

# ESSAY 115: PRIMORDIAL NON-GAUSSIANITY AND LARGE-SCALE STRUCTURE

UNIVERSITY OF CAMBRIDGE

**Supervisors:**

Prof. dr. P. Shellard & Dr. T. Giannantonio

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>From inflation to the CMB</b>	<b>2</b>
2.1	Primordial fluctuations from inflation . . . . .	3
2.1.1	Gaussian fluctuations from slow-roll inflation . . . . .	3
2.1.2	Non-Gaussianities from inflation . . . . .	5
2.2	Anisotropies in the CMB . . . . .	6
<b>3</b>	<b>Non-Gaussianities in the LSS density field</b>	<b>8</b>
3.1	Standard perturbation theory . . . . .	9
3.2	Correlation functions . . . . .	11
3.2.1	Gaussian initial conditions . . . . .	12
3.2.2	Feynman rules . . . . .	13
3.3	The matter power spectrum . . . . .	14
3.4	The matter bispectrum . . . . .	14
3.5	Effective field theory . . . . .	16
<b>4</b>	<b>Probing primordial non-Gaussianities in the LSS</b>	<b>18</b>
4.1	Halo mass function . . . . .	19
4.1.1	Gaussian initial conditions . . . . .	19
4.1.2	Non-Gaussian initial conditions . . . . .	21
4.2	Void abundance . . . . .	22
4.3	Halo biasing . . . . .	22
4.3.1	The local biasing model and the galaxy bispectrum . . . . .	23
4.3.2	Skewness and kurtosis . . . . .	23
4.3.3	Bias scaling . . . . .	24
<b>5</b>	<b>Current constraints and future prospects</b>	<b>25</b>

## 1 Introduction

Our universe contains a vast diversity of intricate structures, ranging from the solar system and the Milky Way to, on a much larger scale, the distribution of galaxies known as the large-scale structure (LSS) or the cosmic web [7, 48, 49]. These structures were formed by gravitational collapse of small density fluctuations in the early universe. Inflationary models<sup>1</sup> predict these primordial fluctuations from fundamental physics [25]. During inflation, small quantum fluctuations in the inflaton field(s) got stretched to superhorizon scales and became classical. After inflation, these fluctuations reentered the horizon as fluctuations in the gravitational potential and density fields. Before recombination the fluctuations oscillated, as matter in the universe was charged and supported by photon pressure. After recombination the universe became neutral, the photon pressure dropped and the fluctuations started to collapse under their own gravity. This furnishes a deep connection between quantum physics in the early universe and the current density

---

<sup>1</sup>Alternative models exist, such as the topological defects models, but a discussion of these is outside the scope of this essay.

field, and the distribution of stars and galaxies in our universe.

A complete understanding of this process would require a theory of both the quantum world and gravitational collapse. At the present time we lack such an understanding. On the one hand, we lack a fundamental theory of inflation and the quantum to classical transition. On the other hand, we do not yet understand the non-linear evolution of the density field and, in particular, the formation of the LSS. However, our current understanding does allow us to make predictions. Given a model of inflation, quantum field theory enables us to predict the statistical nature of the primordial fluctuations in the gravitational potential. Statistical mechanics allows us to evolve these fluctuations till recombination, and perturbation theory and  $N$ -body simulations can be used to model the dynamics after recombination.

Many different inflationary models predict the gravitational potential to be a realization of a nearly scale-invariant Gaussian random field. Over the last years a lot of attention has been put into finding observables which differ between these models. A first observable is a slight tilt in the primordial power spectrum, towards the red or the blue. However, most models predict the tilt to be very small. A second observable are  $B$ -mode polarizations in the Cosmic Microwave Background radiation field (CMB), interpreted as primordial gravitational waves, which would support the inflationary paradigm and determine the energy scale of inflation. A third observable are isocurvature perturbations in the density fields, which can occur in multi-field inflation and would reject simple single-field inflationary models. A fourth observable are deviations from the Gaussian statistic of the primordial fluctuations, known as primordial non-Gaussianities. In this essay we will only consider the fourth observable.

For a long time it was thought that detailed observations of the CMB were the best probe to determine the nature of primordial fluctuations and inflation, as the CMB reflects the density field at recombination. Observations by the WMAP and Planck consortium show that the CMB closely resembles a nearly scale-invariant Gaussian random field [26, 37]. The Planck consortium has measured a slight tilt to the red in the power spectrum, and has only set bounds on the  $B$ -mode polarizations, isocurvature perturbations and amplitude of non-Gaussianities, i.e.,  $f_{NL} = 0.8 \pm 5.0$  (of local type). These constraints still allow for many different inflationary models. New CMB missions are being proposed, but due to cosmic variance, the next generation of CMB observations are not expected to significantly improve these bounds.

A natural next step would be to use LSS survey data to improve our understanding of early universe physics. Non-Gaussianities in the density field at recombination are namely expected to influence the statistics of the LSS. This is a complicated observable, since the density field evolves non-linearly after recombination and erases primordial statistics. We furthermore only observe the density fluctuations via galaxies, which only form in high density regions. However, over the last few years our understanding of both the evolution of the density field and the formation of galaxies has improved, and the next generation of LSS missions will detect tens of thousands of galaxy clusters. This combination of theoretical and observational advancements might lead to tighter constraints on primordial non-Gaussianities.

In this essay we review ways to trace early universe physics with LSS surveys. We in particular concentrate on the theoretical side of the statistics of the large-scale structure. In section 2 we briefly discuss the generation of primordial fluctuations by inflation and the evolution of these fluctuations till recombination. In section 3 we review current developments in standard and effective perturbation theory of large-scale structure formation. We formalize and extend the formalism of Carroll *et al.* [9] to include non-Gaussian statistics. In section 4 we discuss imprints of primordial non-Gaussianities on the galaxy distribution of the LSS, and different methods to detect them. Finally, in section 5 we review current constraints and prospects.

## 2 From inflation to the CMB

Inflation theory predicts that the classical fluctuations in the gravitational potential were formed from quantum fluctuations in the inflaton field(s) during inflation and reentered the horizon after inflation. After reentry, the fluctuations evolved in a photon baryon plasma till recombination. The state of the fluctuations at this moment can be observed in the CMB. In this section we briefly review the predictions of primordial fluctuations by inflation theory and the evolution of these fluctuations in the photon baryon plasma till recombination.

## 2.1 Primordial fluctuations from inflation

The simplest model of inflation is single field slow-roll inflation governed by the action

$$S = \int d^4x \sqrt{-g} \left[ -\frac{1}{2} \partial_\mu \phi \partial^\mu \phi - V(\phi) \right], \quad (1)$$

with the inflaton potential  $V$  and field  $\phi$  minimally coupled to a homogeneous Friedmann-Robertson-Walker (FRW) metric  $g_{\mu\nu}$ , defined by the equation

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = -dt^2 + a^2(t) \left[ \frac{dr^2}{1 - kr^2} + r^2 [\sin^2(\theta) d\phi^2 + d\theta^2] \right], \quad (2)$$

with scale factor  $a$ , curvature  $k$  and  $g$  the determinant of  $g_{\mu\nu}$  [3, 43]. For convenience, we have set Planck's constant and the speed of light in vacuum to unity, i.e.,  $\hbar = c = 1$ . For single field inflation, the theory is completely specified by the potential  $V$ . The inflaton field  $\phi$  corresponds to a particle present in the early universe which led to the epoch of inflation. The fundamental theory describing the field  $\phi$  induces the potential  $V$ . So far, we do not know whether inflation is governed by only one field and if so, which particle corresponds to  $\phi$ . We furthermore lack a fundamental theory describing the inflaton field<sup>2</sup>. For this reason we consider a generic potential leading to sufficient inflation to explain the flatness and horizon problem<sup>3</sup>.

An inflaton potential  $V$  generally has a plateau region in which inflation occurs (see figure 1). The inflaton field slowly rolls down the plateau towards the minimum, where it rapidly oscillates. This is known as slow-roll inflation. Formally a potential  $V$  leads to slow-roll inflation if there exists a region in which the slow-roll parameters  $\epsilon$  and  $\eta$  defined as

$$\epsilon = -\frac{\dot{H}}{H^2}, \quad \eta = \frac{\dot{\epsilon}}{H\epsilon}, \quad (3)$$

satisfy the slow-roll conditions  $\epsilon < 1, |\eta| < 1$ , with Hubble parameter  $H = \frac{\dot{a}}{a}$  and the dot denoting a derivative with respect to time. Using the Friedmann equation, which describes the evolution of a FRW metric subject to Einstein's theory of general relativity, we can approximate the slow-roll parameters in terms of the potential and its derivatives

$$\epsilon_V = \frac{M_{pl}^2}{2} \left( \frac{V'}{V} \right)^2, \quad \eta_V = M_{pl}^2 \frac{|V''|}{V}, \quad (4)$$

with  $M_{pl} = \sqrt{\frac{\hbar c}{8\pi G}} = 2.4 \times 10^{18}$  GeV the Planck mass and the primes denoting derivatives of the potential  $V$  with respect to the field  $\phi$ . In equation (4), we observe that the slow-roll conditions indeed require a plateau region in the potential.

A classical treatment of equation (1) with a potential satisfying the slow-roll conditions leads to an epoch of inflation. Quantum fluctuations around the classical solution lead to fluctuations in the gravitational potential after inflation. As the slow-roll action is non-linear it is usually analysed perturbatively. To leading order in slow-roll parameters  $\epsilon$  and  $\eta$ , the action reduces to a harmonic oscillator leading to Gaussian fluctuations after inflation. These Gaussian fluctuations can be formally represented by a Gaussian random field as discussed in section 2.1.1. Higher order corrections in the action lead to deviations from the Gaussian statistics considered in section 2.1.2. For a derivation of these results see [31] and [11].

### 2.1.1 Gaussian fluctuations from slow-roll inflation

A Gaussian random field is a statistical object of which the realizations are functions having a Gaussian PDF at every point and are completely described by the two-point correlation function [1]. Examples of Gaussian fields can be observed in noise in telephone cables [39], shallow ocean waves and mountains (see figure 2) [30]. From inflation theory it follows that also the primordial fluctuations in the gravitational

<sup>2</sup>There exist several proposals, such as inflation governed by the Higgs field or particles predicted by string theory.

<sup>3</sup>During inflation the scale factor  $a$  must at least grow by a factor  $e^{64}$  in order to explain the observed flatness of our universe and extreme isotropy in the CMB [36].

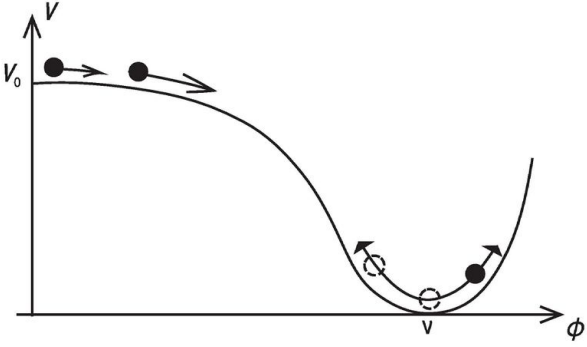


Figure 1: The general form of a slow-roll inflation potential.

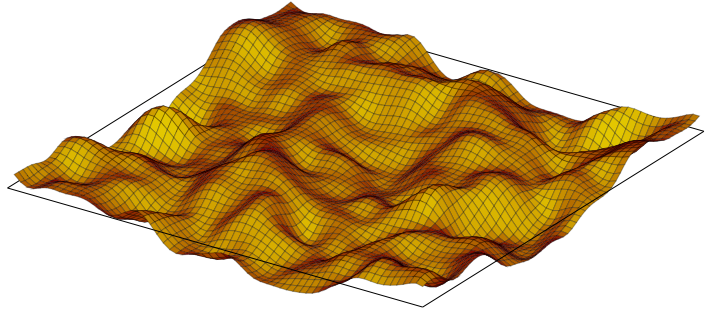


Figure 2: A realization of a two-dimensional Gaussian random field, by Hidding [20].

field are Gaussian distributed [15, 23]. Formally a realization of a Gaussian random field on the volume  $\mathbb{R}^3$  is a function  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  such that the probability density for  $f$  to assume the values  $f_1, \dots, f_n$  in the points  $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^3$  is

$$P[f_1, \dots, f_n] = \frac{e^{-\frac{1}{2} \Delta f_i (M^{-1})^{ij} \Delta f_j}}{[(2\pi)^n \det M]^{-1/2}}, \quad (5)$$

with  $\Delta f_i = f_i - \langle f(\mathbf{x}_i) \rangle$  and  $M = (M_{ij})$  the auto-correlation matrix where  $M_{ij} = \xi(\mathbf{x}_i, \mathbf{x}_j)$ . Here  $\xi$  is the two-point correlation function

$$\xi(\mathbf{x}_i, \mathbf{x}_j) = \langle f(\mathbf{x}_i) f(\mathbf{x}_j) \rangle. \quad (6)$$

Note that the Gaussian random field is indeed fully characterized by  $\xi$ ; higher order correlation functions can be expressed in terms of  $\xi$  by Wick's theorem [35]. Colloquially odd-point correlation functions vanish and even-point correlation functions are equal to the sum of the product of all pairings of the points, i.e.,

$$\langle f(\mathbf{x}_1) f(\mathbf{x}_2) f(\mathbf{x}_3) \rangle = 0, \quad (7)$$

$$\langle f(\mathbf{x}_1) f(\mathbf{x}_2) f(\mathbf{x}_3) f(\mathbf{x}_4) \rangle = \xi(\mathbf{x}_1, \mathbf{x}_2) \xi(\mathbf{x}_3, \mathbf{x}_4) + \xi(\mathbf{x}_1, \mathbf{x}_3) \xi(\mathbf{x}_2, \mathbf{x}_4) + \xi(\mathbf{x}_1, \mathbf{x}_4) \xi(\mathbf{x}_2, \mathbf{x}_3), \quad (8)$$

$$\langle f(\mathbf{x}_1) f(\mathbf{x}_2) f(\mathbf{x}_3) f(\mathbf{x}_4) f(\mathbf{x}_5) \rangle = 0. \quad (9)$$

For fields with statistical homogeneity and isotropy, the two-point correlation function only depends on the distance between points, i.e.,  $\xi(r) = \xi(\mathbf{x}, \mathbf{x} + \mathbf{r})$  with  $r = \|\mathbf{r}\|$ .

Define the Fourier transform  $\hat{f}$  of the function  $f$  as

$$f(\mathbf{x}, t) = \int \frac{d^3k}{(2\pi)^3} e^{-i\mathbf{k}\cdot\mathbf{x}} \hat{f}(\mathbf{k}, t). \quad (10)$$

The statistics of the Fourier modes of a Gaussian random field take the particularly simple form

$$\langle \hat{f}(\mathbf{k}_1) \hat{f}(\mathbf{k}_2) \rangle = (2\pi)^3 \delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2) P_f(k_1) \quad (11)$$

with  $k_1 = \|\mathbf{k}_1\|$  and  $P_f$  the Fourier transform of the two-point correlation function  $\xi$ , known as the power spectrum.

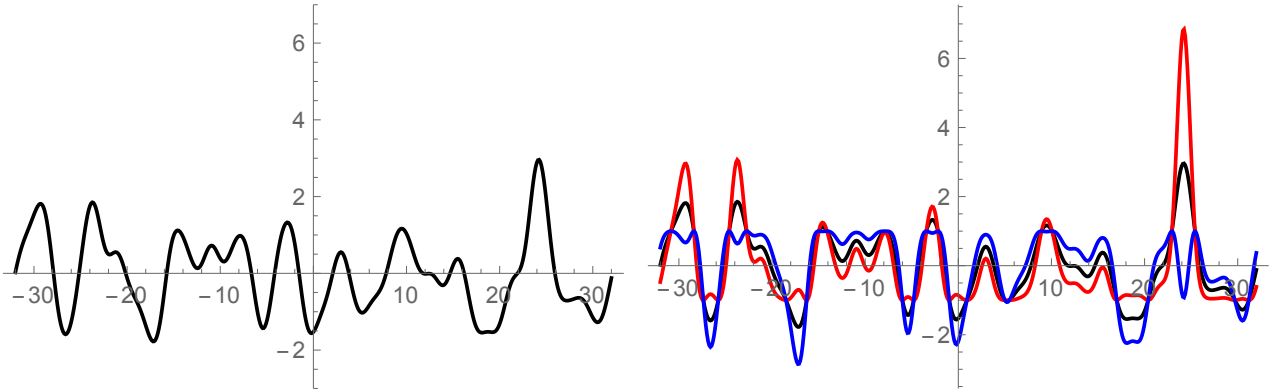
To leading order, single field slow-roll inflation predicts the gravitational potential  $\Phi$  to fluctuate as a Gaussian random field with power spectrum

$$P_\Phi(k) = \frac{1}{8\pi^2 \epsilon} \frac{H^2}{M_{pl}^2} \Big|_{k=aH}, \quad (12)$$

with  $M_{pl}$  the Planck mass. The expression is evaluated at horizon exit, i.e.,  $k = aH$ . Since both the Hubble parameter  $H$  and slow-roll parameter  $\epsilon$  only evolve slowly during slow-roll inflation, the fluctuations are nearly scale-invariant and can be expressed as a power-law, i.e.,

$$P_\Phi(k) = A_s \left( \frac{k}{k_*} \right)^{n_s - 1}. \quad (13)$$

See figure 3a for a realization of a one-dimensional Gaussian random field with a power-law power spectrum. According to the most recent Planck observations, the amplitude  $A_s = (2.196 \pm 0.060) \times 10^{-9}$  at scale  $k_* = 0.05 \text{ Mpc}^{-1}$  and the spectral index  $n_s = 0.9677 \pm 0.0060$  [36].



(a) A realization of a scale-invariant Gaussian random field with spectral index  $n_s = 0.9$  and a Gaussian smoothing function with  $\sigma = 1$  pixels.

(b) Local non-Gaussian fields corresponding to the realization of figure 3a. The red line correspond to  $f_{NL} = 0.5$  and the blue line to  $f_{NL} = -0.5$ .

Figure 3: One-dimensional Gaussian and non-Gaussian field.

### 2.1.2 Non-Gaussianities from inflation

The simplest inflationary models lead to a highly Gaussian primordial gravitational potential. To be precise, Maldacena [31] proved that inflationary models governed by (i) a single scalar field, (ii) with canonical kinetic terms, (iii) which always satisfies slow-roll conditions, (iv) is initially in the Bunch-Davies vacuum, and (v) is in accordance with Einstein's theory of gravity, lead to negligible primordial non-Gaussianities<sup>4</sup>. For a more extensive discussion see [11] and [28].

The violation of any of these conditions can lead to the occurrence of measurable non-Gaussianities. The simplest type of non-Gaussianities is the local type, first described by Komatsu and Spergel [27], which is defined as

$$\Phi(\mathbf{x}) = \Phi_G(\mathbf{x}) + f_{NL} [\Phi_G^2(\mathbf{x}) - \langle \Phi_G^2(\mathbf{x}) \rangle] \quad (15)$$

with  $\Phi_G$  a Gaussian random field and  $f_{NL}$  the non-linear parameter measuring the amplitude of non-Gaussianity (see figure 3b). The effect of the non-linear term is most easily characterized by higher-order correlation functions. We here only consider the three-point function, with the corresponding bispectrum  $B_\Phi$ ,

$$\langle \Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\Phi(\mathbf{k}_3) \rangle = (2\pi)^3 B_\Phi(k_1, k_2, k_3) \delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3). \quad (16)$$

Statistical homogeneity and isotropy imply that the bispectrum only depends on the norm of the difference of the momentum vectors:  $\|\mathbf{k}_1 - \mathbf{k}_2\|$ ,  $\|\mathbf{k}_1 - \mathbf{k}_3\|$ ,  $\|\mathbf{k}_2 - \mathbf{k}_3\|$ . The triangle condition  $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0$  further restricts the interesting configurations of momentum vectors to those forming a closed triangle. Hence the bispectrum is completely determined by its value on a tetrahedron (see figure 4). It is common to write the bispectrum in terms of the shape function  $S_\Phi$  defined as

$$S_\Phi(k_1, k_2, k_3) = \frac{(k_1 k_2 k_3)^2}{N_\Phi} B_\Phi(k_1, k_2, k_3), \quad (17)$$

with  $N_\Phi$  a normalization constant such that  $S_\Phi(k, k, k) = 1$ . For local type non-Gaussianities the shape function is

$$S^{loc}(k_1, k_2, k_3) = \frac{1}{3} \left( \frac{k_1}{k_2 k_3} + \frac{k_2}{k_1 k_3} + \frac{k_3}{k_1 k_2} \right),$$

<sup>4</sup> Slow-roll single field inflation generates a mix of local and equilateral non-Gaussianities with shape function

$$S(k_1, k_2, k_3) \approx (6\epsilon - 2\eta) S^{local}(k_1, k_2, k_3) + \frac{5}{3} \epsilon S^{equil}(k_1, k_2, k_3). \quad (14)$$

For small slow-roll parameters this tends to the local type non-Gaussianities with non-linear parameter  $f_{NL}^{loc} = \frac{5}{12}(1 - n_s) \approx 0.017$ , which is beyond the scope of current measurements. Shape functions characterize non-Gaussianities and are defined later in this section.

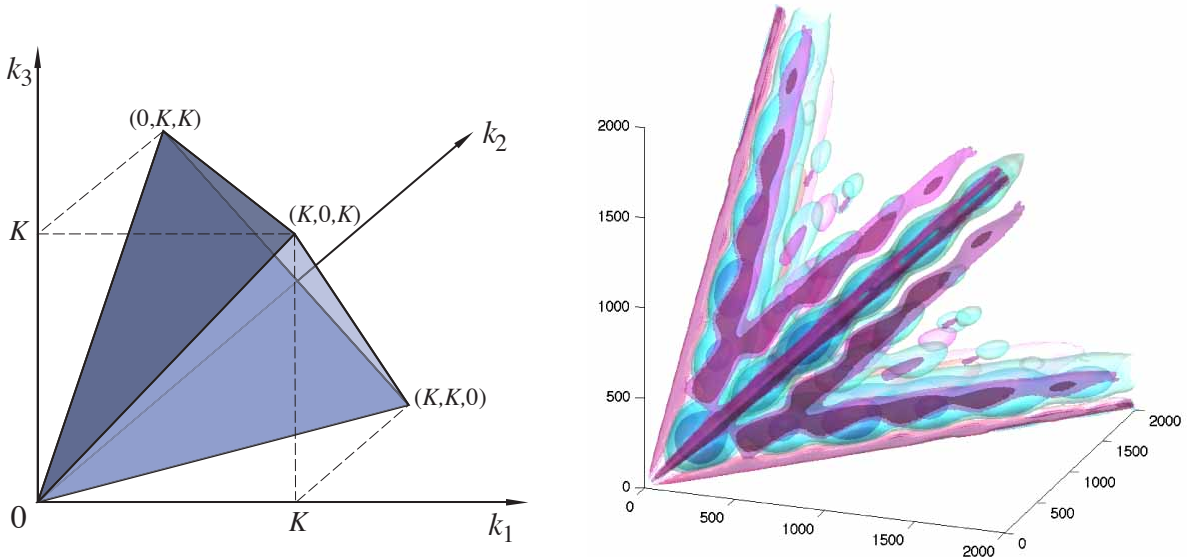


Figure 4: The left plot shows the tetrahedron domain of momentum configurations contributing to the bispectrum. The right plot shows the bispectrum in the CMB corresponding to local primordial local non-Gaussianities normalized with respect to a constant shape function [28].

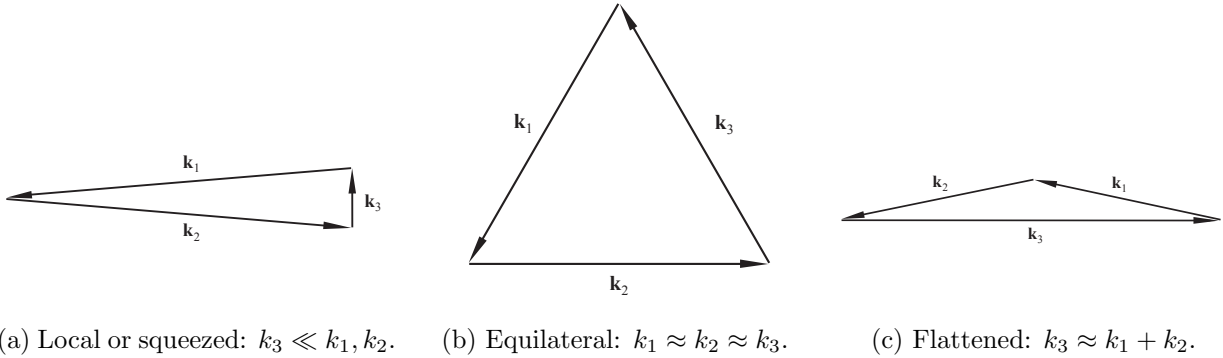


Figure 5: Three triangle configurations [28].

peaking in the squeezed limit configuration  $k_3 \ll k_1 \approx k_2$  (see figure 5). This generally corresponds to non-Gaussianity being generated outside the horizon by potentially multiple inflaton fields.

Theories with higher-order derivative operators, such as DBI inflation [2, 24], lead to the shape function of the equilateral type

$$S^{equil}(k_1, k_2, k_3) = \frac{(k_1 + k_2 - k_3)(k_2 + k_3 - k_1)(k_3 + k_1 - k_2)}{k_1 k_2 k_3}, \quad (18)$$

peaking in the equilateral triangle configuration  $k_1 \approx k_2 \approx k_3$  (see figure 5). When the initial vacuum is different from the Bunch-Davies vacuum, see for example [21], the shape function tends to the flattened type

$$S^{flat}(k_1, k_2, k_3) \propto 6 \left[ \frac{k_1^2 + k_2^2 - k_3^2}{k_2 k_3} + 2 \text{ perm} \right] + \frac{2(k_1^2 + k_2^2 + k_3^2)}{(k_1 + k_2 - k_3)^2 (k_2 + k_3 - k_1)^2 (k_3 + k_1 - k_2)^2}, \quad (19)$$

peaking in the flattened configuration  $k_2 \approx k_3 \approx k_1/2$  (see figure 5). For a more extensive overview of the different kinds of violations and resulting non-Gaussianities, see the reviews by Liguori *et al.* [28] and Chen [11].

## 2.2 Anisotropies in the CMB

The fluctuations in the gravitational potential  $\Phi$  led to density fluctuations  $\delta_0$  which evolved in the photon baryon plasma. To linear order, which is a very good approximation, the density at redshift  $z$  can be

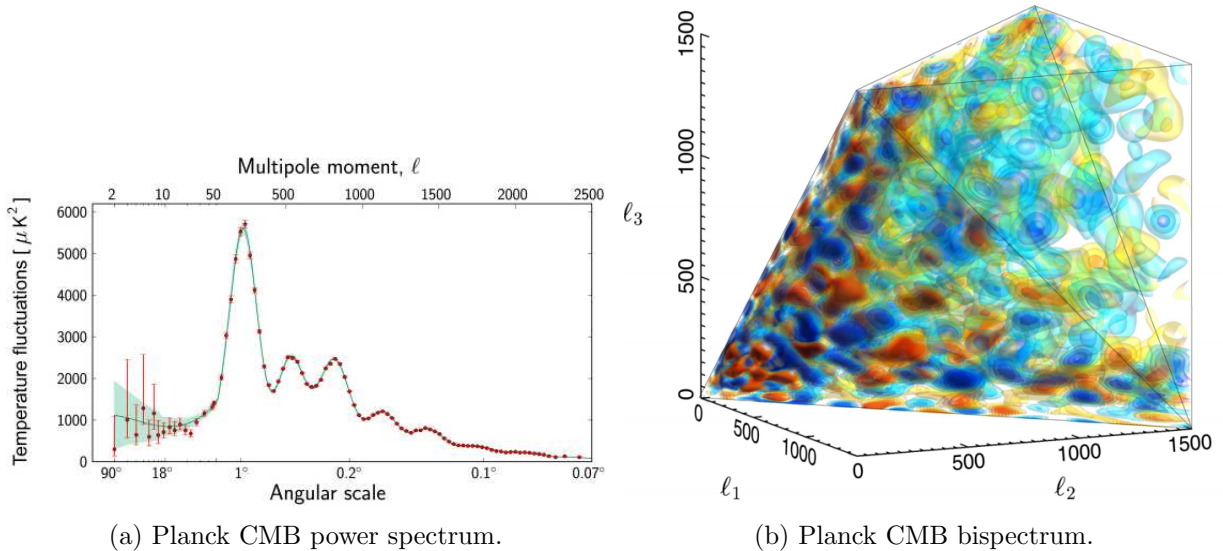


Figure 6: Most recent CMB power spectrum and bispectrum results from the Planck consortium [36].

expressed in terms of the primordial gravitational perturbation as

$$\delta_0(\mathbf{k}, z) = \mathcal{M}(k, z)\Phi(\mathbf{k}) \quad (20)$$

with

$$\mathcal{M}(k, z) = \frac{2}{3} \frac{k^2 T(k) D(z)}{\Omega_m H_0^2}, \quad (21)$$

where  $k^2$  reflects the Poisson equation,  $\Omega_m$  the current matter density,  $D(z)$  the linear growth rate and  $T(k)$  the matter transfer function. The transfer function contains effects like acoustic oscillations leading to Doppler peaks, and silk damping. The function can be obtained by numerical evaluation of the Boltzman equation.

To linear order the density fluctuations at the epoch of last scattering inherit the statistics of the gravitational perturbation, i.e.,

$$\langle \delta_0(\mathbf{k}_1, z) \dots \delta_0(\mathbf{k}_n, z) \rangle = \left[ \prod_{i=1}^n \mathcal{M}(k_i, z) \right] \langle \Phi(\mathbf{k}_1) \dots \Phi(\mathbf{k}_n) \rangle. \quad (22)$$

We conclude that to linear order, the anisotropies of the CMB directly reflect the density perturbation at recombination. By measuring the CMB, we can infer the primordial fluctuations generated during inflation and possibly reheating. In particular, the Planck consortium has found a nearly scale-invariant power spectrum with acoustic peaks (see figure 6a) and a bispectrum (see figure 6b) constraining primordial non-Gaussianities as  $f_{NL}^{local} = 0.8 \pm 5.0$ ,  $f_{NL}^{equil} = -4 \pm 43$  and  $f_{NL}^{orth} = -26 \pm 21$  for local, equilateral and orthogonal type respectively at 68% confidence levels [37].

However, the CMB can only be measured on a two-dimensional sphere surrounding us. This sets a natural limit on the amount of data we can collect to test our theories. The uncertainty of measurables on large scales that cannot be reduced indefinitely by collecting more CMB data is known as cosmic variance. The Planck consortium has nearly reached this limit, while they only have been able to set bounds on the existence of primordial non-Gaussianities. The next generation of CMB surveys are not expected to significantly improve these bounds. The next leap forward might well be the use of LSS data in the quest for non-Gaussianities; LSS surveys provide an immense amount of three dimensional data and might enable us to improve existing bounds.

### 3 Non-Gaussianities in the LSS density field

At recombination, the photons and baryons decouple. Photons stream freely through the universe, while matter fluctuations start to collapse under their own gravity. At this point the density perturbation

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \langle \rho(\mathbf{x}) \rangle}{\langle \rho(\mathbf{x}) \rangle} \quad (23)$$

closely resembles a Gaussian random field. The current large-scale structure (LSS), however, is a highly non-Gaussian field with nontrivial higher order statistics. The non-Gaussianities in this field originate from three different sources [5, 28]:

- *Gravitational instability*: Gravitational collapse is a non-linear process leading to clusters, filaments, voids. Note from equation (23) that  $\delta \rightarrow -1$  in voids while  $\delta \rightarrow \infty$  in clusters. Hence gravitational instability leads to an asymmetry in the PDF of  $\delta$  which implies non-Gaussian statistics.
- *Primordial non-Gaussianities*: Alongside the non-Gaussianities generated after recombination, it is natural to assume that primordial non-Gaussianities in the density field at recombination will influence the level of non-Gaussianity in the current large-scale structure.
- *Halo biasing*: Current large-scale structure surveys observe galaxies, which are used to trace the underlying density field. However, galaxies only form in dense regions and different populations of galaxies are known to trace different features of the density field. Non-linear halo biasing can lead to non-Gaussianities. There exist attempts to measure the density field directly using gravitational lensing. We will here, however, concentrate on surveys aimed at galaxies.

If we are to measure primordial non-Gaussianities from the large-scale structure we need an accurate understanding of these three effects. During the last few years the scientific community has developed more sophisticated perturbation schemes and studied non-Gaussianities in  $N$ -body simulations to improve our understanding of gravitational instability theory and the effect of primordial non-Gaussianities on the LSS density field. Moreover, Dalal *et al.* [13] showed that halo biasing is intrinsically influenced by primordial non-Gaussianities, i.e., the presence of primordial non-Gaussianities influences the formation of galaxies which leads to a significant change in the biasing relation. In this section we review current developments in our understanding of the LSS density field. We discuss standard perturbation theory and effective field theory and compare the theories with perturbative quantum field theory. We follow the discussion by Carroll *et al.* [9] and extend this discussion by deriving Feynman rules, relating the discussion to solutions in an Einstein-de Sitter universe, and by including non-Gaussian initial conditions. We subsequently show how gravitational instability theory and primordial non-Gaussianities lead to non-Gaussian statistics of the matter density. In section 4 we focus on halo biasing and methods to measure primordial non-Gaussianities in LSS surveys.

Large-scale structure formation is a process governed by gravity. While general relativistic approaches to LSS formation exist and our discussion can be extended to include relativistic corrections, we will for simplicity model LSS formation by a single Newtonian pressureless fluid in an expanding FRW background universe. We will neglect radiation and consider baryonic and dark matter as a single fluid. This is a valid approximation on subhorizon scales after the epoch of recombination. The evolution of such a fluid is governed by the continuity equation, Euler equation and Poisson equation, respectively

$$0 = \partial_\tau \delta + \partial_i((1 + \delta)v^i), \quad (24)$$

$$0 = \partial_\tau v^i + \mathcal{H}v^i + \partial_i \psi + v^j \partial_j v^i, \quad (25)$$

$$\nabla^2 \psi = \frac{3}{2} \Omega \mathcal{H}^2 \delta, \quad (26)$$

with conformal time  $\tau$  defined by  $dt = a^{-1}d\tau$ , conformal Hubble parameter  $\mathcal{H} = \frac{a'}{a}$  with the prime denoting a derivative with respect to conformal time, gravitational perturbation  $\psi$ , and density perturbation  $\delta$ . This is a set of coupled non-linear partial differential equations. The density fluctuations induce a gravitational potential via the Poisson equation. The potential drives the flow of matter via the Euler equation,



constraint by the continuity equation.

In order to understand the consequences of these LSS equations, we approximate solutions with perturbation theory. We make the simplifying assumption that the velocity field is a gradient field<sup>5</sup> and work with the divergence  $\theta = \partial_i v^i$ . In Fourier space, the LSS equations can be expressed as

$$\begin{aligned} 0 &= \partial_\tau \delta(\tau, \mathbf{k}) + \theta(\tau, \mathbf{k}) + \int \frac{d^3 q}{(2\pi)^3} \frac{\mathbf{k} \cdot \mathbf{q}}{q^2} \delta(\tau, \mathbf{k} - \mathbf{q}) \theta(\tau, \mathbf{q}), \\ 0 &= \partial_\tau \theta(\tau, \mathbf{k}) + \mathcal{H} \theta(\tau, \mathbf{k}) + \frac{3}{2} \Omega \mathcal{H}^2 \delta(\tau, \mathbf{k}) + \int \frac{d^3 q}{(2\pi)^3} \frac{k^2 \mathbf{q} \cdot (\mathbf{k} - \mathbf{q})}{2q^2(\mathbf{k} - \mathbf{q})^2} \theta(\tau, \mathbf{k} - \mathbf{q}) \theta(\tau, \mathbf{q}), \end{aligned} \quad (27)$$

where we suppress the hat notation for Fourier space variables.

For clarity and in order to make contact with perturbative quantum field theory we bundle the density perturbation and velocity divergence in a vector field

$$\phi(\tau, \mathbf{k}) = \begin{pmatrix} \delta(\tau, \mathbf{k}) \\ \theta(\tau, \mathbf{k}) \end{pmatrix}, \quad (28)$$

and write the system of equations as

$$0 = \mathcal{D}_j^i \phi^j(\tau, \mathbf{k}) - \frac{1}{2} \iint \frac{d^3 q_1 d^3 q_2}{(2\pi)^6} M_{jk}^i(\mathbf{k}, \mathbf{q}_1, \mathbf{q}_2) \phi^j(\tau, \mathbf{q}_1) \phi^k(\tau, \mathbf{q}_2), \quad (29)$$

with  $\mathcal{D}$  containing the first order conformal time derivatives and linear terms and  $M_{jk}^i$  the quadratic interactions. To define  $M$  uniquely, we take  $M$  to be symmetric in its lower indices. For our set of differential equations,  $\mathcal{D}$  and  $M$  are given by

$$\mathcal{D}_j^i = \begin{pmatrix} \partial_\tau & 1 \\ \frac{3}{2} \Omega \mathcal{H}^2 & \partial_\tau + \mathcal{H} \end{pmatrix}, \quad (30)$$

$$M_{ij}^\delta(\mathbf{k}, \mathbf{q}_1, \mathbf{q}_2) = \begin{pmatrix} 0 & -\frac{\mathbf{k} \cdot \mathbf{q}_2}{q_2^2} \\ -\frac{\mathbf{k} \cdot \mathbf{q}_1}{q_1^2} & 0 \end{pmatrix} (2\pi)^3 \delta^{(3)}(\mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2), \quad (31)$$

$$M_{ij}^\theta(\mathbf{k}, \mathbf{q}_1, \mathbf{q}_2) = \begin{pmatrix} 0 & 0 \\ 0 & -\frac{k^2(\mathbf{q}_1 \cdot \mathbf{q}_2)}{q_1^2 q_2^2} \end{pmatrix} (2\pi)^3 \delta^{(3)}(\mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2). \quad (32)$$

Momentum conservation is enforced by the Dirac delta function in  $M$ . The factor  $(2\pi)^3$  comes from the Fourier convention used in this essay. The perturbation theory described below is most transparent if we follow Carroll *et al.* [9] by suppressing the integrals over  $\mathbf{q}_1$  and  $\mathbf{q}_2$ , and time and momentum dependence,

$$0 = \mathcal{D}_j^i \phi^j - \frac{1}{2} M_{jk}^i \phi^j \phi^k. \quad (33)$$

Interactions beyond the quadratic term can be included by adding higher order interactions

$$-\frac{1}{3!} N \phi^3 - \frac{1}{4!} O \phi^4 - \dots \quad (34)$$

However, for our discussion the quadratic interaction suffices.

### 3.1 Standard perturbation theory

Traditionally, the LSS equations (33) are solved with standard perturbation theory. We approximate  $\phi$  by an expansion in terms of a perturbation parameter  $\epsilon$ , i.e.,

$$\phi_{SPT} = \epsilon \phi_{(1)}^i + \epsilon^2 \phi_{(2)}^i + \epsilon^3 \phi_{(3)}^i + \dots \quad (35)$$

---

<sup>5</sup>On large scales vorticity is suppressed by the expansion of the universe.

This is an approximation as some solutions to (33) cannot be expressed as an expansion in  $\epsilon$ . Substituting  $\phi_{SPT}$  in the LSS equations and selecting terms proportional to powers in  $\epsilon$  gives a set of differential equations

$$\mathcal{O}(\epsilon) : \mathcal{D}_j^i \phi_{(1)}^j = 0, \quad (36)$$

$$\mathcal{O}(\epsilon^2) : \mathcal{D}_j^i \phi_{(2)}^j = \frac{1}{2} M_{jk}^i \phi_{(1)}^j \phi_{(1)}^k, \quad (37)$$

$$\mathcal{O}(\epsilon^3) : \mathcal{D}_j^i \phi_{(3)}^j = \frac{1}{2} M_{jk}^i \phi_{(1)}^j \phi_{(2)}^k + \frac{1}{2} M_{jk}^i \phi_{(2)}^j \phi_{(1)}^k, \quad (38)$$

$$\mathcal{O}(\dots) : \dots \quad (39)$$

where we suppress integrals over momentum space corresponding to the interaction  $M$ . Note that the differential equation for  $\phi_{(i)}$  only depends on  $\phi_{(j)}$  with  $j < i$ . We can iteratively solve the differential equations with a retarded Green function  $G$ , satisfying

$$\mathcal{D}_j^i G_k^j(\tau, \tau_{in}) = \delta_k^i \delta(\tau - \tau_{in}), \quad (40)$$

with  $\tau_{in}$  the initial time<sup>6</sup>. For an Einstein-de Sitter universe, spatially flat and matter-dominated with  $\Omega = 1, a \propto \tau^2$  and  $\mathcal{H} = \frac{2}{\tau}$ , the retarded Green function is

$$G(\tau_1, \tau_2) = \begin{pmatrix} \frac{3\tau_1^5 + 2\tau_2^5}{5\tau_1^3 \tau_2^2} & \frac{-\tau_1^5 + \tau_2^5}{5\tau_1^3 \tau_2} \\ -\frac{6(\tau_1^5 - \tau_2^5)}{5\tau_1^4 \tau_2^2} & \frac{2\tau_1^5 + 3\tau_2^5}{5\tau_1^4 \tau_2} \end{pmatrix} \Theta(\tau_1 - \tau_2) \quad (41)$$

with  $\Theta$  the Heaviside step function. The leading order term  $\phi_{(1)}$  is solved by

$$\phi_{(1)}^i(\tau, \mathbf{k}) = G_j^i(\tau, \tau_{in}) \phi_{in}^j(\mathbf{k}), \quad (42)$$

with  $\phi_{in}$  the initial condition at  $\tau = \tau_{in}$ . The initial time can be any time at which the initial conditions are satisfied. In LSS formation it is common to take  $\tau_{in}$  to be the time of recombination. The higher order terms  $\phi_{(2)}$ , and  $\phi_{(3)}$  can be solved using the solution for  $\phi_{(1)}$ :

$$\phi_{(2)}^i = \frac{1}{2} \int_{\tau_{in}}^{\tau} d\tau' G_j^i(\tau, \tau') M_{kl}^j(\tau') \phi_{(1)}^k(\tau') \phi_{(1)}^l(\tau'), \quad (43)$$

$$\phi_{(3)}^i = \int_{\tau_{in}}^{\tau} d\tau' G_j^i(\tau, \tau') M_{kl}^j(\tau') \phi_{(1)}^k(\tau') \phi_{(2)}^l(\tau'). \quad (44)$$

The solution for  $\phi_{(1)}$ ,  $\phi_{(2)}$ , and  $\phi_{(3)}$  can be expressed in terms of Feynman diagrams (see figure 7). In these Feynman diagram, the horizontal and vertical direction respectively represent space and time. The lower and upper horizontal lines represent the initial and final field. The vertical line represents the Green function or propagator in quantum field theory language. The quadratic coupling of the Fourier modes  $M_{kl}^j$  is represented by a vertex. The integral over time  $\tau'$  reflects the fact that the interaction occurred between time  $\tau_{in}$  to  $\tau$ . Couplings beyond the quadratic coupling  $M$  will as in perturbative quantum field theory lead to vertices with more legs. Note that  $n$ -point interactions in the equation of motion will always lead to a vertex with  $n$  legs pointing down and one leg pointing up. This prevents us from drawing loop diagrams, reflecting the fact that our discussion is fully determined by specifying  $\phi_{in}$ .

Using the Feynman diagrams we can easily write any term  $\phi_{(i)}$  in the perturbation expansion of  $\phi$ , by drawing all diagrams with one leg connecting to the upper horizontal line and  $i$  legs connecting to the lower horizontal line. The vertices always have one upper and two lower legs. The symmetry factors work in exact analogy with perturbative quantum field theory. This is equivalent to deriving the differential equation by extending the set of equations (39) and solving the equation by means of the retarded Greens function.

In the large-scale structure literature, standard perturbation theory in an Einstein-de Sitter universe is often written in a slightly different formulation. In an Einstein-de Sitter universe, we can use  $\Omega = 1, a(\tau) \propto$

<sup>6</sup>Formally we are free to choose any Green function. We select the retarded Green function to enforce causality on our solutions.

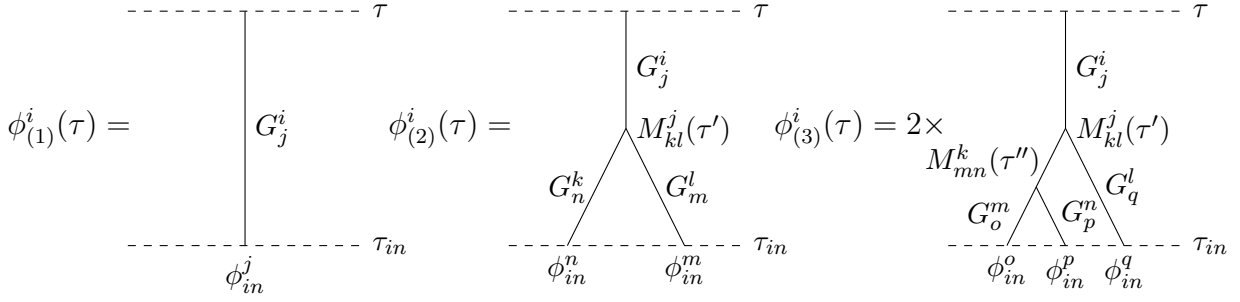


Figure 7: Standard perturbation theory in terms of Feynman diagrams, with all indices and conformal time dependences explicitly expressed. The multiplicative factor 2 in expression of  $\phi_{(3)}$  is a symmetry factor, reflecting different ways of writing the Feynman diagram as in perturbative quantum field theory.

$\tau^2$ ,  $\mathcal{H} = \frac{2}{\tau}$  and factor out a factor  $\mathcal{H}$  from the velocity field to bring equation (27) in a homogeneous form with respect to conformal time  $\tau$ . We for this reason expand  $\delta$  and  $\theta$  in a factorized form

$$\delta(\tau, \mathbf{k}) = \sum_{i=1}^{\infty} \delta_{(i)}(\tau, \mathbf{k}) = \sum_{i=1}^{\infty} a^i(\tau) \tilde{\delta}_{(i)}(\mathbf{k}) \quad (45)$$

$$\theta(\tau, \mathbf{k}) = \sum_{i=1}^{\infty} \theta_{(i)}(\tau, \mathbf{k}) = -\mathcal{H} \sum_{i=1}^{\infty} a^i(\tau) \tilde{\theta}_{(i)}(\mathbf{k}). \quad (46)$$

Substituting the expansion in equations (27) and using that from the continuity equation we can show that the first order corrections in  $\delta$  and  $\theta$  coincide, i.e.,  $\tilde{\delta}^{(1)} = \tilde{\theta}^{(1)}$ , we solve the equations by

$$\tilde{\delta}_{(n)}(\mathbf{k}) = \int d^3 q_1 \dots \int d^3 q_n \delta^{(3)}(\mathbf{k} - \mathbf{q}_1 - \dots - \mathbf{q}_n) F_n(\mathbf{q}_1, \dots, \mathbf{q}_n) \tilde{\delta}_{(1)}(\mathbf{q}_1) \dots \tilde{\delta}_{(1)}(\mathbf{q}_n), \quad (47)$$

$$\tilde{\theta}_{(n)}(\mathbf{k}) = \int d^3 q_1 \dots \int d^3 q_n \delta^{(3)}(\mathbf{k} - \mathbf{q}_1 - \dots - \mathbf{q}_n) G_n(\mathbf{q}_1, \dots, \mathbf{q}_n) \tilde{\delta}_{(1)}(\mathbf{q}_1) \dots \tilde{\delta}_{(1)}(\mathbf{q}_n), \quad (48)$$

with  $F_n$  and  $G_n$  symmetrized kernels containing the non-linear nature of the continuity and Euler equations. In particular, by construction the first order kernels are  $F_1 = G_1 = 1$ , and the second order kernels read

$$F_2(\mathbf{q}_1, \mathbf{q}_2) = \frac{5}{7} + \frac{1}{2} \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1 q_2} \left( \frac{q_1}{q_2} + \frac{q_2}{q_1} \right) + \frac{2}{7} \frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)^2}{q_1^2 q_2^2}, \quad (49)$$

$$G_2(\mathbf{q}_1, \mathbf{q}_2) = \frac{3}{7} + \frac{1}{2} \frac{\mathbf{q}_1 \cdot \mathbf{q}_2}{q_1 q_2} \left( \frac{q_1}{q_2} + \frac{q_2}{q_1} \right) + \frac{4}{7} \frac{(\mathbf{q}_1 \cdot \mathbf{q}_2)^2}{q_1^2 q_2^2}. \quad (50)$$

Higher order kernels can be obtained from a recursion relation for which we refer to Bernardeau *et al.* [5]. Note that these solutions are formally the same as we would have obtained by calculating the Feynman diagrams corresponding to  $\phi_{(n)}$  with an Einstein-de Sitter propagator as in equation (41). One of the advantages of this formulation is that we do not have to worry about evaluating the internal structure of every vertical tree. When considering correlation functions of density perturbations evaluated at the same time, we furthermore do not have to worry about the time component or the velocity field perturbations as these will follow automatically from equation (45). However, when considering more general cosmologies with for example a cosmological constant, effective (imperfect) or relativistic fluid(s) with possible vorticity generation we will have to reside to the more general framework. For simplicity, we will in this essay write the general Feynman diagrams but only give the specific expressions in terms of the kernels.

### 3.2 Correlation functions

So far we assumed the initial conditions  $\phi_{in}$  at time  $\tau_{in}$  to be known and approximated its evolution in time. However, as the density field at recombination originated from quantum fluctuations at inflation, we only know the statistical properties of the initial conditions. Gravitational instability theory can, for

this reason, at most approximate the statistical properties of the LSS. Using standard perturbation theory the two- and three-point correlation function of the LSS are

$$\langle \phi^i \phi^j \rangle = \epsilon^2 \langle \phi_{(1)}^i \phi_{(1)}^j \rangle + \epsilon^3 \left[ \langle \phi_{(1)}^i \phi_{(2)}^j \rangle + \langle \phi_{(2)}^i \phi_{(1)}^j \rangle \right] + \epsilon^4 \left[ \langle \phi_{(1)}^i \phi_{(3)}^j \rangle + \langle \phi_{(3)}^i \phi_{(1)}^j \rangle + \langle \phi_{(2)}^i \phi_{(2)}^j \rangle \right] + \mathcal{O}(\epsilon^5) \quad (51)$$

$$\langle \phi^i \phi^j \phi^k \rangle = \epsilon^3 \langle \phi_{(1)}^i \phi_{(1)}^j \phi_{(1)}^k \rangle + \epsilon^4 \left[ \langle \phi_{(1)}^i \phi_{(1)}^j \phi_{(2)}^k \rangle + \langle \phi_{(1)}^i \phi_{(2)}^j \phi_{(1)}^k \rangle + \langle \phi_{(2)}^i \phi_{(1)}^j \phi_{(1)}^k \rangle \right] + \mathcal{O}(\epsilon^5), \quad (52)$$

where we suppressed the momentum dependence corresponding to the superscript indices. For generic initial conditions, in practice a Gaussian field with a small non-Gaussian contribution, these correlators can be expressed in terms of the  $n$ -point correlation functions or  $n$ -dimensional spectra of the initial conditions. For the two-point correlation function the first few correlators give

$$\langle \phi_{(1)}^i \phi_{(1)}^j \rangle = G_l^i(\tau_i, \tau_{in}) G_m^j(\tau_j, \tau_{in}) \langle \phi_{in}^l(\tau_i) \phi_{in}^m(\tau_j) \rangle = P_{(11)}^{ij}(\tau_i, \tau_j), \quad (53)$$

$$\begin{aligned} \langle \phi_{(1)}^i \phi_{(2)}^j \rangle &= \langle \phi_{(1)}^i(\tau_i) \frac{1}{2} \int_{\tau_{in}}^{\tau_j} d\tau' G_k^j(\tau_j, \tau') M_{lm}^k(\tau') \phi_{(1)}^l(\tau') \phi_{(1)}^m(\tau') \rangle \\ &= \frac{1}{2} \int_{\tau_{in}}^{\tau_j} d\tau' G_k^j(\tau_j, \tau') M_{lm}^k(\tau') B_{(111)}^{ilm}(\tau_i, \tau', \tau') \end{aligned} \quad (54)$$

$$\begin{aligned} \langle \phi_{(1)}^i \phi_{(3)}^j \rangle &= \langle \phi_{(1)}^i(\tau_i) \int_{\tau_{in}}^{\tau_j} d\tau' G_k^j(\tau_j, \tau') M_{lm}^k(\tau') \phi_{(1)}^l(\tau') \phi_{(2)}^m(\tau') \rangle \\ &= \frac{1}{2} \int_{\tau_{in}}^{\tau_j} d\tau' \int_{\tau_{in}}^{\tau'} d\tau'' G_k^j(\tau_j, \tau') M_{lm}^k(\tau') G_n^m(\tau', \tau'') M_{op}^n(\tau'') T_{(1111)}^{ilop}(\tau_i, \tau', \tau'', \tau'), \end{aligned} \quad (55)$$

where the correlators  $\langle \phi_{(1)} \phi_{(1)} \rangle$ ,  $\langle \phi_{(1)} \phi_{(1)} \phi_{(1)} \rangle$ , and  $\langle \phi_{(1)} \phi_{(1)} \phi_{(1)} \phi_{(1)} \rangle$  are written as

$$P_{(11)}^{ij}(\tau_i, \tau_j) = G_l^i(\tau_i, \tau_{in}) G_m^j(\tau_j, \tau_{in}) (2\pi)^3 P_{\phi_{in}}^{lm}(k_i) \delta^{(3)}(\mathbf{k}_i + \mathbf{k}_j), \quad (56)$$

$$B_{(111)}^{ijk}(\tau_i, \tau_j, \tau_k) = G_l^i(\tau_i, \tau_{in}) G_m^j(\tau_j, \tau_{in}) G_n^k(\tau_k, \tau_{in}) (2\pi)^3 B_{\phi_{in}}^{lmn}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_k) \delta^{(3)}(\mathbf{k}_i + \mathbf{k}_j + \mathbf{k}_k), \quad (57)$$

$$\begin{aligned} T_{(1111)}^{ijkl}(\tau_i, \tau_j, \tau_k, \tau_l) &= G_m^i(\tau_i, \tau_{in}) G_n^j(\tau_j, \tau_{in}) G_o^k(\tau_k, \tau_{in}) G_p^l(\tau_l, \tau_{in}) \\ &\quad \times (2\pi)^3 T_{\phi_{in}}^{mnop}(\mathbf{k}_i, \mathbf{k}_j, \mathbf{k}_k, \mathbf{k}_l) \delta^{(3)}(\mathbf{k}_i + \mathbf{k}_j + \mathbf{k}_k + \mathbf{k}_l), \end{aligned} \quad (58)$$

with  $T_{\phi_{in}}^{mnop}$  the full trispectrum. Note that often the connected trispectrum is used which is obtained by subtracting the Gaussian trispectrum as discussed below. For the three-point function the first few terms give

$$\langle \phi_{(1)}^i \phi_{(1)}^j \phi_{(1)}^k \rangle = G_l^i(\tau_i, \tau_{in}) G_m^j(\tau_j, \tau_{in}) G_n^k(\tau_k, \tau_{in}) \langle \phi_{in}^l(\mathbf{k}_i) \phi_{in}^m(\mathbf{k}_j) \phi_{in}^n(\mathbf{k}_k) \rangle = B_{(111)}^{ijk}(\tau_i, \tau_j, \tau_k), \quad (59)$$

$$\begin{aligned} \langle \phi_{(1)}^i \phi_{(1)}^j \phi_{(2)}^k \rangle &= \langle \phi_{(1)}^i(\tau_i) \phi_{(1)}^j(\tau_j) \frac{1}{2} \int_{\tau_{in}}^{\tau_k} d\tau' G_l^k(\tau_k, \tau') M_{mn}^l(\tau') \phi_{(1)}^m(\tau') \phi_{(1)}^n(\tau') \rangle \\ &= \frac{1}{2} \int_{\tau_{in}}^{\tau_k} d\tau' G_l^k(\tau_k, \tau') M_{mn}^l(\tau') T_{(1111)}^{ijmn}(\tau_i, \tau_j, \tau', \tau'), \end{aligned} \quad (60)$$

where we suppress the integrals over momentum space corresponding to the interaction terms. In the remainder of this essay we use  $P_{(11)}^{ij}$ ,  $B_{(111)}^{ijk}$ , and  $T_{(1111)}^{ijkl}$  in a slightly different way as we will discard the factor  $(2\pi)^3 \delta^{(3)}(\dots)$ .

### 3.2.1 Gaussian initial conditions

For Gaussian initial conditions correlation functions containing an odd number of initial fields vanish. Using Wick's theorem the initial trispectrum can be expressed in terms of the initial power spectrum, i.e.,

$$T_{G, \phi_{in}}^{ijkl}(\tau_i, \tau_j, \tau_k, \tau_l) = P_{\phi_{in}}^{ij}(\tau_i, \tau_j) P_{\phi_{in}}^{kl}(\tau_k, \tau_l) + P_{\phi_{in}}^{ik}(\tau_i, \tau_k) P_{\phi_{in}}^{jl}(\tau_j, \tau_l) + P_{\phi_{in}}^{il}(\tau_i, \tau_l) P_{\phi_{in}}^{jk}(\tau_j, \tau_k). \quad (61)$$

The trispectrum for non-Gaussian initial conditions is often written as the Gaussian trispectrum plus the so-called connected trispectrum, i.e.,

$$T_{\phi_{in}}^{ijkl} = T_{G, \phi_{in}}^{ijkl} + T_{c, \phi_{in}}^{ijkl}. \quad (62)$$

Using equation (61) and explicitly writing the momentum dependence, we can evaluate these correlation functions for Gaussian initial conditions in terms of the primordial power spectrum. For example, for the  $\langle \phi_{(1)} \phi_{(3)} \rangle$  term we get

$$\begin{aligned} \langle \phi_{(1)}^i \phi_{(3)}^j \rangle &= (2\pi)^3 \delta^{(3)}(\mathbf{k}_i + \mathbf{k}_j) \int_{\tau_{in}}^{\tau_j} d\tau' \int_{\tau_{in}}^{\tau'} d\tau'' \int \frac{d^3q}{(2\pi)^3} G_k^j(\tau_i, \tau') M_{lm}^k(\tau', \mathbf{k}_j, \mathbf{q}, \mathbf{k}_j - \mathbf{q}) G_n^m(\tau', \tau'') \\ &\times M_{op}^n(\tau'', \mathbf{k}_j - \mathbf{q}, -\mathbf{q}, \mathbf{k}_j) P_{(11)}^{ip}(\tau_i, \tau'', k_i) P_{(11)}^{lo}(\tau', \tau'', q), \end{aligned} \quad (63)$$

where one of the Wick contractions vanishes for our set of differential equations (27).

### 3.2.2 Feynman rules

The calculations above can be systematically expressed in terms of Feynman diagrams with the following Feynman rules. The  $\epsilon^i$  order approximation of the  $n$ -point function with generic initial conditions can be determined as follows:

1. Draw all Feynman diagrams with  $n$  legs connecting to the upper horizontal line and up to  $i$  legs connecting to the lower horizontal line. The interaction vertex has one upper leg and two lower legs.
2. Label all legs with a momentum vector, respecting conservation of momentum at the vertices. Every leg corresponds to a Green function or propagator  $G$ . Label all vertices with a time parameter. Every vertex correspond to a interaction term  $M$  and an integral over time.
3. Connect all lower legs with a comb. The comb corresponds to  $(2\pi)^3$  times the full  $i$ -dimensional spectrum (power spectrum, bi-spectrum, trispectrum, etc.).
4. Multiply by the symmetry factor analogous to perturbative quantum field theory.
5. Integrate over the undetermined momenta, running in the loops, i.e.,  $\int \frac{d^3q}{(2\pi)^3}$ . In perturbative quantum field theory this corresponds to virtual particles.
6. Multiply by an extra Dirac delta function  $\delta^{(3)}(\mathbf{k}_1 + \dots + \mathbf{k}_i)$ .

For Gaussian initial conditions, we can use Wick's theorem expressed for the trispectrum in equation (61) to modify step 3 by: "Diagrams with  $i$  being odd vanish. For even  $i$ , connect the lower legs in all possible pairings with a comb with two legs. The comb corresponds to the power spectrum." For an example see figure 8 corresponding to the  $\langle \phi_{(1)} \phi_{(1)} \rangle$  and  $\langle \phi_{(1)} \phi_{(3)} \rangle$  correlation functions for Gaussian initial conditions evaluated in equation (63). Note that we integrate over the undetermined momentum vector  $\mathbf{q}$  which appears in the power spectrum term  $P_{(11)}^{lo}$ . This can be seen as integrating over the initial power spectrum representing our ignorance of the initial conditions. This is analogous to momentum running in loops in quantum field theory.

Finally for generic initial conditions we can express the correlation function in terms of the connected spectra by performing both the Gaussian rule and the generic rule where the comb with  $i$  legs now correspond to the connected spectrum. This is the notation we will adopt in this essay as it most explicitly describes both the Gaussian and non-Gaussian effect on correlation functions.

For fluctuations evolving in an Einstein-de Sitter universe, we can simplify the calculation considerably using the kernel notation. Note that we only have to consider  $\delta$  and do not have to perform the internal evaluation of the vertical trees. We can identify  $\phi_{(i)}$  with  $\delta_{(i)}$ , and let the coupling  $M$  correspond to a vertex with one upper leg and  $i$  lower legs directly connecting the upper and lower horizontal lines corresponding to a factor

$$F_i(\mathbf{q}_1, \dots, \mathbf{q}_i). \quad (64)$$

We do not identify a time with the vertex and do not integrate over it. Note that we do have to consider the symmetry factor. Neglecting a factor  $(2\pi)^3 \delta^{(3)}(\mathbf{k}_1 + \dots + \mathbf{k}_i)$  gives the LSS density  $i$ -dimensional

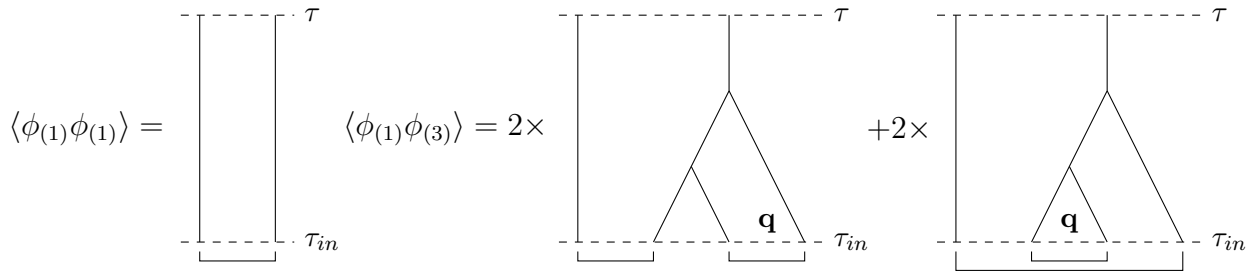


Figure 8: Correlation functions in terms of Feynman diagrams, assuming Gaussian initial conditions. Note that the second Feynman diagram in  $\langle\phi_{(1)}\phi_{(3)}\rangle$  vanishes for our choice of equations of motion.

spectrum. In quantum field theory this procedure is known as ‘amputating the legs’. The  $\langle\delta_{(1)}\delta_{(3)}\rangle$  correlation function for Gaussian initial conditions now becomes

$$\langle\delta_{(1)}(\tau, \mathbf{k}_1)\delta_{(3)}(\tau, \mathbf{k}_2)\rangle = \left[6 \int d^3q F_3(\mathbf{k}, \mathbf{q}, -\mathbf{q})P_{(11)}(k_1)P_{(11)}(q)\right] (2\pi)^3\delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2). \quad (65)$$

For non-Gaussian corrections we would have to add a term containing the initial connected trispectrum. We illustrate this expression in terms of a Feynman diagram in figure 9.

### 3.3 The matter power spectrum

The matter power spectrum in terms of Feynman diagrams to loop order is given in figure 11,

$$\langle\delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\rangle = (2\pi)^3[P_{(11)}(k_1) + P_{(12)}(k_1) + P_{(22)}(k_1) + P_{(13)}(k_1) + \dots]\delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2). \quad (66)$$

The first diagram corresponds to the linearly extrapolated power spectrum. The second diagram corresponds to a terms vanishing for Gaussian initial conditions and gives the first correction for non-Gaussian initial conditions. The third and fourth diagrams correspond to two loop order terms forming from gravitational collapse.

In an Einstein-de Sitter universe, we can express these diagrams in terms of the kernels  $F_n$

$$\langle\delta_{(1)}(\mathbf{k}_1)\delta_{(1)}(\mathbf{k}_2)\rangle = (2\pi)^3P_{(11)}(k_1)\delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2), \quad (67)$$

$$P_{(12)}(k) = 2 \int d^3q F_2(\mathbf{q}, \mathbf{k} - \mathbf{q})B_{(111)}(-\mathbf{k}, \mathbf{q}, \mathbf{k} - \mathbf{q}), \quad (68)$$

$$P_{(22)}(k) = 2 \int d^3q F_2(\mathbf{k} - \mathbf{q}, \mathbf{q})P_{(11)}(|\mathbf{k} - \mathbf{q}|)P_{(11)}(q), \quad (69)$$

$$P_{(13)}(k) = 6 \int d^3q F_3(\mathbf{k}, \mathbf{q}, -\mathbf{q})P_{(11)}(k)P_{(11)}(q). \quad (70)$$

In terms of the growing mode  $D$ , the linear extrapolated power spectrum is  $P_{(11)}(k) = D^2(\tau)P_{in}(k)$ . The ratio  $\beta_\delta = P_\delta^G/P_\delta^{NG}$ , with  $P_\delta^G = P_{(11)} + P_{(22)} + P_{(13)}$  the Gaussian contribution and  $P_\delta^{NG} = P_{(12)}$  the non-Gaussian contribution, is a measure of the non-Gaussianity contribution to the matter power spectrum. Evaluation of  $\beta_\delta$  shows that primordial non-Gaussianities significantly influence the matter power spectrum on small scales, while leaving the matter power spectrum on large scales invariant. The approximation developed in equation (66) in an Einstein-de Sitter universe is in good agreement with  $N$ -body simulation with primordial non-Gaussianities of local type (see figure 10).

### 3.4 The matter bispectrum

The matter bispectrum can be expressed in terms of Feynman diagram, as in figure 12, and

$$\langle\delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3)\rangle = (2\pi)^3[B_{(111)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + B_{(112)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + \dots]\delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3). \quad (71)$$

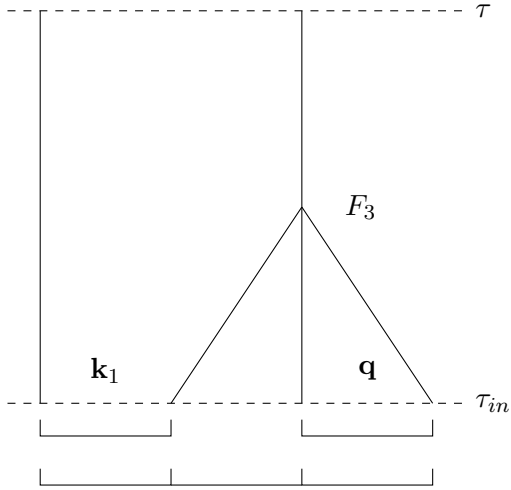


Figure 9: Correlation function  $\langle \delta_{(1)} \delta_{(3)} \rangle$  in terms of Feynman diagrams in the Einstein-de Sitter universe.

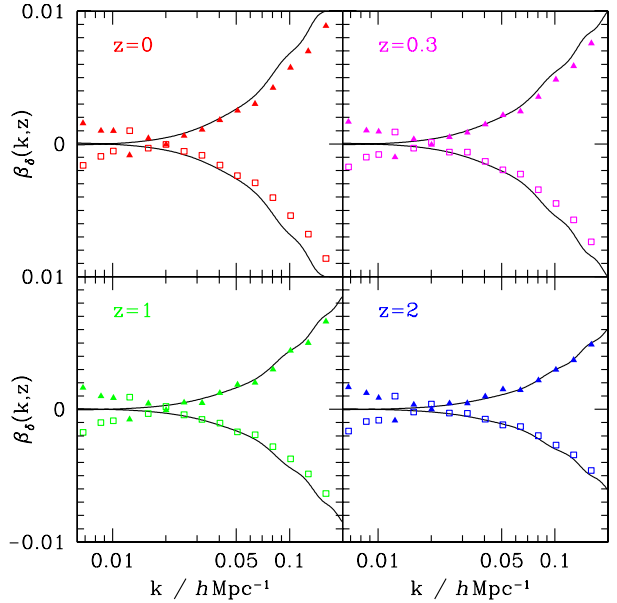


Figure 10: The fraction  $\beta_\delta(k, z) = P_\delta^G(k)/P_\delta^{NG}(k)$  for local non-Gaussianity with  $f_{NL} = \pm 100$ , at different redshifts. The solid line denotes 1-loop standard perturbation theory predictions [14].

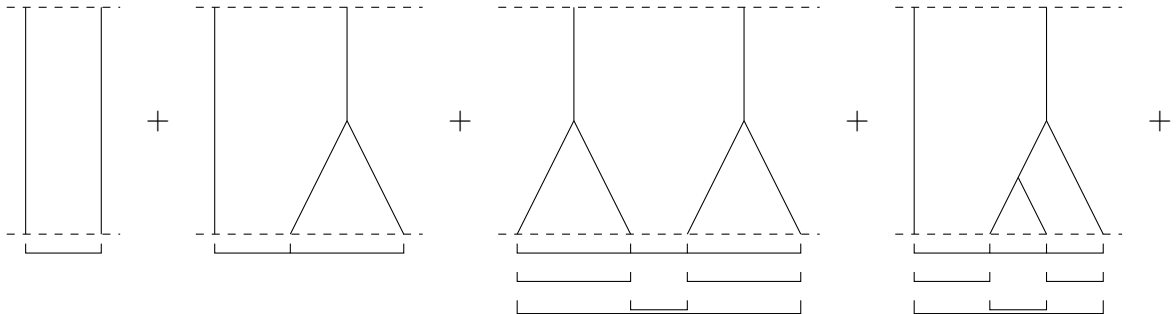


Figure 11: The Feynman diagrams corresponding to the matter power spectrum to sub-leading order.

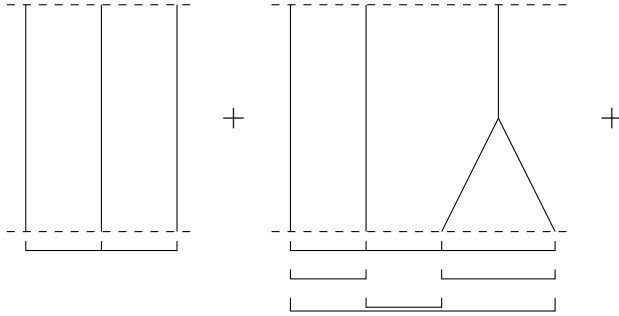


Figure 12: The Feynman diagrams corresponding to the matter bispectrum to sub-leading order.

The first term,  $B_{(111)}$ , vanishes for Gaussian initial conditions and is the first correction due to primordial non-Gaussianities. The second term is due to gravitational instability. In an Einstein-de Sitter universe, we express this term in terms of the kernel  $F_n$ ,

$$B_{(112)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = 2F_2(\mathbf{k}_1, \mathbf{k}_2)P_{(11)}(k_1)P_{(11)}(k_2) + 2 \text{ perm} \\ + \int \frac{d^3q}{(2\pi)^3} F_2(\mathbf{q}, \mathbf{k}_3 - \mathbf{q})T_0(\mathbf{q}, \mathbf{k}_3 - \mathbf{q}, \mathbf{k}_1, \mathbf{k}_2) + 2 \text{ perm}, \quad (72)$$

with  $T_0$  the primordial trispectrum.

The contribution of primordial non-Gaussianities in the matter bispectrum is normally measured in terms of the reduced bispectrum

$$Q(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \frac{B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)}{P_{(11)}(k_1)P_{(11)}(k_2) + P_{(11)}(k_1)P_{(11)}(k_3) + P_{(11)}(k_2)P_{(11)}(k_3)}. \quad (73)$$

For Gaussian initial conditions the reduced bispectrum is to tree-level independent of time, and in the equilateral configuration  $k_1 = k_2 = k_3$  equal to  $4/7$ . Different types of primordial non-Gaussianity lead to different reduced bispectra. See figure 13 for a comparison between perturbation theory and  $N$ -body simulations and different reduced bispectra for different primordial non-Gaussianity types. The 1-loop correction is a reasonable approximation. The reduced bispectrum is a sensitive probe for primordial non-Gaussianities and allows us to distinguish different types of non-Gaussianities (see figure 14). We observe that equation (71) for an Einstein-de Sitter universe is a good approximation and that primordial non-Gaussianities lead to significant corrections to the matter bispectrum on large scales.

### 3.5 Effective field theory

The LSS equations are traditionally solved using standard perturbation theory. This theory works well in the linear and mildly non-linear region. However, it has been shown that it breaks down on short scales  $k > k_{NL} \approx (10 \text{ Mpc})^{-1}$  as the perturbations become large due to non-linear evolution [4, 8, 9]. If perturbation theory breaks down at short scales it also becomes untrustworthy for long wavelengths due to couplings of wavelengths at quadratic order. This problem has recently been addressed with effective field theory. This theory smoothens short-wavelength scales ( $k > k_{NL}$ ), which cause problems in standard perturbation theory, and only model long-wavelength modes ( $k < k_{NL}$ ), which can be evaluated even when standard perturbation theory breaks down. Effective field theory is a common tool in high energy physics in cases where we either do not know the ultra-high energy physics and renormalise the theory or would like to use a simplified version of a known underlying theory.

We here smooth the short-wavelength physics as discussed by Carrol *et al.* [9]. We start by splitting the density and velocity fluctuations in short- and long-wavelength modes. Let the long-wavelength modes be defined as a convolution of the field  $\phi$  with a smoothing function depending on the cut-off  $\Lambda \approx k_{NL}$ ,

$$\phi_L^i(x) = \int d^3y W_\Lambda^i(x-y)\phi(y). \quad (74)$$

Define the short wavelength modes as  $\phi_S = \phi - \phi_L$ . In Fourier space the convolution becomes a multiplication,

$$\phi_L^i(\mathbf{k}_i) = W_\Lambda^i(\mathbf{k}_i)\phi^i(\mathbf{k}_i) \text{ and } \phi_S(\mathbf{k}) = \phi(\mathbf{k}) - \phi_L(\mathbf{k}). \quad (75)$$



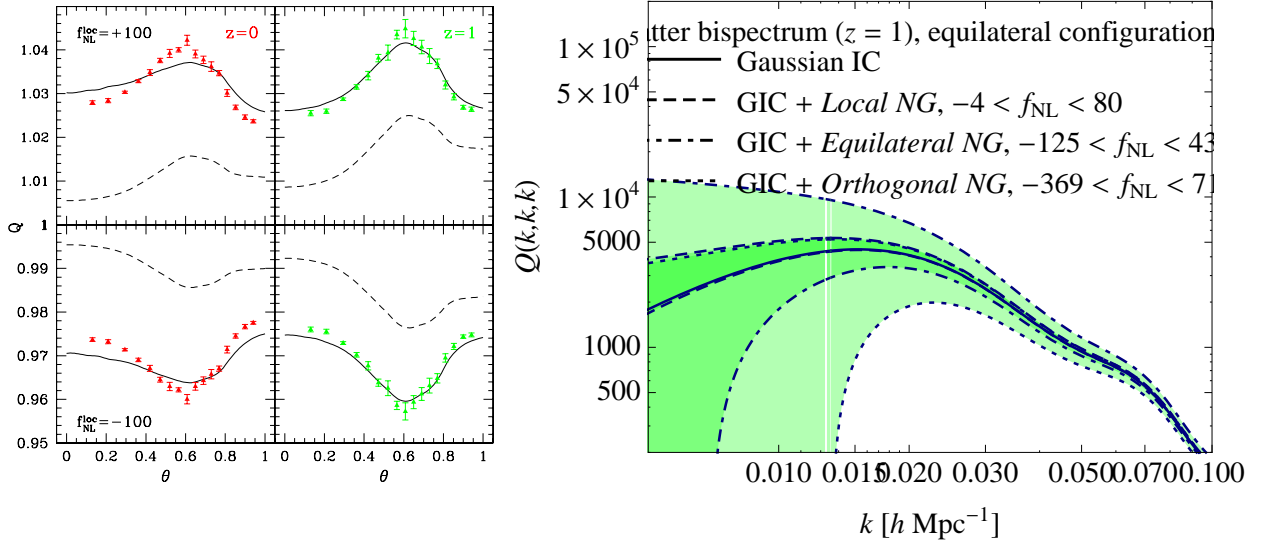
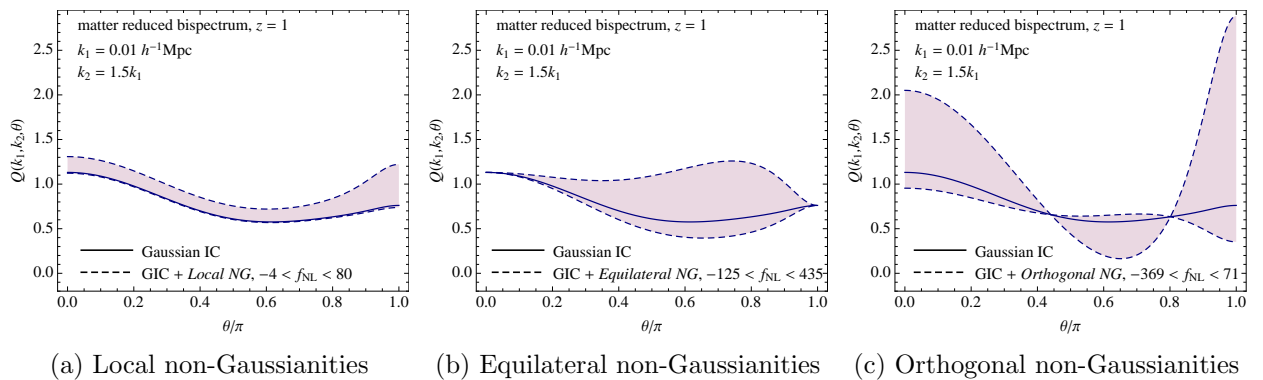


Figure 13: The left plot: reduced matter bispectrum as a function of the angle  $\theta$  between  $\mathbf{k}_1$  and  $\mathbf{k}_2$  with  $k_1 = 0.094h\text{Mpc}^{-1}$  and  $k_2 = 1.5k_1$  for different redshifts (left  $z = 0$ , right  $z = 1$ ) and local non-Gaussianity (upper  $f_{NL} = 100$ , lower  $f_{NL} = -100$ ). The crosses correspond to results from  $N$ -body simulations, the dashed lines to tree-level standard perturbation theory and the solid lines to 1-loop standard perturbation theory [14]. The right plot: reduced matter bispectrum in the equilateral configuration  $k = k_1 = k_2 = k_3$  for different primordial non-Gaussian conditions.



(a) Local non-Gaussianities

(b) Equilateral non-Gaussianities

(c) Orthogonal non-Gaussianities

Figure 14: Reduced matter bispectrum three-level predictions for different types of primordial non-Gaussianities at redshift  $z = 1$ , assuming  $k_1 = 0.01h\text{Mpc}^{-1}$ ,  $k_2 = 1.5k_1$  as a function of the angle  $\theta$  between  $\mathbf{k}_1$  and  $\mathbf{k}_2$ . The solid line is the prediction for Gaussian initial conditions. The dashed line corresponds to the bounds explained in the image corresponding to the 95% confidence limits of the WMAP survey [28].

Using the fact that  $\mathcal{D}$  is diagonal we can write the equation of motion in short- and long-wavelength modes

$$0 = \mathcal{D}_j^i \phi_L^j - \frac{1}{2} W_\Lambda^i M_{jk}^i \phi^j \phi^k = \mathcal{D}_j^i \phi_L^j - \frac{1}{2} W_\Lambda^i M_{jk}^i \phi_L^j \phi_L^k - W_\Lambda^i M_{jk}^i \phi_S^j \phi_L^k - \frac{1}{2} W_\Lambda^i M_{jk}^i \phi_S^j \phi_S^k. \quad (76)$$

$N$ -body experiments have shown that the short-wavelength modes  $\phi_S$  are highly dependent on the long-wavelength modes while the long-wavelength modes are only weakly dependent on the short-wavelength modes [29]. For this reason we can treat the short-wavelength modes as a functional of the long-wavelength modes  $\phi_S[\phi_L]$  and expand  $\phi_S$  in a Taylor series around its value when  $\phi_L = 0$ , i.e.,

$$\phi_S^i = \phi_{S_0}^i(\tau) + \int_{\tau_{in}}^\tau d\tau' \frac{\partial \phi_S^i(\tau)}{\partial \phi_L^j(\tau')} \Big|_{\phi_L=0} \phi_L^j(\tau') + \dots \quad (77)$$

The first term denotes the short wavelength field for vanishing long-wavelength field. The second term represents the first order correction for a nonzero long-wavelength field. Both  $\phi_{S_0}$  and  $\partial \phi_S^i(\tau)/\partial \phi_L^j(\tau')$  are not predicted by effective field theory and should be obtained from  $N$ -body simulations. Note that the second term does not have to be local. To first order we should think about these terms as the viscosity in fluid dynamics.

Upon substituting the Taylor expansion in the equation of motion we obtain the effective equation of motion,

$$0 = \mathcal{D}_j^i \phi_L^j - \frac{1}{2} W_\Lambda^i M_{jk}^i \phi_L^j \phi_L^k - W_\Lambda^i M_{jk}^i \phi_{S_0}^j \phi_L^k - \frac{1}{2} W_\Lambda^i M_{jk}^i \phi_{S_0}^j \phi_{S_0}^k - \quad (78)$$

$$- W_\Lambda^i M_{jk}^i \phi_{S_0}^j \int_{\tau_{in}}^\tau d\tau' \frac{\partial \phi_S^k(\tau)}{\partial \phi_L^l(\tau')} \phi_L^l(\tau') - W_\Lambda^i M_{jk}^i \phi_L^j \int_{\tau_{in}}^\tau d\tau' \frac{\partial \phi_S^k(\tau)}{\partial \phi_L^l(\tau')} \phi_L^l(\tau') + \dots \quad (79)$$

We can interpret this equation as a set of effective interactions  $\tilde{\mathcal{D}}, \tilde{M}, \tilde{O}$  related by the equation

$$0 = \tilde{\mathcal{D}}_j^i \phi_L^j - \frac{1}{2} \tilde{M}_{jk}^i \phi_L^j \phi_L^k - \frac{1}{3!} \tilde{O}_{jkl}^i \phi_L^j \phi_L^k \phi_L^l. \quad (80)$$

We can treat this differential equation as in the standard perturbation theory discussion and find effective approximations for the correlation functions.

Carroll *et al.* [9] have shown that the first effective correction to the equation of motion adds linear term to the equation of motion,

$$\tilde{\mathcal{D}} = \mathcal{D} + C(\tau'|\mathbf{k}), \quad (81)$$

while the interaction remains unaffected. Baumann *et al.* [4] and Carrasco *et al.* [8] have shown that this corresponds to an imperfect fluid by which we can relate  $C$  to the speed of sound, the viscosity, and the heat conductivity coefficients of the fluid

$$C(\tau'|\mathbf{k}) = \begin{pmatrix} \xi^\delta & \xi^\theta \\ k^2 c_s^2 & -k^2 \frac{c_v^2}{\mathcal{H}} \end{pmatrix}, \quad (82)$$

where the parameters  $c_s, c_v, \xi^\delta$  and  $\xi^\theta$  should be obtained from small scale  $N$ -body simulations. For a more sophisticated analysis of the corrections see Carroll *et al.* [9].

## 4 Probing primordial non-Gaussianities in the LSS

Primordial non-Gaussianities can significantly influence the statistics of the LSS matter density field. However, as the matter in our universe largely consists of a not yet observed kind (dark matter), we are unable to map the matter density field directly. Large-scale structure surveys generally survey the galaxy distribution in our universe. In this section we study the effect of primordial non-Gaussianities on bounded objects in which galaxies generally reside. We also shortly discuss their effect on voids.

Historically, the scientific community focused on the mass function of very massive structures [42]. Very massive virialized objects correspond to rare high peaks in the density field, which are sensitive

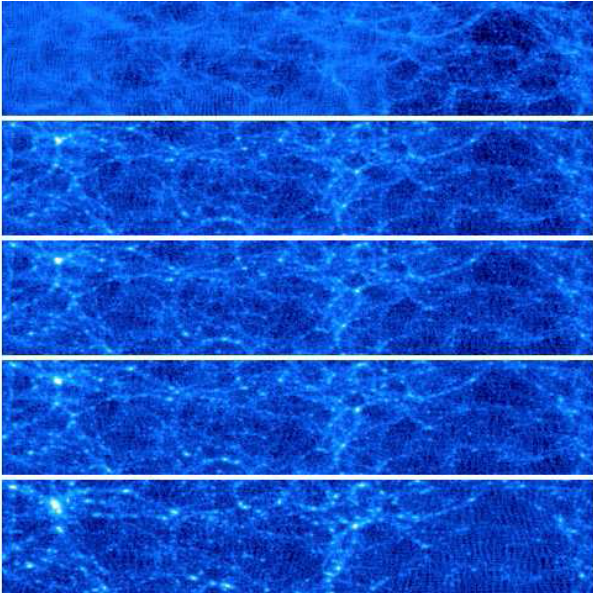


Figure 15: A slice through an  $N$ -body simulation at  $z = 0$  for non-Gaussian initial conditions of local type, with the same phase information, by Dalal *et al.* [13]. From top to bottom  $f_{NL} = -5000, -500, 0, 500, 5000$ . Each slice is  $375h^{-1}Mpc$  wide and  $80h^{-1}Mpc$  high and deep.

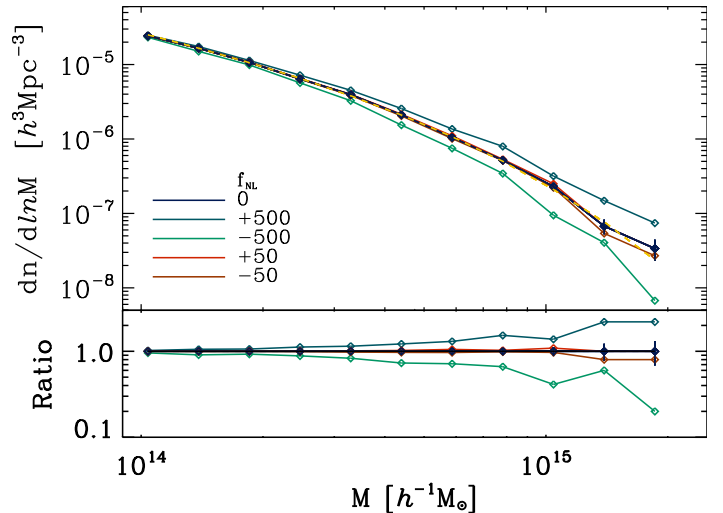


Figure 16: The halo mass function at  $z = 0$  for  $N$ -body simulations with non-Gaussian initial conditions of local type with different  $f_{NL}$ , by Dalal *et al.* [13]. The top plane shows the mass function. The bottom plane shows the ratio of the mass function with respect to the Gaussian case.

to non-Gaussian contributions. For large empty voids we observe a similar effect. We shortly review this effect in sections 4.1 and 4.2. Although these hypotheses are supported by  $N$ -body simulations, the approach in practice suffers from low statistics and selection effects. With the prospect of the next generation of LSS surveys observing and studying tens of thousands of clusters, it will become achievable to measure higher order correlation functions, which we review in section 4.3. In particular, we discuss the local biasing model in combination with the skewness and kurtosis of the galaxy distribution. We subsequently discuss the recently discovered biasing scaling by Dalal *et al.* [13], which has been used to constrain primordial non-Gaussianities.

## 4.1 Halo mass function

The number densities of bounded objects such as halos form a sensitive probe for primordial non-Gaussianities. This can be observed in figure 15 where we plot a slice through an  $N$ -body simulation with varying non-Gaussian initial conditions of the local type. Positive amplitude  $f_{NL}$  leads to an enhanced formation of virialized objects and smaller voids, whereas negative  $f_{NL}$  leads to a suppressed formation of bounded objects and larger voids than the Gaussian case. We can see this effect in more detail in halo mass function  $\frac{dn(M)}{d \ln M}$  denoting the density of halos with mass  $M$  (see figure 16). The halo mass function is in particular sensitive to primordial non-Gaussianities for massive objects.

There exist several fitting formulas for the halo mass function for Gaussian initial conditions such as the extended Press-Schechter formalism [6, 13, 14, 38]. We here briefly review the extended Press-Schechter formalism, and discuss its extensions to understand the effect of primordial non-Gaussianities on bounded objects and voids.

### 4.1.1 Gaussian initial conditions

In Eulerian perturbation theory, to linear order, the density field evolves as

$$\delta(\mathbf{x}, t) = D(t)\delta_{in}(\mathbf{x}), \quad (83)$$

with  $\delta_{in}(\mathbf{x})$  the linearly extrapolated initial density field and  $D(t)$  the growing mode normalized to unity at  $t = t_0$  our current time. This is a valid approximation in the linear regime of structure formation, but

breaks down in the non-linear regime. For the non-linear regime, it is common to use the spherical or elliptic collapse model. According to the spherical collapse model, regions in which the linear approximation satisfies

$$\delta(\mathbf{x}, t) = D(t)\delta_{in}(\mathbf{x}) > \delta_c \approx 1.686 \quad (84)$$

will have collapsed into bounded objects at time  $t$ . Under these approximations, the Press-Schechter formalism approximates the number density of bounded objects. Hence it answers the questions: how many bounded objects, on average, will a volume contain and what is the mass distributions of these objects?

In the Press-Schechter formalism, the initial density field is generally smoothed on a scale  $R$  with a window function  $W(\mathbf{x}, R)$  normalized by  $\int W(\mathbf{x}, R)d^3x = 1$  for all  $R$ . Formally, the density field is convolved with the window function in real space

$$\delta(\mathbf{x}, R) = \int \delta_{in}(\mathbf{x}')W(\mathbf{x} + \mathbf{x}', R)d^3x', \quad (85)$$

and multiplied with its Fourier transform in Fourier space

$$\delta(\mathbf{k}, R) = \delta_{in}(\mathbf{k})W(\mathbf{k}, R), \quad (86)$$

where the latter  $W$  is the Fourier transform of the window function. In practice it is common to use window functions which only depend on the product  $kR$  and write  $W(kR)$ . Examples of window functions are the top hat, Gaussian or so called sharp  $k$ -filter. The sharp  $k$ -filter will be used later on and is defined as

$$W_{sk}(kR) = \Theta(1 - kR), \quad (87)$$

with  $\Theta$  the Heaviside function. For the smoothed field we can define the variance as

$$\sigma^2(R) = \langle \delta^2(\mathbf{x}, R) \rangle = \frac{1}{2\pi^2} \int P_{in}(k)W^2(kR)k^2dk, \quad (88)$$

where we have moved to Fourier space, used the definition of the initial power spectrum  $P_{in}$ , integrated over a Dirac delta function and transformed to spherical coordinates. The smoothing can be related to a mass scale,  $M = \gamma_f \bar{\rho} R^3$  with  $\gamma_f$  a constant depending on the window function. For the Gaussian filter  $\gamma_f = (2\pi)^{3/2}$  and for the sharp  $k$ -filter  $\gamma_f = 6\pi^2$ . We can use this relation to identify smoothed variables  $\sigma(R)$  and  $\delta(\mathbf{x}, R)$  with the mass scale variables  $\sigma_M$  and  $\delta_M(\mathbf{x})$  respectively.

Using the spherical collapse model, it is reasonable to identify the number density of halos with mass  $M$  at time  $t$  with the number density of peaks in  $\delta_M(\mathbf{x})$  with critical value  $\delta_c(t) = \delta_c/D(t)$ . Bardeen et al. [23] derived statistics for the distribution of peaks. This approximation has however a caveat since some peaks become absorbed by another peak on a larger mass scale and should not be identified as a halo. This is known as the cloud-in-cloud problem. Press and Schechter [38] avoided this problem by stating that the probability for  $\delta_M(\mathbf{x}) > \delta_c(t)$  is equal to the mass fraction contained in halos with mass greater than  $M$ , i.e., the mass fraction contained in halos  $F$  at time  $t$  satisfies

$$F[> M, t] = P[\delta_M(\mathbf{x}) > \delta_c(t)], \quad (89)$$

with  $P$  the PDF of the initial density field. Geometrically we can imagine this as a barrier at height  $\delta_c(t)$  which is lowered as time evolves, letting more and more regions collapse. However, as  $\delta_c(t) > 0$  for all  $t$ , the Press-Schechter formalism states that only overdense regions end up in halos. This is unreasonable as underdense regions can be enclosed by overdense regions. Press and Schechter compensated for this caveat by introducing a fudge factor 2,

$$F[> M, t] = 2P[\delta_M(\mathbf{x}) > \delta_c(t)]. \quad (90)$$

Now all mass ends up in halos as  $\delta_c(t)$  approaches 0. Given this fudge factor, the differential mass function for Gaussian initial conditions reads

$$\frac{dn(M, t)}{dM} = \frac{\bar{\rho}}{M} \frac{\partial F[> M, t]}{\partial M} = \sqrt{\frac{2}{\pi}} \frac{\bar{\rho}}{M^2} \text{Exp} \left[ -\frac{\delta_c(t)}{2\sigma_M^2} \right] \left| \frac{d \ln \sigma_M}{d \ln M} \right|. \quad (91)$$

This is often written as

$$\frac{dn(M, t)}{dM} = \frac{\bar{\rho}}{M} f(\nu) \left| \frac{d \ln \nu}{d \ln M} \right|, \quad (92)$$

with  $f(\nu) = \sqrt{\frac{2}{\pi}} \nu e^{-\nu^2/2}$  known as the multiplication factor and dimensionless parameter  $\nu = \frac{\delta_c(t)}{\sigma_M}$ .

The fudge factor proposed by Press and Schechter was explained by Bond *et al.* [6], by their excursion set theory. They studied  $\delta_M(\mathbf{x})$  for a fixed position  $\mathbf{x}$  as a function of mass scale  $M$  for a sharp  $k$ -filter  $W_{sk}$ . For a so-called sharp  $k$ -filter  $W_{sk}$ , the density field

$$\delta_M(\mathbf{x}) = \int d^3k W(kR) \delta_{in}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} = \int_{k < k_c} \delta_{in}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} \quad (93)$$

where  $k_c = 1/R$ . In the limit  $R \rightarrow \infty$ , the density goes to zero. For a Gaussian random field,  $\delta_M(\mathbf{x})$  performs a random walk as  $R$  decreases to 0 since the different Fourier modes of a Gaussian field are independent. The question whether a mass element resides in a halo at time  $t$  can now be identified with the probability for the random walk to cross the boundary  $\delta_c(t)$  before that time. A detailed analysis of the random walk yields the Press-Schechter halo mass function with the fudge factor 2. Window functions other than the sharp  $k$ -filter give correlated walks for which the halo mass function cannot be analytically evaluated this way. Such an analysis generally requires a Monte Carlo estimation. This formalism is generally known as the extended Press-Schechter formalism.

The extended Press-Schechter formalism gives a halo mass function that roughly agrees with numerical simulations. It however overpredicts the high-mass regime while underpredicting the low-mass regime. The approximation can be improved by considering the elliptic collapse model. Unfortunately such a model, even with a sharp- $k$  filters, correspond to a moving barrier problem, i.e.,  $\delta_c(t)$  now also depends on scale  $R$ . We are unable to handle moving boundary problems analytically and again have to perform a Monte Carlo simulations [44].

#### 4.1.2 Non-Gaussian initial conditions

Several methods have been pursued to generalize the extended Press-Schechter argument to non-Gaussian initial conditions. The most naive way to generalize would be to replace the Gaussian PDF in the definition of  $F$  by a non-Gaussian distribution function. Such an attempt faces however a few problems. In the first place, the fudge factor will even for a sharp  $k$ -filter no longer be equal to two, as different  $k$ -modes of the initial density field will be correlated. Secondly, the generalized halo mass function will suffer from the same inaccuracies as the original model.

A less direct generalization has been proposed in which we acknowledge the inaccuracies of the original Press-Schechter formalism and only use its general form, by using a Gaussian halo mass function and multiplying it with the departure from the Gaussian case, i.e.,

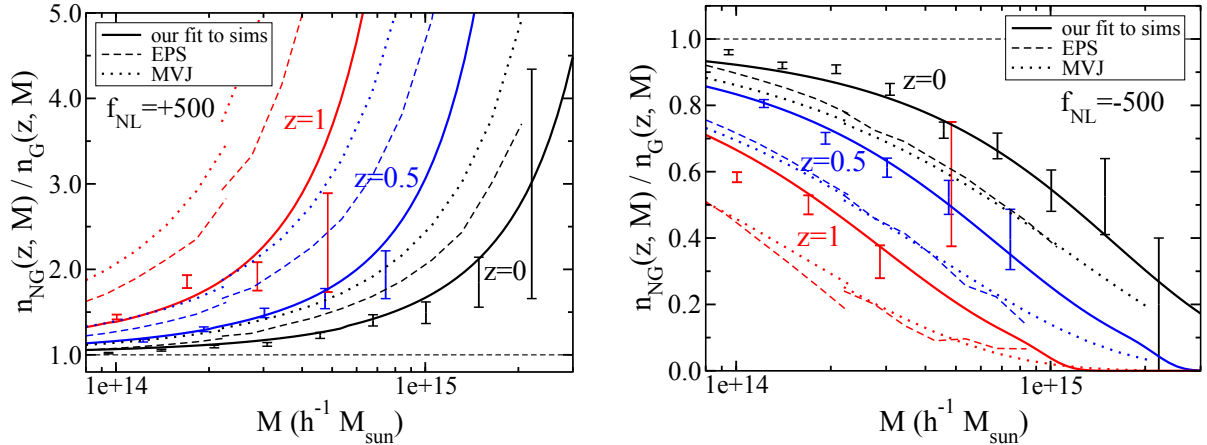
$$n_{NG}(M, t) = \frac{\frac{d}{dM} F_{NG}[> M]}{\frac{d}{dM} F_G[> M]} n_G(M, t) = \frac{f_{NG}(\nu)}{f_G(\nu)} n_G(M, t). \quad (94)$$

This method can either be performed numerically or analytically. Numerically, we can take the Gaussian halo mass function from an  $N$ -body simulation (for example Jenkins *et al.* [22]) and estimate the ratio numerically using several non-Gaussian  $N$ -body simulations [13]. Analytically, Matarrese *et al.* [33] (known as MVJ) approximated the non-Gaussian PDF with an Edgeworth expansion and evaluated the ratio in terms of the skewness

$$\left( \frac{dn}{d \ln M} \right)_{MVJ} = 2 \frac{\bar{\rho}}{M} P_G \frac{\delta_*}{\sigma_M} \left[ \frac{1}{6} \frac{\delta_*^3}{\delta_c} \left| \frac{dS_{3,M}}{d \ln M} \right| + \delta_* \left| \frac{d\sigma_M}{d \ln M} \right| \right], \quad (95)$$

with  $\delta_* = \frac{\delta_c}{\sqrt{1 - S_{3,M} \delta_c/3}}$  and  $S_{3,M} = \frac{\langle \delta_M^3 \rangle}{\sigma_M^4}$ .

We compare the numerical and analytic evaluations of the ratio with the halo mass function of a  $N$ -body simulation in figure 17. We observe that both methods largely agree but overestimate the effect of primordial non-Gaussianities.



(a) The ratio between non-Gaussian and Gaussian halo mass for positive  $f_{NL}$ . (b) The ratio between non-Gaussian and Gaussian halo mass for negative  $f_{NL}$ .

Figure 17: The ratio of a non-Gaussian and Gaussian halo mass function as function of scale. The points with error bars denote numerically estimated values. The solid line represents a fit through the points. The dotted lines represent numerical and analytical approximations. Analysis by Dalal *et al.* [13]

## 4.2 Void abundance

Primordial non-Gaussianities strongly influence the halo mass function of very massive halos. On the other end of the mass distribution we observe in figure 15 that also the number density of large cosmic voids is affected. Generally primordial non-Gaussianities of the local type with positive  $f_{NL}$  lead to smaller voids whereas non-Gaussianities with negative  $f_{NL}$  lead to larger voids than the Gaussian case. This is a useful observation, as voids have a small density perturbations and evolve linearly. The statistics of the initial conditions should for this reason still reside in these voids. See the review by Desjaques and Seljak [14] for an estimate based on analysis along the lines of the excursion set argument.

A joint estimate of primordial non-Gaussianity, using both halo and void abundances would be an interesting development. There are however still several caveats in using voids abundance. There is no unique theoretical definition of a void. Voids can be the region where the density field falls below some threshold, where the eigenvalues of the Hessian of the gravitation potential are negative indicating outflow of matter, or they can be identified by a division of the space by filaments and clusters such as the watershed method [47], or Nexus+ [10]. For an overview of different void detection schemes see [12]. These voids detection methods generally do not agree among each other in  $N$ -body simulations. Given a void region, it is non-trivial to define its size as it can assume irregular shapes. Finally, identifying voids in high redshift surveys is difficult as voids contain a very limited number of galaxies which can be used to trace them.

## 4.3 Halo biasing

We now concentrate on the spatial distribution of the galaxies in the large-scale structure. Galaxies only form in dense regions, and are for that reason a biased tracer of the underlying density field. Hence in order to measure primordial non-Gaussianities in the LSS by means of its spatial distribution, we need to relate the statistics of the matter distribution  $\delta$  to the galaxy distribution

$$\delta_g(\mathbf{x}) = \frac{n_g(\mathbf{x}) - \langle n_g(\mathbf{x}) \rangle}{\langle n_g(\mathbf{x}) \rangle} \quad (96)$$

smoothed on some scale. The relation between the matter and the galaxy distribution, known as the biasing relation, is still not fully understood. We here shortly review the traditional local biasing model and more recent developments suggesting an implicit dependence of galaxy formation on primordial non-Gaussianities.

### 4.3.1 The local biasing model and the galaxy bispectrum

Until recently, it was assumed that the perturbations in the galaxy distribution  $\delta_g$  only depend on the local behaviour of the density field, i.e., there exists a function  $f$  such that

$$\delta_g(\mathbf{x}) = f[\delta(\mathbf{x})]. \quad (97)$$

On large scales the biasing effect is assumed to be small. We can therefore approximate  $f$  by its Taylor expansion in terms of biasing parameters  $b_i$  [16],

$$\delta_g(\mathbf{x}) = f[\delta(\mathbf{x})] = b_1\delta(\mathbf{x}) + \frac{1}{2}b_2\delta(\mathbf{x})^2 + \frac{1}{3!}b_3\delta(\mathbf{x})^3 + \dots \quad (98)$$

The correlation functions of the galaxy perturbation  $\delta_g$  can be expressed in terms of correlation functions of  $\delta$

$$\langle \delta_g(\mathbf{x}_1)\delta_g(\mathbf{x}_2) \rangle = b_1^2\langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2) \rangle + b_1b_2 [\langle \delta(\mathbf{x}_1)\delta^2(\mathbf{x}_2) \rangle + \langle \delta^2(\mathbf{x}_1)\delta(\mathbf{x}_2) \rangle] + \dots \quad (99)$$

$$\langle \delta_g(\mathbf{x}_1)\delta_g(\mathbf{x}_2)\delta_g(\mathbf{x}_3) \rangle = b_1^3\langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\delta(\mathbf{x}_3) \rangle + b_1^2b_2 [\langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\delta^2(\mathbf{x}_3) \rangle + 2 \text{ perm} ] + \dots \quad (100)$$

This implies that the galaxy power spectrum and bispectrum at tree-level are

$$P_g(k) = b_1^2P(k), \quad (101)$$

$$B_g(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = b_1^3B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + b_1^2b_2[P(k_1)P(k_2) + P(k_1)P(k_3) + P(k_2)P(k_3)] \quad (102)$$

with  $P$  and  $B$  the matter power spectrum and bispectrum. In the last expression we did not include the connected trispectrum. The reduced galaxy bispectrum is defined as

$$Q_g(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \frac{B_g(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)}{P_g(k_1)P_g(k_2) + P_g(k_1)P_g(k_3) + P_g(k_2)P_g(k_3)} = \frac{1}{b_1}Q(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) + \frac{b_2}{b_1^2}, \quad (103)$$

with  $Q$  the reduced matter bispectrum possibly containing contributions coming from primordial non-Gaussianities. By measuring the reduced galaxy bispectrum in different triangle configurations we should be able to disentangle the biasing, gravitational instability and primordial source of the non-Gaussianity. Current large-scale structure galaxy catalogue do not provide enough data to accurately measure the galaxy bispectrum. The scientific community has for this reason mainly concentrated on the two-point function. The next generation of high redshift surveys will probe tens of thousands of galaxies and probably might enable us to estimate the three-point function. This would be a more direct probe to primordial non-Gaussianities.

### 4.3.2 Skewness and kurtosis

Early large-scale structure surveys were unable to probe the bispectrum and trispectrum. The early literature for this reason concentrates on the moments of the galaxy distribution. The  $p^{\text{th}}$  order moment is defined in terms of the density field of the galaxy distribution  $\delta_g$  as

$$s_p = \frac{\langle \delta_g^p(\mathbf{x}) \rangle_c}{\langle \delta_g^2(\mathbf{x}) \rangle_c^{p/2}}, \quad (104)$$

where  $\langle \dots \rangle_c$  means the connected part of the correlation function [5]. The third-order and fourth-order moments  $s_3$  and  $s_4$  are known as the skewness and kurtosis. The skewness is given in terms of the bispectrum as

$$\langle \delta_g^3(\mathbf{k}) \rangle_c = \int \frac{d^3k_1}{(2\pi)^3} \frac{d^3k_2}{(2\pi)^3} \frac{d^3k_3}{(2\pi)^3} B_g(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3). \quad (105)$$

For Gaussian initial conditions, Peebles [34] showed that to leading order

$$s_3 = \frac{34}{7}\sigma_g, \quad (106)$$

with  $\sigma_g^2 = \langle \delta_g^2 \rangle$  in linear perturbation theory. A non-Gaussian component in the initial conditions and non-linear biasing will lead to additional components in the skewness. Fry and Gaztanaga [16] showed that, to leading order,

$$s_3 = s_3^{(0)} + \frac{34}{7}\sigma_g + \frac{6b_2}{b_1}\sigma_g, \quad (107)$$

where  $s_3^{(0)}$  is the skewness in the initial conditions and  $b_1$  and  $b_2$  are the bias parameters. This equation can be derived from substituting equation (102) and (101) in equation (105) and equation (104). Note that in this formula the first term corresponds to primordial non-Gaussianities, the second to gravitational instability, and the third term to non-linear halo biasing. For higher order corrections the different sources for non-Gaussianities in the galaxy modes mix and the effect cannot be written as a sum of the different effects. Higher order moments such as the kurtosis can be approximated in a similar fashion from standard perturbation theory. However, note that these results are derived under the assumption of the local biasing model.

It is in general difficult to measure primordial non-Gaussianities from the skewness and kurtosis of the LSS galaxy distribution, as they are two numbers resulting from complex gravitational collapse, non-linear biasing and can differ for different types of primordial non-Gaussianities. Bispectrum analyses contain more information, but require a very large galaxy catalogue. Such a catalogue might be achievable with the next generation of large-scale structure surveys.

### 4.3.3 Bias scaling

Galaxy formation is an extremely complex non-linear process which is still not fully understood. Under which circumstances in the matter density field do stars form? The local biasing model is a natural first approximation. There is no fundamental reason why the galaxy perturbation should only depend on the local value of the density perturbation. Their relation could be far more complex. To leading order in the local biasing model, the galaxy or halo power spectrum is given by

$$P_g(k) = b_1^2 P_{(11)}(k) \quad (108)$$

which does not depend on primordial non-Gaussianities. However, simulations by Dalal *et al.* [13] and later confirmed by Desjacques *et al.* [14], and Grossi *et al.* [19] show a scale-dependent correction to the linear biasing model (see figure 18). The correction is often written as

$$P_g(k) = [b_1 + \Delta b_1(k)]^2 P_{(11)}(k), \quad (109)$$

with the correction for local type primordial non-Gaussianities

$$\Delta b_1(k) = 3f_{NL}(b_1 - 1)\delta_c \frac{\Omega_m H_0^2}{k^2 T(k) D(z)}, \quad (110)$$

where  $\delta_c \approx 1.68$  is the linear extrapolated critical density for spherical collapse. We observe that  $\Delta b_1$  increases with scale and decreases with time as the growth factor  $D(z)$  increases. The correction is proportional to the amplitude of the local type primordial non-Gaussianities and vanishes for unbiased distributions of galaxies.

There currently exist several arguments supporting the scaling of the bias for primordial non-Gaussian fluctuations. Dalal *et al.* [13] originally argued that given the local model

$$\Phi(\mathbf{x}) = \Phi_G(\mathbf{x}) + f_{NL} [\Phi_G^2(\mathbf{x}) - \langle \Phi_G^2(\mathbf{x}) \rangle] \quad (111)$$

the Laplacian of the gravitational potential is

$$\nabla^2 \Phi = \nabla^2 \Phi_G + 2f_{NL} [\Phi_G \nabla^2 \Phi_G + |\nabla \Phi_G|^2]. \quad (112)$$

The density field is proportional to this expression. For peaks in the density field, which would lead to halos and galaxy formation, the term  $-\nabla^2 \Phi_G$  dominates over  $|\nabla \Phi|^2$ . Neglecting the  $|\nabla \Phi|^2$  term we get



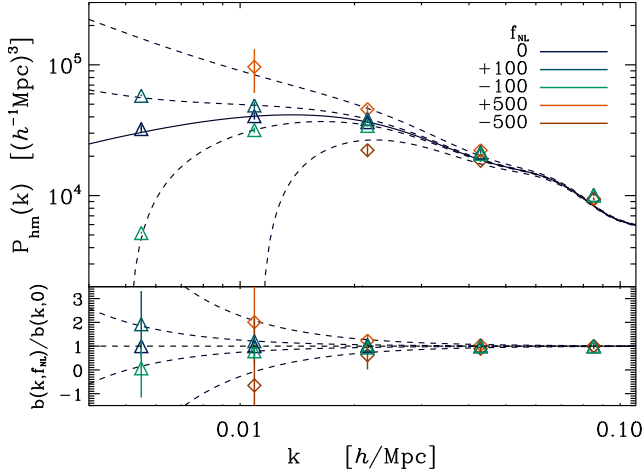


Figure 18: Matter-halo power spectrum in simulations with primordial non-Gaussianities of the local type for different amplitudes  $f_{NL}$ , by Dalal *et al.* [13].

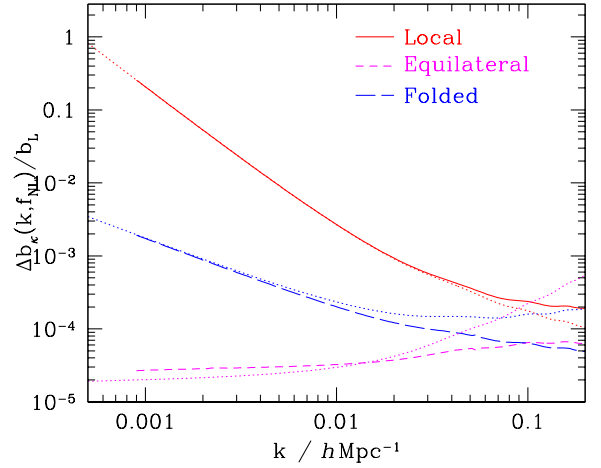


Figure 19: Bias scaling for different type of primordial non-Gaussianities, by Desjacques and Seljak [14].

via the Poisson equation that  $\delta \approx \delta_G [1 + 2f_{NL}\Phi_G]$  with  $\delta_G$  the density field corresponding to the Gaussian potential  $\Phi_G$ . During the evolution of the field,  $\delta$  and  $\delta_G$  grow like  $D(a)$  while  $\Phi_G$  decays like  $a$ , hence

$$\delta \approx \delta_G \left[ 1 + 2f_{NL}\Phi_G \frac{a}{D(a)} \right]. \quad (113)$$

The enhanced or suppressed galaxy formation is due to the correlation of the Gaussian density  $\delta_G$  and potential  $\Phi_G$  which are related via the Poisson equation. For a discussion of the correlation between these fields see Rossi [41]. Now consider a peak of high  $\delta_{pk}$  in the matter density field. In the absence of primordial non-Gaussianities the galaxy density is  $b_L\delta_{pk}$  with  $b_L$  the Lagrangian bias. For nonzero  $f_{NL}$  we have an extra contribution  $2f_{NL}\Phi_G\delta_{pk}a/D(a)$ . Spherical collapse theory predicts the collapse of peaks into virialized objects when the linearly extrapolated density exceeds  $\delta_c \approx 1.68$ . Hence for peaks with the critical density fluctuation  $\delta_{pk} \approx \delta_c$  we get

$$\delta_g = b_L \left[ \delta + 2f_{NL}\Phi_G\delta_c \frac{a}{D(a)} \right]. \quad (114)$$

By switching to the Eulerian bias  $b = 1 + b_L$ , and writing  $\Phi_G$  in terms of  $\delta_G$ , we obtain the relation (110). This calculation should be seen as a crude estimate establishing the scaling of the biasing relation for local type primordial non-Gaussianities. Slosar *et al.* [45] extended the analysis to more general settings using a peak-background split.

All these methods rely on the transformation rule for local type non-Gaussianities and are difficult to translate to generic non-Gaussianities. Other types of non-Gaussianities can however also lead to similar scaling laws. See figure 19 for estimations using  $N$ -body simulations. Matarrese and Verde [32] and Taruya *et al.* [46] showed that for local non-Gaussianities the 1-loop correction to the power spectrum leads to the same scaling relations for large scales. This method can be extended to generic type non-Gaussianities.

## 5 Current constraints and future prospects

In the previous section we discussed the effect of primordial non-Gaussianities on the halo mass function, void abundance and galaxy correlation functions. We estimated the behaviour of these observables and compared them with  $N$ -body simulations. In order to measure or constrain primordial non-Gaussianities we need to compare the estimates with large-scale structure surveys. The detection of primordial non-Gaussianities in the large-scale structure is a rapidly evolving area. So far, several groups have been able to constrain the non-Gaussianities by measuring the bias scaling in the LSS galaxy power spectrum, including [17, 18, 40, 45]. Other LSS observables are still not robust enough to put reliable bounds on primordial non-Gaussianities. The halo mass function is for example still too much affected by systematics and selection effects. We here briefly discuss the results of Giannantonio *et al.* [18].

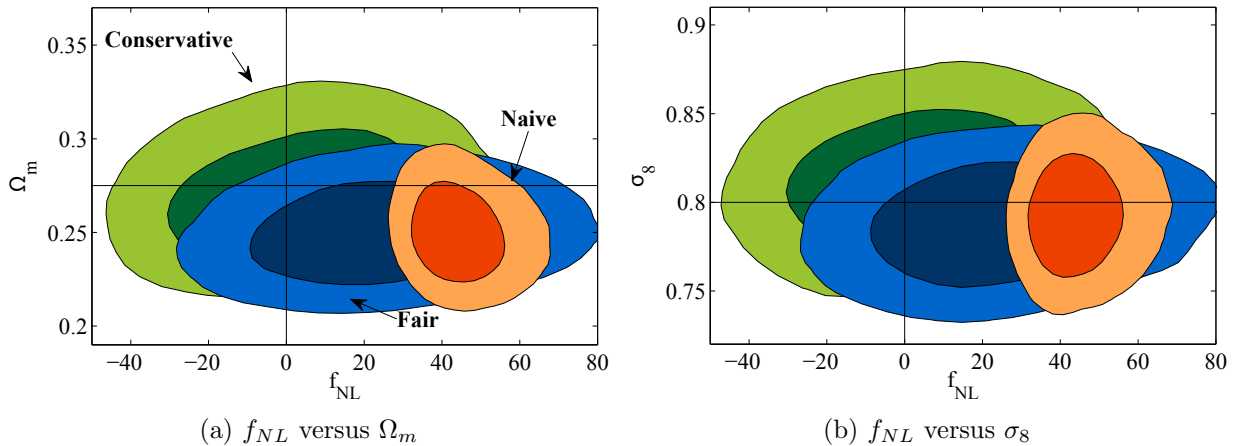


Figure 20: Marginalized probability distributions for  $f_{NL}$  versus  $\Omega_m$  and  $\sigma_8$ , with contours at 68% and 95% confidence levels, by Giannantonio *et al.* [18].

Giannantonio *et al.* [18] estimate  $f_{NL}$  for local type primordial non-Gaussianities by cross-correlating several galaxy surveys amongst each other and with the WMAP cosmic microwave background map. In their analysis they included for the CMB map the integrated Sachs-Wolf effect caused by the evolution of the gravitational potential which generates red-shifts corrections for CMB photons. They in particular correlated the Two Micron All-Sky Survey (2MASS), the galaxy and quasar catalogue of the Sloan Digital Sky Survey (SDSS), the Baryon Oscillation Spectroscopic Survey (BOSS) CMASS sample, the High Energy Astronomy Observatory Program of the X-ray background (HEAO), and the NRAO VLA Sky Survey radio-galaxy catalogue (NVSS). See figure 21 for the different power spectra of the cross-correlations.

In general, large-scale structure surveys suffer from several kinds of contamination by systematic distortions which may significantly bias the primordial non-Gaussianity results. These systematics can be caused by for example observational issues originating in the galaxy such as extinction by dust, or effects caused by our atmosphere such as seeing and airmass. In order to take care of these contaminations, Giannantonio *et al.* [18] divided the data in three subsets. A ‘naïve’ data set containing all the 26 correlations, a ‘fair’ sample containing 25 correlations excluding the auto-correlations of the NVSS and QSO data sets which are supposed to be less reliable, and a ‘conservative’ set containing only the most reliable correlations. By means of a Markov Chain Monte Carlo algorithm, the cosmic parameters are estimated (see figure 20). Note that the ‘naïve’ sample leads to a detection  $31 < f_{NL} < 64$ , while the ‘fair’ and ‘conservative’ subset lead to constraints  $-15 < f_{NL} < 68$ , and  $-36 < f_{NL} < 45$  respectively at a 95% confidence level. We observe that the result strongly depends on the chosen data set. It is for this reason fair to say that the results probably do not imply a detection. However, they do bound the amplitude of local type non-Gaussianities and do so in accordance with the Planck bound on local type primordial non-Gaussianities.

Current bounds on primordial non-Gaussianities by the CMB observations of the Planck consortium are still more stringent than current large-scale structure bounds. However, while current CMB surveys have nearly reached the bound set by cosmic variance, probing primordial non-Gaussianities with large-scale structure surveys is still a relatively new method with a lot of potential. The next generation of large-scale structure missions such as the Dark Energy Survey, Supernova/Acceleration Probe and Large Synoptic Survey Telescope will detect and study tens of thousands of clusters. This will greatly improve the accuracy of the power spectrum analyses of the galaxy distribution. It will moreover enable us to perform bispectrum analyses of the galaxy distribution. For an estimate of the constraining power of the next generation of large-scale structure surveys see [17]. There are also several theoretical developments which would help improve current bounds. In the first place, a better theoretical understanding of the halo mass function and a better definition of voids might make halo and void abundances a competitive source of data to probe primordial non-Gaussianities. Secondly, better halo and void identification algorithms would allow for more reliable comparisons of theoretical studies and large  $N$ -body simulations. Finally, note that the best estimate will likely come from a shared analysis using a range of observables sensitive

to primordial non-Gaussianities. Methods to accurately weight multiple analyses should for this reason be further developed.

## References

- [1] R. J. Adler and J. E. Taylor. *Random Fields and Geometry*. John Wiley & Sons, 2007.
- [2] D. Babich, P. Creminelli, and M. Zaldarriaga. The shape of non-Gaussianities. *J. Cosmology Astropart. Phys.*, 8:9, August 2004.
- [3] D. Baumann. *Part III Applied Mathematics: Cosmology*.
- [4] D. Baumann, A. Nicolis, L. Senatore, and M. Zaldarriaga. Cosmological non-linearities as an effective fluid. *J. Cosmology Astropart. Phys.*, 7:51, July 2012.
- [5] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro. Large-scale structure of the Universe and cosmological perturbation theory. *Phys. Rep.*, 367:1–248, September 2002.
- [6] J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser. Excursion set mass functions for hierarchical Gaussian fluctuations. *ApJ*, 379:440–460, October 1991.
- [7] J. R. Bond, L. Kofman, and D. Pogosyan. How filaments of galaxies are woven into the cosmic web. *Nature*, 380:603–606, April 1996.

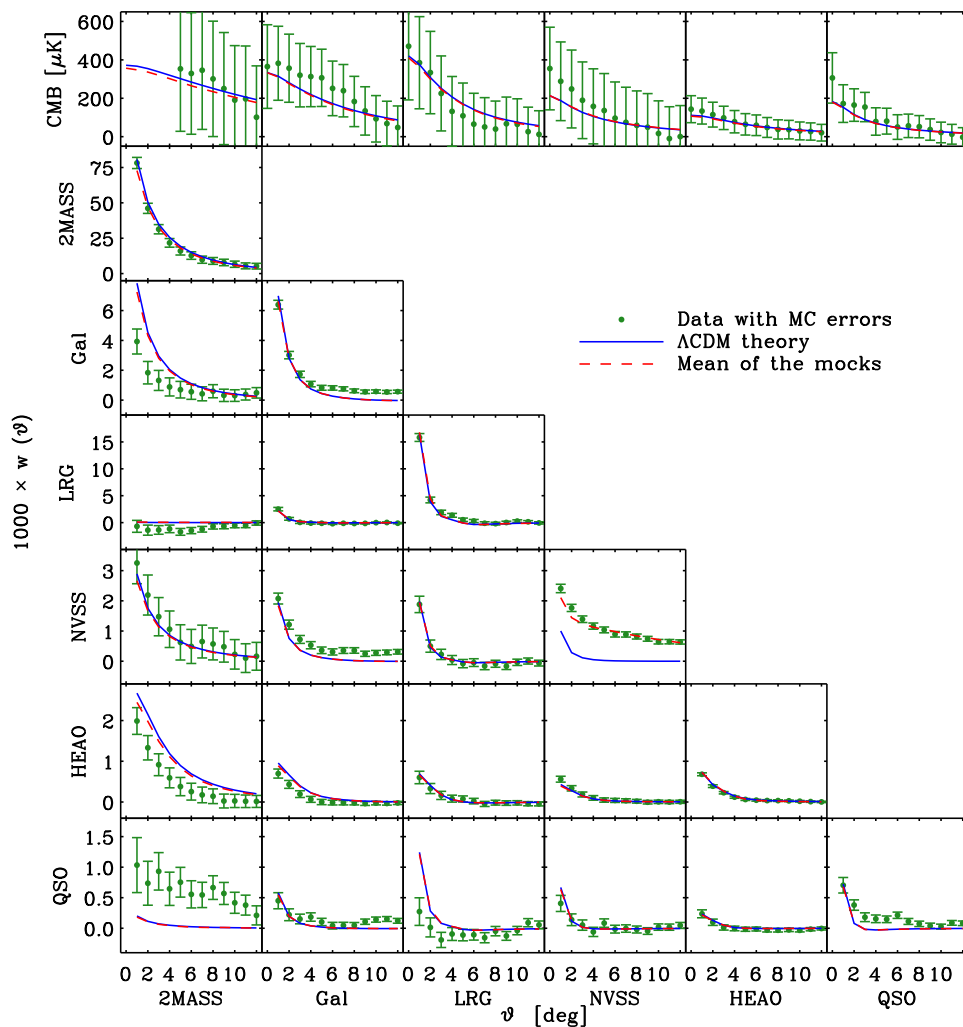


Figure 21: Cross-correlations of several large-scale structure data sets by Giannantonio *et al.* [18].

- [8] J. J. M. Carrasco, M. P. Hertzberg, and L. Senatore. The effective field theory of cosmological large scale structures. *Journal of High Energy Physics*, 9:82, September 2012.
- [9] S. M. Carroll, S. Leichenauer, and J. Pollack. Consistent effective theory of long-wavelength cosmological perturbations. *Phys. Rev. D*, 90(2):023518, July 2014.
- [10] M. Cautun, R. van de Weygaert, and B. J. T. Jones. NEXUS: tracing the cosmic web connection. *MNRAS*, 429:1286–1308, February 2013.
- [11] X. Chen. Primordial Non-Gaussianities from Inflation Models. *Advances in Astronomy*, 2010:72, 2010.
- [12] J. M. Colberg, F. Pearce, C. Foster, E. Platen, R. Brunino, M. Neyrinck, S. Basilakos, A. Fairall, H. Feldman, S. Gottlöber, O. Hahn, F. Hoyle, V. Müller, L. Nelson, M. Plionis, C. Porciani, S. Shandarin, M. S. Vogeley, and R. van de Weygaert. The Aspen-Amsterdam void finder comparison project. *MNRAS*, 387:933–944, June 2008.
- [13] N. Dalal, O. Doré, D. Huterer, and A. Shirokov. Imprints of primordial non-Gaussianities on large-scale structure: Scale-dependent bias and abundance of virialized objects. *Phys. Rev. D*, 77(12):123514, June 2008.
- [14] V. Desjacques and U. Seljak. Primordial non-Gaussianity from the large-scale structure. *Classical and Quantum Gravity*, 27(12):124011, June 2010.
- [15] A. G. Doroshkevich. Spatial structure of perturbations and origin of galactic rotation in fluctuation theory. *Astrophysics*, 6:320–330, October 1970.
- [16] J. N. Fry and E. Gaztanaga. Biasing and hierarchical statistics in large-scale structure. *ApJ*, 413:447–452, August 1993.
- [17] T. Giannantonio, C. Porciani, J. Carron, A. Amara, and A. Pillepich. Constraining primordial non-Gaussianity with future galaxy surveys. *MNRAS*, 422:2854–2877, June 2012.
- [18] T. Giannantonio, A. J. Ross, W. J. Percival, R. Crittenden, D. Bacher, M. Kilbinger, R. Nichol, and J. Weller. Improved primordial non-Gaussianity constraints from measurements of galaxy clustering and the integrated Sachs-Wolfe effect. *Phys. Rev. D*, 89(2):023511, January 2014.
- [19] M. Grossi, L. Verde, C. Carbone, K. Dolag, E. Branchini, F. Iannuzzi, S. Matarrese, and L. Moscardini. Large-scale non-Gaussian mass function and halo bias: tests on N-body simulations. *MNRAS*, 398:321–332, September 2009.
- [20] Johan Hidding. PhD thesis.
- [21] R. Holman and A. J. Tolley. Enhanced non-Gaussianity from excited initial states. *J. Cosmology Astropart. Phys.*, 5:1, May 2008.
- [22] A. Jenkins, C. S. Frenk, S. D. M. White, J. M. Colberg, S. Cole, A. E. Evrard, H. M. P. Couchman, and N. Yoshida. The mass function of dark matter haloes. *MNRAS*, 321:372–384, February 2001.
- [23] N. Kaiser A.S. Szalay J.M. Bardeen, J.R. Bond. The statistics of peaks of gaussian random fields. *Astrophysical journal part 1*, 304.
- [24] T. Kidani and K. Koyama. Non-Gaussianities in DBI inflation with angular motion. *Phys. Rev. D*, 90(2):023515, July 2014.
- [25] C. Kiefer, D. Polarski, and A. A. Starobinsky. Quantum-To Transition for Fluctuations in the Early Universe. *International Journal of Modern Physics D*, 7:455–462, 1998.
- [26] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. R.olta, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *ApJS*, 192:18, February 2011.

- [27] E. Komatsu and D. N. Spergel. Acoustic signatures in the primary microwave background bispectrum. *Phys. Rev. D*, 63(6):063002, March 2001.
- [28] M. Liguori, E. Sefusatti, J. R. Fergusson, and E. P. S. Shellard. Primordial Non-Gaussianity and Bispectrum Measurements in the Cosmic Microwave Background and Large-Scale Structure. *Advances in Astronomy*, 2010:73, 2010.
- [29] B. Little, D. H. Weinberg, and C. Park. Primordial fluctuations and non-linear structure. *MNRAS*, 253:295–306, November 1991.
- [30] M.S. Longuet-Higgins. Mathematical analysis of random noise. *Philosophical transactions of the royal society london A*, 249:321, 1957.
- [31] J. Maldacena. Non-gaussian features of primordial fluctuations in single field inflationary models. *Journal of High Energy Physics*, 5:13, May 2003.
- [32] S. Matarrese and L. Verde. The Effect of Primordial Non-Gaussianity on Halo Bias. *ApJ*, 677:L77–L80, April 2008.
- [33] S. Matarrese, L. Verde, and R. Jimenez. The Abundance of High-Redshift Objects as a Probe of Non-Gaussian Initial Conditions. *ApJ*, 541:10–24, September 2000.
- [34] P. J. E. Peebles. *The large-scale structure of the universe*. 1980.
- [35] M.E. Peskin and D.V. Schroeder. *An Introduction to Quantum Field Theory*. Advanced book classics. Addison-Wesley Publishing Company, 1995.
- [36] Planck Collaboration, R. Adam, P. A. R. Ade, N. Aghanim, Y. Akrami, M. I. R. Alves, M. Arnaud, F. Arroja, J. Aumont, C. Baccigalupi, and et al. Planck 2015 results. I. Overview of products and scientific results. *ArXiv e-prints*, February 2015.
- [37] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, and et al. Planck 2015 results. XVII. Constraints on primordial non-Gaussianity. *ArXiv e-prints*, February 2015.
- [38] W. H. Press and P. Schechter. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *ApJ*, 187:425–438, February 1974.
- [39] S.O. Rice. Mathematical analysis of random noise. *Bell system technology journal*, 23.
- [40] A. J. Ross, W. J. Percival, A. Carnero, G.-b. Zhao, M. Manera, A. Raccanelli, E. Aubourg, D. Bizyaev, H. Brewington, J. Brinkmann, J. R. Brownstein, A. J. Cuesta, L. A. N. da Costa, D. J. Eisenstein, G. Ebelke, H. Guo, J.-C. Hamilton, M. V. Magaña, E. Malanushenko, V. Malanushenko, C. Maraston, F. Montesano, R. C. Nichol, D. Oravetz, K. Pan, F. Prada, A. G. Sánchez, L. Samushia, D. J. Schlegel, D. P. Schneider, H.-J. Seo, A. Sheldon, A. Simmons, S. Snedden, M. E. C. Swanson, D. Thomas, J. L. Tinker, R. Tojeiro, and I. Zehavi. The clustering of galaxies in the SDSS-III DR9 Baryon Oscillation Spectroscopic Survey: constraints on primordial non-Gaussianity. *MNRAS*, 428:1116–1127, January 2013.
- [41] G. Rossi. On the initial shear field of the cosmic web. *MNRAS*, 421:296–307, March 2012.
- [42] R. Scoccimarro, E. Sefusatti, and M. Zaldarriaga. Probing primordial non-Gaussianity with large-scale structure. *Phys. Rev. D*, 69(10):103513, May 2004.
- [43] P. Shallard and A. Challinor. *Part III Applied Mathematics: Advanced Cosmology*.
- [44] R. K. Sheth, H. J. Mo, and G. Tormen. Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes. *MNRAS*, 323:1–12, May 2001.
- [45] A. Slosar, C. Hirata, U. Seljak, S. Ho, and N. Padmanabhan. Constraints on local primordial non-Gaussianity from large scale structure. *J. Cosmology Astropart. Phys.*, 8:31, August 2008.

- [46] A. Taruya, K. Koyama, and T. Matsubara. Signature of primordial non-Gaussianity on the matter power spectrum. *Phys. Rev. D*, 78(12):123534, December 2008.
- [47] M. J. Way, P. R. Gazis, and J. D. Scargle. Structure in the 3D Galaxy Distribution. II. Voids and Watersheds of Local Maxima and Minima. *ApJ*, 799:95, January 2015.
- [48] R. V. D. Weygaert and J. R. Bond. Clusters and the Theory of the Cosmic Web. In M. Plionis, D. Hughes, and O. López-Cruz, editors, *A Pan-Chromatic View of Clusters of Galaxies and the Large-Scale Structure*, volume 740 of *Lecture Notes in Physics*, Berlin Springer Verlag, page 335, 2008.
- [49] I. B. Zel'dovich and S. F. Shandarin. *Large-scale distribution of matter in the universe*, pages 338–352. 1989.