}) - f_{\rm NL}\partial_i\partial_j(J_\psi^{ij}\Psi) \ . \tag{4.4.23}$$

The full solution $\{\delta, \theta\}$ can then be written as a sum of three terms

$$\delta = \delta^{\rm SPT} + \delta^c + \delta^J \quad \text{and} \quad \theta = \theta^{\rm SPT} + \theta^c + \theta^J \ , \tag{4.4.24}$$

where $\{\delta^{\rm SPT}, \theta^{\rm SPT}\}$ is the SPT solution (see Appendix 4.C), while $\{\delta^c, \theta^c\}$ and $\{\delta^J, \theta^J\}$ are the solutions generated by τ_v and τ_n, respectively. In (4.C.12), we expand the SPT solution in powers of the first-order initial condition δ_1. A similar expansion can be defined for the counterterms

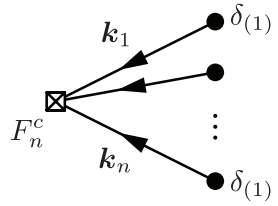
$$\delta^{c,J}(a) = \sum_{n=1}^{\infty} \delta^{c,J}_{(n)}(a) \ , \tag{4.4.25}$$

where $\delta^c_{(n)} \propto (\delta_1)^n$ and $\delta^J_{(n)} \propto (\delta_1)^{n-1}$, and equivalently for θ. In particular, the viscosity counterterm $\delta^c_{(n)}$ can be expressed as the sum of a Gaussian and a non-Gaussian contribution

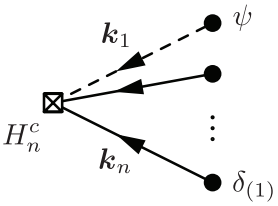
$$\delta^c_{(n)}(\boldsymbol{k}, a) = \\ \int_{\boldsymbol{k}_1} \dots \int_{\boldsymbol{k}_n} (2\pi)^3 \delta_D\big(\boldsymbol{k} - \boldsymbol{k}_{1\dots n}\big)\, F_n^c(\boldsymbol{k}_1, \dots, \boldsymbol{k}_n | a)\, \delta_{(1)}(\boldsymbol{k}_1, a) \dots \delta_{(1)}(\boldsymbol{k}_n, a) + \\ f_{\rm NL} \int_{\boldsymbol{k}_1} \dots \int_{\boldsymbol{k}_n} (2\pi)^3 \delta_D\big(\boldsymbol{k} - \boldsymbol{k}_{1\dots n}\big)\, H_n^c(\boldsymbol{k}_1, \dots, \boldsymbol{k}_n | a)\, \psi(\boldsymbol{k}_1) \dots \delta_{(1)}(\boldsymbol{k}_n, a) \ , \tag{4.4.26}$$

where the kernel functions F_n^c and H_n^c have the following diagrammatical

representations:



$$\equiv \quad F_n^c(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n|a)\,(2\pi)^3\delta_D\big(\boldsymbol{k}-\boldsymbol{k}_{1\ldots n}\big)\ , \tag{4.4.27}$$



$$\equiv \quad f_{\rm NL}\,H_n^c(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n|a)\,(2\pi)^3\delta_D\big(\boldsymbol{k}-\boldsymbol{k}_{1\ldots n}\big)\ . \tag{4.4.28}$$

Our goal in this section is to compute δ^c and δ^J up to second order. Naturally, both δ^c and δ^J have contributions from the Gaussian and non-Gaussian terms of the stress tensor. However, since the Gaussian contributions have already been computed in [45, 52, 73], we will mainly focus our attention on the non-Gaussian contributions. Furthermore, since we are only interested in showing that the counterterms cancel the cutoff dependence arising from the SPT loop diagrams, we will, for simplicity, restrict the analysis to the Einstein-de Sitter cosmology. The extension to ΛCDM will be explained in more detail in the next section.

Viscosity counterterms

- The first-order counterterm $\delta^c_{(1)}$ is the solution to

$$\mathcal{D}_\delta\,\delta^c_{(1)} = -d^2(a)\,\triangle\,\delta_{(1)} - g(a)f_{\rm NL}\,\triangle\,\psi\ . \tag{4.4.29}$$

The Green's functions associated with the operator $\mathcal{D}_\delta$ is defined in Appendix 4.C. It allows us to write the solution of (4.4.29) as

$$\delta^c_{(1)}(\boldsymbol{k},a) = -\xi(a)k^2\delta_{(1)}(\boldsymbol{k},a) - \gamma(a)f_{\rm NL}k^2\psi(\boldsymbol{k})\ , \tag{4.4.30}$$

where we have defined[4]

$$\xi(a) \equiv -\frac{1}{a}\int_{a_{in}}^{a} \mathrm{d}a'\, G_\delta(a,a')\, a' d^2(a') \ , \tag{4.4.31}$$

$$\gamma(a) \equiv -\int_{a_{in}}^{a} \mathrm{d}a'\, G_\delta(a,a')\, g(a') \ . \tag{4.4.32}$$

The non-Gaussian contribution to $\delta^c_{(1)}$ cancels the loop divergence in P_{12}:

$P_{12} + P_{1c} \supset$  $+$

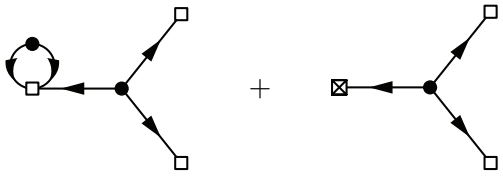
$$= -f_{\rm NL}\left(\frac{4}{21}\sigma^2(\Lambda) + 2\gamma(a)\right) k^2 P_{1\psi}(k) \ = \ \text{finite} \ , \tag{4.4.33}$$

where, for the moment, we have only focused on the divergence proportional to $P_{1\psi}(k)$. (The divergence proportional to $M(k)$ in (4.4.4) will be cancelled by a noise term.) Hence, the cutoff dependence of the one-loop power spectrum is removed if

$$\gamma(\Lambda,a) = -\frac{2}{21}\sigma^2(\Lambda,a) + \text{finite} \quad \Leftrightarrow \quad g(\Lambda,a) = -\frac{1}{3a}\sigma^2(\Lambda,a) + \text{finite} \ , \tag{4.4.34}$$

where we have added the explicit arguments to highlight that both γ and σ are functions of the cutoff Λ and the time a. The finite piece of (4.4.34) will be discussed in §4.4.3.

The Gaussian contribution to $\delta^c_{(1)}$ cancels the loop divergence of $B^{(\rm II)}_{113}$:

$B^{(\rm II)}_{113} + B_{11c} \supset$ 

$$= -\left[\frac{61}{630}\sigma_v^2(\Lambda) + \xi(a)\right](k_1^2 + k_2^2 + k_3^2) B_{111}(k_1,k_2,k_3) \ . \tag{4.4.35}$$

[4]Note that the parameter γ of [45] corresponds to the parameter ξ in this Chapter.

Hence, renormalization requires

$$\xi(\Lambda,a) = -\frac{61}{630}\sigma_v^2(\Lambda,a) + \text{finite} \quad \Leftrightarrow \quad d^2(\Lambda,a) = -\frac{61}{180}\frac{1}{a}\sigma_v^2(\Lambda,a) + \text{finite}\ . \tag{4.4.36}$$

Note that the parameter ξ is fixed by the renormalization of the Gaussian one-loop diagram P_{13} [47].

- The second-order counterterm $\delta^c_{(2)}$ satisfies

$$\mathcal{D}_\delta\, \delta^c_{(2)} = \mathcal{S}^{(2)}_\beta - \mathcal{H}\partial_a(a\mathcal{S}^{(2)}_\alpha) + \tau^{(2)}_v\ , \tag{4.4.37}$$

where $\mathcal{S}^{(2)}_{\alpha,\beta}$ are obtained by replacing one of the δ (or θ) in the nonlinear source terms by their first-order SPT solution and the other one by the first-order counterterm $\delta^c_{(1)}$ (or $\theta^c_{(1)} = -\dot{\delta}^c_{(1)}$). Focusing on the solution generated by the non-Gaussian terms, we find

$$\begin{aligned}\delta^{c,\mathrm{NG}}_{(2)}(\boldsymbol{k},a) &= \int_{a_{in}}^{a} \mathrm{d}a'\ G_\delta(a,a')\left[\mathcal{S}^{(2)}_\beta(a') - \mathcal{H}\partial_{a'}(a'\mathcal{S}^{(2)}_\alpha(a')) + \tau^{(2)}_v(a')\right] \\ &= f_{\mathrm{NL}}\int_{\boldsymbol{p}} H^c_2(\boldsymbol{p},\boldsymbol{k}-\boldsymbol{p}|a)\,\psi(\boldsymbol{p})\delta_{(1)}(\boldsymbol{k}-\boldsymbol{p},a)\ .\end{aligned} \tag{4.4.38}$$

The kernel H^c_2 receives contributions from $\mathcal{S}_{\alpha,\beta}$ and τ_v, and can be written as

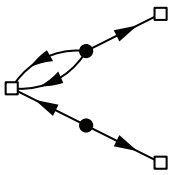
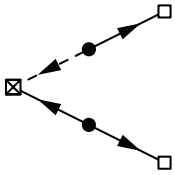
$$\begin{aligned}H^c_2(\boldsymbol{k}_1,\boldsymbol{k}_2|a) \equiv &-\gamma(a)\big[G_\Psi(\boldsymbol{k}_1,\boldsymbol{k}_2) + G_{\alpha\beta}(\boldsymbol{k}_1,\boldsymbol{k}_2)\big] \\ &-\gamma_1(a)G_1(\boldsymbol{k}_1,\boldsymbol{k}_2) - \gamma_2(a)G_2(\boldsymbol{k}_1,\boldsymbol{k}_2)\ ,\end{aligned} \tag{4.4.39}$$

where we have defined

$$\gamma_i(a) \equiv -\frac{1}{a}\int_{a_{in}}^{a} \mathrm{d}a'\ G_\delta(a,a')\,a' g_i(a')\ . \tag{4.4.40}$$

Explicit expressions for the kernels $G_{\{\Psi,\alpha\beta,i\}}$ are given in Appendix 4.C. Let us note that the value of $\gamma(a)$ has already been fixed by the renormalization of the power spectrum and therefore can no longer be adjusted. Substituting (4.4.34) for $\gamma(a)$, the sum of the loop diagram and the counterterm

becomes

$B^{(\mathrm{I})}_{113} + B_{11c} \supset$  $+$  $=$

$$= \left[\left(\frac{74}{567}\sigma^2(\Lambda) + \gamma_1(a) \right) k_3^2 + \left(\frac{52}{135}\sigma^2(\Lambda) + \gamma_2(a) \right) \left(\frac{(\boldsymbol{k}_3 \cdot \boldsymbol{k}_2)^2}{k_2^2} - \frac{1}{3}k_3^2 \right) \right] \times (-f_{\mathrm{NL}}) P_{1\psi}(k_1) P_{11}(k_2) + 5 \text{ perms} \,. \tag{4.4.41}$$

The cutoff dependence is therefore cancelled if

$$\gamma_1(\Lambda, a) = -\frac{74}{567}\sigma^2(\Lambda, a) + \text{finite} \quad \Leftrightarrow \quad g_1(\Lambda, a) = -\frac{74}{63}\frac{1}{a}\sigma^2(\Lambda, a) + \text{finite} \,, \tag{4.4.42}$$

$$\gamma_2(\Lambda, a) = -\frac{52}{135}\sigma^2(\Lambda, a) + \text{finite} \quad \Leftrightarrow \quad g_2(\Lambda, a) = -\frac{52}{15}\frac{1}{a}\sigma^2(\Lambda, a) + \text{finite} \,. \tag{4.4.43}$$

We have shown that every counterterm generated by the viscosity contribution to the stress tensor, τ_v, yields the correct momentum dependence to remove some of the loop divergences in the power spectrum and the bispectrum. Next, we show that the remaining divergences are cancelled by the noise counterterms. This part is for aficionados and can be skipped on a first reading.

Noise counterterms

Noise terms appear both in the initial conditions (4.3.4) and in the stress tensor (4.3.31). They are essential in the renormalization procedure as they remove some divergences in P_{12} and $B^{(\mathrm{I})}_{113}$, as well as all the divergences in $B^{(\mathrm{I})}_{122}$.

As indicated in (4.4.25), the noise term can be expanded in powers of the long-wavelength fluctuations. The n-th order solution $\delta^J_{(n)}$ receives contributions from both Gaussian and non-Gaussian initial conditions:

$$\delta^J_{(n)}(\boldsymbol{k}, a) = \delta^{J,\mathrm{G}}_{(n)}(\boldsymbol{k}, a) + \delta^{J,\mathrm{NG}}_{(n)}(\boldsymbol{k}, a) \,. \tag{4.4.44}$$

Specifically, at one-loop order, we will require the first-order Gaussian contribution

$$\delta^{J,\mathrm{G}}_{(1)}(\boldsymbol{k},a) = k_i k_j N^{ij}_0(\boldsymbol{k},a)\ , \tag{4.4.45}$$

and the first and second-order non-Gaussian contributions

$$\delta^{J,\mathrm{NG}}_{(1)}(\boldsymbol{k},a) = f_{\mathrm{NL}}\, M(k)\,\psi_J(\boldsymbol{k})\ , \tag{4.4.46}$$

$$\delta^{J,\mathrm{NG}}_{(2)}(\boldsymbol{k},a) = f_{\mathrm{NL}}\int_{\boldsymbol{p}}\Big[k_i k_j N^{ij}_{\psi}(\boldsymbol{k}-\boldsymbol{p},a)\psi(\boldsymbol{p}) + 2F_2(\boldsymbol{k}-\boldsymbol{p},\boldsymbol{p})\delta_{(1)}(\boldsymbol{k}-\boldsymbol{p},a)M(p)\psi_J(\boldsymbol{p})\Big]\ , \tag{4.4.47}$$

where

$$N^{ij}_{0,\psi}(\boldsymbol{k},a) \equiv \int_{a_{in}}^{a} \mathrm{d}a'\ G_\delta(a,a')\, J^{ij}_{0,\psi}(\boldsymbol{k},a')\ . \tag{4.4.48}$$

We will now show that these counterterms cancel the divergences proportional to $M(k)$ in both P_{12} and $B^{(\mathrm{I})}_{113}$ and all the divergences in $B^{(\mathrm{I})}_{122}$:

- The divergence proportional to $M(k)$ in P_{12} is renormalized by the counterterm $\delta^J_{(1)}$. More precisely, we have

$$\begin{aligned} P_{12}+P_{JJ} \ &\supset\ -\frac{2}{21}f_{\mathrm{NL}}\sigma^2_\psi(\Lambda)\,k^2M(k) + 2\,\langle\delta^{J,\mathrm{NG}}_{(1)}(\boldsymbol{k},a)\delta^{J,\mathrm{G}}_{(1)}(-\boldsymbol{k},a)\rangle' \\ &=\ f_{\mathrm{NL}}\left[-\frac{2}{21}\sigma^2_\psi(\Lambda) + 2\int_{a_{in}}^{a}\mathrm{d}a'\ G_\delta(a,a')\tilde{\mathcal{J}}_0(a')\right]k^2M(k) \\ &=\ \text{finite}\ , \end{aligned} \tag{4.4.49}$$

  where $\tilde{\mathcal{J}}_0(a')$ was defined in (4.3.34). Hence, the cutoff dependence is eliminated provided that

$$\tilde{\mathcal{J}}_0(\Lambda,a) = -\frac{1}{6\,a}\sigma^2_\psi(\Lambda,a) + \text{finite}\ . \tag{4.4.50}$$

  Similarly, the divergence proportional to $M(k)$ in $B^{(\mathrm{II})}_{122}$ (which is the same as the divergence in P_{12}), is renormalized by the second term in (4.4.47).

- Let us consider the following correlation between the first-order SPT solution and the noise terms

$$\begin{aligned} B_{1JJ} \supset\ & \langle \delta^{\rm SPT}_{(1)}(\boldsymbol{k}_1,a)\delta^{J,{\rm NG}}_{(1)}(\boldsymbol{k}_2,a)\delta^{J,{\rm G}}_{(2)}(\boldsymbol{k}_3,a)\rangle' \\ &+\ \langle \delta^{\rm SPT}_{(1)}(\boldsymbol{k}_1,a)\delta^{J,{\rm G}}_{(1)}(\boldsymbol{k}_2,a)\delta^{J,{\rm NG}}_{(2)}(\boldsymbol{k}_3,a)\rangle' \\ &\equiv\ B^{(A)}_{1JJ} + B^{(B)}_{1JJ} \end{aligned} \tag{4.4.51}$$

The first term, $B^{(A)}_{1JJ}$, cancels the divergence proportional to $M(k)$ in $B^{(\rm I)}_{113}$, while the second term, $B^{(B)}_{1JJ}$, cancels the first divergence in $B^{(\rm I)}_{122}$. We will demonstrate the second fact explicitly and leave the first as an exercise to the reader.

Using (4.4.46) and (4.4.47), we can write

$$\begin{aligned} B^{(B)}_{1JJ} = (k_2)_i(k_2)_j(k_3)_k(k_3)_l\langle N^{ij}_0(\boldsymbol{k}_2,a)N^{kl}_\psi(-\boldsymbol{k}_2,a)\rangle' P_{1\psi}(k_1) \\ + 5 \text{ perms} , \end{aligned} \tag{4.4.52}$$

where $N^{ij}_{0,\psi}$ was defined in (4.4.48). We will assume that the cutoff-dependent part of $J^{ij}_{0,\psi}$ satisfies $J^{ij}_{0,\psi} \propto a$, so that $N^{ij}_{0,\psi} = -\frac{2}{7}aJ^{ij}_{0,\psi}$ and

$$\langle N^{ij}_0(\boldsymbol{k},a)N^{kl}_\psi(-\boldsymbol{k},a)\rangle' = \frac{4}{49}a^2\left[\mathcal{J}(a)\delta^{ij}\delta^{kl} + \hat{\mathcal{J}}(a)\left(\delta^{ik}\delta^{jl}+\delta^{il}\delta^{jk}\right)\right] . \tag{4.4.53}$$

Hence, we find that

$$\begin{aligned} B^{(\rm I)}_{122} + B^{(B)}_{1JJ} \supset\ & f_{\rm NL}\Bigg[\left(\frac{4}{49}\mathcal{J}(a)a^2 - \frac{52}{147}\hat{\sigma}^2(\Lambda)\right)k_2^2k_3^2 + \\ &+\left(\frac{8}{49}\hat{\mathcal{J}}(a)a^2 + \frac{160}{147}\hat{\sigma}^2(\Lambda)\right)(\boldsymbol{k}_2\cdot\boldsymbol{k}_3)^2\Bigg]P_{1\psi}(k_1) + \text{perms} \\ =\ & \text{finite} . \end{aligned} \tag{4.4.54}$$

We see that $B^{(B)}_{1JJ}$ has the right momentum dependence to absorb the first divergence in $B^{(\rm I)}_{122}$. In particular, the cutoff dependence is can-

celled if

$$\begin{aligned}\mathcal{J}(\Lambda,a) =& \frac{13}{3a^2}\tilde{\sigma}^2(\Lambda,a) + \text{finite} \quad \text{and} \\ \hat{\mathcal{J}}(\Lambda,a) =& -\frac{20}{3}\frac{1}{a^2}\hat{\sigma}^2(\Lambda,a) + \text{finite} \,. \end{aligned} \tag{4.4.55}$$

- Finally, we consider

$$\begin{aligned} B_{JJJ} \supset & \langle \delta^{J,\text{NG}}_{(1)}(\boldsymbol{k}_1,a)\delta^{J,\text{G}}_{(1)}(\boldsymbol{k}_2,a)\delta^{J,\text{G}}_{(1)}(\boldsymbol{k}_3,a)\rangle' + \text{perms} \\ &= f_{\text{NL}}\, M(k_1)(k_2)_i(k_2)_j(k_3)_k(k_3)_l \langle \psi_J(\boldsymbol{k}_1) J_0^{ij}(\boldsymbol{k}_2,a) J_0^{kl}(\boldsymbol{k}_3,a)\rangle' \,. \end{aligned} \tag{4.4.56}$$

It is difficult to determine the precise momentum dependence of this correlation function, since the probability distributions of the noise terms are unknown. We will assume that it can be expanded around $\boldsymbol{k}_i = \boldsymbol{0}$, so that we can write

$$\begin{aligned}\langle \psi_J(\boldsymbol{k}_1) J_0^{ij}(\boldsymbol{k}_2,a) J_0^{kl}(\boldsymbol{k}_3,a)\rangle' =& \mathcal{N}(a)\delta^{ij}\delta^{kl} + \hat{\mathcal{N}}(a)\big(\delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk}\big) \\ &+ \mathcal{O}(k_i^2) \,. \end{aligned} \tag{4.4.57}$$

In that case, we get

$$\begin{aligned} B^{(\text{I})}_{122} + B_{JJJ} \supset & \; f_{\text{NL}}\Bigg[\left(-\frac{26}{147}\hat{\sigma}^2_\psi(\Lambda) + \mathcal{N}(a)\right) k_2^2 k_3^2 \\ &+ \left(\frac{80}{147}\hat{\sigma}^2_\psi(\Lambda) + 2\hat{\mathcal{N}}(a)\right)(\boldsymbol{k}_2\cdot\boldsymbol{k}_3)^2\Bigg] M(k_1) + \text{perms} \\ =& \; \text{finite} \,. \end{aligned} \tag{4.4.58}$$

We see that B_{JJJ} has the correct momentum dependence to absorb the cutoff dependence of the second divergence in (4.4.15). More precisely, we have

$$\begin{aligned}\mathcal{N}(\Lambda,a) =& \frac{26}{147}\hat{\sigma}^2_\psi(\Lambda,a) + \text{finite} \quad \text{and} \\ \hat{\mathcal{N}}(\Lambda,a) =& -\frac{40}{147}\hat{\sigma}^2_\psi(\Lambda,a) + \text{finite} \,. \end{aligned} \tag{4.4.59}$$

We have shown that the solutions generated by the stress tensor derived in Section 4.3 have the correct momentum dependence to cancel all divergences coming from SPT loop diagrams. So far, we have focused on the cutoff-dependent parts of the EFT parameters. In the next section, we study their finite (or renormalized) parts.

4.4.3 Renormalized EFT Parameters

Any EFT parameter can be written as the sum of a cutoff-dependent part (the counterterm) and a finite (or "renormalized") part:

$$g(\Lambda, a) = g^{(0)}(\Lambda, a) + g^{(R)}(a) \ . \tag{4.4.60}$$

Ultimately, every long-wavelength observable must be independent of the cutoff. Indeed, once the large-scale correlation functions have been properly renormalized, one can send the cutoff to infinity. In ΛCDM, all loop integrals are convergent and the cutoff can be taken to infinity even *before* renormalization. This may seem to go against the EFT philosophy since we are including modes which are outside the regime of validity of the EFT. However, the mistakes one makes in doing so, can be absorbed into a shift of the renormalized EFT parameters.

The time dependence of the renormalized parameters may be different from that of the counterterms. Moreover, it cannot be computed within the EFT framework, but must be determined from simulations or observations. An exception is a "scaling universe" — i.e. a matter-dominated or Einstein-de Sitter (EdS) cosmology with scale-free initial conditions $P_{11}(k, a_{in}) \propto k^n$. In this special case, an additional symmetry constrains the time dependence of the parameters in the EFT. To see this, we first note that the equations of motion in EdS are invariant under a *Lifshitz scaling*

$$\boldsymbol{x} \mapsto \lambda_x \boldsymbol{x} \quad \text{and} \quad a \mapsto \lambda_a a \ . \tag{4.4.61}$$

The dimensionless power spectrum of the initial conditions, $\Delta^2_{11}(k, a_{in}) \propto a_{in}^2 k^{n+3}$, is also invariant under the Lifshitz scaling iff

$$\lambda_x \equiv \lambda_a^{2/(n+3)} \ . \tag{4.4.62}$$

When (4.4.62) holds, then the Lifshitz scaling maps one EdS solution to another with a different realization of the same statistical initial condition. The linearly-evolved power spectrum can be written in a manifestly self-similar form, $\Delta^2_{11}(k, a) = (k/k_{\rm NL})^{n+3}$, where the nonlinear scale satisfies $k_{\rm NL} \propto a^{-2/(n+3)}$. Nonlinear corrections to Δ^2_{11} appear as higher powers of $k/k_{\rm NL}$ [43, 74]. Assuming that the short-scale fluctuations satisfy the Lifshitz scaling, the terms in the effective stress tensor must also have the right

transformation properties. In particular, the symmetry fixes the combination of powers of k and a which can appear in the equations of motion and therefore constrains the time dependence of the renormalized parameters. For instance, using such an argument,[5] one finds that the speed of sound in a scaling universe must satisfy $c_s^2(a) \propto a^{(1-n)/(n+3)}$ [43].

For non-Gaussian initial conditions, we also need to impose that the higher-point correlation functions are invariant under the Lifshitz scaling (4.4.61) with λ_x and λ_a related by (4.4.62). More specifically, let us assume that the dimensionless bispectrum satisfies

$$\mathcal{B}_{111}\left(\frac{k_1}{\lambda_x}, \frac{k_2}{\lambda_x}, \frac{k_3}{\lambda_x}, \lambda_a\, a_{in}\right) = \lambda_x^m \lambda_a^3\, \mathcal{B}_{111}\left(k_1, k_2, k_3, a_{in}\right) \ , \tag{4.4.63}$$

where we have used that $\mathcal{B}_{111}(k_i, a_{in}) \propto a_{in}^3$. Given (4.4.62), the initial bispectrum is invariant under (4.4.61) iff $m = -\frac{3}{2}(n+3)$. Furthermore, if the primordial Gaussian potential φ_g is scale-invariant, then the field ψ must transform as $\psi(\boldsymbol{x}) \mapsto \lambda_x^{\Delta}\, \psi(\lambda_x \boldsymbol{x})$ under the Lifshitz scaling; cf. eq. (4.3.9). Using that the Lagrangian coordinate transforms as $\boldsymbol{q}(\boldsymbol{x}, a) \mapsto \lambda_x^{-1} \boldsymbol{q}(\lambda_x \boldsymbol{x}, \lambda_a a)$, we find that the field $\Psi(\boldsymbol{x}, a) = \psi(\boldsymbol{q}(\boldsymbol{x}, a))$ transforms as

$$\Psi(\boldsymbol{x}, a) \ \mapsto \ \lambda_x^{\Delta}\, \Psi(\lambda_x \boldsymbol{x}, \lambda_a a) \ . \tag{4.4.64}$$

This implies that the terms in the stress tensor (4.3.27) preserve the Lifshitz scaling iff their time dependence is

$$\left\{ g^{(R)}(a) \ , \ g^{(R)}_{1,2}(a) \right\} \propto a^{(1-n+2\Delta)/(n+3)} \propto \left[D_1(a)\right]^{(1-n+2\Delta)/(n+3)} \ . \tag{4.4.65}$$

In ΛCDM and for general initial conditions, the Lifshitz scaling isn't a symmetry and the initial conditions aren't scale-free. However, in fig. 4.3 we show that the initial matter power spectrum has momentum regimes in which a power law ansatz is a good approximation. This explains why, for Gaussian initial conditions, the ansatz deduced from the Lifshitz scaling symmetry has worked remarkably well [45, 52, 73]. We will extend these results and assume that the EFT parameters have the following time dependence[6]

$$g^{(R)}(a) = \left[\mathcal{H}(a) f(a)\right]^2 \left[D_1(a)\right]^{m_g+1}\, \bar{g}^{(R)} \ , \tag{4.4.66}$$

[5] Let us remark that this argument does not determine the time dependence of the counterterms $g^{(0)}(\Lambda, a)$, since the cutoff introduces a new scale and therefore breaks the Lifshitz scaling symmetry [43].

[6] Notice that we have included a factor $(\mathcal{H}f)^2$ which is not present in [45]. We found that this ansatz improves the accuracy of (4.C.11) by a few percents.

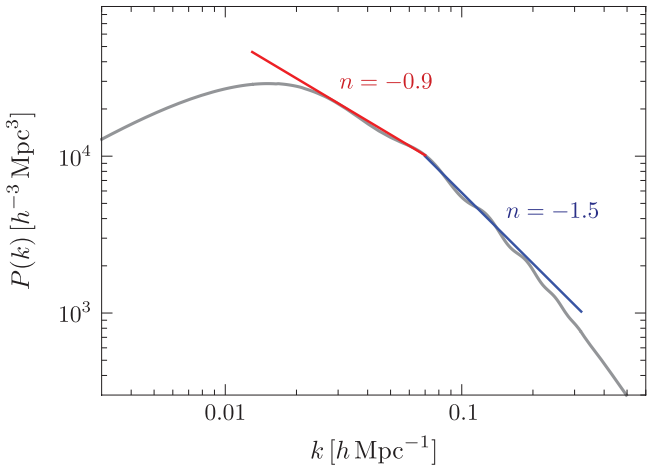


Figure 4.3: Scaling of the initial matter power spectrum. We see that in the momentum regions $[0.02, 0.07]\,h\,\mathrm{Mpc}^{-1}$ and $[0.07, 0.25]\,h\,\mathrm{Mpc}^{-1}$ the power spectrum is well approximated by a power law with $n \approx -0.9$ and $k_{\mathrm{NL}} \approx 0.16\,h\,\mathrm{Mpc}^{-1}$ (red line) and $n \approx -1.5$ and $k_{\mathrm{NL}} \approx 0.23\,h\,\mathrm{Mpc}^{-1}$ (blue line), respectively.

where m_g is a real parameter and $\bar{g}^{(R)}$ is a constant. At one loop, we don't need to know the precise time dependence of $g_{1,2}^{(R)}$ to compute the counterterms. The ansatz (4.4.66) is therefore only required for the parameter $g(a)$. In general, the value of m_g is unknown. In what follows we will assume that the time dependence matches the ansatz of the scaling universe, i.e. we will use $m_g = (1 - n + 2\Delta)/(n + 3)$. From fig. 4.3 we deduce that $n \in [-1.5, -0.9]$ for $k \in [0.02, 0.25]\,h\,\mathrm{Mpc}^{-1}$. For these values of n, we have $m_g \in [0.9 + \Delta, 1.7 + 1.3\,\Delta]$. For our numerical computations, we will choose $m_g = 1.7 + 1.3\,\Delta$. However, we have checked that the counterterms only change by at most a few percents as the value of m_g runs over this interval, so our numerical results will not depend sensitively on this choice.

4.5 Numerical Analysis

In this section, we present a numerical computation of the renormalized power spectrum and bispectrum for non-Gaussian initial conditions. We start, in §4.5.1, with a summary of the different contributions (SPT and

EFT) to the dark matter correlation functions. In §4.5.2, we estimate the relative sizes of the terms in a scaling universe. Finally, in §4.5.3, we make some preliminary observations on the shapes of the various contributions. A more detailed analysis will appear in [60].

4.5.1 Renormalized Correlation Functions

In §4.3.3, we wrote the effective stress tensor of the dark matter as a sum of terms made from the long-wavelength fields and their derivatives

$$\tau_v = \sum_n a_n \mathcal{O}_n \ , \tag{4.5.1}$$

where the coefficients a_n and the corresponding operators $\mathcal{O}_n$ relevant for this thesis are listed in Table 4.1.

$\mathcal{O}_n$	$\triangle\delta$	$\triangle(\delta^2)$	$\triangle(s^2)$	$\partial^i(s_{ij}\partial^j\delta)$	$\triangle\Psi - \partial_i(\delta\partial^i\Psi)$	$\triangle(\Psi\delta)$	$\partial^i\partial^j(\Psi s_{ij})$
a_n	d^2	e_1	e_2	e_3	g	g_1	g_2
α_n	ξ	ϵ_1	ϵ_2	ϵ_3	γ	γ_1	γ_2

Table 4.1: Operators and parameters in the EFT-of-LSS with primordial non-Gaussianities.

In §4.4.2, we showed that these operators gives rise to counterterms which renormalize the one-loop power spectrum and bispectrum of standard perturbation theory. Schematically, the solution can be written as (see §4.4.2 and §4.C.2 for more details)

$$\delta^c(a) \simeq \int_{a_{in}}^{a} \mathrm{d}a' \, G_\delta(a,a')\tau_v(a') \ \simeq \ \sum_n \alpha_n(a)\mathcal{O}_n(a) \ , \tag{4.5.2}$$

where the parameters α_n can be expressed as integrals over time of the corresponding parameters a_n. The SPT solution $\delta^{\mathrm{SPT}}(a)$ and the EFT counterterms $\delta^c(a)$ generate correlation functions for the renormalized dark matter density contrast $\delta(a)$ which depend both on the cosmological parameters of the ΛCDM model $\boldsymbol{\theta} \equiv \{\Omega_m^0, \Omega_\Lambda^0, h, \Delta_\varphi, n_s\}$ and on the EFT parameters $\boldsymbol{\alpha} \equiv \{\xi, \epsilon_{1,2,3}, \gamma, \gamma_{1,2}\}$.

- The dark matter power spectrum can be written as $P = P^{\rm G} + f_{\rm NL} P^{\rm NG}$, where $P^{\rm G}$ arises from the nonlinear evolution of Gaussian initial conditions [47] and $P^{\rm NG}$ is the contribution from non-Gaussian initial conditions. Both $P^{\rm G}$ and $P^{\rm NG}$ can be split into an SPT part and an EFT part. At one loop, we have

$$P^{\rm G}_{\rm SPT} = P_{11} + P_{13} + P_{22} \ , \qquad P^{\rm G}_{\rm EFT} = -2\xi k^2 P_{11} \ , \tag{4.5.3}$$

$$P^{\rm NG}_{\rm SPT} = P_{12} \ , \qquad P^{\rm NG}_{\rm EFT} = -2\gamma k^2 P_{1\psi} \ , \tag{4.5.4}$$

  where $P_{1\psi}(k)$ was defined in (4.4.2).

- The total dark matter bispectrum is $B = B^{\rm G} + f_{\rm NL} B^{\rm NG}$, where $B^{\rm G}$ is the part arising from the nonlinear evolution of Gaussian initial conditions [45, 52], and $B^{\rm NG}$ is the part associated with non-Gaussian initial conditions. Both $B^{\rm G}$ and $B^{\rm NG}$ can be split into an SPT part and an EFT part. At one loop, we have

$$B^{\rm G}_{\rm SPT} = B_{112} + \Big[B_{114} + B^{\rm (I)}_{123} + B^{\rm (II)}_{123} + B_{222} \Big] \ , \tag{4.5.5}$$

$$B^{\rm NG}_{\rm SPT} = B_{111} + \Big[B^{\rm (I)}_{113} + B^{\rm (II)}_{113} + B^{\rm (I)}_{122} + B^{\rm (II)}_{122} \Big] \ , \tag{4.5.6}$$

$$B^{\rm G}_{\rm EFT} = \xi B^{\rm G}_{\xi} + \sum_{i=1}^{3} \epsilon_i B_{\epsilon_i} \ , \tag{4.5.7}$$

$$B^{\rm NG}_{\rm EFT} = \xi B^{\rm NG}_{\xi} + \gamma B_{\gamma} + \sum_{i=1}^{2} \gamma_i B_{\gamma_i} \ . \tag{4.5.8}$$

  The individual bispectra in (4.5.7) and (4.5.8) can be written as

$$B^{\rm G}_{\xi} \equiv -2\big[E_{\alpha\beta}(\boldsymbol{k}_1, \boldsymbol{k}_2) + E_{\delta}(\boldsymbol{k}_1, \boldsymbol{k}_2) \big] P_{11}(k_1) P_{11}(k_2) + 2 \text{ perms} \ , \tag{4.5.9}$$

$$B^{\rm NG}_{\xi} \equiv -(k_1^2 + k_2^2 + k_3^2) B_{111}(k_1, k_2, k_3) \ , \tag{4.5.10}$$

$$B_{\epsilon_i} \equiv -2\, E_i(\boldsymbol{k}_1, \boldsymbol{k}_2)\, P_{11}(k_1) P_{11}(k_2) + 2 \text{ perms} \ , \tag{4.5.11}$$

$$B_{\gamma} \equiv -\big[G_{\alpha\beta}(\boldsymbol{k}_1, \boldsymbol{k}_2) + G_{\Psi}(\boldsymbol{k}_1, \boldsymbol{k}_2) \big] P_{11}(k_1) P_{1\psi}(k_2) + 5 \text{ perms} \ , \tag{4.5.12}$$

$$B_{\gamma_i} \equiv -G_i(\boldsymbol{k}_1, \boldsymbol{k}_2) P_{11}(k_1) P_{1\psi}(k_2) + 5 \text{ perms} \ , \tag{4.5.13}$$

where the kernel functions $E_{..}(\boldsymbol{k}_1, \boldsymbol{k}_2)$ and $G_{..}(\boldsymbol{k}_1, \boldsymbol{k}_2)$ are defined explicitly in Appendix 4.C.

For simplicity, we will take the cosmological parameters to be fixed and allow only the EFT parameters to vary (for a more complete treatment see [60]). Furthermore, we will assume that the sound speed parameter has been measured in the power spectrum and its value is fixed $\xi \equiv 1.5\, h^{-2}\,\mathrm{Mpc}^2$ [45]. The following parts of both the Gaussian and the non-Gaussian bispectra are therefore predicted

$$B_0^{\mathrm{I}} \equiv B_{\mathrm{SPT}}^{\mathrm{I}} + \xi B_\xi^{\mathrm{I}} \ , \quad \mathrm{I} = \mathrm{G}, \mathrm{NG} \ . \tag{4.5.14}$$

The total bispectrum can then be written as

$$\boxed{B = B_0^{\mathrm{G}} + B_c^{\mathrm{G}} + f_{\mathrm{NL}} \left(B_0^{\mathrm{NG}} + B_c^{\mathrm{NG}} \right)} \ , \tag{4.5.15}$$

where we have defined the sum of the additional counterterm contributions as

$$B_c^{\mathrm{G}} \equiv \sum_{i=1}^{3} \epsilon_i B_{\epsilon_i} \ , \tag{4.5.16}$$

$$B_c^{\mathrm{NG}} \equiv \gamma B_\gamma + \sum_{i=1}^{2} \gamma_i B_{\gamma_i} \ . \tag{4.5.17}$$

In general, the coefficients in (4.5.16) and (4.5.17) cannot be predicted, but need to be measured in N-body simulations or in observations. For the numerical results of §4.5.3, we will estimate the sizes of these coefficients by looking at the one-loop divergences that they cancel. In particular, since in our universe loop integrals are convergent, they are in practice evaluated with the cutoff Λ taken to infinity. In that case, the SPT loop diagrams contain integrals of the linear power spectrum extrapolated to scales beyond the linear regime (i.e. $k > k_{\mathrm{NL}}$). This is not a problem, since, as we explained in §4.4.3, this finite error is removed by adjusting the counterterms. More precisely, from (4.4.34) we infer that the term in γ which cancels this known error is

$$|\gamma| \sim \frac{2}{21} \int_{k_{\mathrm{NL}}}^{\infty} \frac{\mathrm{d}p}{2\pi^2} \frac{a_0}{(p/\mu)^\Delta} P_{11}(p) + \cdots \ , \tag{4.5.18}$$

where $k_{\rm NL}$ is the nonlinear scale. We will choose $k_{\rm NL} = 0.2\,h\,{\rm Mpc}^{-1}$, which corresponds to the nonlinear scale found in [45, 52]. For local non-Gaussianity (i.e. $\Delta = 0$), we then find $|\gamma| \sim 1.6\,h^{-2}\,{\rm Mpc}^2$. Note that, for $\Delta \neq 0$, the precise value of μ is unimportant as it cancels in the product γB_γ.

Of course, there is also an unknown, finite contribution to γ that accounts for the effects of the nonperturbative short scales on the long-wavelength dynamics. We will assume that there is no fine-tuned cancellation between this finite part and the term shown in (4.5.18). The result in (4.5.18) then gives an approximate lower bound on the size of the counterterm. Applying the same estimate to γ_1 and γ_2, we find

$$\gamma_1 \approx \frac{37}{27}\gamma \quad \text{and} \quad \gamma_2 \approx \frac{182}{45}\gamma\ . \tag{4.5.19}$$

Let us stress that the estimates in (4.5.18) and (4.5.19) only provide approximated lower bounds and should not be considered precise evaluations.

In §4.5.3, we will present numerical results for three different primordial bispectrum shapes:

- Local non-Gaussianity is obtained by setting $K_{\rm NL} = 1$ in (4.2.8). The primordial bispectrum then is

$$B_\varphi^{\rm local}(k_1,k_2,k_3) = \ 2f_{\rm NL}^{\rm local}A^2\left[\frac{1}{k_1^3k_2^3}+\frac{1}{k_1^3k_3^3}+\frac{1}{k_2^3k_3^3}\right]\ , \tag{4.5.20}$$

  where $A \equiv 2\pi^2\Delta_\varphi^2$. For simplicity, we have written the bispectrum for scale-invariant initial conditions, $n_s = 1$. The latest CMB constraint on the amplitude of the local bispectrum is $f_{\rm NL}^{\rm local} = 1.8 \pm 5.6$ [75].

- Higher-derivative corrections to slow-roll inflation produce equilateral non-Gaussianity [33, 65]. In the effective field theory of inflation [33], this is captured by two cubic operators, $\dot\pi^3$ and $\dot\pi(\partial_i\pi)^2$, for the Goldstone boson of broken time translations, π. Both operators produce very similar equilateral bispectra. For purposes of illustration, we focus on the PNG produced by the operator $\dot\pi^3$, which has a bispectrum of the form

$$B_\varphi^{\rm equil}(k_1,k_2,k_2) = 162\,f_{\rm NL}^{\rm equil}\cdot\frac{A^2}{k_1k_2k_3}\frac{1}{K^3}\ , \tag{4.5.21}$$

where $K \equiv k_1 + k_2 + k_3$. The latest CMB constraint on the amplitude of the equilateral bispectrum is $f_{\rm NL}^{\rm equil} = -9.2 \pm 69$ [75].

- The shape of the bispectrum in quasi-single-field inflation cannot be computed analytically. However, it is well approximated by the following ansatz [66, 76]

$$B_\varphi^{\rm QSFI}(k_1, k_2, k_3) = 18\sqrt{3}\, f_{\rm NL}^{\rm QSFI} \cdot \frac{A^2}{k_1 k_2 k_3} \frac{1}{K^3} \cdot \frac{1}{\sqrt{\kappa}} \frac{N_\nu[8\kappa]}{N_\nu[8/27]} \,, \quad (4.5.22)$$

where $\kappa \equiv k_1 k_2 k_3 / K^3$ and N_ν is the Neumann function of order ν. The parameter ν depends on the mass of the hidden sector field to which the inflaton field couples during inflation. It determines the scaling of the bispectrum in the squeezed limit via $\Delta \equiv \frac{3}{2} - \nu$ in (4.2.10). For purposes of illustration, we choose $\nu = \frac{1}{2}$ corresponding to $\Delta = 1$ which is precisely intermediate between the scaling of local and equilateral non-Gaussianity.

In our numerical computations, we consider a flat ΛCDM cosmology with the following parameters: $\Omega_m^0 = 0.27$, $\Omega_\Lambda^0 = 0.73$, $h = 0.70$. The amplitude of the primordial potential is $\Delta_\varphi^2(k_0) \simeq 8.7 \times 10^{-10}$, defined at the pivot scale $k_0 = 0.002\, h\,\mathrm{Mpc}^{-1}$. The one-loop correlators are computed numerically with Mathematica. For validation purposes, we wrote two independent codes. Both codes yielded results which are in very good agreement.[7]

Our results are summarized in figs. 4.4, 4.5 and 4.6 for local, equilateral and quasi-single-field PNG, respectively. The plots show the SPT and counterterm contributions to the bispectrum separated into Gaussian terms (which are the same in every plot) and non-Gaussian terms. Here, and in the following, the results are evaluated at redshift $z = 0$. Notice that while PNG has a larger effect on larger scales, we have chosen to show the "mildly non-linear" range of scales, where the EFT counterterms can play an important

[7] As a technical aside, let us note that the expression of the one-loop bispectrum in (4.5.6) is ill-suited for numerical evaluation as the different contributions each contain IR divergences. The equivalence principle guarantees that these divergences cancel when all diagrams are summed [74]. However, to avoid cancelling large numbers against each other, it is preferable to rewrite the integrand in a form which makes this cancellation manifest. In Appendix 4.D, we present the "IR-safe" integrands for the non-Gaussian loops.

role. For each type of PNG, we represent the corresponding matter bispectrum in the equilateral configuration (left) and for fixed $k_L \equiv 0.01\,h\,\mathrm{Mpc}^{-1}$ (right). We see that, in the equilateral configuration, the two-loop Gaussian contribution becomes comparable to the non-Gaussian contribution on rather large scales. In §4.5.2, we provide an analytical estimate of the critical scale k_c at which the two-loop Gaussian contribution becomes relevant. On the other hand, we see that the two-loop contribution is much more subdominant in squeezed configurations. This is also the momentum configuration in which the shape of PNG (and more precisely, the scaling Δ of the squeezed limit) can leave an appreciable imprint on the matter bispectrum. In §4.5.3, we discuss the possibility of extracting information about the shape of PNG from the matter bispectrum.

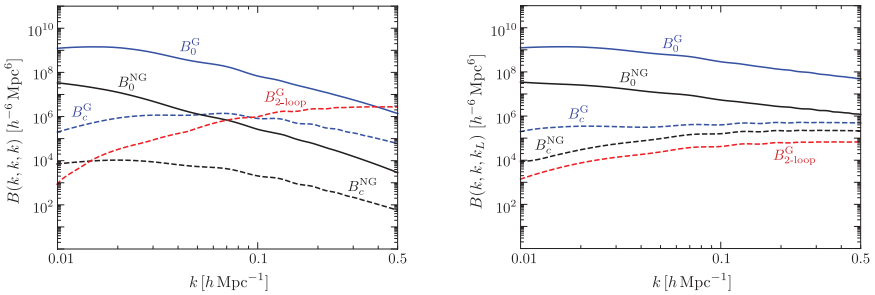


Figure 4.4: Contributions to the bispectrum for local non-Gaussianity with $f_{\mathrm{NL}}^{\mathrm{local}} = 10$, evaluated in the equilateral configuration (*left*) and for fixed $k_L \equiv 0.01\,h\,\mathrm{Mpc}^{-1}$ (*right*). The definitions of $B_0^{\mathrm{G,NG}}$ and $B_c^{\mathrm{G,NG}}$ can be found in (4.5.14), (4.5.16) and (4.5.17).

4.5.2 Estimates in a Scaling Universe

In this section, we determine the critical scale k_c at which the two-loop Gaussian contribution becomes relevant. For purpose of illustration, we will consider local PNG, although our result won't depend sensitively on the precise shape of the non-Gaussianity. To estimate the relative sizes of the various contributions to the power spectrum and bispectrum we consider a scaling universe (i.e. EdS with power-law spectrum $P_\delta(k,\tau_{in}) \propto k^n$). This extends the results of [43, 45] to non-Gaussian initial conditions.

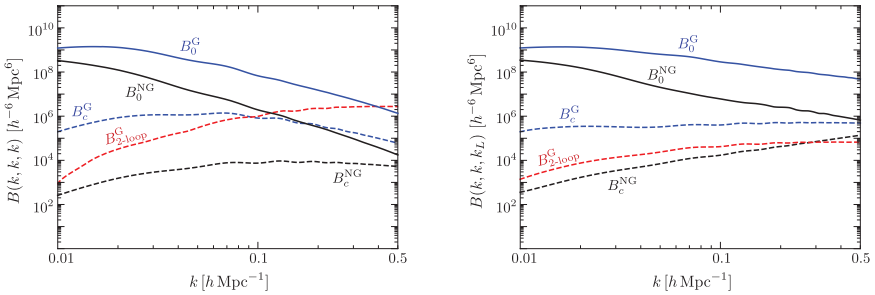


Figure 4.5: Same as fig. 4.4 but for equilateral non-Gaussianity with $f_{\rm NL}^{\rm equil} = 100$.

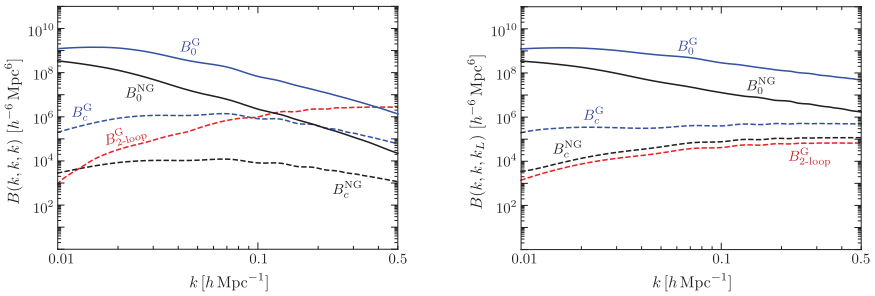


Figure 4.6: Same as fig. 4.4 but for quasi-single-field inflation with $\Delta = 1$ and $f_{\rm NL}^{\rm QSFI} = 100$.

As explained in §4.4.3, in a scaling universe, Lifshitz transformations of the form (4.4.62) leave the initial dimensionless power spectrum unchanged, which can therefore be written as a polynomial in $k/k_{\rm NL}$. Because the evolution is self-similar, the same applies to the nonlinear power spectrum

$$\Delta_\delta^2(k,\tau) \equiv \frac{k^3}{2\pi^2} P_\delta(k,\tau) = \Delta_\delta^2(k/k_{\rm NL}) \ . \tag{4.5.23}$$

Furthermore, the transformation (4.4.63) implies that the initial bispectrum is also self-similar, and that the nonlinear dimensionless bispectrum is therefore a polynomial in $k_i/k_{\rm NL}$:

$$\mathcal{B}_\delta(k_1,k_2,k_3,\tau) \equiv \left(\frac{k_1^3}{2\pi^2}\right)^2 B_\delta(k_1,k_2,k_3,\tau) = \mathcal{B}_\delta\big(k_1/k_{\rm NL}, k_2/k_{\rm NL}, k_3/k_{\rm NL}\big) \ . \tag{4.5.24}$$

The different contributions to the dimensionless power spectrum and bispectrum will scale as different powers of $k/k_{\rm NL}$.

Power Spectrum

Let us first collect the terms that contribute to the power spectrum $\Delta^2_\delta(k/k_{\rm NL})$. The momentum scalings of the Gaussian contributions were derived in [43]:

$$\Delta^2_{\delta,\rm G}(k) \;\subset\; s(n)\left(\frac{k}{k_{\rm NL}}\right)^p \;\leftarrow\; \begin{cases} p = n+3 & \approx 1.5 & \text{tree} \\ p = 2(n+3) & \approx 3.0 & \text{1-loop} \\ p = n+5 & \approx 3.5 & \text{LO vis.} \\ p = 3(n+3) & \approx 4.5 & \text{2-loop} \\ p = 2n+8 & \approx 5.0 & \text{NLO vis.} \\ p = n+7 & \approx 5.5 & \text{NLO h.d.} \\ p = 4(n+3) & \approx 6.0 & \text{3-loop} \\ p = 7 & \approx 7.0 & \text{noise} \end{cases}$$

where 'tree' refers to the linear power spectrum Δ^2_{11}, 'loop' stands for the SPT loop contributions, 'LO vis.' is the leading-order viscosity counterterm defined in (4.5.3), 'NLO vis.' are the next-to-leading-order viscosity counterterms [43], 'NLO h.d.' are next-to-leading-order higher-derivative terms, and 'noise' refers to Δ^2_{JJ}. The overall coefficient $s(n)$ is a (computable) order-one number, which is different for the different contributions (see Appendix A of [43]). For the numerical estimates of the scaling index p we have used $n \approx -1.5$, which corresponds to the scaling of the linear power spectrum in the regime $k \in [0.07, 0.25]\, h\,\text{Mpc}^{-1}$ (see fig. 4.3). The scaling for $k \in [0.02, 0.07]\, h\,\text{Mpc}^{-1}$, which is $n \approx -0.9$, will also be relevant, but hasn't been shown explicitly.

For the non-Gaussian contributions to the power spectrum $\Delta^2_{\delta,\rm NG}(k)$, we

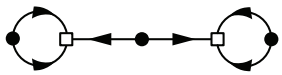
find

$$s(n)\, f_{\rm NL}\Delta_\varphi \left(\frac{k}{k_{\rm NL}}\right)^p \;\leftarrow\; \begin{cases} p = \frac{3}{2}(n+3) & \approx 2.25 & \text{1-loop} \\ p = \frac{1}{2}(n+7)+\Delta & \approx 2.75+\Delta & \text{LO vis.} \\ p = \frac{5}{2}(n+3) & \approx 3.75 & \text{2-loop} \\ p = \frac{7}{2}(n+3) & \approx 5.25 & \text{3-loop} \\ p = \frac{1}{2}(n+13) & \approx 5.75 & \text{noise} \end{cases}$$

where 'LO vis.' refers to the leading order non-Gaussian viscosity counterterm, defined in (4.5.4), 'loop' stands for the SPT loop contributions and 'noise' stands for Δ^2_{JJ}. The scalings of the viscosity counterterm and the noise term are fixed by the k-dependent parts of the UV-limit of Δ^2_{12} (see §4.4.1). We note that the amplitude of the non-Gaussian contributions is suppressed by a factor of $f_{\rm NL}\Delta_\varphi \sim 3\times 10^{-5} f_{\rm NL}$. Higher-loop corrections are suppressed by additional factors of $\Delta^2_{11}(k) = (k/k_{\rm NL})^{n+3}$. As before, the overall coefficient $s(n)$ is a computable order-one number which is different for the different contributions.

We see that the non-Gaussian contributions to the power spectrum are highly suppressed. We should therefore ask when the Gaussian two-loop terms (which we haven't included in our analysis) are of the same order. The two-loop Gaussian contribution is estimated through a single representative term, namely $P^{\rm (I)}_{33}$ [51]:

$$P^{\rm G}_{\text{2-loop}} \;\subset\; P^{\rm (I)}_{33} \;=\; \tag{4.5.25}$$



A rough estimate of the scale k_c at which the two-loop Gaussian contribution equals the one-loop non-Gaussian contribution is

$$\frac{P^{\rm G}_{\text{2-loop}}}{P^{\rm NG}_{\text{1-loop}}} \approx \frac{1}{f_{\rm NL}\Delta_\varphi}\frac{s^{\rm (I)}_{33}}{s_{12}}\left(\frac{k_c}{k_{\rm NL}}\right)^{\frac{3}{2}(n+3)} = 1 \quad\longrightarrow\quad k_c \approx 0.18\, k_{\rm NL}\left(\frac{f_{\rm NL}}{10}\right)^{1/3}, \tag{4.5.26}$$

where $s^{\rm (I)}_{33}$ and s_{12} are the coefficients of $P^{\rm (I)}_{33}$ and P_{12}, respectively. For the numerical estimate in (4.5.26) we have used[8] $n \approx -0.9$, $k_{\rm NL} \approx 0.16\, h\,{\rm Mpc}^{-1}$

[8] These values are applicable to the momentum range $k \in [0.02, 0.07]\, h\,{\rm Mpc}^{-1}$ (see fig. 4.3), which is roughly where the two-loop Gaussian contribution is expected to be larger than the one-loop non-Gaussian contribution (see fig. 4.7).

and $s_{33}^{(\mathrm{I})}/s_{12} \simeq 0.06$. In fig. 4.7, we compare this estimate for $k_c(f_{\mathrm{NL}})$ with the value obtained from a numerical computation in a ΛCDM universe. We see that already on relatively large scales the two-loop Gaussian term is of the same order as the one-loop non-Gaussian term.

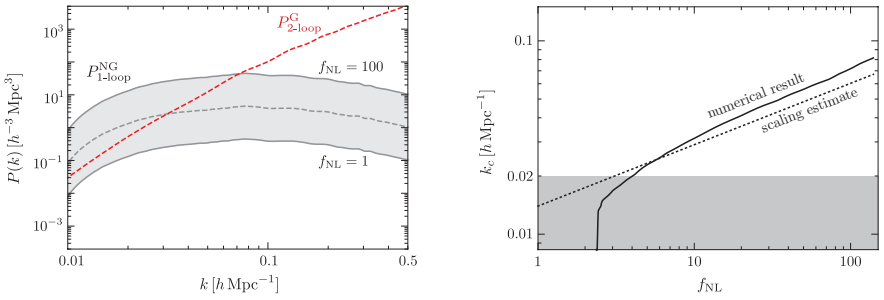


Figure 4.7: Comparison between the 1-loop contribution to the power spectrum for local PNG with $f_{\mathrm{NL}} \in [1, 100]$ (grey band) and an estimate of the 2-loop Gaussian contribution (red dashed line). The right plot shows the critical momentum value k_c at which the two-loop Gaussian contribution becomes larger than the one-loop non-Gaussian contribution as a function of f_{NL}, evaluated numerically (solid line) and using the estimate (4.5.26) (dashed line). The numerical result is for a ΛCDM cosmology, while the estimate is for a scaling universe. The shaded region represents the momentum range where the scaling $n \approx -0.9$ is no longer valid in the real universe.

Bispectrum

For the bispectrum, we focus our attention on the equilateral configuration, $\mathcal{B}_\delta(k, k, k)$, for which the two-loop corrections are largest. For Gaussian initial conditions, the scalings of the various contributions were derived

in [45]:

$$\mathcal{B}_\delta^{\rm G}(k) \;\subset\; s(n)\left(\frac{k}{k_{\rm NL}}\right)^p \;\leftarrow\; \begin{cases} p=2(n+3) & \approx 3.0 & \text{tree} \\ p=3(n+3) & \approx 4.5 & \text{1-loop} \\ p=2(n+4) & \approx 5.0 & \text{NLO vis.} \\ p=4(n+3) & \approx 6.0 & \text{2-loop} \\ p=3n+11 & \approx 6.5 & \text{NNLO vis.} \\ p=2n+10 & \approx 7.0 & \text{NNLO h.d.} \\ p=5(n+3) & \approx 7.5 & \text{3-loop} \\ p=n+10 & \approx 8.5 & \text{noise} \end{cases}$$

where 'NLO vis.' and 'NNLO vis.' refer to the next-to-leading-order and next-to-next-to-leading-order contributions to the Gaussian counterterm, respectively. More precisely, the 'NLO vis.' term corresponds to the term $\mathcal{B}_{\rm EFT}^{\rm G}$ defined in (4.5.7), while the 'NNLO vis.' term corresponds to the one-loop diagrams formed with the leading Gaussian counterterms. The term 'NNLO h.d.' represents higher-derivative contributions and the term 'noise' corresponds to $\mathcal{B}_{1JJ}$.

For the non-Gaussian contributions to the bispectrum $\mathcal{B}_\delta^{\rm NG}(k)$, we find

$$s(n)\, f_{\rm NL}\Delta_\varphi\left(\frac{k}{k_{\rm NL}}\right)^p \;\leftarrow\; \begin{cases} p=\frac{3}{2}(n+3) & \approx 2.25 & \text{tree} \\ p=\frac{5}{2}(n+3) & \approx 3.75 & \text{1-loop} \\ p=\frac{1}{2}(3n+13) & \approx 4.25 & \text{LO vis. (1)} \\ p=\frac{1}{2}(3n+13)+\Delta & \approx 4.25+\Delta & \text{LO vis. (2)} \\ p=\frac{7}{2}(n+3) & \approx 5.25 & \text{2-loop} \\ p=\frac{1}{2}(5n+19)+\Delta & \approx 5.75+\Delta & \text{NLO vis.} \\ p=\frac{1}{2}(3n+17)+\Delta & \approx 6.25+\Delta & \text{NLO h.d.} \\ p=\frac{1}{2}(3n+19) & \approx 7.25 & \text{noise (A)} \\ p=\frac{1}{2}(n+17)+\Delta & \approx 7.75+\Delta & \text{noise (B)} \\ p=\frac{1}{2}(n+23) & \approx 10.75 & \text{noise (C)} \end{cases}$$

where 'LO vis.' and 'NLO vis.' refer to the leading-order and next-to-leading-order contributions to the non-Gaussian counterterm $\mathcal{B}_c^{\rm NG}$, respectively. More precisely, the term 'LO vis. (1)' refers to $\mathcal{B}_\xi^{\rm NG}$ (4.5.10), the term

'LO vis. (2)' refers to $\mathcal{B}_c^{\rm NG}$ (4.5.17), and the term 'NLO vis.' refers to the finite part of the one-loop diagrams formed with one counterterm. The term 'NLO h.d.' represents higher-derivative contributions and the terms 'noise (A)', 'noise (B)' and 'noise (C)' refer to the noise terms $\mathcal{B}_{1JJ}^{(A)}$, $\mathcal{B}_{1JJ}^{(B)}$ and $\mathcal{B}_{JJJ}$ (cf. §4.4.2). As before, the scalings of the viscosity and noise contributions to the bispectrum are determined by the k-dependence of the UV-limit of the loop diagrams they cancel (cf. §4.4.1).

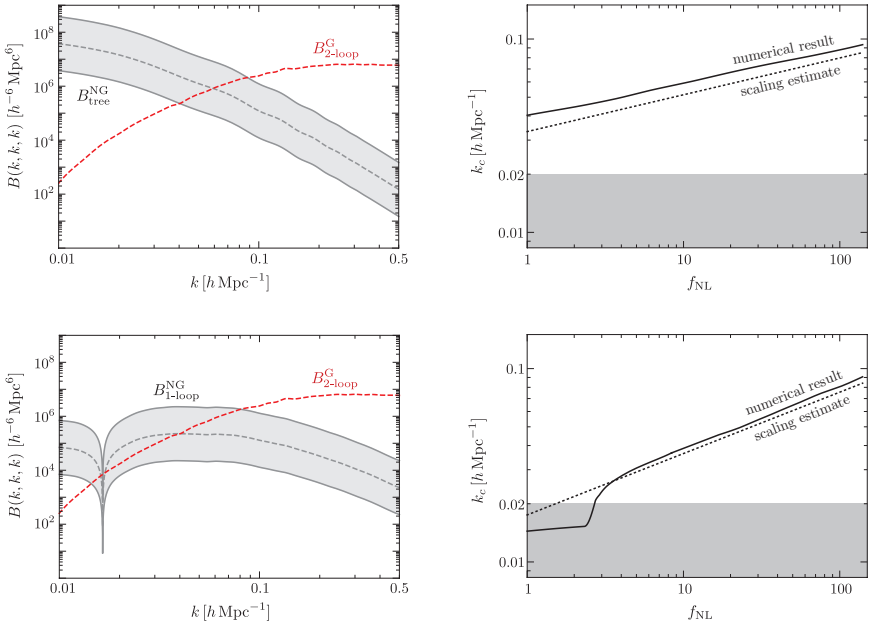
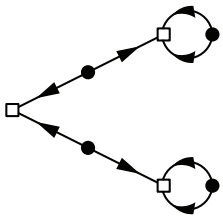


Figure 4.8: *Top*: Comparison between the tree-level contribution to the bispectrum for local PNG with $f_{\rm NL} \in [1, 100]$ (grey band) and an estimate of the 2-loop Gaussian contribution (red dashed line). *Bottom*: Comparison between the 1-loop non-Gaussian contribution (gray band) and the 2-loop Gaussian contribution (red dashed line). The curves in the right plots have the same meaning as in fig. 4.7.

Since the non-Gaussian contributions to the bispectrum already appear at tree level, they are less suppressed than in the power spectrum. Again, we should ask up to which point we can neglect the Gaussian two-loop corrections. To estimate the two-loop corrections, we will consider one rep-

resentative contribution, namely $B^{(\mathrm{I})}_{233}$ [52]:

$$B^{\mathrm{G}}_{\text{2-loop}} \subset B^{(\mathrm{I})}_{233} =$$



(4.5.27)

We can then determine the scales at which the two-loop bispectrum starts to become important relative to tree-level bispectrum B_{111} and the non-Gaussian one-loop contribution $B^{\mathrm{NG}}_{\text{1-loop}}$:

$$\frac{B^{\mathrm{G}}_{\text{2-loop}}}{B^{\mathrm{NG}}_{\text{tree}}} \sim \frac{1}{f_{\mathrm{NL}}\Delta_\varphi}\frac{s^{(\mathrm{I})}_{233}}{s_{111}}\left(\frac{k_c}{k_{\mathrm{NL}}}\right)^{\frac{5}{2}(n+3)} = 1 \longrightarrow k_c \approx 0.32\, k_{\mathrm{NL}}\left(\frac{f_{\mathrm{NL}}}{10}\right)^{1/5}, \tag{4.5.28}$$

$$\frac{B^{\mathrm{G}}_{\text{2-loop}}}{B^{\mathrm{NG}}_{\text{1-loop}}} \sim \frac{1}{f_{\mathrm{NL}}\Delta_\varphi}\frac{s^{(\mathrm{I})}_{233}}{s^{\mathrm{NG}}_{\text{1-loop}}}\left(\frac{k_c}{k_{\mathrm{NL}}}\right)^{\frac{3}{2}(n+3)} = 1 \longrightarrow k_c \approx 0.30\, k_{\mathrm{NL}}\left(\frac{f_{\mathrm{NL}}}{10}\right)^{1/3}. \tag{4.5.29}$$

As before, we have used $n \simeq -0.9$ to estimate the numerical values of $k_c(f_{\mathrm{NL}})$. For the numerical prefactors we have used $s^{(\mathrm{I})}_{233}/s_{111} \simeq 0.11$ and $s^{(\mathrm{I})}_{233}/s^{\mathrm{NG}}_{\text{1-loop}} \simeq 0.01$. In fig. 4.8 these estimates are compared to the exact result of a numerical computation in ΛCDM. Again, we find that two-loop corrections becomes relevant at relatively large scales.

We wish to emphasize that the importance of Gaussian two-loop corrections is maximal in the equilateral configuration. Away from this limit the two-loop amplitude is significantly smaller. In fig. 4.9, we show tree-level and one-loop contributions associated with local non-Gaussianity for fixed $k_L \equiv 0.01\, h\,\mathrm{Mpc}^{-1}$. This time the two-loop correction is subdominant for all scales, even for relatively small values of f_{NL}.

4.5.3 Comments on Primordial Shapes

The squeezed limit of the bispectrum contains interesting information about the spectrum of particles during inflation. In particular, it can reveal the

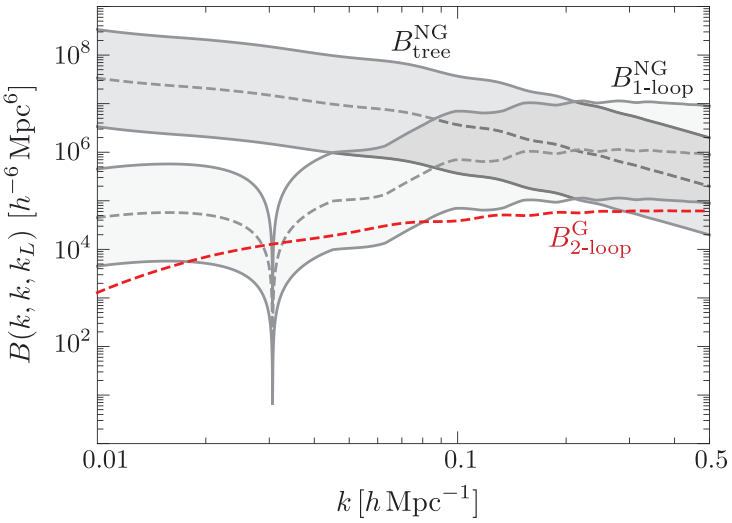


Figure 4.9: Comparison of the non-Gaussian tree-level and one-loop contributions with an estimate of the Gaussian two-loop contribution, evaluated for fixed $k_L \equiv 0.01\,h\,\mathrm{Mpc}^{-1}$.

masses [66, 77] and spins [69] of particles to which the inflaton couples. This requires an accurate measurement of the scaling dimension Δ in the primordial bispectrum (4.2.10). In [76], the detectability of this signature was discussed for future galaxy surveys. However, their analysis only used the tree-level form of the dark matter bispectrum, so we should ask when the nonlinear corrections discussed in this Chapter become important. (See [78] for related observations in the special case of local non-Gaussianity.)

In fig. 4.10, we show the non-Gaussian contributions to the bispectrum $B^{\rm NG}_{\rm SPT}$, normalized with respect to the tree-level bispectrum of the local type $B^{\rm local}_{111}$. We have plotted three different primordial bispectra — local, equilateral and quasi-single-field — whose shapes were defined in §4.5.1. The dashed lines show the different scaling behavior of the tree-level contributions. We see clearly that the one-loop corrections become important at relatively large scales, especially as we go further away from the scaling of local PNG. Moreover, the loop contributions are much less sensitive to the value of Δ. In fact, the one-loop contributions for equilateral and quasi-single-field scale almost the same. This suggests that adding the one-loop

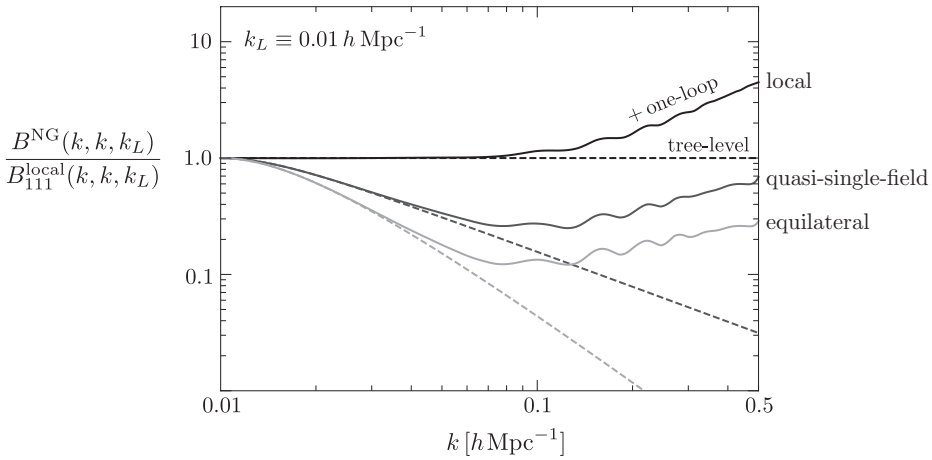


Figure 4.10: Comparison of the non-Gaussian SPT contribution $B_{\mathrm{SPT}}^{\mathrm{NG}}$ for local, equilateral and QSF shape, normalized to the local tree-level bispectrum B_{111}^{local}. Dashed lines refer to non-Gaussian tree-level contributions, while solid lines include one-loop corrections. We see that the loop corrections become relevant on relatively large scales and tend to *decrease* the difference among the three primordial signals.

corrections is important and that it can make it harder to extract Δ (relative to the tree-level expectation). We will study this question in more detail in [60].

4.6 Conclusions

Measurements of primordial non-Gaussianity in the cosmic microwave background are nearly saturated — correlations on large angular scales can't be measured more precisely because of cosmic variance, while Silk damping limits the information that can be extracted from small angular scales. In contrast, constraints on PNG from LSS surveys are only starting to become available and have the potential to improve significantly in the future [31]. However, to extract this information requires pushing the theoretical understanding of gravitational clustering to smaller scales into the mildly nonlinear regime.

Nonlinear gravitational evolution produces two types of effects: First, it

generates non-Gaussianity even if the initial conditions are perfectly Gaussian. This inevitable "background" needs to be characterized precisely before any non-Gaussian "signal" can be discovered. Second, any primordial non-Gaussianity in the initial conditions gets distorted. These distortions can affect how well we can distinguish between distinct types of non-Gaussianity (cf. fig. 4.10). Both of these effects are captured by the EFT-of-LSS. The first was studied in [45, 52], while, in this work, we have developed the tools to address the second.

A number of open questions have been left for future work [60]:

- For small levels of PNG, two-loop Gaussian corrections become comparable to the one-loop non-Gaussian corrections already on relatively large scales. (This is especially true for the power spectrum and the equilateral limit of the bispectrum.) Extending the theoretical treatment to $k \sim 0.1\,h\,\mathrm{Mpc}^{-1}$ (and beyond), therefore requires a more complete treatment of the two-loop corrections for Gaussian initial conditions.

- The finite parts of the EFT contributions cannot be predicted, but have to be measured in N-body simulations or in observations. For Gaussian initial conditions this one-loop matching has been performed in [45, 47, 52]. The one-loop matching for non-Gaussian initial conditions is still outstanding.

- Naively, the number of useful modes in LSS scales as the cube of the maximum wavenumber, k_{max}, at which theory errors are still under control. However, in practice, nonlinear evolution moves information from low-order correlations to higher-order correlations [79, 80]. It remains to be quantified how much information can actually be extracted from measurements of the dark matter bispectrum.

- Most LSS observations only provide access to the statistics of biased tracers (galaxies, halos, etc.) of the underlying dark matter density field. The biasing between these tracers and the dark matter introduces an additional source of nonlinearity. On the other hand, the scale-dependence of the biasing also provides the opportunity of seeing specific imprints of the non-Gaussian initial conditions [81]. A consistent model of nonlinear biasing requires renormalization [53–56]. The

first steps towards defining a self-consistent nonlinear biasing model in the presence of non-Gaussian initial conditions have appeared in [57].

4.A Odd Spin in the Squeezed Bispectrum

In this appendix, we make a few comments about the allowed angular dependence in the squeezed limit of the bispectrum (see also [68, 82]). We will explain why, at leading order, we should expect only even-spin contributions to the squeezed limit.

The following two types of squeezed limits are commonly used in the literature:

$$\lim_{k\to 0} B_\varphi(|\tfrac{1}{2}\boldsymbol{k}+\boldsymbol{p}|, |\tfrac{1}{2}\boldsymbol{k}-\boldsymbol{p}|, k) \ , \tag{4.A.1}$$

$$\lim_{k\to 0} B_\varphi(p', |\boldsymbol{p}'-\boldsymbol{k}|, k) \ . \tag{4.A.2}$$

In this work, we have defined the Legendre expansion of (4.A.1),

$$\lim_{k\to 0} B_\varphi(|\tfrac{1}{2}\boldsymbol{k}+\boldsymbol{p}|, |\tfrac{1}{2}\boldsymbol{k}-\boldsymbol{p}|, k) = \left[\sum_{L,i} a_{L,i} \left(\frac{k}{p}\right)^{\Delta_i} P_L(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{p}})\right] \times \left[1+\mathcal{O}\left(\frac{k^2}{p^2}\right)\right] P_\varphi(p) P_\varphi(k) \ . \tag{4.A.3}$$

Since the left-hand side is invariant under $\boldsymbol{p}\mapsto -\boldsymbol{p}$ (regardless of whether p is large or not), we must have $a_{L,i}=0$ for L odd. Comparing (4.A.1) and (4.A.2), we find that $\boldsymbol{p}'=\boldsymbol{p}+\frac{1}{2}\boldsymbol{k}$, and hence

$$\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{p}} = \hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{p}}' - \frac{k}{2p'}\left(1-(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{p}}')^2\right) + \mathcal{O}\left(\frac{k^2}{(p')^2}\right) \ . \tag{4.A.4}$$

Substituting (4.A.4) into (4.A.3), we get

$$P_L(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{p}}) = P_L(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{p}}') + \frac{k}{2p'}(L+1)\left[P_{L+1}(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{p}}') - (\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{p}}')\,P_L(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{p}}')\right] \ . \tag{4.A.5}$$

This shows that we do generate Legendre polynomials of odd powers in the squeezed limit (4.A.2). However, we also see that they are always subleading since they are forced to come together with a leading Legendre polynomial of even degree.

4.B Coarse Graining in Perturbation Theory

In this appendix, we explicitly integrate out perturbative short scales in the Euler equation and show that all terms consistent with the symmetries are generated in the resulting effective theory.

Suppose that the fluid equations are defined at a scale $\hat{\Lambda} < k_{\rm NL}$ and we wish to integrate out modes in the momentum shell $p \in [\Lambda, \hat{\Lambda}]$. The result is an additional stress tensor in the Euler equation [11]:

$$\tau^{ij} = \frac{1}{8\pi G a^2}\left[2\partial^i\phi_s\partial^j\phi_s - \delta^{ij}(\partial_k\phi_s)^2\right]_\Lambda + \left[\rho v_s^i v_s^j\right]_\Lambda . \tag{4.B.1}$$

We wish to expand τ^{ij} in terms of the long-wavelength fields. We will illustrate the computation for a single term

$$\sigma^{ij} \equiv \left[\partial^i\phi_s\partial^j\phi_s\right]_\Lambda . \tag{4.B.2}$$

For simplicity, we will work in an Einstein-de Sitter universe and compute σ^{ij} to first order in the long-wavelength fluctuations.

4.B.1 Nonlinear Evolution

Since all modes are perturbative, σ^{ij} can be computed in perturbation theory. Below we show that

$$\sigma^{ij}(\boldsymbol{k},\tau) = \int_\Lambda^{\hat{\Lambda}} \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}\left[p^i p^j \sigma_{\boldsymbol{p}}(\boldsymbol{k}) + \int_0^\Lambda \frac{\mathrm{d}^3\tilde{\boldsymbol{k}}}{(2\pi)^3}\,\mathcal{S}_{\boldsymbol{p}}^{ij}(\boldsymbol{k},\tilde{\boldsymbol{k}})\,\sigma_{\boldsymbol{p}}(\boldsymbol{k}-\tilde{\boldsymbol{k}})\,\delta(\tilde{\boldsymbol{k}},\tau) + \cdots\right] , \tag{4.B.3}$$

where we have defined

$$\sigma_{\boldsymbol{p}}(\boldsymbol{k}) \equiv \left[\phi_s(\boldsymbol{p}+\tfrac{1}{2}\boldsymbol{k},\tau_{in})\phi_s(-\boldsymbol{p}+\tfrac{1}{2}\boldsymbol{k},\tau_{in})\right]_\Lambda , \tag{4.B.4}$$

$$\mathcal{S}_{\boldsymbol{p}}^{ij}(\boldsymbol{k},\tilde{\boldsymbol{k}}) \equiv p^i p^j\left[\frac{20}{7} + \frac{(\boldsymbol{k}-\tilde{\boldsymbol{k}})\cdot\tilde{\boldsymbol{k}}}{\tilde{k}^2} - \frac{20}{7}\frac{(\boldsymbol{p}\cdot\tilde{\boldsymbol{k}})^2}{p^2\tilde{k}^2}\right] + \left[p^i\tilde{k}^j + p^j\tilde{k}^i\right]\frac{\boldsymbol{p}\cdot\tilde{\boldsymbol{k}}}{\tilde{k}^2} . \tag{4.B.5}$$

The first term in (4.B.3) simply arises from the linear evolution. The second term comes from replacing one of the ϕ by its second-order solution $\phi_{(2)}$ and

extracting the dependence on the long-wavelength fluctuations. We give the details of the derivation in the following insert.

Derivation. The first-order contribution to (4.B.2) is

$$\frac{4}{9\mathcal{H}^4}\big[\sigma^{ij}\big]_{(1)} = a^3\int_\Lambda^{\hat{\Lambda}}\frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}\,\Pi^{ij}(\boldsymbol{p},\boldsymbol{k})\times \\ \Big[\delta_1(-\boldsymbol{p}+\tfrac{1}{2}\boldsymbol{k})\delta_2(\boldsymbol{p}+\tfrac{1}{2}\boldsymbol{k})+\delta_2(-\boldsymbol{p}+\tfrac{1}{2}\boldsymbol{k})\delta_1(\boldsymbol{p}+\tfrac{1}{2}\boldsymbol{k})\Big]\,, \tag{4.B.6}$$

where we have used $-k^2\phi(\boldsymbol{k}) = \frac{3}{2}\mathcal{H}^2\delta(\boldsymbol{k})$ and defined

$$\Pi^{ij}(\boldsymbol{p},\boldsymbol{k}) \equiv \frac{(p^i-\frac{1}{2}k^i)(p^j+\frac{1}{2}k^j)}{(\boldsymbol{p}-\frac{1}{2}\boldsymbol{k})^2(\boldsymbol{p}+\frac{1}{2}\boldsymbol{k})^2}\,. \tag{4.B.7}$$

Substituting (4.C.16) for δ_2, we get

$$\frac{4}{9\mathcal{H}^4}\big[\sigma^{ij}\big]_{(1)} = 2\,a^3\int_\Lambda^{\hat{\Lambda}}\frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}\int_{\tilde{\boldsymbol{k}}}\Pi^{ij}(\boldsymbol{p},\boldsymbol{k})F_2(\tilde{\boldsymbol{k}},\boldsymbol{p}+\tfrac{1}{2}\boldsymbol{k}-\tilde{\boldsymbol{k}})\,\delta_1(-\boldsymbol{p}+\tfrac{1}{2}\boldsymbol{k})\times \\ \delta_1(\boldsymbol{p}+\tfrac{1}{2}\boldsymbol{k}-\tilde{\boldsymbol{k}})\delta_1(\tilde{\boldsymbol{k}})\,. \tag{4.B.8}$$

There are two different regions of integration for $\tilde{\boldsymbol{k}}$:

(i) The UV region where all three fluctuations δ_1 are short. This contributes at cubic order in an expansion of the short-scale fluctuations. The dependence of these short scales on long-wavelength fluctuations is captured by a primordial trispectrum. Since we are working at first order in non-Gaussianity, we will not keep track of these contributions in the rest of this appendix.

(ii) The region where one of the modes is a long-wavelength fluctuation. This happens either for $|\tilde{\boldsymbol{k}}| < \Lambda$ or for $|\tilde{\boldsymbol{k}}-\boldsymbol{p}-\frac{1}{2}\boldsymbol{k}| < \Lambda$. We assume that these two regions do not overlap substantially and therefore consider them as disjoint. (The error made in using this assumption is captured by higher-derivative terms). Since both of these regions give exactly the same contribution to σ^{ij}, we focus on $\tilde{k} < \Lambda$ and multiply the result by two.

Extracting the long-wavelength mode, the first-order contribution to σ^{ij} be-

comes

$$\frac{4}{9\mathcal{H}^4}\big[\sigma^{ij}\big]_{(1)} \simeq 4a^3\int_0^\Lambda \frac{\mathrm{d}^3\tilde{\boldsymbol{k}}}{(2\pi)^3}\Bigg[\int_\Lambda^{\hat{\Lambda}} \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}\,\Pi^{ij}(\boldsymbol{p},\boldsymbol{k})F_2(\tilde{\boldsymbol{k}},\boldsymbol{p}+\tfrac{1}{2}\boldsymbol{k}-\tilde{\boldsymbol{k}}) \times \delta_1(-\boldsymbol{p}+\tfrac{1}{2}\boldsymbol{k})\delta_1(\boldsymbol{p}+\tfrac{1}{2}\boldsymbol{k}-\tilde{\boldsymbol{k}})\Bigg]\delta_1(\tilde{\boldsymbol{k}})\ . \tag{4.B.9}$$

Next, we shift the variable of integration, $\boldsymbol{p}\mapsto\boldsymbol{p}+\frac{1}{2}\tilde{\boldsymbol{k}}$, and express the two short fluctuations δ_1 in terms of the initial potential $\phi_s(\tau_{in})$. Expanding the integrand in the limit of large $\boldsymbol{p}$, we get the answer in (4.B.3). □

It is instructive to go back to position space. Eq. (4.B.3) then becomes

$$\sigma^{ij}(\boldsymbol{x},\tau) = c_0^{ij}(\boldsymbol{x},\tau) + c_1^{ij}(\boldsymbol{x},\tau)\delta(\boldsymbol{x},\tau) + c^{ij}{}_{kl}(\boldsymbol{x},\tau)\,\partial^k\partial^l\phi(\boldsymbol{x},\tau) + \cdots\ , \tag{4.B.10}$$

which is to be compared to the structure of the stress tensor in (4.3.1). Up to corrections of order ϕ_s^3, we have

$$c_a^{ij}(\boldsymbol{x},\tau) = c_a\,[\partial^i\phi_s(\boldsymbol{q})\partial^i\phi_s(\boldsymbol{q})]_\Lambda\ , \tag{4.B.11}$$

$$c^{ij}{}_{kl}(\boldsymbol{x},\tau) = -\frac{40}{21\mathcal{H}^2}[\partial^i\partial^j\phi_s(\boldsymbol{q})\partial_k\partial_l\partial^{-2}\phi_s(\boldsymbol{q})]_\Lambda + \frac{1}{3\mathcal{H}^2}\Big([\partial^i\phi_s(\boldsymbol{q})\partial_k\phi_s(\boldsymbol{q})]_\Lambda\,\delta_l^j + \text{perms}\Big)\ , \tag{4.B.12}$$

where $\boldsymbol{q}(\boldsymbol{x},\tau)$ is the Lagrangian position of the Eulerian coordinate $\boldsymbol{x}$, $a\in\{0,1\}$ and c_a are unimportant numerical factors. Let us make a few comments:

- We see that the coefficients are in general non-local functions of ϕ_s — in this example, this is the case only for the coefficient $c^{ij}{}_{kl}$. Since the inverse Laplacian in (4.B.12) acts only on short-scale fluctuations, $c^{ij}{}_{kl}$ is a non-local function of ϕ_s within a region of size Λ^{-1} centered around the Lagrangian coordinate $\boldsymbol{q}$.

- Some terms such as θ and $\partial^i v^j$ seem to be missing in (4.B.10). This is just an artefact of working to lowest order in perturbation theory, where these terms are indistinguishable from δ and $\partial^i\partial^j\phi$.

Eq. (4.B.3) shows how the late-time $\sigma^{ij}(\boldsymbol{x},\tau)$ depends on the long-wavelength fluctuations and the initial short-scale fluctuations. So far, this is very general, in the sense that we haven't specified the initial conditions for the short scales. Next, we will determine how these initial fluctuations are modulated by long-wavelength fluctuations in the presence of PNG.

4.B.2 Non-Gaussian Initial Conditions

To determine the dependence of $\sigma_{\boldsymbol{p}}(\boldsymbol{k})$ on the long-wavelength fluctuations, we average (4.B.3) over the short scales. This boils down to replacing $\sigma_{\boldsymbol{p}}(\boldsymbol{k})$ by $\langle\sigma_{\boldsymbol{p}}(\boldsymbol{k})\rangle_s$. We then substitute (4.3.6) into the expression of $\langle\sigma_{\boldsymbol{p}}(\boldsymbol{k})\rangle_s$. The result receives contributions from different spins

$$\langle\sigma_{\boldsymbol{p}}(\boldsymbol{k})\rangle_s = \bar{\sigma}_{\boldsymbol{p}}(\boldsymbol{k}) + f_{\rm NL}\big[\langle\sigma_{\boldsymbol{p}}(\boldsymbol{k})\rangle_s^{[0]} + \langle\sigma_{\boldsymbol{p}}(\boldsymbol{k})\rangle_s^{[2]} + \langle\sigma_{\boldsymbol{p}}(\boldsymbol{k})\rangle_s^{[4]}\big] \ , \tag{4.B.13}$$

where $\bar{\sigma}_{\boldsymbol{p}}(\boldsymbol{k}) \propto \delta(\boldsymbol{k})$ is the Gaussian contribution (which is independent of the long-wavelength fluctuations) and $\langle\sigma_{\boldsymbol{p}}(\boldsymbol{k})\rangle_s^{[L]}$ represent the spin-L contribution

$$\langle\sigma_{\boldsymbol{p}}(\boldsymbol{k})\rangle_s^{[0]} = (\mu/p)^{\Delta}\, P_\phi(p) \times \psi(\boldsymbol{k}) \ , \tag{4.B.14}$$

$$\langle\sigma_{\boldsymbol{p}}(\boldsymbol{k})\rangle_s^{[2]} = (\mu/p)^{\Delta}\, P_\phi(p)\, (\hat{p}_i\hat{p}_j) \times \psi^{ij}(\boldsymbol{k}) \ , \tag{4.B.15}$$

$$\langle\sigma_{\boldsymbol{p}}(\boldsymbol{k})\rangle_s^{[4]} = (\mu/p)^{\Delta}\, P_\phi(p)\, (\hat{p}_i\hat{p}_j\hat{p}_k\hat{p}_l) \times \psi^{ijkl}(\boldsymbol{k}) \ , \tag{4.B.16}$$

where the fields ψ, ψ^{ij} and ψ^{ijkl} were defined in eqs. (4.3.9), (4.3.17) and (4.3.21), respectively. As explained in the main text, higher-spin contributions are captured by higher-order tensors which can only contribute at higher order in the fluctuations and/or derivatives. For this reason, we did not consider them in this work.

4.B.3 EFT Operators

We split σ^{ij} into Gaussian and non-Gaussian contributions

$$\langle\sigma^{ij}\rangle_s = \langle\sigma^{ij}\rangle_s^{\rm G} + \langle\sigma^{ij}\rangle_s^{\rm NG} \ . \tag{4.B.17}$$

To obtain the Gaussian contribution we replace $\langle\sigma_{\boldsymbol{p}}(\boldsymbol{k})\rangle_s$ by its Gaussian contribution $\bar{\sigma}_{\boldsymbol{p}}(\boldsymbol{k})$. The non-Gaussian part gets contributions from the different spins:

$$\langle\sigma^{ij}\rangle_s^{\rm NG} = f_{\rm NL}\big[\langle\sigma^{ij}\rangle_s^{[0]} + \langle\sigma^{ij}\rangle_s^{[2]} + \langle\sigma^{ij}\rangle_s^{[4]}\big] \ . \tag{4.B.18}$$

To compute the spin-L contribution to $\langle\sigma^{ij}\rangle_s$, we replace $\sigma_{\boldsymbol{p}}(\boldsymbol{k})$ in (4.B.3) by $\langle\sigma_{\boldsymbol{p}}(\boldsymbol{k})\rangle_s^{[L]}$. Going to real space, we find

$$\langle\sigma^{ij}\rangle_s^{[0]} = \frac{1}{3}\beta\left[\left(\Psi + \frac{16}{7}\Psi\delta\right)\delta^{ij} + \frac{6}{7}\Psi\partial^i\partial^j\Phi\right] , \tag{4.B.19}$$

$$\langle\sigma^{ij}\rangle_s^{[2]} = \frac{2}{15}\beta\left[\Psi^{ij} + \frac{120}{49}\Psi^{ij}\delta - \frac{20}{49}\Psi^{kl}\partial_k\partial_l\Phi\,\delta^{ij} + \frac{9}{49}\Psi^{k(i}\partial^{j)}\partial_k\Phi\right] , \tag{4.B.20}$$

$$\langle\sigma^{ij}\rangle_s^{[2]} = -\frac{32}{441}\beta\,\Psi^{ijkl}\partial_l\partial_k\Phi , \tag{4.B.21}$$

where we have defined the coefficient

$$\beta(\Lambda,\hat{\Lambda}) \equiv \int_\Lambda^{\hat{\Lambda}} \frac{\mathrm{d}p}{2\pi^2}\, p^4 \frac{P_\phi(p)}{(p/\mu)^\Delta} . \tag{4.B.22}$$

We see that all terms consistent with the symmetries are generated in the effective theory — see eqs. (4.3.27), (4.3.28) and (4.3.29).

4.C Perturbation Theory and Counterterms

In this appendix, we collect a few results from standard perturbation theory (SPT) and derive explicit expressions for the one-loop counterterms in the EFT-of-LSS.

4.C.1 Equations of Motion

The equations of motion for the density contrast δ and the velocity divergence θ are

$$\partial_\tau \delta + \theta \;=\; \mathcal{S}_\alpha \;, \tag{4.C.1}$$

$$(\partial_\tau + \mathcal{H})\theta + \frac{3}{2}\Omega_m \mathcal{H}^2 \delta \;=\; \mathcal{S}_\beta + \tau_\theta \;, \tag{4.C.2}$$

where $\tau_\theta \equiv -\partial_i\big[\rho^{-1}\partial_j \tau^{ij}\big]$, and $\mathcal{S}_{\alpha,\beta}$ are nonlinear source terms

$$\mathcal{S}_\alpha(\boldsymbol{k},\tau) \equiv -\int_{\boldsymbol{p}} \alpha(\boldsymbol{p},\boldsymbol{k}-\boldsymbol{p})\theta(\boldsymbol{p},\tau)\delta(\boldsymbol{k}-\boldsymbol{p},\tau) \;, \tag{4.C.3}$$

$$\mathcal{S}_\beta(\boldsymbol{k},\tau) \equiv -\int_{\boldsymbol{p}} \beta(\boldsymbol{p},\boldsymbol{k}-\boldsymbol{p})\theta(\boldsymbol{p},\tau)\theta(\boldsymbol{k}-\boldsymbol{p},\tau) \;, \tag{4.C.4}$$

with

$$\alpha(\boldsymbol{k}_1,\boldsymbol{k}_2) \equiv \frac{\boldsymbol{k}_1\cdot(\boldsymbol{k}_1+\boldsymbol{k}_2)}{k_1^2} \;, \qquad \beta(\boldsymbol{k}_1,\boldsymbol{k}_2) \equiv \frac{(\boldsymbol{k}_1+\boldsymbol{k}_2)^2}{2}\frac{\boldsymbol{k}_1\cdot\boldsymbol{k}_2}{k_1^2 k_2^2}. \tag{4.C.5}$$

Using the scale factor $a(\tau)$ as the evolution variable, the equations of motion become

$$\overbrace{\mathcal{H}^2\left[-a^2\partial_a^2 + \left(\frac{3}{2}\Omega_m - 3\right)a\partial_a + \frac{3}{2}\Omega_m\right]}^{\mathcal{D}_\delta}\delta = \mathcal{S}_\beta + \tau_\theta - \mathcal{H}\partial_a(a\mathcal{S}_\alpha), \tag{4.C.6}$$

$$\underbrace{\mathcal{H}^2\left[+a^2\partial_a^2 + \left(4 - \frac{3}{2}\Omega_m\right)a\partial_a + (2-3\Omega_m)\right]}_{\mathcal{D}_\theta}\theta = \partial_a(a\mathcal{S}_\beta + a\tau_\theta) - \frac{3}{2}\Omega_m\mathcal{H}\mathcal{S}_\alpha \;. \tag{4.C.7}$$

The linearized equation of motion for δ (obtained by setting the right-hand side of (4.C.6) to zero) has the following growing mode solution

$$\delta_{(1)}(\boldsymbol{k}, a) = D_1(a)\delta_1(\boldsymbol{k}) \ , \tag{4.C.8}$$

where $\delta_1(\boldsymbol{k})$ describes the initial condition and $D_1(a)$ is the linear growth factor

$$D_1(a) = \frac{5}{2}\mathcal{H}_0^2\Omega_m^0\frac{\mathcal{H}}{a}\int_{a_{in}}^{a}\frac{\mathrm{d}a'}{\mathcal{H}^3(a')} \ . \tag{4.C.9}$$

In Einstein-de Sitter, the result in (4.C.9) reduces to $D_1(a) = a/a_{in}$.

Solving (4.C.6) with a delta-function source, $\delta_D(a-a')$, gives the Green's function for the evolution of δ:

$$G_\delta(a, a') \ = \ \Theta(a-a')\,\frac{2}{5}\frac{1}{\mathcal{H}_0^2\Omega_m^0}\frac{D_1(a')}{a'}\left[\frac{D_-(a)}{D_-(a')} - \frac{D_1(a)}{D_1(a')}\right] \ , \tag{4.C.10}$$

and $D_- \equiv \mathcal{H}/(a\mathcal{H}_0) \approx D_1^{-3/2}$. A similar Green's function $G_\theta(a, a')$ exists for θ, but it won't be needed in this work. To a remarkably good approximation, we have

$$\int_{a_{in}}^{a}\mathrm{d}a'\ G_\delta(a, a')\mathcal{H}^2(a')f^2(a')[D_1(a')]^n \approx -\frac{2}{(n-1)(2n+3)}[D_1(a)]^n \ , \tag{4.C.11}$$

where $f \equiv d\ln D_1/d\ln a$. The result in (4.C.11) will be useful below. We have checked that this approximation is valid to better than 2% accuracy, for the values of n considered in this thesis.

4.C.2 Perturbative Solution

For $\delta_{(1)} < 1$, the solution to the nonlinear equations can be written as a series in powers of the initial dark matter contrast δ_1 (and integrals over the Green's function G_δ):

$$\delta(\boldsymbol{k}, a) = \sum_{n=1}^{\infty}\delta_{(n)}(\boldsymbol{k}, a) \ , \qquad \theta(\boldsymbol{k}, a) = -\mathcal{H}f(a)\sum_{n=1}^{\infty}\theta_{(n)}(\boldsymbol{k}, a) \ . \tag{4.C.12}$$

For a non-vanishing stress tensor in (4.C.2), the n-th order solution can be written as

$$\delta_{(n)}(\boldsymbol{k}, a) = \delta_{(n)}^{\mathrm{SPT}}(\boldsymbol{k}, a) + \delta_{(n)}^{c}(\boldsymbol{k}, a) + \delta_{(n)}^{J}(\boldsymbol{k}, a) \ , \tag{4.C.13}$$

where $\delta^{\rm SPT}$ is the SPT result obtained with $\tau_\theta \equiv \tau_v + \tau_n = 0$, δ^c is the solution sourced by the viscosity part of the stress tensor τ_v, while δ^J is sourced by the noise component of the stress tensor τ_n. We first review the SPT solution (§4.C.2) and then derive the expression for the one-loop counterterms (§4.C.2 and §4.C.2).

SPT Solution

The n-th order SPT solution can be written in terms of a convolution of the Green function (4.C.10) and lower-order SPT solutions. Using (4.C.11), one finds that the n-th order solution is proportional to the n-th power of the linear growth factor D_1:

$$\delta^{\rm SPT}_{(n)}(\boldsymbol{k}, a) \approx D_1^n(a)\delta_n(\boldsymbol{k}) \ . \tag{4.C.14}$$

A similar result holds for the velocity divergence

$$\theta^{\rm SPT}_{(n)}(\boldsymbol{k}, a) \approx D_1^n(a)\theta_n(\boldsymbol{k}) \ . \tag{4.C.15}$$

The initial conditions δ_n and θ_n can be written as a convolution of powers of δ_1:

$$\delta_n^{\rm SPT}(\boldsymbol{k}) \ = \ \int_{\boldsymbol{k}_1} \dots \int_{\boldsymbol{k}_n} (2\pi)^3 \delta_D\big(\boldsymbol{k} - \boldsymbol{k}_{1\dots n}\big)\, F_n(\boldsymbol{k}_1, \dots, \boldsymbol{k}_n)\, \delta_1(\boldsymbol{k}_1) \dots \delta_1(\boldsymbol{k}_n) \ , \tag{4.C.16}$$

$$\theta_n^{\rm SPT}(\boldsymbol{k}) \ = \ \int_{\boldsymbol{k}_1} \dots \int_{\boldsymbol{k}_n} (2\pi)^3 \delta_D\big(\boldsymbol{k} - \boldsymbol{k}_{1\dots n}\big)\, G_n(\boldsymbol{k}_1, \dots, \boldsymbol{k}_n)\, \delta_1(\boldsymbol{k}_1) \dots \delta_1(\boldsymbol{k}_n) \ , \tag{4.C.17}$$

where $\boldsymbol{k}_{1\dots n} \equiv \boldsymbol{k}_1 + \dots + \boldsymbol{k}_n$. Explicit expressions for the kernel functions F_n and G_n can be found in [29]. For instance, $F_1 = G_1 = 1$, while the (symmetrized) second-order kernel functions are

$$F_2(\boldsymbol{k}_1, \boldsymbol{k}_2) = \frac{5}{7} + \frac{1}{2}\left(\frac{\boldsymbol{k}_1 \cdot \boldsymbol{k}_2}{k_1^2} + \frac{\boldsymbol{k}_1 \cdot \boldsymbol{k}_2}{k_2^2}\right) + \frac{2}{7}\frac{(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)^2}{k_1^2 k_2^2} \ , \tag{4.C.18}$$

$$G_2(\boldsymbol{k}_1, \boldsymbol{k}_2) = \frac{3}{7} + \frac{1}{2}\left(\frac{\boldsymbol{k}_1 \cdot \boldsymbol{k}_2}{k_1^2} + \frac{\boldsymbol{k}_1 \cdot \boldsymbol{k}_2}{k_2^2}\right) + \frac{4}{7}\frac{(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)^2}{k_1^2 k_2^2} \ . \tag{4.C.19}$$

Viscosity Counterterms

In order to renormalize the one-loop bispectrum, we need to compute δ^c and δ^J up to second order. This requires knowing the stress tensor up to second order. First, let us focus on the viscosity contribution to the stress tensor

$$\tau_v^{(1)} \equiv -d^2(a)\triangle\delta_{(1)} - g(a)f_{\rm NL}\triangle\Psi_{(1)} \ , \tag{4.C.20}$$

$$\begin{aligned}\tau_v^{(2)} \equiv& -d^2(a)\triangle\delta_{(2)} - e_1(a)\triangle(\delta_{(1)})^2 - e_2(a)\triangle(s_{(1)})^2 \\ &- e_3(a)\partial_i(s^{ij}_{(1)}\partial_j\delta_{(1)}) - f_{\rm NL}\Big[g(a)\big[\triangle\Psi_{(2)} - \partial_i(\delta_{(1)}\partial^i\Psi_{(1)})\big] \\ &+ g_1(a)\triangle(\Psi_{(1)}\,\delta_{(1)}) + g_2(a)\partial_i\partial_j(s^{ij}_{(1)}\Psi_{(1)})\Big] \ ,\end{aligned} \tag{4.C.21}$$

where $\delta_{(1,2)}$ and $s^{ij}_{(1)} \equiv \partial^i\partial^j\Phi_{(1)} - \frac{1}{3}\delta^{ij}\delta_{(1)}$ refer to the SPT solutions and *not* the full solution (4.C.13). We have defined the parameter $d^2 \equiv c_s^2 + f(c_{vis}^2 + \hat{c}_{vis}^2)$ as the sum of the sound speed and the viscosity parameter (for more details, see [45, 52]). Note that the velocity divergence θ does not appear in these equations because at second order it is completely degenerate with δ. The field $\Psi(\boldsymbol{x},\tau) \equiv \psi(\boldsymbol{q}(\boldsymbol{x},\tau))$ admits an expansion in powers of the fluctuations, with $\Psi_{(1)} \equiv \psi$ and $\Psi_{(2)} \equiv \boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\Phi$; cf. eq. (4.4.1).

The n-th order counterterms $\delta^c_{(n)}$ can be written as

$$\begin{aligned}\delta^c_{(n)}(\boldsymbol{k},a) = &\int_{\boldsymbol{k}_1}\dots\int_{\boldsymbol{k}_n}(2\pi)^3\delta_D\big(\boldsymbol{k}-\boldsymbol{k}_{1\dots n}\big)\,F_n^c(\boldsymbol{k}_1,\dots,\boldsymbol{k}_n|a)\times \\ &\times\delta_{(1)}(\boldsymbol{k}_1,a)\dots\delta_{(1)}(\boldsymbol{k}_n,a) \quad + \\ + f_{\rm NL}&\int_{\boldsymbol{k}_1}\dots\int_{\boldsymbol{k}_n}(2\pi)^3\delta_D\big(\boldsymbol{k}-\boldsymbol{k}_{1\dots n}\big)H_n^c(\boldsymbol{k}_1,\dots,\boldsymbol{k}_n|a)\,\psi(\boldsymbol{k}_1)\dots\delta_{(1)}(\boldsymbol{k}_n,a).\end{aligned} \tag{4.C.22}$$

where F_n^c and H_n^c are kernel functions that are to be determined. In what follows, we will compute these kernels up to second order.

First order

The first-order counterterm $\delta^c_{(1)}$ is the solution to $\mathcal{D}_\delta\,\delta^c_{(1)} = \tau_v^{(1)}$, and can therefore be written as

$$\delta^c_{(1)}(\boldsymbol{k},a) = -\xi(a)k^2\delta_{(1)}(\boldsymbol{k},a) - \gamma(a)f_{\rm NL}k^2\Psi_{(1)}(\boldsymbol{k}) \ , \tag{4.C.23}$$

where we have defined

$$\xi(a) \equiv -\frac{1}{D_1(a)} \int_{a_{in}}^{a} \mathrm{d}a' \; G_\delta(a,a')\, d^2(a') D_1(a') \;, \tag{4.C.24}$$

$$\gamma(a) \equiv -\int_{a_{in}}^{a} \mathrm{d}a' \; G_\delta(a,a')\, g(a') \;. \tag{4.C.25}$$

We see that the parameters ξ and γ will depend on the time dependence of $d^2(a)$ and $g(a)$. For the one-loop power spectrum, this is not very important, since in the end we just need to fit the values of ξ and γ at a given redshift. However, as we shall see in the next section, the second-order counterterm $\delta^c_{(2)}$ depends on the time dependence of these parameters. A convenient ansatz for the time dependence of the parameters in (4.C.24) and (4.C.25) is

$$d^2(a) = [\mathcal{H}(a) f(a)]^2 [D_1(a)]^{m_d+1}\, \bar{d}^2 \;, \tag{4.C.26}$$

$$g(a) = [\mathcal{H}(a) f(a)]^2 [D_1(a)]^{m_g+1}\, \bar{g} \;, \tag{4.C.27}$$

where $\bar{d}$ and $\bar{g}$ are constants. Using (4.C.11), we then have

$$\xi(a) = \frac{2}{(m_d+1)(2m_d+7)} [D_1(a)]^{m_d+1}\, \bar{d}^2 \;, \tag{4.C.28}$$

$$\gamma(a) = \frac{2}{m_g(2m_g+5)} [D_1(a)]^{m_g+1}\, \bar{g} \;. \tag{4.C.29}$$

Second order

At second order, things are a bit more complicated, since the second-order solution will depend on the precise time dependence of the first-order solution $\delta^c_{(1)}$. Consider the equation of motion for $\delta^c_{(2)}$:

$$\mathcal{D}_\delta\, \delta^c_{(2)} = \mathcal{S}^{(2)}_\beta + \tau^{(2)}_v - \mathcal{H}\partial_a(a\mathcal{S}^{(2)}_\alpha) \;, \tag{4.C.30}$$

where $\mathcal{S}^{(2)}_\alpha$ and $\mathcal{S}^{(2)}_\beta$ are obtained by replacing one of the δ or θ in the convolution by their linear SPT solution and the other by the corresponding linear counterterm. Using (4.C.14) and (4.C.23) for $\delta^{\mathrm{SPT}}_{(1)}$ and $\delta^c_{(1)}$, respectively, and replacing by $\theta^{\mathrm{SPT}}_{(1)}$ and $\theta^c_{(1)}$ by $\theta_{(1)} = -\dot{\delta}_{(1)}$, the solution for $\delta^c_{(2)}$

can be written as

$$\delta^{c,\mathrm{G}}_{(2)}(\boldsymbol{k},a) = \int_{\boldsymbol{p}} F_2^c(\boldsymbol{p},\boldsymbol{k}-\boldsymbol{p}|a)\,\delta_{(1)}(\boldsymbol{p},a)\delta_{(1)}(\boldsymbol{k}-\boldsymbol{p},a)\ , \tag{4.C.31}$$

$$\delta^{c,\mathrm{NG}}_{(2)}(\boldsymbol{k},a) = f_{\mathrm{NL}}\int_{\boldsymbol{p}} H_2^c(\boldsymbol{p},\boldsymbol{k}-\boldsymbol{p}|a)\,\psi(\boldsymbol{p})\delta_{(1)}(\boldsymbol{k}-\boldsymbol{p},a)\ . \tag{4.C.32}$$

We have separated the solution into a Gaussian part $\delta^{c,\mathrm{G}}_{(2)}$ and a non-Gaussian term $\delta^{c,\mathrm{NG}}_{(2)}$. We look at each of these terms in turn.

- **Gaussian contributions**

  The kernel functions F_2^c were computed in [45, 52, 73]:

  $$F_2^c(\boldsymbol{k}_1,\boldsymbol{k}_2|a) = F_2^\tau(\boldsymbol{k}_1,\boldsymbol{k}_2|a) + F_2^{\alpha\beta}(\boldsymbol{k}_1,\boldsymbol{k}_2|a) + F_2^\delta(\boldsymbol{k}_1,\boldsymbol{k}_2|a)\ , \tag{4.C.33}$$

  where F_2^τ is sourced by the nonlinear terms in $\tau_v^{(2)}$, $F_2^{\alpha\beta}$ is sourced by $\mathcal{S}_{\alpha,\beta}$ and F_2^δ is sourced by the second-order SPT solution $\delta^{\mathrm{SPT}}_{(2)}$ which appears in the stress tensor (4.C.21) as $\tau_v^{(2)} \supset -d^2 \triangle \delta^{\mathrm{SPT}}_{(2)}$.

  - The kernel F_2^τ can be written as

    $$F_2^\tau(\boldsymbol{k}_1,\boldsymbol{k}_2|a) = -\sum_{i=1}^{3}\epsilon_i(a)E_i(\boldsymbol{k}_1,\boldsymbol{k}_2)\ , \tag{4.C.34}$$

    where the time-dependent coefficients are

    $$\epsilon_i(a) \equiv -\frac{1}{[D_1(a)]^2}\int_{a_{in}}^{a} \mathrm{d}a'\ G_\delta(a,a')\,[D_1(a')]^2\,e_i(a')\ . \tag{4.C.35}$$

    and the momentum kernels are

    $$E_1(\boldsymbol{k}_1,\boldsymbol{k}_2) \equiv \boldsymbol{k}_{12}^2\ , \tag{4.C.36}$$

    $$E_2(\boldsymbol{k}_1,\boldsymbol{k}_2) \equiv \boldsymbol{k}_{12}^2\left[\frac{(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)^2}{k_1^2k_2^2} - \frac{1}{3}\right]\ , \tag{4.C.37}$$

    $$E_3(\boldsymbol{k}_1,\boldsymbol{k}_2) \equiv \left[-\frac{1}{6}\boldsymbol{k}_{12}^2 + \frac{1}{2}\boldsymbol{k}_1\cdot\boldsymbol{k}_2\left[\frac{\boldsymbol{k}_{12}\cdot\boldsymbol{k}_2}{k_2^2} + \frac{\boldsymbol{k}_{12}\cdot\boldsymbol{k}_1}{k_1^2}\right]\right]\ . \tag{4.C.38}$$

- The kernel $F_2^{\alpha\beta}$ can be written as

$$F_2^{\alpha\beta}(\boldsymbol{k}_1, \boldsymbol{k}_2|a) = -\xi(a) E_{\alpha\beta}(\boldsymbol{k}_1, \boldsymbol{k}_2) \ , \tag{4.C.39}$$

where $\xi(a)$ was defined in (4.C.24), and

$$E_{\alpha\beta}(\boldsymbol{k}_1, \boldsymbol{k}_2) \equiv \frac{1}{2m_d + 9} \Bigg[2\beta(\boldsymbol{k}_1, \boldsymbol{k}_2)(k_1^2 + k_2^2) + \frac{2m_d + 7}{2(m_d + 2)} \Big(\alpha(\boldsymbol{k}_1, \boldsymbol{k}_2) \big(k_2^2 + (m_d + 2) k_1^2 \big) + \{1 \leftrightarrow 2\} \Big) \Bigg] \ . \tag{4.C.40}$$

- Finally, the kernel F_2^{δ} reads

$$F_2^{\delta}(\boldsymbol{k}_1, \boldsymbol{k}_2|a) = -\xi(a) E_{\delta}(\boldsymbol{k}_1, \boldsymbol{k}_2) \ , \tag{4.C.41}$$

where we have defined

$$E_{\delta}(\boldsymbol{k}_1, \boldsymbol{k}_2) = \frac{(m_d + 1)(2m_d + 7)}{(m_d + 2)(2m_d + 9)} \, \boldsymbol{k}_{12}^2 \, F_2(\boldsymbol{k}_1, \boldsymbol{k}_2) \ , \tag{4.C.42}$$

and F_2 is the SPT kernel (4.C.18).

- **Non-Gaussian contributions**

The kernel function H_2^c can also be written as a sum of terms

$$H_2^c(\boldsymbol{k}_1, \boldsymbol{k}_2|a) = H_2^{\tau}(\boldsymbol{k}_1, \boldsymbol{k}_2|a) + H_2^{\alpha\beta}(\boldsymbol{k}_1, \boldsymbol{k}_2|a) + H_2^{\Psi}(\boldsymbol{k}_1, \boldsymbol{k}_2|a) \ , \tag{4.C.43}$$

where H_2^{τ} and $H_2^{\alpha\beta}$ have the same meanings as before, and H_2^{Ψ} is sourced by the term proportional to $g(a)$ in the stress tensor.

- The kernel H_2^{τ} can be written as

$$H_2^{\tau}(\boldsymbol{k}_1, \boldsymbol{k}_2|a) = -\sum_{i=1}^{2} \gamma_i(a) G_i(\boldsymbol{k}_1, \boldsymbol{k}_2) \ , \tag{4.C.44}$$

where the time-dependent coefficients are

$$\gamma_i(a) \equiv -\frac{1}{D_1(a)} \int_{a_{in}}^{a} \mathrm{d}a' \ G_{\delta}(a, a') \, D_1(a') g_i(a') \ , \tag{4.C.45}$$

and the momentum kernels are

$$G_1(\boldsymbol{k}_1, \boldsymbol{k}_2) = \boldsymbol{k}_{12}^2 \ , \tag{4.C.46}$$

$$G_2(\boldsymbol{k}_1, \boldsymbol{k}_2) = \frac{(\boldsymbol{k}_{12} \cdot \boldsymbol{k}_2)^2}{k_2^2} - \frac{1}{3}\boldsymbol{k}_{12}^2 \ . \tag{4.C.47}$$

- The kernel $H_2^{\alpha\beta}$ can be written as

$$H_2^{\alpha\beta}(\boldsymbol{k}_1, \boldsymbol{k}_2|a) = -\gamma(a) G_{\alpha\beta}(\boldsymbol{k}_1, \boldsymbol{k}_2) \ , \tag{4.C.48}$$

where

$$G_{\alpha\beta}(\boldsymbol{k}_1, \boldsymbol{k}_2) \equiv \ \frac{4}{2m_g + 7}\, \beta(\boldsymbol{k}_1, \boldsymbol{k}_2) k_1^2 + \frac{2m_g + 5}{(m_g + 1)(2m_g + 7)} \big[(m_g + 1)\alpha(\boldsymbol{k}_1, \boldsymbol{k}_2) + \alpha(\boldsymbol{k}_2, \boldsymbol{k}_1)\big] k_1^2 \ . \tag{4.C.49}$$

- Finally, the kernel H_2^{Ψ} reads

$$H_2^{\Psi}(\boldsymbol{k}_1, \boldsymbol{k}_2|a) = -\gamma(a) G_{\Psi}(\boldsymbol{k}_1, \boldsymbol{k}_2) \ , \tag{4.C.50}$$

where

$$G_{\Psi}(\boldsymbol{k}_1, \boldsymbol{k}_2) = \frac{m_g(2m_g + 5)}{(m_g + 1)(2m_g + 7)} \left[\boldsymbol{k}_{12}^2 \frac{\boldsymbol{k}_1 \cdot \boldsymbol{k}_2}{k_2^2} - \boldsymbol{k}_{12} \cdot \boldsymbol{k}_1\right] \ . \tag{4.C.51}$$

Noise Counterterms

Up to second order, the noise contributions to the stress tensor are

$$\tau_n^{(1)} \equiv -\partial_i \partial_j J_0^{ij} \ , \tag{4.C.52}$$

$$\tau_n^{(2)} \equiv \partial_i(\delta \partial_j J_0^{ij}) - \partial_i \partial_j (J_1^{ij} \delta_{(1)}) - \partial_i \partial_j (J_{2\ kl}^{ij} s_{(1)}^{kl}) - f_{\rm NL} \partial_i \partial_j (J_\psi^{ij} \Psi_{(1)}) \ . \tag{4.C.53}$$

The n-th order noise counterterm $\delta_{(n)}^J$ can be written as the sum of Gaussian and non-Gaussian contributions

$$\delta_{(n)}^J(\boldsymbol{k}, a) = \delta_{(n)}^{J,\rm G}(\boldsymbol{k}, a) + \delta_{(n)}^{J,\rm NG}(\boldsymbol{k}, a) \ . \tag{4.C.54}$$

The non-Gaussian part receives contributions from both noise terms in the initial conditions (4.3.4) and the noise terms in the stress tensor (4.3.31). We will derive explicit expressions for the counterterms up to second order.

First order

Solving $\mathcal{D}_\delta\, \delta^{J,\mathrm{G}}_{(1)} = \tau^{(1)}_n$, gives the first-order Gaussian solution

$$\delta^{J,\mathrm{G}}_{(1)}(\boldsymbol{k},a) = k_i k_j \int_{a_{in}}^{a} \mathrm{d}a'\; G_\delta(a,a')\, J_0^{ij}(\boldsymbol{k},a') \;\equiv\; k_i k_j N_0^{ij}(\boldsymbol{k},a)\;. \tag{4.C.55}$$

The non-Gaussian contribution to $\delta^{J}_{(1)}$ only comes from the initial conditions. In fact, the solution is obtained by replacing $\delta_{(1)}(\boldsymbol{k},a)$ with $f_{\mathrm{NL}}\, M(k)\psi_J(\boldsymbol{k})$ in the SPT expansion. At first order, we get

$$\delta^{J,\mathrm{NG}}_{(1)}(\boldsymbol{k},a) = f_{\mathrm{NL}}\, M(k)\, \psi_J(\boldsymbol{k})\;. \tag{4.C.56}$$

Second order

The second-order Gaussian contribution is obtained by solving (4.C.6) with τ_θ replaced by the Gaussian contribution to $\tau^{(2)}_n$ and by replacing one of the δ (or θ) in $\mathcal{S}_{\alpha,\beta}$ by their linear SPT solution $\delta^{\mathrm{SPT}}_{(1)}$ and the other by the first-order Gaussian noise contribution $\delta^{J,\mathrm{G}}_{(1)}$ (or $\theta^{J,\mathrm{G}}_{(1)}$). Since we will not require the second-order Gaussian noise counterterm for this work, we will not explicitly compute it here and refer the reader to [45] for an explicit expression.

The second-order non-Gaussian contribution is obtained in a similar way. The contribution coming from the initial noise term ψ_J is obtained by replacing $\delta_{(1)} \mapsto f_{\mathrm{NL}}\, M(k)\psi_J(\boldsymbol{k})$ in the SPT expansion. Adding the term sourced by the non-Gaussian contribution in (4.C.53), we get

$$\begin{aligned}\delta^{J,\mathrm{NG}}_{(2)}(\boldsymbol{k},a) =& f_{\mathrm{NL}} \int_{\boldsymbol{p}} \Big[k_i k_j N^{ij}_\psi(\boldsymbol{k}-\boldsymbol{p},a)\psi(\boldsymbol{p}) \\ &+ 2F_2(\boldsymbol{k}-\boldsymbol{p},\boldsymbol{p})\delta_{(1)}(\boldsymbol{k}-\boldsymbol{p},a)M(p)\psi_J(\boldsymbol{p})\Big]\,,\end{aligned} \tag{4.C.57}$$

where

$$N^{ij}_\psi(\boldsymbol{k},a) \equiv \int_{a_{in}}^{a} \mathrm{d}a'\; G_\delta(a,a') J^{ij}_\psi(\boldsymbol{k},a')\;. \tag{4.C.58}$$

4.D IR-Safe Integrands

The one-loop integrals can have divergences as the loop momentum $\boldsymbol{p}$ or one of the external momenta $\pm\boldsymbol{k}_i$ approach $\boldsymbol{0}$. Although individual diagrams can be IR divergent, the equivalence principle guarantees that the sum will be finite. To avoid delicate cancellations of large integrals (which might affect the precision of the numerical computation) it is useful to define integrands which are well-behaved in the IR. For Gaussian initial conditions, this was done for the power spectrum in [51] and for the bispectrum in [45, 52]. In this appendix, we extend these results to non-Gaussian diagrams.

4.D.1 Power Spectrum

Let us first consider the non-Gaussian contribution to the one-loop power spectrum (4.4.3)

$$P_{12}(k) = 2\int_{\boldsymbol{p}} F_2(\boldsymbol{k}-\boldsymbol{p},\boldsymbol{p})B_{111}(k,|\boldsymbol{k}-\boldsymbol{p}|,p) \equiv \int_{\boldsymbol{p}} p_{12}(\boldsymbol{p},\boldsymbol{k}) \ . \tag{4.D.1}$$

The kernel function F_2 is divergent when either of its argument vanishes [29], cf. eq. (4.C.18). This means that the integrand is divergent in the limits $\boldsymbol{p}\to\boldsymbol{0}$ and $\boldsymbol{p}\to\boldsymbol{k}$. However, these IR divergences are unphysical and cancel in the integral. To make this cancelation manifest at the level of the integrand, we will first map the divergence at $\boldsymbol{p}=\boldsymbol{k}$ into a divergence at $\boldsymbol{p}=\boldsymbol{0}$. To do so, let us split the region of integration as follows

$$P_{12}(k) = \int_{p<|\boldsymbol{k}-\boldsymbol{p}|} p_{12}(\boldsymbol{p},\boldsymbol{k}) \ + \int_{p>|\boldsymbol{k}-\boldsymbol{p}|} p_{12}(\boldsymbol{p},\boldsymbol{k}) \ . \tag{4.D.2}$$

Changing the integration variable in the second integral, $\boldsymbol{p}\mapsto\boldsymbol{k}-\boldsymbol{p}$, we get

$$P_{12}(k) = 2\int_{\boldsymbol{p}} p_{12}(\boldsymbol{p},\boldsymbol{k})\,\Theta(|\boldsymbol{k}-\boldsymbol{p}|-p) \ , \tag{4.D.3}$$

where Θ is the Heaviside function. Although this integrand no longer has a divergence at $\boldsymbol{p}=\boldsymbol{k}$, there is still a divergence at $\boldsymbol{p}=\boldsymbol{0}$. To remove this divergence, we first notice that

$$\lim_{\boldsymbol{p}\to\boldsymbol{0}} F_2(\boldsymbol{k}-\boldsymbol{p},\boldsymbol{p}) = \frac{\boldsymbol{k}\cdot\boldsymbol{p}}{2p^2} + \mathcal{O}(p^0) \ . \tag{4.D.4}$$

This means that the problematic divergence disappears if we make the integrand invariant under the exchange $\boldsymbol{p} \mapsto -\boldsymbol{p}$. We therefore write

$$\tilde{p}_{12}(\boldsymbol{p},\boldsymbol{k}) \equiv p_{12}(\boldsymbol{p},\boldsymbol{k})\,\Theta(|\boldsymbol{k}-\boldsymbol{p}|-p) + p_{12}(-\boldsymbol{p},\boldsymbol{k})\,\Theta(|\boldsymbol{k}+\boldsymbol{p}|-p)\ . \quad (4.D.5)$$

This integrand is IR-safe: it is well behaved in the limits $\boldsymbol{p} \to \boldsymbol{0}$ and $\boldsymbol{p} \to \pm\boldsymbol{k}$.

4.D.2 Bispectrum

In the main text, we have written the non-Gaussian one-loop bispectrum as

$$B^{\rm NG}_{\rm loop} = \int_{\boldsymbol{p}} \left[b^{\rm (I)}_{311}(\boldsymbol{p},\boldsymbol{k}_i) + b^{\rm (II)}_{311}(\boldsymbol{p},\boldsymbol{k}_i) + b^{\rm (I)}_{122}(\boldsymbol{p},\boldsymbol{k}_i) + b^{\rm (II)}_{122}(\boldsymbol{p},\boldsymbol{k}_i) + \text{perms}\right]\ , \quad (4.D.6)$$

where

$$b^{\rm (I)}_{311}(\boldsymbol{p},\boldsymbol{k}_i) \equiv 3\,F_3(\boldsymbol{k}_1+\boldsymbol{p},-\boldsymbol{p},\boldsymbol{k}_2)B_{111}(k_1,p,|\boldsymbol{k}_1+\boldsymbol{p}|)P_{11}(k_2)\ , \quad (4.D.7)$$

$$b^{\rm (II)}_{311}(\boldsymbol{p},\boldsymbol{k}_i) \equiv 3\,F_3(\boldsymbol{k}_1,\boldsymbol{p},-\boldsymbol{p})B_{111}(k_1,k_2,k_3)P_{11}(p)\ , \quad (4.D.8)$$

$$b^{\rm (I)}_{122}(\boldsymbol{p},\boldsymbol{k}_i) \equiv 4\,F_2(\boldsymbol{k}_3+\boldsymbol{p},-\boldsymbol{p})F_2(\boldsymbol{p},\boldsymbol{k}_2-\boldsymbol{p})B_{111}(k_1,|\boldsymbol{k}_3+\boldsymbol{p}|,|\boldsymbol{k}_2-\boldsymbol{p}|)\times \times P_{11}(p)\ , \quad (4.D.9)$$

$$b^{\rm (II)}_{122}(\boldsymbol{p},\boldsymbol{k}_i) \equiv 2\,F_2(\boldsymbol{k}_1,\boldsymbol{k}_2)F_2(\boldsymbol{k}_1-\boldsymbol{p},\boldsymbol{p})B_{111}(k_1,|\boldsymbol{k}_1-\boldsymbol{p}|,p)P_{11}(k_2)\ . \quad (4.D.10)$$

These integrands contain divergences when the loop momentum $\boldsymbol{p}$ approaches zero or $\pm\boldsymbol{k}_i$. Following the same logic as for the power spectrum, we can write the bispectrum (4.D.6) as an integral of two IR-safe integrands

$$B^{\rm NG}_{\rm loop} = \int_{\boldsymbol{p}} \left[b^{(A)}(\boldsymbol{p},\boldsymbol{k}_i) + b^{(B)}(\boldsymbol{p},\boldsymbol{k}_i)\right]\ . \quad (4.D.11)$$

The integrands $b^{(A)}(\boldsymbol{p},\boldsymbol{k}_i)$ and $b^{(B)}(\boldsymbol{p},\boldsymbol{k}_i)$ can be written as

$$b^{(A)}(\boldsymbol{p},\boldsymbol{k}_i) = b^{(A)}_1(\boldsymbol{p},\boldsymbol{k}_i) + b^{(A)}_2(\boldsymbol{p},\boldsymbol{k}_i) + b^{(A)}_3(\boldsymbol{p},\boldsymbol{k}_i)\ , \quad (4.D.12)$$

$$b^{(B)}(\boldsymbol{p},\boldsymbol{k}_i) = b^{(B)}_1(\boldsymbol{p},\boldsymbol{k}_i) + b^{(B)}_2(\boldsymbol{p},\boldsymbol{k}_i)\ , \quad (4.D.13)$$

where we have defined

$$b_1^{(A)}(\boldsymbol{p},\boldsymbol{k}_i) \equiv b_{122}^{(\mathrm{I})}(\boldsymbol{k}_2-\boldsymbol{p},\boldsymbol{k}_i)\Big[\Theta(|\boldsymbol{k}_2-\boldsymbol{p}|-p)\Theta(|\boldsymbol{k}_3+\boldsymbol{p}|-|\boldsymbol{k}_2-\boldsymbol{p}|) + \Theta(|\boldsymbol{k}_3+\boldsymbol{p}|-p)\Theta(|\boldsymbol{k}_2-\boldsymbol{p}|-|\boldsymbol{k}_3+\boldsymbol{p}|)\Big] + 5 \text{ perms} , \quad (4.D.14)$$

$$b_2^{(A)}(\boldsymbol{p},\boldsymbol{k}_i) \equiv 2\, b_{122}^{(\mathrm{II})}(\boldsymbol{p},\boldsymbol{k}_i)\, \Theta(|\boldsymbol{k}_1-\boldsymbol{p}|-p) + 5 \text{ perms} , \quad (4.D.15)$$

$$b_3^{(A)}(\boldsymbol{p},\boldsymbol{k}_i) \equiv 2\, b_{311}^{(\mathrm{I})}(\boldsymbol{p},\boldsymbol{k}_i)\, \Theta(|\boldsymbol{k}_1+\boldsymbol{p}|-p) + 5 \text{ perms} , \quad (4.D.16)$$

$$b_1^{(B)}(\boldsymbol{p},\boldsymbol{k}_i) \equiv \frac{1}{2}\Big[b_{122}^{(\mathrm{I})}(\boldsymbol{p},\boldsymbol{k}_i)\Theta(|\boldsymbol{k}_3+\boldsymbol{p}|-p)\Theta(|\boldsymbol{k}_2-\boldsymbol{p}|-p) + \{\boldsymbol{p}\to-\boldsymbol{p}\}\Big] + 2 \text{ perms} , \quad (4.D.17)$$

$$b_2^{(B)}(\boldsymbol{p},\boldsymbol{k}_i) \equiv b_{311}^{(\mathrm{II})}(\boldsymbol{p},\boldsymbol{k}_i) + 2 \text{ perms} . \quad (4.D.18)$$

It is easy to check that the integrands $b_i^{(A,B)}$ do not have any divergences as $\boldsymbol{p}$ approaches one of the external momenta $\pm\boldsymbol{k}_i$. Moreover, while each of the individual integrands $b_i^{(A,B)}$ is divergent in the limit $\boldsymbol{p}\to\boldsymbol{0}$, the sums $b^{(A)}$ and $b^{(B)}$ are finite.

4.E Notation and Conventions

Symbol	Relation	Meaning	Equation
a		scale factor	
τ	$a\,\mathrm{d}\tau = \mathrm{d}t$	conformal time	
$\mathcal{H}$	$\equiv d\ln(a)/d\tau$	conformal Hubble parameter	
$\mathcal{H}_0$		present value of $\mathcal{H}$	
$\boldsymbol{x}$		comoving coordinate	
$\boldsymbol{q}$		Lagrangian coordinate	
$\boldsymbol{k}$		long momentum	
$\boldsymbol{p}$		short momentum	
Ω_m		matter density in units of the critical density	
Ω_Λ		dark energy density	
h		dimensionless Hubble constant	
ρ		dark matter density	
δ	$\equiv \delta\rho/\rho$	dark matter density contrast	
θ	$\equiv \partial_i v^i$	velocity divergence	
$\delta_{(n)}$		density contrast in SPT at order n	(4.C.12)
$\theta_{(n)}$		velocity divergence in SPT at order n	(4.C.12)
F_n		kernel function in $\delta_{(n)}$	(4.C.16)
G_n		kernel function in $\theta_{(n)}$	(4.C.17)
P_{mn}	$\equiv \langle\delta_{(m)}\delta_{(n)}\rangle'$	power spectrum in SPT	(3.3.13)
B_{lmn}	$\equiv \langle\delta_{(l)}\delta_{(m)}\delta_{(n)}\rangle'$	bispectrum in SPT	(3.3.15)
P_δ	$\equiv \langle\delta\delta\rangle'$	nonlinear dark matter power spectrum	(3.3.13)
B_δ	$\equiv \langle\delta\delta\delta\rangle'$	nonlinear dark matter bispectrum	(3.3.15)

Symbol	Relation	Meaning	Equation
Δ^2_δ		dimensionless power spectrum	(4.5.23)
$\mathcal{B}_\delta$		dimensionless bispectrum	(4.5.24)
ϕ		Newtonian potential	(3.3.3)
Φ	$\triangle\Phi = \delta$	rescaled Newtonian potential	(3.3.3)
φ	$\phi = T(k)\varphi$	primordial potential	(4.2.1)
φ_g		Gaussian primordial potential	(4.2.7)
$\Delta_\varphi(k_0)$	$\equiv 3.0 \times 10^{-5}$	amplitude of the primordial potential	(4.2.3)
k_0		pivot scale	(4.2.3)
$T(k)$		transfer function	(4.2.1)
$M(k)$		transfer function in the Poisson equation	(4.2.2)
D_1		linear growth factor	(4.C.9)
f	$\equiv d\ln D_1/\ln a$	growth rate	(4.3.27)
P_φ		primordial power spectrum	(4.2.3)
B_φ		primordial bispectrum	(4.2.5)
n_s		scalar spectral index	(4.2.3)
s_{ij}	$\equiv \partial_i\partial_j\Phi - \frac{1}{3}\delta_{ij}\triangle\Phi$	tidal tensor	(4.3.27)
ψ		correlation in the initial conditions	(4.3.9)
Ψ	$\Psi(\boldsymbol{x}) \equiv \psi(\boldsymbol{q}(\boldsymbol{x}))$	Eulerian definition of ψ	(4.3.13)
ψ_J		noise term in the initial conditions	(4.3.4)
ψ^{ij}		spin-2 correlation in the initial conditions	(4.3.17)
ψ^{ijkl}		spin-4 correlation in the initial conditions	(4.3.21)
$K_{\rm NL}$		kernel function of the primordial bispectrum	(4.2.7)
a_L		coefficient in the Legendre expansion of $K_{\rm NL}$	(4.2.10)

Symbol	Relation	Meaning	Equation
Δ		scaling dimension in $K_{\rm NL}$	(4.2.10)
$f_{\rm NL}$		amplitude of the primordial bispectrum	(4.2.10)
P_L		Legendre polynomial of order L	(4.2.10)
W_Λ		window function	(3.3.17)
F_Λ		Fourier transform of W_Λ	(3.3.17)
τ^{ij}		stress tensor of the EFT	(3.3.25)
$\langle\tau^{ij}\rangle_s$		viscosity part of τ^{ij}	(4.3.3)
$\Delta\tau^{ij}$		noise part of τ^{ij}	(4.3.3)
J^{ij}_ψ		parameter in $\Delta\tau^{ij}$	(4.3.31)
τ_θ		EFT source in the Euler equation	(4.4.20)
τ_v		viscosity part of τ_θ	(4.4.22)
τ_n		noise part of τ_θ	(4.4.23)
c_s		sound speed	(3.3.30)
d^2	$\equiv c_s^2 + f(c_{vis}^2 + \hat{c}_{vis}^2)$	parameter in τ_v	(4.4.22)
e_i, g, g_i		parameters in τ_v	(4.4.22)
ξ		parameter in $\delta^c_{(1)}$	(4.4.31)
γ		parameter in $\delta^c_{(1)}$	(4.4.32)
ϵ_i		parameter in $\delta^c_{(2)}$	(4.C.35)
γ_i		parameter in $\delta^c_{(2)}$	(4.4.40)
$\mathcal{S}_{\alpha,\beta}$		SPT quadratic source terms	(4.C.3)
$\delta^c_{(n)}$		viscosity counterterm at order n	(4.4.26)
$\delta^J_{(n)}$		noise counterterm at order n	(4.4.44)
F^c_n		kernel function in $\delta^c_{(n)}$	(4.C.33)
H^c_n		kernel function in $\delta^c_{(n)}$	(4.C.43)
G_δ		Green's function for δ	(4.C.10)
$\mathcal{D}_\delta$		evolution operator in the fluid equation	(4.C.6)
$P_{1\psi}$	$\equiv \langle\delta_{(1)}\psi\rangle'$	correlation of $\delta_{(1)}$ and ψ	(4.4.2)
P_{1c}	$\equiv \langle\delta_{(1)}\delta^c_{(1)}\rangle'$	correlation of $\delta_{(1)}$ and $\delta^c_{(1)}$	(4.4.33)

Symbol	Relation	Meaning	Equation
$\sigma^2(\Lambda)$		divergence in P_{12}	(4.4.5)
$\sigma^2_\psi(\Lambda)$		divergence in P_{12}	(4.4.6)
$\hat{\sigma}^2(\Lambda)$		divergence in B_{122}	(4.4.16)
$\hat{\sigma}^2_\psi(\Lambda)$		divergence in B_{122}	(4.4.17)
σ_v^2		velocity dispersion	(4.4.14)
$P^{\rm G}_{\rm SPT}$	$\equiv P_{11} + P_{13} + P_{22}$	Gaussian SPT contributions to P_δ	(4.5.3)
$P^{\rm NG}_{\rm SPT}$	$\equiv P_{12}$	non-Gaussian SPT contribution to P_δ	(4.5.4)
$P^{\rm G}_{\rm EFT}$	$\equiv -2\xi k^2 P_{11}$	Gaussian EFT counterterm	(4.5.3)
$P^{\rm NG}_{\rm EFT}$	$\equiv -2\gamma k^2 P_{1\psi}$	non-Gaussian EFT counterterm	(4.5.4)
$B^{\rm G}_{\rm SPT}$		Gaussian SPT contributions to B_δ	(4.5.5)
$B^{\rm NG}_{\rm SPT}$		non-Gaussian SPT contributions to B_δ	(4.5.6)
$B^{\rm G}_{\rm EFT}$		sum of Gaussian EFT counterterms	(4.5.7)
$B^{\rm NG}_{\rm EFT}$		sum of non-Gaussian EFT counterterms	(4.5.8)
$B^{\rm G}_0$	$\equiv B^{\rm G}_{\rm SPT} + \xi B^{\rm G}_\xi$	Gaussian SPT contributions plus $B^{\rm G}_\xi$	(4.5.14)
$B^{\rm NG}_0$	$\equiv B^{\rm NG}_{\rm SPT} + \xi B^{\rm NG}_\xi$	non-Gaussian SPT contributions plus $B^{\rm NG}_\xi$	(4.5.14)
$B^{\rm G}_c$	$\equiv B^{\rm G}_{\rm EFT} - \xi B^{\rm G}_\xi$	sum of Gaussian counterterms $(-B^{\rm G}_\xi)$	(4.5.16)
$B^{\rm NG}_c$	$\equiv B^{\rm NG}_{\rm EFT} - \xi B^{\rm NG}_\xi$	sum of non-Gaussian counterterms $(-B^{\rm NG}_\xi)$	(4.5.17)

5 Lifting Primordial Non-Gaussianity Above the Noise

5.1 Introduction and summary

Primordial deviations from Gaussianity are key to understand inflation and will soon be tested via a number of ambitious Large Scale Structure (LSS) surveys. It is therefore imperative to understand how late-time LSS observations can be related to the parameters that characterize *primordial non-Gaussianity* (PNG). This relation is complicated and non-linear. The degree to which we will be able to collect further primordial information from LSS survey will eventually be determined by our ability to model this non-linear relation. In this work, we focus on specific source of non-linearities, namely perturbative *matter non-linearities*. These are generated by the sub-horizon gravitational evolution of small initial matter inhomogeneities into larger ones, until the perturbations compete locally with the homogeneous background expansion. For concreteness, we study local, equilateral and quasi-single field non-Gaussianity, since these are well-motivated theoretically and represent signals that are complementary from the point of view of observations. Additional sources of non-linearity are also important, such as for example bias and redshift space distortion. In case of equilateral and quasi single field PNG, these are expected to further worsen our ability to constraint primordial parameters. In this sense, our results can be interpreted as *lower bounds* on the precision of future constraints. For local PNG, it is possible that non-linearities encapsulated in the biasing of tracers, if very well understood, might eventually help us improve on the bounds we find here (see [81] and e.g. [83, 84] for a recent estimate). We will discuss this possibility in subsection 5.4.1.

In our analysis, we will use the Effective Field Theory of Large Scale Structures (EFT of LSS) [11], which builds on Standard Perturbation Theory (SPT) [29], and provides a consistent perturbative approach to describe the evolution of matter distribution. We focus exclusively on the matter bispectrum, since it is a primary probe of PNG that is affected by matter non-linearities. Recent work on PNG and the bispectrum includes [83, 85–90]. Within the EFT approach, the bispectrum generated by the late-time gravitational evolution from otherwise Gaussian initial conditions has been studied in [45, 52]). This contribution plays the role of background noise in PNG searches. The signal, namely the primordial bispectrum, is also affected by gravitational non-linearities. This has been recently studied in [1]. Here, we use these two results and present a Fisher forecast for constraints on

PNG. A key element of our forecast is the inclusion of theoretical error, employing and further developing the recent proposal of [91].

For the convenience of the reader, we collect here our major findings with references to where they are discussed in this Chapter.

- When using the EFT of LSS, the perturbative approach to model matter non-linearities will not prevent upcoming LSS surveys to *improve upon the current bounds* from CMB anisotropies [92] (see Table 5.1).

- Our limited perturbative understanding of matter non-linearities limits the achievable bounds on equilateral non-Gaussianity from planned galaxy surveys to $\sigma(f_{NL}^{eq}) \gtrsim 10$ (see Table 5.1), *far from the theoretically interesting benchmark* $\sigma(f_{NL}^{eq}) \sim 1$ (see e.g. [31, 35] and references therein). We estimate that this remains true even if one included the full two-loop corrections (see Table 5.2). Local non-Gaussianity is more promising, and we find e.g. for Euclid $\sigma(f_{NL}^{loc}) \gtrsim 1$.

- The consistent treatment of short-scale effects within the EFT approach allows one to *improve PNG constraints by a factor of about 3* (see Table 5.3). This relies on two facts. First, the EFT parameters provide a better description of the late-time gravitational non-linearities (the "background" discussed in [45]). Second, for the specification of most upcoming experiments, the EFT parameters are only weakly correlated with PNG, and so marginalizing over them hardly degrades the constraints (see subsection 5.4.2).

- Both the SPT loops and the EFT corrections to primordial non-Gaussianity (the "signal" discussed in [1]) are small and their inclusion does *not* improve the PNG constraints appreciably (see first and second line of Table 5.3).

- We discuss several aspects of the method proposed in [91] to *model the theoretical error* inherent to the perturbative approach. We show that a wrong shape for the theoretical error can lead to a biased estimate for f_{NL} . This happens when it partly underestimates the error. Conversely, a conservatively large theoretical error leads to correct unbiased results. We thoroughly analyze the dependence of the Fisher forecast on the correlation length used in [91], and explain our results with a toy model.

This Chapter is organized as follows. In section 5.2, we review the results for the matter bispectrum in the EFT of LSS accounting for PNG. In section 5.3, we discuss the details of the Fisher forecast with particular emphasis on a consistent treatment of theoretical uncertainties. Section 5.4 is devoted to a discussion of the results of the Fisher forecast on PNG constraints from LSS surveys. We conclude in section 5.5. Several appendices contain more technical results. In Chapter 4.C and Appendix 5.A, one can find all relevant formulae to compute the bispectrum in the EFT of LSS. In Appendix 5.B, we present a detailed discussion of how to consistently account for theoretical errors. Appendix 5.C discusses the issue of binning the data for the Fisher forecast and finally, for the convenience of the reader, all symbols used in this chapter and their meaning were collected in Appendix 4.E.

Conventions Redshift z and conformal spatial coordinates $\boldsymbol{x}$ are used as measures of time and position. We use the following convention for the Fourier transform

$$F(\boldsymbol{x}) = \int_k \tilde{F}(\boldsymbol{k}) e^{i\boldsymbol{k}\cdot\boldsymbol{x}}, \quad \text{where we use the shorthand} \quad \int_k \equiv \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}. \tag{5.1.1}$$

In particular, this implies that we have the following relation between any N-point equal-time correlation function and its spectrum

$$\langle \delta(\boldsymbol{k}_1) \dots \delta(\boldsymbol{k}_N) \rangle = (2\pi)^3 \delta_D(\boldsymbol{k}_1 + \dots + \boldsymbol{k}_N) S(\boldsymbol{k}_1, \dots, \boldsymbol{k}_N), \tag{5.1.2}$$

where we suppressed the time dependence.

For the numerical analysis, we compute the linear power spectrum with CAMB [93], where we assume a standard cosmological model with $\Omega_\Lambda^0 = 0.728$, $\Omega_m^0 = 0.272$, $h = 0.704$, $n_s = 0.967$ and $A_\zeta = 2.46 \times 10^{-9}$.

5.2 Analytical predictions for the bispectrum

In this section, we set the initial conditions for our numerical analysis. In subsection 5.2.2, we introduce the concept of theoretical errors, which are intrinsic to the perturbative approach.

For clarity, we briefly recap our notation for the bispectrum and the relevant parameters,

$$B^{\text{th}} = B^{\text{G}}_{\text{SPT}} + B^{\text{G}}_{\text{EFT}} + f_{\text{NL}} \left(B^{\text{NG}}_{\text{SPT}} + B^{\text{NG}}_{\text{EFT}} \right). \tag{5.2.1}$$

As we will see in section 5.3, for a Fisher forecast we do not need to specify[1] $B^{\rm G}_{\rm SPT}$. The leading order counterterms for Gaussian initial conditions have been computed in [45, 52] and read

$$B^{\rm G}_{\rm EFT} = \xi B^{\rm G}_{\xi} + \sum_{i=1}^{3} \epsilon_i B^{\rm G}_{\epsilon_i}. \tag{5.2.2}$$

For non-Gaussian initial conditions, short modes and long modes are already correlated at the initial time. This leads to additional contributions to the matter bispectrum. To leading order, these are given by [1]

$$B^{\rm NG}_{\rm EFT} = \xi B^{\rm NG}_{\xi} + \gamma B^{\rm NG}_{\gamma} + \sum_{i=1}^{2} \gamma_i B^{\rm NG}_{\gamma_i}. \tag{5.2.3}$$

For convenience, we have adopted the notation of 4, in which all the explicit expressions for the terms appearing in this subsection. Appendix 5.A contains the expressions for our two-loop estimation.

5.2.1 Primordial non-Gaussianity

To evaluate the non-Gaussian contributions to (5.2.1), we need to specify the primordial bispectrum. In this Chapter, we study the constraints on three types of primordial non-Gaussianity: local [94], equilateral [95] and quasi-single field [66]. In terms of the primordial potential ϕ, the primordial bispectra are given by the following shapes

$$B^{\rm loc}_{\phi}(k_1,k_2,k_3) = 2 f^{\rm loc}_{\rm NL} \left(P_\phi(k_1)P_\phi(k_2) + \text{perm}\right), \tag{5.2.4a}$$

$$B^{\rm eq}_{\phi}(k_1,k_2,k_3) = 162 f^{\rm eq}_{\rm NL} A^2_\phi \frac{1}{k_1k_2k_3K^3}, \tag{5.2.4b}$$

$$B^{\rm qsf}_{\phi}(k_1,k_2,k_3) = 18\sqrt{3} f^{\rm qsf}_{\rm NL} A^2_\phi \frac{1}{k_1k_2k_3K^3} \frac{N_\nu(8\kappa)}{\sqrt{\kappa} N_\nu(8/27)}. \tag{5.2.4c}$$

Here we define $K = k_1 + k_2 + k_3$, and $\kappa = k_1k_2k_3/K^3$. Moreover, N_ν is the Neumann function of order ν and we choose $\nu = \frac{1}{2}$. The normalization of

[1] On the other hand, we do need to specify the SPT contributions to the power spectrum to compute the cosmic variance. Assuming it is dominated by the linearly evolved matter power spectrum, we do not have to specify additional 'EFT' parameters.

the primordial power spectrum[2] is given by $A_\phi = 1.72 \cdot 10^{-8}$. To linearly evolve these to the late time matter bispectrum[3] B_{111}, we use the transfer function $M(k,z)$, defined by

$$\delta_1(k,z) = M(k,z)\phi(k), \quad \text{with} \quad M^2(k,z) \equiv \frac{k^3 P_{11}(k,z)}{A_\phi \left(\frac{k}{k_\star}\right)^{n_s-1}}. \tag{5.2.5}$$

Here $k_\star = 0.0028\ h\text{Mpc}^{-1}$ and $n_s = 0.967$. This means we have

$$B_{111}(k_1,k_2,k_3,z) = M(k_1,z)M(k_2,z)M(k_3,z)B_\phi(k_1,k_2,k_3). \tag{5.2.6}$$

All relevant higher order non-Gaussian contributions to the bispectrum can be found in 4 and 5.A.

5.2.2 Theoretical error

By definition, any results from perturbation theory are approximate - there is always an intrinsic theoretical error, typically estimated within perturbation theory itself. The true bispectrum is therefore given by

$$B^{true} = B^{th} + B^{er}, \tag{5.2.7}$$

where B^{th} is the perturbative theoretical prediction given in (5.2.1), and B^{er} represents the theoretical error. The strength of a well defined perturbation theory is of course that the error can be estimated within the perturbation theory itself.
In our case, there are in principle two perturbative schemes employed. First, we assume perturbative primordial non-Gaussianity. This means we assume the primordial potential can be schematically expanded as

$$\varphi_p = \varphi_p^{\text{G}} + f_{\text{NL}}\varphi_p^{\text{G}} \star \varphi_p^{\text{G}} + \dots . \tag{5.2.8}$$

Here φ_p^G is a Gaussian field and $\star$ denotes a convolution in Fourier space. This means we are effectively expanding in $f_{\text{NL}}\varphi_p \sim f_{\text{NL}}\sqrt{A_\phi}$, which is indeed very small according to current bounds. Hence we will not worry about corrections to this approximation for the rest of the Chapter.

[2] Note that we define $A_\phi = 2\pi^2 \frac{9}{25} A_\zeta$.
[3] See appendix 4 for relevant notation.

Second, the EFT of LSS relies on the smallness of density perturbations on large scales, consistently taking into account our ignorance of short scale physics. Effectively, this comes down to an expansion in $k/k_{\rm NL}$ [11]. As argued in [45], the most relevant correction to B^{th} is the two-loop bispectrum. Since we have not computed the full two-loop bispectrum, we are forced to make an educated guess about its size and shape. One way to do this was proposed in [91], and relies on the scaling universe results of [43]. Here we use instead a different estimate. Unless indicated otherwise, we estimate the two-loop bispectrum by adding up the absolute values of the two two-loop diagrams we can compute, namely the so-called reducible diagrams, which we indicate by B_{332}. An explicit expression for B_{332} is again given in appendix 5.A. We compare our estimate to the scaling estimate of [91] in appendix 5.B.5.

The importance of keeping track of the theoretical error for forecasts has recently been stressed in [91], and we build on their approach. Qualitatively, one expects not to be able to learn much about $f_{\rm NL}$ from bispectrum configurations for which B^{er} is larger than the non-Gaussian signal. To get an idea of the configurations for which this is the case, we plot the one-loop expressions for the three types of non-Gaussianities we consider (with $f_{\rm NL} = 1$) and B_{332} as a function of scale for both squeezed and equilateral configurations in Figure 5.1. For reference we also plotted the one-loop Gaussian contribution to the bispectrum. As expected, for local PNG we can push to smaller scales in the squeezed configuration than for equilateral PNG. Note also that the naive k_{max}, beyond which we do not expect to gain any more signal, is *configuration dependent*. A detailed discussion on how to incorporate this theoretical error in a Fisher analysis is given in section 5.2.2, which proceeds along the lines of [91]. In appendix 5.B, we present further investigation of the validity of this method of treating the theoretical error.

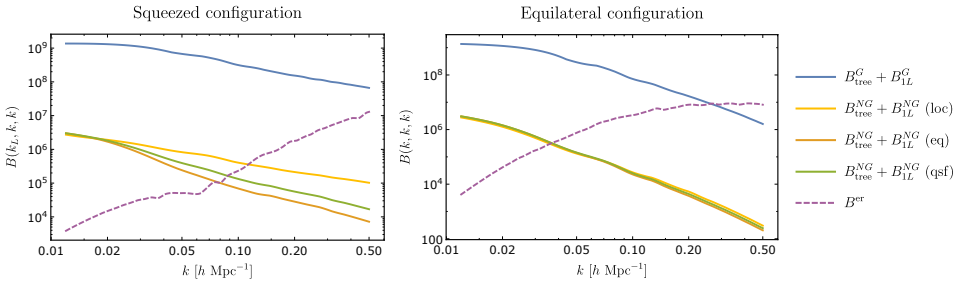


Figure 5.1: The SPT contributions to the bispectrum in the squeezed (left) and equilateral (right) configurations. The blue solid line denotes the Gaussian tree-level and one-loop contributions. The yellow, orange and green lines denote the one loop non-Gaussian contribution for $f_{\mathrm{NL}} = 1$ for local, equilateral and quasi-single-field PNG, respectively. The dashed purple line corresponds to our order-of-magnitude estimate for the Gaussian two-loop correction B_{332}. In the squeezed configurations (left), we chose $k_L = 0.012\ h\mathrm{Mpc}^{-1}$.

5.3 Fisher analysis

In this section, we outline our method to forecast constraints on primordial non-Gaussianity. We have in mind a Gedankenexperiment that provides us with the matter distribution in space and time up to some maximal redshift. In this highly idealized scenario, we determine to what extent our inability to analytically describe the non-linear gravitational collapse of matter limits the information we can extract on primordial perturbations. We proceed along the lines of [85, 86, 96]. The outcome of the analysis for various surveys is presented in the next section.

5.3.1 Assumptions and approximations

For the convenience of the reader, we summarize the assumptions and approximations we make in the Fisher analysis.

- We assume we are given an idealized *survey of the late time dark matter density field*, instead of that of some biased tracer. This allows us to answer the question of whether further progress is needed in the modeling of the dark matter distribution to strengthen current bounds on PNG using upcoming LSS surveys.

- The idealized dark matter survey is characterized by a redshift range and the fraction of the sky covered. We divide the survey in redshift bins, to which we assign a fixed time that is equal to the mean redshift of the bin. Hence, we only need to know $z_{\rm bin}$ to predict the power spectrum and bispectrum. Observational redshift errors are neglected.

- We assume that each redshift bin can be approximated by a cube. Then we just need the volume of the bin $V(z_{\rm bin})$ to account for cosmic variance.

- We compute correlation functions only within each bin. This does not seem to be a big drawback in the case of equilateral PNG. Instead, for local PNG, this might cause an unnecessary loss of information. We will discuss this issue elsewhere.

- We include shotnoise in the analysis to correctly remove weight from the higher redshift bins. For this, we use the specifications of specific upcoming surveys. We discuss this in section 5.4.1.

- We assume that the bispectra for different configurations are uncorrelated with each other. This means that we approximate the *bispectrum covariance matrix as diagonal*. In [96] it has been checked that this approximation works fine for the scales $k \leq 0.3h{\rm Mpc}^{-1}$ at redshift zero. We assume it holds up to $k \leq 0.4h{\rm Mpc}^{-1}$, since for local PNG we still gain signal up to this scale, as we see in Figure 5.9. Moreover, we assume that only the linear power spectrum determines the covariance matrix (see subsection 5.3.4 for more details). Finally, we neglect covariance due to observational effects, such as survey geometry and mask.

Importantly, we parameterize the theoretical error by treating the higher loop corrections to the bispectrum as a source of noise, which we integrate out. This contributes to the effective covariance matrix. Our estimate of the typical size of the two-loop corrections is given by B_{332}, defined in Appendix 5.A.

The time-dependence of the counterterms has been chosen to match the one loop diagrams they are supposed to renormalize [97] (see also [98] for a related discussion). All the time dependence is absorbed in the contributions to the bispectrum, so that all the theoretical parameters become

time-independent (see the definitions in 4). This means that we are measuring the same theoretical parameters in each redshift bin.
We need to discretize the bispectrum in order to compute the Fisher matrix. We will use logarithmic bins instead of linear bins, since we do not fully trust linear binning. We refer the reader to the appendix 5.C for more details. Finally, we do not marginalize over the standard cosmological parameters, but fix their values.

5.3.2 Statistical analysis and likelihoods

Let us briefly introduce the basics of the statistical analysis, based on the comprehensive review [99]. The most easily calculable statistical object is often the probability of a certain data output $\boldsymbol{x}$, given some statistical model $p(\boldsymbol{x}|\theta)$ that depends on a vector of parameters θ. For instance, one might have in mind a Gaussian model for the probability of the data depending on $\theta = (\mu, \sigma)$,

$$p(\boldsymbol{x}|\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\boldsymbol{x}-\mu)^2}{2\sigma^2}}, \tag{5.3.1}$$

where a sum over data points is assumed.[4]

In our case, we are interested in the opposite question: given a data set, what do we learn about the parameters of the model? To answer this, we invoke Bayes' theorem, which relates the posterior probability for the parameters to the probability above,

$$p(\theta|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\theta)p(\theta)}{p(\boldsymbol{x})}, \tag{5.3.2}$$

where $p(\theta)$ is the prior probability of the parameters (e.g. from a previous experiment), and $p(\boldsymbol{x})$ is the evidence, which for our purpose is merely a normalization factor. The relevant quantity is therefore the so called likelihood function

$$\mathcal{L}(\boldsymbol{x}|\theta, \text{priors}) \equiv p(\boldsymbol{x}|\theta)p(\theta). \tag{5.3.3}$$

[4] In a more elaborate setup, one might consider non-diagonal covariance among the data points. As argued, we assume diagonal covariance everywhere.

Given some data set, the parameters are then estimated by maximizing the likelihood function. For an unbiased estimator, this gives best fit parameter values $\theta_0 = \langle\theta\rangle$.

To quantify the error on the best fit parameter values, we need to go one step further. We assume that we can expand the log-likelihood function around the maximum,

$$\log \mathcal{L}(\boldsymbol{x}|\theta) = \log \mathcal{L}(\boldsymbol{x}|\theta_0) + \frac{1}{2}\,(\theta_i - \theta_{i,0})\,\frac{\partial \log \mathcal{L}}{\partial\theta_i \partial\theta_j}\,(\theta_j - \theta_{j,0}) + \dots\,, \tag{5.3.4}$$

such that the Hessian matrix $H_{ij} = \frac{\partial \log \mathcal{L}}{\partial\theta_i \partial\theta_j}$ quantifies the sensitivity of the likelihood to the parameters.

Finally, in the absence of actual data, it is still useful to forecast the sensitivity of a certain measurement to the parameters. To that end, one defines the Fisher matrix as the expectation value of the Hessian matrix,

$$F_{ij} = \langle H_{ij} \rangle\,. \tag{5.3.5}$$

Usually, we are interested in errors on a certain parameter after *marginalizing* over the other parameters. This means we take into account that different parameters might have a degenerate effect on the measurement, making it harder to identify the error on a single parameter. For a Gaussian likelihood, one can prove that the marginalized errors are given by,

$$\sigma_i^2 = \left(F^{-1}\right)_{ii}\,. \tag{5.3.6}$$

In this work we use this definition for the estimate of the forecasted error on a certain parameter. The unmarginalized error is obtained by simply inverting the Fisher element itself F_{ii}^{-1}. A toy model example of the application of this procedure can be found in 5.B.3, which also explains the way we treat the theoretical error in the model. Next, we present the actual Fisher matrix for out forecast.

5.3.3 Fisher matrix

In a Fisher forecast, one computes the expected curvature around the maximum of the likelihood. The likelihood is given by

$$\mathcal{L}(\text{data}|\Theta,\ \text{priors}) = \frac{1}{\sqrt{(2\pi)^N \det(C_B)}} P_{\text{prior}}(\Theta) \times \exp\left(-\frac{1}{2}\sum_{\hbar p} \Delta B(\hbar,\Theta) C_B^{-1}(\hbar, q, \Theta) \Delta B(p,\Theta)\right), \tag{5.3.7}$$

where Θ denote the set of theoretical parameters and N is the number of datapoints. We suppressed the time dependence. Here we use the Dutch calligraphic lower case symbols $\hbar$ as a shortcut for a triplet of wavenumbers on which the bispectrum depends, i.e. $\hbar = (k_1, k_2, k_3)$. Furthermore, ΔB corresponds to the difference between estimator and theoretical prediction $\Delta B(\hbar,\Theta) = \hat{B}(\hbar) - B(\hbar,\Theta)$ and C_B is the covariance matrix of the bispectrum $C_B = \langle \Delta B \Delta B \rangle$.

Neglecting the theoretical error for the moment, the theoretical prediction for the bispectrum is given in equation (5.2.1). The 8 parameters we include in the Fisher analysis are therefore $\{f_{\text{NL}},\ \xi,\ \epsilon_1,\ \epsilon_2,\ \epsilon_3,\ \tilde{\gamma},\ \tilde{\gamma}_1,\ \tilde{\gamma}_2\}$[5]. The parameter ξ also appears in the power spectrum and has been measured to be[6] 0.98 $h^{-2}\text{Mpc}^2$ [97]. Therefore, we can put a sharp prior on this parameter. The other parameters are unknown, but we expect them to be of the same order of magnitude (see [1, 45] for naive numerical estimates). Therefore, we can use a fiducial value of zero and a Gaussian prior with variance of 10.

For simplicity, we assume that all priors are Gaussian with covariance

[5] Here we denote $\tilde{\gamma} = f_{\text{NL}} \cdot \gamma$ and similarly for γ_i, so that the bispectrum is linear in all parameters. This is convenient for the Fisher analysis, as this makes the result independent of the fiducial values of the parameters. On the other hand, we effectively assume that the one-loop non-Gaussian counterterms have amplitudes independent of f_{NL}. Later in this Chapter we will find that these counterterms are negligibly small, therefore, this will not affect our results.

[6] Previous measurement gave (1.62 ± 0.03) $h^{-2}\text{Mpc}^2$[73] and (1.5 ± 0.03) $h^{-2}\text{Mpc}^2$ [45], but neglected two-loop corrections.

matrix C_Θ. Then the Fisher matrix is given by (see e.g. [99])

$$F_{ij} \equiv -\left\langle \frac{\partial^2 \log(\mathcal{L})}{\partial\Theta_i \partial\Theta_j} \right\rangle = \frac{1}{2}\mathrm{Tr}\left[C_B^{-1} C_{B,i} C_B^{-1} C_{B,j}\right] + B_i^T C_B^{-1} B_j + \left(C_\Theta^{-1}\right)_{ij}. \tag{5.3.8}$$

As we will see in a moment, our approximation for the covariance matrix does not depend on the theoretical parameters. Writing out the time dependence explicitly, the Fisher matrix simplifies to

$$F_{ij}(z) = \sum_{k,k'} B_i^T(k,z) C_B^{-1}(k,k',z) B_j(k',z) + \left(C_\Theta^{-1}\right)_{ij}(z) \tag{5.3.9}$$

for each redshift bin. Since the bispectrum is linear in all parameters - taking into account that ξ has been measured - the Fisher matrix is independent of the fiducial value of the theoretical parameters[7], which is very convenient for the analysis. To combine the results from the different redshift bins, we use that the parameters are the same in each bin, since we have fixed their time dependence. This time dependence is chosen to match the time dependence of the divergences they are supposed to cancel, motivated by [97]. The explicit expression can be found in Appendix 4.C. Therefore, we can compute the constraints on f_{NL} by summing the Fisher matrices and then marginalizing over the EFT parameters i.e.,

$$\sigma(f_{\mathrm{NL}}) = \sqrt{\left(\sum_z F_{ij}(z)\right)^{-1}_{11}}, \tag{5.3.10}$$

where we assumed that the entry of the Fisher matrix belonging to f_{NL} is the first. Note that we have not included cross correlations between bins, which means we might be throwing away valuable information. This effect will be studied elsewhere [100].

[7] To be more precise, the Fisher matrix is independent of the fiducial values of the parameters *to good approximation*. We choose a fiducial value for ξ of zero and in Section 5.4 we either specify a prior with $\sigma = 1$ for ξ or no prior at all. Therefore, the Fisher matrix has some dependence on the choice of ξ, but it will come exclusively from the non-Gaussian counterterm proportional to ξ. Again, since the non-Gaussian counterterms turn out to be extremely small, we expect this not to affect the results. Moreover, we have checked this explicitly by changing its fiducial value to 1.

5.3.4 Covariance matrix

To evaluate the Fisher matrix, we need to know the covariance matrix. Let us shortly review the derivation of the bispectrum covariance matrix. The estimator of the bispectrum is given by [101]

$$\hat{B}(k_1,k_2,k_3,z) = \frac{1}{V(z)V_{123}}\int_{\boldsymbol{q}_1}\int_{\boldsymbol{q}_2}\int_{\boldsymbol{q}_3}\delta(\boldsymbol{q}_1,z)\delta(\boldsymbol{q}_2,z)\delta(\boldsymbol{q}_3,z) \times \delta_D^3(\boldsymbol{q}_1+\boldsymbol{q}_2+\boldsymbol{q}_3), \tag{5.3.11}$$

with $V(z)$ the volume of the bin. The integration is over logarithmic bins centered around the given wavenumbers, i.e. $\ln(|\boldsymbol{q}_i|) \in [\ln(k_i)-\frac{1}{2}\Delta\ln(k), \ln(k_i)+\frac{1}{2}\Delta\ln(k)]$. Moreover, V_{123} corresponds to the following k-space volume squared

$$V_{123} = \int_{\boldsymbol{q}_1}\int_{\boldsymbol{q}_2}\int_{\boldsymbol{q}_3}\delta_D(\boldsymbol{q}_1+\boldsymbol{q}_2+\boldsymbol{q}_3) \approx \frac{8\pi^2}{(2\pi)^9}k_1^2k_2^2k_3^2\sinh^3(\Delta\ln k). \tag{5.3.12}$$

This approximation becomes exact when we consider 'internal' bins, but it fails on the 'edge' bins. In the numerical analysis we compute the exact value of V_{123} for each bin, see appendix 5.C for more details. This allows us to compute the bispectrum covariance matrix, namely

$$C_B(\hbar,\hbar',z) = \left\langle\left(\hat{B}(\hbar,z)-B(\hbar,z)\right)\left(\hat{B}(\hbar',z)-B(\hbar',z)\right)\right\rangle \tag{5.3.13}$$

$$\approx \frac{s_{123}}{(2\pi)^3V(z)V_{123}}P(k_1,z)P(k_2,z)P(k_3,z)\delta_{\hbar,\hbar'}. \tag{5.3.14}$$

There is a factor s_{123} which counts the number of non-vanishing contractions when computing $\langle\hat{B}(\hbar)\hat{B}(\hbar)\rangle$, which depends on the type of triangle that the triplet $\hbar$ forms. As each contraction comes with a delta function, this counting factor equals 6, 2 or 1 for equilateral, isosceles and scalene triangles respectively. If we include shotnoise in the covariance matrix, we replace $P(k_i,z)$ with $P(k_i,z)+\frac{1}{\bar{n}}$ in equation (5.3.14), where $\bar{n}$ is the effective number density for the density contrast. This will be explained more when we include shotnoise in Section 5.4.1.

In this expression for the covariance matrix, we completely neglected higher order corrections beyond the power spectrum, making it approximately diagonal. In [96] it has been checked that this approximation works

fine for the scales we are considering. The off-diagonal terms become important exactly when the higher order corrections to the power spectrum become important, since they are of the same order. Therefore, in order to be consistent, we only take into account the linear contribution P_{11} to the power spectrum $P(k_i, z)$. In particular, this means that the covariance matrix is independent of the theoretical parameters.

5.3.5 Theoretical error as nuisance parameters

To account for the theoretical error inherent to the perturbative expansion, we parameterize the bispectrum as

$$B(\hbar) = B^{\rm th}(\hbar) + n(\hbar) B^{\rm er}(\hbar)\,, \tag{5.3.15}$$

$$B^{\rm th}(\hbar) = B^{\rm G}_{\rm SPT}(\hbar) + B^{\rm G}_{\rm EFT}(\hbar) + f_{\rm NL}\left[B^{\rm NG}_{\rm SPT}(\hbar) + B^{\rm NG}_{\rm EFT}(\hbar)\right], \tag{5.3.16}$$

where $B^{\rm th}$ represents the theoretical prediction up to some order in perturbation theory as before, and $B^{\rm er}$ is the estimate of the theoretical error. Following [91], we introduce a series of nuisance parameters $n(\hbar)$, one per bin in k-space. The reason we implement the theoretical error this way, instead of proposing some k_{max} is that, as discussed in subsection 5.2.2, k_{max} depends on where the theoretical error and the signal become comparable. This complicates the analysis in two ways. First, k_{max} is configuration dependent, and second, it depends on the fiducial value of $f_{\rm NL}$, which makes finding the error on $f_{\rm NL}$ a recursive problem. In the approach we take, the set of theoretical parameters thus becomes $\Theta = \{n(\hbar)\}_\hbar \cup \{f_{\rm NL},\ \xi,\ \epsilon_1,\ \epsilon_2,\ \epsilon_3,\ \tilde{\gamma},\ \tilde{\gamma}_1,\ \tilde{\gamma}_2\}$. Since the bispectrum remains linear in all parameters, expression (5.3.9) for the respective block of the Fisher matrix still applies.

We assume that the true corrections to the bispectrum are of similar size as $B^{\rm er}(\hbar)$. Therefore, we put a Gaussian prior on the parameters $n(\hbar)$, with mean zero and variance one. Moreover, we expect the correction to have a smooth shape, which varies not too rapidly within the contours defined by $B^{\rm er}(\hbar)$. Therefore, the coefficients should have non-negligible cross correlations. Since we would like to have an increasing correlation for nearby points, we include cross-correlations with a typical correlation length as follows

$$N_{\alpha\beta} = \exp\left(-\frac{\sum_i \ln(|k^i_\alpha / k^i_\beta|)}{l}\right). \tag{5.3.17}$$

We replaced the label $\hbar_\alpha$ of the nuisance parameter of a given bin with the index α, so that we can reserve latin indices for the other theoretical parameters. Moreover, in order not to confuse this covariance matrix with the covariance matrix of the bispectrum C_B, we denote it as $N_{\alpha\beta}$. Note that we choose $\sigma_\alpha = 1$ for all α's. Here, l denotes the logarithmic correlation length. We could have also chosen a quadratic correlation length, similar to modeling it as a random field [102]. The reason we opted for this form is that here the inverse matrix is very sparse, which is convenient for numerical purposes. Since our final results are quite insensitive to the correlation length (see 5.B.3), we do not believe this choice affects the results very much.

Since we introduced a set of new nuisance parameters, we should write down the full Fisher matrix $F_{\mu\nu}$ and invert it

$$F^{-1}_{\mu\nu} = \begin{pmatrix} F_{\alpha\beta} & F_{\alpha j} \\ F_{i\beta} & F_{ij} \end{pmatrix}^{-1} = \begin{pmatrix} \bullet & \bullet \\ \bullet & (F_{ij} - F_{i\gamma}F^{-1}_{\gamma\delta}F_{\delta j})^{-1} \end{pmatrix}, \tag{5.3.18}$$

where we use latin indices for the parameters $\{f_{\rm NL},\ \xi,\ \epsilon_1,\ \epsilon_2,\ \epsilon_3,\ \tilde{\gamma},\ \tilde{\gamma}_1,\ \tilde{\gamma}_2\}$ and greek indices from the early alphabeth for the theoretical error parameters $\{n(\hbar_\alpha)\}_\alpha$. We did not write out explicitly the other entries, since we are only interested in the effective Fisher matrix, after marginalizing over the nuisance parameters coming from the theoretical uncertainty. To compute the effective Fisher matrix, we need to know $F_{\alpha\beta}$ and $F_{\alpha i}$. Since the derivative of the bispectrum with respect to the nuisance parameters Θ_α is only non-zero for the corresponding bin, these contributions to the Fisher matrix are particularly simple. We have

$$F_{\alpha\beta} = B^{\rm er}(\hbar_\alpha)C_B^{-1}(\hbar_\alpha,\hbar_\beta)B^{\rm er}(\hbar_\beta) + N^{-1}_{\alpha\beta} \equiv D_{\alpha\beta} + N^{-1}_{\alpha\beta}\,, \tag{5.3.19}$$

and similarly

$$\begin{aligned} F_{i\beta} &= \sum_{\hbar} B_i(\hbar)C_B^{-1}(\hbar,\hbar_\beta)B^{\rm er}(\hbar_\beta)\,; \\ F_{\alpha j} &= \sum_{\hbar} B^{\rm er}(k_\alpha)(\hbar_\alpha)C_B^{-1}(\hbar_\alpha,\hbar)B_j(\hbar). \end{aligned} \tag{5.3.20}$$

This allows us to compute $F^{\rm eff}_{ij}$ by using (5.3.18). After some algebraic manipulations, we can rewrite it in the simple form

$$F^{\rm eff}_{ij} = \sum_{\hbar,p} B_i(\hbar)\left(N^{\rm eff}(\hbar,p) + C_B(\hbar,p)\right)^{-1} B_j(p) + \left(C^{-1}_\Theta\right)_{ij}, \tag{5.3.21}$$

with $N^{\rm eff}_{\alpha\beta} = B^{\rm er}(k_\alpha) N_{\alpha\beta} B^{\rm er}(k_\beta)$. Again, the time dependence has been suppressed. In Appendix 5.B we present two alternative derivations of(5.3.21) and provide further detail.

In the next subsection we show the effectiveness of the current treatment of the theoretical error. However, we believe the interpretation of this method, and its relation to the actual situation, is a subtle matter. In particular, in Appendix 5.B.3 we argue by means of a simple toy model that this way of dealing with the theoretical error is certainly not the right way in the extremes of zero and maximal correlation among the parameters. Namely, on the one hand the theoretical error acts as shot noise per bin for zero correlation length, whereas for maximal correlation it acts as a single coefficient multiplying a fixed shape, effectively reducing the uncertainty about its shape to one number. Neither of these cases correspond to the way we believe the theoretical error should act. At the same time, Appendix 5.B.3 shows that the effect of the correlation length on the results is very mild. This suggests that the main reason our method works so well is that our ansatz for the error is a much steeper function of k than the signal, so that the size of the error is much more important than its shape. Thus, even though our treatment of the theoretical error seems to work for the current case, we recommend a conservative use of the method. In this spirit, we use the correlation lenght that gives the most pessimistic results for the analysis, which we found to be $l \sim 0.5$.

5.3.6 Testing the effect of the theoretical error

To test the method of integrating out the theoretical error, we study its effect on the constraints on f_{NL} in a χ^2-analysis. To that end, we compare two types of analyses, one which includes the theoretical error as outlined above, and one which does not. We generate a fake dataset with no primordial non-Gaussianity to test the theory. Our datapoints are given by

$$D(\hbar) = B_{112}(\hbar) + E_b(\hbar) + \text{cosmic noise}, \tag{5.3.22}$$

where we add some random noise, with variance equal to the cosmic variance, to each point. We consider a survey at redshift $z = 0$ with volume $V = 10\ (h^{-1}\text{Gpc})^3$, and restrict $(k_1,\ k_2,\ k_3)$ to be the central values of the binned range $[0.001, 1]\ h\text{Mpc}^{-1}$, where we take 27 logarithmic sized bins.

The additional contribution to the bispectrum is given by

$$E_b(k_1, k_2, k_3) = 3B_{112}(k_1, k_2, k_3) \left(\frac{k_1 + k_2 + k_3}{3k_{NL}} \right)^{(3+n)l}, \tag{5.3.23}$$

with $n = -1.4$, $k_{NL} = 0.45$ and $l = 2$. This is exactly the ansatz for the two-loop contribution to the bispectrum used in [91] and it is based on scaling universes [43]. In appendix 5.B.5, we compare the ansatz for the higher loop corrections E_b with our ansatz B_{332}. As theoretical model for the bispectrum, we use

$$B^{\text{th}}(\hbar) = f_{NL} \cdot B_{111}(\hbar) + B_{112}(\hbar). \tag{5.3.24}$$

We now consider two analyses. In the first analysis, we neglect theoretical errors and take only cosmic variance into account. In the second analysis, we use our ansatz for the higher order corrections, namely B_{332}, and we account for both theoretical error and cosmic variance. In order to find the best fit value for f_{NL}, we minimize χ^2_B, which is given by

$$\chi^2_B = \Big(D(\hbar) - B^{\text{th}}(\hbar)\Big) \Big(C_B(\hbar, \wp) + N^{\text{eff}}(\hbar, \wp)\Big)^{-1} \Big(D(\wp) - B^{\text{th}}(\wp)\Big) + \text{const.}, \tag{5.3.25}$$

see (5.B.13) in Appendix 5.B. In the first case, we set N^{eff} to zero. Minimizing χ^2_B yields

$$\text{Est}(f_{NL}) = \frac{B_{111}(\hbar) \left(C_B(\hbar, \wp) + N^{\text{eff}}(\hbar, \wp)\right)^{-1} \left(D(\wp) - B_{112}(\wp)\right)}{B_{111}(\hbar) \left(C_B(\hbar, \wp) + N^{\text{eff}}(\hbar, \wp)\right)^{-1} B_{111}(\wp)}, \tag{5.3.26}$$

and taking another derivative with respect to f_{NL} allows us to compute the standard deviation

$$\sigma(f_{NL}) = \left(B_{111}(\hbar) \Big(C_B(\hbar, \wp) + N^{\text{eff}}(\hbar, \wp)\Big)^{-1} B_{111}(\wp) \right)^{-1/2}. \tag{5.3.27}$$

With the best-fit value of f_{NL}, we can evaluate $\left(\chi^2_B\right)_{\text{red}}$ and the p-value, which are given by

$$\left(\chi^2_B\right)_{\text{red}} = \frac{\chi^2_B}{N} \quad \text{and} \quad p\text{-value} = 1 - \text{CDF}_{\chi^2}(N, \chi^2_B), \tag{5.3.28}$$

with CDF_{χ^2} the cumulative distribution function of the χ^2-distribution, and $N = N_{\mathrm{bins}} - N_{\mathrm{dofs}} - 1$ the number of datapoints minus one minus the number of fitting parameters. The p-value takes values between 0 and 1. It gives the probability of finding a higher value for χ^2_B if it was drawn from a χ^2-distribution. Therefore, it should take values around to 0.5. If the p-value is very close to zero, then the proposed theory vector is not a good description of the data. If the p-value is close to one, then either one is overfitting the data, or the estimate for the noise is too pessimistic.

In Figure 5.2, we plot the estimate for f_{NL} with errorbars, $\left(\chi^2_B\right)_{\mathrm{red}}$ and the p-value, as given in equations (5.3.26), (5.3.27) and (5.3.28) respectively, for the two analyses. In the left panel, we show both the results for the analysis in which the higher order corrections are neglected, and the analysis in which we use B_{332} as an ansatz. In the right panel, we use $10 \times B_{332}$ as error estimate to make sure that our ansatz is always bigger than the true value of the higher order corrections. One can check that E_b has a different shape than B_{332}. For instance, in the equilateral configuration, E_b is smaller than B_{332} on small scales (more optimistic). On the other hand, on large scales in the equilateral configuration, and in the squeezed limit, it tends to be larger than B_{332} (more pessimistic). Upon multiplying the latter by a factor 10, we find a robust, conservative estimate (see 5.B.5).

In the left panel of Figure 5.2, we see that if we neglect the theoretical error (blue lines and contours), we get the wrong value for the best fit value for f_{NL}, because higher order corrections are mistakingly interpreted as signal. Fortunately, the p-value singles out where the theoretical description fails. Taking this into account, we get a reliable estimate for f_{NL}, albeit with larger errorbars, since we have to stop already at a relatively small k_{max}. From the analysis that accounts for the theoretical error (red lines and contours), it seems we can continue the analysis to a higher k_{max}. However, the result we get for f_{NL} is biased, i.e. it is more than 5σ away from the actual value. The problem is that, in certain configurations of the bispectrum, our ansatz takes smaller values than the actual value in the data. This tends to decrease the p-value. At the same time, the p-value increases in the configurations where the theoretical error is overestimated. The interplay of these two effects can lead to a p-value, which is neither too small nor too large, and this gives rise to a biased estimate. Hence, if one wants to use the p-value as diagnostic for k_{max} and avoid biased results, it is important to have a fairly good understanding of the form of the theoretical noise. Alternatively, one can

work with an ansatz which is consistently *underestimating* the theoretical error. In this case, the p-value should go to zero rapidly, as soon as the theoretical error kicks in. As a double check, we did the analysis using $0.1E_b$ as ansatz instead, which indeed gives unbiased results similarly to the case where we neglect the theoretical error altogether. In general, when performing a datafit, if one has insufficient information about the higher order corrections, it is therefore safer not to integrate out the theoretical error at all. Summarizing, *assuming the wrong shape for the theoretical error might lead to a false detection of primordial non-Gaussianity.*

In the right panel (green lines and contours), we show the same results, where now the ansatz is always more pessimistic than the actual theoretical error ($10 \times B_{332} > E_b$). In this case, the estimate for f_{NL} is equal to the real value within 2σ. The p-value now is very *large* and it would naively tells us to stop at some smaller k_{max}. However, since we are obviously not overfitting the data, this reflects the fact that our ansatz for the theoretical error is too pessimistic. Therefore, we can safely evaluate the estimate for f_{NL} and the corresponding errorbar at the highest value of k_{max}, where the errorbar is frozen to a finite value. As expected, we find better errorbars than in the case we neglect the theoretical error. This shows that integrating out the theoretical error helps constraining f_{NL}, as long as one is careful to take a conservative enough ansatz. In the next section, we take B_{332} as ansatz for the theoretical error[8].

[8] We checked that our results for $\sigma(f_{NL})$ change with less than a factor of 2, when we use $10B_{332}$ instead.

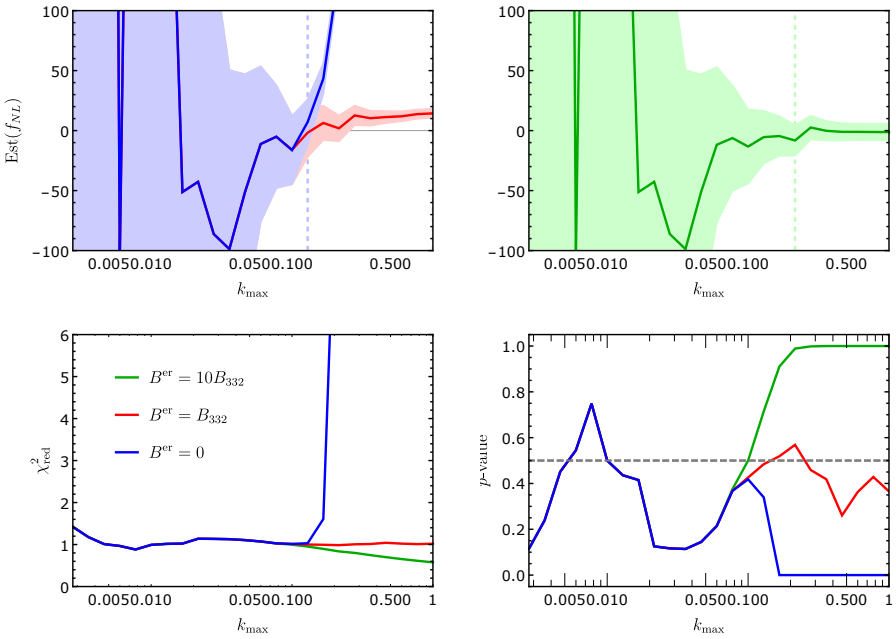


Figure 5.2: The figure shows the results from the χ-squared analysis for the data and theory given in equations (5.3.22) and (5.3.24). In the left panel, we show the results for both the analysis in which the higher order corrections are neglected (blue) and the one in which we use B_{332} as ansatz (red). In the right panel, we use $10B_{332}$ as ansatz (green) instead. In the upper panels, we show the best fit value for f_{NL} (solid line) as function of k_{max} and the lighter coloured regions correspond to the 2σ errorbars. The dashed vertical line corresponds to the largest value for k_{max} where the p-value is still between 0.01 and 0.99. The second and third row show $\left(\chi^2_B\right)_{\rm red}$ and the p-value as function of k_{max}.

5.4 Results

In this section, we present the main results of our analysis. First, we give $\sigma(f_{\rm NL})$ for various surveys, comparing our results to [91] and [84]. Next, we study the correlations among the EFT parameters for relevant surveys. Furthermore, we address the question of how much better the constraints would be, if we were able to compute the two-loop bispectrum. Finally, we show that the EFT of LSS clearly outperforms SPT in the constraints on $f_{\rm NL}$, where we assume the EFT contributions to the bispectrum are part of the theoretical error in SPT.

5.4.1 Constraints as function of $z_{\rm max}$

In this subsection, we compute the constraints on f_{NL} as function of maximum redshift for surveys similar to the ones studied in [91] and [84]. This allows us to study the effects of shotnoise and to compare our results with theirs. Furthermore, we show the effect of marginalizing over the EFT parameters for these surveys.

A large redshift survey (comparison with Baldauf et al. 2016)

First, to compare with [91], we focus on local and equilateral PNG. In the following, we list the specifications of the survey and the particular assumptions we make (in addition to general assumptions for the Fisher analysis discussed in subsection 5.3.1).

- We assume a survey with maximum redshift of $z = 5$. We divide the survey in redshift bins of equal volume, where the first bin runs from $z = 0$ to $z = 1$. We assume the survey covers 20000 deg^2, which means that, with our choice of cosmological parameters, the volume of each bin is given by $V = 26.5(h^{-1}\text{Gpc})^3$. We approximate each bin to be a cube, so that $k_{\rm min} \approx 0.002\ h\text{Mpc}^{-1}$. The maximum redshifts of all the redshift bins are given by[9] $\{1.00, 1.39, 1.71, 2.02, 2.31, 2.60, 2.89, 3.19,$

[9] Here, and in the next section, we make a particular choice of redshift bins. Larger redshift bins imply that we can include more configurations of the bispectrum in the analysis, in particular more configurations in the squeezed limit. At the same time, we fix the redshift of the bin to be the mean redshift, therefore, larger redshift bins imply a smaller maximum redshift. Therefore, the choice of binning might affect the final result. We will study a bin-independent approach in [100].

$3.49, 3.80, 4.11, 4.44, 4.78, 5.13\}$

- We restrict $(k_1,\ k_2,\ k_3)$ to the values in the binned range $[0.002, 1]\ h\text{Mpc}^{-1}$, where we use 15 logarithmic bins per decade[10].

- At high redshift, the late-time non-Gaussian background is particularly small and so the PNG signal becomes comparatively more pronounced. One the other hand, at high redshift, there are also much fewer tracers and this degrades our ability to measure the distribution of matter. To be able to capture this fact, we introduce a shotnoise that mimics what happens for example in galaxy surveys. For the purpose of comparison, we adopt the same convention for shotnoise as in [91], being the inverse of the effective number density of galaxies

$$\bar{n} = b_1^2 n_0 (1+z)^\alpha, \tag{5.4.1}$$

  where we correct for the galaxy bias $b_1 = 2$, since the shotnoise in [91] applies to galaxies, while here it has been translated to the dark matter density field. We choose $n_0 = 10^{-3} h^3 \text{Mpc}^{-3}$ and $\alpha = -1$.

- When we marginalize, we assume for each EFT coefficient a Gaussian prior with $\sigma = 10$, except for the EFT parameter ξ for which we take $\sigma = 1$.

- For B^{er} we consider both B_{332} and the ansatz E_b given in [91] (see (5.3.23)).

The results for $\sigma(f_{NL})$ are shown in Figure 5.3. We show the effects of the ansatz for B^{er}, shotnoise, and marginalization over the EFT parameters.

We can compare our results with those found in [91] by looking at the unmarginalized results, using their ansatz E_b for B^{er}. Thus, we should compare our dashed green lines with their dotted red lines in the last columns of their Figure 6 and 7. We indeed find a reasonably good agreement, given the fact that our analyses are not identical (different sky coverage, redshift bins and cosmological parameters, and moreover the translation of their shotnoise to ours is not perfect as we only took into account b_1). This check confirms that our code runs as expected.

[10] This is a little less than 45 bins over the full range, corresponding to $O(3000)$ triangles.

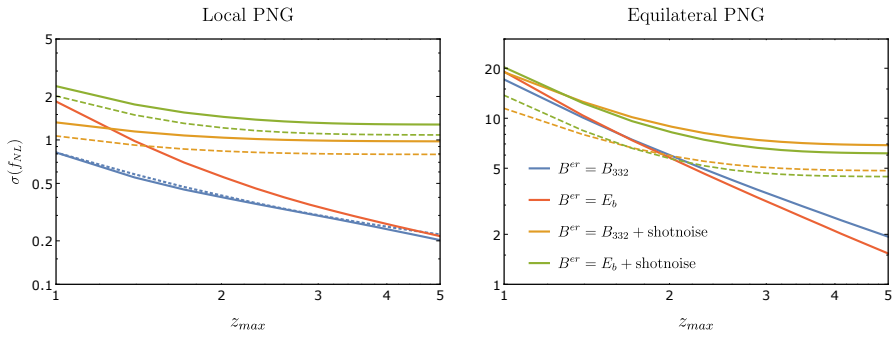


Figure 5.3: In these plots we show $\sigma(f_{NL})$ as function of maximum redshift $z_{\max}$ for local PNG (left) and equilateral PNG (right). We include shotnoise with the same specifications as [91] (orange and green lines). For the green lines we used E_b as ansatz for B_{2L}^G whereas for the other lines we used B_{332}. The blue solid line shows the result without shotnoise. The solid lines show the marginalized results assuming Gaussian priors for the EFT parameters. The dashed lines correspond to the unmarginalized result with the same color. Finally, the dotted blue line in the left panel corresponds to the curve $\sigma(z_1)/\sqrt{N}$ with N the number of redshift bins.

Let us now study the effect of the EFT parameters. The solid lines in Figure 5.3 correspond to marginalization over the EFT parameters, where we assume a Gaussian prior with $\sigma = 10$ for each EFT coefficient, except for the EFT parameter ξ, for which we take $\sigma = 1$. We see that *the results for local PNG are almost unaffected by marginalizing over the EFT parameters.* The constraints on equilateral PNG weaken by a factor of about two, however.

Using B_{332} as an ansatz for $B^{\rm er}$, we find slightly more optimistic results for local PNG, and slightly worse results for equilateral PNG, as compared with [91]. This can be understood from the comparison between B_{332} and E_b, shown in appendix 5.B.5. Local PNG peaks in the squeezed limit, and B_{332} is more optimistic than E_b in this configuration. On the other hand, equilateral PNG peaks in the equilateral configuration, and in this configuration E_b is more optimistic.

If we neglect shotnoise, we find that the differences at low redshifts are even bigger for the two ansätze. The difference is largest for local PNG, since

B_{332} is an order of magnitude bigger than E_b in the squeezed configuration, whereas the difference in the equilateral configuration is only of the order of a few. However, at higher $z_{\max}$ the differences disappear. This might seem a bit strange at first. However, we should stress that what we find here is not the true result in case of zero shotnoise. It turns out that with no shotnoise, we can always gain information in the ultra squeezed limit, and at higher redshift we can actually go to higher k than our choice $k_{\max} = 1h\text{Mpc}^{-1}$. Since $\sigma(f_{NL})$ does not freeze before we reach $k_{\max}$, this explains why the red and blue curves can get close for high $z_{\max}$. In appendix 5.B.5, we show these statements explicitly.

For comparison, note that if we were to always gain signal up to the same $k_{\max}$ in each redshift bin, we would find roughly the same $\sigma(f_{NL})$ for each redshift bin[11]. Therefore, combining all the redshift bins, we should find a $1/\sqrt{N_{\text{bins}}}$ behavior if we neglect shotnoise. The dotdashed blue line in the figure corresponds to $\sigma(z_1)/\sqrt{N_{\text{bins}}}$, which indeed resembles the blue solid line quite well. This provides another indication that we can go up to higher $k_{\max}$. Interestingly, we find that we can also go to much smaller scales for equilateral PNG than suggested by scaling estimates for $k_{\max}$ (see for instance [91]). The squeezed limit allows us to extract more information, also for equilateral PNG.

If we include shotnoise, it correctly cuts off the signal before we reach $k_{\max}$, so these results are reliable. However, one should keep in mind that, for more optimistic galaxy number densities, we might still extract more information from the ultra-squeezed limit. For the number densities we consider at the higher redshifts, shotnoise is the dominant source of noise. This is why we do include shotnoise in our analysis.

Current and upcoming surveys (comparison with Tellarini et al. 2016)

Next, we compare with [84], using the specifications of the surveys Euclid [103], BOSS [104], eBOSS [105] and DESI [106]. We have to consider the emission line galaxies (ELG), the luminous red galaxies (LRG) and quasars (QSO) separately, as they are measured at different redshifts, and have dif-

[11] To good approximation: the entry of the Fisher matrix, corresponding to f_{NL}, will scale as $F \sim D(a)^6/D(a)^6 \sim 1$, if we neglect the loop corrections to B_{NG}. Then, forgetting about the marginalization over the EFT parameters, we find the same $\sigma(f_{NL})$ in each redshift bin, since we took the bins so that they have the same volume.

ferent number densities and bias coefficients. The specifications and assumptions are as follows.

- For the precise number densities and bias coefficients as function of redshift we refer to appendix D of [84]. Moreover, one can find here the fraction of sky covered by each survey.
- As before, we divide each survey in equal sized redshift bins. The boundaries of all the redshift bins are given by

$$\begin{aligned}
\text{eBOSS (ELG)} &: \{0.6, 0.8, 0.95, 1.09, 1.21\} && V_{\text{bin}} = 2.8(h^{-1}\text{Gpc})^3, \\
\text{DESI (ELG)} &: \{0.1, 0.6, 0.79, 0.94, 1.07, \\
&\qquad 1.19, 1.3, 1.4, 1.5, 1.59, \\
&\qquad 1.69, 1.78\} && V_{\text{bin}} = 5.4(h^{-1}\text{Gpc})^3, \\
\text{Euclid (ELG)} &: \{0.6, 1., 1.28, 1.53, 1.75, 1.97\} && V_{\text{bin}} = 14.0(h^{-1}\text{Gpc})^3, \\
\text{eBOSS (LRG)} &: \{0.6, 0.75, 0.87, 0.98\} && V_{\text{bin}} = 2.0(h^{-1}\text{Gpc})^3, \\
\text{DESI (LRG)} &: \{0.1, 0.6, 0.79, 0.94, 1.07\} && V_{\text{bin}} = 5.4(h^{-1}\text{Gpc})^3, \\
\text{BOSS (LRG)} &: \{0., 0.4, 0.52, 0.61, 0.68, \\
&\qquad 0.75, 0.8\} && V_{\text{bin}} = 1.3(h^{-1}\text{Gpc})^3, \\
\text{eBOSS (QSO)} &: \{0.6, 1., 1.28, 1.53, 1.75, \\
&\qquad 1.97, 2.17\} && V_{\text{bin}} = 6.6(h^{-1}\text{Gpc})^3, \\
\text{DESI (QSO)} &: \{0.1, 0.8, 1.08, 1.31, 1.51, \\
&\qquad 1.7, 1.89\} && V_{\text{bin}} = 11.0(h^{-1}\text{Gpc})^3.
\end{aligned}$$

- We use k_{min} determined by the volume of each bin. Moreover, we choose the same binning of the k-range as in Section 5.4.1.
- The ansatz for shotnoise is now $\bar{n}(z) = b_1^2(z)n(z)$, where we correct for galaxy bias, similar as before.
- When we marginalize, we take the same prior as before. We assume for each EFT coefficient a Gaussian prior with $\sigma = 10$, except for the EFT parameter ξ for which we take $\sigma = 1$.
- As ansatz for the higher order corrections we use $B^{\text{er}} = B_{332}$.

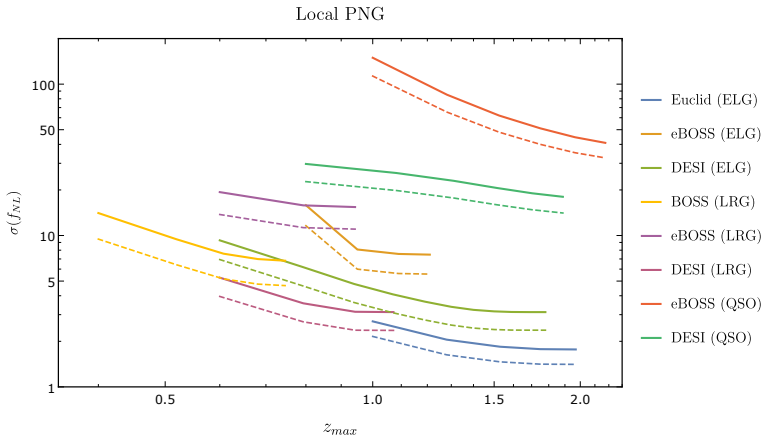


Figure 5.4: We show $\sigma(f_{NL})$ as function fo $z_{\max}$ for local PNG. We use the specification from Euclid (blue), BOSS (yellow), eBOSS (orange, purple and red) and DESI (green/yellow, pink and green). We show both the marginalized (solid lines) and unmarginalized results (dashed lines).

The results for local PNG are shown in Figure 5.4. We plot $\sigma(f_{NL})$ as function fo $z_{\max}$ for the four surveys. We show both the marginalized and unmarginalized results. After combining the different galaxy catalogs of a survey, we get $\sigma(f_{NL})$ for each survey, as summarized in Table 5.1a for local PNG and in Tables 5.1b and 5.1c for equilateral and quasi-single field PNG respectively.

We compare our results with [84] by looking at our unmarginalized results in Table 5.1a and their results in the last column in Table 1 of their paper. We find much weaker constraints, varying from 4 to 8 times smaller. This can be explained by the fact that we account for the theoretical error, which freezes the errorbars. Therefore, including the theoretical error gives rise to more conservative constraints. Moreover, scale-dependent bias could actually help us constrain local PNG also in the bispectrum. In fact, redoing the analysis for Euclid up to $k_{\max}(z) = 0.1D(z)$, as used in [84], with the same specifications, except that we ignore the theoretical error, gives $\sigma(f_{NL})$ equal to 0.57, 0.71 and 1.3 (unmarginalized, including and neglecting priors, respectively). This is roughly a factor three improvement from the results in Table 5.1a. The fact that this still does not challenge the results from [84] seems to indicate that scale-dependent bias helps to improve the constraints

$\sigma(f_{NL}^{\text{loc}})$	unmarg.	with prior	no prior
BOSS	4.67	6.81	17.3
eBOSS	4.91	6.6	14.15
Euclid	1.41	1.77	3.66
DESI	1.66	2.18	4.68

(a) Local PNG

$\sigma(f_{NL}^{\text{eq}})$	unmarg.	with prior	no prior
BOSS	16.89	29.86	37.99
eBOSS	17.25	26.88	33.4
Euclid	7.46	11.37	13.66
DESI	7.18	11.4	13.48

(b) Equilateral PNG

$\sigma(f_{NL}^{\text{qsf}})$	unmarg.	with prior	no prior
BOSS	12.57	23.65	27.26
eBOSS	13.1	21.43	23.49
Euclid	5.52	8.92	9.74
DESI	5.37	8.98	9.66

(c) Quasi-single-field PNG

Table 5.1: The final $\sigma(f_{NL})$ for local, equilateral and quasi-single field PNG for each survey, combining all expected galaxy catalogues. For the marginalized $\sigma(f_{NL})$, we put a Gaussian prior on each EFT coefficient with $\sigma = 10$, except for the EFT parameter ξ, for which we take $\sigma = 1$. In the last row, we also show the marginalized results, without prior on the EFT coefficients.

on primordial non-Gaussianity.

From the combined catalogs, we ultimately find $\sigma(f_{NL}^{\rm loc}) = 1.8$, $\sigma(f_{NL}^{\rm eq}) = 11.4$ and $\sigma(f_{NL}^{\rm qsf}) = 8.9$, with priors on the EFT parameters, assuming the surveys are not independent. These results do not change dramatically if we do not put priors on the EFT parameters. If the surveys are independent, the constraints improve approximately with a factor $1/\sqrt{2}$ upon combining Euclid and DESI.

5.4.2 Correlation coefficients

To gain intuition on how much the EFT parameters affect the constraints on f_{NL} for local, equilateral and quasi-single field PNG, we compute the correlation coefficients between parameters θ_i and θ_j. These are defined as

$$r_{ij} = \frac{F_{ij}^{-1}}{\sqrt{F_{ii}^{-1} F_{jj}^{-1}}}.$$

The correlation coefficient takes a value between 1 (perfectly correlated) and -1 (perfectly anti-correlated). In particular, the parameters are perfectly correlated with themselves. In Figure 5.5, we plot the absolute value of the correlation coefficients for each pair of parameters. We make the following assumptions

- We use the same binning in k-space as in Section 5.4.1.
- We use the redshift binning and shotnoise from Euclid, as given in Section 5.4.1.
- As ansatz for the theoretical error we use $B^{\rm er} = B_{332}$.
- We do not marginalize over the EFT parameters. The marginalized results are quoted in the text below.

We find that the groups of parameters $\{\xi,\ \epsilon_1,\ \epsilon_2,\ \epsilon_3\}$ and $\{\gamma,\ \gamma_1,\ \gamma_2\}$ have strong correlation among themselves. The correlation between f_{NL} and the other parameters is, however, small.

For local PNG, we find f_{NL} is mainly correlated with γ, γ_1 and ϵ_3, with correlation coefficients 0.44, -0.43 and 0.29 respectively. The other correlation coefficients are in absolute value smaller than 0.2. If we include a

Gaussian prior on the EFT parameters, with the same variances as before, the correlation coefficients become 0.12, 0.08 and 0.14.

In case of equilateral PNG, we find f_{NL} has appreciable correlation with ξ, ϵ_1, ϵ_3 and γ_2 ,with correlationcoefficients 0.43, -0.39, -0.47 and -0.27 respectively. Including priors on the EFT parameters, we find they become 0.14, -0.05, -0.30 and approximately zero. This could motivate further study on the Gaussian EFT coefficients, see for instance [107]. It is surprising that f_{NL}^{eq} is not extremely degenerate with ξ, since the latter comes with an additional k^2 scaling, similar to equilateral non-Gaussianity. It turns out, however, that the full shapes are sufficiently distinct. This will make it easier to constrain equilateral PNG from the bispectrum than naively thought. Then, for quasi-single-field PNG, we find that f_{NL} is mostly correlated with ϵ_3 and γ, with correlation coefficients -0.26 and 0.31. Including the priors, they reduce to -0.18 and approximately zero.

Summarizing, although the ignorance about EFT parameters does affect the final result, for reasonable priors this is only a small effect, especially for local PNG. This indeed agrees with what is seen in Figures 5.3 and 5.4, and Table 5.1a.

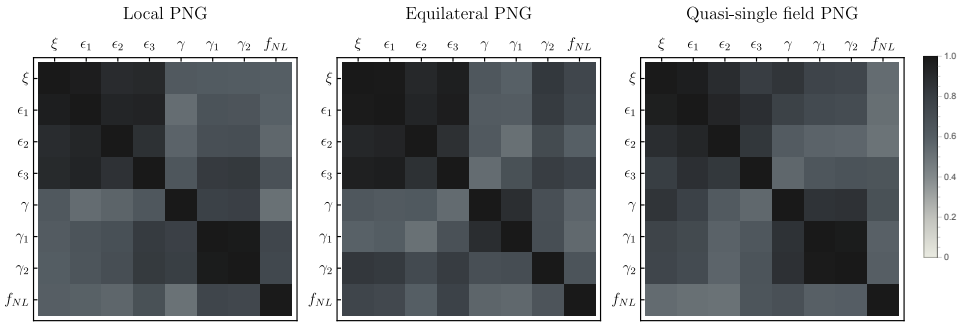


Figure 5.5: In these plots we show the correlation coefficients r_{ij} for each pair of theoretical parameters. We include shotnoise with the same specifications as Euclid and included all information up to redshift $z_{\text{max}} = 2$. A value of 1 (black) corresponds to perfectly correlated or anti-correlated. A value of 0 (white) corresponds to no correlation.

5.4.3 Higher loop corrections

We can ask ourselves whether it is useful to compute the bispectrum up to two loops in gravitational non-linearities. Note that our analysis does not depend on the actual value of the two-loop bispectrum, as there is no theoretical parameter in front. This means we can simply assume that we have computed all diagrams, neglect the counterterms, and assume the theoretical error is given by the SPT three loop bispectrum B^G_{3L}. Again, we do not know what it is, so we have to make an ansatz for it. Here we use the ansatz for the higher loop corrections from [91], given in (5.3.23), since it is easy to compute[12]. For a rough estimate this should suffice. We estimate B^G_{3L} using scaling universes [43]. We choose[13] E_b with $n = -1.5$, $k_{NL} = 0.5\ h\text{Mpc}^{-1}$ and $l = 3$ (see (5.3.23)). Using the specifications from Euclid, we perform the Fisher analysis with both B_{332} and B_{3L} as order of magnitude estimates for the noise. We collect the result in Figure 5.6. This shows that the constraints would improve if one computed the two-loop corrections to the bispectrum. The precise values are given in Table 5.2, where we consider all surveys again. The tighest constraints on local, equilateral and quasi-single field PNG improve with a factor 1.2, 1.3 and 1.3 respectively with this particular choice for B^G_{3L}. If it turns out we can get constraints on PNG close to the theoretical benchmarks, it would then be worth computing the two-loop corrections. It might be time consuming, but otherwise much cheaper than doubling the survey volume.

[12] An alternative - more in line with our two loop ansatz - would be to compute the reducible diagram of B_{433} as order of magnitude estimate of B^G_{3L}. However, we point out that using only one diagram is dangerous. For instance, in [1] we considered only one of the two reducible two loop diagrams in our qualitative analysis, namely B^I_{332}. In the squeezed limit, the two loop contribution turns out to be much larger if we include B^{II}_{332}.

[13] We choose a larger k_{NL} than for the two loop estimate, since this scale determines when the three loop correction equals the lower order corrections. Since we are doing a perturbative expansion, we assume this happens at a smaller scale than when the two loop correction becomes equal to its lower order corrections. Moreover, each loop will scale as k/k_{NL} to the power $3 + n$, where n will be of order of the scaling of the power spectrum at the non-linear scale k_{NL}. The power spectrum is steeper on smaller scales, therefore we take a more negative value for n. Each loop has this scaling, so we have to take $l = 3$, in case of three loops.

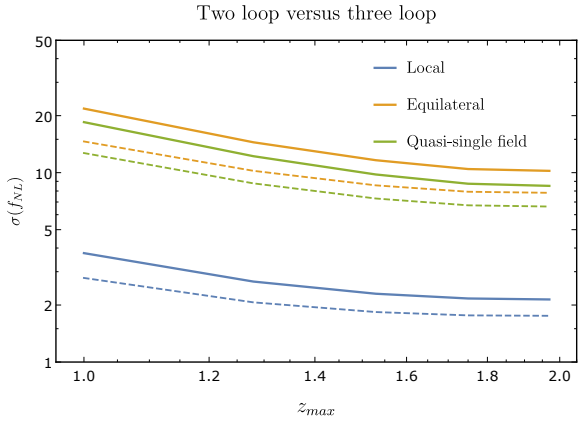


Figure 5.6: In this plot, we show $\sigma(f_{NL})$ as function of $z_{\max}$ for local (blue), equilateral (orange) and quasi-single field (green) PNG. We use both B_{2L} (solid lines) and B_{3L} (dashed lines) as order of magnitude estimation for the theoretical error. We use the specifications of Euclid in the analysis.

$\sigma(f_{NL}^{\rm loc})$	2 loop	3 loop
BOSS	8.73	6.05
eBOSS	7.12	6.07
Euclid	2.14	1.75
DESI	2.62	2.09

(a) Local PNG

$\sigma(f_{NL}^{\rm eq})$	2 loop	3 loop
BOSS	27.8	19.14
eBOSS	22.99	18.44
Euclid	10.22	7.83
DESI	10.2	7.81

(b) Equilateral PNG

$\sigma(f_{NL}^{\rm qsf})$	2 loop	3 loop
BOSS	23.66	16.65
eBOSS	19.15	15.59
Euclid	8.52	6.62
DESI	8.46	6.6

(c) Quasi-single-field PNG

Table 5.2: The final $\sigma(f_{NL})$ for equilateral and quasi-single-field PNG (left and right), for each survey, combining all expected galaxy catalogues. For the marginalized $\sigma(f_{NL})$, we put a Gaussian prior on each EFT coefficient with $\sigma = 10$, except for the EFT parameter ξ, for which we take $\sigma = 1$. In the last row, we also show the marginalized results without prior on the EFT coefficients.

5.4.4 EFT of LSS versus SPT

In the EFT of LSS we are forced to include free parameters over which we have to marginalize. Above, we saw that this marginalization weakens the constraints on $f_{\rm NL}$, be it only mildly. One might therefore wonder how much the improvement actually is over a more conservative approach, in which one uses only SPT results for $B^{\rm th}$ and moves all other gravitational contributions to the theoretical error. In this section we confirm that the EFT approach always performs sizably better. We consider a couple of options, for different choices of $B^{\rm th}$ and $B^{\rm er}$, and compute the constraints. We use the specifications from Euclid as given in the previous section.

Table 5.3 shows our results, which include the usual Gaussian priors for the EFT parameters whenever they are included in $B^{\rm th}$. The second and third columns give the theoretical description of the bispectrum $B^{\rm th}$ and what we consider to be the unknown $B^{\rm er}$. For the latter, we sum the absolute values of all contributions indicated in the table. The EFT contributions, except for ξ, are multiplied by a factor 10, consistent with the priors we chose when we included them in $B^{\rm th}$. The first row of Table 5.3 corresponds to the results we find in the 'with prior' columns of the 'Euclid' rows of Tables 5.1a, 5.1b and 5.1c.

We find that including the non-Gaussian counterterms does not improve the bounds on f_{NL}. This could have been anticipated from the qualitative results in Figure 5 - 7 of [1]. We see that the counterterms are negligible in many configurations. Apparently, they are negligible in most configurations. However, using the EFT for the Gaussian part of the bispectrum performs significantly improves the results compared with just the SPT predictions. We find that the *EFT of LSS improves the constraints on PNG approximately by a factor 3*. Finally, neglecting the one-loop non-Gaussian contribution to the bispectrum makes only about a 10% difference. This is consistent with the observation that the non-Gaussian counterterms are not very important, as the non-Gaussian one loop correction itself is not very relevant.

Approach	$B^{\rm th} = B_0^G + \dots$	$B^{\rm er} = B_{332} + \dots$	$\sigma(f_{NL}^{\rm loc})$	$\sigma(f_{NL}^{\rm eq})$	$\sigma(f_{NL}^{\rm qsf})$
EFT (G+NG)	$+B_0^{NG} + B_{EFT}^G + B_{EFT}^{NG}$		1.77	11.37	8.92
EFT G+SPT NG	$+B_0^{NG} + B_{EFT}^G$	$B_{\rm 2L}^G + B_{EFT}^{NG}$	1.78	11.37	8.92
SPT (G+NG)	$+B_0^{NG}$	$+B_{EFT}^{NG} + B_{EFT}^G$	6.11	27.61	21.76
SPT (G+NG tree)	$+B_{\rm tree}^{NG}$	$+B_{EFT}^{NG} + B_{EFT}^G + B_{1L}^{NG}$	7.17	30.58	24.23

Table 5.3: We show the constraints on primordial non-Gaussianity of the local, equilateral and quasi-single-field type (last three columns). In the first row, we use the EFT of LSS for both the Gaussian and non-Gaussian part of the bispectrum ('EFT (G+NG)'). In the second row, we only use the EFT for the Gaussian part of the bispectrum, and include the non-Gaussian counterterms in the higher order corrections ('EFT G + SPT NG'). Then, in the third row, we use only the SPT for describing the bispectrum ('SPT (G+NG)'). In the last row, we only include the tree level non-Gaussian contribution to the bispectrum ('SPT (G+NG tree)'). The second and third column denote all the contributions to the theoretical description of the bispectrum $B^{\rm th}$, and the higher order corrections $B^{\rm er}$ respectively.

5.5 Discussion and Outlook

In this work, we have presented how the EFT of LSS helps us improve the constraints on primordial non-Gaussianities (PNG), using the matter bispectrum as observable. We have accounted for intrinsic theoretical uncertainties in the perturbative description, and studied in details their modeling in a Fisher forecast.

Our main results are given in Table 5.3. The forecasted values for $\sigma(f_{NL})$ for the local, equilateral and quasi-single field types of PNG are presented. Moreover, we show that the EFT approach improves the constraints on PNG by almost a factor 3 with respect to the results from SPT.

Limitations Let us first discuss the limitations of these results. We would like to compare the constraints we find with theoretically interesting benchmarks and constraints coming from the CMB. However, we should be careful in making a direct comparison, as there are other sources of non-linearities and noise that we have not accounted for in our analysis. First, we have modeled the matter bispectrum. To relate it to the observed galaxy bispectrum, we have to include galaxy bias and redshift space distortions. These introduce new uncertainties, leading to worse constraints. However, considering

the results found in [83, 84], scale-dependent bias might actually *improve* the constraints on (only) local PNG, since it enhances the non-Gaussian signal in the bispectrum. Second, except for shotnoise, we neglected all observational sources of noise. Survey geometry and survey mask may increase the errorbars as well. For instance the authors of [85] found that the errorbars increased by a factor of 4-5. Errors in determining the redshift of galaxies are another source of observational noise. Third, we made some simplifications in the Fisher analysis itself, such as neglecting the covariance between different points of the bispectrum. Combining this with the covariance induced by the survey geometry could further increase the errorbars by a factor of 8 [108].

Improvements On the other hand, there are also ways the constraints could be improved. First, we have used the specifications of Euclid to get a reasonable estimate for the limitations due to shotnoise. This determined our final forecasted result for $\sigma(f_{NL})$. It might well be possible, in a more futuristic survey, to optimize the number densities of galaxies and redshift range to be more suitable for constraining PNG (see for instance [83]). Moreover, we should perform a joint analysis of all large scale structure surveys. We have assumed for simplicity that we can do as well as the single best survey, which turned out to be Euclid for the four surveys we considered. In principle, we can do better if the surveys are not all precisely overlapping. Similarly, we should combine the results from different observables. For instance, we should perform a joint analysis of the power spectrum and bispectrum. This could improve the results by a factor of 2 for local PNG. Combining the results found in [84], and using the multitracer technique proposed in [109] instead, which could improve upon the constraints from the power spectrum by a factor of about 7. In addition, the trispectrum might turn out to be an important probe for non-Gaussianity, since linear theory works for a larger range of scales compared to the bispectrum [110]. The one loop corrections to the trispectrum in the EFT of LSS have recently been computed in [111]. Last, we divided the full redshift range in smaller redshift bins, and only considered correlations within each redshift bin. If we also include correlations among galaxies separated by a larger distance along the line of sight, we might extract more information from a given survey. We will discuss the issue elsewhere [100]. Finally, our focus here was on near future galaxy

surveys, but of course our results will be relevant in the future also for 21 cm survey (see e.g. [112]).

One of our main results is that the EFT approach helps constraining PNG. The improvement comes completely from the EFT corrections to the late time gravitational non-linear evolution of matter. Both the SPT loops and EFT corrections to the primordial non-Gaussian signal, discussed in [1], do not help much improving the constraints.

Comparing our results with the theoretically interesting benchmark $\sigma(f_{NL}^{\rm eq}) \sim 1$, we see that it does not look promising for equilateral PNG. Even with zero shotnoise, as we can see in Figure 5.3, we barely touch the theoretical targets. Our lack of understanding of matter non-linearities is already an important obstacle to reach $\sigma(f_{NL}^{\rm eq}) \sim 1$. The same applies to quasi-single field PNG. Additional sources of non-linearities such as bias and redshift space distortions will make things worse. On the other hand, for local primordial non-Gaussianity, things look more promising. Matter non-linearities can be modeled well enough to get close to $\sigma(f_{NL}^{\rm loc}) \sim 1$ from large scale structure experiments.

We can ask whether N-body simulations can help reaching better constraints on primordial non-Gaussianity. As pointed out in [91], using end to end simulations, without any perturbative input, will most likely be insufficient to reach $\sigma(f_{NL}^{\rm eq}) \sim 1$. The reason is that simulations do not solve the exact problem but make a series of approximations, such as for example the particle mesh and tree approaches to solve Poisson equation, finite size effects and approximate initial conditions. Currently, simulations reach approximately 1% precision [113, 114]. Heuristically, looking at our Figure 5.1, we see that the PNG signal we are trying to extract is much smaller than that, so large improvements in the precision of simulations are needed. Alternatively, one can use N-body simulations to determine the EFT parameters[14]. We can then look directly at the unmarginalized columns in Table 5.1c. We see that, even in the very optimistic case that all relevant EFT parameters at one loop are fixed, $\sigma(f_{NL}^{\rm eq})$ still remains around 7.

Theoretical error Another goal of this Chapter is to clarify some aspects of the modeling of theoretical uncertainties in forecasting observational con-

[14] In fact, in our analysis, we assumed that a one EFT parameter, ξ, was fixed by fitting the power spectrum to simulations.

straints, and, eventually, in analyzing data (see Section 5.3.6 and Appendix 5.B). We introduced the concept of correlation length in Section 5.3.5, along the lines of [91]. In Appendix 5.B, we argued that the choice of correlation length in integrating out the theoretical error is subtle and no "right" choice can be established a priori. However, in our particular analysis, we hardly find any dependence on the correlation length (see Figure 5.7). In future studies, with different observables and different perturbative approaches, we believe that an analysis on the choice of correlation length should be always performed.

In Figure 5.2 we have seen that assuming the wrong shape for the theoretical error can lead to biased results in a χ^2-analysis. Therefore, if we want to fit to data, we need good estimates for the higher order corrections. For instance by using estimates from N-body simulations, or alternatively, by computing additional two-loop diagrams.

Outlook Our work can be extended and improved in different ways.

- Instead of dividing the survey volume in redshift bins and only consider correlations within these bins, it would be interesting to see how much we gain including all possible cross-correlations across redshift bins. This is work in progress [100].

- It would be interesting to perform a similar Fisher analysis with an updated study of covariance effects due to geometry, masking and non-Gaussian gravitational evolution.

- We should join all forces. It would be interesting to do a joint analysis of multiple observations, such as the CMB, LSS surveys and the 21 cm observations. Moreover, all the different LSS surveys should be combined to have maximum constraining power. Furthermore, the results from the power spectrum, bispectrum and trispectrum should be combined too. Finally, on the theory side, one should also try to use results from N-body simulations as soon as our perturbative description starts to break down, i.e. when the theoretical error becomes dominant.

5.A Ansatz Two Loop Bispectrum

As an ansatz for the two loop bispectrum we compute the two reducible two loop diagrams, given by [45],

$$B^{\mathrm{I}}_{332} = 2F_2(\boldsymbol{k}_1, \boldsymbol{k}_2)\frac{P_{13}(k_1)}{2}\frac{P_{13}(k_2)}{2} + 2 \text{ cycl. perms} \tag{5.A.1a}$$

$$B^{\mathrm{II}}_{332} = 6\frac{P_{13}(k_3)}{2}\int_{\boldsymbol{p}} F_3(-\mathbf{p}, \boldsymbol{p}-\boldsymbol{k}_2, -\boldsymbol{k}_3)F_2(\boldsymbol{p}, \boldsymbol{k}_2-\boldsymbol{p}) \times P_{11}(p)P_{11}(|\boldsymbol{p}-\boldsymbol{k}_2|) + 5 \text{ perms}, \tag{5.A.1b}$$

with

$$P_{13}(k) = 6P_{11}(k)\int_{\boldsymbol{p}} F_3(\mathbf{k}, \boldsymbol{p}, -\boldsymbol{p})P_{11}(p). \tag{5.A.2}$$

As an estimate for the theoretical error we use

$$B_{332} = |B^{\mathrm{I}}_{332}| + |B^{\mathrm{II}}_{332}|. \tag{5.A.3}$$

5.B Theoretical noise

This appendix contains the details of the implementation of the theoretical error and further investigates some issues related to it. First, we give the intermediate steps to derive (5.3.21) and provide an alternative derivation of the effective Fisher matrix in the presence of theoretical error. Then, we study the effect of the correlation length, both by means of a toy model and by running the analysis for several correlations lengths. Finally, we discuss in more detail the effect of the two possible ansätze for B^{er}.

5.B.1 Derivation of (5.3.21)

Let us first show how to go from equations (5.3.18), (5.3.19) and (5.3.20) to the effective Fisher matrix given in (5.3.21). We will need to use the Woodbury matrix identity several times, which relates the inverse of sums of matrices to their individual inverses

$$(A+B)^{-1} = A^{-1} - A^{-1}\left(A^{-1}+B^{-1}\right)^{-1}A^{-1}. \tag{5.B.1}$$

By using this identity, we can rewrite (5.3.19) as

$$F^{-1}_{\alpha\beta} = (N^{-1}_{\alpha\beta} + D_{\alpha\beta})^{-1} = D^{-1}_{\alpha\beta} - D^{-1}_{\alpha\gamma}\left(N + D^{-1}\right)^{-1}_{\gamma\delta} D^{-1}_{\delta\beta}. \tag{5.B.2}$$

This allows us to compute

$$\begin{aligned}
F_{i\gamma}F^{-1}_{\gamma\delta}F_{\delta j} &= \sum_{\hbar,p} B_i(\hbar)C^{-1}(\hbar,\hbar_\gamma)B^{\mathrm{er}}(k_\gamma)\left(D^{-1} - D^{-1}(N + D^{-1})^{-1}D^{-1}\right)_{\gamma\delta} \\
&\quad \times B^{\mathrm{er}}(k_\delta)C^{-1}(\hbar_\delta, p)B_j(p) \\
&= \sum_{\hbar,p} B_i(\hbar)\delta_{\hbar,\hbar_\alpha}\left(C^{-1}(\hbar_\alpha,\hbar_\beta) - \frac{1}{B^{\mathrm{er}}(k_\alpha)}(N + D^{-1})^{-1}_{\alpha\beta}\frac{1}{B^{\mathrm{er}}(k_\beta)}\right) \\
&\quad \times \delta_{\hbar_\beta,p}B_j(p) \\
&= \sum_{\hbar_\alpha,\hbar_\beta} \frac{B_i(\hbar_\alpha)}{B^{\mathrm{er}}(k_\alpha)}\left(D - (N + D^{-1})^{-1}\right)_{\alpha\beta}\frac{B_i(\hbar_\beta)}{B^{\mathrm{er}}(k_\beta)} \\
&= \sum_{\hbar_\alpha,\hbar_\beta} \frac{B_i(\hbar_\alpha)}{B_\alpha}\left(D(N^{-1} + D)^{-1}D\right)_{\alpha\beta}\frac{B_i(\hbar_\beta)}{B_\beta}.
\end{aligned}$$

Here summation over the Greek indices is understood. Upon applying the Woodbury identity again, the effective Fisher matrix then becomes

$$\begin{aligned}
F^{\mathrm{eff}}_{ij} &= \sum_{\hbar_\alpha,\hbar_\beta} \frac{B_i(\hbar_\alpha)}{B_\alpha}\left(D - D(N^{-1} + D)^{-1}D\right)_{\alpha\beta}\frac{B_j(\hbar_\beta)}{B_\beta} + \left(C^{-1}_\Theta\right)_{ij} \\
&= \sum_{\hbar_\alpha,\hbar_\beta} \frac{B_i(\hbar_\alpha)}{B_\alpha}\left(N + D^{-1}\right)^{-1}_{\alpha\beta}\frac{B_j(\hbar_\beta)}{B_\beta} + \left(C^{-1}_\Theta\right)_{ij} \\
&= \sum_{\hbar_\alpha,\hbar_\beta} B_i(\hbar_\alpha)\left(N^{\mathrm{eff}} + C_B\right)^{-1}_{\alpha\beta}B_j(\hbar_\beta) + \left(C^{-1}_\Theta\right)_{ij},
\end{aligned} \tag{5.B.3}$$

with $N^{\mathrm{eff}}_{\alpha\beta} = B^{\mathrm{er}}(k_\alpha)N_{\alpha\beta}B^{\mathrm{er}}(k_\beta)$.

5.B.2 Alternative derivation of the effective Fisher matrix

Next, we present a slightly different derivation of the effective Fisher matrix, by marginalizing at the level of the likelihood function. Let us first expand

$\chi^2 = -2\log(\mathcal{L})$ in the nuisance parameters Θ_α. We would like to expand about some value $\bar{\Theta}_\alpha$ to get

$$\begin{aligned}\chi^2(\Theta_i, \Theta_\alpha) =& \chi^2(\Theta_i, \bar{\Theta}_\alpha) + (\Theta_\alpha - \bar{\Theta}_\alpha)\chi^2_\alpha(\Theta_i, \bar{\Theta}_\alpha) \\ &+ \frac{1}{2}(\Theta_\alpha - \bar{\Theta}_\alpha)(\Theta_\beta - \bar{\Theta}_\beta)\chi^2_{\alpha\beta}(\Theta_i, \bar{\Theta}_\alpha),\end{aligned} \tag{5.B.4}$$

where summation over repeating indices is understood, and the index α on χ^2 denotes a derivative with respect to the corresponding nuisance parameter. It is an equality, since the variables are Gaussian distributed. We can rewrite this expression in more compact notation as

$$\chi^2 = \chi^2_0 + \delta\Theta^\alpha X_\alpha + \frac{1}{2}\delta\Theta^\alpha\delta\Theta^\beta Y_{\alpha\beta}, \tag{5.B.5}$$

where χ^2_0 is the chi-squared we would get if we ignored the presence of the nuisance parameters Θ_α. By completing the square and adding some prior information on the nuisance parameters (i.e. a covariance matrix), we can integrate them out to get an effective chi-squared. In other words, we would like to evaluate the following integral

$$\begin{aligned}\int d^N\Theta_\alpha \exp\left(-\frac{1}{2}\left(\chi^2_0 + \delta\Theta^\alpha X_\alpha + \frac{1}{2}\delta\Theta^\alpha\delta\Theta^\beta Y_{\alpha\beta}\right)\right) \\ \times \exp\left(-\frac{1}{2}\delta\Theta^\alpha (N^{-1})_{\alpha\beta}\delta\Theta^\beta\right).\end{aligned} \tag{5.B.6}$$

where $N_{\alpha\beta}$ is the covariance matrix of the theoretical error parameters. The integration results into

$$\sqrt{(2\pi)^N \cdot \det((\tfrac{1}{2}Y + N^{-1})^{-1})}\ \exp\left(-\frac{1}{2}\chi^2_0 + \frac{1}{4}X_\gamma \left(Y + 2N^{-1}\right)^{-1}_{\gamma\delta} X_\delta\right), \tag{5.B.7}$$

and therefore,

$$\chi^2_{\text{eff}} = \chi^2_0 - \frac{1}{2}X_\gamma \left(Y + 2N^{-1}\right)^{-1}_{\gamma\delta} X_\delta + \ln\left(\det((\tfrac{1}{2}Y + N^{-1}))\right). \tag{5.B.8}$$

Please note that all these terms do in general depend on Θ_i. Since the joint probability distribution of all parameters is a multivariate Gaussian,

we know that Y is independent of Θ_i and X only depends linearly on Θ_i. In that case, we get

$$\left(\chi^2_{\rm eff}\right)_{ij} = \left(\chi^2_0\right)_{ij} - \frac{1}{2} X_{i\gamma} \left(Y + 2N^{-1}\right)^{-1}_{\gamma\delta} X_{\delta j}. \tag{5.B.9}$$

The full Fisher matrix is given by

$$F_{\mu\nu} = \begin{pmatrix} \frac{1}{2}Y_{\alpha\beta} + N^{-1}_{\alpha\beta} & \frac{1}{2}X_{\alpha j} \\ \frac{1}{2}X_{i\beta} & F_{ij} \end{pmatrix} \tag{5.B.10}$$

where we have to evaluate the matrices at the maximum likelihood value of the parameters. This means that the effective chi-squared is given by

$$\left(\chi^2_{\rm eff}\right)_{ij} = \left(\chi^2_0\right)_{ij} - 2F_{i\gamma}F^{-1}_{\gamma\delta}F_{\delta j}, \tag{5.B.11}$$

or, in other words, the effective Fisher matrix for the theoretical parameters is given by

$$F^{\rm eff}_{ij} = F_{ij} - F_{i\gamma}F^{-1}_{\gamma\delta}F_{\delta j}. \tag{5.B.12}$$

This is what we found before in the main text.

Alternatively, starting from (5.B.9) we can write down immediately the expression for the effective likelihood

$$\begin{aligned}
\mathcal{L}^{\rm eff} &= \frac{1}{\sqrt{\det\left(\frac{1}{2}Y + N^{-1}\right)}} \exp\left[-\frac{1}{2}\left(\chi^2_0 - X_\gamma(\tfrac{1}{2}Y + N^{-1})^{-1}_{\gamma\delta}X_\delta\right)\right] = \\
&= \frac{1}{\sqrt{\det\left(D + N^{-1}\right)}} \exp\Bigg[-\frac{1}{2}\sum_{\hbar,p}\Delta B(\hbar)\Bigg(C_B^{-1}(\hbar,p) - C_B^{-1}(\hbar,\hbar_\gamma) \\
&\qquad B_{2L}(k_\gamma)(D+N^{-1})^{-1}_{\gamma\delta}B^{\rm er}(k_\delta)C_B^{-1}(\hbar_\delta,p)\Bigg)\Delta B(p)\Bigg] \\
&= \frac{1}{\sqrt{\det\left(D + N^{-1}\right)}} \exp\left[-\frac{1}{2}\sum_{\hbar_\alpha,\hbar_\beta}\Delta B(\hbar_\alpha)\left(C_B + N^{\rm eff}\right)^{-1}_{\alpha\beta}\Delta B(\hbar_\beta)\right].
\end{aligned} \tag{5.B.13}$$

Here the difference between the data and theory vector ΔB is evaluated at the fiducial values for the nuisance parameters $\bar{\Theta}_\alpha$ (i.e. at zero in our

case). Taking now two derivatives with respect to the remaining theoretical parameters Θ_i, we find the effective Fisher matrix

$$F_{ij}^{\rm eff} = \sum_{\hbar_\alpha,\hbar_\beta} B_i(\hbar_\alpha) \left(C_B + N^{\rm eff}\right)^{-1}_{\alpha\beta} B_j(\hbar_\beta) \tag{5.B.14}$$

with $N^{\rm eff}_{\alpha\beta} = B^{\rm er}(k_\alpha) N_{\alpha\beta} B^{\rm er}(k_\beta)$.

5.B.3 Theoretical error - a toy model

In our approach, the theoretical error on the value of the bispectrum is modeled in the following way. For every bin, we introduce a nuisance parameter that is drawn from a Gaussian distribution with average zero and variance set by the estimated size of the theoretical error for that bin. Importantly, we allow for non-vanishing correlations among these nuisance parameters, i.e. we allow for a non-diagonal covariance matrix for them. The purpose of this appendix is to show that both the limit of zero and maximal correlation among the parameters have a clear interpretation, neither of which resembles the way we think the theoretical error should act. To be more precise, we prove, by means of a simple toy model that still captures the essence of the real analysis, the intuitive statements that:

- for zero correlation length, the theoretical error just acts as shot noise per bin;
- for maximal correlation length, the theoretical error acts as some free coefficient multiplying a fixed shape function, which by definition we think is the wrong function.

Toy model

We consider measuring some observable d a total of N times and collecting the data d_i. Our model is $d_i = x + e_i$, with x a Gaussian random variable with variance σ_x^2, whose average, $\bar{x}$, we would like to determine as well as possible. The e_i are additional Gaussian variables that represent the systematic error or theoretical uncertainty in every measurement. Their averages and variances are $\bar{e_i}$ and $\sigma^2_{e_i}$, respectively. One can think of this scenario as determining the average weight of a group of people, knowing

that their weights are Gaussian distributed with variance σ_x^2, where we use a different weighing scale with a systematic error $\bar{e_i}$ and some uncertainty in the measurement characterized by $\sigma_{m_i}^2$ every time we weigh someone. Since the e_i are uncorrelated with x, this leads to the likelihood

$$\log L = -\sum_{i=1}^{N} \frac{(d_i - \bar{x} - \bar{e_i})^2}{2\sigma_d^2}, \tag{5.B.15}$$

where $\sigma_d^2 = \sigma_x^2 + \sigma_m^2$, assuming for convenience that $\sigma_{m_i} = \sigma_m$ (the arguments below do not depend on this assumption). Without any prior information on the systematic errors, they are completely degenerate with $\bar{x}$, so we do not expect to be able to learn anything about $\bar{x}$ in this case. This can be verified using a Fisher analysis. We have

$$F_{ab} = \begin{pmatrix} \frac{N}{\sigma_d^2} & \frac{1}{\sigma_d^2}\overrightarrow{1}^T \\ \frac{1}{\sigma_d^2}\overrightarrow{1} & \frac{1}{\sigma_d^2}1_{N\times N} \end{pmatrix}, \tag{5.B.16}$$

where $a, b = \bar{x}, \bar{e_i}$. Since we are ignorant about the systematic errors, we compute the marginalized error on $\bar{x}$,

$$\sigma_{\bar{x},marg}^2 = \left(F^{-1}\right)_{\bar{x}\bar{x}}, \tag{5.B.17}$$

which can be computed using the block matrix inversion formula (see also (5.3.18)):

$$\text{let} \quad F = \begin{pmatrix} A & B^T \\ B & D \end{pmatrix}, \tag{5.B.18}$$

then

$$\sigma_{\bar{x},marg}^2 = (A - B^T D^{-1} B)^{-1} = \left(\frac{N}{\sigma_d^2} - \frac{1}{\sigma_d^2}\overrightarrow{1}^T 1_{N\times N}\overrightarrow{1}\right)^{-1} = \frac{1}{0}, \tag{5.B.19}$$

as expected. In a realistic situation we do have some prior information about the systematic errors. Here, and in the body of the Chapter, we assume they are also Gaussian random variables with some variance $\sigma_{\bar{e_i}}^2$. Moreover, we allow for nontrivial correlations among the $\bar{e_i}$, which for the scales could

mean they were produced by the same machine for instance. This means we obtain the updated likelihood

$$\log L = -\sum_{i=1}^{N} \frac{(d_i - \bar{x} - \bar{e_i})^2}{2\sigma_d^2} - \bar{e_i} \left(C^{-1}\right)_{ij} \bar{e_j}, \tag{5.B.20}$$

where

$$C_{ij} = < \bar{e_i} \bar{e_j} > . \tag{5.B.21}$$

In the following, we investigate the effect of zero and maximal correlation length on $\sigma^2_{\bar{x},marg}$.

Zero correlation

Let us first assume zero correlation among the systematic errors, leading to a diagonal covariance matrix,

$$C_{ij} = \sigma_i^2 \delta_{ij}. \tag{5.B.22}$$

In terms of the weighing scales this could mean all scales really come from different companies with uncorrelated systematic errors. We now show that in this case the ignorance about the systematic errors acts as shot noise per bin; it simply updates the variance of the measurements $\sigma_d^2 \to \sigma_{d_i}^2$. For any covariance matrix, the Fisher matrix is

$$F = \begin{pmatrix} \frac{N}{\sigma_d^2} & \frac{1}{\sigma_d^2} \overrightarrow{1}^T \\ \frac{1}{\sigma_d^2} \overrightarrow{1} & \frac{1}{\sigma_d^2} \delta_{ij} + (C^{-1})_{ij} \end{pmatrix}. \tag{5.B.23}$$

Then, using the block matrix inversion formula, zero correlation leads to an error

$$\sigma^2_{\bar{x},marg} = \left(\frac{N}{\sigma_d^2} - \frac{1}{\sigma_d^4} \overrightarrow{1}^T \frac{1}{\frac{1}{\sigma_d^2} + \frac{1}{\sigma_i^2}} \delta_{ij} \overrightarrow{1} \right)^{-1} \tag{5.B.24}$$

$$= \left[\sum_{i=1}^{N} \left(\frac{1}{\sigma_d^2} - \frac{1}{\sigma_d^4} \frac{1}{\frac{1}{\sigma_d^2} + \frac{1}{\sigma_i^2}} \right) \right]^{-1} \tag{5.B.25}$$

$$= \left[\sum_{i=1}^{N} \left(\frac{1}{\sigma_d^2 + \sigma_i^2} \right) \right]^{-1} \equiv \left[\sum_{i=1}^{N} \frac{1}{\sigma_{d_i}^2} \right]^{-1}, \tag{5.B.26}$$

which is the same error one gets from assuming the likelihood function

$$\log L = -\sum_{i=1}^{N} \frac{(d_i - \bar{x})^2}{2\sigma_{d_i}^2}. \tag{5.B.27}$$

This shows that indeed for zero correlation the systematic errors acts as shot noise per bin. In particular, this means the error on $\bar{x}$ can be made arbitrarily small by increasing the number of measurements (if the $\sigma_{\bar{e_i}}$ do not grow too fast for additional measurements). The intuitive reason is of course that in this model we expect the systematic errors to average out to zero in the long run. This is clearly not what is expected of the theoretical error in the measurement of the bispectrum.

Maximal correlation

Next we assume maximal correlation, which by definition means

$$C_{ij} =< \bar{e_i}\bar{e_j} >= \sqrt{< \bar{e_i}^2 >}\sqrt{< \bar{e_j}^2 >} = \sigma_i\sigma_j. \tag{5.B.28}$$

Since this matrix has rank one (all columns are multiples of the same vector), it is not invertible in more than one dimension. One way to deal with this is to introduce a regulator, such as a small matrix $\epsilon\delta_{ij}$, to break the degeneracy. Using our block matrix inversion formula, this is however not necessary. In the notation of ((5.B.18)), we wish to compute D^{-1}. Let us write $D = S + C^{-1}$, where $S = \frac{1}{\sigma_d^2}\delta_{ij}$. The Woodbury identity then gives

$$(S + C^{-1})^{-1} = S^{-1} - S^{-1}(S^{-1} + C)^{-1}S^{-1}. \tag{5.B.29}$$

Hence, we have to compute the inverse of $S^{-1} + C$, where $S^{-1} = \sigma_d^2\delta_{ij}$, and $C = \sigma_i\sigma_j$. Conveniently, since C is of the form $\overrightarrow{\sigma}(\overrightarrow{\sigma})^T$, we can use the Sherman-Morrison formula to compute the inverse

$$(S^{-1} + C)^{-1} = S - \frac{S(\sigma_i\sigma_j)S}{1 + \sigma_i S_{ij}\sigma_j}. \tag{5.B.30}$$

Plugging this into the previous formula, we find

$$D^{-1} = (S + C^{-1})^{-1} = S^{-1} - S^{-1} + \frac{\sigma_i\sigma_j}{1 + \frac{(\sum\sigma_i^2)}{\sigma_d^2}} = \frac{\sigma_i\sigma_j}{1 + \frac{(\sum\sigma_i^2)}{\sigma_d^2}}. \tag{5.B.31}$$

This finally leads to the error on $\bar{x}$:

$$\sigma^2_{\bar{x},marg} = \left[\frac{N}{\sigma_d^2} - \frac{\frac{(\sum \sigma_i)^2}{\sigma_d^2}}{\sigma_d^2 + \left(\sum \sigma_i^2\right)}\right]^{-1}. \tag{5.B.32}$$

In order to interpret this result, let us rewrite this expression as follows

$$\sigma^2_{\bar{x},marg} = \left[\frac{N}{\sigma_d^2} - \frac{\left(\sum \frac{\sigma_i}{\sigma}\right)^2}{\sigma_d^4}\frac{1}{\frac{1}{\sigma^2} + \frac{\sum\left(\frac{\sigma_i}{\sigma}\right)^2}{\sigma_d^2}}\right]^{-1}, \tag{5.B.33}$$

where we have introduced the dimensionful parameter σ to keep the dimensions clean. Now observe that we get exactly the same error on $\bar{x}$ from the following likelihood function

$$\log L = -\sum_{i=1}^{N} \frac{(d_i - \bar{x} - \frac{\sigma_i}{\sigma}\bar{e})^2}{2\sigma_d^2} - \frac{\bar{e}^2}{2\sigma^2}. \tag{5.B.34}$$

whose Fisher matrix is

$$F = \begin{pmatrix} \frac{N}{\sigma_d^2} & \frac{1}{\sigma_d^2}\sum \frac{\sigma_i}{\sigma} \\ \frac{1}{\sigma_d^2}\sum \frac{\sigma_i}{\sigma} & \frac{1}{\sigma^2} + \frac{\sum\left(\frac{\sigma_i}{\sigma}\right)^2}{\sigma_d^2} \end{pmatrix}, \tag{5.B.35}$$

This means that the maximal correlation case is equivalent to having a single, unknown parameter multiplying a known 'shape' function σ_i/σ. In terms of the weighing scales this would mean that we know in advance exactly the ratios between the systematic errors of the scales. In terms of the bispectrum this would mean that we claim to know the theoretical error is exactly some number times the two loop estimate we put in, which it is clearly not. Finally note that if we choose all σ_i to be equal, which for convenience we take to be σ, we find

$$\sigma^2_{\bar{x}} = \left[\frac{N}{\sigma_d^2} - \frac{N^2}{\sigma_d^4}\frac{1}{\frac{1}{\sigma^2} + \frac{N}{\sigma_d^2}}\right]^{-1} = \frac{\sigma_d^2}{N} + \sigma^2, \tag{5.B.36}$$

meaning the error on $\bar{x}$ can never get below the uncertainty in the degenerate parameter $\bar{e}$. In terms of the weighing problem this makes perfect sense,

as this case is equivalent to simply using one and the same scale for every measurement. In this case we never expect to beat the unknown systematic error in the scale. In terms of the bispectrum this shows the importance of the relation between the shapes of the non-Gaussian signal and the theoretical error. In fact, in the maximal correlation limit we treat the theoretical error exactly the same as the EFT terms.

Conclusions

From the above example it is clear that in neither limit the implementation of the theoretical error is completely satisfactory. Moreover, if the shapes are not too similar, the estimates from both limits are probably too optimistic. For this reason we recommend a conservative use of the method. In particular, we choose to use the correlation length that gives the weakest constraints on $f_{\rm NL}$, as we show in the next subsection.

5.B.4 Choice of correlation length

In order to find the most conservative correlation length to work with, we ran a test computation of $\sigma(f_{\rm NL})$. We did this at redshift zero with $k_{\rm min} = 0.001\ h{\rm Mpc}^{-1}$ and $k_{\rm max} = 1\ h{\rm Mpc}^{-1}$, where we divide the k-range in 9 (blue), 15 (orange), 27 (green), 45 (red) and 81 (purple) bins. We find that the weakest constraints are obtained for $l \approx 0.5$, see Figure 5.7. This is therefore the value we take for the analysis.

Remarkably, the error is actually very insensitive to the correlation length, despite the very different nature of the effect of small and large correlation length. We believe the reason for this to be the fact that our ansatz for the theoretical error is a much steeper function of k than the non-Gaussian signal. The transition from the k's for which the error is negligible to the region where it is completely dominant is therefore very small, and the shape of the error is therefore not very important in this case.

Another observation is that the error keeps increasing as we increase the correlation length beyond 10 decades, whereas the k's we consider only run over a couple of decades. This makes the nuisance parameters almost maximally correlated for all these large correlation lengths. At the moment, we have no good explanation for the fact that the error seems to keep improving, other than it being a numerical fluke, perhaps related to the inversion

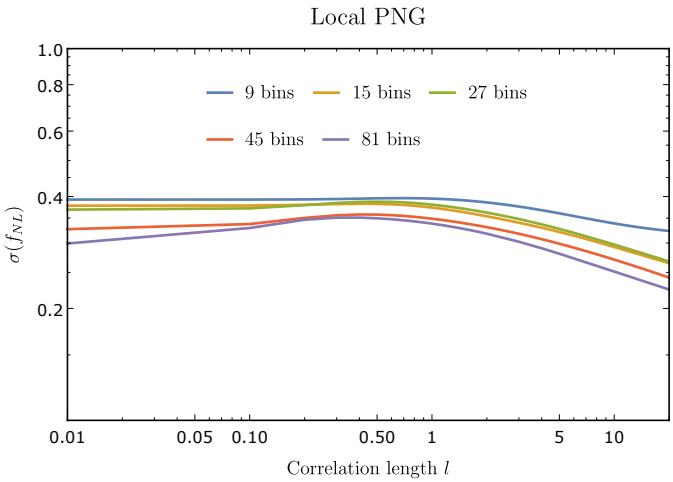


Figure 5.7: A test computation of $\sigma(f_{NL})$ at redshift $z = 0$ including theoretical error as function of correlation length. We choose $k_{\text{min}} = 0.001\ h\text{Mpc}^{-1}$ and $k_{\text{max}} = 1\ h\text{Mpc}^{-1}$ where we divide the k-range in 9 (blue), 15 (orange), 27 (green), 45 (red) and 81 (purple) bins.

of the correlation matrix.

5.B.5 Ansätze for higher loop corrections

We compare the ansätze for the higher loop corrections in Figure 5.8. It is a zoom-in of Figure 5.1, where we now show in addition the ansatz used in [91], for the two and three loop contribution to the bispectrum. Please note that in [1], we only showed one of the two reducible two loop diagrams contributing to B_{332}. Therefore, the plots look different now, in particular in the squeezed configuration of the bispectrum. We see that in the squeezed configuration, the ansatz B_{332} is an order of magnitude smaller than E_b. This explains why we have to multiply B_{332} by a factor 10 in section 5.3.6 to get reliable results. Furthermore, we note that at redshift zero, B_{332} allows one to go to higher k_{max} in the squeezed configuration, whereas E_b allows one to go further in the equilateral configuration (for f_{NL} bigger than 10). This explains why using B_{332} as an ansatz gives more optimistic results for local PNG, whereas E_b gives more optimistic results for equilateral PNG (see section 5.4.1). Keep in mind also that the time dependence of the theoretical

error terms is different from the signal, making the signal stronger at higher redshifts.

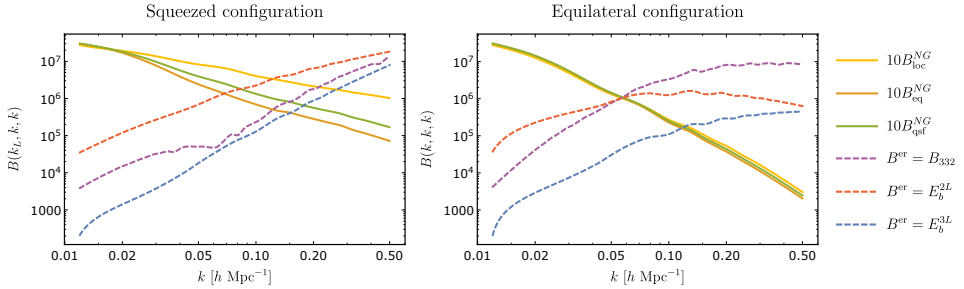


Figure 5.8: Comparison of the ansätze for the higher loop corrections. We plot E_b from equation (5.3.23) for two and three loops (yellow and green dashed lines) versus B_{332} (green dashed line) defined in equation (5.A.1). For the two loop ansatz using E_b we take $n = -1.4$, $k_{\rm NL} = 0.45h\mathrm{Mpc}^{-1}$ and $l = 2$ and for the three loop ansatz we use $n = -1.5$, $k_{\rm NL} = 0.50h\mathrm{Mpc}^{-1}$ and $l = 3$. We compare these ansätze with the non-Gaussian contribution to the bispectrum up to one loop with $f_{NL} = 10$ for local, equilateral and quasi-single-field PNG (red, blue and purple solid lines). In the left panel we compare the different contributions in the configuration $B(k_L, k, k)$ where we varied k and fixed $k_L = 0.012h\mathrm{Mpc}^{-1}$. The smaller k the more squeezed the configuration is. In the right panel we show the equilateral configuration $B(k, k, k)$.

Next, we consider $\sigma(f_{NL})$ as function of $k_{\rm max}$ at various redshifts in Figure 5.9. We do not include shotnoise, but we integrate out the theoretical error. The result for local PNG is shown in the left panel. We see that at redshift zero, the signal freezes out at some $k_{\rm max} < 1h\mathrm{Mpc}^{-1}$. Furthermore, in agreement with what we expect from Figure 5.8, we see that using B_{332} as ansatz for the two loop corrections gives more optimistic results. More specifically, $\sigma(f_{NL})$ is a factor 5 smaller. At higher redshifts, we find that $\sigma(f_{NL})$ does not freeze when we reach $k_{\rm max} = 1h\mathrm{Mpc}^{-1}$. We think this is due to the fact that we keep gaining information as we go to more and more squeezed configurations. This is also important for equilateral PNG, shown in the right panel, even though, $\sigma(f_{NL})$ does freeze out in this case. Interestingly, compared to scaling estimates for $k_{\rm max}$ for equilateral PNG (see for instance [91]) we find that we can go to much smaller scales than

naively thought. The squeezed limit allows us to extract more information, also for equilateral PNG. The fact that $k_{\max} = 1h\text{Mpc}^{-1}$ is not large enough to ensure that $\sigma(f_{NL})$ is frozen when we ignore shotnoise explains the results we find in section 5.4.1.

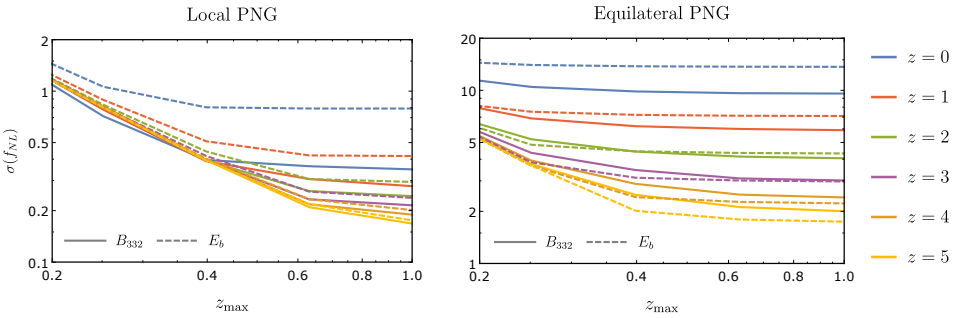


Figure 5.9: We show $\sigma(f_{NL})$ as function of $k_{\max}$ using B_{332} and E_b as ansatz for the two loop corrections (solid and dashed lines). In the left panel we show the results for local PNG and in the right panel for equilateral PNG. The redshift takes values between $z = 0$ and $z = 5$. We use $k_{\min} = 0.001h\text{Mpc}^{-1}$ and $V = (2\pi/k_{\min})^3$ at each redshift.

5.C Choice of binning and volume of the bins

In this appendix, we motivate the decision of section 5.3.4 to use logarithmic binning and exactly computed values of V_{123}.

5.C.1 Exact computation of V_{123}

We will now explain how to compute V_{123} exactly and systematically by dividing all the bins in 'interior' and 'edge' bins. Moreover, the selection of bins is now determined by whether it contains at least some valid triangles instead of the usual selection rule that the central point should be a triangle. Recall that V_{123} is defined as

$$V_{123} = \int_{\boldsymbol{q}_1} \int_{\boldsymbol{q}_2} \int_{\boldsymbol{q}_3} \delta_D(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3). \tag{5.C.1}$$

We choose logarithmic binning, i.e. we have

$$q_i \equiv |\boldsymbol{q}|_i \in \left[k_i e^{-\frac{1}{2}\Delta \ln k}, k_i e^{\frac{1}{2}\Delta \ln k}\right]. \tag{5.C.2}$$

The integrand above only depends on the relative orientations of the vectors and their lengths. Fixing $\boldsymbol{q}_1$ along the $\hat{z}$-direction and $\boldsymbol{q}_2$ to be in the (x, z)-plane, their relative orientation is given by $\theta_{12} = \theta_2$. Now the lengths of these vectors, together with c_{12}, the cosine of θ_{12}, completely determine $\boldsymbol{q}_3$. The length of $\boldsymbol{q}_3$ is then restricted to be in $\left[k_3 e^{-\frac{1}{2}\Delta \ln k}, k_3 e^{\frac{1}{2}\Delta \ln k}\right]$, which means

$$c_{12} \in [-1,1] \cap \left[\frac{\left(k_3 e^{-\frac{1}{2}\Delta \ln k}\right)^2 - q_1^2 - q_2^2}{2q_1 q_2}, \frac{\left(k_3 e^{\frac{1}{2}\Delta \ln k}\right)^2 - q_1^2 - q_2^2}{2q_1 q_2}\right], \tag{5.C.3}$$

where q_1 and q_2 also take values within their bin. Then, if $[-1, 1]$ contains the range on the right for all values of q_1 and q_2, we are dealing with an 'interior bin', and we get

$$\int dc_{12} dq_1 dq_2 \; q_1^2 q_2^2 = k_1^2 k_2^2 k_3^2 \sinh^3(\Delta \ln k). \tag{5.C.4}$$

Finally, accounting for the fact that we fixed θ_1, $\phi_{1,2}$ and the factors of $(2\pi)^3$ we find

$$V_{123} = \frac{1}{(2\pi)^9} 8\pi^2 k_1^2 k_2^2 k_3^2 \sinh^3(\Delta \ln k). \tag{5.C.5}$$

This approximation breaks down when the two ranges are partly overlapping, in which case we have an 'edge bin'.
let us evaluate V_{123} more precisely. We have seen that the integral simplifies to

$$V_{123} = \frac{8\pi^2}{(2\pi)^9} \int_{q_1} \int_{q_2} \int_{c_{12}} q_1^2 q_2^2 \tag{5.C.6}$$

where the c_{12} is restricted to be in the range given above. This integral can therefore be rewritten as

$$V_{123} = \frac{8\pi^2}{(2\pi)^9}\int_{q_1}\int_{q_2} \tfrac{1}{2}q_1 q_2 \max\left[0, \left(\min\left[(q_1+q_2)^2, k_3^2 e^{\Delta\ln k}\right] - \max\left[(q_1-q_2)^2, k_3^2 e^{-\Delta\ln k}\right]\right)\right].$$

In other words we integrate over the overlap

$$\left[(q_1-q_2)^2, (q_1+q_2)^2\right] \cap \left[k_3^2 e^{-\Delta\ln k}, k_3^2 e^{\Delta\ln k}\right]. \tag{5.C.7}$$

There are multiple possibilities:

- The overlap is always zero.
  This happens whenever $|q_1+q_2|_{\max} \leq k_3 e^{-\frac{1}{2}\Delta\ln k}$ or $|q_1-q_2|_{\min} \geq k_3 e^{\frac{1}{2}\Delta\ln k}$. This means we should exclude the cases $k_3 \geq (k_1+k_2)e^{\Delta\ln k}$ and $k_3 < k_2$. The latter is already excluded since we have $k_1 \leq k_2 \leq k_3$. The first leads to a constraint to select the bins, namely
  $$k_3 < (k_1+k_2)e^{\Delta\ln k}. \tag{5.C.8}$$

- The first range always contains the second range.
  This happens when $|q_1-q_2|_{\max} \leq k_3 e^{-\frac{1}{2}\Delta\ln k}$ ánd $|q_1+q_2|_{\min} \geq k_3 e^{\frac{1}{2}\Delta\ln k}$. So we need both
  $$\begin{cases} k_3 \geq k_2 e^{\Delta\ln k} - k_1 \\ k_3 \leq (k_1+k_2)e^{-\Delta\ln k} \end{cases}. \tag{5.C.9}$$
  In this case the volume takes the simple form
  $$V_{123} = \frac{8\pi^2}{(2\pi)^9} k_1^2 k_2^2 k_3^2 \sinh^3(\Delta\ln k). \tag{5.C.10}$$

- Any other type of overlap.
  For the other cases we have to compute the actual volume of the bin.

We will numerically perform the integral given above. This is when one of the two options below is satisfied

$$\begin{cases} k_3 < k_2 e^{\Delta \ln k} - k_1 \\ k_3 > (k_1 + k_2) e^{-\Delta \ln k} \end{cases} \tag{5.C.11}$$

Not for all these edge bins the central point has to be a triangle, since there are some cases for which $k_3 > k_1 + k_2$, considering the second inequality. Thus, we can either decide to define another point in the bin to represent the central triangle or we can merge these bins with one of their neighbors. In the first case a valid central triangle in the bin (k_1, k_2, k_3) is given by

$$\left(k_1 e^{\frac{k_3}{2(k_2+k_1)}}, k_2 e^{\frac{k_3}{2(k_2+k_1)}}, k_3 e^{-\frac{k_3}{2(k_2+k_1)}} \right) \tag{5.C.12}$$

The other option is to merge (i.e. we add the volumes) the bin with one of its neighbors, which has the advantage that we never have to change the representing triangle of a given bin. For practical reasons, we choose this option. We implement this by merging each bin $\boldsymbol{k}$ for which $k_3 > k_1 + k_2$ with the bin $\boldsymbol{p}$ which has $p_1 = k_1$, $p_2 = k_2$ and p_3 the biggest value below or equal to $k_1 + k_2$.

5.C.2 Logarithmic versus linear binning

Let us now compare logarithmic binning with linear binning. We will show two examples of a computation of a Fisher matrix and show that the linear binning might cause problems.
We assume the following form for the Fisher matrix

$$F = \sum_{k_1,k_2,k_3} f(k_1, k_2, k_3) \frac{V_{123}}{s_{123}}, \tag{5.C.13}$$

for some function $f(k_1, k_2, k_3)$. We will consider a 'local'-type function $f^{\rm loc}$ and an 'equilateral'-type function $f^{\rm eq}$. The local function corresponds to assuming the late time power spectrum scales as $P(k) \sim k^{-3}$, where F represents the (f_{NL}, f_{NL})-component of the Fisher matrix for local PNG. Forgetting about the right normalization, this gives

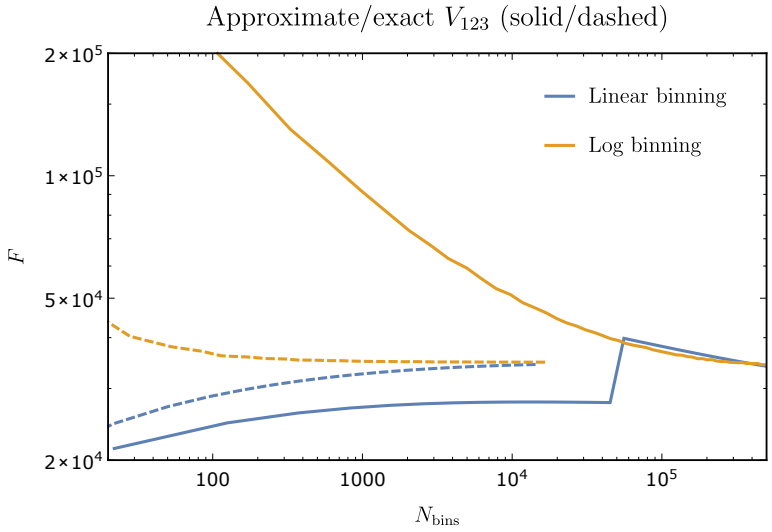
$$f^{\rm loc}(k_1, k_2, k_3) = \left(\frac{k_1^{3/2}}{k_2^{3/2} k_3^{3/2}} + \frac{k_2^{3/2}}{k_1^{3/2} k_3^{3/2}} + \frac{k_3^{3/2}}{k_2^{3/2} k_1^{3/2}} \right)^2 . \tag{5.C.14}$$

Similarly, we can define a function that corresponds to equilateral PNG

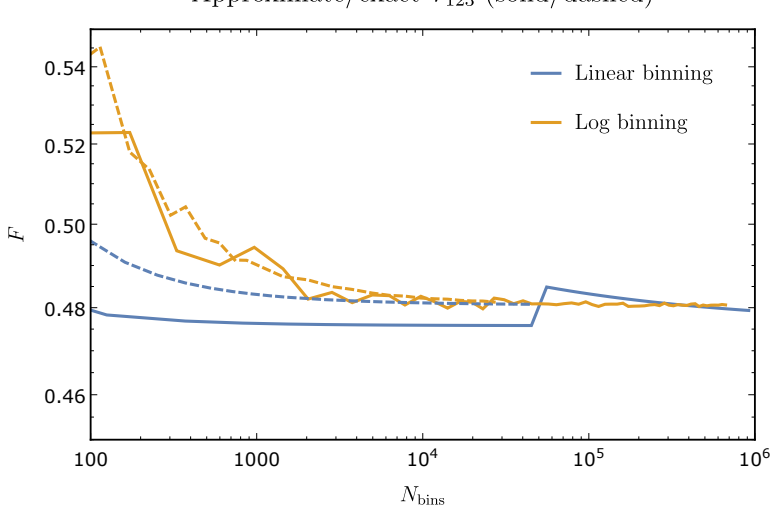
$$f^{\mathrm{eq}}(k_1, k_2, k_3) = \frac{k_1 k_2 k_3}{(k_1 + k_2 + k_3)^6}. \tag{5.C.15}$$

We compute F over a range $k \in [0.003, 0.5]$ $h\mathrm{Mpc}^{-1}$ for both logarithmic and linear bins and for both the approximate and exact computation of V_{123}. In Figure 5.10 we plot F as function of number of bins (number of triangles) considered.
First of all, in Figure 5.10a, restricting ourselves to about 10.000 bins, we see that if we use the approximate value for V_{123} (the usual assumption), the linear bins seem to converge quickly to the asymptotic value. The logarithmic bins seem to converge much slower. However, as we keep increasing the number of triangles, suddenly the graph of the linear bins jumps to the graph of the logarithmic bins. This shows that if we would have trusted the linear binning for a smaller number of bins we would have gotten the wrong result. This is quite unexpected and alarming as it seems we cannot always trust linear binning! If we now change to the exact V_{123} we see that both linear and logarithmic binning converge much faster and both to the same value. In fact, it turns out that they reach one percent agreement for about 15.000 triangles. Let us try to understand why this happens. As we are summing over a function which peaks in the squeezed limit we do in fact get most signal from k-triplets which satisfy $k_3 \sim k_2 \gg k_1$. In particular the edge bins will contribute an important part to the final result. We know that precisely for these bins the approximate value for V_{123} does not work, which is probably why the results improve dramatically when using the exact V_{123}. Now one can still wonder why the linear binning performs so badly in this case. A reason might be that we are sampling the values for k_1 much better in case of logarithmic binning. However one could argue exactly the opposite, namely that linear binning samples the values of k_2 and k_3 much better. We have not found a convincing argument why linear binning fails, this remains an open question. As we are also studying the Fisher matrix for local PNG, we decided to stick to logarithmic binning. Even with the exact value of V_{123} the result converges quite slowly for the local function. In order to be within a couple of percent of the actual outcome of the Fisher analysis we need quite some triangles. For the analysis therefore we divide the k-axis over three logarithmic decades in 45 bins. By this we mean that if for instance each k_i from the triplet can take values in

the range $[0.001, 1]$ $h\mathrm{Mpc}^{-1}$, it can take one of the 45 logarithmically separated values.
We did the same analysis for the equilateral function. In Figure 5.10b we see again a jump of the graph corresponding to linear binning. This time we do not expect to gain most signal from the edge bins. However, when we use the exact value for V_{123} everything seems to be fine again. The jump takes place at a comparable value of N_{bins}. The graph of the logarithmic binning remains a bit wiggly, but we find one percent agreement between linear and logarithmic binning already for 1000 triangles. For equilateral PNG we therefore divide the k-axis over three logarithmic decades in either 27 or 45 bins.



(a) Local function f^{loc}



(b) Equilateral function f^{eq}

Figure 5.10: Computation of F as function of the number of triangles N_{bins} for (a) f^{loc} and (b) f^{eq} as given in equations (5.C.14) and (5.C.15). We show the results for both the approximate (solid) and exact (dashed) expression for V_{123}. Moreover we denote the results from linear binning with a blue line and logarithmic binning with a orange line.

6 Divergence of Perturbation Theory in Large Scale Structures

6.1 Introduction

Exact results in physics are few and far between. Perturbation theory is often the only analytical tool available for both qualitative understanding and quantitative predictions. The dynamics of Large Scale Structures (LSS) is no exception. Because perturbation theory is at the core of most analytic LSS predictions [29], it is essential to thoroughly understand its regime of validity and have accurate quantitative estimates of its eventual departure from the exact result. In this work, we make progress in this direction by highlighting the divergence of perturbation theory for LSS correlators in some highly symmetric configurations.

The divergence of perturbation theory in Quantum Field Theory is very familiar to high energy theorist. Already more than 60 years ago, in less than two pages and with only two equations, Dyson showed that perturbation theory for Quantum Electrodynamics (QED) cannot converge and is at best asymptotic [115]. His very elegant argument[1] goes as follows. Physical quantities must be analytic functions of the QED coupling constant α. Perturbative approximations are given by a power series in α. This series must converge in the complex α plane within a ball of radius the distance to the closest singularity. For arbitrarily small but negative α we expect the vacuum to be unstable towards the quantum creation of a large number of pairs of oppositely charged particles. All equally charged particles can be bunched together reducing the energy (since $\alpha < 0$) and hence satisfying energy conservation. So any $\alpha < 0$ predicts an unstable ground state, which is infinitely different from the free, $\alpha = 0$ theory. The radius of convergence is therefore vanishing.

This elegant argument is intrinsically quantum mechanical in that it relies on pair creation out of the vacuum (violating instantaneous energy conservation in accordance with Heisenberg's uncertainty principle). Yet, the non-convergence of perturbation theory is much more general and ubiquitous even in classical systems with $\hbar = 0$. Intuitively, perturbation theory fails to converge whenever the perturbative solutions of the deterministic equations have finite radius of convergence *and* some averaging process, such as the QFT path integral or some stochastic average, probes solutions that lie outside that convergence region. To develop some intuition, let us consider

[1] The title of this Chapter is an homage to Dyson's classic contribution.

the following toy model. Assume δ is some random variable (not a space-dependent field). Let us assume that δ is related to a Gaussian variable δ_L in a non-linear way. For concreteness and relevance to LSS studies, we take this relation to be the solution of 0+1D gravitational collapse[2]:

$$\delta = \frac{\lambda\delta_L}{1-\lambda\delta_L}\,, \tag{6.1.1}$$

where, to emphasize that the perturbation expansion in small δ_L, we introduced[3] a dummy "coupling constant" λ. We will compute the variance of δ. The discussion for other cumulants is analogous. By its definition[4]

$$\begin{aligned}\langle\delta^2\rangle_\lambda &= \int_{-\infty}^{\infty}\frac{d\delta_L}{\sigma_L\sqrt{2\pi}}e^{-\frac{\delta_L^2}{2\sigma_L^2}}\delta^2 \\ &= \int_{-\infty}^{\infty}\frac{d\delta_L}{\sigma_L\sqrt{2\pi}}e^{-\frac{\delta_L^2}{2\sigma_L^2}}\left(\frac{\lambda\delta_L}{1-\lambda\delta_L}\right)^2\,,\end{aligned} \tag{6.1.2}$$

where we introduced the variance σ_L^2 of the Gaussian random variable δ_L. In more realistic examples, σ_L is a function of scale, but for the moment we neglect this complication, i.e. we have zero spatial dimensions.

The integral (6.1.2) does not converge because of the divergence at $\lambda\delta_L = 1$. This pathology is peculiar to the zero-dimensional case and does not play a role in more realistic cases such as 1+1D and 3+1D dynamics, so we just regularize it by hand, considering

$$\frac{\delta_L}{1-\lambda\delta_L} \to \frac{1}{\epsilon}\arctan\left[\frac{\epsilon\delta_L}{1-\lambda\delta_L}\right]\,, \tag{6.1.3}$$

for some small but finite ϵ. For any finite $\epsilon > 0$, the variance $\langle\delta^2\rangle_{\epsilon\lambda}$ is finite. For small ϵ, the perturbative expansion of δ around $\lambda = 0$ is independent of ϵ and so it is the same as for $\epsilon = 0$ (up to $\mathcal{O}(\epsilon)$ corrections). We therefore neglect $\mathcal{O}(\epsilon)$ corrections in the rest of the discussion and focus on the perturbative expansion in λ. Each order in perturbation theory around $\lambda = 0$

[2]The time dependence $\delta(t)$ can be easily added by using $\delta_L(t) = t^{2/3}\delta_p$ for some constant δ_p. Since it is irrelevant for this argument we simply drop it.

[3]This can be thought of as the coupling constant of all non-linear terms in the fluid equations (6.2.3), which are all quadratic in perturbations.

[4]The vacuum average $\langle 1\rangle$ is identically 1 both exactly and in perturbation theory, since λ does not appear in the usual way of computing the integral.

is finite and the series has the factorial growth typical of asymptotic series[5]. Using

$$\left(\frac{\delta_L}{1-\lambda\delta_L}\right)^2 = \delta_L^2 \sum_n^\infty (1+n)\,(\lambda\delta_L)^n\,, \tag{6.1.4}$$

$$\langle \delta_L^{2m} \rangle = \left(\frac{\sigma_L^2}{2}\right)^m \frac{(2m)!}{m!}\,, \tag{6.1.5}$$

we find

$$\langle \delta^2 \rangle_\lambda^{\rm PT} = \sum_{m=0}^\infty (1+2m) \left(\frac{\lambda\sigma_L}{\sqrt{2}}\right)^{2m} \frac{\sigma_L^2}{2} \frac{(2m+2)!}{(m+1)!}\,. \tag{6.1.6}$$

One can the use Stirling formula to expand this for large m

$$\langle \delta^2 \rangle_\lambda^{{\rm PT},m} \sim 4\sqrt{2} m^2 \sigma_L^2 \left(\frac{2m\lambda^2\sigma_L^2}{e}\right)^m\,. \tag{6.1.7}$$

The perturbative calculation therefore starts diverging at order $n = 2m \simeq e/(\lambda^2\sigma_L^2)$ and therefore

$$\langle \delta^2 \rangle_\lambda^{\rm PT} \neq \langle \delta^2 \rangle_\lambda\,. \tag{6.1.8}$$

Summarizing, we have proven that the perturbative series (6.1.6) does not converge for any finite value of λ, it has zero radius of convergence. The series is nevertheless asymptotic to the right (regularized) answer because at every finite perturbative order n one has

$$\lim_{\lambda\to 0} \left[\langle \delta^2 \rangle_\lambda^{\rm PT,n} - \langle \delta^2 \rangle_\lambda\right] = 0\,. \tag{6.1.9}$$

What happened? The non-linear relation (6.1.1) between δ and δ_L admits a perturbative approximation around $\lambda = 0$ (equivalently $\delta_L = 0$, but for extra clarity we formulate it in terms of the fictitious coupling constant) that has a finite radius of convergence $|\lambda\delta_L| < 1$. But the average in (6.1.2) extends all the way to $\delta_L = \infty$. The result is that every perturbative correction to

[5] Here we use the adjective asymptotic to refer to non-convergent asymptotic series, as it is often done in the physics literature, even though of course convergent series are also asymptotic.

the variance contains an error coming from the exponentially damped tails of the integral. This error grows with the perturbative order because of the growth of the order of the polynomial approximation in (6.1.4). In more realistic cases such as 1+1D or 3+1D dynamics, we believe it is still true that the perturbative solution of the deterministic equations of motion has finite radius of convergence. The exact solutions to planar and spherical collapse we discuss in section 6.2 support this idea. One therefore generically expects perturbation theory to be divergent (and asymptotic) also in more realistic cases. In this work, we show explicitly that this is the case for real space correlators and count-in-cell statistics in 1+1D. Perhaps surprisingly, but as anticipated in [116], perturbation theory instead converges in Fourier space. We are certainly not the first to investigate the convergence of perturbation theory for LSS [116–128], and we refer to the relevant literature in due course.

It is important to stress that the non-convergence of perturbation theory we discuss in this work has nothing to do with the improvements advocated by the EFT of LSS [11]. This is easily seen since the EFT corrections arise from smoothing short scale dynamics and hence disappear as we take the short scale power to zero. The non-convergence we discuss here instead does not disappear in this limit. More intuitively, the EFT of LSS corrections captures the effect of the non-perturbative short scales on large scales. Non-convergence of perturbation theory instead results exclusively from large scales, with arbitrarily small power. We come back to this point in section 6.5.3.

There are actually two conceptually distinct ways in which perturbation theory can fail to approximate some desired result: it might not converge, as we have just seen, or perturbation theory might converge to a result that is not the right one. The second situation arises also in all realistic LSS computation. In both Eulerian and Lagrangian approaches one cannot fully capture multistreaming (but see [129] for an exception) and therefore even if perturbation theory converged, it would not describe the correct physical result. While we originally attempted to makes progress in this direction as well, we have been able only to derive rough estimates for the non-perturbative corrections coming from multi-streaming. We have collected them with some general remarks in section 6.5.

Before diving into the derivation of our results, it is important to explain why one should care about non-perturbative results, since their exponentially small amplitude is typically trumped by larger perturbative correc-

tions. There are several reasons. First, since the dawn of time, analytical approaches to LSS have been trying to push predictions closer and closer to the non-linear scale, where all perturbative approximations break down. Around the non-linear scale perturbative and non-perturbative corrections both become of order one! Which one is largest might depend on numerical factors that are impossible to predict a priori. Therefore, a conservative estimate of the theoretical error of perturbative methods, as advocated recently in [2, 91], should include non-perturbative corrections as well. Second, some non-perturbative corrections might break symmetries that are respected by perturbative terms. Tunneling in quantum mechanics for example is invisible to perturbation theory (see e.g. [130]). In the context of LSS, scale dependent bias [81] is a relevant example: it is a non-perturbative effect (since the tracers of interest are non-perturbative objects) that cannot be mimicked by standard, late-time gravitational evolution as consequence of the equivalence principle. It would be nice to find other non-perturbative observables with an equivalent sensitivity to primordial initial conditions. Last but not least, a physicist has so few occasions to glimpse at what lies beyond perturbation theory that any chance should be taken advantage of.

For the convenience of the reader we summarize here our main results.

- The convergence of 1D Standard Perturbation Theory (SPT) to the Zel'dovich result (ZA), which is exact in 1+1D before shell crossing, was recently established by McQuinn and White [116] for ΛCDM-like initial conditions. We review their derivation and formalize one technical but crucial step. We stress that convergence relies on the (realistic) assumption that the variance of the displacement is finite. In fact, SPT is shown to diverge for scaling universes with a negative spectral tilt $P_L \propto k^n$ with $-1 < n \leq 0$ [131] (see subsection (6.3.3)). We generalize the convergence result for ΛCDM-like initial conditions to the bispectrum and for non-Gaussian initial conditions.

- We prove analytically and verify numerically that instead, perturbation theory does *not* converge for the real space equivalent, namely the correlation function. The technical reason is that the Fourier transform integral cannot be interchanged with the infinite sum over perturbative contributions. More intuitively, we show that the reason for non-convergence is a non-perturbative tail contribution to the correlation function, similar to the toy model above. We argue that the

non-convergence is related to the breakdown of our theory in case of rare fluctuations.

- The relevance and potential non-perturbativity of tails of the probability distribution function for the average density δ_R in a cell of radius R has been noted in various places, e.g. [123, 132]. We prove and verify numerically that, in the context of their 1+1D equivalent, there is indeed a finite radius of convergence for perturbation theory for this PDF. We show that any perturbative computation of cumulants is therefore asymptotic. Again, we highlight the analogy with the toy model.

- Along the way, we present a new derivation of this count-in-cell PDF in 1D, which is unitary with unit mean by construction.

The rest of this Chapter is organized as follows. In section 6.2, we collect standard results about exact solutions for gravitational collapse. We discuss the radius of convergence of perturbation theory show how nonlinear transformations can improve convergence. Section 6.3 contains the main results about the convergence of perturbation theory for Fourier space correlators and the non-convergence for real space correlators. Section 6.4 is dedicated to the construction of the 1D count-in-cell PDF, with details in Appendix 6.D, and a proof of the finite radius of convergence of perturbation theory. In section 6.5 we draw some qualitative conclusions about the existence and relevance of non-perturbative effects in real and Fourier space. We conclude in Section 6.6.

6.2 Exact and perturbative classical solutions to gravitational collapse

In this section we discuss perturbation theory of the classical equations of motion for LSS and its convergence properties. This discussion is logically separated from the discussion of statistical/quantum correlators, which we postpone to the following sections. In the following, we review some well-known exact solutions to gravitational collapse in 1D in the context of Newtonian cosmology (see subsection 3.3.1). We follow mostly the review part of [116]. One can think of 1D (from now on we drop the '+1' time dimension)

gravitational collapse as a more symmetric version of 3D collapse, in which the density field is only allowed to vary in one direction, say along x. It is thus the problem of the evolution of 2D-homogeneous and isotropic sheets of matter, with density contrast

$$\delta(x) = \frac{\rho(x)}{\bar{\rho}} - 1, \tag{6.2.1}$$

where x is just a number in this case and we omitted the time dependence. Moreover, we restrict to an Einstein-de Sitter spacetime background, for which

$$\bar{\rho}(a) = \bar{\rho}(a_i) \left(\frac{a_i}{a}\right)^3. \tag{6.2.2}$$

The equations of motion are obtained by imposing the symmetries of the fluid equations that are assumed to hold in Standard Perturbation Theory (SPT) (for a review see [29]). Except in section 6.3.3, we neglect Effective Field Theory (EFT) corrections [11], since introducing these terms should not change our results qualitatively, but considerably complicates the algebraic manipulations. To consistently neglect them, we exponentially damp the initial power spectrum, such that all fields can be thought of as smoothed fields. The equations of motion are then

$$\begin{aligned} \partial_\tau \delta + \theta &= -\nabla(\delta v)\,, \\ \partial_\tau \theta + \mathcal{H}\theta + 4\pi G a^2 \bar{\rho}\delta &= -\nabla(v\nabla v), \end{aligned} \tag{6.2.3}$$

where v is the velocity field, $\theta = \nabla v$, and τ and $\mathcal{H}$ are the conformal counterparts of the time coordinate and the Hubble rate, respectively. Here we have taken the gradient of the Euler equation without loss of generality, as in 1D there are no vector modes, and we used the Poisson equation to get rid of the Newtonian potential ϕ:

$$\Delta\phi = 4\pi G a^2 \bar{\rho}\delta. \tag{6.2.4}$$

As we review below, the Zel'dovich approximation is the exact solution to 1D gravitational collapse before shell-crossing, and its implications for the density field are straightforward. We show that the same solution holds for the evolution of the average density in cylindrical cells, which are effectively

1-dimensional cells. This analysis is very similar to spherical collapse. Given these exact solutions, we investigate the convergence of the perturbative solutions. We show that the perturbative solutions in real space have a finite radius of convergence. We comment on how non-linear transformations can provide non-perturbative improvements in the convergence.

6.2.1 Zel'dovich solution

The Zel'dovich solution [133] is exact before shell crossing in 1D [134, 135]. This follows from the fact that in 1D Newtonian Gravity, force is independent of distance. The equation for the gradient of the displacement field, giving the displacement of a fluid element from its initial position q (see 3.3.27 and the subsequent line), turns out to be linear

$$\nabla_q \left[\Psi''(q) + \mathcal{H}\Psi'(q)\right] = 4\pi G a^2 \bar{\rho} \nabla_q \Psi, \tag{6.2.5}$$

where primes denote derivatives with respect to conformal time, and we used

$$1 + \delta(x) = \int dq \delta_D[x - q - \Psi(q)] = \frac{1}{1 + \nabla_q \Psi}\bigg|_{x=q+\Psi(q)}, \tag{6.2.6}$$

which in Fourier space reads

$$\delta(k) = \int dq\, e^{-ikq} \left(e^{-ik\Psi(q)} - 1\right). \tag{6.2.7}$$

One can check that this definition of $\delta(x)$ indeed yields a solution to the Euler-Poisson system (6.2.3), (6.2.4) for $\Psi(q, a) = a/a_i \Psi(q, a_i)$, which solves (6.2.5). In particular, upon the identification

$$(\partial_\tau|_x + v\partial_x) = \partial_\tau|_q\,, \tag{6.2.8}$$

where we indicated what is kept fixed when performing the time-derivative, $\delta(q)$ satisfies the Lagrangian equation

$$\delta''(q) + \mathcal{H}\delta'(q) - 2\frac{\delta'^2(q)}{1 + \delta(q)} = 4\pi G\bar{\rho}\delta(q)(1 + \delta(q)). \tag{6.2.9}$$

Introducing the linear order density $\delta_L(q, \tau) \equiv -\nabla_q \Psi(q, \tau)$, we write

$$1 + \delta(x, \tau) = \frac{1}{1 - \delta_L(q, \tau)}. \tag{6.2.10}$$

A few comments are in order. First, observe that the analytic properties for the solution for δ in real space seem different from those in Fourier space. We come back to this issue below. Second, the Lagrangian equation (6.2.9) is actually also found for the evolution of the density in finite cells. We derive this in Appendix 6.A within Newtonian cosmology. Third, as expected, (6.2.10) breaks down for overdensities when the density blows up, which is precisely when shell-crossing occurs. On the other hand, this solution is well defined at all times for underdensities, whose density asymptotes to -1.

Perturbative solution

A more elaborate analysis of SPT for the 1D Euler-Poisson system (6.2.3), (6.2.4) was done in [116], in which they showed that the Fourier kernels [136] obtained from SPT are equivalent to the ones obtained by expanding the ZA solution (6.2.7). Here we are more modest, and just consider the 'Lagrangian' problem of the evolution of the density in a fluid element or cell (6.2.9). Formally, one can solve (6.2.9) (or its cosmological time equivalent, (6.A.9)) perturbatively, using a Green's function method

$$\delta = \delta_L + \int dt' G(t,t') \left[2\frac{\dot{\delta}^2}{1+\delta} + 4\pi G\bar{\rho}\delta^2 \right], \tag{6.2.11}$$

where we have selected the growing mode linear solution, and

$$D_t G(t,t') = \delta_D(t-t'), \tag{6.2.12}$$

for the linear differential operator D_t in (6.A.9). The perturbative solution is then obtained by iteratively plugging the lower order solutions into the nonlinear terms. In an Einstein-de Sitter universe, this leads to a power series in δ_L. However, since we already know the full solution (6.2.10), this has to coincide with a simple expansion of the exact solution in δ_L, leading to the following perturbative solution:

$$\delta_{PT}^{(n)} = \sum_{i=1}^{n} c_i \delta_L^i, \tag{6.2.13}$$

for some constants c_i, and $c_1 = 1$.

Convergence

Since the exact solution (6.2.10) has a nice analytic form, we can directly apply standard results from complex analysis about the convergence properties of the perturbative series [6]. In particular, the radius of convergence around the origin ($\delta_L = 0$) is given by the distance to the nearest pole, which is in our case is $\delta_L = 1$. For overdensities, this makes sense, as this is the point beyond which the exact solution breaks down as well. In other words, the series converges for overdensities all the way to $\delta = +\infty$. However, this radius of convergence also implies that, for underdensities, the series only converges up to $\delta_L = -1$, for which $\delta = -1/2$, whereas the exact solution sensibly extends all the way to $\delta = -1$.

One way to visualize the performance of perturbation theory is to plot $\delta_{PT}^{(n)}$ against the exact solution for all times. So we plot the following points

$$\left\{ \left(\delta(a), \delta_{PT}^{(n)}(a) \right) \| a \in [-\infty, \infty] \right\} \tag{6.2.14}$$

in Figure 6.1. The non-convergence beyond $\delta = -1/2$ is clearly visible.

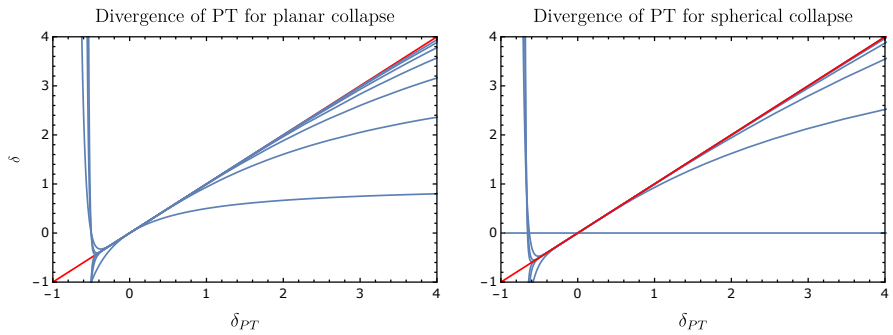


Figure 6.1: The parametric plots shows δ versus δ_{PT}. The blue lines show δ_{PT} from linear to 18-loop order in steps of three loops. The red reference line is the simple diagonal $\{\delta, \delta\}$. These plots show the divergence of perturbation theory beyond $\delta < -0.5$ for planar collapse (6.2.10) (left) and beyond $\delta < -0.684$ for spherical collapse (6.2.28) (right).

[6] This is an example of the idea put forth in footnote 5 of [38].

Nonlinear transformations and improved convergence

As noted above[7], *the radius of convergence of the perturbative expression is smaller than the radius for which the densities are physically well defined, which is the reason perturbation theory does not converge for all physically relevant densities.* One could ask if nonlinear transformations could fix this problem. Here we show that the answer is yes. One option is to choose an invertible, analytic function over the whole real axis, whose range is at least $(-1, \infty)$. Trivial examples are, on top of the linearizing transformation (6.2.10),

$$\delta(a) = e^{\lambda(a)} - 1 \quad \text{or} \quad \delta(a) = 2e^{\lambda(a)} - 2. \tag{6.2.15}$$

This corresponds to a nonlinear transformation of the perturbation parameter of the form $\lambda = -\log(1-\delta_L)$. Perturbation theory in λ now converges for all physically meaningful values, namely in the whole interval $\delta \in \{-1, \infty\}$.

Two comments are in order. First, this example shows that not all perturbation schemes are equivalent. In particular there can very well be non-linear transformations that substantially improve the convergence of perturbation theory. Second, the improvement of convergence above was only possible because we knew the full result and could therefore guess the correct nonlinear transformation. It is not clear whether in more complicated cases, such as the full 3D dynamics, this can be achieved. For empirical attempts in this direction see [137–142] for logarithmic and Gausianizing transformations, and [143, 144] for clipping procedures, in which large overdensities are taken out of the ensemble averages.

6.2.2 Spherical collapse

Let us now consider the collapse of a spherical overdensity in a spherically symmetric universe (see, e.g. [4, 145]). The density contrast is a function of radius R only:

$$\delta(R) = \frac{\rho(R)}{\bar{\rho}} - 1. \tag{6.2.16}$$

Throughout, we assume the background density is the EdS one, (6.2.2). We are interested in the evolution of the density inside a spherical cell. The

[7] Results in this subsection were obtained in collaboration with Gabriele Trevisan.

total mass inside the cell is

$$M = 4\pi \int_0^R dr r^2 \rho(r). \tag{6.2.17}$$

Then, by spherical symmetry, Gauss' law yields the following flux perpendicular to the surface,

$$4\pi G M = 4\pi R^2 \nabla_r \phi, \tag{6.2.18}$$

leading of course to the spherical collapse equation

$$\ddot{R} = -\frac{GM}{R^2}. \tag{6.2.19}$$

Rewriting equation (6.2.17) in terms of the average density contrast,

$$\delta_R = \frac{3}{R^3} \int_0^R dr r^2 \delta(r), \tag{6.2.20}$$

we can express R in terms of the average density

$$\delta_R = \frac{3M}{4\pi R^3 \bar{\rho}} - 1. \tag{6.2.21}$$

Plugging this into (6.2.19), we obtain the evolution equation for the average density in this spherical cell:

$$\ddot{\delta}_R + 2H\dot{\delta}_R - \frac{4}{3}\frac{\dot{\delta}_R^2}{1+\delta_R} = 4\pi G \bar{\rho} \delta_R (1+\delta_R). \tag{6.2.22}$$

Once again, one finds the same equation for infinitesimal volume elements in Lagrangian coordinates from the fluid equations in a spherically symmetric setup. Note that this equation, as the planar one, only depends on the density, there is no explicit mention of mass or scale. This is a consequence of the fact that the spherical collapse equation is symmetric under rescalings that leave M/R^3 - the density - fixed. Similarly, for planar collapse, rescalings that leave M/R - the 1D density, up to the Hubble expansion in the orthogonal directions - fixed, are a symmetry.

Exact solution

Despite its similarity to the planar case, the solution to (6.2.22) is only known in parametric form. Moreover, depending on whether the initial density perturbation is positive or negative, the form of the solution is slightly different. For overdensities, one can check that (6.2.19), and therefore (6.2.22) are solved by

$$R = A(1 - \cos\eta); \quad t = B(\eta - \sin\eta) + C, \tag{6.2.23}$$

provided

$$\frac{A^3}{B^2} = GM, \tag{6.2.24}$$

and $\eta \in [0, 2\pi]$, as can be seen from the expression for R. For underdensities, we find

$$R = A(\cosh\eta - 1); \quad t = B(\sinh\eta - \eta) + C, \tag{6.2.25}$$

with the same restriction on the constants A and B and $\eta \in [0, \infty]$ this time. The solutions are found by plugging this into (6.2.21), which for overdensities becomes

$$\delta = \frac{3M}{4\pi A^3(1 - \cos\eta)^3\bar{\rho}} - 1. \tag{6.2.26}$$

Note that this expression requires the time dependence of $\bar{\rho}$, which in an EdS universe is given by

$$\bar{\rho} = 3M_{pl}^2H^2 = \frac{1}{6\pi Gt^2}. \tag{6.2.27}$$

Hence,

$$\delta = \frac{9}{2}\frac{GMt^2}{A^3(1 - \cos\eta)^3} - 1 = \frac{9}{2}\frac{t^2}{B^2(1 - \cos\eta)^3} - 1. \tag{6.2.28}$$

To find $\delta(t)$, we need to invert the relation between t and η. This gives the solution for δ as a function of two constants, B and C, as it should for a second order differential equation. For a more familiar interpretation of these constants in terms of the growing and decaying mode, we need to

restrict ourselves to the small density regime. As we show in appendix 6.B, this is given by

$$\delta_L = \frac{3}{10}\left(\frac{9}{2}\right)^{1/3}\left(\frac{t}{B}\right)^{2/3} + \frac{2C}{t}. \tag{6.2.29}$$

These are indeed the familiar growing and decaying modes, parametrized by B and C respectively. This makes manifest that for adiabatic initial conditions, we should set $C = 0$. At the same time, this clarifies the range of validity of the growing mode solution. The initial conditions are set by

$$\delta_i = \frac{3}{10}\left(\frac{9}{2}\right)^{1/3}\left(\frac{t_i}{B}\right)^{2/3}. \tag{6.2.30}$$

As argued before, the solution for overdensities only makes sense up to $\eta = 2\pi$ - the point of collapse. This means

$$2\pi = \frac{t}{B} = \left(\frac{3}{10}\right)^{3/2}\left(\frac{9}{2}\right)^{1/2}\frac{t}{t_i}\delta_i^{3/2}. \tag{6.2.31}$$

In other words, the solution is well defined up to the present for initial conditions that satisfy

$$\delta_i < \delta_c\left(\frac{t_i}{t_0}\right)^{2/3} = \delta_c\frac{a_i}{a_0}, \tag{6.2.32}$$

where

$$\delta_c = \frac{10}{3}\left(\frac{9}{2}\right)^{1/3}(2\pi)^{2/3} \approx 1.686. \tag{6.2.33}$$

One can check that, up to a minus sign, the growing and decaying mode are the same for underdensities. In the underdense case, the fully non-linear, growing mode solution is well defined for all initial conditions (larger than -1) and all times.

Perturbative solution and convergence

The perturbative solution to the equation of motion still leads to a series expansion in δ_L, which we can obtain from the exact solution as follows. We

are looking for a solution of the form

$$\delta(t) = \sum_i^n c_i \delta_L^i + \mathcal{O}(\delta_L^{n+1}) = \sum_i^n \tilde{c}_i \left(\frac{t}{B}\right)^{2i/3} + \mathcal{O}\left(\left(\frac{t}{B}\right)^{2/3(n+1)}\right). \tag{6.2.34}$$

We chose to keep the parameter B explicitly, so that the both the left hand side (see (6.2.28)) and the right are functions of η only. Expanding in η and matching order by order then allows us to solve for the $\tilde{c}_i$, which are directly related to the c_i.

Once again, we can study the convergence of this series, previously also discussed in [38, 146]. Observe that the series breaks down for overdensities when the density blows up, which is precisely when $\delta_L = \delta_c$. From the Green's function approach, it is clear that the perturbative solution has to be the same for over- and underdensities; the only thing that distinguishes between them is whether δ_L is positive or negative. This can of course be checked explicitly applying the above logic to the underdense solution. Thus, the same rules of complex analysis tell us that the series for underdensities only converges up to the point where $\delta_L^u = -\delta_c$, which by definition corresponds to

$$\frac{t_c}{B} = 2\pi. \tag{6.2.35}$$

The critical η parameter is then found from the relation between t and η for underdensities:

$$\sinh \eta_c - \eta_c = \frac{t_c}{B} \quad \Longrightarrow \quad \eta_c \approx 2.915. \tag{6.2.36}$$

Plugging this back into the full solution, we find that perturbation theory only converges in the range $-0.684 < \delta < +\infty$. Again, in Figure 6.1 we plot δ_{PT} versus δ for spherical collapse to visualize these statements.

Improved convergence

Similar to the 1D case, there are ways to improve the radius of convergence of perturbation theory, knowing the full solution[8]. In this case, a neat example

[8] Results in this subsection were obtained in collaboration with Gabriele Trevisan.

is found from observing that the underdensity solution is obtained from the overdensity solution by rotating in the complex plane $\eta \to i\eta$. This immediately tells us that the perturbative expansion of δ in terms of η converges to $|\eta| = 2\pi$, as this is the radius of convergence for the overdense solution. This is a much larger value than η_c we found above. In fact, the final underdensity at this value of η is $\delta = -0.984$. So, once again, a non-linear transformation of the expansion parameter from δ to η does enlarge the physical radius of convergence of the theory (to the range $-0.984 < \delta < +\infty$). The analogous non-linear transformation in realistic 3D cases can also be searched for heuristically [137, 138, 143, 144].

6.3 (non-)Convergence of PT for 1D correlators

In this and the following sections, we move away from the discussion of the classical solutions of the equations of motion and delve into the computation of stochastic/quantum correlators. There are already various studies on the reach of perturbation theory for large scale structures in the literature. The relevance of halos in this context was stressed in [118–121]. The reach of PT was further analyzed in [122–124], and the relevance of shell crossing was studied in [49, 125, 126]. Finally, a generic perturbative expression including EFT corrections for the power spectrum was tested numerically in [127]. In this section we analytically test these ideas by studying the convergence of PT for 1D correlators. Our main finding here is that, for ΛCDM-like initial conditions, SPT [29] converges to the correct power spectrum and bispectrum both for Gaussian and non-Gaussian initial conditions. The convergence of SPT for the power spectrum was to a large extent already established by McQuinn and White in [116]. In subsection 6.3.2, we review their derivation and extend it marginally by rigorously justifying their final step, namely that one can safely exchange the integral over initial positions with the infinite perturbative sum. While this seemingly minor technical assumption is justified for the power spectrum of LCDM, it is actually invalid in a few relevant cases. In fact, for scaling universes, $P(k) = Ak^n$, SPT diverges for $-1 < n < 0$, while it converges for $n > 0$ [131] (see subsection 6.3.3). More importantly, the exchange of sum and integral is not allowed for real space correlators and leads to the non-convergence of SPT, a new result which we discuss in the section 6.3.6.

6.3.1 Prerequisites

We start by collecting the ingredients necessary for the derivations below. The key mathematical observations that lead to our results are explained first. Then we define the Zel'dovich approximation (ZA) for correlators and recall some properties of the initial conditions of our universe that are important for what follows.

Mathematical prerequisites

We will see that the ZA allows us to write all observable as some integral, of the form

$$\hat{O}(k,\sigma^2) = \int dq\, f(q,k,\sigma^2), \tag{6.3.1}$$

where σ^2 is a dimensionless parameter representing the size of the linear power spectrum. We wish to answer the question whether perturbation theory in this parameter resums to the ZA result. This relies on two steps:

- Can we write f as a convergent power series in σ^2?
- Can we interchange the order of integral and sum?

To show the subtleties of this second step, consider the Fourier transform of a Gaussian:

$$\hat{O}(k,\sigma^2) = \int dq\, e^{iqk} \frac{1}{\sqrt{2\pi}} e^{-\frac{q^2\sigma^2}{2}} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{k^2}{2\sigma^2}}. \tag{6.3.2}$$

Clearly, the final expression for $\hat{O}$ is non-analytic around $\sigma^2 = 0$, whereas the integrand in the middle step is analytic (in fact it can be extended to an entire function). Apparently,

$$\int dq \sum_i f_i(q,k) \left(\sigma^2\right)^i \neq \sum_i \int dq\, f_i(q,k) \left(\sigma^2\right)^i, \quad \text{for} \quad f(q,k,\sigma^2) = e^{iqk} \frac{1}{\sqrt{2\pi}} e^{-\frac{q^2\sigma^2}{2}}, \tag{6.3.3}$$

where f_i are the series coefficients. In fact, the integrals over the f_i, which are simple power laws in this case, only make sense as a distribution. The

resulting expression on the right hand side is then a sum of derivatives of the Dirac-delta distribution

$$\sum_i \int dq\, f_i(q,k) \left(\sigma^2\right)^i = \sum_i \delta_D^{(2i)}(k) c_i \sigma^{2i}, \tag{6.3.4}$$

for some c_i. One can now check that the left and right hand side are *not* equal as a distribution acting on test functions φ, since

$$\begin{cases} \int dk\, \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{k^2}{2\sigma^2}} \varphi(k) = \int dk\, \sum_i \delta_D^{(2i)}(k) c_i \sigma^{2i} \varphi(k) & \text{for analytic } \varphi \\ \int dk\, \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{k^2}{2\sigma^2}} \varphi(k) \neq \int dk\, \sum_i \delta_D^{(2i)}(k) c_i \sigma^{2i} \varphi(k) & \text{for non-analytic } \varphi, \end{cases}$$

Thus interchanging sum and integral is not allowed, even if interpret functions as distributions. This is the crucial question in all of our derivations below. The mathematical criterion to decide whether this is possible is Fubini-Tonelli's theorem, which, if we interpret the sum as an integral over a sum of block functions, states:

Theorem 1. *(Fubini-Tonelli:) If* $\int dq \sum_i |f_i(q)| < \infty$*, then* $\int dq \sum_i f_i(q) = \sum_i \int dq f_i(q)$*.*

The assumption of the theorem fails to hold in the Gaussian Fourier transform example because the integral of the sum of the absolute values of the PT terms is infinite. In the computation of LSS correlators, the theorem can be used, with some minor subtlety, for *Fourier space* observables, as discussed below. However, the exchange of sum and integral is not allowed for the Fourier transform of the power spectrum, i.e. the correlation function. To prove the non-convergence of PT for the correlation function, we assume the following.

- *Assumption:* If the Taylor series of some integrand f does not converge for some range of the integral,

$$\sum_i f_i(q,k) \left(\sigma^2\right)^i \neq f(q,k,\sigma^2), \text{ for some } q, \tag{6.3.5}$$

and

$$\left(\frac{\partial}{\partial\sigma^2}\right)^i \Big|_{\sigma^2=0} \int dq\, f(q,k,\sigma^2) = \int dq\, \left(\frac{\partial}{\partial\sigma^2}\right)^i \Big|_{\sigma^2=0} f(q,k,\sigma^2) \tag{6.3.6}$$

is well defined, then the series expansion for $\hat{O} = \int dq\, f$ does not converge generically.

To argue for this assumption, let us adopt the shorthand notation $\sum_{i=0}^{N} \hat{O}_i = \hat{O}_N$, and $\sum_{i=0}^{N} f_i = f_N$. Then, according to our assumptions, and since a finite sum can always be interchanged with an integral, $\hat{O}_N = \int dq\, f_N(q)$. Then, in order for the $\hat{O}_N$ to converge to $\hat{O}$, we need $\lim_{N\to\infty} \int dq\, f_N(q) = \int dq\, f(q)$. This means that if the series diverges in an uncontrollable way for a range of q, for instance because of the occurrence of an (complex) essential singularity in the domain of integration, the integral generically also diverges. For a general review of asymptotic series see [147].

Cosmological prerequisites: ZA and SPT for 1D correlators

To define the statistics of the displacement $\psi(q)$ it is easiest to invert the Zel'dovich relation

$$\delta(x) = \int d^d q \delta_D \left(x - q - \psi(q)\right) , \tag{6.3.7}$$

and expand it to linear order, finding

$$\psi(q) = \int_k e^{ikq} \frac{ik}{k^2} \delta_L(k) \,, \tag{6.3.8}$$

where δ_L is the linear order density and the integral is over $dk/(2\pi)$. Therefore

$$\langle \psi(q)\psi(0)\rangle = \int_k e^{ikq} \frac{P_L(k)}{k^2} \,. \tag{6.3.9}$$

The variance of the 1D Zel'dovich displacement, which is exact before shell crossing, can then be computed

$$\sigma^2(q) = \langle [\psi(q) - \psi(0)]^2 \rangle = \int_0^\infty \frac{dk}{\pi} \frac{2P_L(k)}{k^2} \left[1 - \cos\left(kq\right)\right] \tag{6.3.10}$$

$$= \sigma_\infty^2 - \sigma_q^2(q) \,, \tag{6.3.11}$$

where we defined

$$\sigma_\infty^2 \equiv \int_0^\infty \frac{dk}{\pi} \frac{2P_L(k)}{k^2} > 0 \,, \tag{6.3.12}$$

$$\sigma_q^2(q) \equiv \int_0^\infty \frac{dk}{\pi} \frac{2P_L(k)}{k^2} \cos\left(kq\right) = \int_{-\infty}^\infty \frac{dk}{\pi} \frac{P_L(k)}{k^2} e^{iqk} \,, \tag{6.3.13}$$

and assumed that

$$\sigma_q^2(q=0) = \sigma_\infty^2 < \infty\,. \tag{6.3.14}$$

This property holds in a ΛCDM universe where both the UV and IR part of the integral converge[9]. As we show in subsection 6.3.2 by marginally extending the proof of [116], SPT converges to the ZA power spectrum. On the other hand, there exist (less realistic) cases in which $\sigma(q)^2$ is unbounded, as for example in scaling universes, $P_L(k) = Ak^n$ for some n. In subsection 6.3.3 we discuss the finding of S. Foreman [131] that SPT for the power spectrum can both converge or diverge depending on n. Notice that $P(k) = P(-k)$ and so $\sigma^2(q)$ is real, as evident from (6.3.10). Also, despite its name, σ_q^2 does *not* need to be positive, unlike σ_∞^2.

A standard result in Fourier analysis guarantees that if a function is square integrable, the Fourier transform vanishes at least as fast as q^{-1} for large q. If the function is also continuous, then the Fourier transform vanishes as q^{-2}. For a ΛCDM-like universe we have that $P(k)/k^2$ is a continuous, square integrable function and therefore σ_q^2 vanishes for large q as q^{-2}, justifying the name of σ_∞^2.

Throughout the Chapter, we support our analytical results with some plots of the observables in question. For simplicity, and in order to be least sensitive to multistreaming, we use the following linear power spectrum as initial condition in all our plots

$$P_L(k) = \frac{4\times 10^4}{\pi} k^2 e^{-\frac{k^2}{0.05^2}}, \tag{6.3.15}$$

where k and the power spectrum are in units of inverse megaparsec. Its corresponding dimensionless variance for the average density in cells of size R is given by

$$\sigma^2_{R,\mathrm{lin}}(R) = \int_0^\infty \frac{1}{\pi}\left(\frac{2}{kR}\right)^2 \sin^2\left(\frac{kR}{2}\right) P_L(k) = \frac{2\times 10^4}{\pi^{3/2}} \frac{1-e^{-\frac{R^2}{1600}}}{R^2} = \frac{\sigma^2(R)}{R^2}. \tag{6.3.16}$$

To get an idea of the size of the perturbation parameter in this work, we plot $\sigma^2_{R,\mathrm{lin}}$ and the dimensionless power spectrum, $kP(k)/2\pi$ in 1D, in figure 6.2.

[9] For small k, $kP_L = k^3 P_{3D} \sim k^4$, and so $P_L \sim k^3$

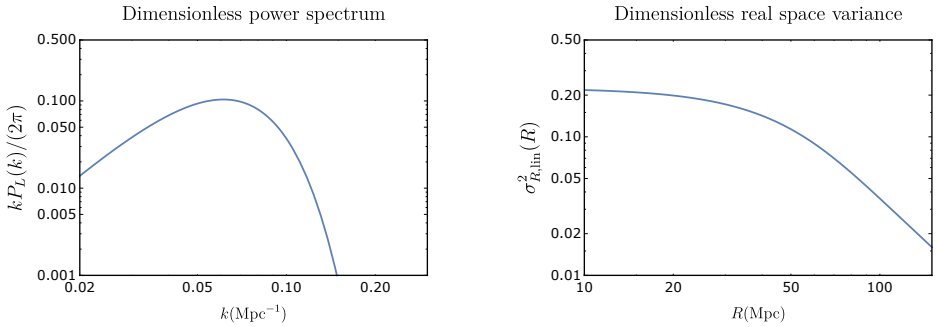


Figure 6.2: The plot shows the initial conditions that are used throughout this Chapter in two forms. In particular it shows that the dimensionless variance (6.3.16) is significantly less than unity on all scales.

6.3.2 Convergence of SPT for the power spectrum for ΛCDM-like universes

Let us start our discussion considering a ΛCDM-like 1D power spectrum as discussed above. Recall that the ZA expression for the power spectrum is [116, 148]

$$P_{ZA}(k) = \int_{-\infty}^{\infty} dq e^{-ikq} \left[e^{-k^2\sigma^2(q)/2} - 1 \right]. \tag{6.3.17}$$

Using the decomposition $\sigma^2(q) = \sigma^2_\infty - \sigma^2_q(q)$, this can be rewritten as

$$P_{ZA}(k) = \int_{-\infty}^{\infty} dq e^{-ikq} \left[e^{-k^2\sigma^2_\infty/2} \left(e^{k^2\sigma^2_q(q)/2} - 1 \right) + \left(e^{-k^2\sigma^2_\infty/2} - 1 \right) \right]. \tag{6.3.18}$$

The Fourier transform of the last term in brackets should be interpreted as a distribution, in which case we get

$$\int_{-\infty}^{\infty} dq e^{-ikq} \left(e^{-k^2\sigma^2_\infty/2} - 1 \right) = \left(e^{-k^2\sigma^2_\infty/2} - 1 \right) \delta_D(k) = 0. \tag{6.3.19}$$

Thus we are left with

$$P_{ZA}(k) = \int_{-\infty}^{\infty} dq e^{-ikq} \left[e^{-k^2\sigma^2_\infty/2} \left(e^{k^2\sigma^2_q(q)/2} - 1 \right) \right]. \tag{6.3.20}$$

For given q, the remaining term in brackets can be written as its Taylor series in σ^2, where σ_∞ and σ_q count at the same order:

$$e^{-k^2\sigma_\infty^2/2}\left(e^{k^2\sigma_q^2(q)/2}-1\right)=\sum_{n=1}^{\infty}\left[\sum_{j=1}^{n}\frac{(k^2\sigma_q^2/2)^j}{j!}\frac{(-k^2\sigma_\infty^2/2)^{n-j}}{(n-j)!}\right]. \quad (6.3.21)$$

In order to use Fubini-Tonelli, observe that

$$\begin{aligned}\sum_{n=1}^{\infty}\left|\sum_{j=1}^{n}\frac{(k^2\sigma_q^2/2)^j}{j!}\frac{(-k^2\sigma_\infty^2/2)^{n-j}}{(n-j)!}\right| &\leq \sum_{n=1}^{\infty}\sum_{j=1}^{n}\frac{|k^2\sigma_q^2/2|^j}{j!}\frac{|k^2\sigma_\infty^2/2|^{n-j}}{(n-j)!}\\ &= e^{k^2\sigma_\infty^2/2}\left(e^{k^2|\sigma_q^2(q)|/2}-1\right), \quad (6.3.22)\end{aligned}$$

where we have used that $\sigma_\infty^2>0$, but σ_q^2 is not necessarily positive. Thus,

$$\begin{aligned}\int_{-\infty}^{\infty}dq\sum_{n=1}^{\infty}\left|\sum_{j=1}^{n}e^{-ikq}\frac{(k^2\sigma_q^2/2)^j}{j!}\frac{(-k^2\sigma_\infty^2/2)^{n-j}}{(n-j)!}\right| &\leq \int_{-\infty}^{\infty}dqe^{k^2\sigma_\infty^2/2}\left(e^{k^2|\sigma_q^2(q)|/2}-1\right)\\ &= e^{k^2\sigma_\infty^2/2}\int_0^{\infty}dq\left(e^{k^2|\sigma_q^2(q)|/2}-1\right)\\ &<\infty\,,\end{aligned}$$

where in the last step we used that σ_q^2 goes to zero at least as q^{-2} as $q\to\infty$. We conclude that Fubini-Tonelli can indeed be applied. Thus we find that SPT converges to the ZA expression,

$$\begin{aligned}P_{ZA}(k) &= \int_{-\infty}^{\infty}dqe^{-ikq}\sum_{n=1}^{\infty}\left[\sum_{j=1}^{n}\frac{(k^2\sigma_q^2/2)^j}{j!}\frac{(-k^2\sigma_\infty^2/2)^{n-j}}{(n-j)!}\right]\\ &= \sum_{n=1}^{\infty}\int_{-\infty}^{\infty}dqe^{-ikq}\left[\sum_{j=1}^{n}\frac{(k^2\sigma_q^2/2)^j}{j!}\frac{(-k^2\sigma_\infty^2/2)^{n-j}}{(n-j)!}\right]\\ &= \sum_{n=1}^{\infty}P_n(k)=P_{SPT}(k)\,. \quad (6.3.23)\end{aligned}$$

To confirm this analytic result we can plot the power spectrum for initial condition (6.3.15). The result is shown in Figure 6.3.

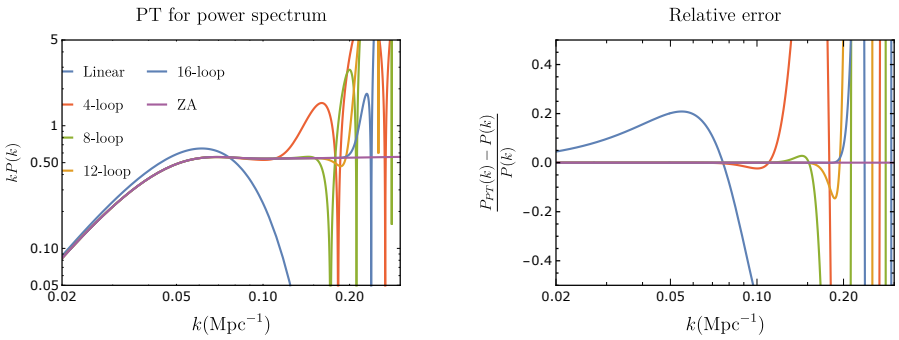


Figure 6.3: The plot shows the ZA power spectrum (purple continuous line) and some of its SPT approximations (full plot on left and relative error on right). For given k, one can reach arbitrary precision by going to high enough orders in perturbation theory, which is the hallmark of a convergent series. We use (6.3.15) as initial condition.

6.3.3 Divergence of dimensionally regulated SPT for the power spectrum for scaling universes

Let us consider now scaling universes [4], $P_L(k) = Ak^n$, for some $n > -1$. These avoid the conclusion of the previous subsection where we assumed $\sigma^2 < \infty$. The easiest case is actually $n = 0$, so we discuss it first and then move to arbitrary n. For $n = 0$, the variance of the displacement can be computed analytically from (6.3.10) to be $\sigma^2(q) = A|q|$. The ZA power spectrum follows from (6.3.17):

$$P_{ZA}(k) = \frac{2A}{4 + k^2A^2}\,. \tag{6.3.24}$$

SPT is an expansion in A and therefore has a finite radius of convergence $|kA| = 1$. For some given A, SPT converges on large scales but diverges on short scales. Moving on to arbitrary n, we are forced to regulate the infinities. We quote the result of S. Foreman [131], who used dimensional regularization to obtain

$$P_{L\text{-loop}}(k) = \tag{6.3.25}$$
$$\frac{\pi^{L/2}}{(L+1)!}\frac{\Gamma(\frac{1}{2}(n-1))^{L+1}\Gamma(\frac{1}{2}(L-Ln-n+2))}{\Gamma(1-\frac{n}{2})^{L+1}\Gamma(\frac{1}{2}(-L+Ln+n-1))}\left(\frac{k}{k_{NL}}\right)^{L(1+n)}P_L(k)\,,$$

where L is the number of loops and $A = 2\pi k_{NL}^{-1-n}$. Note that this result refers to the cut-off independent contribution, a.k.a. the "finite part" of the loop correction. To study the convergence of this series, let us focus on the Taylor coefficients a_L in the loop expansion

$$P_{L\text{-loop}}(k) = a_L \left(\frac{k}{k_{NL}}\right)^{L(1+n)} P_L(k)\,. \tag{6.3.26}$$

For large L one finds

$$a_L \simeq \left(\frac{c_n}{L^n}\right)^L \tag{6.3.27}$$

where c_n is some n-dependent constant. We hence conclude that for $-1 < n < 0$ the series diverges for any k. In this case the a_L's have indeed the typical growth encountered in divergent asymptotic expansions, which stop approaching the exact result at some k-dependent order $L_{\rm opt}$. On the other hand, the series converges for $n > 0$ for any k, since a_L decreases rapidly with L. The boundary case $n = 0$ was discussed above and has somewhat hybrid behavior converging only for some range of scales.

Note that the dimensionally regulated results above show that the divergence of PT we discuss in this Chapter is indeed unrelated to the loop-divergences that are renormalized in the EFT of LSS [43]. We further discuss this point in 6.5.3.

6.3.4 Convergence of SPT for the Bispectrum

The bispectrum in 1D can be intuitively expressed as

$$B(k_1, k_2) = V^{-1}\langle|\delta(k_1)\delta(k_2)\delta(-k_1-k_2)|\rangle, \tag{6.3.28}$$

where V is the 1D spatial volume. In Appendix 6.C, we derive this expression more formally from symmetries. Plugging in the ZA expressions for the density, we obtain

$$\begin{aligned} B(k_1, k_2) = V^{-1}\langle \int dq_{123} e^{-ik_1 q_1} e^{-ik_2 q_2} e^{i(k_1+k_2)q_3} \\ \times \left(e^{-ik_1\Psi_1} - 1\right)\left(e^{-ik_2\Psi_2} - 1\right)\left(e^{i(k_1+k_2)\Psi_3} - 1\right)\rangle, \end{aligned} \tag{6.3.29}$$

where $\Psi_i \equiv \Psi(q_i)$. To guide our intuition, we can already observe that this expression goes to zero as any of the displacements goes to zero. Working out the terms in brackets, we find

$$\begin{aligned} B(k_1,k_2) = V^{-1}\int dq_{123} e^{ik_1(q_3-q_1)} e^{ik_2(q_3-q_2)} \times \\ \Big[\langle e^{ik_1(\Psi_3-\Psi_1)} e^{ik_2(\Psi_3-\Psi_2)}\rangle - \langle e^{-ik_1\Psi_1} e^{-ik_2\Psi_2}\rangle \\ - \langle e^{-ik_1(\Psi_1-\Psi_3)} e^{-ik_2\Psi_3}\rangle + \langle e^{-ik_1\Psi_1}\rangle \\ - \langle e^{-ik_2(\Psi_2-\Psi_3)} e^{-ik_1\Psi_3}\rangle + \langle e^{-ik_2\Psi_2}\rangle \\ - \langle e^{-i(k_1+k_2)\Psi_3}\rangle - 1 \Big]. \end{aligned} \tag{6.3.30}$$

Now it's just a matter of correlating. For this we use the following observations. First, we use the cumulant expansion theorem, which is easily proved in this context. Namely, all expectation values are of the form

$$\langle e^{i\vec{\Psi}\cdot\vec{a}}\rangle = \int d\Psi_1 d\Psi_2 d\Psi_3 e^{i\vec{\Psi}\cdot\vec{a}} \frac{1}{\sqrt{\det 2\pi C}} e^{\frac{1}{2}\vec{\Psi}^T C^{-1}\vec{\Psi}}, \tag{6.3.31}$$

where $\vec{\Psi}^T = (\Psi_1, \Psi_2, \Psi_3)$, $\vec{a}$ is some Ψ-independent vector, and C the multivariate Gaussian correlation matrix. Completing the square in the exponent then allows us to evaluate the Gaussian integral, and we are left with

$$\langle e^{i\vec{\Psi}\cdot\vec{a}}\rangle = e^{-\frac{1}{2}\vec{a}^T C\vec{a}}. \tag{6.3.32}$$

Finally, we observe that

$$\vec{a}^T C\vec{a} = \langle \Big(\vec{a}\cdot\vec{\Psi}\Big)^2\rangle, \tag{6.3.33}$$

proving the cumulant expansion theorem in this case. Evaluating the two point correlations, we write

$$2\langle\Psi(q_1)\Psi(q_2)\rangle = \sigma_q^2(q_1-q_2) = -\sigma^2(q_1-q_2) + \sigma_\infty^2, \tag{6.3.34}$$

with $\sigma_\infty^2 = 2\langle\Psi(q_1)\Psi(q_1)\rangle$. Note that this means the three expectation values in (6.3.30) that only depend on a single coordinate evaluate to a constant.

They can therefore be taken out of the integral and the remaining integral can then be computed using a change of coordinates to yield (up to potential factors of 2π)

$$V^{-1}\int dq_{123}e^{ik_1(q_3-q_1)}e^{ik_2(q_3-q_2)} = V^{-1}\delta_D(0)\delta_D(k_1)\delta_D(k_2) = \delta_D(k_1)\delta_D(k_2). \tag{6.3.35}$$

Because of these δ-functions, we can set the momenta of the single-coordinate-expectation-values to zero at the start, and replace the full integral with unity. Using similar changes of coordinates, and using the same trick we used in (6.3.18) and (6.3.19), we finally find the following expression for the bispectrum

$$\begin{aligned}B(k_1,k_2) = V^{-1}\int dq_{123}e^{ik_1(q_3-q_1)}e^{ik_2(q_3-q_2)}\times\\ \Big[e^{-\frac{1}{2}\left[k_1^2\sigma^2(q_1-q_3)+k_2^2\sigma^2(q_2-q_3)+2k_1k_2\left(\sigma_\infty^2-\sigma_q^2(q_1-q_3)-\sigma_q^2(q_2-q_3)+\sigma_q^2(q_1-q_2)\right)\right]}\\ - e^{-\frac{1}{2}\left[k_1^2\sigma_\infty^2+k_2^2\sigma_\infty^2+2k_1k_2\sigma_q^2(q_1-q_2)\right]} - e^{-\frac{1}{2}\left[k_1^2\sigma^2(q_1-q_3)+k_2^2\sigma_\infty^2-2k_1k_2\left(\sigma_q^2(q_1-q_3)-\sigma_\infty^2\right)\right]}\\ - e^{-\frac{1}{2}\left[k_2^2\sigma^2(q_2-q_3)+k_1^2\sigma_\infty^2-2k_1k_2\left(\sigma_q^2(q_2-q_3)-\sigma_\infty^2\right)\right]}+2\Big].\end{aligned} \tag{6.3.36}$$

Changing coordinates to $\tilde{q}_1 = q_1 - q_3$, $\tilde{q}_2 = q_2 - q_3$, $\tilde{q}_3 = q_3$, and dropping the tildes (note that this transformation has unit determinant), we obtain

$$\begin{aligned}B(k_1,k_2) = \int dq_{12}e^{-ik_1q_1}e^{-ik_2q_2}\times\\ \Big[e^{-\frac{1}{2}\left[k_1^2\sigma^2(q_1)+k_2^2\sigma^2(q_2)+2k_1k_2\left(\sigma_\infty^2-\sigma_q^2(q_1)-\sigma_q^2(q_2)+\sigma_q^2(q_1-q_2)\right)\right]}\\ - e^{-\frac{1}{2}\left[k_1^2\sigma_\infty^2+k_2^2\sigma_\infty^2+2k_1k_2\sigma_q^2(q_1-q_2)\right]} - e^{-\frac{1}{2}\left[k_1^2\sigma^2(q_1)+k_2^2\sigma_\infty^2-2k_1k_2\left(\sigma_q^2(q_1)-\sigma_\infty^2\right)\right]}\\ - e^{-\frac{1}{2}\left[k_2^2\sigma^2(q_2)+k_1^2\sigma_\infty^2-2k_1k_2\left(\sigma_q^2(q_2)-\sigma_\infty^2\right)\right]}+2\Big].\end{aligned} \tag{6.3.37}$$

Let us first focus on the constant terms in the exponents. When one is interested in its contribution to the bispectrum, one is free to replace the

constant 2 with any function of k_1 and k_2 that equals unity when the k_i are zero. In this case it is convenient to replace it with

$$2 \to e^{k_1 k_2 \sigma_\infty^2} + e^{\frac{1}{2}(k_1^2+k_2^2)\sigma_\infty^2}. \tag{6.3.38}$$

Schematically, the perturbative expansion of the integrand looks like

$$\sum_n \frac{1}{n!} \left(\sum_i (f_i(q) + f_{i,\infty})^n \right), \tag{6.3.39}$$

where $\lim_{q\to\infty} q^2 f_i(q) = 0$, and $\sum_i f_{i,\infty} = 0$ because of the δ-function replacement above. The sum over i denotes the five terms of the integrand. In order to use Fubini-Tonelli, we wish to bound

$$\sum_n \frac{1}{n!} \left| \sum_i^{i_{max}} (f_i(q) + f_{i,\infty})^n \right|. \tag{6.3.40}$$

Now, for q large, by definition $f_i(q) < f_{i,\infty}$. Let $C = \max_i\{|f_{i,\infty}|\}$, and $g(q) = \max_i\{|f_i(q)|\}$. Then we can approximate the above sum as

$$\sum_n \frac{1}{n!} \left| \sum_i^{i_{max}} (f_i(q) + f_{i,\infty})^n \right| \leq g(q) \sum_n \frac{1}{n!} i_{max} 2^n C^{n-1} = g(q) i_{max} \frac{1}{C} e^{2C}. \tag{6.3.41}$$

Since this goes to zero faster than q^2 as $q_i \to \infty$, we conclude that

$$\int dq_{12} \sum_n \frac{1}{n!} \left| b_n(q_1, q_2) \right| < \infty, \tag{6.3.42}$$

where b_n is the n-th order term in the expansion of the integrand. Since the absolute value of the oscillating factors is obviously also bounded, this argument holds for the bispectrum as well, such that we conclude that Fubini-Tonelli indeed applies.

In conclusion, we can exchange the sum with the integral in (6.3.37) and this shows that the SPT bispectrum converges to the ZA bispectrum, in analogy with what happens for the power spectrum.

6.3.5 Convergence of SPT for the Power Spectrum for NG initial conditions

Here we are interested in the convergence of the power spectrum for NG initial conditions. Of course, the answer to this question depends on the type of initial conditions. We specify the assumptions we make below, which hold for perturbative non-Gaussianity. Using the cumulant expansion theorem as in equation (2.27) of [116], we find an expression of the form

$$P^{NG}(k) = \int_{-\infty}^{\infty} dq e^{-ikq} \left[e^{-k^2\sigma^2(q)/2 + ik^3\sigma_3(q)/3! - k^4\sigma_4(q)/4! + \dots} - 1 \right] , \quad (6.3.43)$$

where $\sigma_{3,4,\dots}$ characterize the type of primordial non-Gaussianity. Let us now assume the exponent can be decomposed as $f_{k,\infty} + \sum_i f_{k,i}(q)$, such that $\lim_{q\to\infty} f_{k,i}(q) = 0$. Once again, we can rewrite the 1 in the integrand as $e^{f_{k,\infty}}$. Then

$$P^{NG}(k) = \int_{-\infty}^{\infty} dq e^{-ikq} e^{f_{k,\infty}} \left[e^{\sum_i f_{k,i}(q)} - 1 \right] . \quad (6.3.44)$$

One has to be a bit careful about the expansion, since clearly the $f_{k,i}$ count at different orders in PT. However, at every order in PT we get some sum of terms, whose absolute value is always smaller than the sum of the absolute values of the individual terms. Resumming these absolute values is easy. If

$$P^{NG}(k) = \int_{-\infty}^{\infty} dq \sum_n p_n^{NG}(k,q), \quad (6.3.45)$$

we get

$$\int_{-\infty}^{\infty} dq \sum_n |p_n^{NG}(k,q)| \leq e^{|f_{k,\infty}|} \int_{-\infty}^{\infty} dq \left[e^{\sum_i |f_{k,i}(q)|} - 1 \right] . \quad (6.3.46)$$

From this we easily conclude the condition for convergence of PT: if all $f_{k,i}(q)$ satisfy $\lim_{q\to\infty} q^2 f_{k,i}(q) = 0$, then the expression above converges. In this case, the Fubini-Tonelli theorem applies and SPT converges to ZA, even for non-Gaussian initial conditions.

6.3.6 Non-convergence of SPT for real space correlation function

The two-point correlation function is obtained by Fourier transforming the power spectrum [149],

$$\xi(r) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{ikr} P(k). \tag{6.3.47}$$

Applying this to (6.3.20), we just need to evaluate Gaussian integrals, and find

$$1 + \xi(r) = \int_{-\infty}^{\infty} dq \frac{1}{\sqrt{2\pi\sigma^2(q)}} e^{-\frac{(r-q)^2}{2\sigma^2(q)}}. \tag{6.3.48}$$

Obviously, even if we forget about the square root factors, the Taylor expansion of the integrand around $\sigma^2(q) = 0$ does not converge to the right result. Hence, a convergence proof such as the one above does not exist. In fact, in the following, we argue that SPT does not converge to ZA. To that end, we show that this expression satisfies the conditions (6.3.5),(6.3.6). More precisely, we show that upon a change of variables, the integrand contains an essential singularity in the domain of integration. By defining $q = \sigma_\infty x + r$, we can write (6.3.48) as[10]

$$1 + \xi(r) = \int \frac{dx}{\sqrt{2\pi}} \frac{1}{\sqrt{\sigma^2(\sigma_\infty x + r)/\sigma_\infty^2}} e^{-\frac{x^2}{2} \frac{1}{\sigma^2(\sigma_\infty x + r)/\sigma_\infty^2}}. \tag{6.3.49}$$

Note that since $\sigma^2(q)$ and σ_∞^2 are of the same order in our perturbation parameter, we only need to expand with respect to the σ_∞^2 in the argument. More explicitly, let's write

$$\sigma^2(q)/\sigma_\infty^2 = 1 - \sigma_q^2/\sigma_\infty^2 = 1 - f(q), \tag{6.3.50}$$

[10] One might worry about the validity of this change of coordinates when $\sigma_\infty = 0$ (which is in fact the point we wish to perturb around). However, it suffices to show that (6.3.48) and (6.3.49) are identical as a function of r and σ_∞. This is obvious for finite σ_∞, so we just need to investigate both functions in the limit $\sigma_\infty \to 0$. In order to see this, observe that in this limit, the integrand in (6.3.48) simply becomes a $\delta_D(r-q)$, and the integrand in (6.3.49) just a Gaussian with constant variance $\sigma^2(r)/\sigma_\infty^2$ (this ratio does not scale with σ_∞). Fortunately, integrating over both just gives unity, such that, as expected, in this limit the correlation function is just zero. Thus we have shown that these expressions are indeed equivalent. We choose to analyze the latter.

for which typically $|f(q)| \leq 1$. The perturbative series is then obtained in the standard way:

$$\xi_n(r) = \frac{1}{n!}\left(\frac{\partial}{\partial\sigma_\infty}\right)^n \int \frac{dx}{\sqrt{2\pi}} \frac{1}{\sqrt{1-f(\sigma_\infty x + r)}} e^{-\frac{x^2}{2}\frac{1}{1-f(\sigma_\infty x+r)}}\Bigg|_{\sigma_\infty=0}. \tag{6.3.51}$$

One can check that this gives the same series as before *upon interchanging the order of the integral and derivatives.*

In order to use assumptions (6.3.5),(6.3.6), we analyze the integrand. More precisely, we show below that the point $\sigma_\infty = -r/x$ is an essential singularity in the complex σ-plane. The series therefore diverges for $\sigma_\infty > |r/x|$. Conversely, this means that for given σ_∞, the series diverges for $x > |r/\sigma_\infty|$. Moreover, the integral of derivatives of the integrand evaluated at $\sigma = 0$ are well defined. Hence we conclude that SPT diverges for the correlation function, given the non-convergence proof below.

Proof. Let us define

$$I(x,\sigma,r) = \frac{1}{\sqrt{1-f(\sigma_\infty x + r)}} e^{-\frac{x^2}{2}\frac{1}{1-f(\sigma_\infty x+r)}}, \tag{6.3.52}$$

we ask whether

$$\sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{\partial}{\partial\lambda}\right)^n I(x,\lambda,r)\Big|_{\lambda=0} \sigma^n \stackrel{?}{=} I(x,\sigma,r), \tag{6.3.53}$$

for every x, and given r. Let us now define a slight modification of this function

$$\tilde{I}(x,\lambda,r) \equiv I(x,\sigma/x,r) = \frac{1}{\sqrt{1-f(\lambda + r)}} e^{-\frac{x^2}{2}\frac{1}{1-f(\lambda+r)}}. \tag{6.3.54}$$

This has the property that

$$\sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{\partial}{\partial\lambda}\right)^n \tilde{I}(x,\lambda,r)\Big|_{\lambda=0} (\sigma x)^n = \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{\partial}{\partial\lambda}\right)^n I(x,\lambda,r)\Big|_{\lambda=0} \sigma^n. \tag{6.3.55}$$

This is useful for the following reason. Since $f(0) = 1$ by definition, and assuming f to be smooth and symmetric around zero, this means

$$\left(\frac{\partial}{\partial\lambda}\right)^n \tilde{I}(x,\lambda,r)\bigg|_{\lambda=-r} = 0, \tag{6.3.56}$$

so this function is certainly non-analytic in λ at $\lambda = -r$. The radius of convergence for $\tilde{I}$ around 0 is at most r, since this point constitutes an essential singularity in the complex plane. In fact, it is very reminiscent of the expansion of $e^{-1/(1+x)^2}$ around zero, which has radius of convergence 1. Thus we conclude that for $|x| > r/\sigma$,

$$\sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{\partial}{\partial\lambda}\right)^n \tilde{I}(x,\lambda,r)\bigg|_{\lambda=0} (\sigma x)^n \neq I(x,\sigma,r) \quad \text{for } |x| > r/\sigma. \tag{6.3.57}$$

But we can rewrite the left hand side back in its original form (this is a strict equality that follows from its definition), and conclude

$$\begin{aligned}&\sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{\partial}{\partial\lambda}\right)^n x^n \tilde{I}(x,\lambda,r)\bigg|_{\lambda=0} \sigma^n = \\ &= \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{\partial}{\partial\lambda}\right)^n I(x,\lambda,r)\bigg|_{\lambda=0} \sigma^n \neq I(x,\sigma,r) \quad \text{for } |x| > r/\sigma.\end{aligned} \tag{6.3.58}$$

□

Thus we conclude that perturbation theory for the correlation function diverges. Assuming convergence for $|x| < r/\sigma$, we can estimate the error as the contribution to the integral of the 'tails': the collection x, for which $|x| > r/\sigma$. Since f goes to zero for large arguments, these tail contributions are roughly exponentially suppressed by $e^{-(r/\sigma)^2}$, because of the exponent in the integrand. To be more precise, upon expanding, every term will be evaluated at $\sigma_\infty = 0$. This means the exponential suppression is $e^{-\frac{x^2}{2\sigma^2(r)/\sigma_\infty^2}}$. Plugging in the relevant value of x, we find that the non-perturbative error is indeed exponentially suppressed in the dimensionless variance

$$\text{NP-error} \sim e^{-\left(\frac{r}{\sigma(r)}\right)^2}. \tag{6.3.59}$$

We can verify numerically the non-convergence using the initial power spectrum (6.3.15). Figure 6.4 shows the comparison for the correlation function of the ZA result and the SPT approximations. For example, around $r = 90$ Mpc, we see that the 5 loop result is much closer to ZA than the 9 loop one. At those scales the series stops diverging somewhere between 5 and 9 loops. The divergence of perturbation theory at larger scales shows up at higher order in PT because the variance is smaller. For example, at $r \sim 110$ Mpc, PT is getting closer to the right answer up to 9 loops but then start diverging somewhere between 9 and 13 loops.

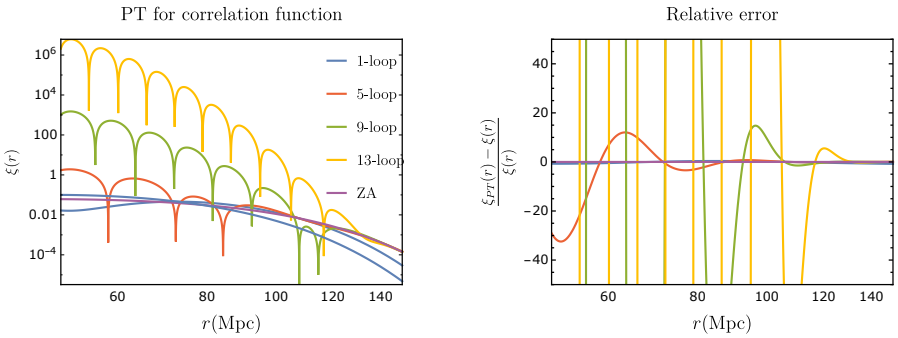


Figure 6.4: The plot shows the ZA correlation function and some of its SPT approximations. At any given r, perturbation theory stops improving and starts diverging from the exact result at high enough orders. At larger scales the divergence sets in at a higher loop order. This provides numerical evidence for the non-convergence of the SPT series for real space correlators, for arbitrarily large r. We use (6.3.15) as initial power spectrum.

6.3.7 (non-)Convergence of RPT for the correlation function

As we have seen previously, the power spectrum admits a convergent SPT expansion. Let us now assume that the exact correlation function exists and is finite

$$\xi(x) = \int_k e^{ikx} P(k) = \int_k e^{ikx} \sum_{n=1}^{\infty} P_n(k) < \infty \,. \tag{6.3.60}$$

Notice that, since $P_n(k) = P_n(-k)$ for every n, we can also use that the Fourier transform reduces to the cosine transform

$$\xi(x) = \int_k \cos(kx)\, P(k) = \int_k \cos(kx) \sum_{n=1}^{\infty} P_n(k) < \infty\,. \tag{6.3.61}$$

Then, a candidate perturbative expansion is given by

$$\xi(x) \stackrel{?}{=} \sum_{n=1}^{\infty} \xi_n(x) = \sum_{n=1}^{\infty} \int_k e^{ikx} P_n(k)\,. \tag{6.3.62}$$

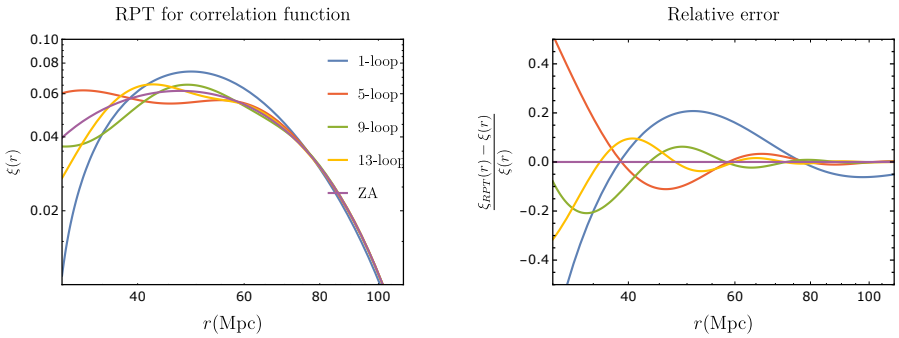


Figure 6.5: The plot shows the ZA correlation function and its RPT approximations. To the order we consider, there is no sign of divergence by increasing loop order, although the improvement from 9 to 13 loops is small. Lacking an analytical proof, we consider the (non-)convergence of RPT inconclusive. We use (6.3.15) as initial condition.

This perturbative expansion converges if we are allowed to exchange the sum with the integrand. For SPT, we proved in subsection 6.3.6 that this is not the case. One obstacle was that $P_n(x)$ are not all positive. To this end, let us consider Renormalized Perturbation Theory (RPT) [150]. At least for LCDM[11]

$$P_n^{\mathrm{RPT}}(k) = e^{-\frac{k^2}{2}\sigma_\infty^2} \frac{1}{n!} \int_q e^{-ikq} \left[\frac{k^2}{2}\sigma_q^2(q)\right]^n > 0\,. \tag{6.3.63}$$

[11] Note that there is a wrong minus sign in eq (2.39) of [116]

Then, indeed, all P_n are positive, as shown in [150]. Thus we can use Fubini-Tonelli theorem to invert the sum with the integral, provided that

$$\int_k \sum_{n=1}^{\infty} |\cos(kx)\, P_n^{\mathrm{RPT}}(k)| < \infty\,. \tag{6.3.64}$$

A sufficient condition for this to be true is $\xi(0) < \infty$, since

$$\begin{aligned}\int_k \sum_{n=1}^{\infty} |\cos(kx)\, P_n^{\mathrm{RPT}}(k)| &= \int_k |\cos(kx)\,| \sum_{n=1}^{\infty} P_n^{\mathrm{RPT}}(k) < \int_k \sum_{n=1}^{\infty} P_n^{\mathrm{RPT}}(k) \\ &= \int_k P(k) = \xi(0)\,,\end{aligned} \tag{6.3.65}$$

It turns out, however, that $\xi(0)$ is unbounded. This is due to the q^{-1} singularity in the integrand of its expression, which is hard to overcome as $\sigma^2(q) \sim q^2$ as $q \to 0$ by construction. Since we do not believe the presence of the $|\cos(kx)|$ factor changes the divergence of the integral, a proof along these lines seems out of reach. One could also ask what the difference between SPT and RPT is in terms of the non-convergence proof for SPT. Going back to (6.3.49) and the following discussion, we see that RPT is effectively expanding in the size of f. Then, interestingly, f is by definition bound to be less than or equal to unity (since $\sigma^2(q) \geq 0$), meaning the essential singularity is just barely part of the integrand; we do not go beyond it. It is therefore not possible to conclude that perturbation theory for the integral diverges along these lines either. We plot the performance of RPT for the same initial conditions in Figure 6.5. Once again, we believe the result is inconclusive.

6.4 The 1D count-in-cell PDF: non-convergence of PT for cumulants

In this section we discuss cumulants obtained from the probability distribution function (PDF) for the density averaged in cells of a certain fixed radius R. This observable recently attained some renewed interest in the 3D context, see e.g. [132, 151], but has been object of study for quite some time [152–162]. Here we show that in 1D, it is possible to use the exact

solution for planar collapse, see Appendix 6.A, to compute this PDF exactly up to shell crossing events (which are limited for our choice of initial conditions (6.3.20)). As we will see, the application of the construction of the PDF in 3D, which relies on spherical collapse, is easily applied to 1D. Most approaches [132, 156, 159, 162] agree on the exponential behavior of the PDF, but there is still some confusion about the prefactor [162], which is related to normalization constraints on the PDF: it should be unitary and have vanishing first moment. Typically normalization is enforced by hand [157, 158, 160, 162]. Here we present a new derivation of the so called Lagrangian space PDF, and show how, at least in 1D, one can also obtain the final Eurlerian space PDF from first principles, which automatically satisfies the unitarity and mass conservation contraints. Moreover, we show how the tails of this distribution are beyond the reach of perturbation theory, implying a non-perturbative error in the computation of any cumulant. This was also observed in the 3D context in [123]. We formalize the argument for non-convergence and highlight the relation to the non-perturbative error for the correlation function obtained in 6.3.6.

6.4.1 Conservation of probability

Let us start with the definition of the PDF for the density averaged in cells. The question the PDF should answer is the following:

- If one picks a random point x in space, what is the probability $P[\bar{\delta}_R]d\bar{\delta}_R$ that the average density between x and $x+R$, which we call $\delta_R(x)$, is in the range $[\bar{\delta}_R, \bar{\delta}_R + d\bar{\delta}_R]$?

Assuming ergodicity, an equivalent but more useful way to phrase the question is

- What fraction of the spatial volume $x \in V$ has the property that $\delta_R(x)$ is in the range $[\bar{\delta}_R, \bar{\delta}_R + d\bar{\delta}_R]$?

One can try to compute the answer to this question in two steps. The first step is to find how the above property translates into a property in the initial conditions: if $\delta_R(x)$ is in the range $[\bar{\delta}_R, \bar{\delta}_R + d\bar{\delta}_R]$, what does this mean for $\delta_{i,R_i}(q(x))$? Here $q(x)$ is the initial position of the fluid element that ends up at x and δ_i, R_i are to be found from the cell dynamics. Since we know the statistics of the initial conditions, we can then compute fraction of the initial

volume with this property. The second step is to compute how the volume fraction $q \in V_i$ with this property changes as a function of time. Combining these steps allows us to find the PDF.

6.4.2 Lagrangian space PDF

For simplicity and clarity, we start by forgetting about the second step, which has the same main features in terms the qualitative conclusions we like to draw from the PDF. We call it the Lagrangian PDF, and it gives the following probability:

- What fraction of the initial volume $q \in V_i$ has the property that at the final time $\delta_R(x(q))$ is in the range $[\bar{\delta}_R, \bar{\delta}_R + d\bar{\delta}_R]$?

There are two ways to obtain this Lagrangian PDF. The first is more elaborate and insightful, and, to the best of our knowledge, it is novel. The second is more elegant mathematically and can already be found in the literature [29, 159, 162]. We present them both here.

Method 1

Let us use the equations of motion to consider what the range $[\bar{\delta}_{R,f}, \bar{\delta}_{R,f} + d\bar{\delta}_{R,f}]$ maps into in the initial conditions. The equations are the (radius independent) equation for the evolution of the density, supplemented with a mass conservation equation,

$$\begin{aligned} 1 + \delta_f &= \frac{1}{1 - \delta_L}, \\ (1 + \delta_f) R_f &= R_i, \end{aligned} \tag{6.4.1}$$

where $\delta_L = \frac{a}{a_i}\delta_i$. This means that this range in δ at fixed R_f, maps into a line in $\{\delta, R_i\}$-space in the initial conditions (in the linear approximation). We care about its slope s:

$$s = \frac{d\delta_L}{dR_i} = \frac{d\delta_L}{d\delta_f}\frac{d\delta_f}{dR_i} = \frac{1}{(1 + \delta_f)^2 R_f}. \tag{6.4.2}$$

Thus the question becomes what the probability is that, given some point q in the initial conditions, the function $\delta_{R_i}(q) \equiv \delta_L(R_i)$ (it better be q-independent) as a function of R_i crosses the infinitesimal line element l

between the points $\{R_i, \delta_L\}$ and $\{R_i + d\delta_L/s, \delta_L + d\delta_L\}$ somewhere. Here $d\delta_L = d\delta_f/(1+\delta_f)^2$. For this, we need the following (correlated) ingredients:

1. What is the initial probability distribution for the density in cell of length R_i, $P_{R_i}(\delta_L)$?

2. What is the initial probability distribution for the derivative $\mu \equiv \frac{d\delta_L(R_i)}{dR_i}$, given $\delta_L(R_i)$?

In other words, we need their joint PDF. Fortunately, we can assume Gaussian initial conditions, for which the joint PDF is merely a two-dimensional Gaussian, determined by the correlators $\langle \delta^2_{R_i} \rangle, \langle \delta_{R_i} \mu \rangle$, and $\langle \mu^2 \rangle$. Formally, we now have to integrate over all combinations of δ_{R_i} and μ such that the line $\delta_{R_i} + \lambda\mu$, where $\lambda \in \mathbf{R}$, crosses l. And multiply its probabilities. The computation and corresponding approximations we leave to the appendix. The result for small enough variances however agrees with the much simpler expression we find next. The result is schematically given by

$$P_{R_f}(\delta_f) = \text{Prefactor}(\delta_f) P_G\left(\delta_L(\delta_f), \sigma_{R_i(\delta_f, R_f)}\right), \tag{6.4.3}$$

where the prefactor is some non-exponential function of δ_f, and P_G is the Gaussian probability density. The main qualitative features are determined by the Gaussian.

Method 2

The key observation for this method is that the fundamental variable for the Lagrangian PDF, $y \equiv \delta_L/\sigma_{R_i}$, is in fact Gaussian distributed [159, 162]. Thus,

$$P(y)dy = P_G(y, \sigma = 1)dy, \tag{6.4.4}$$

meaning

$$P^L_{R_f}(\delta_f)d\delta_f = P_G\left(\delta_L/\sigma_{R_i}\right) \times \left(\frac{d\delta_L}{d\delta_f}\frac{1}{\sigma_{R_i}} - \frac{\delta_L}{\sigma^2_{R_i}}\frac{d\sigma_{R_i}}{d\delta_f}\right) d\delta_f, \tag{6.4.5}$$

where the expression on the right should be interpreted as a function of R_f and δ_f, using the mapping. This gives the same prefactor as the first method for small variances, corrections to which are negligible in our case.

6.4.3 Eulerian space PDF

Even though the transition from Lagrangian to Eulerian space does not qualitatively change the PDF too much, and is therefore not too relevant for the purpose of this Chapter, we discuss it here for two reasons. First, we want to verify the statement that indeed the transition does not change the qualitative behavior too much. Second, as advertised above, there has been some discussion in the literature about the prefactor for the count-in-cell PDF in 3D, which is very similar in spirit to our 1D PDF. Our derivation of the Eulerian prefactor might inspire a new approach in that context as well.

Method 1 to derive the Lagrangian PDF gives us a way to calculate the volume fraction satisfying the given property as an integral over the probability of all initial conditions that satisfy that property. The Eulerian density is obtained from the Lagrangian one by multiplying it by the ratio of the final volume to the initial volume that satisfies this property. Remember that the initial volume we are talking about is the collection of $q \in V_{in}$ satisfying a certain property. Assuming some continuity conditions, this is indeed a volume. Namely, if $\bar{q}$ satisfies the property, then at least an infinitesimal region around $\bar{q}$ does also. The evolution of this volume element is then determined by the local density at $\bar{q}$ in the standard Zel'dovich manner:

$$r = \frac{dV_f\left(\bar{x}(\bar{q})\right)}{dV_i\left(\bar{q}\right)} = 1 - \delta_L(\bar{q}). \tag{6.4.6}$$

The probability density can thus be thought of as a weighted (by the joint probability) integral over all initial conditions that satisfy the Lagrangian condition, multiplied by this ratio r. Schematically, we can thus write

$$P_{R_f}(\delta_f)d\delta_f = \int_{I|_{\text{Lagrangian property}}} P_{MVG}\left[\delta_L(q), \delta_{R_i}(q), \mu(q)\right]\left(1 - \delta_L(q)\right), \tag{6.4.7}$$

where MVG stand for multi-variate Gaussian (as the initial conditions are Gaussian), I stands for initial conditions, which in this case this means integration over the random variables $\{\delta_L(q), \delta_{R_i}(q), \mu(q)\}$. The sub-text indicates the restriction on them. Even though we wrote the q dependence everywhere, statistical homogeneity guarantees the answer will not depend on it. Note that the multi-variate Gaussian depends on all non-vanishing cross correlations among the arguments as well. The detailed expression for

this PDF can be found in appendix 6.D. In Figure 6.6 we plot the PDF for a couple of radii, and check its normalization conditions, which are all correct at the subpercent level.

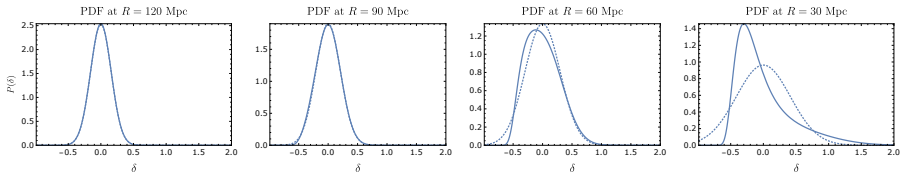


Figure 6.6: The plot shows the PDF (6.D.8) as a function of the final density evaluated at four different scales, respectively $R = 120, 90, 60, 30$ Mpc, including the linear, Gaussian PDF in dotted black for reference. Its norm and mean evaluate to $(1.006, 3 \times 10^{-5}), (1.005, 2 \times 10^{-4}), (1.004, 7 \times 10^{-4}), (1.004, 2 \times 10^{-3})$, respectively. We use (6.3.15) as initial condition.

6.4.4 Perturbation theory and convergence

Perturbation theory for the PDF proceeds similar to perturbation theory for the correlation function. As in the previous sections, perturbation theory here means a series in the amplitude of the primordial, linear power spectrum. As indicated in appendix 6.D, we can write the PDF solely as a function of δ_L,

$$P_{R_f}(\delta_f)d\delta_f = \text{Prefactor}(\delta_L) P_G\left(\delta_L, \sigma_{R_i(\delta_L,R_f)}\right) d\delta_L, \tag{6.4.8}$$

where $\sigma^2_{R_i(\delta_L,R_f)} = C_{22}(\delta_L)$. This PDF is, a priori, not well defined beyond $\delta_L = 1$, but the only sensible way to extend it is to set it to zero for larger values of the density, such that effectively we only integrate up to $\delta_L = 1$, which corresponds to $\delta = \infty$. The perturbative expansion is most easily obtained by introducing an artificial small parameter λ that multiplies the variance, and subsequently changing variable to $\tilde{\delta} = \delta_L/\lambda$. Due to the nature of the Gaussian distribution, λ drops out in several places, and the PDF becomes

$$P_{R_f}(\delta_f)d\delta_f = \text{Prefactor}(\lambda\tilde{\delta}, R_f) P_G\left(\tilde{\delta}, C_{22}(\lambda\tilde{\delta}, R_f)\right) d\tilde{\delta}. \tag{6.4.9}$$

The perturbative expansion is then obtained by expanding this function around $\lambda = 0$, before evaluating it at $\lambda = 1$. The convergence properties are most easily understood by realizing that both the prefactor and $C_{22}(\lambda\tilde{\delta}, R_f)$ in (6.4.9) are actually functions of $(1-\lambda\tilde{\delta})$. Let us forget about the prefactor for simplicity. It is instructive to write out the Gaussian part more explicitly. Let us denote $\sigma^2_{R_i} = h(R_i)$, which, through (6.4.1), can be written as $h(R_f/(1-\delta_L))$. The Gaussian then becomes

$$P_G\left(\tilde{\delta}, C_{22}(\lambda\tilde{\delta}, R_f)\right) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{h(R_f/(1-\lambda\tilde{\delta}))}} e^{-\frac{x^2}{2}\frac{1}{h(R_f/(1-\lambda\tilde{\delta}))}}. \quad (6.4.10)$$

Now note that for large R_i, $\sigma^2_{R_i}$ typically goes as R_i^{-n} for $n > 1$, implying $h(R_f/(1-\lambda\tilde{\delta})) \sim (1-\lambda\tilde{\delta})^n/R_f^n$. This means that, analogously to the expression for the real space correlation function, the perturbative expansion has an essential singularity at $\lambda = 1/\tilde{\delta}$. Conversely, setting λ to unity, this means perturbation theory diverges beyond $|\tilde{\delta}| > 1$. Since cumulants are obtained by integrating over the full PDF, perturbation theory for all cumulants is generically divergent (once again, no divergences appear in the limit $\lambda \to 0$, so we can intechange derivatives with respect to λ and the integral). We plot the perturbative approximations to the PDF for $R_f = 80$ Mpc in Figure 6.7. This is similar in spirit to the asymptotic nature of Edgeworth expansions of PDFs [163].

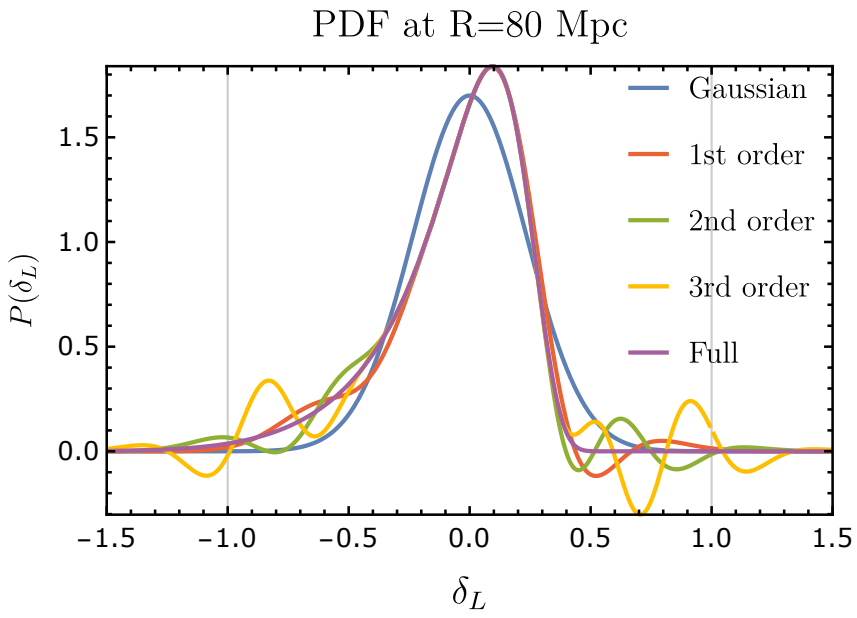


Figure 6.7: The plot shows the PDF for the average over cells of radius $R = 80$ Mpc (purple continuous line) as well as some of its perturbative approximations. Perturbation theory captures the PDF pretty well around the peak and one can see that including higher order terms extends the region where PT agrees with the full result. The convergence is not expected to improve beyond radius of convergence, which is indicated by the grey vertical lines. We use (6.3.15) as initial power spectrum.

6.5 General properties of the non-perturbative error

The previous sections contained concrete, explicit examples of the reach of perturbation theory for particular observables in particular settings. In contrast, we now speculate on the qualitative lessons we can learn from those examples. To highlight the generic nature of non-perturbative effects in real space, we review some well known facts about the halo model in appendix 6.E, whose non-perturbative effects have been estimated in, e.g. [118, 119].

6.5.1 Non-perturbative errors in real space

From the examples studied above, we can get some idea on a lower limit on the size of the non-perturbative error in real space. At least for the count-in-cell statistics and the correlations function, we found that observables $\mathcal{O}$ can be written qualitatively as an integral of the form

$$\mathcal{O}(r) \sim \int \frac{dx}{\sqrt{2\pi}} \frac{1}{\sqrt{g_r(1-\lambda x)}} e^{-\frac{x^2}{2}\frac{1}{g_r(1-\lambda x)}}, \tag{6.5.1}$$

for some function g_r, satisfying $g_r(0) = 0$. As a reminder, λ is a fictitious coupling constant which we expand around $\lambda = 0$, and which should be set to unity at the end of the calculation. We argued that the non-perturbative error comes from the $|x| \geq |1/\lambda|$ tails, since this point constitutes an essential singularity in the complex plane. This can be seen by approaching the point in the complex direction $\lambda x = 1 \pm i\epsilon$, where $\epsilon \to 0$. Thus, the error can be estimated as the tail contribution to each term in perturbation theory (PT):

$$\int_{\frac{1}{\lambda}}^{\infty} dx\, h(r,x) \frac{1}{\sqrt{2\pi g_r(1)}} e^{-\frac{x^2}{2g_{r(1)}}}, \tag{6.5.2}$$

for $h(r,x)$ some polynomial in x. At higher orders in PT the polynomial h becomes larger and larger, increasing the error on the tails. This can be seen explicitly in Figure 6.8, where we show the PT approximations of the integrand of the correlation function $\xi(r = 80\ \text{Mpc})$, (6.3.51). For clarity, we have subtracted the 0-th order Gaussian, which is responsible for the 1 in $1 + \xi(r)$. We have indicated the values beyond which the integrand does not converge to the true answer anymore, given by $x = \pm r/\sigma_\infty$. Since PT is asymptotic for the correlation function, the optimal approximation is obtained at some n_{opt}, beyond which the series starts diverging away from the true answer. We show how this optimal appriximation can be obtained numerically in Figure 6.9. We plot the PT errors, $\xi^{\text{PT}} - \xi$, for increasing PT orders as function of distance. We also compare this error to a rough estimate (black thick line) obtained from the dimensionless integrand of the tail of the distribution

$$\text{NP error} \sim \frac{1}{\sqrt{2\pi\tilde{\sigma}^2}} e^{-\frac{1}{2\tilde{\sigma}^2}}, \tag{6.5.3}$$

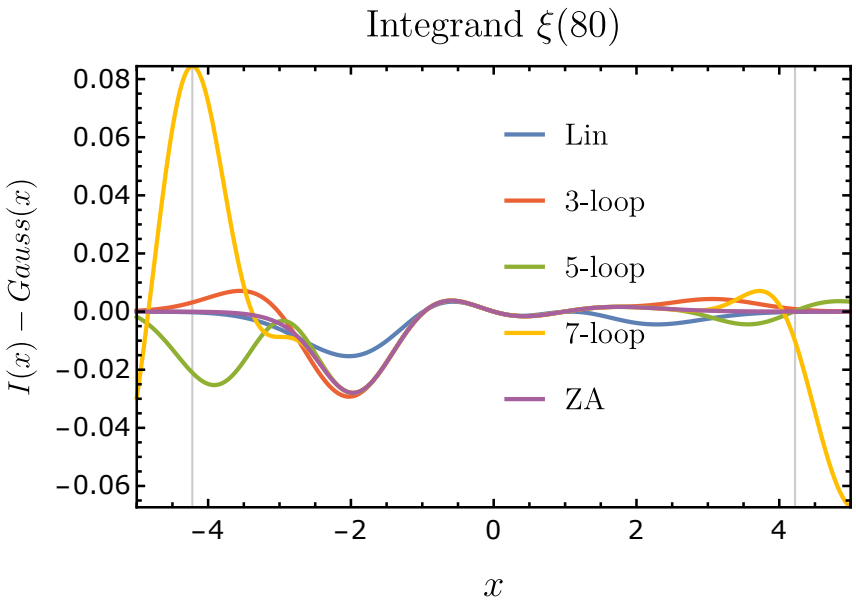


Figure 6.8: Here we plot the perturbative integrand for $\xi(80)$ minus the Gaussian whose integral evaluates to unity, and compare to the full result in red. Similar to the perturbative expression for the PDF, we see that perturbation theory performs well around the origin, but diverges wildly beyond the radius of convergence, $x = \pm r/\sigma_\infty$, which we plot as vetical grey lines. We use (6.3.15) as initial power spectrum.

where we used that $\lambda^2 g_r(1) = \sigma^2(r)/r^2 \equiv \tilde{\sigma}^2$. From Figure 6.9 we see that this estimate of the non-perturbative error increasingly underestimates the actual error for $r > 45Mpc$, presumably because of the large contribution from the polynomial term h. Physically, we believe the poor performance of PT on large scales is due to the relatively large RMS displacement on those scales for our initial condition (6.3.15). Using the notation and decomposition of [116], the SPT 1-loop correlation function can be written

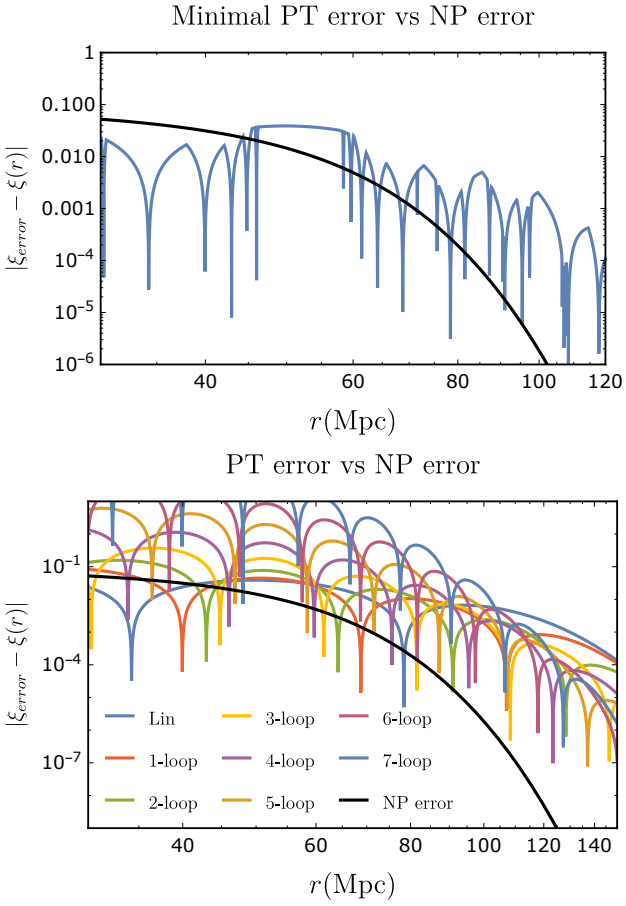


Figure 6.9: The upper plot shows the minimal error one can get from the asymptotic perturbation theory by choosing the order that best approaches the exact result. The lower panel shows that for the scales plotted, the error indeed starts growing beyond 6 loops, so that we can trust the minimal error from this plot on these scales. The estimated NP error is smaller than the actual error on scales $r > 45$ Mpc, but within a factor 10 on mildly nonlinear scales ($50 < R < 80$). Our estimate is orders of magnitude smaller than the actual error on larger scales. We believe this is due to the relatively large RMS displacement for our initial condition (6.3.15).

as[12]

$$\xi_{\text{SPT}}^{\text{1-loop}} = \xi_L(r) + \overbrace{3\xi_L^2(r)}^{\text{growth}} - \overbrace{4\xi_L'(x)\int_r^\infty dx\,\xi_L(x)}^{\text{dilation}} + \overbrace{\frac{\sigma^2(r)}{2}\xi_L''(x)}^{\text{RMS displ.}}. \tag{6.5.4}$$

The reason we have such a large RMS displacement, despite the absence of a BAO feature in our initial conditions is the fact that the linear variances are not simple power laws either, due to the exponential cutoff. This introduces another scales in the problem, which turns out to be picked up by the RMS displacement term. The large effect on large scales is illustrated in Figure 6.10.

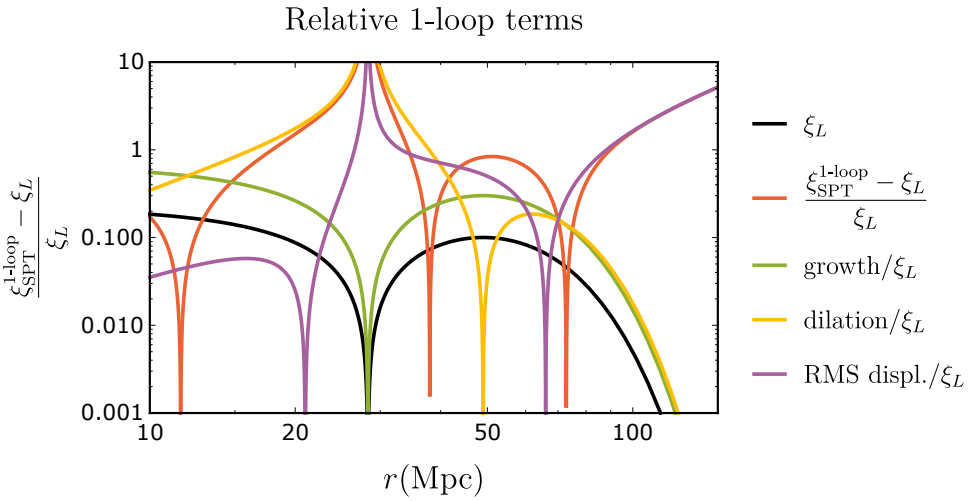


Figure 6.10: The plot shows the size of the various contributions (6.5.4) to $\xi_{\text{1-loop}} - \xi_L$ for our initial condition (6.3.15). We plot the 1-loop terms devided by the linear term and compare to the linear term, since in a typical perturbation theory the 1-loop terms are of order ξ_L^2. On large scales this assumption crearly does not hold in this case. On mildly nonlinear scales the RMS displacement term is still one order of magnitude larger than naively expected.

Finally, we stress that in all cases at least one of the tail contributions signals new physics, meaning they are probably hard to overcome analytically.

[12] Correcting the sign of the dilation term in (4.4) of [116].

For the PDF, the overdensity tail comes from the fact that large overdensities collapse, which is not captured by PT. Hence, a proper treatment of the tails requires a proper understanding of the collapsing process, which typically requires physics beyond the fluid approximation, such as a halo model, see appendix 6.E for references. Similarly, the essential singularity in the integrand of the expression for the correlation function (6.3.49) occurs when $q = \sigma_\infty x + r \leq 0$, which can be interpreted as the contribution to the correlation function at distance r from particles that were initially separated by distance $q \leq 0$, meaning it computes the contribution of particles that have crossed paths. These are beyond the reach of the fluid description and are also not properly captured by the ZA. A recent attempt to go beyond this limitation can be found in [129].

6.5.2 Non-perturbative effects in Fourier space

Here we note that the relation between non-perturbative effects in real and Fourier space is subtle. In particular, we investigate the analyticity in both cases. The results described here are a simple application of Paley-Wiener theorems. Suppose there is some error contribution to the correlation function

$$\xi(r) = \xi_{SPT}(r) + \xi_{NP}(r), \tag{6.5.5}$$

whose asymptotic behavior is

$$\lim_{r\to\infty} \xi_{NP}(r) = e^{-\frac{1}{\sigma_R^2(r)}}, \tag{6.5.6}$$

where σ_R^2 is the dimensionless variance. Now also suppose (realistically) that for large r, $\sigma_R^2(r) \sim r^{-n}$, for $n > 1$. Then we can prove that such a term leads to an analytic contribution to the power spectrum for the following reason. Because of linearity

$$P(k) \supset \int_r e^{ikr}\xi_{NP}(r) = \int_r \sum_n \frac{1}{n!}(ikr)^n \xi_{NP}(r). \tag{6.5.7}$$

In order to use Fubini-Tonelli, we observe that

$$\int_r \sum_n |\frac{1}{n!}(ikr)^n \xi_{NP}(r)| = \int_r \sum_n \frac{1}{n!}(kr)^n |\xi_{NP}(r)| = \int_r e^{kr}|\xi_{NP}(r)| < \infty, \tag{6.5.8}$$

by our assumption on the behavior of $\xi_{NP}(r)$ for large r. Thus we conclude that

$$P(k) \supset \int_r e^{ikr}\xi_{NP}(r) = \sum_n \frac{1}{n!}(ik)^n \int_r (r)^n \xi_{NP}(r), \tag{6.5.9}$$

meaning it is perfectly analytic in k.

Conversely, if we are looking for errors that lead to non-analytic in k type functions in Fourier space of the form e^{-1/k^n} times some polynomial, we need all derivatives with respect to k to vanish at $k = 0$. Assuming that the moments of the correlation function exist, this means that the real space equivalent of this correction has vanishing moments

$$\forall n \geq 0, \quad \int_r r^n \xi_{Error}(r) = 0. \tag{6.5.10}$$

Such functions exist but, as shown above, cannot fall off at infinity too rapidly[13]. We summarize these statements in the following table.

Real space	vs	Fourier space
Faster than e^{-r} as $r \to \infty$	$\Longrightarrow$	Analytic in k
Zero moment functions (less than exponential, but more than polynomial fall off at infinity)	$\Longleftrightarrow$	Vanishing derivatives at $k = 0$ (the right to left implication makes sense for interchangeable derivative and integral only)

[13] Even though the examples we studied above all led to non-perturbative corrections to the correlation function that fall off exponentially at large r, it is not hard to understand that non-perturbative physics can alter the correlation function in a manner that does not fall off exponentially at large r as well. As an example, consider just putting a rectangular bump of width dk around k_{nl} in the power spectrum: $P(k) = P_{old}(k) + Rect(k_{nl}, dk)$. The Fourier transform of the latter term is $\propto \cos(k_{nl}r)\cos(rdk)/(rdk)$. Then it is not inconceivable that non-perturbative physics, which displaces particles around the nonlinear scale, completely changes the size of this bump in a non-calculable way. This therefore changes the respective term in the correlation function, which does not fall off exponentially at infinity. Moreover, it is an example of a contribution to the power spectrum that is non-analytic in k (and which is not multiplied by k^4). More generally, we expect smooth, but non-analytic at $k = 0$, approximations to the Dirac-delta function to have similar behavior.

6.5.3 Relation non-pertubative error and EFT-terms

One might wonder how the effects beyond perturbation theory that we discuss in this Chapter are related to short scale non-perturbative physics that is captured by the EFT of LSS. The fact that they are qualitatively different can be understood as follows. Observe that the non-perturbative effects we are discussing exist even if we cut off the power spectrum on short scales, such that there is zero power left. In that case, counterterms become irrelevant, but the effect we describe remains. Physically, the EFT is designed to capture the effect of unknown, short scale physics on long scales. The effect we are describing, however, are in a sense intrinsic limitations of the long scale physics *itself*. In the first place, the mere asymptotic nature of perturbation theory can put a limit on the level of precision with which we can describe long scale physics, even in the absence of any coupling to short modes. Moreover, we argued that since long scale observables are statistical in nature, they always rely on a 'tail contribution' (rare events) which are beyond perturbative understanding even on large scales. Nonetheless, since in a ΛCDM universe the EFT contributions on mildly nonlinear scales appear to be much larger than the non-perturbative effects we are referring to in this Chapter, the EFT of LSS is still very useful [47].

6.6 Conclusion and outlook

In this Chapter we made progress towards a better understanding of the reach of perturbation theory for large scale structures and the relevance of non-perturbative (NP) effects. In the context of 1D physics, we proved the asymptotic nature of perturbation theory for the real space correlation function and count-in-cell cumulants on all scales. In both cases, the proof is based on the presence of an essential singularity in the domain of the defining integral. Interestingly, this singularity has physical significance, signaling the collapse of cells or multi-streaming events. This adds to the intuition about NP effects in the halo model, which we review in Appendix 6.E. Altogether, *this suggests that there is indeed a floor to the reach of perturbative approaches to LSS based on fluid dynamics*. For our initial condition (6.3.15), we found that in real space the best possible perturbative approximation is typically worse than a naive guess for the error based on

the size of the tail of the integrand (6.5.2),

$$\xi_{\rm NP}(r) \sim \frac{1}{\sqrt{2\pi}} \frac{r}{\sigma(r)} e^{-\frac{r^2}{2\sigma(r)^2}} \quad \text{(naive estimate)}, \tag{6.6.1}$$

where r/σ is dimensionless. It would be interesting to study the implication for these non-perturbative effects for the position and shape of the BAO feature.

Conversely, we found that 1D SPT is convergent for the Fourier space observables we considered: the power spectrum for both Gaussian and non-Gaussian initial conditions, and the bispectrum for Gaussian initial conditions (assuming ΛCDM-like initial conditions). We expect this to extend to other Fourier space observables as well. Convergence seems to be a peculiarity of 1D dynamics and it is not expected to extended to 3D. In fact, the authors of [117] already showed that the 3-loop SPT correction to the power spectrum on large scales displays properties compatible with an asymptotic series.

We stress this does not yet imply that perturbation theory converges to the physically correct answer. In fact, even neglecting EFT corrections, we expect non-perturbative effects of the form

$$\text{NP error power spectrum} \sim \frac{1}{\sqrt{2\pi\Delta^2(k)}} e^{-\frac{1}{2\Delta^2(k)}}, \tag{6.6.2}$$

where Δ is the dimensionless Fourier space variance, to play a role for the power spectrum as well. The intuitive reason is once again the statistical nature of cosmological correlators. This means that even on large scales there are always some rare events for which the dimensionless quantity $k\delta(k)$ is larger than unity and the fluid description fails. The inevitable statistical contribution of these rare events to cosmological correlators is then expected to be exponentially suppressed. An analytic or quantitative understanding of this statement for Fourier space observables remains an open problem.

The importance of a solid control over theoretical errors in predictions for LSS observables was argued in [2, 91]. One of the hopes would therefore be to provide some sort of fitting function for NP effects on mildly nonlinear scales. Unfortunately, it is not clear a priori if there are any symmetries or other physical arguments to do so, despite significant effort in this direction in the literature. Here, we presented a novel construction for the count-in-cell PDF in 1D, which could be useful for the construction of its 3D analogue.

A thorough understanding of theoretical errors in these approaches seems to still be lacking.

Finally, the most useful step going forward is probably a combined effort of both analytical and numerical searches for NP effects, as advocated in [127].

6.A Planar collapse

Here we show that the Lagrangian equation (6.2.9) is also found for the evolution of the density in finite cells. We derive this within Newtonian cosmology. Let us consider the evolution of a cylinder, whose symmetry axis is along the x-direction with physical radius a and length R. In the Newtonian approximation, we can use Gauss' law in physical coordinates

$$\int_V \nabla^2\phi = \int_S \left(\vec{\nabla}\phi\right)\cdot\vec{n}, \tag{6.A.1}$$

where V is the volume of the cell, S is the surface, and $\vec{n}$ is the surface normal vector. Using the Poisson equation,

$$\nabla^2\phi = 4\pi G\rho(x), \tag{6.A.2}$$

this gives

$$4\pi GM = \int_S \left(\vec{\nabla}\phi\right)\cdot\vec{n}, \tag{6.A.3}$$

where and M is the total mass inside the cylinder. We now wish to evaluate the surface integral. In order to compute the flux through the side of the cylinder, observe that shuffeling matter in the x-direction cannot induce a relative force between particles in the orthogonal directions. The flux through the side of the cylinder is therefore equal to what it is in the completely homogeneous case, and its motion is according to the average expansion of the universe. We can therefore interpret a as the scale factor, with corresponding dynamics. The cylinder surface integral then becomes

$$\begin{aligned}\int_S \left(\vec{\nabla}\phi\right)\cdot\vec{n} &= 2\pi aR\left(\nabla_a\phi\right) + \pi a^2\left(\nabla_x\phi(x_1+R) - \nabla_x\phi(x_1)\right)\\ &= -2\pi aR\ddot{a} - \pi a^2\ddot{R},\end{aligned} \tag{6.A.4}$$

where x_1 should be considered as a label for this cell, tracking the left outer edge. We define the average density in this cell as

$$\delta_R(x_1) = \frac{1}{R}\int_{x_1}^{x_1+R} \delta(x)dx. \tag{6.A.5}$$

Mass conservation inside the cell then relates the average density to the size of the cell,

$$R = \frac{M}{\pi a^2 \bar{\rho}(1+\delta_R)}. \tag{6.A.6}$$

Plugging this into the integrated Poisson equation

$$4\pi G M = -2\pi a R\ddot{a} - \pi a^2 \ddot{R}, \tag{6.A.7}$$

and using the Friedmann equations for a matter dominated universe,

$$\begin{cases} H = \frac{\dot{a}}{a} = \frac{8\pi G}{3}\bar{\rho} \\ \dot{H} + H^2 = \frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\bar{\rho}, \end{cases} \tag{6.A.8}$$

this yields the following equation for the evolution of the average density:

$$\ddot{\delta}_R + 2H\dot{\delta}_R - 2\frac{\dot{\delta}_R^2}{1+\delta_R} = 4\pi G\bar{\rho}\delta_R(1+\delta_R). \tag{6.A.9}$$

As noted before, the solution to this equation is found from the nonlinear transformation (6.2.10).

6.B Growing and decaying modes in spherical collapse

The growing and decaying modes in the spherical collapse solution are obtained by taking η and C, and therefore t, small. Let us first define

$$y \equiv \frac{t-C}{B} = \eta - \sin\eta. \tag{6.B.1}$$

Then, in the small η limit, we find to second order

$$y = \frac{\eta^3}{6} - \frac{\eta^5}{120}. \tag{6.B.2}$$

This we can invert perturbatively to find to second order

$$\eta = \left(6y + \frac{(6y)^{2/3}}{20}\right)^{1/3} = (6y)^{1/3} + \frac{y}{10}. \tag{6.B.3}$$

Expanding the denominator in (6.2.28), to second order we find

$$\delta = \frac{9t^2}{2B^2}\left(\frac{8}{\eta^6} + \frac{2}{\eta^4}\right) - 1. \tag{6.B.4}$$

Plugging in the expression for η in terms of y, the second order expression becomes

$$\delta = \frac{9t^2}{2B^2}\left(\frac{2}{9y^2} + \frac{1}{5y(6y)^{1/3}}\right) - 1. \tag{6.B.5}$$

Finally, we substitute back the expression for y in terms of t and C, and expand to first order in C, thereby assuming $C \ll t \ll 1$, which yields

$$\delta_L = \frac{3}{10}\left(\frac{9}{2}\right)^{1/3}\left(\frac{t}{B}\right)^{2/3} + \frac{2C}{t}. \tag{6.B.6}$$

6.C Displacement field symmetries and the bispectrum

The equation of motion for the stochastic field Ψ, (6.2.5), is invariant under two separate symmetries

$$Q_1 \quad : \qquad \Psi(q) \to \psi'(q) \equiv \Psi(q + c_1(t))\,, \tag{6.C.1}$$

$$Q_2 \quad : \qquad \Psi(q) \to \psi'(q) \equiv \Psi(q) + c_2(t)\,, \tag{6.C.2}$$

for arbitrary functions of time $c_i(t)$. It is straightforward to see that both these transformations correspond to translations of δ, namely (in d-dimensions)

$$\delta(x) = \int d^d q \delta_D\left(x - q - \Psi(q)\right) \to \delta'(x) \equiv \delta(x + c_1 - c_2)\,. \tag{6.C.3}$$

Therefore, for $c_1 = c_2 = c$, the linear combination $Q_+ \equiv Q_1 + Q_2$ leaves δ exactly invariant (not covariant), while the orthogonal combination $Q_- \equiv Q_1 - Q_2$ induces a translation $\delta(x) \to \delta(x + 2c)$.

Now there are two different situations:

- *If and only if we are interested in correlators of* δ, we can perform a Q_+ transformation without changing the value of the δ correlator. Also, assuming statistical homogeneity, we can also perform a Q_- transformation, without changing the value of the δ correlator. Hence, we use both Q_1 and Q_2 at will to simplify our δ correlators.

- *If instead we also care about correlators of* Ψ *that do not combine into a* δ, things are trickier. In principle one can invert (6.C.3) exactly getting $\Psi(\delta)$ and then again both Q_+ and Q_- annihilate correlators. In practice though we are not able to invert (6.C.3) but only at linear order, which is a good approximation if the initial perturbations are small $|\Psi|, |\delta| \ll 1$. This linearized inversion, (6.3.8), assumes $|\psi| \ll 1$ and therefore breaks Q_2, but not Q_1. So, when we compute Ψ correlators using this approximation, or equivalently declaring the initial displacement power spectrum (6.3.9), we can perform Q_1 without changing the value of the correlator by invoking statistical homogeneity, but we cannot perform Q_2 transformations.

Let us move on to compute the power spectrum, but inverting the usual definition

$$\langle \delta(\vec{k})\delta(\vec{k}')\rangle \equiv (2\pi)^d\, \delta_D^{(d)}\left(\vec{k}+\vec{k}'\right) P(k)\,, \tag{6.C.4}$$

namely

$$P(k) = \int_{\vec{k}'} \langle \delta(\vec{k})\delta(\vec{k}')\rangle\,. \tag{6.C.5}$$

Using the Fourier transform of (6.C.3), namely

$$\delta(k) = \int dq e^{-ik[q+\Psi(q)]}\,, \tag{6.C.6}$$

one finds three types of terms

$$P(k) = \int_{k'qq'} e^{-i(kq+k'q')}\left[\langle e^{-i(k\Psi+k'\Psi')}\rangle - \langle e^{-ik\Psi}\rangle - \langle e^{-ik\Psi}\rangle + 1\right]. \tag{6.C.7}$$

Using the symmetry Q_2

$$\langle e^{-ik\Psi}\rangle = \langle e^{-ik(\Psi-\Psi)}\rangle = 1\,. \tag{6.C.8}$$

One might be slightly worried about choosing the parameter $c(t)$ of the Q_2 transformation to be a random variable, rather than just a function. This can in principle be justified by looking at the PDF formulation of the correlators, but we won't do it here. Using both Q_1 and Q_2 we can re-write

$$\langle e^{-i(k\Psi+k'\Psi')}\rangle = \langle e^{-ik(\Psi(q-q')-\Psi(0))-ik'(\Psi(0)-\Psi(0))}\rangle = \langle e^{-ik\Delta\Psi(q-q')}\rangle, \tag{6.C.9}$$

with

$$\Delta\Psi(q) \equiv \Psi(q) - \Psi(0)\,. \tag{6.C.10}$$

These two tricks lead to the standard result

$$P(k) \;\; = \;\; \int dq\, e^{-ikq}\left[\langle e^{-ik\Delta\Psi(q)}\rangle - 1\right]\,. \tag{6.C.11}$$

The calculation of the bispectrum proceeds very similarly starting from the definition

$$B(k_1,k_2) \equiv \int_{k_3} \langle\delta(k_1)\delta(k_2)\delta(k_3)\rangle\,. \tag{6.C.12}$$

The result is the same as one would have guessed from the get go, (6.3.28).

6.D 1D count-in-cell PDF

Here we provide the detailed formulas and final expression for the PDF (6.4.7), which we repeat here for convenience

$$P_{R_f}(\delta_f)d\delta_f = \int_{I|_{\text{wih Lagrangian property}}} P_{MVG}\left[\delta_L(q),\delta_{R_i}(q),\mu(q)\right](1-\delta_L(q)). \tag{6.D.1}$$

Let us first specify the multivariate Gaussian:

$$P_{MVG}\left[\delta_L(q),\delta_{R_i}(q),\mu(q)\right] = \frac{1}{\sqrt{(2\pi)^3|C|}} e^{-\frac{1}{2}2\vec{v}^T C^{-1}\vec{v}}, \tag{6.D.2}$$

where $\vec{v}^T = (\delta_L(q), \delta_{R_i}(q), \mu(q))$, and $C = \langle \vec{v}\vec{v}^T \rangle$. More explicitly, the symmetric covariance matrix is

$$C_{ij} = \begin{pmatrix} \int_k P_L(k) & \int_k \frac{\sin(kR)}{kR} P_L(k) & -\frac{1}{R}C_{12} + \frac{1}{R}\xi_L(R) \\ C_{12} & \int_k \left(\frac{2\sin(\frac{kR}{2})}{kR}\right)^2 P_L(k) & -\frac{1}{R}C_{22} + \frac{1}{R}C_{12} \\ C_{13} & C_{23} & \frac{1}{R^2}C_{22} + \frac{1}{R^2}C_{11} - \frac{1}{R^2}C_{12} \end{pmatrix}, \tag{6.D.3}$$

which is positive definite. Since the domain of integration only restricts the integrals over δ_R and μ, we can perform the integral over δ_L already at this point, yielding

$$\begin{aligned} &\int d\delta_L P_{MVG}\left[\delta_L, \delta_{R_i}, \mu\right](1 - \delta_L(q)) = \\ &\qquad = P_{biVG}\left[\delta_{R_i}, \mu\right]\left(1 - \frac{\delta_R\left(\frac{C_{12}}{C_{22}} - \frac{C_{13}C_{23}}{C_{22}C_{33}}\right) + \mu\left(\frac{C_{13}}{C_{33}} - \frac{C_{12}C_{23}}{C_{22}C_{33}}\right)}{1 - \frac{C_{23}^2}{C_{22}C_{33}}}\right), \end{aligned} \tag{6.D.4}$$

where P_{biVG} denotes the bivariate Gaussian distribution. Next we specify how we restrict the domain of the remaining integral. Since we are dealing with infinitesimals, we can use linear approximations everywhere. Therefore, we wish to integrate over all initial δ_R and μ, such that the line $\delta_R + \lambda\mu$ crosses $\left[\{R, \bar{\delta}_L\}, \{R + d\bar{\delta}_L/s, \bar{\delta}_L + d\bar{\delta}_L\}\right]$, which is a line element.This comes down to the following restriction:

$$\begin{cases} \text{if} \quad \delta_R > \bar{\delta}_L, & \mu < s\left(1 + \frac{\bar{\delta}_L - \delta_R}{d\bar{\delta}_L}\right) \\ \text{if} \quad \delta_R < \bar{\delta}_L, & \mu > s\left(1 + \frac{\bar{\delta}_L - \delta_R}{d\bar{\delta}_L}\right) \end{cases}. \tag{6.D.5}$$

It is in this case convenient to rewrite $\delta_R = \bar{\delta}_L + f d\bar{\delta}_L$, and change variables from δ_R to f, such that the PDF becomes proportional to $d\bar{\delta}_L$, as it should, and the above condition becomes cleaner. In fact, we can replace δ_R with $\bar{\delta}_L$, forgetting about the $f d\bar{\delta}_L$ correction, everywhere apart from (6.D.5), since the latter is the only place where the infinitesimal drops out. One might worry that the integral over f could spoil the smallness of this term. The reason we do not have to worry about this is that for large f, (6.D.5)

guarantees we only integrate over $|\mu| \gg 1$, whose contribution to the integral is exponentially small. The remaining multivariate Gaussian integral, with the mentioned restriction, can then be written as

$$P_{R_f}(\delta_f)d\delta_f = P_G(\bar{\delta}_L, C_{22}) \left(\int_0^\infty df \int_{-\infty}^{s(1-f)} d\mu + \int_{-\infty}^0 df \int_{s(1-f)}^\infty d\mu \right) \times$$
$$\times P_G\left(\mu, \tilde{\sigma}^2, m\right) \left(1 - \frac{\bar{\delta}_L \left(\frac{C_{12}}{C_{22}} - \frac{C_{13}C_{23}}{C_{22}C_{33}} \right) + \mu \left(\frac{C_{13}}{C_{33}} - \frac{C_{12}C_{23}}{C_{22}C_{33}} \right)}{1 - \frac{C_{23}^2}{C_{22}C_{33}}} \right), \quad (6.D.6)$$

where $\tilde{\sigma}^2 = C_{33}(1-\kappa^2)$, $m = \bar{\delta}_L \kappa \sqrt{\frac{C_{33}}{C_{22}}}$ is the mean of the Gaussian, and $\kappa^2 = C_{23}^2/(C_{22}C_{33})$. The nontrivial integrals can now be approximated by

$$\left(\int_0^\infty df \int_{-\infty}^{s(1-f)} d\mu + \int_{-\infty}^0 df \int_{s(1-f)}^\infty d\mu \right) \times P_G\left(\mu, C_{33}(1-\kappa^2), m\right)$$
$$\approx 1 - \frac{\mu}{s}$$
$$\left(\int_0^\infty df \int_{-\infty}^{s(1-f)} d\mu + \int_{-\infty}^0 df \int_{s(1-f)}^\infty d\mu \right) \times P_G\left(\mu, C_{33}(1-\kappa^2), m\right) \mu$$
$$\approx m - \frac{m^2}{s} - \frac{\tilde{\sigma}^2}{s}. \quad (6.D.7)$$

These approximations are valid up to exponentially small corrections in $(s-m)^2/\tilde{\sigma}^2$, which lead to sub-percent corrections to the PDF. Thus, the final formula for the PDF is

$$P_{R_f}(\delta_f)d\delta_f = P_G(\bar{\delta}_L, C_{22}) \times \quad (6.D.8)$$
$$\times \left(1 - \frac{m}{s}\right) \left(1 - \frac{\bar{\delta}_L \left(\frac{C_{12}}{C_{22}} - \frac{C_{13}C_{23}}{C_{22}C_{33}} \right) + \left(m - \frac{\tilde{\sigma}^2}{s-m} \right) \left(\frac{C_{13}}{C_{33}} - \frac{C_{12}C_{23}}{C_{22}C_{33}} \right)}{1 - \frac{C_{23}^2}{C_{22}C_{33}}} \right).$$

Using the planar collapse equations, one can write $R_i(R_f, \delta_f)$, or $R_i(R_f, \delta_L)$, and write the PDF in terms of both variables δ_f and δ_L as one pleases. The former is the observationally relevant formulation, but the latter is more convenient to compare with perturbation theory.

6.E Non-perturbative terms in the halo model

The halo model relies on the same idea we used to construct the PDF above [164–166]. The difference is that in the halo model, densities above a certain threshold collapse to form halos that have a mass dependent spatial profile (e.g. the NFW profile [167]). The abundance of halos of a certain mass, which in the halo model translates into a certain initial radius R_i, is therefore given by the tail of the initial Gaussian distribution, with variance $\sigma^2_{R_i}$. This appendix does not contain original work, but simply highlights that the 1-halo term is a non-perturbative contribution to the correlation function similar to the non-perturbative errors we study in the main text.

Correlation function

The 1-halo contribution to the real space two-point correlation function in the halo model is given by

$$\xi_{1h}(r) = \frac{1}{\bar{\rho}^2}\int dm\, m^2 n(m) \int dx\, \lambda_m(x)\lambda_m(x+r) \tag{6.E.1}$$

We care about the contribution of this term at large r. First observe that

$$\frac{m^2 n(m)}{\bar{\rho}}\frac{dm}{m} = \sqrt{\frac{2}{\pi}} e^{-\frac{\nu^2}{2}} d\nu; \quad \nu = \frac{\delta_c}{\sigma(m)}, \tag{6.E.2}$$

and $\sigma(m)$ is given by the initial variance σ_{R_i}, and the relation $\bar{\rho}(a_i)R_i = m$. The 1-halo term thus becomes

$$\begin{aligned}\xi_{1h}(r) &= \frac{1}{\bar{\rho}}\int dm\, m\sqrt{\frac{2}{\pi}} e^{-\frac{\nu^2}{2}}\frac{d\nu}{dm}\int dx\, \lambda_m(x)\lambda_m(x+r) \\ &\equiv \frac{1}{\bar{\rho}}\int dm\, m\sqrt{\frac{2}{\pi}} e^{-\frac{\nu^2}{2}}\frac{d\nu}{dm} f(r,m).\end{aligned} \tag{6.E.3}$$

Now, irrespective of details of f and $\sigma(m)$, it is already clear that this term is non-perturbative in the amplitude of the linear variance. Namely, derivatives of the integrand with respect to σ are identically zero at $\sigma = 0$ for all m. Since these derivatives are well-behaved, we can interchange derivative and the integral over m, and conclude that the derivatives of this contribution to the correlation function are indeed non-analytic, but

obviously non-vanishing. We can try to go a bit further and estimate its behavior as a function of r for large r. For typical halo profiles λ_m, larger mass halos will extend over a larger region of space. This qualitatively causes $f(r, m)$ to only have support for $m > M(r)$, where $M(r)$ is some function that computes the mass threshold. Then the integral can be written as

$$\xi_{1h}(r) = \frac{1}{\bar{\rho}} \int_{M(r)}^{\infty} dm\, m \sqrt{\frac{2}{\pi}} e^{-\frac{\nu^2}{2}} \frac{d\nu}{dm} f(r, m). \tag{6.E.4}$$

Then, for large mass, the variance typically goes as some negative power law: $\sigma^2(m) \sim m^{-n}$, for $n > 0$. Thus for large r,

$$\xi_{1h}(r \to \infty) = \frac{1}{\bar{\rho}} \int_{M(r)}^{\infty} dm\, m \sqrt{\frac{2}{\pi}} e^{-cm^n} \frac{d\nu}{dm} f(r, m). \tag{6.E.5}$$

Due to the exponential cutoff, we can therefore very roughly estimate the large r behavior of this expression as

$$\log \xi_{1h}(r \to \infty) \propto -M^n(r), \tag{6.E.6}$$

which is the dominant term if this function is a power law. Generically, this is not obvious, but one can for instance naively assume the size of the halo is given by the Lagrangian radius, derived from the initial mass-scale relation, in which case the power law is linear. This therefore gives another argument for the exponentially suppressed contribution of non-perturbative terms at large distances.

Power spectrum

As is well-documented , the 1-halo term does not necessarily lead to non-analyticity in k for the power spectrum [168]. From its expression

$$P_{1h}(k) = \frac{1}{\bar{\rho}^2} \int dm\, m^2 n(m) |\lambda_m(k)|^2, \tag{6.E.7}$$

we see that, similar to the correlation function, this expression is non-perturbative in the linear variance, since $n(m)$ is non-perturbative, while at the same time the integral is well-behaved in the $\sigma \to 0$ limit. In order to learn about the analytic properties in terms of k, observe that derivatives in

the integrand can be interchanged with the integral over m. It is not a priori clear whether Fubini-Tonelli can be applied here, since we haven't specified $\lambda_m(k)$. We can still make progress understanding the low-k limit though. First, we recall that $\lambda_m(k)$ must go to unity as $k \to 0$, since

$$\lambda_m(k) = \int dx\, e^{ikx} \lambda_m(x), \tag{6.E.8}$$

and λ_m is normalized such that its integral over space is unity. Therefore the power spectrum gets a constant contribution as $k \to 0$, which forms the classic contradiction [120, 121, 168, 169] between the halo model and Peebles's argument about the k^4 scaling of non-perturbative effects from small scales at low k in the power spectrum as a consequence of mass and momentum conservation [4, 41]. We will not go into this further at this point. Finally, we note that if $\lambda_m(k)$ contained some non-analytic piece,

$$\lambda_m(k) \underset{k\to 0}{=} 1 + \text{analytic terms } + \#e^{-\frac{1}{k^2}}, \tag{6.E.9}$$

it would not be captured by any Taylor expansion in small k of the power spectrum. We leave it to future work to see whether such terms exist, and what their relation to Peebles's argument is.

7 Conclusion and Outlook

Since the main Chapters of the thesis (4, 5, 6) each contain an elaborate conclusion and outlook, we here stick to the main findings and a broad outline to conclude this thesis. For a more detailed discussion of the results, we refer the reader to the conclusions of the respective Chapters.

Chapter 4 lays the theoretical framework for the treatment of primordial non-Gaussianity in large scale structure. We showed how the addition of a new, well defined set of operators to the equations of motion for matter density perturbations renders the theory perturbatively renormalizable and consistent as an effective theory. As we show in Chapter 5, these new operators are actually irrelevant for the extraction of a PNG signal from the late time matter distribution. This is due to the smallness of PNG, as well as the fact that mass and momentum conservation constrain the backreaction of these new operators on the matter distribution to be down by two derivatives. This is no longer the case if we look at the biased tracers that are actually observed. The theory developed here was extended to the theory of biased tracers in [170]. Especially the more local type of new operators are expected to be of great observational relevance in the form of scale dependent bias, see for instance [83, 84].

The main findings of Chapter 5 were summarized in Table 5.3. It shows that the EFT modeling of dark matter is in principle sufficient to improve over current bounds for PNG, but insufficient to reach some acclaimed theoretically interesting benchmarks (see [31]). Moreover, we highlighted the importance of a proper treatment of the theoretical error, and argued that our results are pretty robust against a modified treatment of the nuisance parameter correlation length in particular.

The upshot of the results in Chapter 6, apart from mathematical results on the convergence of perturbation theory for certain LSS observables, is of a more qualitative nature: it suggests a lower bound for the theoretical error in *any* perturbative treatment to LSS. Further theoretical and numerical studies are required to learn more about the scale-dependence and size of these terms in the three-dimensional context. For other work in this di-

rection, see [127]. We presented a first suggestion for the non-perturbative error as a term exponentially small in the reciprocal of the dimensionless variance, which is the SPT expansion parameter.

Short term extensions of our theory include the integration of the EFT with PNG into a final theory for redshift-space observables. For developments along these lines, see e.g. [58, 171]. A subsequent forecast similar to ours for dark matter is a crucial next step. One might also wonder about the combined effort of analytical methods as we have presented here, and numerical solutions to the Vlasov-Poisson system. First, let us stress that the EFT perspective could be relevant for numerical error analysis as well, as noted in [127]. Furthermore, our forecast shows that the shapes of the EFT corrections to SPT are sufficiently distinct from the signal that the EFT coefficients can be efficiently extracted from the survey itself. For optimal results, or other signal shapes, however, extraction of the EFT coefficients from N-body simulations could be useful, a framework for which was for instance suggested in [172].

Finally, improvements on the theoretical control over non-perturbative errors is both interesting and challenging. The importance depends on the question asked, but rapidly becomes more relevant as theories are required to describe physics close to the nonlinear scale. In the absence of a more thorough understanding of non-perturbative effects in LSS, we advocate keeping in mind at least our lower-limit estimate for its effect in parameter estimation, along the lines of 5 and [91].

Bibliography

[1] V. Assassi, D. Baumann, E. Pajer, Y. Welling, and D. van der Woude, "Effective theory of large-scale structure with primordial non-Gaussianity," *JCAP* **1511** (2015) 024, arXiv:1505.06668 [astro-ph.CO].

[2] Y. Welling, D. van der Woude, and E. Pajer, "Lifting Primordial Non-Gaussianity Above the Noise," *JCAP* **1608** no. 08, (2016) 044, arXiv:1605.06426 [astro-ph.CO].

[3] E. Pajer and D. van der Woude, "Divergence of Perturbation Theory in Large Scale Structures," arXiv:1710.01736 [astro-ph.CO].

[4] P. Peebles, *The Large-Scale Structure of the Universe*. Princeton University Press, Princeton, New Jersey, 1980.

[5] A. Einstein, "Cosmological Considerations in the General Theory of Relativity," *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* **1917** (1917) 142–152.

[6] E. P. Hubble, "Extragalactic nebulae," *Astrophys. J.* **64** (1926) 321–369.

[7] G. Lemaitre, "A Homogeneous Universe of Constant Mass and Growing Radius Accounting for the Radial Velocity of Extragalactic Nebulae," *Annales Soc. Sci. Bruxelles A* **47** (1927) 49–59. [Gen. Rel. Grav.45,no.8,1635(2013)].

[8] M. S. Longair, *Galaxy Formation*. Astronomy and Astrophysics Library. Springer, Heidelberg, Germany, 2008. http://link.springer.com/book/10.1007%2F978-3-540-73478-9.

[9] T. Buchert *et al.*, "Is there proof that backreaction of inhomogeneities is irrelevant in cosmology?," *Class. Quant. Grav.* **32** (2015) 215021, arXiv:1505.07800 [gr-qc].

[10] S. R. Green and R. M. Wald, "A simple, heuristic derivation of our no backreaction results," *Class. Quant. Grav.* **33** no. 12, (2016) 125027, arXiv:1601.06789 [gr-qc].

[11] D. Baumann, A. Nicolis, L. Senatore, and M. Zaldarriaga, "Cosmological Non-Linearities as an Effective Fluid," *JCAP* **1207** (2012) 051, arXiv:1004.2488 [astro-ph.CO].

[12] S. Weinberg, *Cosmology*. 2008. http://www.oup.com/uk/catalogue/?ci=9780198526827.

[13] D. Baumann, "Inflation," in *Physics of the large and the small, TASI 09, proceedings of the Theoretical Advanced Study Institute in Elementary Particle Physics, Boulder, Colorado, USA, 1-26 June 2009*, pp. 523–686. 2011. arXiv:0907.5424 [hep-th]. https://inspirehep.net/record/827549/files/arXiv:0907.5424.pdf.

[14] S. Dodelson, "Coherent phase argument for inflation," *AIP Conf. Proc.* **689** (2003) 184–196, arXiv:hep-ph/0309057 [hep-ph]. [,184(2003)].

[15] P. Creminelli, S. Dubovsky, D. Lpez Nacir, M. Simonovi, G. Trevisan, G. Villadoro, and M. Zaldarriaga, "Implications of the scalar tilt for the tensor-to-scalar ratio," *Phys. Rev.* **D92** no. 12, (2015) 123528, arXiv:1412.0678 [astro-ph.CO].

[16] N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space*. Cambridge Monographs on Mathematical Physics. Cambridge Univ. Press, Cambridge, UK, 1984. http://www.cambridge.org/mw/academic/subjects/physics/theoretical-physics-and-mathematical-physics/quantum-fields-curved-space?format=PB.

[17] V. Mukhanov and S. Winitzki, *Introduction to quantum effects in gravity*. Cambridge University Press, 2007. http://www.cambridge.org/us/catalogue/catalogue.asp?isbn=0521868343.

[18] S. Weinberg, "Adiabatic modes in cosmology," *Phys. Rev.* **D67** (2003) 123504, arXiv:astro-ph/0302326 [astro-ph].

[19] E. Pajer and S. Jazayeri, "Systematics of Adiabatic Modes: Flat Universes," arXiv:1710.02177 [astro-ph.CO].

[20] S. Endlich, A. Nicolis, and J. Wang, "Solid Inflation," *JCAP* **1310** (2013) 011, arXiv:1210.0569 [hep-th].

[21] R. Bravo, S. Mooij, G. A. Palma, and B. Pradenas, "A generalized non-Gaussian consistency relation for single field inflation," arXiv:1711.02680 [astro-ph.CO].

[22] B. Finelli, G. Goon, E. Pajer, and L. Santoni, "Soft Theorems For Shift-Symmetric Cosmologies," arXiv:1711.03737 [hep-th].

[23] V. Assassi, D. Baumann, and D. Green, "Symmetries and Loops in Inflation," *JHEP* **1302** (2013) 151, arXiv:1210.7792 [hep-th].

[24] V. Assassi, D. Baumann, and D. Green, "On Soft Limits of Inflationary Correlation Functions," *JCAP* **1211** (2012) 047, arXiv:1204.4207 [hep-th].

[25] K. Hinterbichler, L. Hui, and J. Khoury, "An Infinite Set of Ward Identities for Adiabatic Modes in Cosmology," *JCAP* **1401** (2014) 039, arXiv:1304.5527 [hep-th].

[26] **Planck** Collaboration, P. A. R. Ade *et al.*, "Planck 2015 results. XIII. Cosmological parameters," *Astron. Astrophys.* **594** (2016) A13, arXiv:1502.01589 [astro-ph.CO].

[27] S. Weinberg, "Must cosmological perturbations remain non-adiabatic after multi-field inflation?," *Phys. Rev.* **D70** (2004) 083522, arXiv:astro-ph/0405397 [astro-ph].

[28] E. Pajer, G. L. Pimentel, and J. V. S. Van Wijck, "The Conformal Limit of Inflation in the Era of CMB Polarimetry," *JCAP* **1706** no. 06, (2017) 009, arXiv:1609.06993 [hep-th].

[29] 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, arXiv:astro-ph/0112551 [astro-ph].

[30] J. M. Maldacena, "Non-Gaussian Features of Primordial Fluctuations in Single-Field Inflationary Models," *JHEP* **0305** (2003) 013, arXiv:astro-ph/0210603 [astro-ph].

[31] M. Alvarez *et al.*, "Testing Inflation with Large-Scale Structure: Connecting Hopes with Reality," arXiv:1412.4671 [astro-ph.CO].

[32] P. Creminelli and M. Zaldarriaga, "Single field consistency relation for the 3-point function," *JCAP* **0410** (2004) 006, arXiv:astro-ph/0407059 [astro-ph].

[33] C. Cheung, P. Creminelli, A. L. Fitzpatrick, J. Kaplan, and L. Senatore, "The Effective Field Theory of Inflation," *JHEP* **0803** (2008) 014, arXiv:0709.0293 [hep-th].

[34] D. Baumann and D. Green, "Equilateral Non-Gaussianity and New Physics on the Horizon," *JCAP* **1109** (2011) 014, arXiv:1102.5343 [hep-th].

[35] D. Baumann, D. Green, and R. A. Porto, "B-modes and the Nature of Inflation," *JCAP* **1501** no. 01, (2015) 016, arXiv:1407.2621 [hep-th].

[36] R. Flauger, M. Mirbabayi, L. Senatore, and E. Silverstein, "Productive Interactions: heavy particles and non-Gaussianity," *JCAP* **1710** no. 10, (2017) 058, arXiv:1606.00513 [hep-th].

[37] D. Baumann, "Cosmology," http://www.damtp.cam.ac.uk/user/db275/Cosmology/Lectures.pdf.

[38] R. A. Porto, L. Senatore, and M. Zaldarriaga, "The Lagrangian-space Effective Field Theory of Large Scale Structures," *JCAP* **1405** (2014) 022, arXiv:1311.2168 [astro-ph.CO].

[39] L. Senatore and M. Zaldarriaga, "The IR-resummed Effective Field Theory of Large Scale Structures," *JCAP* **1502** no. 02, (2015) 013, arXiv:1404.5954 [astro-ph.CO].

[40] L. Senatore and G. Trevisan, "On the IR-Resummation in the EFTofLSS," arXiv:1710.02178 [astro-ph.CO].

[41] L. Mercolli and E. Pajer, "On the Velocity in the Effective Field Theory of Large-Scale Structures," *JCAP* **1403** (2014) 006, arXiv:1307.3220 [astro-ph.CO].

[42] C. P. Burgess, "Introduction to Effective Field Theory," *Ann. Rev. Nucl. Part. Sci.* **57** (2007) 329–362, arXiv:hep-th/0701053 [hep-th].

[43] E. Pajer and M. Zaldarriaga, "On the Renormalization of the Effective Field Theory of Large-Scale Structures," *JCAP* **1308** (2013) 037, arXiv:1301.7182 [astro-ph.CO].

[44] M. Mirbabayi, F. Schmidt, and M. Zaldarriaga, "Biased Tracers and Time Evolution," arXiv:1412.5169 [astro-ph.CO].

[45] T. Baldauf, L. Mercolli, M. Mirbabayi, and E. Pajer, "The Bispectrum in the Effective Field Theory of Large-Scale Structure," arXiv:1406.4135 [astro-ph.CO].

[46] L. Senatore, "Bias in the Effective Field Theory of Large-Scale Structures," arXiv:1406.7843 [astro-ph.CO].

[47] J. J. M. Carrasco, M. P. Hertzberg, and L. Senatore, "The Effective Field Theory of Cosmological Large-Scale Structures," *JHEP* **1209** (2012) 082, arXiv:1206.2926 [astro-ph.CO].

[48] S. Inagaki, "On the density correlations of fluctuations in an expanding universe. I," *Astronomical Society of Japan* **28** (1976) 77–87.

[49] S. Pueblas and R. Scoccimarro, "Generation of Vorticity and Velocity Dispersion by Orbit Crossing," *Phys. Rev.* **D80** (2009) 043504, arXiv:0809.4606 [astro-ph].

[50] F. Bernardeau, "The Evolution of the Large-Scale Structure of the Universe: Beyond the Linear Regime," arXiv:1311.2724 [astro-ph.CO].

[51] J. J. M. Carrasco, S. Foreman, D. Green, and L. Senatore, "The Two-Loop Matter Power Spectrum and the IR-Safe Integrand," *JCAP* **1407** (2014) 056, arXiv:1304.4946 [astro-ph.CO].

[52] R. E. Angulo, S. Foreman, M. Schmittfull, and L. Senatore, "The One-Loop Matter Bispectrum in the Effective Field Theory of Large-Scale Structures," arXiv:1406.4143 [astro-ph.CO].

[53] P. McDonald, "Clustering of Dark Matter Tracers: Renormalizing the Bias Parameters," *Phys.Rev.* **D74** (2006) 103512, arXiv:astro-ph/0609413 [astro-ph].

[54] P. McDonald and A. Roy, "Clustering of Dark Matter Tracers: Generalizing Bias for the Coming Era of Precision LSS," *JCAP* **0908** (2009) 020, arXiv:0902.0991 [astro-ph.CO].

[55] F. Schmidt, D. Jeong, and V. Desjacques, "Peak-Background Split, Renormalization, and Galaxy Clustering," *Phys.Rev.* **D88** no. 2, (2013) 023515, arXiv:1212.0868 [astro-ph.CO].

[56] V. Assassi, D. Baumann, D. Green, and M. Zaldarriaga, "Renormalized Halo Bias," *JCAP* **1408** (2014) 056, arXiv:1402.5916 [astro-ph.CO].

[57] R. Angulo, M. Fasiello, L. Senatore, and Z. Vlah, "On the Statistics of Biased Tracers in the Effective Field Theory of Large-Scale Structures," arXiv:1503.08826 [astro-ph.CO].

[58] L. Senatore and M. Zaldarriaga, "Redshift Space Distortions in the Effective Field Theory of Large-Scale Structures," arXiv:1409.1225 [astro-ph.CO].

[59] M. Lewandowski, A. Perko, and L. Senatore, "Analytic Prediction of Baryonic Effects from the EFT of Large-Scale Structures," arXiv:1412.5049 [astro-ph.CO].

[60] V. Assassi, D. Baumann, E. Pajer, Y. Welling, and D. van der Woude, *work in progress*.

[61] A. Lewis, A. Challinor, and A. Lasenby, "Efficient Computation of CMB Anisotropies in Closed FRW Models," *Astrophys.J.* **538** (2000) 473–476, arXiv:astro-ph/9911177 [astro-ph].

[62] F. Schmidt and M. Kamionkowski, "Halo Clustering with Non-Local Non-Gaussianity," *Phys.Rev.* **D82** (2010) 103002, arXiv:1008.0638 [astro-ph.CO].

[63] J. Chluba, J. Hamann, and S. P. Patil, "Features and New Physical Scales in Primordial Observables: Theory and Observation," arXiv:1505.01834 [astro-ph.CO].

[64] E. Komatsu and D. N. Spergel, "Acoustic Signatures in the Primary Microwave Background Bispectrum," *Phys.Rev.* **D63** (2001) 063002, arXiv:astro-ph/0005036 [astro-ph].

[65] M. Alishahiha, E. Silverstein, and D. Tong, "DBI in the Sky," *Phys.Rev.* **D70** (2004) 123505, arXiv:hep-th/0404084 [hep-th].

[66] X. Chen and Y. Wang, "Quasi-Single-Field Inflation and Non-Gaussianities," *JCAP* **1004** (2010) 027, arXiv:0911.3380 [hep-th].

[67] D. Green, M. Lewandowski, L. Senatore, E. Silverstein, and M. Zaldarriaga, "Anomalous Dimensions and Non-Gaussianity," *JHEP* **1310** (2013) 171, arXiv:1301.2630.

[68] M. Shiraishi, E. Komatsu, M. Peloso, and N. Barnaby, "Signatures of Anisotropic Sources in the Squeezed-Limit Bispectrum of the Cosmic Microwave Background," *JCAP* **1305** (2013) 002, arXiv:1302.3056 [astro-ph.CO].

[69] N. Arkani-Hamed and J. Maldacena, "Cosmological Collider Physics," arXiv:1503.08043 [hep-th].

[70] N. Barnaby, R. Namba, and M. Peloso, "Observable Non-Gaussianity from Gauge Field Production in Slow-Roll Inflation, and a Challenging Connection with Magnetogenesis," *Phys.Rev.* **D85** (2012) 123523, arXiv:1202.1469 [astro-ph.CO].

[71] M. Shiraishi, "Parity Violation of Primordial Magnetic Fields in the CMB Bispectrum," *JCAP* **1206** (2012) 015, arXiv:1202.2847 [astro-ph.CO].

[72] M. Shiraishi, D. Nitta, S. Yokoyama, and K. Ichiki, "Optimal Limits on Primordial Magnetic Fields from CMB Temperature Bispectrum of Passive Modes," *JCAP* **1203** (2012) 041, arXiv:1201.0376 [astro-ph.CO].

[73] 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, arXiv:1310.0464 [astro-ph.CO].

[74] R. Scoccimarro and J. Frieman, "Loop Corrections in Nonlinear Cosmological Perturbation Theory," *Astrophys.J.Suppl.* **105** (1996) 37, arXiv:astro-ph/9509047 [astro-ph].

[75] **Planck** Collaboration, P. Ade *et al.*, "Planck 2015 Results. XVII. Constraints on Primordial Non-Gaussianity," arXiv:1502.01592 [astro-ph.CO].

[76] E. Sefusatti, J. R. Fergusson, X. Chen, and E. Shellard, "Effects and Detectability of Quasi-Single-Field Inflation in the Large-Scale Structure and Cosmic Microwave Background," *JCAP* **1208** (2012) 033, arXiv:1204.6318 [astro-ph.CO].

[77] D. Baumann and D. Green, "Signatures of Supersymmetry from the Early Universe," *Phys.Rev.* **D85** (2012) 103520, arXiv:1109.0292 [hep-th].

[78] E. Sefusatti, M. Crocce, and V. Desjacques, "The Matter Bispectrum in N-body Simulations with Non-Gaussian Initial Conditions," *Mon.Not.Roy.Astron.Soc.* **406** (2010) 1014–1028, arXiv:1003.0007 [astro-ph.CO].

[79] C. D. Rimes and A. J. Hamilton, "Information Content of the Nonlinear Matter Power Spectrum," *Mon.Not.Roy.Astron.Soc.* **360** (2005) L82–L86, arXiv:astro-ph/0502081 [astro-ph].

[80] M. Crocce and R. Scoccimarro, "Memory of Initial Conditions in Gravitational Clustering," *Phys.Rev.* **D73** (2006) 063520, arXiv:astro-ph/0509419 [astro-ph].

[81] N. Dalal, O. Dore, D. Huterer, and A. Shirokov, "The Imprints of Primordial Non-Gaussianities on Large-Scale Structure: Scale-Dependent Bias and Abundance of Virialized Objects," *Phys.Rev.* **D77** (2008) 123514, arXiv:0710.4560 [astro-ph].

[82] A. Lewis, "The Real Shape of Non-Gaussianities," *JCAP* **1110** (2011) 026, arXiv:1107.5431 [astro-ph.CO].

[83] O. Doré *et al.*, "Cosmology with the SPHEREX All-Sky Spectral Survey," arXiv:1412.4872 [astro-ph.CO].

[84] M. Tellarini, A. J. Ross, G. Tasinato, and D. Wands, "Galaxy bispectrum, primordial non-Gaussianity and redshift space distortions," arXiv:1603.06814 [astro-ph.CO].

[85] R. Scoccimarro, E. Sefusatti, and M. Zaldarriaga, "Probing primordial non-Gaussianity with large - scale structure," *Phys. Rev.* **D69** (2004) 103513, arXiv:astro-ph/0312286 [astro-ph].

[86] E. Sefusatti and E. Komatsu, "The bispectrum of galaxies from high-redshift galaxy surveys: Primordial non-Gaussianity and non-linear galaxy bias," *Phys. Rev.* **D76** (2007) 083004, arXiv:0705.0343 [astro-ph].

[87] T. Baldauf, U. Seljak, and L. Senatore, "Primordial non-Gaussianity in the Bispectrum of the Halo Density Field," *JCAP* **1104** (2011) 006, arXiv:1011.1513 [astro-ph.CO].

[88] D. Jeong and E. Komatsu, "Primordial non-Gaussianity, scale-dependent bias, and the bispectrum of galaxies," *Astrophys. J.* **703** (2009) 1230–1248, arXiv:0904.0497 [astro-ph.CO].

[89] G. Tasinato, M. Tellarini, A. J. Ross, and D. Wands, "Primordial non-Gaussianity in the bispectra of large-scale structure," *JCAP* **1403** (2014) 032, arXiv:1310.7482 [astro-ph.CO].

[90] N. Roth and C. Porciani, "Can we really measure f_{NL} from the galaxy power spectrum?," *Mon. Not. Roy. Astron. Soc.* **425** (2012) L81–L85, arXiv:1205.3165 [astro-ph.CO].

[91] T. Baldauf, M. Mirbabayi, M. Simonovi, and M. Zaldarriaga, "LSS constraints with controlled theoretical uncertainties," arXiv:1602.00674 [astro-ph.CO].

[92] **Planck** Collaboration, P. A. R. Ade *et al.*, "Planck 2015 results. XVII. Constraints on primordial non-Gaussianity," arXiv:1502.01592 [astro-ph.CO].

[93] A. Lewis and S. Bridle, "Cosmological parameters from CMB and other data: a Monte- Carlo approach," *Phys. Rev.* **D66** (2002) 103511, astro-ph/0205436.

[94] N. Bartolo, S. Matarrese, and A. Riotto, "Nongaussianity from inflation," *Phys. Rev.* **D65** (2002) 103505, arXiv:hep-ph/0112261 [hep-ph].

[95] P. Creminelli, "On non-Gaussianities in single-field inflation," *JCAP* **0310** (2003) 003, arXiv:astro-ph/0306122 [astro-ph].

[96] E. Sefusatti, M. Crocce, S. Pueblas, and R. Scoccimarro, "Cosmology and the Bispectrum," *Phys. Rev.* **D74** (2006) 023522, arXiv:astro-ph/0604505 [astro-ph].

[97] 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.* **D92** no. 12, (2015) 123007, arXiv:1507.02256 [astro-ph.CO].

[98] S. Foreman, H. Perrier, and L. Senatore, "Precision Comparison of the Power Spectrum in the EFTofLSS with Simulations," arXiv:1507.05326 [astro-ph.CO].

[99] A. Heavens, "Statistical techniques in cosmology," arXiv:0906.0664 [astro-ph.CO].

[100] E. Pajer and Y. Welling, "Time dependence in correlation functions,".

[101] R. Scoccimarro, S. Colombi, J. N. Fry, J. A. Frieman, E. Hivon, and A. Melott, "Nonlinear evolution of the bispectrum of cosmological

perturbations," *Astrophys. J.* **496** (1998) 586, arXiv:astro-ph/9704075 [astro-ph].

[102] C. E. Powell, "Generating Realisations of Stationary Gaussian Random Fields by Circulant Embedding." https://www.nag.co.uk/doc/techrep/pdf/tr1_14.pdf.

[103] **EUCLID** Collaboration, R. Laureijs *et al.*, "Euclid Definition Study Report," arXiv:1110.3193 [astro-ph.CO].

[104] **BOSS** Collaboration, K. S. Dawson *et al.*, "The Baryon Oscillation Spectroscopic Survey of SDSS-III," *Astron. J.* **145** (2013) 10, arXiv:1208.0022 [astro-ph.CO].

[105] K. S. Dawson *et al.*, "The SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Overview and Early Data," *Astron. J.* **151** (2016) 44, arXiv:1508.04473 [astro-ph.CO].

[106] **DESI** Collaboration, M. Levi *et al.*, "The DESI Experiment, a whitepaper for Snowmass 2013," arXiv:1308.0847 [astro-ph.CO].

[107] M. Garny, T. Konstandin, R. A. Porto, and L. Sagunski, "On the Soft Limit of the Large Scale Structure Power Spectrum: UV Dependence," *JCAP* **1511** no. 11, (2015) 032, arXiv:1508.06306 [astro-ph.CO].

[108] E. Sefusatti and R. Scoccimarro, "Galaxy bias and halo-occupation numbers from large-scale clustering," *Phys. Rev.* **D71** (2005) 063001, arXiv:astro-ph/0412626 [astro-ph].

[109] U. Seljak, "Extracting primordial non-gaussianity without cosmic variance," *Phys. Rev. Lett.* **102** (2009) 021302, arXiv:0807.1770 [astro-ph].

[110] L. Verde and A. F. Heavens, "On the trispectrum as a Gaussian test for cosmology," *Astrophys. J.* **553** (2001) 14, arXiv:astro-ph/0101143 [astro-ph].

[111] D. Bertolini, K. Schutz, M. P. Solon, and K. M. Zurek, "The Trispectrum in the Effective Field Theory of Large Scale Structure," arXiv:1604.01770 [astro-ph.CO].

[112] A. Cooray, "21-cm Background Anisotropies Can Discern Primordial Non-Gaussianity," *Phys. Rev. Lett.* **97** (2006) 261301, arXiv:astro-ph/0610257 [astro-ph].

[113] K. Heitmann, M. White, C. Wagner, S. Habib, and D. Higdon, "The Coyote Universe I: Precision Determination of the Nonlinear Matter Power Spectrum," *Astrophys. J.* **715** (2010) 104–121, arXiv:0812.1052 [astro-ph].

[114] A. Schneider, R. Teyssier, D. Potter, J. Stadel, J. Onions, D. S. Reed, R. E. Smith, V. Springel, F. R. Pearce, and R. Scoccimarro, "Matter power spectrum and the challenge of percent accuracy," *JCAP* **1604** no. 04, (2016) 047, arXiv:1503.05920 [astro-ph.CO].

[115] F. J. Dyson, "Divergence of perturbation theory in quantum electrodynamics," *Phys. Rev.* **85** (1952) 631–632.

[116] M. McQuinn and M. White, "Cosmological perturbation theory in 1+1 dimensions," *JCAP* **1601** no. 01, (2016) 043, arXiv:1502.07389 [astro-ph.CO].

[117] D. Blas, M. Garny, and T. Konstandin, "Cosmological perturbation theory at three-loop order," *JCAP* **1401** no. 01, (2014) 010, arXiv:1309.3308 [astro-ph.CO].

[118] N. Afshordi, "How well can (renormalized) perturbation theory predict dark matter clustering properties?," *Phys. Rev.* **D75** (2007) 021302, arXiv:astro-ph/0610336 [astro-ph].

[119] P. Valageas, "Accuracy of analytical models of the large-scale matter distribution," *Phys. Rev.* **D88** no. 8, (2013) 083524, arXiv:1308.6755 [astro-ph.CO].

[120] I. Mohammed and U. Seljak, "Analytic model for the matter power spectrum, its covariance matrix, and baryonic effects," *Mon. Not. Roy. Astron. Soc.* **445** no. 4, (2014) 3382–3400, arXiv:1407.0060 [astro-ph.CO].

[121] U. Seljak and Z. Vlah, "Halo Zeldovich model and perturbation theory: Dark matter power spectrum and correlation function,"

Phys. Rev. **D91** no. 12, (2015) 123516, arXiv:1501.07512 [astro-ph.CO].

[122] P. Valageas, "Using the Zeldovich dynamics to test expansion schemes," *Astron. Astrophys.* (2007) , arXiv:0706.2593 [astro-ph]. [Astron. Astrophys.476,31(2007)].

[123] P. Valageas, "Quasi-linear regime and rare-event tails of decaying Burgers turbulence," *Phys. Rev.* **E80** (2009) 016305, arXiv:0905.1910 [cond-mat.stat-mech].

[124] S. Tassev and M. Zaldarriaga, "Estimating CDM Particle Trajectories in the Mildly Non-Linear Regime of Structure Formation. Implications for the Density Field in Real and Redshift Space," *JCAP* **1212** (2012) 011, arXiv:1203.5785 [astro-ph.CO].

[125] P. Valageas, "Impact of shell crossing and scope of perturbative approaches in real and redshift space," *Astron. Astrophys.* **526** (2011) A67, arXiv:1009.0106 [astro-ph.CO].

[126] P. Valageas, "Dynamics of gravitational clustering I. building perturbative expansions," *Astron. Astrophys.* **379** (2001) 8, arXiv:astro-ph/0107015 [astro-ph].

[127] T. Baldauf, E. Schaan, and M. Zaldarriaga, "On the reach of perturbative descriptions for dark matter displacement fields," *JCAP* **1603** no. 03, (2016) 017, arXiv:1505.07098 [astro-ph.CO].

[128] T. Tatekawa, "Improving of the Lagrangian perturbative solution for cosmic fluid: Applying Shanks transformation," *Phys. Rev.* **D75** (2007) 044028, arXiv:astro-ph/0605250 [astro-ph].

[129] P. McDonald and Z. Vlah, "Large-scale structure perturbation theory without losing stream crossing," arXiv:1709.02834 [astro-ph.CO].

[130] M. Mariño, *Instantons and Large N*. Cambridge University Press, 2015. http://www.cambridge.org/mw/academic/subjects/physics/theoretical-physics-and-mathematical-physics/instantons-and-large-n-introduction-non-perturbative-methods-quantum-field-theory?format=HB.

[131] S. Foreman, "Notes on Perturbation Theory in 1+1 Dimensions,". Unpublished manuscript.

[132] F. Bernardeau, C. Pichon, and S. Codis, "Statistics of cosmic density profiles from perturbation theory," *Phys. Rev.* **D90** no. 10, (2014) 103519, arXiv:1310.8134 [astro-ph.CO].

[133] Ya. B. Zeldovich, "Gravitational instability: An Approximate theory for large density perturbations," *Astron. Astrophys.* **5** (1970) 84–89.

[134] S. F. Schandarin and Ya. B. Zeldovich, "The Large scale structure of the universe: Turbulence, intermittency, structures in a selfgravitating medium," *Rev. Mod. Phys.* **61** (1989) 185–220.

[135] V. Sahni and P. Coles, "Approximation methods for nonlinear gravitational clustering," *Phys. Rept.* **262** (1995) 1–135, arXiv:astro-ph/9505005 [astro-ph].

[136] B. Grinstein and M. B. Wise, "On the Validity of the Zeldovich Approximation," *Astrophys. J.* **320** (1987) 448.

[137] M. C. Neyrinck, I. Szapudi, and A. S. Szalay, "Rejuvenating the matter power spectrum: restoring information with a logarithmic density mapping," *Astrophys. J.* **698** (2009) L90–L93, arXiv:0903.4693 [astro-ph.CO].

[138] M. C. Neyrinck, I. Szapudi, and A. S. Szalay, "Rejuvenating Power Spectra II: the Gaussianized galaxy density field," *Astrophys. J.* **731** (2011) 116, arXiv:1009.5680 [astro-ph.CO].

[139] P. Coles and B. Jones, "A Lognormal model for the cosmological mass distribution," *Mon. Not. Roy. Astron. Soc.* **248** (1991) 1–13.

[140] X. Wang, M. Neyrinck, I. Szapudi, A. Szalay, X. Chen, J. Lesgourgues, A. Riotto, and M. Sloth, "Perturbation Theory of the Cosmological Log-Density Field," *Astrophys. J.* **735** (2011) 32, arXiv:1103.2166 [astro-ph.CO].

[141] J. Carron, "On the incompleteness of the moment and correlation function hierarchy as probes of the lognormal field," *Astrophys. J.* **738** (2011) 86, arXiv:1105.4467 [astro-ph.CO].

[142] A. Hall and A. Mead, "Perturbative Gaussianizing transforms for cosmological fields," arXiv:1709.03924 [astro-ph.CO].

[143] F. Simpson, J. B. James, A. F. Heavens, and C. Heymans, "Clipping the Cosmos: The Bias and Bispectrum of Large Scale Structure," *Phys. Rev. Lett.* **107** (2011) 271301, arXiv:1107.5169 [astro-ph.CO].

[144] F. Simpson, A. F. Heavens, and C. Heymans, "Clipping the cosmos. II. Cosmological information from nonlinear scales," *Phys. Rev.* **D88** no. 8, (2013) 083510, arXiv:1306.6349 [astro-ph.CO].

[145] V. Mukhanov, *Physical Foundations of Cosmology*. Cambridge University Press, Oxford, 2005. http://www-spires.fnal.gov/spires/find/books/www?cl=QB981.M89::2005.

[146] V. Sahni and S. Shandarin, "Behavior of Lagrangian approximations in spherical voids," *Mon. Not. Roy. Astron. Soc.* **282** (1996) 641, arXiv:astro-ph/9510142 [astro-ph].

[147] C. M. Bender and S. A. Orszag, "Advanced mathematical methods for scientists and engineers," *McGraw-Hill* (1978) .

[148] K. B. Fisher and A. Nusser, "The nonlinear redshift space power spectrum: omega from redshift surveys," *Mon. Not. Roy. Astron. Soc.* **279** (1996) L1, arXiv:astro-ph/9510049 [astro-ph].

[149] J. Carlson, B. Reid, and M. White, "Convolution Lagrangian perturbation theory for biased tracers," *Mon. Not. Roy. Astron. Soc.* **429** (2013) 1674, arXiv:1209.0780 [astro-ph.CO].

[150] M. Crocce and R. Scoccimarro, "Renormalized cosmological perturbation theory," *Phys. Rev.* **D73** (2006) 063519, arXiv:astro-ph/0509418 [astro-ph].

[151] C. Uhlemann, S. Codis, C. Pichon, F. Bernardeau, and P. Reimberg, "Back in the saddle: Large-deviation statistics of the cosmic log-density field," *Mon. Not. Roy. Astron. Soc.* **460** no. 2, (2016) 1529–1541, arXiv:1512.05793 [astro-ph.CO].

[152] F. Bernardeau, "The Gravity induced quasi-Gaussian correlation hierarchy," *Astrophys. J.* **392** (1992) 1–14.

[153] F. Bernardeau, "Skewness and Kurtosis in large scale cosmic fields," *Astrophys. J.* **433** (1994) 1, arXiv:astro-ph/9312026 [astro-ph].

[154] F. Bernardeau and L. Kofman, "Properties of the cosmological density distribution function," *Astrophys. J.* **443** (1995) 479–498, arXiv:astro-ph/9403028 [astro-ph].

[155] F. Bernardeau, "The Effects of smoothing on the statistical properties of large scale cosmic fields," *Astron. Astrophys.* **291** (1994) 697, arXiv:astro-ph/9403020 [astro-ph].

[156] F. Bernardeau and P. Valageas, "Construction of the one-point pdf of the local aperture mass in weak lensing maps," *Astron. Astrophys.* **364** (2000) 1, arXiv:astro-ph/0006270 [astro-ph].

[157] L. Kofman, E. Bertschinger, J. M. Gelb, A. Nusser, and A. Dekel, "Evolution of one point distributions from Gaussian initial fluctuations," *Astrophys. J.* **420** (1994) 44–57, arXiv:astro-ph/9311028 [astro-ph].

[158] Z. A. M. Protogeros and R. J. Scherrer, "Local Lagrangian approximations for the evolution of the density distribution function in large scale structure," *Mon. Not. Roy. Astron. Soc.* **284** (1997) 425, arXiv:astro-ph/9603155 [astro-ph].

[159] P. Fosalba and E. Gaztanaga, "Cosmological perturbation theory and the spherical collapse model: Part 1. Gaussian initial conditions," *Mon. Not. Roy. Astron. Soc.* **301** (1998) 503–523, arXiv:astro-ph/9712095 [astro-ph].

[160] Y. Ohta, I. Kayo, and A. Taruya, "Cosmological density distribution function from the ellipsoidal collapse model in real space," *Astrophys. J.* **608** (2004) 647–662, arXiv:astro-ph/0402618 [astro-ph].

[161] E. Gaztanaga and R. A. C. Croft, "Predictions for the clustering properties of the lyman-alpha forest - I. one-point statistics," *Mon. Not. Roy. Astron. Soc.* **309** (1999) 885, arXiv:astro-ph/9811480 [astro-ph].

[162] T. Y. Lam and R. K. Sheth, "Perturbation theory and excursion set estimates of the probability distribution function of dark matter, and a method for reconstructing the initial distribution function," *Mon. Not. Roy. Astron. Soc.* **386** (2008) 407, arXiv:0711.5029 [astro-ph].

[163] E. Sellentin, A. H. Jaffe, and A. F. Heavens, "On the use of the Edgeworth expansion in cosmology I: how to foresee and evade its pitfalls," arXiv:1709.03452 [astro-ph.CO].

[164] A. Cooray and R. K. Sheth, "Halo models of large scale structure," *Phys. Rept.* **372** (2002) 1–129, arXiv:astro-ph/0206508 [astro-ph].

[165] U. Seljak, "Analytic model for galaxy and dark matter clustering," *Mon. Not. Roy. Astron. Soc.* **318** (2000) 203, arXiv:astro-ph/0001493 [astro-ph].

[166] C.-P. Ma and J. N. Fry, "Deriving the nonlinear cosmological power spectrum and bispectrum from analytic dark matter halo profiles and mass functions," *Astrophys. J.* **543** (2000) 503–513, arXiv:astro-ph/0003343 [astro-ph].

[167] J. F. Navarro, C. S. Frenk, and S. D. M. White, "The Structure of cold dark matter halos," *Astrophys. J.* **462** (1996) 563–575, arXiv:astro-ph/9508025 [astro-ph].

[168] F. Schmidt, "Towards a self-consistent halo model for the nonlinear large-scale structure," *Phys. Rev.* **D93** no. 6, (2016) 063512, arXiv:1511.02231 [astro-ph.CO].

[169] P. Valageas and T. Nishimichi, "Combining perturbation theories with halo models," *Astron. Astrophys.* **527** (2011) A87, arXiv:1009.0597 [astro-ph.CO].

[170] V. Assassi, D. Baumann, and F. Schmidt, "Galaxy Bias and Primordial Non-Gaussianity," *JCAP* **1512** no. 12, (2015) 043, arXiv:1510.03723 [astro-ph.CO].

[171] A. Perko, L. Senatore, E. Jennings, and R. H. Wechsler, "Biased Tracers in Redshift Space in the EFT of Large-Scale Structure," arXiv:1610.09321 [astro-ph.CO].

[172] A. Barreira and F. Schmidt, "Responses in Large-Scale Structure," *JCAP* **1706** no. 06, (2017) 053, arXiv:1703.09212 [astro-ph.CO].

Nederlandse Samenvatting

Het uiteindelijke doel van de kosmologie, en daarmee ook van dit proefschrift, is het begrijpen van het ontstaan van het heelal. Het gaat daarbij niet zozeer om de details van de ingewikkelde dynamiek van een specifiek sterrenstelsel, als wel om de algemene begincondities die nodig waren om het ontstaan van ons huidige, indrukwekkende heelal mogelijk te maken.

In zekere zin is deze zoektocht vergelijkbaar met de zoektocht naar de condities waaronder de omslagtekening van dit proefschrift is ontstaan. We zouden ons bijvoorbeeld kunnen afvragen of de kunstenares, Janet Towbin, links- of rechtshandig is. Of hoe groot het vel was waarop ze de tekening maakte.

Voor veel van dat soort vragen ligt het voor de hand een statistische analyse te doen. Laten we bijvoorbeeld even aannemen dat linkshandigen gemiddeld vaker de neiging hebben een rechte lijn enigszins naar links af te laten buigen. Om dit te testen zouden we dan alle lijnen in de tekening nauwkeurig langs een liniaal kunnen leggen en kijken of ze vaker naar links dan naar recht afbuigen en daaruit concluderen of Janet linkshandig is.

Om achter de grootte van het vel te komen kunnen we iets vergelijkbaars doen. Laten we nu aannemen dat de tekening is opgebouwd uit kleinere deelstructuren die te tekenen zijn zonder de arm op te tillen. Dat zou zich dan waarschijnlijk vertalen naar een terugkerend patroon (ik herken op de achterzijde bijvoorbeeld een terugkerend cirkelpatroon) ter grootte van die typische afstand (zeg een paar centimeter). Door die typische afstand te vergelijken met de totale lengte van de tekening zouden we nu een schatting kunnen maken van de totale grootte van het vel.

In het geval van de tekening zijn dit natuurlijk maar een paar wilde suggesties, waarvan ik geen idee heb of ze ergens op slaan. Gelukkig weet ik dat in het geval van het heelal een stuk beter. Het specifieke patroon waar we in dit proefschrift naar op zoek zijn heet Primordial non-Gaussianity (PNG) en vertelt ons direct iets over de condities waaronder het heelal is ontstaan. Het grootste deel van dit proefschrift gaat echter over een extra complicatie die optreedt in het geval van het heelal: als gevolg van de zwaartekracht

vervormt het plaatje zoals dat in het begin is 'getekend' aanzienlijk.

Het is dus zaak deze vervorming door de zwaartekracht zo goed mogelijk ongedaan te maken. Daar gaat de Effectieve Veldentheorie van de Groteschaalstructuur (EFT of LSS) over. Iets populairder wordt deze theorie ook wel het 'gesmolten-chocolade-model van het heelal' genoemd om de volgende reden. Om de PNG patronen te herkennen is het niet nodig heel precies alle details van de evolutie van het heelal te begrijpen. Op eenzelfde manier is het om de beweging van gesmolten chocolade (of gewoon water, of welke vloeistof dan ook) goed te begrijpen niet nodig de details van de individuele moleculen te begrijpen. Het idee is dus dat we, als we maar genoeg uitzoomen, het geheel aan sterrenstelsels ook als een soort continue, deinende vloeistof kunnen beschrijven. Interessant genoeg blijkt deze vloeistofbeschrijving goed te werken en gedraagt het heelal zich inderdaad als gesmolten chocolade!

Hieronder staat in meer detail beschreven wat de EFT of LSS en PNG zijn en welke bijdrage ik precies met mijn onderzoek heb geleverd.

De Effectieve Veldentheorie van de Groteschaalstructuur

Het centrale begrip in de Effectieve Veldentheorie van de Groteschaalstructuur (EFT of LSS) is de materiedichtheid: hoeveel materie vinden we per kubieke meter op een bepaalde plek. Om dit begrip enigszins zinnig te maken is het echter nodig een gemiddelde te nemen. De dichtheid op punt x kan bijvoorbeeld gevonden worden door alle massa (in de vorm van bijvoorbeeld sterrenstelsels) binnen een straal r rondom punt x op te tellen, en te delen door het volume van een bol met straal r. De EFT of LSS gaat over de evolutie van de dichtheidsverdeling in de ruimte, ook wel het dichtheidsveld genaamd.

Om de evolutie daarvan goed te beschrijven is het bovendien nodig het snelheidsveld te introduceren: met een vergelijkbaar gemiddelde definiëren we de snelheid op punt x door het gemiddelde van de snelheden van alle materie in een straal r rondom dit punt te nemen.

Vervolgens is het de vraag welke vergelijkingen de evolutie van dit dichtheids- en snelheidsveld beschrijven. Zoals vaak in de natuurkunde is, enigszins paradoxaal, de beste manier om te leren hoe een systeem verandert juist door te vragen wat er tijdens de evolutie hetzelfde blijft. We vragen ons af wat de zogenaamde behouden grootheden zijn. In ons geval zijn er twee cru-

ciaal: behoud van massa en behoud van impuls. Behoud van massa betekent dat als de dichtheid op ene plek kleiner wordt, hij op een andere plek moet toenemen - massa kan niet verdwijnen, slechts verplaatsen. Dit geeft direct de eerste vergelijking van de EFT of LSS. Dan impulsbehoud. Dit is eigenlijk niks anders dan de tweede hoofdwet van Newton: om de impuls (snelheid keer massa) te veranderen is een kracht nodig. Voor de evolutie van het dichtheidsveld is er nu maar één kracht relevant: de zwaartekracht. Dit geeft de tweede, en laatste vergelijking om de evolutie van het dichtheidsveld te beschrijven.

De interessante observatie is nu dat deze vergelijkingen precies de vergelijkingen van een vloeistof zijn. Net als voor een vloeistof kunnen we dus vragen of ook het heelal gekarakteriseerd wordt door een voortplantingssnelheid en viscositeit. Tot de introductie van de EFT of LSS werden deze snelheids- en viscositeitsparameter verwaarloosd, wat inderdaad een goede benadering is op de allergrootste schalen, maar slechter wordt naarmate men ook dichtheidsverschillen op kleinere schalen probeert te beschrijven.

De kracht van de EFT of LSS is dat hij a priori aanneemt dat een deel van de theorie (het gedrag op kleine schaal) nou eenmaal buiten ons bereik ligt en te ingewikkeld is goed te beschrijven. Het effect van deze onwetendheid op grote schalen blijkt echter beperkt en op een hele specifieke manier te karakteriseren. Kernidee is hierbij lokaliteit: de details van wat er lokaal op kleine schaal gebeurt kunnen nooit het dichtheidsveld ver weg significant beïnvloeden. Omgekeerd wordt wat op kleine schaal gebeurt alleen beïnvloed door de zogeheten lokale observabelen. De natuurkundige symmetrieën vertellen ons dan heel precies wat lokaal te observeren is. In dit geval zijn dat alle termen gemaakt van tweede en hogere afgeleiden van de Newtonpotentiaal en producten daarvan.

Wiskundig leidt dit tot een reeks nieuwe termen in de vloeistofvergelijkingen, met onbekende coëfficiënten ervoor. Zo komen we tot een theorie die consistent alle onbekendheden in de theorie parametriseert, maar wel ten koste van de invoering van enkele onbekende parameters in de theorie. Laat ik hier echter benadrukken dat dit laatste onvermijdelijk is: er is nou eenmaal feedback van de ingewikkelde dynamiek op kleine schaal die we niet kunnen uitrekenen. Een volledige en overzichtelijke parametrisatie van hun effect is het maximaal haalbare. En dat is precies wat de EFT of LSS voor ons doet.

Niet-Gaussische begincondities in de Effectieve Veldentheorie van de Groteschaalstructuur

Het stuk hierboven gaat over de vergelijkingen voor de evolutie van het dichtheidsveld. Dat zegt nog niks over de begincondities: met welke dichtheidsverdeling begon het allemaal? Daar gaat deze sectie over.

In zekere zin is het hele doel van een goede beschrijving van de dichtheidsverdeling juist om deze begincondities te achterhalen. De statistische eigenschappen daarvan vertellen ons namelijk iets over het ontstaan van het heelal. Dit eerste deel van het proefschrift gaat echter over net iets anders: voor zogenaamde niet-Gaussische begincondities (afgekort PNG, naar het Engels) moeten zelfs de bewegingsvergelijkingen enigszins worden aangepast. De uitleg daarvan is het doel van dit stuk.

Voor zover we nu weten zijn de begincondities Gaussisch. Dit betekent dat alle statistische informatie over de massaverdeling te geven is in termen van de correlatie tussen de dichtheid op twee verschillende plekken. Deze correlatie is als volgt te berekenen. Neem een gegeven afstand r. Neem nu twee punten in het universum met onderlinge afstand r en vermenigvuldig de dichtheid. Herhaal dit nu voor alle paren van punten in het universum met onderlinge afstand r. Het gemiddelde hiervan is de zogenaamde tweepuntscorrelatie voor afstand r. In een Gaussisch universum zou een vergelijkbare oefening met drie punten altijd nul opleveren.

Voor ons heeft de Gaussische aanname nog een ander belangrijk gevolg: het betekent dat de kleine-schaal-begincondities op plaats x onafhankelijk zijn van wat er een stukje verderop op locatie y gebeurt. Zoals in de vorige subsectie beschreven, is dit essentieel voor de EFT of LSS: de onbekende details op plaats x mogen alleen afhankelijk zijn van wat er lokaal op plaats x gebeurt. De volgende vraag doet zich dus voor: kan dit 'lokaliteitsprincipe' doorbroken worden voor niet-Gaussische begincondities?

Het antwoord hierop is ja: voor een interessante deelverzameling van modellen voor het begin van het heelal kunnen de begincondities op plaats x afhangen van de dichtheid op punt y. Om precies te zijn, ze kunnen lokaal afhangen van de Newtonpotentiaal zelf, iets wat in strijd lijkt met het equivalentieprincipe van Einstein dat zegt dat de Newtonpotentiaal lokaal niet te observeren is. Dit is echter precies het punt: voor zogenaamde meerderevelden-modellen van het begin van het heelal wordt deze (versimpelde) versie van het equivalentieprincipe gebroken. De vraag die ons werk in hoofdstuk

4 beantwoordt is dan ook: hoe moet dit gebrek aan lokaliteit in de theorie worden ingebouwd?

We laten zien dat dit goed mogelijk is, met name vanwege het feit dat de niet-lokaliteit echt alleen door de begincondities kan komen, wat ons een duidelijk voorschrift geeft van hoe ze verder meegenomen dienen te worden. We laten in dit werk zien hoe de nieuwe theorie net zo volledig en consistent is als de oorspronkelijke EFT of LSS, al moeten er een paar nieuwe, onbekende coëfficiënten aan de theorie worden toegevoegd.

Het zichtbaar maken van niet-Gaussische begincondities tegen een achtergrond vol ruis

Hoofdstuk 5 is een numerieke analyse van hoe goed we met onze theorie verwachten dat realistische toekomstige experimenten in staat zijn ons iets te vertellen over de niet-Gaussische begincondities (PNG). Het kernwoord daarbij is informatie: omdat niet-Gaussische begincondities hoe dan ook een heel klein statistisch effect op de uiteindelijke dichtheidsverdeling hebben, is het essentieel dat we heel veel data over de verdeling van sterren kunnen verzamelen, en vervolgens heel goed kunnen ontwarren welk gedeelte echt door de begincondities komt en welke aspecten in de data 'ruis' zijn. Ruis is hierbij alles dat lijkt op het signaal (PNG), maar dat niet is. Een strijd tussen signaal en ruis dus.

De ruis heeft twee oorzaken. Ten eerste is er altijd een theoretische onzekerheid door het feit dat we storingsrekening gebruiken; we weten a priori dat er kleine correcties zijn die we verwaarlozen, maar die kunnen lijken op het PNG signaal. Ten tweede zijn er simpelweg statistische variaties, waar geen enkele theorie vat op heeft, die kunnen lijken op het signaal. Onze analyse vertelt ons, voor een geïdealiseerd scenario waarin we het dichtheidsveld in de ruimte direct kunnen meten (wat niet zo is), hoe goed we deze ruis van het interessante signaal kunnen scheiden.

We vergelijken de prestaties van de EFT of LSS met zijn voorganger: SPT. Tot slot bekijken we hoe relevant de uitbreiding van de theorie zoals beschreven in hoofdstuk 4 is.

Onze resultaten zijn als volgt. In ons geïdealiseerde scenario kunnen we het signaal net wat beter achterhalen dan tot nu toe uit de kosmische achtergrondstraling mogelijk was. Het lijkt echter moeilijk echt interessante theoretische grenzen te kunnen testen. Verder vinden we dat de EFT of LSS

grofweg een factor 3 beter is dan SPT. De uitbreiding naar niet-Gaussische begincondities uit hoofdstuk 4 lijkt voor deze analyse echter irrelevant, vooral omdat de correcties heel klein zijn. Tot slot gaan we dieper in op de rol van de theoretische fout. We beargumenteren dat het cruciaal is deze mee te nemen, suggereren hoe dit te doen en wijzen op de subtiele manier waarop dit een effect kan hebben.

De divergentie van storingsrekening voor de groteschaalstructuur

Het laatste werk van dit proefschrift is in zekere zin geïnspireerd door de theoretische fout hierboven: hoe klein kunnen we die maken?

Om die vraag te beantwoorden bleek het zaak een stap terug te doen naar de essentie van storingsrekening. Zoals de meeste natuurkundige theorieën is de EFT of LSS een storingstheorie: er is een klein getal (zoals bijvoorbeeld de typische grootte van een cluster van sterrenstelsels gedeeld door de veel grotere afstand r waarover we iets willen zeggen) dat het antwoord beïnvloedt. De voorspelling van de waarde van een grootheid wordt nu gegeven als een zogenaamde reeks in dit kleine getal. Laten we het getal x noemen. Dan is een reeks bijvoorbeeld $x + 3x^2 + x^3/3 + \cdots$ enzovoorts. De correcties in deze reeks lijken steeds kleiner te worden om uiteindelijk steeds preciezer het uiteindelijke antwoord te benaderen, maar dit blijkt niet altijd het geval. Zo niet, dan spreken wiskundigen van een asymptotische reeks. De vraag is dus waar sprake van is in het geval van de EFT of LSS.

In dit hoofdstuk beschouwen we het simpelere probleem van structuurformatie in een heelal met maar één ruimtelijke dimensie, zodat bepaalde berekeningen uitgevoerd kunnen worden die anders niet mogelijk zijn. In deze context laten we vervolgens zien dat voor sommige grootheden het antwoord inderdaad asymptotisch van aard is: er is een ondergrens aan hoe precies storingsrekening ooit kan zijn. Voor enkele andere grootheden bewijzen we echter juist dat de reeks mooi convergeert. Dat laatste lijkt in onze ogen niet te betekenen dat dit ook het fysisch juiste antwoord is, maar daarvoor is verder onderzoek nodig.

Voor de asymptotische gevallen geven we bovendien fysische intuïtie voor het gebrek aan convergentie van de reeks, die te maken blijkt te hebben met statisch onwaarschijnlijke, doch onvermijdelijke gebeurtenissen. De statistische aard van onze berekeningen heeft namelijk tot gevolg dat er altijd een kleine kans is dat zelfs op grote schalen iets geks gebeurt. Er kan bijvoor-

beeld ergens op grote schaal een hele grote dichtheid worden aangetroffen, die onder de invloed van zijn zwaartekracht ineenstort om een ingewikkeld cluster te vormen. In dit geval lijkt het gedrag op geen enkele manier op een vloeistof en werkt de EFT of LSS simpelweg niet. Ook de bovengenoemde reeks gaat in dat geval uit van een verkeerde theorie en kan dus nooit dit soort effecten beschrijven. Wat ons echter redt is dat dit op grote schaal maar heel weinig gebeurt en dus vaak te verwaarlozen is. Overigens, het feit dat dit op kleine schaal wel gebeurt is juist geen probleem. De EFT of LSS is namelijk precies ontworpen om onze onwetendheid van de kleine schalen volledig te parametriseren.

Hoe deze logica zich vertaalt naar de 3-dimensionale wereld is nog niet duidelijk, maar het principe lijkt onveranderd. Dit is in zekere zin slecht nieuws voor de zoektocht naar PNG, aangezien er een ondergrens lijkt te zijn aan de precisie van de EFT of LSS en gerelateerde storingsrekeningen.

Acknowledgments

Let me start off by thanking John Stout for warning me about the magnitude of finally sitting and writing one's acknowledgments. It is indeed quite an honor to be able to look back on four years in science and thank everyone who made it such a great experience.

First and foremost I would like to thank my supervisor Enrico. I feel extremely fortunate to have been your student for these four years. As a scientist, you have been an example and an inspiration. I always admired and greatly enjoyed your clarity of thought, both as a speaker and in one-on-one conversations. Your vision and enthusiasm has been a constant source of excitement to me and the rest of the group, and I am sure you will be greatly missed at the ITF. Moreover, I am very grateful for your active effort to help me get the most out of my PhD in every respect, whilst never losing track of the human side of things, and allowing me the freedom to find my own path. I sincerely hope to be able to take along everything I learned in my future endeavors.

Stefan, thank you for guiding me and Enrico through this process as my promotor.

I am also very thankful to Drazen and Tomislav for your guidance during my first steps in cosmology. I will never forget the joy and excitement I felt back then, and I am sure I would have never been able to get to this point without it. Drazen, it has been great to have you around so often to enjoy your friendship and wisdom.

I am very happy to acknowledge my collaborators Daniel, Valentin, Yvette, Giovanni and Sadra for sharing their intelligence and friendliness throughout. Giovanni and Sadra, I truly enjoyed discussing our projects, physics, and life in general all those times. Yvette, what a journey it has been. From talking physics till 10 pm in the Martini Hotel, through understanding the world as an EFT in Sao Paulo, to actually producing some science here in Utrecht, I probably owe most of what I understand about physics to talking with you.

During these four years, I found that somehow I really like cosmologists.

They always seemed both passionate and down to earth - perhaps this is what thinking about the universe does to people. Pavel, Stefano, Bogumila, Leihua, Anja, Zahra, Angelos, Silvije, Sander, Jan, John, Garrett, Luca, Bernardo, Georgos, Cora, Roberto, Gabriele, Alex, Atsuhisa, I hope you all enjoyed being part of this group as much as I did.

My life at the ITF has been extremely pleasant due to the great support of Joost, Wanda and Olga, as well as the vibrant scientific climate that I was grateful to be a part of.

As was hopefully mutual, I always felt a special bond with all PhD's and young postdocs here at the ITF. I hope the welcoming atmosphere you had to offer will stay around here forever. Tara, Jiansen, thank you for being the kindest office mates I can imagine. Bram, thanks for being my go-to-guy when it comes to the final stages of the PhD and life beyond.

I will never forget my experiences in Sao Paulo and Rio in 2015. Without a doubt this was the highlight of my PhD. Meeting many of my fellow DRSTP PhD colleagues for the first time, feeling the excitement of all of us wanting to learn everything about the frontiers of theoretical physics, and having some of the best teachers made the atmosphere irresistible. Nick, Watse, Sabrina, Jacopo, Eva, Francesca, Sam, Laurens, Manus, thank you for also making it the best holiday I ever had.

I would like to thank my yearclub, my friends at the TC and everyone else for shaping my life outside of physics. Guido and Eric, thank you for being the perfect bridge between life in and outside of academia. I couldn't think of a more fitting way to celebrate our friendship than to have you as my paranimfs.

Papa, mama, Imre, Caspar en Iris, bedankt voor jullie liefde en steun. Ik zou niet weten wat ik zonder zou moeten.

Tot slot wil ik dit proefschrift graag opdragen aan mijn beide opa's. Zij hebben mij laten zien dat wetenschap iets is om verliefd op te worden. En dat die liefde altijd gedeeld moet worden.

About the author

I was born on October 13th, 1989 in The Hague, The Netherlands. This is where I completed primary school at "OBS de Houtrust" in 2002, and high school at the "Maerlant Lyceum" in 2008. In 2007 I won the fifth place at the Dutch Physics Olympiad, and accordingly represented The Netherlands at the International Physics Olympiad 2007 in Isfahan, Iran.

I completed the TWIN Bachelor's program in Physics and Mathematics at Utrecht University in 2011, with a thesis on invariant manifolds of normally hyperbolic fixed points of dynamical systems under the supervision of Dr. Heinz Hanssmann. From 2011 to 2014 I studied theoretical physics at the Institute for Theoretical Physics (ITF) in Utrecht, which I completed with a thesis on the "Late-time Quantum Backreaction of a Nonminimally Coupled Massless Scalar on FLRW" under the supervision of Dr. Tomislav Prokopec and under daily supervision of and in collaboration with Dr. Drazen Glavan. In the meantime, I studied abroad at the University of Wisconsin - Madison in the fall semester of 2012, during which I took courses on astronomy and advanced solid state physics.

From September 2014 I worked as a PhD candidate in the group of Dr. Enrico Pajer at the ITF in Utrecht. I mostly studied the evolution of density perturbations in the late universe from the point of view of effective field theory, and investigated if and how this theory can teach us about some interesting aspects of the initial conditions of the universe. The results of this work are summarized in this thesis.

During the last couple of years, I greatly enjoyed spending a big part of my free time on the Olympos tennis courts as a member of the student tennis club TC de Uithof. Finally, I recently developed some interest in writing for a more general public, and I was very proud to win the Robbert Dijkgraaf Essayprijs 2017 with an essay on science in Trumpian times.