EFFECTS OF PRIMORDIAL NON-GAUSSIANITY IN THE CMB AND
THE LARGE-SCALE STRUCTURE

A Dissertation in
Physics
by
Saroj Adhikari

© 2016 Saroj Adhikari

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2016
The dissertation of Saroj Adhikari was reviewed and approved* by the following:

Sarah Shandera  
Assistant Professor of Physics  
Dissertation Adviser  
Chair of Committee

Irina Mocioiu  
Associate Professor of Physics

Yuexing Li  
Assistant Professor of Astronomy

Donghui Jeong  
Assistant Professor of Astronomy

Richard Robinett  
Professor of Physics  
Director of Graduate Studies

*Signatures are on file in the Graduate School.
Abstract

We study the imprints and effects of non-Gaussian primordial perturbations in the cosmic microwave background (CMB) and the large-scale structure (LSS) of the Universe.

Primordial non-Gaussianity, if detected, has the potential to discriminate between different classes of inflationary models. For example, inflationary models with just a single degree of freedom can be ruled out if long wavelength modes are found to be correlated with shorter wavelength modes.

The effects of primordial non-Gaussianity can be seen in the CMB and the LSS—the two important observational probes of the primordial statistics and therefore the early universe. This thesis consists of studies in both of these observational probes. First, in the CMB front, we study the effect of primordial non-Gaussianity on the statistical anisotropy of the CMB temperature fluctuations. This is motivated by recent CMB results that indicate such statistical anisotropies at large CMB scales. We show that the probability of such observations increase in the case that primordial fluctuations are non-Gaussian. For this, we derive and use a general framework to describe statistical anisotropies in the power spectrum in the presence of primordial non-Gaussianity. We expect the formalism to be of use beyond the specific application that we have done.

Second, in the large-scale structure front, we study the effect of primordial non-Gaussianity generated from two fields on the mass function of massive cluster of galaxies, scale-dependent halo bias, and large-scale stochasticity. We derive physically motivated analytic expressions for these observables and calibrate the derived expressions on dedicated N-body simulation results. Thus, the final outputs of the investigation are simulation calibrated semi-analytic formulas that can be used to constrain primordial non-Gaussianity beyond the simple single-field local-type that has been mostly studied up until this point. We also discuss the “position-dependent bispectrum” as an effective method of measuring the squeezed limit of a primordial trispectrum using large-scale structure galaxy surveys.
# Table of Contents

List of Figures vii

List of Tables x

Acknowledgments xi

## Chapter 1
Background and Introduction 1

1.1 Standard model of cosmology: the ΛCDM Universe 2

1.2 Cosmic inflation 4

1.2.1 What is cosmic inflation? 5

1.2.2 Generation of primordial perturbations 6

1.3 What we observe: statistics in the CMB and the LSS 8

1.3.1 Cosmic microwave background (CMB) 8

1.3.2 Large-scale structure (LSS) 11

## Chapter 2
Introduction to Non-Gaussianity 15

2.1 Non-Gaussian statistics 15

2.2 Non-Gaussian primordial fluctuations 18

2.3 The hunt for primordial non-Gaussianity 18

## Chapter 3
Effects of Primordial Non-Gaussianity: I. CMB 20

3.1 The CMB power asymmetry 20

3.1.1 Introduction 20

3.1.2 The local-variance method 20

3.1.3 Simulations and data 23

3.1.4 Results 25

3.1.5 Discussion and Summary 33

3.2 Non-Gaussianity and the power asymmetry 34

3.2.1 Introduction 34

3.2.2 Illustrating the connection between non-Gaussianity and isotropy 37
Appendix B

Calculations for LSS clustering statistics .............................. 124
B.1 Integrals for $\langle \delta_R^n \rangle_c$ ........................................ 124
B.2 Truncation and error ......................................................... 128
B.3 Calculations for bias and stochasticity ................................. 130
B.4 Galaxy trispectrum expressions ........................................... 134

Bibliography ........................................................................... 136
# List of Figures

1.1 The monopole (top), dipole (middle) and the temperature fluctuations after monopole and dipole subtraction (bottom) of the CMB measured by the COBE satellite. ........................................ 2

1.2 The theoretical angular power spectrum $C_\ell$ from the best fit cosmological parameters obtained from the Planck 2015 mission. ............. 10

1.3 Matter density field $\delta(x)$ and halos .................................................. 13

2.1 A squeezed triangular configuration of bispectrum. ............... 16

2.2 A squeezed trispectrum configuration. .......................... 17

3.1 An illustration of power asymmetry in simulated CMB maps ........... 21

3.2 The hemispherical asymmetry in $C_\ell$s generated by our modulation model. 24

3.3 Foreground cleaned maps at channels 143 GHz and 217 GHz. ........... 26

3.4 Local variance dipole analysis for $r_{\text{disk}} = 1,4$ degrees. ............... 27

3.5 Local variance dipole analysis for $r_{\text{disk}} = 0.25,0.18$ degrees. ............... 28

3.6 The component of local variance dipole amplitudes in the direction $(l,b) = (264^\circ,48^\circ)$ using disk of radius $r_{\text{disk}} = 0.18^\circ,8^\circ$ after removing large scales features up to $l_{\text{min}} = 600$. .................................................. 29

3.7 Here we plot some of the directions for the local variance dipoles in our analysis in addition to the CMB dipole direction (black square) and the large scale power asymmetry direction (black cross). ......................... 30

3.8 Small-scale power asymmetry in the direction $(l,b) = (218^\circ,-20^\circ)$ before subtracting the contribution from the Doppler dipole. ......................... 32

3.9 Small-scale power asymmetry in the direction $(l,b) = (218^\circ,-20^\circ)$ after subtracting the Doppler contribution. ......................... 32

3.10 The expected higher order modulations for $f_{\text{NL}} = 500,250$ (red square, open green square). .................................................. 43

3.11 Test of the monopole modulation formula Eq.(3.42) for the local non-Gaussian model with $f_{\text{NL}}$ specified in the figure. ......................... 49

3.12 The distribution of power asymmetry $A_i$ (in a particular direction $d_i$) and the amplitude of power asymmetry $A$ measured in 10000 simulated CMB skies. .................................................. 50
3.13 The posterior probability distribution of $|f_{\text{NL}}|$ values for different observed amplitudes $A$ of power asymmetry.  
3.14 The posterior probability distribution of $|f_{\text{NL}}|$ after combining the bispectrum constraints at large scales ($\ell \lesssim 100$, $f_{\text{NL}} = -100 \pm 100$) with the power asymmetry constraint for the given value of $A$.  
3.15 The $p$-value for different values of asymmetry amplitudes $A$ (i.e. the probability of obtaining an asymmetry amplitude equal to or greater than $A$) in a local non-Gaussian model as a function of the value of $|f_{\text{NL}}|$.  
3.16 The expected dipolar modulation of the power spectrum for large amplitude local-, orthogonal- and equilateral-type non-Gaussianities.  
3.17 The expected amplitude of the power asymmetry from superhorizon modes ($k_\ell < \pi/r_{\text{cmb}}$), as a function of the minimum wavenumber $k_{\ell,\text{min}}$ considered to compute $\langle g_{1M}^2 \rangle^{0.5}$, for large-amplitude local-, orthogonal- and equilateral-type non-Gaussianities.  
3.18 The effect of a simple scale-dependent local non-Gaussian model, $f_{\text{NL}}(\ell) = 50(\ell/60)^{-0.64}$, in the bispectrum (top), the power asymmetry amplitude (middle), and the modulation of power spectrum amplitude (bottom).  
3.19 The posterior distribution of $|f_{\text{NL}}|$ for $A_{\text{obs}} = 0.055$ and different assumed values of $A_{0,\text{obs}}$ and $N_{\text{extra}}$.  
3.20 The Bayesian evidence as a function of $N_{\text{extra}}$.  
4.1 The simulation results and semi-analytic predictions for the hierarchical and feeder simulations.  
4.2 The simulation results and the mass function predictions for F122 model, and the fractional error in the semi-analytic predictions for the F215 and F122 models compared to the simulation results.  
4.3 The simulation results and semi-analytic mass function prediction for: (i) left: a feeder simulation with $M_3 \approx 0.198$ ($f_2 = 1.16$), and (ii) right: a mixed scaling simulation with $M_3 \approx 0.112$ ($f_2 = 1.06$).  
4.4 $M_3$ dependence of $f_2$ defined in Eq.(4.31) and Eq.(4.32) for feeder and hierarchical mass functions respectively.  
4.5 Results from our simulations for both hierarchical and feeder scalings as density plots.  
4.6 The relative difference between the semi-analytic predictions (Eq.(4.32 or Eq.(4.31)) and the simulation results as density plots.  
4.7 Constraint on hierarchical-type and feeder-type non-Gaussianity (parameterized by $M_3$) models from X-ray cluster data.  
4.8 The bias (left) and stochasticity (right) at large scales for the Gaussian simulations at $z = 0$ and $z = 1$ using halos in the mass range $4.83 \times 10^{13} h^{-1} M_\odot \leq M \leq 9.55 \times 10^{13} h^{-1} M_\odot$.  
4.9 The simulation results for the bias at large scales and the large-scale stochasticity using $H500$ and $H99$ models.
4.10 The bias and stochasticity at large scales for the feeder models F677 and F215 at $z = 0, 1, 2$ using halos in the mass range $4.83 \times 10^{13} h^{-1} M_{\odot} \leq M \leq 9.55 \times 10^{13} h^{-1} M_{\odot}$.

4.11 The bias $P_{hm}/P_{mm}$ (left) and stochasticity $r^2(k)$ (right) at large scales for the models F70 and M997 at $z = 0, 1, 2$ using halos in the mass range $4.83 \times 10^{13} h^{-1} M_{\odot} \leq M \leq 9.55 \times 10^{13} h^{-1} M_{\odot}$.

4.12 The various reduced integrated trispectra, $u^{(i)}_R$ (the expressions are given in Eq.(4.70) and Eq.(4.72)) for a large spherical sub-volume with radius $R = 400 \text{Mpc}/h$ at $z = 1.0$. We have taken $b_1 = 1.95, b_2 = 0.5, b_3 = 0.1$.

4.13 The galaxy number density as a function of the redshift assumed in the Fisher matrix calculations.

4.14 The smoothed two-point correlation function $\xi_R(r)$ as a function of the comoving distance $r$ (normalized by $\xi_R(0)$, for three smoothing scales $R = 200, 100, 20 \text{Mpc}/h$).

4.15 Fisher forecast ellipses for two of $(f_{NL}^{\text{local}}, b_1, b_2)$ marginalized over the other, assuming SPHEREx survey volume and other parameters given above in the figure.

4.16 Fisher forecast ellipses for two of $(A_{PD}, b_1, b_2)$ marginalized over the other and $b_3$ (which is not shown), assuming SPHEREx survey volume.

A.1 Example of a realization of a primordial potential $\Phi$ (Gaussian).

A.2 Some useful correlations among quantities in the Gaussian and non-Gaussian numerical maps.

B.1 The cumulant estimates for the various $I_{ij}(M)$ intgral from the Monte-Carlo approach.

B.2 Left: we plot the maximum value of $M_3$ for which the PDF (for various truncations $N$) produces results within 20% error for $\nu$ specified on the x-axis to $\nu_{\text{max}} = 2.1^{0.7}$ for feeder scaling of higher moments. Right: same as left but for hierarchical scaling and with $\nu_{\text{max}} = 2.2^{0.7}$.

B.3 Left: The error of $N = 5$ truncation for different values of $M_3$ for feeder scaling. Right: same as left but for hierarchical scaling.

B.4 Numerical test of the approximation (B.10) for $F_{R,2}^{(3)}$ at $R = 8 \text{Mpc}/h$. 

ix
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Summary of the Doppler dipole detection results.</td>
<td>31</td>
</tr>
<tr>
<td>3.2</td>
<td>Savage-Dickey Density Ratio (SDDR) for different observed values of</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>dipole power modulations at large scales.</td>
<td></td>
</tr>
<tr>
<td>3.3</td>
<td>SDDR for $A = 0.055$ and different observed values of monopole power</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>modulations $A_0$ at large scales.</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>Parameter space of our simulations.</td>
<td>75</td>
</tr>
<tr>
<td>4.2</td>
<td>Large-scale bias and stochasticity for simulations with Gaussian</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>initial conditions.</td>
<td></td>
</tr>
<tr>
<td>4.3</td>
<td>Values of the bias coefficients $b_\phi$ and $b_\psi$ measured by</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>cross correlating the halo density field with the corresponding</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\phi$ and $\psi$ components in the linear density field.</td>
<td></td>
</tr>
<tr>
<td>4.4</td>
<td>Fisher forecast results for $f_{NL}$.</td>
<td>107</td>
</tr>
<tr>
<td>4.5</td>
<td>Fisher forecast results for $A_{PD}$.</td>
<td>109</td>
</tr>
</tbody>
</table>
Acknowledgments

I would like to thank my adviser Sarah Shandera for guiding me throughout the works done in this thesis. Further, I am indebted to her for encouraging me to explore research topics and allowing me to work on topics that interest me the most. I would like to express my sincere gratitude to the members of my dissertation committee for their support and review of my thesis. I would also like to express my gratitude to my research collaborators: Sarah Shandera, Neal Dalal (UIUC), Adrienne Erickcek (UNC), Donghui Jeong and Adam Mantz (Stanford).

I would like to take this opportunity to thank my previous research mentors—my undergraduate mentors Balraj Menon (UCA), Stephen Addison (UCA) and Larry Weaver (KSU) with whom I had learned physics beyond the undergraduate textbooks, and Steven Heppelmann with whom I worked (as a graduate student) in 2011 and learned about the basics of high energy data analysis. I then decided to switch to cosmology, and I am thankful to both him and Richard Robinett (who had suggested that I go talk to my current adviser) for useful suggestions during this difficult period of switching fields and advisers.

I am very thankful to the Institute for Gravitation and the Cosmos, and the Department of Physics at Penn State. In particular, I am thankful to all the scientific members and the administrative staffs for their great support during my years at the institute and in the department.

I would like to express my gratitude to all my wonderful friends at the institute. In particular, I had the pleasure of sharing office space in Whitmore with Aruna, Suddho and Elliot for a number of years. I am also beyond thankful to all my non-physics friends! In particular, I must especially thank my friends with whom I have shared apartments in the US over the course of last many years.

During the time when the work for this thesis was being done, I was fortunate to get external and departmental supports at various times. I am thankful to each of these supports: the NASA Astrophysics Theory Grant No. NNX12AC99G, the Frymoyer Honors Fellowship, the Downsbrugh Fellowship, and the Duncan Fellowship.

Some of the numerical work done for the thesis made use of the computing resources provided by Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number OCI-1053575. In addition, I am thankful to have been able to use the computing resources provided by the Institute
for CyberScience at the Pennsylvania State University.

Finally, I am always thankful of my wonderful, loving and supportive parents and sisters.
Chapter 1

Background and Introduction

Modern cosmology rests on many technological developments that have resulted in new ways to observe the cosmos. Among these observational probes, in the recent years, are the cosmic microwave background (CMB) experiments that have provided to us a great deal of knowledge about the early universe. CMB provides us with a snapshot of the density fluctuations from nearly the largest scales that we can directly probe using photons. These observations can be explained in a well-established model of modern cosmology. This is often remarked as the standard model of cosmology. In addition to the CMB, we find that the standard model is remarkably successful in explaining many of the observations of density fluctuations from the structure that formed later than when the snapshot of the CMB was generated—the so-called large-scale structure (LSS).

The standard model is the six parameter Λ CDM model, augmented with an initial condition for the statistics of the primordial fluctuations from the theory of cosmic inflation. Cosmic inflation is an early period of super-luminal exponential expansion that explains a number of peculiarities in the standard big-bang model, in addition to providing a mechanism for the generation of primordial perturbations.

One of the main tools to study the period of cosmic inflation is primordial non-Gaussianity—deviation from Gaussian statistics of the primordial fluctuations. Effects of primordial non-Gaussianity, if large enough, can be observed in the CMB temperature and polarization spectra and in the large-scale structure of the universe. Some of these effects are the topics of discussion in this thesis. CMB and LSS are the two important observational areas that provide signatures of both the early and late-time evolution of the universe.

In this chapter, we will now provide a brief introduction to the Λ CDM model and the theory of cosmic inflation. In the next chapter, we will then discuss how deviations from Gaussian statistics could have been generated in the early universe. This will then
be followed by the main body of the thesis where we will provide detail studies of some of the effects of primordial non-Gaussianity in the CMB and the LSS of the universe.

1.1 Standard model of cosmology: the ΛCDM Universe

The universe at large scales is homogeneous and isotropic. Many observational tests, direct and indirect, have been performed to test the homogeneity and isotropy of the observable universe at large scales—both at the level of the background density and at the level of the density fluctuations.

For example, the 3K microwave background is extremely isotropic [1]. Since the dipole gets most of its contribution from the Doppler effect due to our own motion with respect to the CMB rest frame, the anisotropies appear roughly at the level of $10^{-5}$ [2] (as evident from the ratio of the temperature fluctuations to that of the CMB monopole: $18\mu K/2.728K \approx 7 \times 10^{-6}$ in Figure 1.1). In the recent years, the WMAP [3] and the Planck [4] satellite missions have obtained great measurements of these CMB anisotropies.

Observations also tell us that the observable universe is expanding. The evolution of this expanding, homogeneous and isotropic universe is well-explained by the ΛCDM model: in which the Λ is used to denote a cosmological constant and CDM stands for
cold dark matter; cosmological observations require these components in addition to the better-known ordinary baryonic matter. The late-time energy budget of the Universe is dominated by these two components (with almost 70% of the present energy density dominated by dark energy). The energy densities of different components (dark matter, dark energy, radiation etc) evolve differently in an expanding universe. Therefore, although the present universe is dominated by dark energy, we find that, in the past, there was a matter-dominated era and a radiation-dominated era.

The metric that describes an expanding, homogeneous and isotropic universe is the famous FLRW (Friedmann-Lemaître-Robertson-Walker) metric:

\[ ds^2 = -dt^2 + a^2(t) \left[ \frac{dr^2}{1 - kr^2} + r^2 d\Omega^2 \right] . \]  

(1.1)

where \( k \) is the spatial curvature parameter; if \( k > 0 \) the universe is closed, if \( k < 0 \) the universe is open and if \( k = 0 \) then the universe if flat. Two Friedmann equations (which are derived from Einstein’s field equations) describe the evolution of the scale factor \( a(t) \). These are

\[ H^2 + \frac{k}{a^2} = \frac{8\pi G}{3} \rho \]  

(1.2)

\[ \ddot{a} = -\frac{4\pi G}{3} (\rho + 3p) a \]  

(1.3)

where \( H = \dot{a}/a \) is the Hubble constant, \( \rho \) the total energy density, and \( p \) is the pressure. The first of the above equation is commonly referred to as the Friedmann equation. The various contributions to the total energy densities at various epochs come from: matter \( \rho_m \), radiation \( \rho_r \) and dark energy or cosmological constant \( \rho_\Lambda \). The values of these energy densities can be probed through observations. And, these energy densities evolve differently with time: \( \rho_m = \rho_{m,0} a^{-3} \), \( \rho_r = \rho_{r,0} a^{-4} \), and \( \rho_\Lambda = \rho_{\Lambda,0} \). The critical density is defined as \( \rho_c = 3H^2/(8\pi G) \). We can now define the energy density as a fraction of the critical density:

\[ \Omega = \Omega_m + \Omega_r + \Omega_\Lambda = \Omega_{m,0} a^{-3} + \Omega_{r,0} a^{-4} + \Omega_{\Lambda,0} . \]

and rewrite the Friedmann equation as

\[ 1 = \Omega - \Omega_k \]  

(1.4)

\[ H^2 = H_0^2 \left[ \frac{\Omega_{m,0}}{a^3} + \frac{\Omega_{r,0}}{a^4} + \Omega_{\Lambda,0} - \frac{\Omega_{k,0}}{a^2} \right] \]  

(1.5)

where \( \Omega_k = k/(a^2H^2) \).

The flatness problem: From the above equation, we see that at early times \( (a \ll 1) \), the energy density from the curvature \( k \) can be neglected. In that case, \( \Omega_{\text{early}} \approx 1 \). The
measured value for $\Omega_{\text{today}}$ is also close to unity. With this fact, and the measured values of the fractional energy densities $\Omega_{m,0}, \Omega_{r,0}, \Omega_{\Lambda,0}$, we can then work out the evolution of $\Omega(a)$. The easiest way to fit the two observations, $\Omega_{\text{early}} \approx \Omega_{\text{today}} \approx 1$, is to postulate an exactly flat universe with $k = 0$. Then, the Friedmann universe will remain flat at all times. But, with a lack of a physical mechanism/principle for this, let us assume that $\Omega_{k,0} \approx 10^{-3}$ within the observational bounds. Then, we can use the above equation to calculate the evolution of $\Omega(a)$. Under the FLRW evolution, we find that

$$\Omega(a) - 1 = \frac{\Omega_{k,0}}{a^2}$$

(1.6)

i.e. for $|\Omega_{k,0}| < 10^{-3}$, we require $|\Omega(a) - 1| < 10^{-19}$ for $a = 10^{-8}$ (around the time when big-bang nucleosynthesis (BBN) is occurring). If we go further in the past, then to get a non-zero small value of $|\Omega_{k,0}|$ within the observational bounds requires that $\Omega(a)$ is even closer to unity. In the lack of a physical mechanism, this is a fine-tuning problem called the flatness problem. We should note, however, that there are arguments that the flatness problem is not a fine-tuning problem in the standard big bang cosmology [5].

The horizon problem: The horizon problem is more universally acknowledged as a genuine problem in an expanding universe than the flatness problem [6]. The horizon problem stems from the observation that the CMB is extremely isotropic, and therefore the universe at large scales is extremely homogeneous. However, we also know that the universe is expanding. If we go back in time, we find that horizons of two photons sufficiently separated were never in causal contact. This in itself is not a problem, but then the extreme isotropy of the CMB will be surprising as microphysics could not have occurred between two points not in one others horizon [7]. This is the horizon problem of the standard big bang cosmology.

These puzzles are naturally solved by postulating a period of exponential expansion in the very early universe—the period of cosmic inflation discussed next.

1.2 Cosmic inflation

The theory of cosmic inflation was proposed originally to solve some of the problems in the standard big bang picture—in particular the horizon and the flatness problems [8]. (The original proposal had some flaws and was soon improved upon by [9, 10].) It postulates a rapidly accelerating expansion phase in the very early history of the Universe. The flatness problem is then solved by the fact that any curvature present before the onset of inflation becomes small during the exponential expansion of the universe. This is because the curvature after the exponential expansion is:

$$k_{\text{end}} = \left( \frac{a_{\text{begin}} H_{\text{begin}}}{a_{\text{end}} H_{\text{end}}} \right)^2 k_{\text{begin}}.$$
During inflation, the Hubble rate is approximately constant: $H_{\text{end}} \approx H_{\text{begin}}$, so that the curvature $k_{\text{end}}$ is exponentially suppressed. In his original proposal [8], Guth also discusses the monopole problem and the horizon problem, which are also solved by invoking enough number of efoldings:

$$\text{N} = \ln \left( \frac{a_{\text{end}}}{a_{\text{begin}}} \right)$$

of expansion during inflation. The number of efoldings required to solve the horizon problem, however, is enough to solve the other two [11]. In addition to solving a number of problems in the big bang picture, the theory of inflation also provides a natural mechanism for the creation of the primordial fluctuations. We will discuss this in the next section. First, let us introduce the basics of inflation.

### 1.2.1 What is cosmic inflation?

In the canonical single-field slow roll (SFSR) model of inflation, we postulate the presence of a homogeneous scalar field, $\phi(x)$, in the very early universe. The particle associated with this field is called the inflaton. The simplest model can be described by a scalar field theory with just this scalar field, by the action

$$S_{\phi} = \int d^4 x \sqrt{-g} \left[ -\frac{1}{2} g^{\mu \nu} \partial_\mu \phi \partial_\nu \phi - V(\phi) \right] \tag{1.7}$$

The energy density and pressure are then given by:

$$\rho = \frac{1}{2} \dot{\phi}^2 + V(\phi)$$

$$p = \frac{1}{2} \dot{\phi}^2 - V(\phi) \tag{1.8}$$

The Friedmann equations, then, are

$$\rho_c = \frac{3H^2}{8\pi G} = \frac{1}{2} \dot{\phi}^2 + V(\phi)$$

$$\ddot{\phi} = - \left[ 3H \dot{\phi} + V'(\phi) \right] \tag{1.9}$$

where $H = \frac{\dot{a}}{a}$ is the Hubble parameter. From the second Friedmann equation Eq.(1.3), we can see that $\ddot{a} > 0$ (accelerated expansion) is possible when $\rho + 3p < 0$. One of the ways to obtain the sufficient negative pressure is through a scalar field. To see this, we find that, from Eq.(1.8),

$$\rho + 3p = 2\dot{\phi}^2 - 2V(\phi)$$

and therefore

$$\ddot{a} = \frac{8\pi G}{3} \left[ V(\phi) - \dot{\phi}^2 \right] a \tag{1.10}$$
By simply requiring that the potential dominates over $\dot{\phi}^2$, $\ddot{a} > 0$ can be achieved. Further, if the energy density $\rho \approx V(\phi)$ is approximately constant, a phase of exponential expansion can be achieved:

$$a(t) = a_{\text{begin}} e^{H(t-t_{\text{begin}})} \quad t_{\text{end}} > t > t_{\text{begin}}$$

(1.11)

The requirements for the potential and therefore the energy density to be nearly constant can be quantified in terms of suitable conditions for the potential $V(\phi)$ and its derivatives $V'(\phi), V''(\phi)$. The conditions require the first and second derivatives to be small compared to the value of the potential. These are the slow-roll conditions. To quantify the slow-roll conditions, two slow-roll parameters are defined [12] and required to be small:

$$\epsilon = \frac{1}{16\pi G} \left[ \frac{V'(\phi)}{V(\phi)} \right]^2$$

(1.12)

$$\eta = \frac{1}{8\pi G} \frac{V''(\phi)}{V(\phi)}$$

(1.13)

In addition to solving some of the problems in the standard big bang expansion, inflation also provides a mechanism for the generation of small primordial fluctuations, which are then observed as anisotropies in the CMB.

### 1.2.2 Generation of primordial perturbations

One of the huge successes of the theory of cosmic inflation is its natural amplification of quantum fluctuations in the primordial universe. These primordial fluctuations then act as the seed of the structure that we observe in the Universe. The theory of cosmic inflation can generate these initial inhomogeneities—the primordial perturbations. Let us now briefly discuss this process.

Early works that describe the creation of near scale-invariant primordial density perturbations in the context of cosmic inflation are [7, 13]. The near scale-invariance part is important: in fact, the argument that primordial density perturbations have to be scale invariant precedes the theory of inflation [14, 15, 16]. During inflation, the quantum fluctuations of a homogeneous scalar field (inflaton) generate a near scale-invariant spectrum of density perturbations [7]; the fluctuations are generated on smaller scales and get expanded to superhorizon scales during the inflationary expansion. These stretched modes eventually reenter the horizon during the FLRW expansion phase and we observe them as a near scale-invariant power spectrum.

With the scalar field theory given by Eq.(1.7), one can study the perturbation to the scalar field $\delta \phi(x,t)$ defined as [17]

$$\phi(x,t) = \phi^{(0)}(t) + \delta \phi(x,t)$$
and obtain a scale-invariant power spectrum:

\[ P_{\delta\phi} = \frac{H^2}{2k^3} \]  \hspace{1cm} (1.14)

The power spectrum for any field \( \delta\phi \) is defined as:

\[ \langle \delta\phi(k)\delta\phi(k') \rangle = (2\pi)^3 \delta_D(k + k') P_{\delta\phi}(k) \]  \hspace{1cm} (1.15)

The perturbation \( \delta\phi \) is not a gauge invariant quantity. It is useful to define a gauge invariant quantity \( \zeta \) called the curvature perturbation [18, 7] that is conserved outside the horizon [19], and is related to \( \delta\phi \) in a spatially flat slicing as:

\[ \zeta = -\frac{H}{\dot{\phi}(0)} \delta\phi \]  \hspace{1cm} (1.16)

where the dot is a derivative with respect to the coordinate time \( t \). Then the power spectrum of the curvature perturbation can be written in terms of the slow roll parameter \( \epsilon \) as

\[ P_\zeta = \frac{2\pi GH^2}{\epsilon k^3} \bigg|_{aH = k} \]  \hspace{1cm} (1.17)

We do not provide further details of this calculation and direct the interested reader to [17] whose discussion we follow, or any other textbook on modern cosmology. More importantly, from the above equation, we can relate the energy scale during inflation (given by the approximately constant Hubble parameter \( H \)), and the slow roll parameter \( \epsilon \) to the power spectrum of the gauge invariant curvature perturbation. These can then be related to the perturbations in the gravitational potential (\( \Phi \)), which are used to compute the CMB angular power spectrum and the large-scale structure power spectrum.

Observationally, we cannot determine both \( H \) and \( \epsilon \) from the measurement of CMB temperature fluctuations and the matter density fluctuations. We can only measure the amplitude and the spectral tilt of the primordial spectrum. Usually, therefore, we write the power spectrum of the primordial curvature perturbation as:

\[ P_{\zeta}(k) = \frac{2\pi^2}{k^3} A_\zeta \left( \frac{k}{k_0} \right)^{n_s-1} \]  \hspace{1cm} (1.18)

or equivalently, we can define the dimensionless curvature perturbation as

\[ P_\zeta(k) = \frac{k^3 P_{\zeta}(k)}{2\pi^2} = A_\zeta \left( \frac{k}{k_0} \right)^{n_s-1} \]  \hspace{1cm} (1.19)

where \( A_\zeta \) is the amplitude of the curvature perturbation and \( n_s \) is the spectral index.
The constraints on these parameters from the Planck mission 2015 data release (Planck temperature data combined with Planck lensing) are [20]:

\[
A_\zeta = (2.139 \pm 0.063) \times 10^{-9} \\
n_s = 0.968 \pm 0.006
\]

Further, the power spectrum of the curvature perturbation is directly related to the power spectrum of the Bardeen potentials after inflation as \( P_\Phi(k) = P_\Psi(k) = (4/9)P_\zeta(k) \).

1.3.1 Cosmic microwave background (CMB)

The observation of a background of thermal radiation by Penzias and Wilson [1] was promptly proposed as the cosmic blackbody radiation [21] as a consequence of an expanding universe. The cosmic microwave background spectrum was measured and shown to have a blackbody spectrum by the FIRAS instrument of the Cosmic Background Explorer (COBE) satellite mission [22]. Further, anisotropies in the cosmic microwave background were also detected by the COBE mission [2]. See Figure 1.1. More recently, improved sensitivities of the WMAP [3] and the Planck [4] satellite missions have provided us CMB anisotropies (both temperature and polarization) data on much smaller scales.

As a function of the direction on the sky from the point of observation \( \hat{n} \), the tiny temperature fluctuations, \( \frac{\Delta T}{T} \), about the mean temperature, \( T_{\text{cmb}} \), can be written as:

\[
T(\hat{n}) = T_{\text{cmb}} \left[ 1 + \frac{\Delta T}{T}(\hat{n}) \right] \tag{1.20}
\]

These fluctuations originate from the primordial fluctuations discussed in the last section. It is convenient to decompose the temperature fluctuations of the CMB into spherical harmonics:

\[
\frac{\Delta T}{T}(\hat{n}) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\hat{n}), \tag{1.21}
\]
with the spherical harmonics normalization given by

\[ \int d\Omega_n Y^*_{\ell m}(\hat{n}) Y_{\ell' m'}(\hat{n}) = \delta_{\ell \ell'} \delta_{mm'}. \]  

(1.22)

Therefore, the multipole coefficients \( a_{\ell m} \) of the CMB temperature fluctuations are given by

\[ a_{\ell m} = \int d\Omega_n Y^*_{\ell m}(\hat{n}) \frac{\Delta T}{T}(\hat{n}). \]  

(1.23)

**The CMB power spectrum:** The variance of these multipole coefficients gives the CMB angular power spectrum \( C_\ell \)'s:

\[ \langle a_{\ell m} a^*_{\ell' m'} \rangle = \delta_{\ell \ell'} \delta_{mm'} C_\ell. \]  

(1.24)

The delta function above is a statement of isotropy; non-zero off-diagonal correlation implies violation of isotropy. Strictly speaking the averaging should be performed for \( a_{\ell m} \) values in different realizations of CMB skies. However, we have access to just one. The isotropy of the CMB temperature fluctuations comes to our rescue: in that case for each multipole \( \ell \) the different \( 2\ell + 1 \) \( a_{\ell m} \) multipole coefficients can be treated as independent.

Observationally, therefore,

\[ \hat{C}_\ell = \frac{1}{2\ell + 1} \sum_{m=\ell}^{\ell} |a_{\ell m}|^2 \]  

(1.25)

**Cosmic variance:** The finite number of \( a_{\ell m} \) for each \( \ell \) in the average above means that there is a limit to how precisely we can determine the value of each \( C_\ell \). Since \( a_{\ell m}^2 \) are \( \chi^2_{2\ell+1} \) distributed, the variance for each \( \hat{C}_\ell \) measurement is:

\[ \sigma^2(\hat{C}_\ell) = \frac{2C_\ell^2}{2\ell + 1} \]  

(1.26)

**The Sachs-Wolfe Effect:** In the inflationary picture, the largest scales have only recently re-entered the horizon, therefore, they probe the initial density perturbations directly. This gives rise to the Sachs-Wolfe [23] plateau at large CMB scales. The smaller scales that have been inside the horizon for longer have been affected by the late-time evolution of the universe, and therefore one needs to account for these to predict the spectrum observed in the CMB. These are encoded in the CMB transfer functions \( g_\ell(k) \).
Figure 1.2: The theoretical angular power spectrum $C_\ell$ from the best fit cosmological parameters obtained from the Planck 2015 mission. On the top we plot $C_\ell$s directly, whereas on the bottom we have shown $D_\ell = \ell(2\ell + 1)C_\ell/2\pi$ which is the quantity that is generally plotted when displaying the angular power spectrum results. Further, we have included the Sachs-Wolfe angular power spectrum (blue dashed) that does not have any of the acoustic oscillations, rescaled to match the quadrupole $C_2$ with the full angular power spectrum.

$$C_\ell = \int \frac{dk}{k} g_\ell(k)P_\phi(k)$$  \hspace{1cm} (1.27)

For Sachs-Wolfe temperature fluctuations ($\Delta T = -\Phi/3$ for adiabatic fluctuations [24]) and when the Bardeen potential $\Phi$ is a Gaussian field ($\Phi(x) = \phi(x)$), the Sachs-
Wolfe $a_{\ell m}$ is given by [25]:

$$a_{\ell m} = -\frac{4\pi}{3} i \int \frac{d^3k}{(2\pi)^3} \phi(k) j_\ell(kx) Y_{\ell m}(\hat{k}).$$  (1.28)

From this, we can obtain the Sachs-Wolfe angular power spectrum:

$$C_{\ell}^{SW} = \frac{4\pi}{9} \int_0^\infty \frac{dk}{k} j_\ell^2(kx) P_\phi(k),$$  (1.29)

where we have used the dimensionless power spectrum $P_\phi$ defined as $P_\phi(k) = 2\pi^2 P_\phi(k)/k^3$.

We conclude the discussion of some basic quantities in the CMB by showing the angular power spectrum $C_\ell$s from the 2015 data release of the Planck satellite mission\(^1\) in Figure 1.2. We show plots for both the angular power spectrum $C_\ell$s and the weighted angular power spectrum $D_\ell$s, which are defined as

$$D_\ell = \frac{\ell(\ell+1)C_\ell}{2\pi}.$$

Further, we also plot the Sachs Wolfe angular power spectrum, which ignores all the late-time effects on the temperature fluctuations.

1.3.2 Large-scale structure (LSS)

We observe regions with extreme matter density—for example galaxies—and at the same time regions with very small matter density i.e. voids. This rich structure evolves due to gravity from the small primordial inhomogeneities. The regions with higher density of matter get denser and the regions with lower matter density get less dense as time progresses. This paradigm is called gravitational instability, which is quite simple but beautifully explains the overdense and underdense distributions of matter that we observe, given that we start with an initial distribution that is not completely homogeneous.

As with the primordial perturbations, it is easier to work out the evolution of the matter field (perturbation about the mean matter density) $\delta(x)$ in Fourier space. The matter density field is related to the gravitational potential $\Phi(k)$ by the Poisson’s equation $k^2\Phi(k) = 4\pi G \rho_m a^2 \delta(k)$ [17]. The time evolution of the potential at late times $\Phi(k,a)$ depends on the wavenumber $k$ and also on the energy content of the Universe (and therefore on the scale factor). The dependence on the wavenumber $k$ is encoded in the matter transfer function $T(k)$ that describes the evolution of a mode after entering the horizon. The time dependence is encoded in the growth function $D(a)$. Therefore,

\(^1\)The values of the $C_\ell$s and other public data from the Planck mission can be obtained from the Planck Legacy Archive (PLA) website: http://www.cosmos.esa.int/web/planck/pla.
we can relate a primordial potential $\Phi(k)$ to $\delta(k, z)$ as

$$
\delta(k, z) = \alpha(k, z)\Phi(k)
$$

$$
\alpha(k, z) = \frac{2k^2T(k)D(z)}{3H_0^2\Omega_m}
$$

(1.30)

Here we are using the redshift $z = 1/a - 1$ instead of the scale factor $a$. Therefore, we can relate the primordial power spectrum to the matter power spectrum as

$$
P(k, z) = \alpha^2(k, z)P_\Phi(k)
$$

(1.31)

**Tracers of the density field and bias:** Generally, we do not observe the matter density field directly—what we observe are the galaxies and other such “tracers of the density field.” These trace the underlying matter density field i.e. we expect larger galaxies to form where the underlying matter overdensity is higher. See Figure 1.3 for a visualization of a slice of a matter density field from a simulation. Similarly, we expect correlation between the clustering statistics of the underlying matter density and the clustering statistics of the tracers [26, 27]. To leading order, therefore,

$$
\delta_g \propto \delta
$$

$$
P_g(k) \propto P(k)
$$

(1.32)

We do not expect the proportionality constant to be exactly unity. Let us call the constant $b$: $\delta_g = b\delta$ (linear bias). The bias $b$ depends on the details of the formation of the galaxy type, the redshift, the mass etc [28]. Further, the bias is scale dependent on smaller scales and tends to a constant value on larger (linear) scales [29]. Also, if we are looking at smaller scales, the linear bias approximation breaks down and we will have to consider higher-order bias terms.

**Smoothing and window functions:** In Figure 1.3, the density field of the square slice has been computed in a $32 \times 32$ grid. That is, we do not have information on density variations on scales smaller than about 18.75 Mpc/h. Therefore, the field visualized in the plots is actually the smoothed density contrast field: $\delta_R(x)$, which is simply a convolution of a window function with the matter density field.

$$
\delta_R(x) = \int d^3x' W_R(x - x')\delta(x')
$$

(1.33)

The choice for the windows function depends on the context. For example, in section 4.2, we will use the real-space spherical top hat:

$$
W_R(x) = \text{constant, if } |x| < R
$$
Figure 1.3: Matter density field $\delta(x)$ and halos. In the left, we plot the matter density field of a $(600 \text{ Mpc}/h)^2$ slice from a N-body simulation. This is superimposed with the halos (black circles) detected on the right plot. Larger circles correspond to more number of halos near the grid point. We have used a sample of halos with masses in the range $(4.83 \leq M \leq 9.55) \times 10^{13} h^{-1} M_\odot$. In the right plot, we can visually see that the halos follow the underlying matter density field and therefore are biased tracers of the density field. The simulation used to make the plots above used Gaussian initial conditions; more details about the simulations can be found in Chapter 4. The top panel shows the density field and the halos formed at an earlier time $z = 1$ than the bottom panel $z = 0$. We can therefore see the effect of gravitational instability at work: the denser regions get denser. This results in a greater number of dark matter halos being formed at $z = 0$.

$$0, \text{ if } |x| \geq R$$

In Fourier space, the convolution theorem implies that the smoothed density field is a product of the field in Fourier space and the windows function in Fourier space: $\delta_R(k) = W_R(k)\delta(k)$. For a spherical top-hat function, the Fourier transform is:

$$W_R(k) = \frac{3\sin(kR) - 3(kR)\cos(kR)}{(kR)^3}$$
with the constant in Eq.(1.34) given by $1/V_R$ if we normalize the window function as \( \int d^3x W(x) = 1 \). Alternatively, one can set the constant to unity in which case the Fourier space window function expression above should be multiplied by $V_R$.

**Correlation function:** The quantities we have defined so far are in Fourier space. It is generally easier to work in Fourier space as modes of different wavelengths are decoupled, whereas the corresponding real-space quantities are correlated. For example, the Fourier transform of the power spectrum is the two-point correlation function in real space: \( \langle \phi(x)\phi(x + r) \rangle = \xi(r) \). And,

\[
\xi(r) = \int \frac{d^3k}{(2\pi)^3} P(k)e^{ik\cdot r} \tag{1.36}
\]

Similarly, we can define the correlation function of smoothed quantities:

\[
\xi_R(r) = \int \frac{d^3k}{(2\pi)^3} W^2_R(k) P(k)e^{ik\cdot r} \tag{1.37}
\]

We will make use of the smoothed correlation function in section 4.4 in discussing correlation between two spherical subvolumes of radius $R \text{Mpc}/h$, separated by $r \text{Mpc}/h$. 


Chapter 2

Introduction to Non-Gaussianity

2.1 Non-Gaussian statistics

In this chapter, we review the language used to describe primordial non-Gaussianity. We have already defined the power spectrum in the previous section, which contains all the information for a Gaussian distribution. Here, we will review the power spectrum and then introduce higher-order correlation functions. We will primarily work in Fourier space. Our Fourier convention is:

\[
\phi(x) = \int \frac{d^3k}{(2\pi)^3} e^{i k \cdot x} \phi(k); \\
\phi(k) = \int d^3x e^{-i k \cdot x} \phi(x). \tag{2.1}
\]

where $\phi(x)$ is a perturbation field (real) with zero mean $\langle \phi(x) \rangle = 0$; for example, it could denote fluctuations in the primordial density field. Generally, we will use $\delta(x)$ to denote late-time matter density fluctuations in the large-scale structure. Following the usual practice in the cosmology literature, we will denote both the real-space and Fourier-space quantities by the same name. The reader should understand whether a quantity is in real space or Fourier space through its argument.

**The reality condition:** It follows from the reality of $\phi(x)$ that

\[ \phi(-k) = \phi^*(k) \]

**The power spectrum:** The power spectrum $P(k)$ is related to the two-point correlation function in Fourier space as follows

\[ \langle \phi(k)\phi(k') \rangle_c = (2\pi)^3 \delta_D(k + k') P(k). \tag{2.2} \]
Figure 2.1: A squeezed triangular configuration of bispectrum. In the squeezed configuration, one of the momenta is much smaller than the other two: $q = |q| \ll k = |k| \approx |k - q|$. In case of a large amplitude bispectrum in this configuration, one expects that the small-scale power spectrum (larger wavenumber) $P(k)$ is affected by the large-scale modes $q$.

The subscript $c$ stands for connected correlation function. The delta function signifies translational invariance. Further, rotational invariance implies that the power spectrum only depends on the magnitude of the wavevector $k$ i.e. $P(k) = P(|k|)$. The power spectrum (two-point statistics) completely describes a Gaussian distribution. Higher order correlations are either zero (odd number of fields), or can be written as a function of the two-point correlations (even number of fields). Therefore, we will now define a few higher order correlation functions in Fourier space that we will to study non-Gaussian statistics.

The bispectrum: The bispectrum $B(k_1, k_2, k_3)$ is defined as follows

$$\left\langle \phi(k_1)\phi(k_2)\phi(k_3) \right\rangle_c = (2\pi)^3 \delta_D(k_1 + k_2 + k_3)B(k_1, k_2, k_3). \quad (2.3)$$

As in the case of the power spectrum, the delta function here is a consequence of translational invariance. It is useful to notice that $k_1 + k_2 + k_3 = 0$ defines a triangle. For example, see Figure 2.1 in which we have taken $k_1 = k, k_2 = q, k_3 = -k - q$.

The squeezed bispectrum: The squeezed limit of a bispectrum has been studied quite extensively in the literature for a number of reasons. First, such a coupling of very long wavelength modes and short wavelength modes cannot be produced in the canonical single-field slow roll inflationary models [30]. Therefore, any detection of such a primordial signal presents a strong case towards multi-field dynamics during inflation.

Second, numerically it is easy to generate a bispectrum that peaks in this squeezed configuration; this can be done in real space as follows [25]:

$$\Phi_{NG}(x) = \phi_G(x) + f_{NL} \left( \phi_G^2(x) - \left\langle \phi_G^2(x) \right\rangle \right) \quad (2.4)$$
where $f_{NL}$ is the non-linearity parameter that determines the level of non-Gaussianity. The non-Gaussianity above is local in real space; therefore, the model is usually called the local ansatz or the local model of non-Gaussianity. The Fourier transform $\Phi_{NG}(k)$ will have a bispectrum $\left\langle \Phi_{NG}(k_1)\Phi_{NG}(k_2)\Phi_{NG}(k_3) \right\rangle$ that peaks in the squeezed limit.

The trispectrum: The trispectrum $T(k_1, k_2, k_3, k_4)$, similarly is defined as

$$\left\langle \phi(k_1)\phi(k_2)\phi(k_3)\phi(k_4) \right\rangle_c = (2\pi)^3 \delta_D(k_1 + k_2 + k_3 + k_4) T(k_1, k_2, k_3, k_4) \quad (2.5)$$

Similar to the case of the bispectrum, the squeezed limit of primordial trispectra have been studied extensively. And, one easy method to generate a trispectrum that peaks in the squeezed configuration is

$$\Phi_{NG}(x) = \phi_G(x) + g_{NL} \left[ \phi_G^3(x) - 3\phi_G \left\langle \phi_G^2(x) \right\rangle \right] \quad (2.6)$$

Dimensionless moments: It is sometimes useful to define the dimensionless moments to compare the amplitude of non-Gaussianity at different orders and their relative scaling.

$$\mathcal{M}_n = \frac{\left\langle \Phi(x)^n \right\rangle}{\left\langle \Phi(x)^2 \right\rangle^{\frac{n}{2}}} \quad (2.7)$$

We will make use of $\mathcal{M}_n$’s mostly in Section 4.2 of Chapter 4. $\mathcal{M}_3$ is also called the reduced skewness, $\mathcal{M}_4$ the reduced kurtosis and so on.

Note that the dimensionless moments are more useful to indicate the actual level of non-Gaussianity than the $f_{NL}$ and $g_{NL}$ parameters. For example, if we generate the
two non-Gaussian models given in Eq.(2.4) and Eq.(2.6) with $f_{\text{NL}} = 10$ and $g_{\text{NL}} = 10^5$ and $\sigma^2 = \langle \phi_G^2 \rangle = 10^{-9}$, the dimensionless moments are: $\mathcal{M}_3 \approx 6f_{\text{NL}}\sigma \approx 0.002$ and $\mathcal{M}_4 = 48g_{\text{NL}}\sigma^2 \approx 0.005$.

### 2.2 Non-Gaussian primordial fluctuations

Now, let us briefly discuss inflationary scenarios (distinguishing between single-source and multi-source) that generate different set of primordial non-Gaussianity.

A useful way of distinguishing between single and multi-field inflation is to consider if the same field is responsible for the background dynamics and the superhorizon curvature perturbations [31]. A model is single source if a single field is responsible for both the background inflationary dynamics (exponential expansion) and the generation of superhorizon curvature perturbations. In these models correlations between long and short wavelength modes are absent. The three point function of single field inflationary models is worked out in [32]. To summarize the relevant important finding, the three point function in the squeezed limit $q \ll k_1, k_2 \approx k$ follows the consistency relation:

$$\langle \zeta(q)\zeta(k_1)\zeta(-q-k) \rangle \approx -\langle \zeta(q)\zeta(-q) \rangle k \frac{d}{dk} \langle \zeta(k_1)\zeta(-k_1-q) \rangle$$

$$\Rightarrow B(q,k,-k-q) \approx -n_s P(q)P(k)$$ (2.8)

i.e. $f_{\text{NL}} = n_s - 1$, which is a very small level of non-Gaussianity; much smaller than the current best limits on $f_{\text{NL}}$ and smaller than the expected constraints from future surveys. More importantly, later works have shown that the contribution from this squeezed limit of the primordial bispectrum vanishes in the CMB and large-scale structure bispectra [33]. However, allowing for a field different than the inflaton to generate superhorizon curvature perturbations after the end of inflation can generate coupling between long and short wavelength modes and therefore local-type non-Gaussianity [34, 35, 36]. Therefore, a detection of such a mode coupling in the squeezed limit (long-short wavelengths) in the primordial perturbations is indicative of multi-field dynamics during inflation [37, 38]. See, however, [39].

### 2.3 The hunt for primordial non-Gaussianity

The simplest method to measure the effect of primordial non-Gaussianity is to directly measure higher-order statistics beyond the one and two-point functions. The excess signal that cannot be accounted by the expected signal from Gaussian primordial fluctuations could then be a signal of primordial non-Gaussianity. In the absence of such a signal, we can put limits on the amount of primordial non-Gaussianity allowed by the data. The direct search for the bispectrum in the CMB [25, 40] has been performed
both by the WMAP and Planck missions constraining the amplitude of non-Gaussianity for various bispectrum templates.

Bispectrum measurements with the large-scale structure [41] have just begun [42, 43] and have not been used to constrain primordial non-Gaussianity. However, the coupling between long and short wavelength modes have been shown to generate scale-dependent bias at large scales [44]. This effect has been used to constrain primordial non-Gaussianity [45, 46, 47, 48], although the constraints are not as strong as the CMB constraints. In the future, with larger volume galaxy surveys, the constraints from large-scale structure is expected to be an order of magnitude better than the current CMB constraints [49].

The next two chapters contain the main body of this thesis. The two sections of chapter 3 are based on [50] and [51] respectively. We show that statistical anisotropies in the CMB can occur in the presence of long-short mode couplings. Therefore, statistical anisotropies can be used to improve the signal in non-Gaussianity searches; conversely, we may be able to tell whether the data is suggestive of a non-Gaussian origin for observed statistical anisotropies in the CMB power spectrum. The sections 4.2 and 4.3 of Chapter 4 are based primarily on [52], where we study in detail the effects of single and two-field local-type non-Gaussian initial conditions on the mass function, bias and stochasticity of halos. Section 4.2 also consists of some results on non-Gaussianity measurements from [53]. Section 4.4 is based on work in progress; we study the position-dependent power-spectrum and the position-dependent bispectrum as efficient tools to measure non-Gaussianity from future galaxy surveys.
Chapter 3

Effects of Primordial Non-Gaussianity: I. CMB

The goal of this chapter is to discuss some of the effects of primordial non-Gaussianity in the cosmic microwave background. In particular, we will focus on the CMB temperature fluctuations and discuss how non-Gaussianity that couples superhorizon modes with the modes that are observed in the CMB can generate a hemispherical power asymmetry in the CMB.

There are two main parts to this chapter. First in Section 3.1, we introduce the CMB hemispherical power asymmetry anomaly using a local-variance method. Then in Section 3.2, we will discuss in detail the relation between non-Gaussian statistics and such a power asymmetry in the presence of superhorizon modes.

3.1 The CMB power asymmetry

3.1.1 Introduction

A number of studies have been performed that show approximately $3\sigma$ to $3.5\sigma$ hemispherical power asymmetry at large scales in WMAP and Planck CMB temperature fluctuations [54, 55, 56, 57, 58]. See Figure 3.1 to see an illustration (using a random CMB realization) of how a dipole modulation in temperature fluctuations generates a power asymmetry.

3.1.2 The local-variance method

Recently, [59] used a conceptually simple pixel space local variance method to demonstrate the presence of asymmetry at large scales in WMAP and Planck data at a significance of at least $3.3\sigma$; they find that none of the 1000 isotropic Planck FFP6 simulations had a local variance dipole amplitude equal to or greater than that found in data for disk radii $6^\circ \leq r_{\text{disk}} \leq 12^\circ$. In this method, a local variance map ($m_r$) is
Figure 3.1: An illustration of power asymmetry in simulated CMB maps. Here we have shown two modulations of a Gaussian and isotropic realization of a HEALPIX generated CMB map (top) using a dipole modulation of $A_T = 0.1$ (middle) and $A_T = 0.2$ (bottom) at all scales. The modulations are done in pixel space and therefore in temperature (Eq.(3.4). The corresponding modulation of the power spectrum is $2A_T$. 
generated at a smaller HEALPIX [60] resolution $N_{\text{side}}$ from a higher resolution CMB temperature fluctuations map by computing temperature variance inside disks of certain radius $r_{\text{disk}}$ centered at the center of each pixel of the HEALPIX map with resolution $N_{\text{side}}$. Isotropic simulations are used to get the expected mean map ($\bar{m}_r$), and each map is normalized:

$$m_r^n = \frac{m_r - \bar{m}_r}{\bar{m}_r}.$$  \hspace{1cm} (3.1)

Then, we obtain local variance dipoles by fitting for dipoles in each of these normalized maps (both simulations and Planck data) using the HEALPIX `remove_dipole` module with inverse variance (of the 1000 simulated local variance maps) weighting. We will denote the amplitude of a local variance dipole as $A_{LV}$. The local variance dipole obtained from data is then compared to the distributions obtained from simulations.

In this work, we make use of the local variance method and extend the results obtained in [61] to include smaller disk radii. The authors in [61] focused on large scale power asymmetry, and therefore only looked at large values of $r_{\text{disk}}$. After confirming their results at large disk radii ($r_{\text{disk}} \geq 4^\circ$), we perform the same local variance dipole analysis using smaller disk sizes. We find that, for smaller disk radii, the contribution of the Doppler dipole becomes increasingly significant. The Doppler dipole in the local variance map is an expected signal because of our velocity with respect to the CMB rest frame. While the direction and magnitude of the CMB dipole has been known from previous CMB experiments [62], the Doppler dipole signal in the temperature fluctuations is rather weak and reported only recently by the Planck Collaboration [63] using harmonic space estimators [64, 65]. We work towards detecting the expected Doppler dipole in the local variance maps after removing large scale features from the maps. Our goal therefore is two fold: first, extend the local variance dipole study of the hemispherical power asymmetry to smaller disk radii and second, use the method of local variance to detect the Doppler dipole whose amplitude is much smaller but is expected to contribute at all angular scales.

Before presenting the details of the analysis and results, we would like to point out and clarify a difference between our analysis and that of Akrami et al. They used 3072 disks ($N_{\text{side}} = 16$ healpix map) for all sizes of disks they considered ($r_{\text{disk}} \geq 1^\circ$). However, we find that for $r_{\text{disk}} = 1^\circ$ and $2^\circ$, 3072 disks are not enough to cover the whole sky. Therefore, we use $N_{\text{side}} = 32$ (12288 disks) for $r_{\text{disk}} = 2^\circ$ and $N_{\text{side}} = 64$ (49152 disks) for $r_{\text{disk}} = 1^\circ$. Once we do this, we find that, unlike the results in Akrami et al. (see Fig 2(a) and Table 1 in [61]), none of our 1000 isotropic simulations produce a local variance dipole amplitude larger than that of our foreground cleaned channel maps, for $r_{\text{disk}} = 1, 2^\circ$. In fact, the effect of the anomalous dipole (with respect to the isotropic case) can be observed for even smaller angular disk radii (see Figure 3.5). This result, however, is not surprising because the local variance dipoles computed at smaller
disk radii get some contribution from the large scale anomalous power asymmetry at low $\ell$.

Next, in section 3.1.3 we discuss in some detail the Planck CMB data and simulations used in this work. Then, in section 3.1.4 we present our results for both the anomalous dipole and the Doppler dipole, followed by discussions and a summary of our results in section 3.1.5.

3.1.3 Simulations and data

For our analysis, we use 1000 realizations of FFP6 Planck simulations ¹ that include lensing and instrumental effects. The FFP6 simulations (CMB and noise realizations separately) are provided at each Planck frequency. We generate local variance maps from which we obtain local variance dipole distributions following the method briefly introduced in the previous section (described in more detail in [61]). Since the FFP6 simulations for CMB and noise processed through the component separation procedure are not yet publicly available, we will generate foreground cleaned CMB channel maps at two frequencies 143 GHz and 217 GHz using the SEVEM method [66, 67]. We use the same four template maps as the Planck Collaboration; these are difference maps of two near frequency channels: (30-44) GHz, (44-70) GHz, (545-353) GHz, and (857-545) GHz. Before generating the four templates, we smooth the larger frequency channel map to the resolution of the smaller frequency channel map:

$$a_{\ell m}^{\text{large}} \rightarrow a_{\ell m}^{\text{large}} \left( \frac{B_{\ell}^{\text{small}}}{B_{\ell}^{\text{large}}} \right),$$

where $B_{\ell}$ is the beam transfer function for the given frequency channel map. See Appendix C of [66] for details of this method. To summarize: the cleaned 143 GHz or 217 GHz map $d_\nu^{\text{clean}}$, is obtained by subtracting from the uncleaned channel map $d_\nu$, a linear combination of the templates $t_j$:

$$d_\nu^{\text{clean}} = d_\nu - \sum_{j=1}^{n_t} \alpha_j t_j$$  \hspace{1cm} (3.2)

The coefficients $\alpha_j$’s are obtained by minimizing the variance of $d_\nu^{\text{clean}}$ outside of the mask used for our analysis. Once the linear coefficients for the four template maps are obtained, we use the same coefficients to process combinations of noise FFP6 simulations in order to generate noise maps for our foreground cleaned simulated maps. We process the Planck maps and the simulated maps identically in each step. In addition to the isotropic local variance maps (obtained from isotropic realizations of CMB plus noise),

¹http://crd.lbl.gov/groups-depts/computational-cosmology-center/c3-research/cosmic-microwave-background/cmb-data-at-nersc/
Figure 3.2: The hemispherical asymmetry in $C_\ell$s generated by our modulation model that modulates large scales (1.25° smoothing) with an amplitude $A_T = 0.073$ in real space, in addition to the dipolar modulation at all scales due to the CMB Doppler effect. This plot was generated using 1000 simulations for the 217 GHz channel.

we obtain two new sets of simulated local variance maps from two other models which are obtained from the isotropic realizations as explained below:

- **Doppler model**: The expected dipolar temperature modulation from the Doppler effect, with an amplitude $0.00123 b_\nu$ along the direction $(l, b) = (264°, 48°)$ [62]. We take $b_\nu = 1.96$ for the 143 GHz map and $b_\nu = 3.07$ for the 217 GHz map [63]. The dipolar modulation is generated in pixel space simply as:

$$\frac{\Delta T}{T}_{\text{dop}} (\hat{n}) = \left[ 1 + 1.23 \times 10^{-3} b_\nu (\hat{n} \cdot \hat{p}) \right] \frac{\Delta T}{T}_{\text{iso}} (\hat{n})$$  \hspace{1cm} (3.3)

- **Modulation model**: The Doppler model as described above plus a large angle modulation, corresponding to a modulation at smoothing fwhm = 1.25°, along the direction $(l, b) = (218°, -20°)$, with an amplitude $A_T = 0.073$ [56]. The map obtained is:

$$\frac{\Delta T}{T}_{\text{mod}} (\hat{n}) = \left[ \frac{\Delta T}{T}_{\text{dop}} (\hat{n}) + A_T (\hat{n} \cdot \hat{p}) \frac{\Delta T}{T}_{\text{smoothed dop}} (\hat{n}) \right]$$ \hspace{1cm} (3.4)

in which the modulation is only applied to the Doppler model after smoothing at fwhm = 1.25°. Here we have used the notation $A_T$ for the amplitude of modulation.
for temperature fluctuations in the above equation to distinguish it from the dipole modulation amplitude in local variance maps $A_{LV}$. Our choice of the smoothing scale $\text{fwhm} = 1.25^\circ$ is guided by our attempt to fit the obtained local variance asymmetry in data for different $r_{\text{disk}}$ that we have considered. For example, we find that a larger $\text{fwhm} = 5^\circ$ modulation does not reproduce the local variance asymmetry in data for $r_{\text{disk}} = 1^\circ$ or smaller. Similarly, using a smaller $\text{fwhm}$ produces local variance asymmetry distributions with amplitudes too large to be consistent within $2\sigma$ with that seen in data for some disk radii ($r_{\text{disk}} = 1^\circ$ and $4^\circ$, for example). Note that when we apply modulation at large scales in this manner, we are using a filter that will suppress modulation at scales much smaller than the smoothing $\text{fwhm}$ specified. Therefore, by construction, the modulation is only generated at large scales. In Figure 3.2, we plot the asymmetry generated by our modulation model in harmonic space (up to $l = 600$); the quantity plotted is:

$$A_{\ell} = 2 \frac{C_{\ell}^+ - C_{\ell}^-}{C_{\ell}^+ + C_{\ell}^-} \quad (3.5)$$

where $+$ is taken to be the hemisphere centered at $(l, b) = (218^\circ, -20^\circ)$ and $-$ is the opposite hemisphere. The $C_{\ell}$s for the relevant hemisphere is computed by masking the rest of the sky.

**Masks:** For each map, we use the union mask U73 from the Planck Collaboration. This is the union of the qualitative masks for the four different component separation techniques employed by the Planck Collaboration, and leaves approximately 73 percent of sky unmasked. The template fitting in our SEVEM-like cleaning procedure is also done in the region outside this mask. The resulting clean maps at 143 GHz and 217 GHz (with the U73 mask) are shown in Figure 3.3.

### 3.1.4 Results

First, let us comment that we get results consistent with [61] for $r_{\text{disk}} = (4^\circ$ to $16^\circ$), in terms of the significance of the magnitude, and the direction of the anomalous dipole. Also, our modified set of simulations that includes temperature modulation at large angular scales produces dipole distributions consistent with data. All our results presented in this section were analyzed using the foreground cleaned 217 GHz maps unless otherwise stated. For large scale results i.e. relating to the anomalous dipole, we get similar results using the 143 GHz maps.

#### 3.1.4.1 Local variance dipole for $r_{\text{disk}} = 1^\circ, 2^\circ$

For $r_{\text{disk}} = 1^\circ, 2^\circ$, our results in Figure 3.4 indicate that the effect of the Doppler dipole is increasing compared to larger disk radii. Further, the results also show that the
modulation model that we have chosen (modulated at large scales corresponding to a smoothing of fwhm = 1.25°, with an amplitude $A_T = 0.073$) is consistent with data, while the no modulation (isotropic) model is disfavored at approximately 3.5σ. Note that the distribution of $A_{LV}$ is not exactly Gaussian, but we assume the distribution to be Gaussian to obtain standard deviation significance values as rough guides.

Next, we extend the analysis to apply to even smaller disk radii.

3.1.4.2 Local variance dipole for sub-degree disk radii

We will now consider disks of radii $r_{\text{disk}} = 0.25°$ and 0.18°. This requires that we increase the number of disks to cover the whole sky. We therefore use disks centered on pixels of a healpix map of $N_{\text{side}} = 256$ (786432 disks). Figure 3.5 shows our results for these two disk sizes. From the figures, it is clear that for these disk radii, the effect of the Doppler dipole cannot be neglected when computing the significance of the dipole amplitude obtained in the data. If we do not include the Doppler effect, then the local variance dipole in the Planck map is anomalous at approximately 3.8σ for both $r_{\text{disk}} = 0.25°$ and 0.18°, whereas with respect to the Doppler model distribution the significance drops to about 2σ. Also, the direction of the dipole detected from

Figure 3.3: Foreground cleaned maps at channels 143 GHz and 217 GHz. Planck U73 mask is used in both maps.
Figure 3.4: Local variance dipole analysis for $r_{\text{disk}} = 1.0^\circ$. The vertical line and the text describe the measurements from Planck maps. The histograms are obtained from the simulations for three cases— isotropic, Doppler and large scale modulation models—as labeled. For $r_{\text{disk}} = 2^\circ$ (not shown), we get local variance dipole of amplitude $A_{LV} = 0.026$ (2.8$\sigma$) in the direction $(l, b) = (203^\circ, -1^\circ)$. The significance values in the above figures are computed using the Doppler model distributions. If instead we use the isotropic distributions, we obtain values of 3.5$\sigma$, 3.4$\sigma$ and 4.0$\sigma$ for $r_{\text{disk}} = 1^\circ$, 2$^\circ$ and 4$^\circ$ respectively.

Planck maps gets closer to the Doppler dipole as the disk radius is decreased. This is illustrated in Figure 3.7. However, as seen in Figure 3.5, the effect is still small and the distributions of the isotropic simulations and the Doppler model simulations overlap quite a bit. We will now investigate if, by removing large scale features from the temperature fluctuation maps, it is possible to separate the Doppler modulated local variance dipole distribution from the isotropic one, and therefore measure the Doppler dipole in the Planck maps.
Figure 3.5: Local variance dipole analysis for $r_{\text{disk}} = 0.25, 0.18$ degrees. The vertical line and the text describe the measurements from Planck maps. The histograms are obtained from the simulations for our three different models as labeled. The significance values in the figures are calculated using the Doppler model distributions. If instead we use the isotropic distributions, the corresponding significances become $3.78\sigma$ and $3.82\sigma$ respectively.

3.1.4.3 The Doppler dipole in local variance maps

We remove large scale features by simply filtering a map using a high-$\ell$ filter i.e. set the low-$\ell$ $a_{\ell m}$ values to zero in multipole space. Since we are looking for an expected signal that is a vector (i.e. to test the significance of the measurement we need to look at both the direction and magnitude of the dipole signals), we will determine the distribution of the component of the dipoles obtained in the direction of the known CMB dipole. This was also the approach taken by the Planck Collaboration [63]; the quantity whose
Figure 3.6: The component of local variance dipole amplitudes in the direction \((l, b) = (264^\circ, 48^\circ)\) using disk of radius \(r_{\text{disk}} = 0.18^\circ, 8^\circ\) after removing large scales features up to \(\ell_{\text{min}} = 600\). In both cases, one obtains the Doppler signal at nearly \(3\sigma\), while no effect is seen in the orthogonal directions (not shown in figures). The results for our large scale modulation model are not shown above but they mostly coincide with the Doppler models (as we have removed the large scale features), with a very small right shift as expected.

distribution they plot, called \(\beta_{\parallel}\), is the component of their estimator in the CMB dipole direction.

In Figure 3.6, we show results for our local variance analysis using disk radius \(r_{\text{disk}} = 0.18^\circ\) but using CMB temperature maps with \(a_{\ell m} = 0\), for \(\ell \leq \ell_{\text{min}} = 600\) i.e. large scale features removed up to \(\ell = 600\). The signal in the direction parallel to the CMB dipole is consistent with the Doppler model distribution while the isotropic model is disfavored at \(3.2\sigma\). When using a larger disk radius \(r_{\text{disk}} = 8^\circ\) (also shown in Figure
We have repeated this analysis using other disk radii, different $\ell_{\text{min}}$ and the 143 GHz channel map. The results are summarized in Table 3.1. Unexpectedly, we obtain higher values of amplitude for the 143 GHz channel but in most cases the dipole amplitude in the direction of CMB dipole is within $2\sigma$ of the expected distribution obtained from the Doppler model. The only exception being the case of $\ell_{\text{min}} = 900$, 143 GHz, $r_{\text{disk}} = 0.18^\circ$ in which the signal in data is at $3.1\sigma$ from the Doppler model distribution. In all cases, no amplitude in excess of $2\sigma$ is obtained in two directions orthogonal to the CMB dipole and with each other (see Figure 2 of [63]).

The local variance dipole directions obtained for various cases are shown in Figure 3.7.

3.6.4 Small-scale power asymmetry

We can also investigate the presence of power asymmetry at small scales in the direction $(l, b) = (218^\circ, -20^\circ)$, in our small scale maps that have $a_{\ell m}$’s up to $\ell = 600$ set to zero. A similar study of the power asymmetry at small scales (but in multipole space) was performed in [57] using foreground cleaned SMICA maps; they report no such small scale asymmetry after accounting for a number of effects, including estimates
Table 3.1: Summary of the Doppler dipole detection results. The significance of detection is computed with respect to the corresponding isotropic distribution for local variance amplitudes (1000 simulations) in the CMB dipole direction.

<table>
<thead>
<tr>
<th>$l_{\text{min}}$</th>
<th>$r_{\text{disk}}$</th>
<th>217 GHz</th>
<th>143 GHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>600</td>
<td>0.18</td>
<td>0.0042 (3.2σ)</td>
<td>0.0057 (3.4σ)</td>
</tr>
<tr>
<td>900</td>
<td>0.18</td>
<td>0.0027 (2.9σ)</td>
<td>0.0047 (4.1σ)</td>
</tr>
<tr>
<td>600</td>
<td>8.0</td>
<td>0.0048 (2.7σ)</td>
<td>0.006 (2.4σ)</td>
</tr>
<tr>
<td>900</td>
<td>8.0</td>
<td>0.0042 (2.6σ)</td>
<td>0.0071 (2.9σ)</td>
</tr>
</tbody>
</table>

of power asymmetry due to the Doppler effect. They report an upper bound on the modulation amplitude of 0.0045 (95%) at these scales. Another important previous work on constraining hemispherical power asymmetry at smaller scales is [68] using quasars that reports $-0.0073 < A_T < 0.012$ (95%) at $k_{\text{eff}} \approx 1.5 h \text{Mpc}^{-1}$.

When repeating our measurements shown in Figure 3.6, but now in the direction $(l, b) = (218^\circ, -20^\circ)$, we find that the local variance dipole amplitude component in the data is within $1\sigma$ expectation from both isotropic and Doppler model distributions. This is shown in Figure 3.8. Each value in the Doppler model distribution is basically shifted right from the isotropic distribution (see the inset in Figure 3.8); this shift is proportional to the cosine of the angle between the direction of large scale asymmetry and the direction of the CMB Doppler dipole. Therefore, we subtract this shift to get the value of the intrinsic local variance power asymmetry. Using $r_{\text{disk}} = 0.18^\circ$, we obtain (at 2σ):

$$A_{LV} = (0.71 \pm 3.0) \times 10^{-3}$$  \hspace{1cm} (3.6)

We show the Doppler-subtracted small-scale power asymmetry results in Figure 3.9.

However, we can see from our results for the Doppler model distributions that the relation between $A_T$ of a dipole modulated model and the most likely value of $A_{LV}$ obtained from the corresponding local variance maps is not quite simple. This can be directly seen in Figure 3.6 in which the Doppler model had $A_T = 3.07 \times 0.00123 \approx 0.0038$ whereas the obtained distributions for $A_{LV}$ are different for the two cases with different $r_{\text{disk}}$. To estimate the $A_T$ constraint from the $A_{LV}$ constraint in eqn 3.6, let us assume that the relationship between them is linear for small values of $A_T$ up to approximately $A_T = 0.0038$ (for a given frequency channel, $r_{\text{disk}}$ and $\ell_{\text{min}}$). Then, for the case of the 217 GHz channel maps, $r_{\text{disk}} = 0.18^\circ$ and $\ell_{\text{min}} = 600$, we can use the correspondence between $A_T$ and $A_{LV}$ of the distributions from simulations in Figure 3.6, and translate...
Figure 3.8: Small-scale power asymmetry in the direction \((l, b) = (218^\circ, -20^\circ)\) before subtracting the contribution from the Doppler dipole. The inset shows the distribution for the difference between local variance dipole amplitudes for the Doppler model and the isotropic model. This value \(0.00075 \pm 0.00001(2\sigma)\) is subtracted from the power asymmetry in data to obtain the constraint for the intrinsic power asymmetry.

Figure 3.9: Small-scale power asymmetry in the direction \((l, b) = (218^\circ, -20^\circ)\) after subtracting the Doppler contribution. We also show the corresponding \(A_T\) on the top horizontal axis, in addition to \(A_{LV}\) on the bottom axis; the fact that these are different is discussed in the text. The error bands show the expected 1\(\sigma\) and 2\(\sigma\) errors based on the variance of the isotropic distribution.
the constraint in eqn 3.6 to (at 2σ):

\[ A_T = (0.8 \pm 3.5) \times 10^{-3} \]  

(3.7)

To further check our translation between \( A_{LV} \) and \( A_T \), we repeated this small scale analysis with a \( A_{T, \text{input}} = 0.0008 \) modulation model and recovered the intrinsic local variance dipole amplitude obtained in eqn 3.6. Using other larger values of \( r_{\text{disk}} \) or the 143 GHz channel map, we obtained slightly weaker but consistent constraints for the small-scale power asymmetry \( A_T \) in the direction of the large scale hemispherical power asymmetry.

### 3.1.5 Discussion and Summary

In this section, we have used local variance maps to study the power asymmetry in Planck temperature anisotropy maps, extending the work done in [61] that used the same method to smaller scales. We have shown that the effect of the Doppler dipole is small for local variance dipole measurements at large disk radii (\( r_{\text{disk}} \gtrsim 4^\circ \)); we find that the peak of the Doppler model distribution for the 217 GHz case is less than 0.5σ away from that of the isotropic distribution for the \( r_{\text{disk}} = 4^\circ \) result. The difference gets even smaller for larger values of \( r_{\text{disk}} \). Further, we have also shown that the Doppler dipole can be measured to a moderate significance using this method. The first measurement of the Doppler signal in the temperature fluctuations was done by the Planck Collaboration in [63]. The method of local variance is a pixel space method, so our measurement is complimentary to that done by the Planck team using harmonic space estimators sensitive to the correlations between different multipoles induced by the Doppler effect. However, local variance dipoles are not sensitive to the aberration effect and therefore our method is less sensitive than the estimators that consider the aberration effect in addition to the dipolar power modulation.

For our large scale asymmetry analysis, we find similar results using both 217 GHz channel maps (reported in figures in this section) and 143 GHz channel maps. This is in agreement with previous works [58, 69] that have investigated channel dependence of hemispherical power asymmetry in WMAP maps.

Throughout our analysis, we included a large scale modulation model in which temperature anisotropies were modulated only for scales smoothed (Gaussian smoothing) at \( \text{fwhm}=1.25 \) degrees and found that the anomalous power asymmetry seen at all \( r_{\text{disk}} \) analyzed using the local variance method is consistent with this simple phenomenological model. This tells us about the scale dependence of the power asymmetry seen in data as our modulation model generates a scale dependent hemispherical asymmetry in \( C_\ell s \) (see Figure 3.2).

We summarize important aspects of our work and results in the following points:

- We have verified the hemispherical power asymmetry results obtained in [61], with
the exception of $r_{\text{disk}} = 1, 2$ degrees for which we have identified the need to use more disks in order to cover the whole sky.

- Once we use smaller disk radii in our local variance analysis, the effect of the dipolar Doppler modulation becomes increasingly important which can be directly observed from our dipole amplitude distributions.

- After removing large scale features up to $\ell = \ell_{\text{min}} = \{600, 900\}$, we could detect the expected dipolar Doppler modulation in Planck temperature anisotropy maps at a significance of approximately $3\sigma$.

- We have obtained a constraint on dipolar modulation amplitude at small scales ($\ell > 600$) in the direction of the large scale anomalous hemispherical power asymmetry as $A_T = 0.0008 \pm 0.0035(2\sigma)$.

We expect this work to be useful for a better understanding of the power asymmetries that are known to exist in the CMB data. In particular, we hope that our work will shed more light on local variance statistics in CMB maps which is already being used to compare theoretical models of power asymmetry with data [70]. While a satisfying theoretical explanation for the power asymmetry anomaly still lacks in the literature, several interesting proposals exist [71, 72, 73, 74, 75, 76, 77]. The statistical fluke hypothesis still remains a possibility too [78]. More careful studies of both data and theoretical possibilities are important since the implications of progress in any direction consists of learning more about the fundamental statistical assumptions that we make about the universe at large scales.

### 3.2 Non-Gaussianity and the power asymmetry

#### 3.2.1 Introduction

It is very tempting to try to use large-scale features of the primordial fluctuations, such as the hemispherical power asymmetry, as a clue toward primordial physics. Such signals are both intriguing (maybe they say something about the beginning of inflation) and statistically unfortunate (i.e., not all that unlikely to be a feature of a particular realization of Gaussian, isotropic fluctuations). Many studies of the observed power asymmetry and its statistical significance have been reported using the WMAP data [54, 58, 79] and the Planck data [56, 57, 59, 80, 81, 82]. A variety of possible explanations for this asymmetry have been discussed (see e.g. [71]): several of the most intriguing ideas use superhorizon fluctuations to generate the asymmetry, either by using non-Gaussianity to couple them to observable perturbations [83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 70, 99, 100, 101, 102, 103, 104, 105] or by postulating some different primordial physics that precedes the usual slow-roll inflation.
Alternatively, one can postulate scenarios that are fundamentally anisotropic on the largest scales [110, 111, 112, 113, 114]. It is quite general that if the primordial fluctuations are non-Gaussian, the likelihood of observing statistical anisotropies in our cosmic microwave background (CMB) changes. Although one might expect that isotropy and Gaussianity are independent criteria for the statistics of the primordial fluctuations, this distinction is not actually clear when we only have access to a finite volume of the universe [115, 116, 117]. If the primordial fluctuations are non-Gaussian, the observed large-scale “discrepancies” from the simplest isotropic, power law, Gaussian fluctuations need not be a signal of a special scale, time, or feature during the primordial (inflationary) era. Instead, they may simply be a consequence of cosmic variance in a universe larger than the volume we currently observe, filled with non-Gaussian fluctuations.

Here, we present a single framework to calculate the distribution of expected deviations from isotropy in our observed sky from any model with non-Gaussian primordial fluctuations. This framework incorporates most successful proposals for generating the power asymmetry, even some that were not originally formulated as non-Gaussian models. The reason is that if the assumption of statistical isotropy is maintained, any explanation of the temperature power asymmetry can be modeled by assuming a fluctuation in a long wavelength modulating field that couples to some cosmological parameter relevant for determining the CMB power spectrum (the fundamental constants, the scalar spectral index and the inflaton decay rate, for example [110, 71, 111, 93, 99, 118]). Since the asymmetry is well modeled by a (small) range of shorter scales all coupling to a longer scale this is, by definition, a sort of non-Gaussianity.

Using our analytic framework together with numerical realizations, we show that isotropy violations in our observable universe are more likely in models where subhorizon modes are coupled to longer wavelength modes. In some models they are so much more likely that the extent to which our observed sky is isotropic can be interpreted as a constraint on non-Gaussianity.

In addition, we find that scale dependence of such modulations is a generic outcome of non-Gaussianity beyond the standard local ansatz. This result is important because a major hurdle in model building for the power anomalies has been their scale dependence, as the observations suggest a very sharp decrease in the power asymmetry amplitude at smaller scales [57, 119, 50]. Scale dependence of the type observed occurs whenever modes closer to the present Hubble scale couple more strongly to superhorizon modes than very short wavelength modes do. This can be the case even for scale-invariant bispectra. For example, since equilateral-type non-Gaussianity peaks when the momenta configurations are nearly equal, the modulation from an equilateral-type bispectrum is biggest at large scales and quickly dies off at smaller CMB scales.

Further, we will see that any non-Gaussianity used to explain the scale-dependent power asymmetry will also produce a scale-dependent modulation of the power-spectrum.
monopole. This means that the observed low power on large scales and the dipole power asymmetry can both provide evidence in favor of non-Gaussian fluctuations. To quantify the significance of these signals, we perform a parameter estimation of the non-Gaussian amplitude and shape using both monopole and dipole modulations, and estimate Bayesian evidences against the Gaussian model.

Finally, we will show that scale-dependent non-Gaussianity eliminates the need for enhanced inhomogeneity on superhorizon scales to generate the observed asymmetry. The most common method for introducing a dipolar power modulation is to postulate the existence of a large-amplitude superhorizon fluctuation in a spectator field during inflation that then alters the power spectrum on smaller scales via local-type non-Gaussianity. (A multiple-field model of inflation, or one that otherwise breaks the usual consistency relation, is required because superhorizon perturbations in the standard inflaton field cannot generate an asymmetry [120].) Dubbed the Erickcek-Kamionkowski-Carroll (EKC) mechanism [83], this approach has been expanded upon and refined several times since its inception [84, 96, 89, 87, 90, 91, 86, 94, 92, 95, 88, 97, 98, 70, 101, 99, 100, 102, 103, 104, 105]. In these analyses, WMAP and Planck bounds on local-type non-Gaussianity forced the amplitude of the superhorizon perturbation to be much larger than predicted by an extrapolation of the observed primordial power spectrum to larger scales. Although possible origins for this large-amplitude fluctuation have been proposed, such as a supercurvature perturbation in an open universe [91], a bounce prior to inflation [102], a non-vacuum initial state [98], topological defects [90, 70], or deviations from slow-roll inflation [94, 92], it is largely taken as an ad hoc addition to the inflationary landscape. The scale dependence of the asymmetry often requires an additional elaboration to the theory, either in the form of scale-dependent non-Gaussianity [98, 84, 90, 70, 101, 103, 102] or isocurvature fluctuations [84, 92, 97].

Dropping the idea of a special large superhorizon fluctuation and instead starting with scale-dependent non-Gaussianity changes this picture. Importantly, the constraints on the amplitude of the non-Gaussianity on large scales is rather weak; the Planck bound of $f_{\text{NL}}^{\text{local}} = 2.5 \pm 5.7$ at 68% CL [121] assumes a scale-invariant bispectrum. To generate the observed power asymmetry, we only need non-Gaussianity on the scales that are asymmetric. For example, if we consider CMB multipoles up to $\ell \approx 100$, the WMAP 5 year data give $f_{\text{NL}} = -100 \pm 100$ [122]. (This constraint has not appreciably changed since WMAP5: see also Figure 11 of [121] for the most recent plot of Planck’s $\ell_{\text{max}}$-dependent constraints.) Many arbitrary choices for $\ell_{\text{max}}$ in the range $[40, 600]$ (and different binning schemes to study the scale dependence) can be found in the literature for computing the large-scale power asymmetry amplitude and its significance [54, 57, 56, 81]. We will use $\ell_{\text{max}} = 100$ as our fiducial value to compute the asymmetry for our numerical tests and statistical analysis later. This also allows us to use the WMAP5 large-scale bispectrum constraints computed with $\ell_{\text{max}} = 100$.

The weaker constraint on $f_{\text{NL}}$ implies that it is possible to generate the observed
asymmetry without enhancing the amplitude of superhorizon fluctuations, as was noted by [85], which considered the power-spectrum modulations generated by anisotropic bispectra. We note, however, that the dipolar bispectrum considered there does not generate a power asymmetry because the $\vec{k} \rightarrow -\vec{k}$ symmetry of the power spectrum forbids any modulation to the power spectrum from bispectra that depend on odd powers of the angle between the long and short modes; this point was clarified in [123]. In this work, we will focus on isotropic, scale-dependent non-Gaussianity (although our framework can easily be extended to include fundamentally anisotropic models), and we will show that this is sufficient to generate the observed asymmetry without enhancing the amplitude of superhorizon fluctuations.

The plan of this section is as follows. Next, we use the usual local ansatz to demonstrate the validity of our analytic calculations of the statistical anisotropies expected in non-Gaussian scenarios and to illustrate several of the key conceptual points relating non-Gaussianity and anisotropies. We also show how our framework encompasses the EKC mechanism and demonstrate that exotic superhorizon perturbations are not required to generate the observed power asymmetry. In section 3.2.3, we test our analytic calculations for monopole and dipole power modulations using numerical realizations of CMB maps. We also introduce and discuss our parameter estimation and model comparison methods. In section 3.2.4, we investigate non-Gaussianity beyond the local ansatz and in particular, consider a representative model that generates features that are fully consistent with constraints on the isotropic power spectrum and bispectrum. We discuss and summarize important aspects of our work and conclude in section 3.2.5. The appendixes contain technical details.

3.2.2 Illustrating the connection between non-Gaussianity and isotropy

We assume that at some early time (after reheating but prior to the release of the cosmic microwave background radiation) a large volume of the Universe ($V_L$) contains adiabatic fluctuations described by isotropic but non-Gaussian statistics. To compare predictions of this model with observations, we are interested in the statistics of the fluctuations in smaller volumes, $V_S \ll V_L$, that correspond in size to our presently observable Hubble volume. We will first consider the usual local model with constant $f_{NL}$ for simplicity; we will present more general results in a later section.

3.2.2.1 The local model

Suppose the Bardeen potential $\Phi$ is a non-Gaussian field described by the local model:

$$\Phi(x) = \phi(x) + f_{NL} \left( \phi(x)^2 - \langle \phi(x)^2 \rangle \right),$$

where $\phi(x)$ is a Gaussian random field. When the large volume is only weakly non-Gaussian, the power spectrum observed in our sky, $P_{\phi,S}(k, x)$, will be related to the
mean power spectrum in the large volume, $P_φ(k)$, by

$$P_{Φ,S}(k, x) = P_φ(k) \left[ 1 + 4f_{NL} \int \frac{d^3 k_ℓ}{(2π)^3} \frac{\phi(k_ℓ)e^{ik_ℓ \cdot x}}{\phi(k_ℓ)} \right], \quad (3.9)$$

where the radial integration for $k_ℓ$ is confined to $|k_ℓ| < π/r_{ cmb}$ when the CMB spectrum is the quantity of interest. Our conventions for the power spectrum are stated in Appendix A.1, and Appendix A.2 provides the derivation of this equation. Any particular model for generating the fluctuations in volume $V_L$ should provide a well-motivated lower bound on the $k_ℓ$ integral (e.g., from the duration of inflation). However, not all shifts to local statistics are sensitive to the full range of the integral; for local-type non-Gaussianity, as we will show later, only the monopole receives contributions from all super-Hubble modes.

The power spectrum $P_φ$ and amplitude of non-Gaussianity, $f_{NL}$, appearing on the right-hand side of Eq.(3.9) are those defined in the large volume. However, looking ahead to the result for the dipole modulation from the local ansatz, Eq.(3.22), $P_φ$ and $f_{NL}$ will be shifted to the observed values. In particular, the local non-Gaussianity that generates the observed power asymmetry is only the portion that violates Maldacena’s consistency relation [32] and is zero in single-clock inflation [120]. Since the observed values are ultimately the relevant quantities in the analysis, we will not increase the complexity of the notation to distinguish the large and small volume parameters, except in the appendixes.

The field $ϕ(k_ℓ)$ appearing in the integral is no longer stochastic but consists of the particular realization of the field that makes up the background of a particular Hubble volume. The power spectrum in Eq.(3.9) can depend on position $x$ within $V_S$ because an individual realization (local value) of the fluctuations $ϕ(k_ℓ)$ can be non-zero. If we consider the average statistics in the large volume (equivalent to averaging over all regions of size $V_S$), then the term proportional to $f_{NL}$ in Eq.(3.9) above averages to zero since $⟨ϕ(k_ℓ)⟩_{V_L} = 0$. In that case we recover the isotropic power spectrum of $V_L$. Finally, keep in mind that Eq.(3.9) is still more general than our actual CMB sky: it provides the statistics from which to draw realizations of our observed modes. Any single sky realization will still be subject to the usual cosmic variance that affects the values of small $ℓ$ modes and that can generate a power asymmetry even for volumes where the term proportional to $f_{NL}$ is zero.

It is also interesting to consider a two-field extension of Eq.(3.8):

$$Φ(x) = φ(x) + σ(x) + f_{NLσ} \left( σ(x)^2 - ⟨σ(x)^2⟩ \right), \quad (3.10)$$

where both $φ(x)$ and $σ(x)$ are Gaussian random fields and are uncorrelated. In this
case, the power spectrum observed in our sky, $P_{\Phi,s}(k,x)$, is

$$P_{\Phi,S}(k,x) = P_{\Phi}(k) \left[ 1 + 4\xi f_{NL} \int \frac{d^3k}{(2\pi)^3} \sigma(k)e^{ik\cdot x} \right],$$

(3.11)

where $P_{\Phi}(k) = P_{\varphi}(k) + P_{\sigma}(k)$ is the mean power spectrum in the large volume and $\xi = P_{\Phi,\sigma}/P_{\Phi}$ is the fraction of power in the $\sigma$ field. We will only consider cases with weak non-Gaussianity ($f_{NL}^2 \sigma P_{\sigma} \ll 1$ and $P_{\Phi,\sigma} \approx P_{\sigma}$). The amplitude of the local-type bispectrum for this weakly non-Gaussian two-field model is given by $f_{NL} = \xi^2 f_{NL}^2 \sigma$.

Therefore, the inhomogeneous power spectrum in terms of the observed $f_{NL}$ and the fraction of power $\xi$ is

$$P_{\Phi,S}(k,x) = P_{\Phi}(k) \left[ 1 + 4\frac{f_{NL}}{\xi} \int \frac{d^3k}{(2\pi)^3} \sigma(k)e^{ik\cdot x} \right].$$

(3.12)

A scale-dependent power fraction $\xi(k)$ is a natural way to generate a scale-dependent power asymmetry, as can be seen from the inhomogeneous term in Eq.(3.11); if $\xi(k)$ decreases for large values of $k$, the modulation of the power spectrum will decrease as well. Such mixed-perturbation models also have other potentially observable consequences: $\xi$ affects the tensor-to-scalar ratio and contributes to large-scale stochasticity in the power spectra of galaxies [124].

### 3.2.2.2 Effect on the CMB sky

The imprint of the inhomogeneous power spectrum given by Eq.(3.9) on the CMB can be described in terms of a multipole expansion:

$$P_{\Phi}(k,\hat{n}) = P_{\varphi}(k) \left[ 1 + f_{NL} \sum_{LM} g_{LM} Y_{LM}(\hat{n}) \right],$$

(3.13)

where $Y_{LM}$ is a spherical harmonic, and $\hat{n}$ is the direction of observation on the last scattering surface. To find the expansion coefficients $g_{LM}$ we make use of the plane wave expansion

$$e^{ik\cdot x} = 4\pi \sum_{LM} i^L j_L(k\ell x)Y^*_{LM}(\hat{k}\ell)Y_{LM}(\hat{n}),$$

(3.14)

where $j_L$ is a spherical Bessel function of the first kind, and $x = x\hat{n}$ specifies the position of the observed fluctuation: for the CMB, $x = r_{cmb}$ is the comoving distance to the last scattering surface. Eq.(3.9) then implies that

$$g_{LM} = 16\pi i^L \int_{|k\ell|<\pi/x} \frac{d^3k}{(2\pi)^3} j_L(k\ell x)\phi(k\ell) Y^*_{LM}(\hat{k}\ell).$$

(3.15)
The quantity \( g_{LM} \) has a fixed value in any single volume \( V_S \), but when averaged over all small volumes in \( V_L \), \( \langle g_{LM} \rangle_{V_L} = 0 \). The expected covariance, on the other hand, is non-zero:

\[
\langle g_{LM} g_{L'M'}^* \rangle_{V_L} = 256\pi^2 (-1)^{L'+L+L'} \int_{|k|<\pi/x} \frac{d^3k}{(2\pi)^3} j_L(kx) j_{L'}(kx) P_\phi(k) Y_{LM}^*(k\hat{\ell}) Y_{L'M'}(k\hat{\ell});
\]

\[
= \frac{32}{\pi} \delta_{LL'} \delta_{MM'} \int_0^{\pi/x} \frac{dk}{k} k^2 \ell^2 j_L^2(kx) P_\phi(k);
\]

\[
= 64\pi \delta_{LL'} \delta_{MM'} \int_0^{\pi/x} \frac{dk}{k} k^2 \ell^2 (kx) P_\phi(k),
\]

(3.17)

where in the last line, we have defined the dimensionless power spectrum as \( P_\phi(k) = k^3 F_\phi(k)/(2\pi^2) \). We have again used the subscript \( V_L \) to indicate the ensemble average is over the values of \( g_{LM} \) in the full volume \( V_L \). Note that both the individual values of \( g_{LM} \) and their variance depend on the size of the small volume through the upper limit of integration in Eqs. (3.15) and (3.16). While the mean statistics in the large volume cannot depend on the scale for the small volume, the variance of the statistics observed in sub-volumes generically does. It is now straightforward to study the monopole and dipole contributions from non-Gaussian cosmic variance to the modulated component of the power spectrum in a small volume.

In the case of the two-field extension, using Eq.(3.12) in the definition of the modulation moments Eq.(3.13) gives

\[
\langle g_{LM} g_{L'M'}^* \rangle_{V_L} = \frac{64\pi \delta_{LL'} \delta_{MM'}}{\xi^2} \int_0^{\pi/x} \frac{dk}{k} k^2 \ell^2 (kx) P_\phi(k);
\]

\[
= \frac{1}{\xi} \langle g_{LM} g_{L'M'}^* \rangle_{V_L,\xi=1}.
\]

That is, for the same amplitude of non-Gaussianity observed in the \( \Phi \) field, the variance of the non-Gaussian modulations increases by a factor of \( 1/\xi \) compared to the single source (\( \xi = 1 \)) local model.

### 3.2.2.3 Monopole modulation \((L = 0)\)

The power-spectrum amplitude shift, \( A_0 \), in the parametrization of Eq.(3.13) is:

\[
P_\Phi(k) = P_\phi(k) [1 + A_0]
\]

\[
= P_\phi(k) \left[ 1 + f_{NL} \frac{g_{00}}{2\sqrt{\pi}} \right],
\]

(3.19)

where \( A_0 \) can be either positive or negative, but has a lower bound \( A_0 \geq -1 \). From the discussion above and Eq.(3.16), it is clear that \( g_{00} \) is Gaussian distributed with zero
mean and variance given by:

\[ \langle g_{00}^2 \rangle = 64\pi \int \frac{dk_\ell}{k_\ell} \left[ \frac{\sin(k_\ell x)}{k_\ell x} \right]^2 P_\phi(k_\ell). \]  

(3.20)

Therefore, the distribution of the monopole power modulation amplitude \( A_0 \) also follows a normal distribution, for small values of \( A_0 \), \(|A_0| \ll 1\), with zero mean and standard deviation:

\[ \sigma_{f_{NL}^{\text{mono}}} = \frac{1}{2\sqrt{\pi}} |f_{NL}| \langle g_{00}^2 \rangle^{1/2}. \]  

(3.21)

The expression for \( \langle g_{00}^2 \rangle \) is sensitive to the infrared limit of the integral. That is, all super-Hubble modes can contribute. Interesting aspects of cosmic variance arising from this term, including effects on the observed non-Gaussianity in small volumes have been subjects of investigation in [125, 126, 127, 128, 129, 130, 131]. In particular, the observed value of \( f_{NL} \) is, in general, shifted from the mean value in the large volume.

For a constant \( f_{NL} \), the effect of the monopole modulation is to change the power-spectrum amplitude on all scales and therefore is not observationally distinguishable from the “bare” value of the power-spectrum amplitude. For scale-dependent non-Gaussianity, there is a scale-dependent power modulation, which can generically be interpreted as shifting the spectral index in the small volume away from the mean value in the large volume. In cases where the amplitude of non-Gaussianity is small (and consistent with zero) at small scales (large \( \ell \)), the power-spectrum amplitude from those scales can be taken as \( P_\phi(k) \), and then one can look for monopole modulation at large scales for which the non-Gaussianity constraints are not as strong. The large-scale power suppression anomaly [132, 133] is exactly such a situation. We will return to this point in more detail in Section 3.2.4.

### 3.2.2.4 Dipole modulation \((L = 1)\)

The dipole modulation of the power spectrum in the parametrization of Eq.(3.13) is given by:

\[ P_\Phi(k, \hat{n}) = P_\phi(k) \left[ 1 + f_{NL} \sum_{M=-1,0,1} g_{1M} Y_{1M}(\hat{n}) \right]. \]  

(3.22)

Since we are interested in the dipole modulation of the observed power spectrum in the CMB sky, the above equation should be obtained from Eq.(3.9) by absorbing the (unobservable) monopole shift to the observed power spectrum. Then, on the right-hand side of Eq.(3.22), \( P_\phi(k) \) is the observed isotropic power spectrum and \( f_{NL} \) is the observed amplitude of local non-Gaussianity within our Hubble volume. See Eq. (A.17) and the discussion there for details. (Appendix A.3 contains the corresponding expression in terms of bipolar spherical harmonics.) The \( g_{1M} \) coefficients are Gaussian distributed
with zero mean and a variance

\[ \langle g_1M^* g_1M \rangle = 64\pi \int \frac{dk_\ell}{k_\ell} \left[ \frac{\sin(k_\ell x)}{(k_\ell x)^2} - \frac{\cos(k_\ell x)}{k_\ell x} \right]^2 P_\phi(k_\ell). \]

If we pick a direction \( \vec{d}_i \) in which to measure the dipole modulation \( A_i \) such that

\[ P_\Phi(k) = P_\phi(k) \left[ 1 + 2A_i \cos \theta \right], \quad (3.23) \]

where \( \cos \theta = \hat{d}_i \cdot \hat{n} \), then the contribution to the dipole from the non-Gaussianity is \( A_i^{NG} = \frac{1}{4} \sqrt{\frac{3}{\pi}} f_{NL} g_{10} \), which is normally distributed with mean zero and standard deviation:

\[ \sigma_{f_{NL}} = \frac{1}{4} \sqrt{\frac{3}{\pi}} |f_{NL}| \langle g_{10}^2 \rangle^{1/2}. \quad (3.24) \]

In the two-field model Eq.(3.10), using Eq.(3.18), the standard deviation gets modified:

\[ \sigma_{f_{NL}} = \frac{1}{\sqrt{\xi}} [\sigma_{f_{NL}}]_{\xi=1}. \quad (3.25) \]

This shows that for \( \xi < 1 \) (e.g. a mixed inflaton-curvaton model), it is easier to generate the hemispherical power asymmetry with a small value of \( f_{NL} \). However, there is a minimal value of \( \xi \) that can generate a power asymmetry of a given amplitude: the requirement of weak non-Gaussianity in the non-Gaussian field \( \sigma(x) \) in Eq.(3.12) demands that \( \xi \gtrsim A_i^{NG} \).

The above discussion of the distribution of the dipole asymmetry \( A_i \) assumes that we measure \( A_i \) in a fixed direction \( \vec{d}_i \). However, we have no a priori choice of direction \( \vec{d}_i \) in most situations. This is especially true when considering a power asymmetry that is generated by the random realization of superhorizon perturbations as opposed to a single exotic perturbation mode. Therefore, observations of dipole power modulations are necessarily reported using the amplitude of dipole modulation in the direction of the maximum modulation. To obtain that amplitude, we can consider any three orthonormal directions \( (d_1, d_2, d_3) \) on the CMB sky and measure the corresponding three dipole modulation amplitudes \( (A_1, A_2, A_3) \) in the three corresponding orthonormal directions for each sky. The amplitude of modulation for the CMB sky (simulated or observed) is then given by \( A = (A_1^2 + A_2^2 + A_3^2)^{1/2} \). Clearly then, \( A \) follows the \( \chi \) distribution with three degrees of freedom (also known as the Maxwell distribution). In the Section 3.2.3.1, we will directly test the distributions and parameters obtained in this section using numerical realizations of CMB maps.
3.2.2.5 Higher multipole modulations

The anisotropic modulation of the power spectrum is expected to continue to higher multipoles in the presence of non-Gaussianity. However, as shown in Figure 3.10, the expected value of the modulation gets smaller quickly for higher multipoles $L$. The corresponding expected variance of higher multipole modulations for Gaussian CMB maps, however, is only weakly dependent on $L$. See, for example, Figure 2(d) of [59]. There is no evidence for modulation at higher order multipoles in the Planck temperature anisotropies data (see Figure 34 of [56]). In statistical analysis of the kind we discuss later in Section 3.2.3.4, it may, nevertheless, be useful to add higher multipole modulations (at least the quadrupole $L = 2$) at large scales as it may provide increased evidence for or against non-Gaussian mode coupling. An approximate constraint on $f_{NL}g_{2M}$ may be obtained from the result for the Fourier space quadrupole modulation constraint in [67]. There are two possible scenarios: (i) the expected amplitude of the modulation is larger than that from cosmic variance in the Gaussian case; in this scenario, the lack of observation of such a modulation in the data will disfavor the non-Gaussian model that is used to explain the monopole and dipole modulations. (ii) The expected amplitude of the modulation is within the cosmic variance from the Gaussian case, in which case the data are not discriminatory for or against the...
non-Gaussian model.

3.2.2.6 Connection to prior work

Before proceeding, we pause to connect Eq.(3.22) to the EKC mechanism to better see how scale-dependent non-Gaussianity eliminates the need for enhanced perturbations on superhorizon scales. In the EKC mechanism, a single superhorizon perturbation mode in a spectator field during inflation is responsible for generating the asymmetry; the original proposal used a curvaton field [83], but later work extended the mechanism to any source of non-Gaussian curvature fluctuations [86, 105]. For example, consider a field $\sigma$ that generates a curvature perturbation $\zeta = N_\sigma \delta \sigma$. A superhorizon (SH) sinusoidal fluctuation in $\sigma$,

$$\sigma_{\text{SH}}(\vec{x}) = \sigma_L \cos \left( \vec{k}_L \cdot \vec{x} + \theta \right), \quad (3.26)$$

will generate a dipolar power asymmetry in the curvature power spectrum [105]:

$$P_\zeta(k,\vec{x}) = \bar{P}_\zeta(k) \left[ 1 - \frac{12}{5} f_{\text{NL}}(\vec{k}_L \cdot \vec{x}) N_\sigma \sigma_L \sin \theta \right] \quad (3.27)$$

to first order in $k_L x$. In terms of the Bardeen potential, $\Phi = (3/5) \zeta$, the power asymmetry is

$$P_\Phi(k,\vec{x}) = P_\phi(k) \left[ 1 - 4 f_{\text{NL}}(\vec{k}_L \cdot \vec{x}) \Phi_L \sin \theta \right], \quad (3.28)$$

where $\Phi_L \equiv (3/5)N_\sigma \sigma_L$. The Fourier transform of the superhorizon fluctuation given by Eq. (3.26) is

$$\Phi_{\text{SH}}(\vec{q}) = (2\pi)^3 \frac{\Phi_L}{2} \left[ e^{i\theta} \delta_D(\vec{k}_L - \vec{q}) + e^{-i\theta} \delta_D(\vec{k}_L + \vec{q}) \right]. \quad (3.29)$$

Inserting this expression into Eq.(3.15) for $g_{LM}$ implies that

$$g_{10} = -16\pi\Phi_L \frac{1}{6} \sqrt{\frac{3}{\pi}} k_L x \sin \theta + O(k_L^2 x^2). \quad (3.30)$$

With this expression for $g_{10}$, Eq. (3.22) matches Eq. (3.28) to first order in $k_L x$. Therefore, we see that the EKC mechanism can be described by our framework.

For a single superhorizon mode, Eq. (3.28) implies that the non-Gaussian contribution to the dipole is

$$A_{\text{NG}} = 2 |f_{\text{NL}}| \Delta \Phi, \quad (3.31)$$

where $\Delta \Phi = \Phi_L k_L x \sin \theta$ is the variation of $\Phi$ across the surface of last scatter. Since $k_L x < 1$, $\Phi_L > \Delta \Phi$. The RMS amplitude of $\Phi$ values given by extrapolating the
observed value of $P_\zeta$ to larger scales is

$$\Phi_{\text{rms}} \simeq \sqrt{P_\phi}, \quad (3.32)$$

where $P_\phi \simeq 8 \times 10^{-10}$ [20]. It follows that the amplitude of the superhorizon mode $\Phi_L$ is bounded from below as

$$\frac{\Phi_L}{\Phi_{\text{rms}}} \gtrsim \frac{A}{2|f_{\text{NL}}|\sqrt{P_\phi}}. \quad (3.33)$$

If $A = 0.06$ (and is entirely due to the non-Gaussianity) and $|f_{\text{NL}}| < 100$, then $\Phi_L/\Phi_{\text{rms}} > 10.6$, which implies that the superhorizon mode must be at least a 10σ fluctuation. This is why Eq.(3.26) was not originally considered to be part of the inflationary power spectrum, but rather a remnant of pre-inflationary inhomogeneity or a domain-wall-like feature in the curvaton field [83]. It was then necessary to consider the imprint this enhanced superhorizon mode would leave on large-scale temperature anisotropies in the CMB through the Grishchuk-Zel’dovich (GZ) effect [134]. Although the curvature perturbation generated by Eq.(3.26) does not generate an observable dipolar anisotropy in the CMB [135, 136, 137], it does contribute to the quadrupole and octupole moments, and observations of these multipoles severely constrain models that employ the EKC mechanism [136, 83].

However, if we relax our upper bound on $f_{\text{NL}}$ to 270 or 500, Eq.(3.33) indicates that a 4σ or 2σ fluctuation, respectively, could generate an asymmetry with $A = 0.06$. The odds of generating the observed asymmetry are also improved by accounting for the fact that there are three spatial dimensions, which provide three independent opportunities for a large-amplitude fluctuation. We will see in Section 3.2.3.4 that considering the combined contributions of several superhorizon modes and accounting for the red tilt of the primordial power spectrum further increases the probability of generating the observed asymmetry, to the point that the $p$-value for $A = 0.06$ increases to greater than 0.05 for $|f_{\text{NL}}| \gtrsim 300$. Thus, if the perturbations on large scales are sufficiently non-Gaussian, there is no need to invoke enhanced superhorizon perturbations to generate the observed power asymmetry.

In the absence of an enhancement of the superhorizon power spectrum, the variance of the quadrupole moments and octupole moments in the CMB will not be altered. Consequently, we do not expect significant constraints on such models from the GZ effect. We note though that the specific realization of modes outside our sub-volume will still source quadrupole and octupole anisotropies in the CMB. For realizations that generate a large power asymmetry, the GZ contribution to these anisotropies would likely be larger than expected from theoretical predictions of $C_2$ and $C_3$ and aligned with the power asymmetry. However, this effect may be difficult to disentangle from the monopole power modulation described in Section 3.2.2.3, and we leave a detailed analysis of this observational signature to future work.

Furthermore, the ratio $\Delta \Phi/\Phi_{\text{rms}}$ for a given value of $A$ and $f_{\text{NL}}$ can be significantly
reduced if we consider mixed Gaussian and non-Gaussian perturbations. Using the mixed perturbation scenario introduced in Eq. (3.10), (with $\xi = P_{\Phi,\sigma}/P_{\Phi}$),

$$\frac{\Phi_{L,\sigma}}{\Phi_{\text{rms},\sigma}} > \frac{A\sqrt{\xi}}{2|f_{\text{NL}}|\sqrt{P_\phi}} > \frac{A^{3/2}}{2|f_{\text{NL}}|\sqrt{P_\phi}}.$$  

(3.34)

In the last inequality, we employ the fact that $\xi > A$ is required for the non-Gaussianity in the $\sigma(x)$ field to be weak enough to make the $\mathcal{O}(f_{\text{NL}}^2 P_{\sigma})$ contribution to $P_{\Phi,\sigma}(k)$ negligible. In this case, $\Phi_{L}$ could be sourced by a 1$\sigma$ fluctuation in the $\sigma(x)$ field if $|f_{\text{NL}}| = 270$. The possibility of using a mixed curvaton-inflaton model to generate a scale-dependent asymmetry using a single large superhorizon perturbation was explored in [84].

### 3.2.3 Statistical Anisotropy in the CMB Power Spectrum

#### 3.2.3.1 Numerical tests

We now present numerical tests of our analytic expressions for the dipole power modulation in the case of local non-Gaussianity. We will work in the Sachs-Wolfe (SW) regime: we only consider

$$\frac{\Delta T}{T} = -\frac{\Phi}{3}.$$  

(3.35)

Therefore, for local non-Gaussianity, Eq.(3.8), the temperature fluctuation is given by:

$$\left|\frac{\Delta T}{T}\right|_{f_{\text{NL}}} = \left|\frac{\Delta T}{T}\right|_{\text{gaus}} - 3f_{\text{NL}} \left[\left|\frac{\Delta T}{T}\right|_{\text{gaus}}^2 - \left\langle \left|\frac{\Delta T}{T}\right|_{\text{gaus}}^2 \right\rangle \right].$$  

(3.36)

We generate 10000 simulated Gaussian SW CMB skies using

$$C_\ell = \frac{4\pi}{9} \int_0^\infty \frac{dk}{k} \mathcal{P}_\phi(k) j_\ell^2(kx)$$  

(3.37)

for $0 \leq \ell \leq 300$; the primordial power spectrum is given by

$$\mathcal{P}_\phi(k) = A_\phi \left( \frac{k}{k_0} \right)^{n_s-1}$$  

(3.38)

with $A_\phi = 7.94 \times 10^{-10}$ and $n_s = 0.965$ (from Planck TT,TE,TE+lowP column in Table 3. of [20]), and $k_0 = 0.05$ Mpc$^{-1}$ as the pivot scale.

Then it is easy to generate non-Gaussian Sachs-Wolfe CMB temperature maps using Eq.(3.36) for a constant $f_{\text{NL}}$. Unlike most CMB analyses, we will keep the dipole variance term $C_1$. A non-zero $C_1$ is used to model the dipolar anisotropy in density fluctuations on the scale of the observable universe (from the perspective of the large volume $V_L$). However, note that the $C_1$ we use is not what we would measure for the
CMB dipole, even if we assume that the dominant contribution to the measurement of the dipole from our local motion [63] has been subtracted out. This is because, for adiabatic fluctuations, the leading-order contribution to the observed CMB dipole from superhorizon perturbations exactly cancels the Doppler dipole generated by the superhorizon perturbations [136, 137].

It is convenient to set the monopole $C_0$ to zero for the purpose of studying dipole modulations; otherwise, the cosmic variance power asymmetry (i.e. the contribution that is not due to local non-Gaussianity) will be different for the weakly non-Gaussian realization compared to the Gaussian realization from which it is generated. Therefore, in this section, we use numerical realizations with non-zero $C_0$ values only when testing the monopole modulation formula. The expression for $C_0$ is infrared divergent, so we assume an infrared cutoff $k_{\text{min}}$; the same cutoff scale is used to compute the expected amount of monopole power modulation $A_0$. Numerically,

$$C_0 = \frac{4\pi}{9} \int_{k_{\text{min}}}^{\infty} \frac{dk}{k} P_\phi(k) \left[ \frac{\sin(kx)}{kx} \right]^2. \quad (3.39)$$

The $k_{\text{min}}$ cutoff can be related to the number of superhorizon e-folds of inflation (if interpreted as such) as

$$N_{\text{extra}} = \ln \left[ \frac{(\pi/r_{\text{cmb}})}{k_{\text{min}}} \right].$$

The above integral gets most of its contribution from $k < \pi/r_{\text{cmb}}$, and therefore can be well approximated by [127]:

$$C_0 \approx \frac{4\pi}{9} A_\phi \left( \frac{\pi}{k_0 r_{\text{cmb}}} \right)^{n_s-1} \left[ 1 - e^{-\frac{(n_s-1)N_{\text{extra}}}{n_s-1}} \right] \quad (3.40)$$

for $n_s \neq 1$, and $(4\pi/9)A_\phi N_{\text{extra}}$ for $n_s = 1$. Additional details about our numerical results can be found in Appendix A.4.

### 3.2.3.2 Monopole modulation ($L = 0$)

The normally distributed monopole shift amplitude $A_0$ for a local non-Gaussian model is given by Eq.(3.21), and can be written in terms of $C_0$ using Eq.(3.39) as:

$$\sigma_{f_{\text{NL}}}^{\text{mono}} \approx 6|f_{\text{NL}}| \sqrt{\frac{C_0}{\pi}}. \quad (3.41)$$

The necessary infrared cutoff has already been set by the value of $N_{\text{extra}}$ in Eq.(3.40) to compute $C_0$. Although it is not possible to observe monopole modulations for a constant local $f_{\text{NL}}$, we can test the expected modulations assuming a value of $N_{\text{extra}}$. 
The probability distribution of the shift for any $f_{NL} \neq 0$ is

$$p_N(A_0, \sigma_{\text{mono}}) = \frac{1}{\sigma_{\text{mono}} \sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{A_0}{\sigma_{\text{mono}}} \right)^2 \right], \quad (3.42)$$

where the variance has contributions from the Gaussian realization and the non-Gaussian coupling to the realization of long wavelength modes:

$$(\sigma_{\text{mono}})^2 = (\sigma_{f_{NL}})^2 + (\sigma_G)^2.$$ 

For our numerical tests, $\sigma_{G}^{\text{mono}}$ is the variance of $A_0$ measured in Gaussian CMB maps. The quantity we measure for $A_0$ from each realization of CMB maps is:

$$A_0 = \frac{1}{\sum_{\ell=2}^{\ell_{\text{max}}} (2\ell + 1)} \sum_{\ell=2}^{\ell_{\text{max}}} (2\ell + 1) \left[ \frac{C_\ell - C_\ell^{\text{true}}}{C_\ell^{\text{true}}} \right], \quad (3.43)$$

where $C_\ell^{\text{true}}$ are the input angular power-spectrum values used to obtain the set of numerical CMB maps, and $C_\ell$ is the angular power spectrum of a particular realization of that set of CMB maps, and $\ell_{\text{max}} = 100$. In Figure 3.11, we plot the distribution of $A_0$, using Eq.(3.42), for $f_{NL} = 0, 50, 100$, along with the distribution obtained from the numerically generated Sachs-Wolfe CMB maps.

### 3.2.3.3 Dipole modulation ($L = 1$)

A dipole modulation of the power spectrum defined as in Eq.(3.23) generates a hemispherical power asymmetry with the same amplitude $2A_i$. Therefore, we will look at the quantity:

$$A_i = \frac{1}{\sum_{\ell=2}^{\ell_{\text{max}}} (2\ell + 1)} \sum_{\ell=2}^{\ell_{\text{max}}} (2\ell + 1) \frac{\Delta C_\ell}{2C_\ell}, \quad (3.44)$$

with $\ell_{\text{max}} = 100$ and $\Delta C_\ell = C_\ell^+ - C_\ell^-$, where $+$ and $-$ refer to two hemispheres in some direction $d_i$. We will consider $A_i$s in three orthonormal directions $d_1, d_2, d_3$ on the sky. Each $A_i$ in a particular direction $d_i$ is normally distributed with zero mean. The variance $\sigma_G^2$ can be measured from the numerical realizations of Gaussian Sachs-Wolfe CMB maps and depends on the CMB multipoles used in Eq.(3.44) and the value of the $C_\ell$s. For non-Gaussian maps, the distribution of the $A_i$s have an increased variance given by: $\sigma^2 = \sigma_{f_{NL}}^2 + \sigma_G^2$, where $\sigma_{f_{NL}}$ is given by Eq.(3.24). The power asymmetry dipole amplitude for each CMB sky is then $A = (A_1^2 + A_2^2 + A_3^2)^{\frac{1}{2}}$. The probability distribution function (pdf) of $A$ is the $\chi$ distribution (or the Maxwell distribution):

$$p_{\chi}(A, \sigma) = \sqrt{\frac{2}{\pi}} \frac{A^2}{\sigma^3} \exp \left[ -\frac{A^2}{2\sigma^2} \right], \quad (3.45)$$
Figure 3.11: Test of the monopole modulation formula Eq.(3.42) for the local non-Gaussian model with $f_{NL}$ specified in the figure. The dotted blue line for the $f_{NL} = 0$ (Gaussian) model is the best-fit normal distribution to the distribution obtained from numerically generated Sachs-Wolfe CMB maps, while the other two curves (dashed green, $f_{NL} = 50$ and solid red, $f_{NL} = 100$) are obtained using Eq.(3.42). The value of the CMB monopole $C_0$ is set using Eq.(3.40) with $N_{extra} = 50$. The measurement of $A_0$ is bounded below i.e. $A_0 \geq -1$. The normal distribution is an excellent fit for small modulations i.e. $|A_0| \ll 1$, whereas the distribution is positively skewed for larger values of $|A_0|$.

where $\sigma = \sqrt{\sigma_{f_{NL}}^2 + \sigma_G^2}$. Figure 3.12 shows that the distribution of asymmetry amplitudes obtained from the CMB realizations agree extremely well with the $\chi$ distribution given above. Note that only $\sigma_G$ is measured from the numerical maps; $\sigma_{f_{NL}}$ is directly computed for a value of $f_{NL}$ using Eq.(3.24).

3.2.3.4 Statistical analysis

In this section, we present examples of how we can perform a statistical analysis using the results from the previous section for the distributions of power modulation on the Sachs-Wolfe CMB sky. While direct comparison of the amplitudes obtained in our Sachs-Wolfe CMB realizations with the reported values of power asymmetry $A$ is not possible, we can make a connection between our simpler case and the asymmetry in the observed CMB sky by using the $p$-value of the asymmetry. For a given measurement of $A$ and the normalized pdf for $A$, $p(A)$ [which in our model depends on $f_{NL}$, see Eq.(3.45)], the $p$-value is simply given by

$$p - \text{value}(A) = \int_A^{\infty} p(A')dA';$$
Figure 3.12: Top: The distribution of power asymmetry $A_i$ (in a particular direction $d_i$) measured in 10000 simulated CMB skies as described in the text. From each simulated map, three $A_i$ values are generated in three orthonormal directions in the sky. The dotted blue line for the Gaussian CMB maps is the best-fit normal distribution curve; this gives us the Gaussian cosmic variance standard deviation $\sigma_G$. The curves (dashed green and solid red) for the non-Gaussian models are normal distributions with zero mean and variance given by $\sigma^2 = \sigma_G^2 + \sigma_{f_{NL}}^2$, where $\sigma_{f_{NL}}$ is computed using Eq.(3.24). Bottom: The distribution of the amplitude of power asymmetry $A$ from the simulated Gaussian and non-Gaussian CMB maps. The lines are the computed Maxwell distributions for corresponding $\sigma$ values and match the distributions obtained in numerical realizations very well. Note that the figures above were generated for the single source local model. If we used the mixed inflaton-curvaton two-field extension with the curvaton power fraction $\xi < 1$, we would get the above distributions for smaller values of $f_{NL}$. For example, for $\xi = 0.25$, the distributions shown for $f_{NL} = 500$ above would be generated by a smaller $f_{NL} = \sqrt{0.25} \times 500 = 250$. 
Figure 3.13: The posterior probability distribution of $|f_{NL}|$ values for different observed amplitudes $A$ of power asymmetry. Only large-scale CMB modes $\ell \leq 100$ are used to compute the cosmic variance pdf for $A$. $A = 0.055$ corresponds to a $p$-value of 0.001 in our $f_{NL} = 0$ numerical maps. This is about 3.3σ, approximately equal to some of the reported significance of the power asymmetry anomaly [59, 138]. Although our formula for the expected asymmetry becomes less accurate for larger values of $f_{NL}$ we have checked that it approximates the numerical results quite well even for $f_{NL} = 2000$. Therefore, the shape of the posterior distributions obtained above in the $f_{NL}$ window shown will not be affected by the inaccuracy of the formula at larger $f_{NL}$ values. However, the change in the distribution for larger $f_{NL}$ can change the normalization.

i.e. it gives the probability that the observed value of the asymmetry amplitude is greater than some threshold value $A$. We find that an asymmetry amplitude of $A = 0.055$ is approximately 3.3σ, i.e. a $p$-value of 0.001 with respect to the distribution of $A$ obtained in our Gaussian Sachs-Wolfe CMB maps. This is approximately equal to some of the more recent reports for the significance of the hemispherical power asymmetry [59, 138]. Therefore, we will use $A = 0.055$ as the value of the asymmetry when making connections with the observations of the anomaly.

When we have a measurement of the power asymmetry amplitude $A$, we can write the likelihood for $f_{NL}$ as $L(f_{NL}|A) = p_x(A, \sigma)$, whose expression is given in Eq.(3.45). From this likelihood, we can infer the posterior distribution for $f_{NL}$ given a measurement of $A$. We can interpret the statistics in different ways:

- We can use any power asymmetry as a signal of local non-Gaussianity. Using only the large-scale CMB multipoles ($\ell \leq 100$), for a given value of $A$, we can obtain the posterior distribution for $|f_{NL}|$ (averaged) for the corresponding range of scales. In Figure 3.13, we plot the $f_{NL}$ posterior for a few values of the asymmetry $A$, assuming a uniform prior on $|f_{NL}|$. 


We can combine the large-scale bispectrum constraints on $f_{NL}$ with the constraints from the power asymmetry $A$. For this, we use a rough estimate of $f_{NL}$ for $\ell \lesssim 100$ of $f_{NL} = -100 \pm 100(1\sigma)$ (estimated from Figure 2 of [122]). We assume that the $f_{NL}$ posterior from WMAP is a normal distribution. However, since the power asymmetry is only sensitive to the magnitude of $f_{NL}$ and not the sign, we use the folded normal distribution (given $f_{NL}$ measured $= \mu \pm \sigma$):

$$p_{\text{fold}}(f_{NL}) = \begin{cases} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{f_{NL} - \mu}{\sigma} \right)^2} & f_{NL} \geq 0 \\ e^{-\frac{1}{2} \left( \frac{f_{NL} + \mu}{\sigma} \right)^2} & f_{NL} < 0. \end{cases}$$

Then, we multiply the above pdf with $L(f_{NL}|A)$ to get the combined likelihood from which we can get the posterior for $f_{NL}$ after normalizing. We show the combined posterior distribution of $f_{NL}$ for a few power asymmetry amplitudes $A$ in Figure 3.14.

Although the power asymmetry data alone ($A \approx 0.055$) prefers $|f_{NL}| \approx 500$ as the most likely value (or $|f_{NL}| \approx 200$ when the bispectrum constraints are also applied), the probability of an asymmetry increases whenever $f_{NL} \neq 0$. In Figure 3.15 we quantify how the probability of an observed dipole modulation changes as $|f_{NL}|$ increases. For example, the $p$-value of 0.001 for $A = 0.055$ changes to 0.046 for $|f_{NL}| = 265$ (which is within the $2\sigma$ window of the large-scale bispectrum constraint). In other words, even an amplitude of non-Gaussianity well below $|f_{NL}| \approx 500$ renders the observed asymmetry less “anomalous.”

### 3.2.3.5 Bayesian evidence

The previous section demonstrated that the power asymmetry data can be used to constrain non-Gaussian models, and that the amplitude of the observed asymmetry is less “anomalous” when non-Gaussianity is included. However, we also need to ask whether the data are such that the non-Gaussian model is preferred over the Gaussian.

To compare the posterior odds for different models $M_i$, given the data $\tilde{y}$, we compute the Bayes factor

$$B_{12} = \frac{p(\tilde{y}|M_1)}{p(\tilde{y}|M_2)},$$

where the factors in the numerator and denominator are the model likelihoods for models 1 and 2 respectively. In the simplest comparison, we take $f_{NL}$ as the only parameter of the models. The data we consider include the measured amplitude of the power asymmetry and the CMB constraint on the amplitude of the local bispectrum.
Figure 3.14: The posterior probability distribution of $|f_{NL}|$ after combining the bispectrum constraints at large scales ($\ell \lesssim 100$, $f_{NL} = -100 \pm 100$) with the power asymmetry constraint for the given value of $A$.

Figure 3.15: The $p$-value for different values of asymmetry amplitudes $A$ (i.e. the probability of obtaining an asymmetry amplitude equal to or greater than $A$) in a local non-Gaussian model as a function of the value of $|f_{NL}|$. For $A = 0.055$, we see that the significance goes below $3\sigma$ around $|f_{NL}| \approx 100$. For the two-field extension of the local model (with a curvaton power fraction $\xi$) described in the text, the x-axis should be labeled $|f_{NL}|/\sqrt{\xi}$. Then, for $\xi < 1$, a smaller bispectrum amplitude $|f_{NL}|$ (by a factor of $\sqrt{\xi}$ than labeled in the figure above) is required to achieve a $p$–value shown in the figure above.
on large angular scales. For an introduction to Bayesian statistical methods applied to cosmology, see for example [139].

The non-Gaussian model reduces to the isotropic Gaussian model for \( f_{NL} = 0 \) (while the probability of the power asymmetry remains non-zero). In that case the evaluation of the Bayes factor can be simplified and a direct Bayesian model comparison can be done using the Savage-Dickey density ratio (SDDR) [140, 139]. The SDDR is given by,

\[
B_{01} = \left. \frac{p(\theta_i | \vec{y}, M_1)}{\pi(\theta_i | M_1)} \right|_{\theta_i = \theta_i^*}
\]

where \( M_1 \) is the more complex model (non-Gaussian in our case) that reduces to the simpler model \( M_0 \) (Gaussian) when the set of parameters \( \theta_i \) goes to \( \theta_i^* \) \( (f_{NL} \rightarrow f_{NL} = 0) \). Here, \( p(\theta_i | \vec{y}, M_1) \) is the posterior for \( f_{NL} \) (plotted in Figure 3.14) and \( \pi(\theta_i | M_1) \) represents the prior for the parameter in the complex model \( M_1 \). Our current case only has one parameter \( (f_{NL}) \) and one datum (the dipolar asymmetry \( A \)). While there may be other interesting possibilities to consider for the prior probability of \( f_{NL} \), we illustrate the calculation of \( B_{01} \) above using the constraint on the parameter \( f_{NL} \) from large-scale bispectrum measurements as reported by the WMAP and Planck missions as the prior.

In Table 3.2, we list SDDR for a few values of \( A_{\text{obs}} \). For the prior, we have used the folded normal distribution for \( f_{NL} = -100 \pm 100 \) which is a rough estimate of \( f_{NL} \) for the largest scale i.e. up to \( \ell = 100 \) from [122]. Note that the only value from the prior pdf that is used to compute the SDDR is \( f_{NL} = 0 \), so the above consideration from large-scale \( f_{NL} \) constraints is the same as using a uniform prior for \( |f_{NL}| \) in the range \( (0, \frac{1}{\pi(f_{NL} = 0)}) \approx (0, 207) \). If the prior range is expanded, then the magnitude of \( B_{01} \) increases thereby reducing the evidence for non-zero \( f_{NL} \). For \( A_{\text{obs}} = 0.055 \) (whose \( p \)-value roughly corresponds to observed \( A \)), the strength of evidence for a non-zero \( f_{NL} \) is between weak and moderate in the empirical (Jefferys’) scale [141] quoted, for example, in Table 1 of [139].

The results in Table I show that the data we have used, at least in this simple analysis, show no more than a weak preference for the non-Gaussian model. A more thorough analysis is unlikely to change this conclusion very much: in [142], the authors use earlier studies of the power asymmetries in the WMAP data to put the best possible Bayesian evidence of \( \ln B_{01} \approx -2.16 \) corresponding to odds \( (\leq 9 : 1, \text{ weak support}) \). The method to compute the maximum possible Bayesian evidence is based on Bayesian calibrated \( p \)-values [143]. The \( p \)-value used from the data analysis of the 3-year WMAP maps was \( p = 0.01 \) [144]. If one instead used \( p = 0.001 \), which is approximately the level of significance from various more recent analyses of Planck and WMAP temperature anisotropy maps, the best possible Bayesian evidence in favor of the \( A \neq 0 \) anisotropic model becomes \( \ln B_{01} \approx -4.0 \) corresponding to the odds \( (\leq 50 : 1) \). A \( p \)-value of \( 0.0003 \) (about \( 3.6\sigma \)) is necessary to obtain a best possible \( \ln B_{01} \approx -5.0 \), which implies
### Table 3.2: Savage-Dickey Density Ratio (SDDR) for different observed values of dipole power modulations at large scales. The $p-$values listed above are computed with respect to the distribution of $A$ values from Gaussian CMB maps. A reasonable value to compare to various reports of the observed hemispherical power asymmetry is a $p-$value of 0.001 ($\approx 3.3\sigma$) i.e. $A = 0.055$. Values of $|\ln B_{01}| = 5.0, 2.5$ suggest strong and moderate evidence, respectively [139]. In our convention above, a negative value for the logarithm of the Bayes factor means the evidence is in favor of the more complex non-Gaussian model $\mathcal{M}_1$.

<table>
<thead>
<tr>
<th>$A_{\text{obs}}$</th>
<th>$p$-value</th>
<th>SDDR ($B_{01}$)</th>
<th>$\ln B_{01}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.02</td>
<td>0.5511</td>
<td>1.1362</td>
<td>0.128</td>
</tr>
<tr>
<td>0.04</td>
<td>0.0381</td>
<td>0.6174</td>
<td>-0.482</td>
</tr>
<tr>
<td>0.05</td>
<td>0.0043</td>
<td>0.3211</td>
<td>-1.136</td>
</tr>
<tr>
<td>0.055</td>
<td>0.001</td>
<td>0.2012</td>
<td>-1.603</td>
</tr>
<tr>
<td>0.06</td>
<td>0.0003</td>
<td>0.1123</td>
<td>-2.186</td>
</tr>
</tbody>
</table>

3.2.4 Beyond the local ansatz

The previous section discussed in detail the effect of local-type non-Gaussianity with constant $f_{\text{NL}}$ that couples a gradient (induced by superhorizon modes) across the CMB sky to the observable modes. We demonstrated that a dipolar asymmetry is expected in models with local-type non-Gaussianity. Of course, local non-Gaussianity as the source of the asymmetry is only compatible with the data if we restrict ourselves to the largest scales. Both the amplitude of the asymmetry and the amplitude of non-Gaussianity must sharply decrease on smaller scales. The question then is whether there is a different model of non-Gaussianity that is consistent with all observational constraints and generates the observed asymmetry in detail. If so, does the current data favor this model over the isotropic, Gaussian assumption? Could future data ever favor such a model?

To address some of these questions, we will first demonstrate that scale-dependent modulations are a generic feature of non-Gaussian models other than the local model. We will then construct a scenario that is more likely to be preferred by the data by considering scale-dependent local non-Gaussianity. Finally, we will provide examples of evaluating the Bayesian evidence for this scenario. Although a model of non-Gaussianity beyond the local ansatz may add more parameters, if the model has other consequences in the data we might hope to find more evidence for it. This is particularly true if the model has measurable effects on smaller scales, where the usual cosmic variance for Gaussian models is smaller.
3.2.4.1 Power asymmetry from general bispectra

We can easily extend the inhomogeneous power spectrum calculation in the presence of local non-Gaussianity to other bispectrum shapes. For example, consider that a Fourier mode of the Bardeen potential is given by [131]:

$$\Phi(k) = \phi(k) + \frac{f_{\text{NL}}}{2} \int \frac{d^3q_1}{(2\pi)^3} \int d^3q_2 \phi(q_1)\phi(q_2) N_2(q_1, q_2, k)\delta_D(k - q_1 - q_2) + \ldots$$  \hspace{1cm} (3.48)

where as before $\phi$ is a Gaussian field. The kernel $N_2$ can be chosen to generate any desired bispectrum and the dots represent terms higher order in powers of $\phi$ (which generate tree-level $n$-point correlations). Considering only the generic quadratic term, the power spectrum in sub-volumes can be computed as in the case of the local bispectrum (see Appendix A.2), and we get $P_{\Phi,S}(k, x)$:

$$P_{\phi}(k) \left[ 1 + 2f_{\text{NL}} \int \frac{d^3k_\ell}{(2\pi)^3} \phi(k_\ell)N_2(k_\ell, -k, k)e^{ik_\ell \cdot x} \right].$$  \hspace{1cm} (3.49)

From the form of the above equation, one can see that a $k$-dependent power modulation is a feature of non-local non-Gaussianity i.e. the $k$ dependence of the kernel $N_2$ is carried by the modulated component of the power spectrum in the small volume. The kernels for local, equilateral and orthogonal bispectrum templates are [131]:

$$N_{2\text{local}} = 2;$$
$$N_{2\text{ortho}} = \frac{4k^2 - 2kk_\ell}{k^2};$$
$$N_{2\text{equil}} = \frac{2k^2_\ell}{k^2}.$$  \hspace{1cm} (3.50)

If one uses the kernel for equilateral- or orthogonal-type non-Gaussianities, then the monopole shifts are not infrared divergent. However, the magnitudes of modulation (both monopole and dipole) are smaller compared to the local case i.e. a very large amplitude of $f_{\text{NL}}^{\text{local}}$ or $f_{\text{NL}}^{\text{ortho}}$ is necessary for the effect to be interesting. For example, we plot the expected modulation amplitude for local-, equilateral- and orthogonal-type non-Gaussianities in Figure 3.16. In Figure 3.17, we illustrate that for local, orthogonal and equilateral bispectra, the power asymmetry is generated by perturbation modes that lie just outside the horizon. The quasi-single field model [145] may also be interesting to consider: it has a scale-independent bispectrum with a kernel that varies between the local and equilateral cases depending on the mass of an additional scalar field coupled to the inflaton.

In general then, if the power asymmetry is coming from mode coupling, the fact that the observed asymmetry falls off on small scales implies that shorter scales are
more weakly coupled to superhorizon modes than larger scales are. This is possible with either a scale-independent bispectrum (as the equilateral and orthogonal cases above demonstrate) or with a scale-dependent bispectrum.

### 3.2.4.2 Generating a scale-dependent power asymmetry

To match the observed scale dependence of the power asymmetry anomaly, the strength of coupling of subhorizon modes to the long wavelength background must be scale dependent. The relevant scale dependence in this context can be fully parametrized by introducing two bispectral indices that capture the scale dependence in our observable volume and a more general coupling strength to the long wavelength modes:

\[
P_{\Phi,S}(k,x) = P_{\phi}(k) \left[ 1 + 4f_{\text{NL}}(k_0) \left( \frac{k}{k_0} \right)^{n_f} \right. \\
\times \left. \int \frac{d^3k_\ell}{(2\pi)^3} \left( \frac{k_\ell}{k_0} \right)^{\alpha} \phi(k_\ell)e^{i(k_\ell \cdot x)} \right].
\]  

Here \( n_f < 0 \) turns off any power asymmetries on shorter scales. The parameter \( \alpha < 0 \) enhances the sensitivity of the model to infrared modes (as used in [85, 102]). In the case \( \alpha \leq -1 \), the dipole asymmetry would be infrared divergent in a universe with a scale-invariant or red-tilt power spectrum. Notice that scale-invariant bispectra always have \( n_f = -\alpha \), so any scale-invariant bispectrum that increases IR sensitivity also increases
the expected asymmetry on smaller scales. Finally, although we have used $f_{NL}$ to label
the coefficient above, a similar expression can be derived from higher-order correlation
functions (e.g., to capture the effects of $g_{NL}$ [104]). Previous discussions of the power
asymmetry from scale-dependent non-Gaussianity include [84, 86, 90, 101, 98, 103, 102].

In principle, additional data could eventually constrain all of the parameters intro-
duced above (or at least their values on subhorizon scales). However, here we will
consider only one additional measurement (the large-scale power suppression) and so
we will restrict our attention to the case with just one additional parameter. We take
a local-shape bispectrum with an amplitude that depends on the scale of the short
wavelength mode as

$$f_{NL}(k) = f_{NL}^0 \left( \frac{k}{k_0} \right)^{n_{fNL}}. \quad (3.52)$$

In terms of the parameters in Eq.(3.51), $n_f = n_{fNL}$ and $\alpha = 0$.

In addition to a scale-dependent power asymmetry and a scale-dependent bispectrum
amplitude, a scale-dependent local non-Gaussianity also generates a scale-dependent
modulation of the power-spectrum amplitude. This is the monopole power modulation
($L = 0$) discussed in Section 3.2.2.2 (a similar point was made in [101]). When local-type
non-Gaussianity has a scale-independent amplitude, the power spectrum amplitude is

Figure 3.17: The expected amplitude of the power asymmetry from superhorizon modes
($k_L < \pi/r_{cmb}$), as a function of the minimum wavenumber $k_{L,\text{min}}$ considered to compute $\langle g_{IM} \rangle^{0.5}$,
for large-amplitude local-, orthogonal- and equilateral-type non-Gaussianities. We can see that
most of the contribution is due to the modes with wavelengths almost the size of the observable
universe i.e. within one e-fold. We get similar behaviors for the monopole modulations (not
shown in the figure) except for the case of local non-Gaussianity. For local non-Gaussianity, the
monopole modulation gets contributions from arbitrarily small $k$ modes and $\langle g_{00} \rangle_{\text{local}}$ becomes
infrared divergent.
modulated similarly at all scales, thereby making the effect unobservable. However, the scale-dependent case is more interesting as it makes the power modulation scale dependent and therefore an observable effect. This can be easily seen by generalizing Eq.(3.19) for the case of \( f_{\text{NL}}(k) \):

\[
P_\Phi(k) = P_\phi(k) \left[ 1 + f_{\text{NL}}(k) \frac{g_{00}}{2\sqrt{\pi}} \right].
\]

(3.53)

As previously discussed, \( g_{00} \) is normally distributed with zero mean and the variance \( \langle g_{00}^2 \rangle \) requires a cutoff to limit contributions from arbitrarily large modes, which we have earlier parametrized as the number of superhorizon efolds of inflation. While the above formula is only valid for small modulations \( A_0 \), Figure 3.11 shows that the formula is quite accurate at least up to \(|A_0| \approx 0.3\). There is an additional subtlety because, in the presence of a monopole shift, the observed bispectrum on large scales will not be exactly of the form in Eq.(3.52). However, the difference is small for small \( A_0 \).

Consider a simple example of scale-dependent local non-Gaussianity given by

\[
\begin{align*}
    n_{f_{\text{NL}}} &= -0.64, \\
    k_0 &= 60(\pi/r_{\text{cmb}}), \\
    f_{\text{NL}}^0 &= 50.
\end{align*}
\]

(3.54)

In multipole space, one can approximate

\[
\begin{align*}
    f_{\text{NL}}^0 &= f_{\text{NL}}(\ell = 60) = 50 \\
    f_{\text{NL}}(\ell) &= 50(\ell/60)^{-0.64}.
\end{align*}
\]

(3.55)

The numbers above are chosen to facilitate comparison with the scale-dependent modulation model results in [82] (see Table I therein). In Figure 3.18, we plot \( f_{\text{NL}}, \sigma_{f_{\text{NL}}} \) (the expected 1\( \sigma \) amplitude of the power asymmetry in a particular direction due to \( f_{\text{NL}} \)), and \( \sigma_{f_{\text{NL}}^{\text{mono}}} \) (the expected 1\( \sigma \) shift in the amplitude of the power spectrum due to \( f_{\text{NL}} \)), as a function of the multipole number \( \ell \). The purpose of the figure is to illustrate how scale-dependent local non-Gaussianity can produce more than one signature in the CMB. Therefore, we have not included the effect of the Gaussian cosmic variance, which would add variance to both the power asymmetry and monopole modulation amplitudes; this becomes important when performing parameter estimation of \( f_{\text{NL}}^0 \) and \( n_{f_{\text{NL}}} \). While such a full parameter estimation analysis is beyond the scope of this work, we illustrate in the next section that, in a simplified context, adding the \( A_0 \) constraint can be useful in some situations.

Notice that the scale-dependent non-Gaussian model of Eq.(3.52) also generates
asymmetries in the spectral index:

$$\frac{d \ln P}{d \ln k} \bigg|_{k_0} = (n_s - 1) + n_fNL \frac{\Delta P(k_0, x)}{1 + \Delta P(k_0, x)},$$

(3.56)

where \((n_s - 1)\phi\) is the spectral index of the Gaussian field and

$$\Delta P(k_0, x) = 4f_{NL}^0 \int \frac{d^3k_\ell}{(2\pi)^3} \phi(k_\ell) e^{ik_\ell \cdot x}$$

(3.57)

is the super cosmic variance contribution. As before, \(\Delta P\) may be expanded in multipole moments. A spatial modulation of the spectral index was used in [71] as an alternative way to produce the power asymmetry. Here we see that such a modulation is a natural consequence of non-Gaussian scenarios that generate a scale-dependent power asymmetry.

### 3.2.4.3 Statistical analysis

The Planck 2013 analysis reported a power deficit at \(\ell \lesssim 40\) with a statistical significance of approximately \(2.5\sigma - 3\sigma\) [132], while the more recent Planck 2015 analysis [146] shows...
a slightly lower statistical significance to the deficit. Although that measurement alone
does not require a new model, the scale-dependent non-Gaussian scenario generically
generates a power modulation. Therefore, the data should be included in constraining
the model. If we include a measurement of the monopole modulation \( A_0 \) in addition to
the dipole power asymmetry amplitude \( A \), the combined likelihood for \( f_{NL} \) is:

\[
\mathcal{L}(f_{NL}|A, A_0) = p_x(A, \sigma) p_{\text{fold}}(A_0, \sigma^{\text{mono}}).
\]  

(3.58)

In Figure 3.19, we show examples in which, in addition to the power-asymmetry
amplitude \( A \), we also consider the monopole-modulation amplitude \( A_0 \). For simplicity,
we assume that the result of a scale-dependent non-Gaussianity produces some large
(statistically significant at \( \approx 3.3\sigma \), for example) power asymmetry at large scales \( \ell \leq 100 \)
and which becomes smaller in magnitude and therefore less significant at small scales.
In Figure 3.19, therefore, we always take \( A = 0.055 \) and we have also added the
estimate of \( f_{NL} \approx -100 \pm 100 \) at these scales. This simplification allows us to use
the Gaussian cosmic variance from our constant \( f_{NL} \) CMB realizations and simply
augment the analysis in the previous section with the added information from a possible
monopole modulation. This is because we expect the scale-dependent non-Gaussianity
to produce a monopole modulation \( A_0 \) at large scales, the variance of which depends on
the strength of non-Gaussianity at the scales used to obtain the monopole modulation
amplitude \( A_0 \). We will define the measured \( A_0 \) as in Eq.(3.43) \( \hat{A}_0 \) = \( \langle A \rangle \), i.e. a weighted average
over multipoles \( 2 < \ell \leq 100 \). Depending on the values of \( f_{NL}^0 \) and \( n_{fNL} \), a scale-
dependent non-Gaussianity may generate significant monopole modulation at higher
multipoles. But, we will restrict our analysis to a single value \( A_0 \) obtained from the
range \( 2 < \ell \leq 100 \). Thus, our simple analysis is only sensitive to the average amplitude
of non-Gaussianity at these scales and cannot constrain \( n_{fNL} \) for which we will need to
consider multiple bins of \( A_0 \) in the data.

For \( A_0 \) [in Eq.(3.43)] generated at large scales from a scale-dependent non-Gaussianity,
\( C_{\ell}^{\text{true}} \) is the angular power spectrum at large multipoles where the non-Gaussianity and
therefore the effect of superhorizon modes on the observed power spectrum is small. In
the case of the Planck satellite measurement of \( A_0 \) in [132], the best-fit angular power
spectrum, which is also dominated by the larger-\( \ell \) modes, is taken as the \( C_{\ell}^{\text{true}} \).

In Figure 3.19, we see that for \( A_0 = 0.02 \), which is about \( 1.5\sigma \) in the distribution of
\( A_0 \) values for our Gaussian CMB realizations, the posterior distribution shows more
support for the Gaussian model. However, for a larger value \( A_0 = 0.04 \) (about \( 2.9\sigma \)),
the addition of the \( A_0 \) data favors the non-Gaussian model for a superhorizon efolds
of \( N_{\text{extra}} = 50 \). For a larger number of superhorizon efolds \( N_{\text{extra}} = 100 \), the support
for the non-Gaussian model again decreases because of the increased variance in the
prediction from the non-Gaussian model itself. See Figure 3.20 for a plot of how the
evidence changes with \( N_{\text{extra}} \). (However, some care should be taken in extrapolating
Figure 3.19: The posterior distribution of $|f_{NL}|$ for $A_{\text{obs}} = 0.055$ and different assumed values of $A_{0,\text{obs}}$ and $N_{\text{extra}}$. Note that the use of $A_0$ changes the shape of the posterior distribution, and generates a peak at around $f_{NL} \approx 10$ for $A_0 = 0.04, N_{\text{extra}} = 50$ (blue dotted line) in addition to the peak at $|f_{NL}| \approx 200$ from the use of $A = 0.055$ and the bispectrum constraints (black dashed line). For a larger $N_{\text{extra}} = 100$ (green solid line), the peak shifts towards $|f_{NL}| = 0$ as the variance of the prediction of the non-Gaussian model increases; the evidence for the non-Gaussian model also decreases.

We have only shown examples of how using $A_0$ (i.e. power modulation from scale-dependent non-Gaussianity) can change the posterior probability (and hence the evidence) of scale-dependent non-Gaussianity in a simplified situation. A more involved analysis (i.e. including larger $\ell$ modes and measuring the scale dependence of $f_{NL}$ from bispectrum measurements with the latest data) is necessary to see how useful the inclusion of $A_0$, in addition to the power asymmetry amplitude, proves to be in the

<table>
<thead>
<tr>
<th>$A_{0,\text{obs}}$</th>
<th>$N_{\text{extra}}$</th>
<th>SDDR ($B_{01}$)</th>
<th>$\ln B_{01}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>-</td>
<td>0.2012</td>
<td>-1.603</td>
</tr>
<tr>
<td>0.02</td>
<td>50</td>
<td>1.107</td>
<td>0.102</td>
</tr>
<tr>
<td>0.04</td>
<td>50</td>
<td>0.0664</td>
<td>-2.712</td>
</tr>
<tr>
<td>0.04</td>
<td>100</td>
<td>0.1475</td>
<td>-1.914</td>
</tr>
</tbody>
</table>

Table 3.3: SDDR for $A = 0.055$ and different observed values of monopole power modulations $A_0$ at large scales. See the text for a discussion on why the evidence increases and decreases for different values of $A_0$ and $N_{\text{extra}}$. 

this result since our analytic expressions are not valid when the non-Gaussianity in the large volume is strong.) The SDDR values for these cases are shown in Table 3.3.
actual CMB data. In addition, we have chosen a simple model that may capture the features seen in the data but that may not be easily achievable from an inflationary point of view [147].

3.2.5 Summary and Conclusion

We have performed a systematic study of the power asymmetry expected in the CMB if the primordial perturbations are non-Gaussian and also exist, with no special features, on scales larger than we can observe. We have derived an expression to compute the expected deviations from isotropy of the observed two-point function due to mode coupling with our specific realization of superhorizon modes, which generate anisotropy across our Hubble volume. Although we have focused our analysis on local non-Gaussianity, our method is quite general for describing deviations from statistical isotropy in a finite sub-volume of an isotropic (but non-Gaussian) large volume.

Exploiting the fact that local non-Gaussianity naturally produces a power asymmetry, we have shown how the observed asymmetry can be used for parameter estimation of the non-Gaussian amplitude $f_{NL}$. We have also combined $f_{NL}$ constraints from bispectrum and power-asymmetry measurements to evaluate the Bayesian evidence for $f_{NL} \neq 0$. In our simple examples, we find that the observed CMB power asymmetry only provides weak evidence for non-Gaussianity on large scales.

Many previous works that propose mechanisms to generate the CMB power asymmetry have required a departure from the simplest extrapolation of our observed statistics to larger scales in addition to non-Gaussianity. The most popular new feature is a large-amplitude superhorizon fluctuation that has an origin beyond inflation. Generating the observed scale dependence of the power asymmetry requires an additional

Figure 3.20: The Bayesian evidence as a function of $N_{\text{extra}}$. 
elaboration, either in the form of scale-dependent non-Gaussianity or isocurvature fluctuations. However, our results show that scale-dependent non-Gaussianity of the local type is sufficient to generate a dipole power asymmetry at large scales without invoking exotic superhorizon fluctuations. A value of $f_{NL}$ consistent with the rather weak large-scale-only constraint (but well above the scale-independent bound) is enough to make the observed power asymmetry no longer anomalous. It is also worth noting that scale-dependent bispectra are only required for the local ansatz: scale-independent non-local bispectra can generate a scale-dependent power asymmetry. All that is required to qualitatively match the observed scale dependence of the asymmetry is that smaller scales couple more weakly than large scales do to near-Hubble scale modes.

For scale-dependent non-Gaussianity, we have demonstrated that a large-scale suppression or enhancement of the isotropic power spectrum is expected. Consequently, observations of a scale-dependent shift in the amplitude of the power spectrum can provide additional evidence in favor of non-Gaussianity. Although we have not found that the observed power deficit at large scales is sufficient to say the data strongly prefer the non-Gaussian model, we have also not performed a thorough analysis nor forecast the improvement possible with future data. Our parameter estimation and model comparison works were performed in a simplified setting to illustrate that a more careful look at the CMB data is necessary and useful.

Our results strongly suggest that to evaluate the observational evidence for models of the cosmological perturbations, departures from Gaussianity and from statistical isotropy in the data can and should be considered together. Furthermore, in weighing whether the large-scale CMB anomalies are evidence of something beyond slow-roll inflation lurking just outside the horizon, we should be sure to account for the possibility of non-Gaussian cosmic variance contributions to anisotropy, which do not require the introduction of additional physics.
Chapter 4

Effects of Primordial
Non-Gaussianity: II. LSS

4.1 Introduction

4.1.1 Background and motivation

Recently, the Planck satellite mission reported tight constraints on primordial non-Gaussianity from measurements of the Cosmic Microwave Background (CMB) [148, 121]. They find the amplitude of the local, equilateral and orthogonal type bispectrum to be consistent with zero. While these results show that the primordial fluctuations were remarkably Gaussian, they still leave room for interesting signatures of primordial physics to be found in statistics beyond the power spectrum: minimal, single field models predict non-Gaussianity that is about two orders of magnitude below these constraints. Non-Gaussianity is such an informative tool for inflationary physics that it is crucial to push observational bounds as far as possible. A number of forecasts have shown that the use of clustering data from future large scale structure surveys can constrain $f_{\text{NL}}^\text{local}$ with $\sigma(f_{\text{NL}}) \approx (1 - 10)$ [149, 150, 151, 152, 153, 154]. In addition, the galaxy bispectrum is a promising way to probe many shapes to the $f_{\text{NL}} \sim \mathcal{O}(1 - 10)$ level [155, 156, 157]. The constraints from large scale structure (LSS) —whether consistent with the CMB measurements or in tension — will provide interesting and useful complementary results.

In the inflationary scenario, non-Gaussian signatures in the primordial density fluctuations depend on the details of interactions of the inflaton or other fields relevant for generating the perturbations; therefore, any detection (or the lack) of primordial non-Gaussianity tells us about the dynamics of those fields. Furthermore, if non-Gaussianity is detected, we expect to see patterns in the correlations that are consistent with perturbation theory (in contrast to apparently independent statistics at each order, for example). The simplest such pattern is in the relative amplitudes of higher order statistics, or, how the amplitudes of the correlation functions of the gravitational
potential, $\langle \Phi^n(x) \rangle$, scale with order $n$. The relative scaling of higher moments falls into a fairly narrow range of behaviors in inflationary models [31].

Thanks to the non-linearity of structure formation, the statistics of objects in the late universe contain information about the entire series of higher order moments of the initial density fluctuations. This is especially straightforward to see in number counts of galaxy clusters [158]. While information about higher moments is clearly non-trivial to extract when non-Gaussianity is weak, some constraints with current data are already affected by assumptions about higher order moments. For example, [159] reports constraints on primordial non-Gaussianity using a sample of 237 X-ray clusters from the ROSAT All-Sky Survey [160]. Their analysis clearly suggests that it is important to take the scaling of moments into account when deriving constraints on $f_{\text{NL}}$ from cluster number counts. The results presented in this section will be useful for further studies in the same direction. Tighter constraints could likely be achieved with current data by combining the X-ray clusters with SZ detected clusters that have been separately used to constrain non-Gaussianity [161, 162, 163]. In addition, the eROSITA survey [153] and third generation surveys detecting clusters via the Sunyaev-Zel’dovich effect are expected to provide much larger samples of clusters in the next few years. It will be interesting to revisit the non-Gaussian analysis when that data becomes available. Other large scale structure statistics, including the power spectrum and bispectrum of galaxies, also contain contributions from higher order primordial statistics. Constraints on non-Gaussianity from those observables are still being developed, and will ultimately be very powerful (constraints obtained to date include [164, 45, 165, 166, 46, 167, 168, 169, 48]). To make full use of them it is important to understand the signatures of the full spectrum of non-Gaussian models predicted from inflation.

In the next two sections, we will use a slight variant of the popular local ansatz of primordial non-Gaussianity in order to generate two different scalings of the $n$-point functions and study how the scalings can be distinguished. For the usual local ansatz with hierarchical scaling, many groups have already performed simulations [170, 171, 172, 173, 174, 175] and found that the semi-analytic prescription of [158] for the non-Gaussian mass function works well. Here we will also test the validity of the mass function proposed in [31] and used in [159] for another scaling that is well motivated by inflationary models. In addition to looking at the mass function, we will also look at a signature of the shape (momentum dependence) of local type non-Gaussianity using the scale dependent halo bias [176], and at the stochastic bias on large scales [124]. Subleading contributions to the bias and leading contributions to the stochastic bias are sensitive to non-Gaussianity beyond the skewness.

In section 4.4, we will study an efficient method (by dividing a large survey into subvolumes) to measure the squeezed limit of any primordial trispectrum that generates non-zero bispectra in biased subvolumes.
4.1.2 Model

In the usual local ansatz, the Bardeen potential $\Phi(x)$ has non-Gaussian statistics thanks to a contribution from a non-linear, local function of a Gaussian field $\psi_G(x)$:

$$\Phi(x) = \psi_G(x) + f_{NL}(\langle \psi_G(x)^2 \rangle - \langle \psi_G(x) \rangle^2) .$$  \hspace{1cm} (4.1)

Here $f_{NL}$ parametrizes the size of the non-linear term and the level of non-Gaussianity. Non-Gaussianity of this type is usually thought of as produced by a light field that is not the inflaton [177, 178, 179, 180, 181, 182], and in fact cannot be generated by single field inflation proceeding along the attractor solution with modes in the Bunch-Davies vacuum [37, 33].

The two-point statistics (the power spectrum) of the fluctuations are well measured from CMB observations for about three decades in scale [183, 184, 185]. The amplitude of the three point function gives the skewness of the distribution of $\Phi$, and is also well-constrained by the CMB [148]. Higher order correlation functions are more difficult to measure in the data. We would like to see if we can get some handle on the structure of these higher order statistics of $\Phi$. For this purpose, we define the dimensionless moments $M_n$:

$$M_n = \frac{\langle \Phi(x)^n \rangle_c}{\langle \Phi(x)^2 \rangle_c^{n/2}}$$  \hspace{1cm} (4.2)

where the subscript $c$ indicates that we take the connected part of the $n$-point function.

Two scalings: For the local model given by Eq. (4.1), the moments scale approximately as

$$M_n^{\text{hier}} \approx A_n \left( \frac{M_3}{6} \right)^{n-2} , \quad n > 2$$  \hspace{1cm} (4.3)

where $A_n = 2^{n-3}n!$ comes out of combinatorics. The numerical coefficients in this scaling are not quite precise because of the difference in integrals over momenta at each order $n$ (see Appendix B.1), but the parametric dependence on the amplitude of the skewness and the total power is fixed. This is the behavior of the moments (which we label ‘hierarchical’ scaling) for $f_{NL}$ not too large i.e. when $2f_{NL}^2 \sigma^2 \ll 1$, where $\sigma^2 = \langle \psi_G(x)^2 \rangle$. However, if $2f_{NL}^2 \sigma^2 \gg 1$, the moments scale differently:

$$M_n^{\text{feeder}} \approx B_n \left( \frac{M_3}{8} \right)^{n/3} , \quad n > 2$$  \hspace{1cm} (4.4)

where $B_n = 2^{n-1}(n-1)!$. In the single source case, this is far too non-Gaussian to be consistent with observations. Therefore, we consider the following two source model

...
\[ \Phi(\vec{x}) = \phi_G(\vec{x}) + \psi_G(\vec{x}) + \tilde{f}_{NL} \left( \psi_G(\vec{x})^2 - \langle \psi_G(\vec{x})^2 \rangle \right) \]  

(4.5)

where the two Gaussian fields \( \psi_G(\vec{x}) \) and \( \phi_G(\vec{x}) \) are uncorrelated i.e. \( \langle \phi_G(\vec{x}) \psi_G(\vec{x}) \rangle = 0 \).

To express the correlations in the gravitational potential \( \Phi \) in terms of the amplitude of fluctuations in just one of the source fields we define

\[ q = \frac{P_{\psi,G}(k)}{P_{\psi,G}(k) + P_{\phi}(k)}, \]  

(4.6)

the ratio of the contribution of the Gaussian part of the field \( \psi \) to the total Gaussian power. \( P_{\psi,G}(k) \) is defined by

\[ \langle \psi_G(\vec{k}) \psi_G(\vec{k}') \rangle = (2\pi)^3 \delta_D^3(\vec{k} + \vec{k}') P_{\psi,G}(k). \]  

(4.7)

For simplicity, we assume both Gaussian components have constant, identical spectral indices but different amplitudes:

\[ P_{\phi}(k) = 2\pi^2 P_{\phi}(k)/k^3 \]  

\[ P_{\phi}(k) = A_{\phi} \left( \frac{k}{k_0} \right)^\gamma, \]  

(4.8)

and similarly for \( P_{\psi,G} \) (with the same index \( \gamma \) but a different amplitude \( A_{\psi} \)). With this choice, \( q \) is a scale-independent constant.

The power spectrum, bispectrum, and trispectrum in our model are:

\[ P_\Phi(k) = \left[ \frac{1}{q} + \tilde{f}_{NL}^2 I_1(k) P_{\psi,G}(k) \right] P_{\psi,G}(k) \]  

\[ = \left[ \frac{1}{1 - q} + \left( \frac{q \tilde{f}_{NL}}{1 - q} \right)^2 I_1(k) P_{\phi}(k) \right] P_{\phi}(k) \]  

(4.9)

\[ B(k_1, k_2, k_3) = 2 \tilde{f}_{NL} \left[ P_{\psi,G}(k_1) P_{\psi,G}(k_2) + 2 \text{ perm} \right] \]  

\[ + 2 \tilde{f}_{NL}^3 \left[ \int \frac{d^3 \vec{p}}{(2\pi)^3} P_{\psi,G}(p) P_{\psi,G}(\vert \vec{k}_1 - \vec{p} \vert) P_{\psi,G}(\vert \vec{k}_2 + \vec{p} \vert) + 3 \text{ perm} \right] \]  

(4.10)

\[ T(k_1, k_2, k_3, k_4) = 2 \tilde{f}_{NL}^2 \left[ P_{\psi,G}(k_1) P_{\psi}(k_2) P_{\psi,G}(\vert \vec{k}_1 + \vec{k}_3 \vert) + 23 \text{ perm} \right] \]  

\[ + \tilde{f}_{NL}^4 \left[ \int \frac{d^3 \vec{p}}{(2\pi)^3} P_{\psi,G}(p) P_{\psi,G}(\vert \vec{k}_1 - \vec{p} \vert) P_{\psi,G}(\vert \vec{k}_2 + \vec{p} \vert) P_{\psi,G}(\vert \vec{k}_3 + \vec{p} \vert) + 47 \text{ perm} \right] \]  

(4.11)
where we have defined $I_1(k)$ as

$$I_1(k) = \int_{k_{\min}}^{k_{\max}} \frac{du}{k} \int_{-1}^{1} d\mu \left[ u^{\gamma-1}(1 + u^2 + 2\mu u)^{\frac{\gamma-3}{2}} \right]$$  \hspace{1cm} (4.12)$$

Here, $k_{\min} = 2\pi/L$ is the infrared cutoff for a boxsize of $L$. In general, we do not know of the size of the universe beyond our observable universe, but for our purposes, the simulation box size $L$ is the natural choice. $k_{\max}$ is the scale leaving the horizon at the initial epoch [187]. The above integrals converge for large values of $k_{\max}$. For the computations to compare with simulations we set $k_{\max} = N_p^{1/3}k_{\min}$, where $N_p$ is the number of particles in a simulation. To arrive at the expressions quoted above, we have used

$$\int \frac{d^3\vec{p}}{(2\pi)^3} P_\psi(p)P_\psi(|\vec{k} - \vec{p}|) = \frac{1}{2} I_1(k)P_\psi(k)P_\psi(k).$$  \hspace{1cm} (4.13)$$

In Equations (4.9), (4.11) and (4.11) we have included terms that are usually sub-dominant in the case of single field, weakly non-Gaussian local ansatz. In our model, these terms are important when the field $\psi$ is strongly non-Gaussian. To discuss the observational constraints on $(q, \tilde{f}_{NL})$, let’s consider

$$P_\phi(k) \approx 10^{-9},$$  \hspace{1cm} (4.14)$$

$$I_1(k) \approx 10,$$  \hspace{1cm} (4.15)$$

$$f_{NL} \lesssim \mathcal{O}(10).$$  \hspace{1cm} (4.16)$$

Observational constraint from small non-Gaussianity can be satisfied by making

$$(q^2\tilde{f}_{NL}) \lesssim \mathcal{O}(10),$$  \hspace{1cm} (4.17)$$

$$(q\tilde{f}_{NL})^3 \lesssim \mathcal{O}(10^9),$$  \hspace{1cm} (4.18)$$

in which case the Gaussian contributions $(P_\phi(k) + P_{\psi,G}(k))$ dominate the total power spectrum in Eq.(4.9) as well. Notice that the non-Gaussian contribution to the power can shift the spectral index slightly, so that when the $\psi$ field is strongly non-Gaussian the measured spectral index is close to, but not identical to, the spectral index of the Gaussian components$^1$, $n_s - 1 \neq \gamma$.

If $q \ll 1$ and $(q\tilde{f}_{NL})^3 \ll 10^9$, then one can generate the feeder scaling in Eq.(4.4) without being inconsistent with the current observations of power spectrum and bounds

$^1$The new, integral terms have slightly different shapes than the usual terms. The difference, however, is small—approximately described by $\ln kL$ which has a weak dependence on $k$. These terms are also infrared divergent. For the purpose of comparing with the results from N-body simulations, the box size $L$ of the simulation provides a natural cutoff [173]. We will only look at quantities well enough inside the volume that the arbitrary size $L$ doesn’t affect our results.
on non-Gaussianity. The feeder scaling dominates when the condition $q f_{NL}^2 \gg 10^9$ is satisfied. This scaling, or a hybrid between hierarchical and feeder scaling, arises in particle physics scenarios where a second, non-Gaussian field couples to the inflaton and provides an extra source for the fluctuations [31, 188]. However, those scenarios differ from the model here because they most often generate bispectra not of the local type\footnote{One can also argue on statistical grounds that our observable universe is unlikely to have local non-Gaussianity of the type written in Eq.(4.5) with large $f_{NL}$ if inflation lasts much longer than 55 e-folds [189, 190].}. Here we are using the two field, local model primarily as a phenomenological tool, easy to implement in N-body simulations, to generate the feeder-type scaling of moments rather than as the output of a particular inflation model. The mass function is sensitive only to the integrated moments, not the shape, so this is a useful test of how different scalings affect the number of objects in a non-Gaussian cosmology.

4.2 Mass-function of massive halos

4.2.1 Abundance and clustering statistics

In this section we present the analytic predictions for the effect of locally non-Gaussian, two source initial conditions on the abundance and clustering of dark matter halos.

At large scales, the evolution of the density contrast is well described by linear perturbation theory, and the density contrast in Fourier space at redshift $z$ is given by

$$\delta(\vec{k}, z) = \alpha(k, z) \Phi(\vec{k}),$$

where

$$\alpha(k, z) = \frac{2}{3} \frac{k^2 T(k) D(z)}{H_0^2 \Omega_m}.$$  (4.19)

Here $T(k)$ is the transfer function, $D(z)$ is the growth function, $H_0$ is the Hubble scale today, and $\Omega_m$ is the energy density in matter today (compared to critical density). The smoothed density contrast, similarly, is

$$\delta_R(\vec{k}, z) = W_R(k) \alpha(k, z) \Phi(\vec{k}),$$

where

$$W_R(k) = \frac{3 \sin(kR) - 3(kR) \cos(kR)}{(kR)^3}.$$  (4.20)

is the smoothing function (here the Fourier transform of the real space top-hat). Note that we will generally suppress the $z$ dependence in $\alpha(k, z)$ and $\delta(\vec{k}, z)$, and usually write $\alpha(k)$ and $\delta(\vec{k})$ only. We can now compute the connected n-point functions of the smoothed density contrast in real space:

$$\langle \delta_R(\vec{x})^n \rangle_c(z) = \int \frac{d^3 \vec{k}_1}{(2\pi)^3} \cdots \int \frac{d^3 \vec{k}_n}{(2\pi)^3} \langle \delta_R(\vec{k}_1, z) \cdots \delta_R(\vec{k}_n, z) \rangle_c.$$  (4.21)
and therefore the dimensionless moments of smoothed density fields:
\[ \mathcal{M}_{n,R} = \frac{\langle \delta_{R}(\vec{x})^n \rangle_c}{\langle \delta_{R}(\vec{x})^2 \rangle^{n/2}}. \]

Note that the dimensionless moments are redshift independent. Eq. (4.21) can be separated into two components: (i) of \( O(\tilde{f}_{NL}^{-2}) \) and (ii) of \( O(\tilde{f}_{NL}) \), corresponding to contributions to the hierarchical and feeder scalings respectively. Some of these integrals, with a brief summary of our method to evaluate them, are given in Appendix B.1.

### 4.2.1.1 Mass function for non-Gaussian primordial fluctuations

We follow previous studies of non-Gaussian mass functions in that we calculate the ratio \( R = \frac{n_{NG}(M,z)}{n_G(M,z)} \) of the number density in the presence of non-Gaussianity, \( n_{NG} \), to the number density for Gaussian initial conditions, \( n_G \), for a particular halo mass \( M \) at some redshift \( z \) using the Press-Schechter formalism [191]. The fractional volume \( F(M) \) of dark matter in the collapsed structures (halos) is proportional to \( \int_{\delta_c}^{\infty} P(\delta_M) d\delta_M \), where \( \delta_c \) is the critical value of the smoothed density contrast \( \delta_M \) above which the dark matter in a region collapses to form halos and \( P(\delta_M) \) is a probability density function (PDF). Here we have written the smoothing scale in terms of the mass \( M \) rather than the smoothing radius \( R \); they are simply related by \( M = \frac{4}{3} \pi R^3 \rho_m \), where \( \rho_m \) is the mean matter density of the universe. Then, the number density (mass function) is given by:

\[
\frac{d n}{d M} = -\frac{d F(M)}{d M} \times \frac{\rho_m}{M}
\]

For a Gaussian PDF, one can easily perform the integration to find a prediction for the mass function. However, the result is known to be only an approximation and in practice Gaussian mass functions are calibrated on simulations [192].

To apply the method above in case of non-Gaussian initial conditions we need a non-Gaussian PDF. The Petrov expansion [193] (which generalizes the Edgeworth expansion [194, 195]) expresses a non-Gaussian PDF as a series in the cumulants of the distribution. In terms of \( \mathcal{M}_{n,R} \)'s, this is:

\[
P(\nu, R) = \frac{e^{-\nu^2/2}}{\sqrt{2\pi}} \left[ 1 + \sum_{s=1}^{\infty} \sum_{k_m} H_{s+2r}(\nu) \prod_{m=1}^{s} \frac{1}{k_m!} \left( \frac{\mathcal{M}_{m+2,R}}{(m+2)!} \right)^{k_m} \right]
\]

where
\[
\nu = \frac{\delta_M}{\sigma_M}, \sigma_M = \sqrt{\langle \delta_M^2 \rangle}
\]
and $H_n$'s are the Hermite polynomials defined as

$$H_n(\nu) = (-1)^n e^{\nu^2/2} \frac{d^n}{d\nu^n} e^{-\nu^2/2}.$$ 

The second sum is over the non-negative integer members of the set $\{k_m\}$ that satisfy

$$k_1 + 2k_2 + \cdots + sk_s = s.$$  \hspace{1cm} (4.24)

For each set $r \equiv k_1 + k_2 + \cdots + k_s$. This series can be integrated term by term to obtain $F(M)$, and with Eq.(4.22) gives ratio of non-Gaussian mass function to Gaussian mass function \[159\]

$$\frac{n_{NG}}{n_G} \approx 1 + \frac{F_{h,f}^h(M)}{F_0^h(M)} + \frac{F_{h,f}^f(M)}{F_0^f(M)} + \cdots$$  \hspace{1cm} (4.25)

where

$$F_0^h(M) = -\frac{\nu_c^2(M)}{\sqrt{2\pi}} e^{-\frac{1}{2}\nu_c^2(M)^2}, \hspace{0.5cm} \nu_c = \frac{\delta_c}{\sigma_M}.$$  \hspace{1cm} (4.26)

The two superscripts $h, f$ indicate that the set of terms that are of the same order depends on whether higher order cumulants have hierarchical $(h)$ or feeder $(f)$ scaling. Formally grouping terms assuming the scalings in Eq(4.3) or Eq.(4.4) are exact gives (for $s \geq 1$)

$$F_s^{h'}(\nu) = F_0^f \sum_{\{k_m\}_h} \left\{ H_{s+2r} \prod_{m=1}^s \frac{1}{k_m!} \left( \frac{M_m + 2R}{(m+2)!} \right)^{k_m} + H_{s+2r-1} \frac{\sigma}{\nu} \frac{d}{d\sigma} \left[ \prod_{m=1}^s \frac{1}{k_m!} \left( \frac{M_m + 2R}{(m+2)!} \right)^{k_m} \right] \right\}$$  \hspace{1cm} (4.27)

$$F_s^{f'}(\nu) = F_0^f \sum_{\{k_m\}_f} \left\{ H_{s+2r} \prod_{m=1}^s \frac{1}{k_m!} \left( \frac{M_m + 2R}{(m+2)!} \right)^{k_m} + H_{s+2r-1} \frac{\sigma}{\nu} \frac{d}{d\sigma} \left[ \prod_{m=1}^s \frac{1}{k_m!} \left( \frac{M_m + 2R}{(m+2)!} \right)^{k_m} \right] \right\}.$$  \hspace{1cm} (4.28)

The prime stands for the derivative with respect to the halo mass $M$. In the hierarchical case, the sets $\{k_m\}_h$ still satisfy Eq.(4.24), while for feeder scaling the $\{k_m\}_f$ are the sets of non-negative integer solutions to

$$3k_1 + 4k_2 + \cdots + (s+2)k_s = s + 2.$$
Eq.(4.25) assumes that the two point statistics of the smoothed linear density contrast (i.e \( \langle \delta^2_M \rangle \)) are the same for the non-Gaussian and the Gaussian cases (not that the Gaussian contributions to \( \langle \delta^2_M \rangle \) are the same). However, this is difficult to maintain at all scales in our simulations. We require the two point clustering statistics to match at a particular scale \( R = 8\text{Mpc}/h \), but do not correct for the shift to the spectral index coming from the non-Gaussian term in the power spectrum. As a result, on scales other than \( R = 8\text{Mpc}/h \) we need to make a distinction between \( F'_{0,\text{NG}}(M) \) for the non-Gaussian cosmology and \( F'_{0,G}(M) \) for the Gaussian cosmology. In that case, there is an extra (mass or scale dependent) factor \( f_1(M) \) in the ratio \( \frac{\text{NG}}{\text{NG}} \), where

\[
f_1(M) = \frac{F'_{0,\text{NG}}(M)}{F'_{0,G}(M)} = \frac{v_{c,\text{NG}}(M)}{v_{c,G}(M)} e^{-\frac{1}{2}(v^2_{c,\text{NG}}-v^2_{c,G})}
\]

This factor is typically quite close to one. For example, in the mass range \((4 \times 10^{13} < M < 2 \times 10^{15})h^{-1}M_\odot\) at \( z = 1 \), \( 0.995 \lesssim f_1(M) \lesssim 1.001 \) for single field \( f_{\text{NL}} = 500 \) case. For our smallest feeder scaling simulation (with \( M_3 \approx 0.020 \)), in this mass scale and redshift range, \( 0.92 \lesssim f_1(M) \lesssim 1.02 \). The factor deviates away from unity more at larger redshifts and at mass scales far from \( M \approx 1.61 \times 10^{14}h^{-1}M_\odot \).

In addition, the derivation above assumes the same constant of proportionality between \( F(M) \) and \( \int_0^\infty P(\delta_M)d\delta_M \) regardless of the level of non-Gaussianity. The standard Press-Schechter constant of proportionality is two, but for the non-Gaussian case it is reasonable to fix the constant by requiring

\[
\bar{\rho} = \int_0^\infty M \frac{dn}{dM} dM.
\]

Gaussian and non-Gaussian cosmologies with identical \( \sigma_8 \) will have slightly different normalization factors. This factor shifts further away from 2 as the level of non-Gaussianity in the initial conditions is increased. Integrating various truncations of the expanded PDFs indicates we expect a difference from 2 between about 0.5% and 2% for the amplitudes of non-Gaussianity we consider in this work. So, we will introduce an extra factor \( f_2 \) that multiplies our analytical mass function to fit with the simulation results.

### 4.2.2 Simulations

The simulations for this project were performed using the popular GADGET-2 code \[196\]. The initial conditions were generated using second order Lagrangian perturbation theory (2LPT) \[197\]. We used the code from \[197\] for generating local type (single field) non-Gaussian initial conditions using 2LPT and modified the code—discussed next—to generate initial conditions for our two field model.
4.2.2.1 Initial conditions

First, two Gaussian random fields ($\phi_G(\vec{x})$ and $\psi_G(\vec{x})$) were generated using the power spectrum of our fiducial Gaussian cosmology ($n_s = 0.96, \sigma_8 = 0.8, \Omega_m = 0.27, \Omega_\Lambda = 0.73$). The amplitude of the power spectrum for the $\psi_G(\vec{x})$ and $\phi_G(\vec{x})$ fields were multiplied by the appropriate factors $q$ and $(1 - q)$ respectively. Then, the $\psi_G(\vec{x})$ field was squared and multiplied by $\tilde{f}_{NL}$. The total non-Gaussian field $\Phi(\vec{x})$ was obtained by adding the three components as in Eq. (4.5). We ensured $\sigma_8$ of the generated non-Gaussian field was that of our specified cosmology for all of our parameter sets by renormalizing the $\Phi(\vec{x})$ field by a factor of $(\sigma_8 / \sigma_8,\phi)^{0.5}$. These are the only modifications done to the 2LPT code. The rest of the code generated the required displacement and velocity fields as usual. The redshift of all our initial conditions is $z_{\text{start}} = 49$.

4.2.2.2 N-body simulations

The public version of the GADGET-2 code was used to perform cosmological collisionless dark matter only simulations. All simulations were done with $(1024)^3$ particles in a cube of side $2400h^{-1}\text{Mpc}$. This gives the mass of a single particle to be $9.65 \times 10^{11}h^{-1}\text{M}_\odot$. The force softening length was set to be 5% of the inter-particle distance. Simulations with Gaussian initial conditions ($q = 0$ and $\tilde{f}_{NL} = 0$) were also performed with the same seeds as the $\phi$ field (as it has the dominant contribution for small $q$) to compare with our feeder models; another set of Gaussian simulations was performed with $q = 1$ and $\tilde{f}_{NL} = 0$ to compare with the hierarchical simulations. For each set of parameters listed in Table 4.1, we ran four simulations with different seeds. All simulation results reported in this section are average over the four simulations. Similarly, the errors reported are the 1σ standard deviation of the four simulations. The AHF halo finder [198] was used to identify halos which were then used to get the mass function of dark matter halos and power spectra of halos (halo-matter cross power spectrum and halo-halo autospectrum). In all our analyses, we only use halos with number of dark matter particles $N_p \geq 50$.

4.2.2.3 Parameter space of simulations

Since our method of generating the feeder scaling produces a slightly different bispectrum shape than the hierarchical case, the scale dependence of $M_3$ for the two scalings also differ. For comparison purposes, we will define $f_{NL}^{\text{eff}}$ at the scale of $R = 8\text{Mpc}/h$ corresponding to a halo mass of $1.61 \times 10^{14}h^{-1}\text{M}_\odot$, as the ratio:

$$f_{NL}^{\text{eff}} = \frac{M_3(q, \tilde{f}_{NL})}{M_3(q = 1, \tilde{f}_{NL} = 1)} \bigg|_{M = 1.61 \times 10^{14}h^{-1}\text{M}_\odot}$$

(4.30)

In Table 4.1, we list the parameter sets of $(q, \tilde{f}_{NL})$ that we have simulated. Notice that these parameter sets are not consistent with the small $f_{NL}^{\text{local}}$ reported by the Planck
Table 4.1: Parameter space of our simulations. In the first column, we name our models following a simple naming convention. The first letter of the name stands for the type of scaling of the model: F stands for feeder scaling, H stands for hierarchical scaling, M stands for mixed scaling (when neither component is negligible). The number following is the approximate f_{eff}^NL, also listed in the second column. For example: F70 means that the scaling of the model is feeder and has f_{eff}^NL = 70. The quantity q in the third column is defined in Eq.(4.6), and gives the ratio of power in the Gaussian part of the $\psi$ contribution ($P_{\psi,G}$), to the total Gaussian power ($P_{\psi,G} + P_{\phi}$). The second last column is the dimensionless skewness $M_3$ computed at a halo mass scale $M = 1.61 \times 10^{14} h^{-1} M_\odot$. The last column $M_{3,f}/M_{3,h}$ is the ratio of the dimensionless skewness $M_3$ from the feeder contribution to that of the hierarchical contribution and indicates the relative importance of the feeder term.

<table>
<thead>
<tr>
<th>Name</th>
<th>f_{eff}^NL (Eq:(4.30))</th>
<th>q</th>
<th>f_{NL}</th>
<th>$M_3$</th>
<th>$M_{3,f}/M_{3,h}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(M993)</td>
<td>993</td>
<td>0.1</td>
<td>50000</td>
<td>0.290</td>
<td>2.144</td>
</tr>
<tr>
<td>(F677)</td>
<td>677</td>
<td>0.00005</td>
<td>$10^8$</td>
<td>0.198</td>
<td>4287</td>
</tr>
<tr>
<td>(H500)</td>
<td>500</td>
<td>1</td>
<td>500</td>
<td>0.145</td>
<td>0.0027</td>
</tr>
<tr>
<td>(M384)</td>
<td>384</td>
<td>0.11925</td>
<td>20620</td>
<td>0.112</td>
<td>0.5089</td>
</tr>
<tr>
<td>(F215)</td>
<td>215</td>
<td>0.00003</td>
<td>$10^8$</td>
<td>0.063</td>
<td>2923</td>
</tr>
<tr>
<td>(F122)</td>
<td>122</td>
<td>0.00003</td>
<td>$8 \times 10^7$</td>
<td>0.036</td>
<td>1933</td>
</tr>
<tr>
<td>(H99)</td>
<td>99</td>
<td>1</td>
<td>99</td>
<td>0.029</td>
<td>0.0001</td>
</tr>
<tr>
<td>(F70)</td>
<td>70</td>
<td>0.00003</td>
<td>$6.5 \times 10^7$</td>
<td>0.020</td>
<td>1303</td>
</tr>
</tbody>
</table>

mission from bispectrum measurements. However, it is necessary to use parameter sets with larger values of f_{eff}^NL in order to get useful results from N-body simulations.

4.2.3 Mass function results and discussion

The hierarchical scaling has been considered a number of times already with the prediction from the Edgeworth-series formalism [158] providing good fit to the outputs from simulations [170, 171, 172, 173, 174, 175]. Our focus will be on the feeder type scaling.

In the following sections, we will use the mass function truncated up to $M_5$. From error analysis (Appendix B.2), we see that we gain little by adding higher terms for the feeder mass function. To this order, for the feeder case, the ratio of non-Gaussian to Gaussian mass function is:

$$
\left( \frac{n_{NG}}{n_G} \right)_{\text{feed}} = f_1(M) f_2^{\text{feed}} \left[ 1 + \frac{M_3}{3!} H_3(\nu_c) - \frac{M'_3}{3! \nu_c^3} H_2(\nu_c) + \frac{M_4}{4!} H_4(\nu_c) - \frac{M'_4}{4! \nu_c^4} H_3(\nu_c) + \frac{M_5}{5!} H_5(\nu_c) - \frac{M'_5}{5! \nu_c^5} H_4(\nu_c) \right]
$$

(4.31)

where we have chosen to rewrite all the quantities in terms of the halo mass $M$ rather
than $\sigma$ and $R$; the $M$ dependence of the moments $M_n$’s and $\nu_c$ has been suppressed for clarity, as is done throughout the section. We use this expression to fit to the results from N-body simulations.

Similarly, for hierarchical scaling of moments, the expression for the non-Gaussian mass function including terms up to $O(M^5)$ is:

\[
\left( \frac{n_{NG}}{n_G} \right)_{\text{hier}} = f_1(M) f_2 \left[ 1 + \left( \frac{M_3 H_3(\nu_c)}{3!} - \frac{M_3' H_2(\nu_c)}{3! \nu_c} \right) + \left( \frac{M_4 H_4(\nu_c)}{4!} + \frac{M_3 H_6(\nu_c)}{2 \times 3! \times 3!} - \frac{M_4' H_3(\nu_c)}{4! \nu_c} - \frac{M_3 M_3' H_5(\nu_c)}{3! \times 3! \nu_c} \right) \right. \\
+ \left( \frac{M_5 H_5(\nu_c)}{5!} - \frac{M_5' H_4(\nu_c)}{5! \nu_c} + \frac{M_3 M_3 M_4 H_7(\nu_c)}{4! 3!} + \left( \frac{M_3}{3!} \right)^3 H_9(\nu_c) \right. \\
- \left( \frac{M_4 M_3' + M_3 M_4' H_6(\nu_c)}{3! 4! \nu_c} - \frac{1}{3!} \left( \frac{M_3}{3!} \right)^2 H_9(\nu_c) \right. \\
+ \left. \frac{M_3 M_3' M_3'' H_8(\nu_c)}{3! 3! \nu_c} \right] 
\]

where the $M$ dependence of $M_n$ and $\nu_c$ has been suppressed for clarity. In the above mass function formulae for feeder scaling Eq.(4.31) and for hierarchical scaling Eq.(4.32), the expression for $f_1(M)$ is calculated for a given model using Eq.(4.29) and the $f_2$ factors are fit for each of our simulation results. Both of these factors approach unity for small non-Gaussianity. To compute the cumulants necessary to calculate the dimensionless moments $M_n$ in the above formulae, we have used the Monte-Carlo method described in [173]. See Appendix B.1 for details.

Let us now compare and discuss the results from simulations and calculations from the Edgeworth series formulation. First, as a check of our simulations, we did a purely hierarchical scaling parameter set—single field, $f_{NL} = 500$ simulation—which can be compared directly with previous works. Then, for the two source case, we study the parameter sets listed in Table 4.1 which allow feeder scaling as well as mixed scaling (i.e., neither term in Eqs.(4.11),(4.11) is negligible). In this section, we will present the simulation results and the corresponding Edgeworth mass functions. We analyzed our mass function results with a simple two parameter ($\delta_c, f_2$) chi-square minimization procedure. The errors reported in the best fit values increase the reduced chi-square of the fit by unity when added to the best fit values. We find that all simulations ($H99, H500, F70, F122, F215$) prefer a reduced $\delta_c \approx 1.4 - 1.5$ and different values for $f_2$, which are of the expected size and increase appropriately with $M_3$. We will discuss the dependence of this factor on $M_3$ and the type of scaling later.

First, let us present our mass function fits. We use $\delta_c = 1.46$ obtained by performing chi-square minimization together for $H99$, $H500$, $F70$, $F122$, $F215$ by forcing the same $\delta_c$ but allowing overall rescaling factors ($f_2$) for each case. This is a reasonable procedure if one interprets shifting $\delta_c$ as allowing departures from the assumption of spherical collapse, which should be relatively independent of the level of non-Gaussianity. However,
the normalization requirement

\[ \tilde{\rho} = \int_0^\infty M \frac{dn}{dM} dM \]

(which has not been analytically enforced) suggests that \( f_2 \) should depend on the level of non-Gaussianity. The top panel of Figure 4.1 has our mass function results for two simulations with hierarchical scalings.

The results for two feeder simulations are presented in the bottom panels of Figure 4.1. As discussed in Appendix B.2, the errors on our truncated feeder mass function are large compared to the hierarchical case with comparable \( M_3 \). Also, the error becomes large at relatively small \( \nu \) even for the feeder scaling case with \( M_3 \approx 0.03 \). Therefore, based on our error evaluation, we do not expect that our feeder mass function describes the simulation results well for the more massive halos or at higher redshifts for which \( \nu_c \gtrsim 3 \). We find that the feeder mass function formula Eq.(4.31) fits well our simulation results for F70 and F122 (see the left panel of Figure 4.2 for the F122 simulation results). Consistent with the error analysis of feeder mass function, Eq.(4.31) fits to F215 with \( M_3 \approx 0.063 \) are not equally good. With a larger \( M_3 \approx 0.198 \), the F677 simulation is clearly not well fit by our truncated feeder mass function (see Figure 4.3). Finally, Figure 4.1 also shows the difference between the simulations and Eq.(4.31) assuming \( \delta_c = 1.686, f_2 = 1, \) and that the moments scale exactly as in Eq.(4.4). That is, the bottom-most panels illustrate how calibrating on simulations shifts the purely analytic expectations for the non-Gaussian mass function. The dominant effect among these three factors is that of \( \delta_c \); change in \( \delta_c \) affects the non-Gaussian mass functions starting at \( O(M_3) \). On the other hand, using different scalings of higher moments only change the expressions starting at \( O(M_4) \) and the \( f_2 \) factors modify the mass functions at a few percent level at most (see Figure 4.4).

In Figure 4.2, the right hand panel shows the fractional difference between the semi-analytic expression Eq.(4.31) (with the cumulants \( M_n \) measured in the realizations and \( \delta_c, f_2 \) fit) and the simulation results for the two feeder models: F122 and F215. We have plotted the fractional difference as a function of \( \nu_c = \delta_c / \sigma_M \) and the plotted points include simulation results from all three redshifts \( z = 0, 1, 2 \). The result can be interpreted as the error in the semi-analytic approach and qualitatively correlates with our analytic error analysis of the PDF: (i) the error at low \( \nu_c \) is small, (ii) the error increases for higher \( \nu_c \) but the error is smaller for smaller \( M_3 \).

On a different note, by looking at the non-Gaussian mass function results from simulations only, we can verify that the F70 model is more non-Gaussian than the H99 model (compare top and bottom plots in Figure 4.1), even though the skewness of F70 is smaller than the skewness of H99. Similarly, we also see that the F215 model has comparable amount of non-Gaussianity as the H500 model. This verifies that the non-Gaussian mass function is sensitive to the total non-Gaussianity and the scaling of
Figure 4.1: Top: The simulation results and semi-analytic prediction Eq.(4.32) for hierarchical simulations. Left panel: $M_3 = 0.145$ ($f_{NL} = 500$) with $f_2 = 1.042$. Right panel: $M_3 = 0.029$ ($f_{NL} = 99$) with $f_2 = 1.009$. Bottom: The simulation results and semi-analytic prediction Eq.(4.31) for feeder simulations. Left panel: $M_3 = 0.063$ ($f_{eff}^{NL} = 215$) and $f_2 = 1.043$. Right panel: $M_3 = 0.020$ ($f_{eff}^{NL} = 70$) with $f_2 = 1.012$. In addition, we also plot the fractional difference between the simulation results and the prediction from the truncated mass function Eq.(4.31) but using ($\delta_c = 1.686, f_3 = 1$), and assuming that $M_n$ scale as Eq.(4.4). A negative value means that the analytic mass function overpredicts the simulation result.
Figure 4.2: Left: The simulation results and the mass function prediction Eq.(4.31) for F122 model: $M_3 = 0.036 \, (f_{NL}^{\text{eff}} = 122)$ with $f_2 = 1.021$. Right: the fractional error in the semi-analytic predictions for F215 and F122 models compared to simulation results. The plotted points include results from all three redshift values that we have looked at i.e. $z = 0, 1, 2$.

higher moments in the initial conditions.

The right hand panel of Figure 4.3 is a mixed scaling simulation with $M_3 = 0.112$, and approximately a third of the contribution to the $M_3$ coming from the feeder component. For the theory curve we used both Eq.(4.31) and Eq.(4.32) for the corresponding $M_3$ components but forced the same $f_2$. As with the case with other simulations containing feeder scaling, the high $\nu_c$ mass function results are not described well by the Edgeworth mass function.

All in all, our analysis of the truncation error and the Edgeworth fits to the mass functions from simulations were qualitatively consistent with each other. Hence, we find that the error evaluation of the PDF is a good indicator of the accuracy of the Edgeworth (or Petrov) mass function. However, to fit the simulations well we needed to rescale the analytic ratio of non-Gaussian to Gaussian mass function by an extra $M_3$ dependent parameter, $f_2$, for both scalings. This rescaling seems reasonable to enforce the same total matter density regardless of level of non-Gaussianity. In Figure 4.4, we plot our best fit $f_2$ values as a function of $M_3$ for both scalings. We find a simple linear relation between $f_2$ and $M_3$: 

\[ f_2^{\text{hier}} = 1.0 + 0.29 M_3 \]  \hspace{1cm} (4.33)  
\[ f_2^{\text{feed}} = 1.0 + 0.66 M_3 \]  \hspace{1cm} (4.34)
Figure 4.3: The simulation results and semi-analytic mass function prediction for: (i) left: a feeder simulation with $M_3 \approx 0.198$ ($f_2 = 1.16$), and (ii) right: a mixed scaling simulation with $M_3 \approx 0.112$ ($f_2 = 1.06$).

Figure 4.4: $M_3$ dependence of $f_2$ defined in Eq.(4.31) and Eq.(4.32) for feeder and hierarchical mass functions respectively. The data points are best fit $f_2$ obtained from simulations and to obtain the best fit lines we require $f_2 = 0$ for $M_3 = 0$. 
Figure 4.5: Results from our simulations for both hierarchical and feeder scalings as density plots. The quantity plotted is $\log_{10}(n_{NG}/n_G)_{\text{sim}}$. The points overlayed on the plots are the data points used to obtain the density plots. The red solid line shows the maximum trusted $M_3$ when allowing for 20% error in the PDF from $\nu_c$ to some $\nu_{c,\text{max}}$. The maximum trusted $M_3$ line for the hierarchical case lies outside of the plot range. See Appendix B.2 for details of how these curves were obtained.

Figure 4.6: The relative difference between the semi-analytic predictions (Eq. (4.32 or Eq. (4.31)) and the simulation results as density plots. See Eq. (4.35) for the precise definition of the quantity plotted. We have omitted simulation results for which the uncertainty is larger than 20%. The important observation to note here is that the relative difference is quite small i.e. $\ll 0.2$, below the solid red line.
Analytic suggestions for non-Gaussian mass functions have also been obtained using excursion set methods [199, 200, 174, 201, 202, 203, 204]. It would be interesting to compare the predictions for additional corrections to Eq.(4.25) from those methods to these simulations, particularly the results for $f_2$ shown in Figure 4.4.

To finish up our discussion of mass function results, in Figure 4.5 we show the non-Gaussian mass function results (from simulations) in a two dimensional density plot on a $\nu_c - M_3$ plane, for both hierarchical and feeder scalings. We plot the quantity $\log_{10} \left( [n_{NG}/n_G]_{\text{sim}} \right)$. In the same plots, we overlay the maximum $M_3$ that we can trust the PDF 4.23, $M_{3,\text{max}}$, as a function of $\nu_c$; see error analysis in Appendix B.2 for details. In Figure 4.6, we have plots similar to those in Figure 4.5 but now we are plotting the relative difference of our mass function predictions (from 4.32 and 4.31) from the simulation results. The quantity plotted is:

$$R_{SA} = \frac{\left| \frac{(n_{NG}/n_G)_{\text{sim}} - (n_{NG}/n_G)_{\text{semi-analytic}}}{(n_{NG}/n_G)_{\text{sim}}} \right|}{(4.35)}$$

These plots summarize our results for the mass function (for simulations with $M_3 < 0.08$). Each isolated region in the density plots represents one simulation, which can be easily mapped to the exact simulation (in Table 4.1) by looking at the $M_3$ value. Figure 4.5 simply shows our simulation mass function results. But from Figure 4.6, we can see that the magnitude of relative difference of our semi-analytic mass functions to the simulation results is generally less than ten percent ($R_{SA} \lesssim 0.1$) for the values of $\nu_c$ and $M_3$ that are calculated to be trustworthy by evaluating the series truncation error of the PDF 4.23. This is telling us that our semi-analytical mass function fits are consistent with the truncation error analysis.

### 4.2.4 Constraints from X-ray clusters

The semi-analytic mass functions [Eq.(4.32) and Eq.(4.31)] have been used in [53] to constraint the amplitude of $f_{NL}$ using galaxy clusters for the two models. See Figure 4.7.

### 4.2.5 Conclusion

Our simulations show that the Petrov expansion gives a good approximation to the non-Gaussian mass function when the amount of non-Gaussianity is small enough. The criteria for “small enough” depends on the dimensionless skewness $M_3$ and the scaling of moments. We have verified that for the same level of skewness, the feeder scaling is more non-Gaussian than the hierarchical scaling by a straightforward comparison of simulation outputs of the two cases. For the parameter space we probed, we were able to show that the truncation error evaluation of the Petrov PDF, Eq.(4.23), correlated well with the degree to which the truncated non-Gaussian mass function fitted the simulation
results. This gives us further confidence in the use of the truncated Petrov (Edgeworth) mass function for various analyses such as that of [205, 159] to put constraints on primordial non-Gaussianity using number counts of objects. We make progress towards calibrating the mass function formulae Eq.(4.32) and Eq.(4.31) for both scalings by using an extra parameter $f_2$, in addition to verifying that a reduced $\delta_c$ is preferred, which is consistent with previous simulation studies of non-Gaussian mass functions with non-Gaussianity of local type and hierarchical scaling of higher moments. The
effect this calibration might have on previous analysis of cluster constraints is illustrated in the bottom panels of Figure 4.1.

4.3 Large-scale bias and stochasticity

4.3.1 Large-scale bias

On large scales, where density fluctuations are in the linear regime, the clustering of halos is expected to trace the clustering of the underlying matter field. The proportionality constant relating halo clustering to matter clustering is called the halo bias. Local type non-Gaussianity can modify halo bias, compared to Gaussian universes, by coupling the amplitude of short wavelength modes to that of long wavelength modes. Since the coupling occurs in the gravitational potential field (with constant amplitude), rather than in the density field, local type non-Gaussianity introduces a new, scale-dependent term relating the power spectrum of halos to the power spectrum of the linear dark matter field. On large scales (small wave numbers) the non-Gaussian term can dominate and the analytic prediction for the bias is relatively simple.

The potential use of the halo (or galaxy) bias as a probe of primordial non-Gaussianity was first demonstrated in [176]. An analytic derivation capturing the first order effect of non-Gaussian initial conditions had been presented much earlier in [206] and clarified and improved following the Dalal et al result in [207, 208, 209]. The halo bias in models with two sources for the primordial fluctuations was considered in [124, 210], and in [211] which gives some theoretical predictions for the model we consider here. In addition, [212] previously performed N-body simulations for a model where the kurtosis was larger than in the single field case by a factor of two. That work corresponds to our scenario with \( q = 0.5 \). Here we are primarily interested in values of \( q \) that are very small so that the scaling is feeder type, but we also consider cases with \( q \approx 0.1 \) which have intermediate scaling. Measurements of the power spectra of several different galaxy populations have been used to place constraints on primordial non-Gaussianity, at roughly the \( |\sigma(f_{NL})| \sim \mathcal{O}(25 - 200) \) level, depending on the population and treatment of systematic errors [164, 45, 165, 166, 46, 167, 168, 169, 48].

The general form of the large-scale bias for our two-field model Eq.(4.5) can be calculated using the peak-background split formalism [45, 212]. Appendix B.3 has the peak-background-split calculation that results in Eq.(B.15) as the expression of bias. In case of small, local non-Gaussianity \( (q \ll 1 \text{ for the case with a second, strongly non-Gaussian field}) \), the expression for large-scale bias reduces to:

\[
P_{hm}(k) = \left[ b_{\text{all sources}} + \frac{2\delta_c(b_{\text{NG source}} - 1)}{\alpha(k)} \left( q^2 \hat{f}_{NL} + (q f_{NL})^3 P_{\phi}(k) I_1(k) \right) \right] P_{mm}(k).
\]  

(4.36)
Here we have used subscripts on the bias coefficients to emphasize that the scale-independent term depends on all sources of the fluctuations \((b_{\text{all sources}})\), while the scale-dependent term depends only on those sources with a primordial non-Gaussian component \((b_{\text{NG source}})\). Recall that in the more frequently quoted expression with a single, non-Gaussian source, these two bias coefficients are both equal to the Gaussian bias plus scale-independent corrections proportional to the level of non-Gaussianity. When multiple sources are present, the first bias coefficient can be split into terms which are include the Gaussian bias for each source, while the bias coefficient in the second, scale-dependent term is to lowest order the Gaussian bias for the non-Gaussian source. In addition, Eq.(4.36) demonstrates that the large scale halo bias is a probe of \(q^2 f_{\text{NL}}\) for the hierarchical scaling and a probe of \((q f_{\text{NL}})^3\) for the feeder scaling. In other words, the dominant contribution to the non-Gaussian bias is proportional to the amplitude of the bispectrum as expected.

To compare the analytic expressions against simulation results, we will use the size of the simulation volume to truncate integrals in the infrared and will not fit the power spectrum on scales very close to the simulation box size \((k \lesssim 0.007 \, \text{hMpc}^{-1})^3\).

### 4.3.2 Large-scale stochasticity

Using the peak-background-split method, we can also calculate the halo-halo power spectrum \(P_{hh}(k)\) in terms of the underlying matter distribution \(P_{mm}(k)\). We can then define the large-scale stochasticity, or cross-correlation coefficient as

\[
r^2(k) = \frac{P_{hm}^2(k)}{P_{hh}(k)P_{mm}(k)}
\]  

(4.37)

The calculations and the corresponding expressions for \(P_{hm}(k)\) and \(P_{hh}(k)\) are given in Appendix B.3. Peak-background split calculations suggest that in the large scale limit (small \(k\)), \(r^2(k)\) gives the fraction of power in the initial conditions from contributions of the field with non-Gaussianity (i.e. \(1 - (P_{\phi}(k)/P_\Phi(k))\)). For small non-Gaussianity, in case of hierarchical scaling, this is \(\approx q\), while in case of feeder scaling, this is \(\approx (q f_{\text{NL}})^2 P_\phi(k)\). For the single field cases, there is no large-scale stochasticity \((r^2(k) = 1)\). We will test these expectations against results from simulations. Large-scale stochasticity for the Gaussian case and non-Gaussian case with hierarchical scaling was discussed in [212]; their comparison with simulations showed a small discrepancy between their analytic expression and the simulation results. Our results below indicate one possible source for at least some of that discrepancy.

---

3First, for \(k \lesssim 0.007 \, \text{hMpc}^{-1}\) i.e. modes approaching the scale of the box size of our simulations, the sample variance is large. Second, the bias for feeder scaling depends on \(I_1(k)\) which has a sharply decreasing behavior near \(k_{\text{min}}\) of the simulation; this effect runs with the box size. If we are interested in scales near the \(k_{\text{min}}\) \((k \lesssim 0.007 \, \text{hMpc}^{-1})\) for our feeder scalings, then we will either have to do simulations with larger box sizes, or generate initial conditions differently such that when averaged over many realizations of the initial conditions we don’t get this feature.
Appendix B.3. These expressions written in terms of the total matter power spectrum

Table 4.2: Large-scale bias and stochasticity for simulations with Gaussian initial conditions. Simple chisquare fitting was performed to obtain the best fit values. The errors in $b_g$ and $r_g^2$ are computed such that the reduced chisquare increases by unity when adding the error to the best fit values. We will use the same procedure to obtain best fit values and the corresponding error for our non-Gaussian bias and stochasticity results.

<table>
<thead>
<tr>
<th>$z$</th>
<th>Mass range of halos used</th>
<th>$b_g$</th>
<th>$r_g^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$(1.93 \leq M \leq 3.85) \times 10^{14} h^{-1} M_\odot$</td>
<td>$3.36 \pm 0.09$</td>
<td>$1.13 \pm 0.06$</td>
</tr>
<tr>
<td></td>
<td>$(0.965 \leq M \leq 1.92) \times 10^{14} h^{-1} M_\odot$</td>
<td>$2.57 \pm 0.04$</td>
<td>$1.08 \pm 0.03$</td>
</tr>
<tr>
<td></td>
<td>$(4.83 \leq M \leq 9.55) \times 10^{13} h^{-1} M_\odot$</td>
<td>$1.92 \pm 0.04$</td>
<td>$1.03 \pm 0.02$</td>
</tr>
<tr>
<td>1</td>
<td>$(0.965 \leq M \leq 1.92) \times 10^{14} h^{-1} M_\odot$</td>
<td>$5.83 \pm 0.24$</td>
<td>$1.01 \pm 0.10$</td>
</tr>
<tr>
<td></td>
<td>$(4.83 \leq M \leq 9.55) \times 10^{13} h^{-1} M_\odot$</td>
<td>$4.37 \pm 0.11$</td>
<td>$1.04 \pm 0.05$</td>
</tr>
<tr>
<td>2</td>
<td>$(4.83 \leq M \leq 9.55) \times 10^{13} h^{-1} M_\odot$</td>
<td>$8.68 \pm 0.44$</td>
<td>$0.84 \pm 0.29$</td>
</tr>
</tbody>
</table>

4.3.3 Bias and stochasticity results

Next we present simulation results and fit to the analytical predictions for large-scale bias and large-scale stochasticity. Before considering the non-Gaussian results, we present the results from our Gaussian simulations in Figure 4.8. Both the bias and the stochastic bias (after the $1/\bar{n}$ shot noise correction to $P_{hh}(k)$) are constant for $k < 0.04$ hMpc$^{-1}$, consistent with lowest order predictions for Gaussian fluctuations. Further, the large-scale stochasticity parameter $r^2(k)$ is predicted to approach unity for Gaussian initial conditions. Figure 4.8 (right plot) is consistent with large-scale stochasticity being constant for Gaussian initial conditions; however, the constant value is slightly greater than unity at $r_g^2 = 1.03 \pm 0.02$ when halos in the mass range $4.83 \times 10^{13} h^{-1} M_\odot \leq M \leq 9.55 \times 10^{13} h^{-1} M_\odot$ are used at $z = 0$. For our other $z = 0$ samples, the value of $r_g^2$ increased when using halos of larger mass. The values of the fitted parameters (bias and stochasticity) for our Gaussian simulations are summarized in Table 4.2. We also note that the shot noise corrections are large for some of our halo samples. For example, the ratio of shot noise correction to the uncorrected halo power spectrum is largest for the halo sample at $z = 2$ (Gaussian simulations) in Table 4.2; at a reference wavenumber $k \approx 0.02$ hMpc$^{-1}$, this ratio is $\approx 0.8$. Typically, this ratio is smaller ($\approx 0.3 - 0.5$) for the samples at smaller redshift in Table 4.2. Also, note that for many of our non-Gaussian models, the fractional shot noise contribution decreases to $\approx 0.3$ even for the $z = 2$ sample as there are more halos in the non-Gaussian samples at higher redshifts. Some progress has been made recently towards a better understanding and modeling of large-scale stochasticity [213], beyond the usual shot noise correction, for Gaussian initial conditions. In this work, however, we are focusing on the non-Gaussian effect only.

We have derived expressions for $P_{hm}(k)$ and $P_{hh}(k)$ for our two source model in Appendix B.3. These expressions written in terms of the total matter power spectrum
Figure 4.8: The bias (left) and stochasticity (right) at large scales for the Gaussian simulations at \( z = 0 \) and \( z = 1 \) using halos in the mass range \( 4.83 \times 10^{13} h^{-1} M_\odot \leq M \leq 9.55 \times 10^{13} h^{-1} M_\odot \). The values of large scale Gaussian bias and stochastic bias obtained through the fits can be found in Table 4.2.

Figure 4.9: Left: The simulation result for bias \( P_{hm}/P_{mm} \) at large scales for the hierarchical simulations at \( z = 0 \) using halos in the mass range \( 9.65 \times 10^{13} h^{-1} M_\odot \leq M \leq 1.92 \times 10^{14} h^{-1} M_\odot \) and the corresponding best fit using Eq.(4.38). For each set of simulation data, we also include the curve obtained by using \( b_g \) from Gaussian simulation instead of the fitted \( b_\psi \). For the case of \( H500 \) shown in the figure, \( b_\psi = 1.86 \pm 0.03 \) compared to \( b_g = 1.92 \pm 0.04 \). Right: The corresponding large-scale stochasticity simulation results Eq.(4.37) and the best fit constant values.
\[ P_{mm}(k) = \alpha(k)^2 P_{\Phi}(k) \]
and the constant halo bias coefficients \((b_\phi \text{ and } b_\psi)\) for the two independent fields \(\phi\) and \(\psi\) are:

\[
\frac{P_{hm}(k)}{P_{mm}(k)} = \frac{b_\phi (1 - q)}{1 + \tilde{f}^2_{\text{NL}} I_1(k) q P_{\psi,G}} \\
+ \left( b_\psi + 2 \delta_c(b_\psi - 1) \frac{\tilde{f}_{\text{NL}} \sigma_{\phi}^2}{\alpha(k) \sigma_R^2} \right) \left( q + \tilde{f}^2_{\text{NL}} I_1(k) q P_{\psi,G}(k) \right) \left( 1 + \tilde{f}^2_{\text{NL}} I_1(k) q P_{\psi,G}(k) \right)
\] (4.38)

and,

\[
\frac{P_{hh}(k)}{P_{mm}(k)} = \frac{b_\phi^2 (1 - q)}{1 + \tilde{f}^2_{\text{NL}} I_1(k) q P_{\psi,G}} \\
+ \left( b_\psi + 2 \delta_c(b_\psi - 1) \frac{\tilde{f}_{\text{NL}} \sigma_{\phi}^2}{\alpha(k) \sigma_R^2} \right)^2 \left( q + \tilde{f}^2_{\text{NL}} I_1(k) q P_{\psi,G}(k) \right) \left( 1 + \tilde{f}^2_{\text{NL}} I_1(k) q P_{\psi,G}(k) \right)
\] (4.39)

For each model (listed in Table 4.1) we have measured the auto and cross correlations in the matter and halo fields, using halos in the same mass bins and redshifts shown in Table 4.2. To check how well the analytic expressions above fit the simulation results, we measured the bias coefficients \(b_\phi\) and \(b_\psi\) for each case by cross-correlating the halo density field for each sample with

i. the \(\phi_{G}(\vec{x})\) part of the linear density field, and

ii. the \(\psi_{G}(\vec{x}) + \tilde{f}_{\text{NL}} \psi_{G}(\vec{x})^2\) part of the linear density field.

For the correlation with the Gaussian field, we expect a constant large-scale bias \(b_\phi\).

For the second case we expect a scale-independent piece and a scale dependent term. The expression for this bias is obtained using Eq.(B.11) by cross correlating \(\delta_h\) and \(\delta_{\psi,\text{NG}}\), to get:

\[
\frac{P_{h\psi}}{P_{\psi\psi}} = b_\psi + \frac{2 \delta_c(b_\psi - 1) \tilde{f}_{\text{NL}} \sigma_{\phi}^2}{\alpha(k) \sigma_R^2}
\] (4.40)

Equations (4.38) and (4.39) clearly give the expected result for the Gaussian case in the \(q = 0\) limit; in this case only the \(\phi\) field contributes to the initial density field. We obtain the single field non-Gaussian (hierarchical, feeder, or mixed) limit for \(q = 1\), in which case there is no contribution from the \(\phi\) field to the initial density field. Further, in case of single field hierarchical models,
\[
\left( \frac{1 + \hat{f}_{NL}^2 \mathcal{P}_{\psi,G}(k) I_1(k)}{1 + \hat{f}_{NL}^2 q \mathcal{P}_{\psi,G}(k) I_1(k)} \right) = 1, \quad \text{and} \quad \frac{\sigma_{R,\psi}^2}{\sigma_R^2} \approx q \quad (4.41)
\]

and we recover the known results for the local ansatz. Using the general expressions in Eq.(4.38) and Eq.(4.39) and the best fit bias parameters for each source field, we will compare the semi-analytical predictions for the bias and the stochasticity parameter \(r^2(k)\) to our simulation results.

Now, let us discuss the single field hierarchical scenario first. Note that we use \(\delta_c = 1.46\), the best fit \(\delta_c\) from our mass function fits. We find that a different best fit \(b_\psi\) is preferred compared to the corresponding Gaussian bias \(b_g\) measured from Gaussian simulations. This can be clearly seen in Figure 4.9, especially in the case of \(\mathcal{H}500\). We checked this to be true for other halo samples listed in Table 4.2. In general, we find the best fit \(b_\psi\) is less than the corresponding \(b_g\); this is consistent with the picture that bias decreases as mass function increases [214]. From Figure 4.9 (right plot), we can also verify that the single field non-Gaussian cases do not produce excess stochasticity than the Gaussian case as predicted. For the two more massive halos samples at \(z = 0\), for which the Gaussian stochasticity itself deviates from unity, we find similar level of deviation from unity in the non-Gaussian case.

In Figure 4.10, we show bias and stochasticity results for two of our feeder cases. We find that for the cases in which the corresponding \(r^2_{\text{gaus}}\) is consistent with unity, the simulation results for the stochasticity are described quite well by the analytic expression. In Figure 4.11, we show another example of a feeder case \(F70\) and an example for a mixed case \(M997\). In Table 4.3, we list the values \(b_\psi\) and \(b_\phi\) used in the plots.

We have only shown bias and stochasticity results for three samples out of the six samples listed in Table 4.2. Let us comment on the results obtained for the samples for which results are not presented here. First, we find that the bias \(P_{hm}/P_{mm}\) results agree with the analytic prediction for all the samples. Similarly, the stochasticity \(r^2(k)\) for the \(z = 1\) sample with mass range \(9.65 \times 10^{13} h^{-1} M_\odot \leq M \leq 1.92 \times 10^{14} h^{-1} M_\odot\) has excellent fits to the simulation data for all the feeder and mixed models. However, we find that the stochasticity \(r^2(k)\) results for the two samples at \(z = 0\) with mass ranges: (i) \((0.965 \leq M \leq 1.92) \times 10^{14} h^{-1} M_\odot\) and (ii) \((1.93 \leq M \leq 3.85) \times 10^{14} h^{-1} M_\odot\) show deviations from the analytic predictions, roughly at the same level as the deviation shown by corresponding Gaussian \(r^2_{\text{gaus}}\) from unity (at \(\approx 10\%\) level).

Reference [212] found that the stochastic bias predictions for one case of a two field model were off by roughly a factor of 0.7. For comparison, we define their \(b_0\) in our
**Figure 4.10**: Top: The bias $P_{hm}/P_{mm}$ (left) and stochasticity $r^2(k)$ (right) at large scales for the feeder simulation F677 at $z = 0, 1, 2$ using halos in the mass range $4.83 \times 10^{13} h^{-1} M_{\odot} \leq M \leq 9.55 \times 10^{13} h^{-1} M_{\odot}$. The analytical curves for the bias is obtained using Eq.(4.38) for which the bias coefficients $b_\psi$ and $b_\phi$ are measured by cross-correlating each halo sample with the linear density field contribution from the $\phi$ and $\psi$ fields respectively. The analytical curves for the stochasticity is obtained using Eq.(4.38) and Eq.(4.39) in Eq.(4.37) using the same $b_\phi$ and the same $b_\psi$ as the corresponding bias curve on the left plot. Bottom: Same as top but for the model F215 using the same halo samples. All the analytical curves for the bias are consistent with the simulation results. For the stochasticity $r^2(k)$, the most discrepant case in the above plots (the F677, $z = 0$ sample) is off by about five percent.
Figure 4.11: Top: The bias $P_{hm}/P_{mm}$ (left) and stochasticity $r^2(k)$ (right) at large scales for the feeder simulation F70 at $z = 0, 1, 2$ using halos in the mass range $4.83 \times 10^{13} h^{-1} M_\odot \leq M \leq 9.55 \times 10^{13} h^{-1} M_\odot$. The analytical curves for the bias are obtained using Eq.(4.38) for which the bias coefficients $b_\psi$ and $b_\phi$ are measured by cross-correlating each halo sample with the linear density field contribution from the $\phi$ and $\psi$ fields respectively. The analytical curves for the stochasticity are obtained using Eq.(4.38) and Eq.(4.39) in Eq.(4.37) using the same $b_\phi$ and the same $b_\psi$ as the corresponding bias curve on the left plot. Bottom: same as top but for the model M997 using the same halo samples.
Table 4.3: Values of the bias coefficients $b_\phi$ and $b_\psi$ measured by cross correlating the halo density field with the corresponding $\phi$ and $\psi$ components in the linear density field. In all redshifts $z = 0, 1, 2$ listed above, the mass range for the halo samples used was $4.83 \times 10^{13}h^{-1}M_\odot \leq M \leq 9.55 \times 10^{13}h^{-1}M_\odot$. These same halo samples are used for the bias and stochasticity plots shown in Figures 4.10 and 4.11.

\[
\begin{array}{cccc}
\text{Model} & \text{Redshift, } z & b_\phi & b_\psi \\
F677 & 0 & 1.76 \pm 0.05 & 1.85 \pm 0.05 \\
F215 & 0 & 1.83 \pm 0.05 & 2.08 \pm 0.13 \\
F70 & 0 & 1.88 \pm 0.04 & 2.25 \pm 0.23 \\
M997 & 0 & 1.87 \pm 0.08 & 1.95 \pm 0.06 \\
F677 & 1 & 3.34 \pm 0.19 & 4.15 \pm 0.09 \\
F215 & 1 & 3.96 \pm 0.14 & 5.04 \pm 0.10 \\
F70 & 1 & 4.18 \pm 0.11 & 5.09 \pm 0.23 \\
M997 & 1 & 2.95 \pm 0.09 & 3.83 \pm 0.06 \\
F677 & 2 & 4.49 \pm 0.31 & 9.24 \pm 0.17 \\
F215 & 2 & 6.36 \pm 0.63 & 14.36 \pm 0.48 \\
F70 & 2 & 7.88 \pm 0.55 & 15.08 \pm 0.96 \\
M997 & 2 & 4.16 \pm 0.46 & 7.18 \pm 0.11 \\
\end{array}
\]

For their particular model, $b_0 \approx \frac{1}{2}(b_\phi + b_\psi)$, and their model assumes $b_0 = b_\phi = b_\psi$. However, our results indicate that we may have to relax this assumption in general two source non-Gaussian scenarios. We expect better match between simulation results and peak background split calculation in their work too, once $b_\psi$ and $b_\phi$ are measured separately, in addition to accounting for the $O(\tilde{f}^3_{NL})$ term in the bispectrum, and the $O(f^4_{NL})$ term in the trispectrum.

### 4.3.4 Conclusion

The peak-background-split calculations are good fits to the non-Gaussian bias results ($P_{hm}(k)/P_{mm}(k)$) from simulations—both for the hierarchical models (previously done a number of times) and for the feeder models. Moreover, we also presented results for large scale stochastic bias from our simulations which, in addition to the halo-matter bias, also depends on $P_{hh}(k)/P_{mm}(k)$. We found that, for two source scenarios, both the bias and stochastic bias calculations work well once we allow for different bias coefficients for the two independent fields. These bias coefficients, namely $b_\phi$ and $b_\psi$ in our notation,
were measured by cross-correlating the halo density field with the linear density field contributions from $\phi$ and $\psi$ fields separately. The analytical calculations deviated from the simulation results for the stochastic bias only for halo samples whose values of stochasticity were inconsistent with unity in the Gaussian simulations. Therefore, for these halo samples, we expect the need to account for other contributions to the stochasticity that have to be taken into account even in the Gaussian case. While we have not investigated this issue in detail, additional studies would be worthwhile since a better understanding of stochasticity is useful for cosmological applications of galaxy surveys.

4.4 Position-dependent bispectrum

4.4.1 Introduction

A key property of any correlation function in the density fluctuations is the degree to which the local statistics can differ from the global statistics due to coupling between local Fourier modes and long wavelength background modes. For example, the density power spectrum amplitude in subolumes of a survey may be correlated with the local overdensity of the sub-volume [215]. This observable goes by the name of “position-dependent power spectrum” and is a measure of an integrated bispectrum that gets most of its contribution from the squeezed limit of the bispectrum. It is a probe both of non-linear structure formation and of primordial correlations in the curvature fluctuations. The position-dependent power spectrum is easier to measure than directly measuring the bispectrum, and a measurement of the position-dependent correlation function in real space has been recently performed from the SDSS-III data in [216].

In this work, we consider the generalization of the position-dependent power spectrum to higher order correlations. In particular, we focus on the position-dependent bispectrum, which is a measure of an integrated trispectrum. We will limit this initial analysis to position-dependence in the amplitude of the equilateral configuration both for simplicity and as an interesting configuration to probe qualitative properties of any primordial non-Gaussianity. We obtain the expected constraints on a large family of primordial trispectra (including $g_{\text{NL}}^{\text{local}}$) as well as on bias parameters using the Fisher formalism for the proposed SPHEREx (Spectro-Photometer for the History of the Universe, Epoch of Reionization, and Ices Explored) [49] galaxy survey.

Constraints on a few trispectrum shapes exist from both CMB and large-scale structure observations. Given the computational difficulty of searching for an arbitrary primordial trispectrum, constraints have instead been placed on just a few theoretically motivated examples. One useful case is the “local” model, where the non-Gaussian field, $\Phi_{\text{NG}}(\mathbf{x})$, is a non-linear but local function of a Gaussian random field, $\phi_G(\mathbf{x})$. The standard local “$g_{\text{NL}}$” trispectrum is generated by a term proportional to $\phi_G^3(\mathbf{x})$. The Planck mission has constrained the amplitude of this trispectrum $g_{\text{NL}}^{\text{local}} = (-9.0 \pm 7.7) \times$
Constraints from SDSS photometric quasars using the scale-dependent bias \([44]\) give \(|g_{\text{NL}}^{\text{local}}| \lesssim 2 \times 10^5 \,[47]\). However, one of the primary reasons the local type non-Gaussianity is interesting is that it significantly couples long and short wavelength modes. Any evidence of such a coupling has two important implications: it would introduce an additional source of cosmic variance in connecting observations to theory \([189, 130, 127]\), and it would rule out single-clock inflation \([217]\). While the local ansatz is an excellent example of this phenomena, it is of course not the unique example. Constraints on the position dependence of the equilateral bispectrum test for long-short mode-coupling more generally, as we will detail below. In addition, looking at the equilateral configuration is interesting since higher order correlation functions need not be consistent with single clock inflation even if the average bispectrum is.

Further, measurements of galaxy bispectrum have been done by the SDSS collaboration \([42, 43]\). Given the increasing computational difficulty in directly measuring higher order statistics, we investigate the possibility of using position-dependent bispectrum to measure the squeezed-limit of the four-point function.

This section is structured as follows. In the next subsection we introduce the idea of position-dependent \(n\)-point functions, starting with a review of the position-dependent power spectrum studied in detail in \([215, 216]\). After that, we will discuss and derive expressions for position-dependent bispectrum and angle-averaged integrated trispectrum. In 4.4.4 we discuss the signal and noise (using large-scale structure) for the measurement of primordial trispectrum amplitudes by using this method, which will be followed by discussion of our Fisher forecast method. We will report and discuss the results of our Fisher forecasts in 4.4.6, and conclude in 4.4.7.

### 4.4.2 Position-dependent N-point functions

#### 4.4.2.1 Position-dependent power spectrum

Consider a large volume in which the density fluctuation field \(\delta(x)\) is defined. Now, consider spherical sub-volumes with radius given by \(R\). (Although note that as in \([215]\), it is useful to divide into cubic sub-volumes when testing analytic results against data from N-body simulations.) The density field and power spectrum in a sub-volume centered at \(x_R\) are given by:

\[
\delta(k)_{x_R} = \int d^3x \delta(x)W_R(x-x_R)e^{-i\mathbf{x} \cdot \mathbf{k}}
\]

\[
= \int \frac{d^3q}{(2\pi)^3} \delta_{k-q}W_R(q)e^{-i\mathbf{x} \cdot \mathbf{q}}
\]

\[
P(k)_{x_R} = \frac{1}{V_R} \int \frac{d^3q_1}{(2\pi)^3} \int \frac{d^3q_2}{(2\pi)^3} \delta_{k-q_1}\delta_{k-q_2} W_R(q_1)W_R(q_2)e^{-i\mathbf{x} \cdot (\mathbf{q_1} + \mathbf{q_2})}
\]
where $V_R$ is the volume of the sub-volume and $W_R(q)$ is the window function in Fourier space. In this work, we will use the spherical top-hat as the window function, which is defined in real space as:

$$W_R(x) = \begin{cases} 1, & \text{if } |x| \leq R \\ 0, & \text{if } |x| > R \end{cases} \quad (4.45)$$

The correlation between the local power spectrum and the local mean density in each sub-volume ($\bar{\delta}_{xL} = (1/V_R)\delta(k = 0)_{xR}$), gives an integrated bispectrum

$$\langle P(k)_{xR} \bar{\delta}_{xL} \rangle = \frac{1}{V^2_R} \int \frac{d^3q_1}{(2\pi)^3} \int \frac{d^3q_3}{(2\pi)^3} W_R(q_1)W_R(-q_1)W_R(q_3)B(k - q_1, -k + q_{13}, -q_3) \equiv iB_R(k). \quad (4.46)$$

where $q_{13} \equiv q_1 + q_3$. See [215] for the details. For modes well within the sub-volume, $k \gg \pi/R$, the above expression is dominated by the squeezed limit of the bispectrum and simplifies to:

$$iB_R(k) \approx \frac{1}{V^2_R} \int \frac{d^3q}{(2\pi)^3} W^2_R(q)B(k, -k + q, -q) \quad (4.47)$$

where we have also used the Fourier transform of the equality $W^2_R(x) = W_R(x)$. In the squeezed limit $k \gg \pi/R$, the squeezed limit approximation Eq.(4.47) produces the same result as Eq.(4.46) for any separable bispectrum of the form [215]

$$B(k_1, k_2, k_3) = f(k_1, k_2, \hat{k}_1 \cdot \hat{k}_2)P(k_1)P(k_2) + 2 \text{ perm.}$$

In addition, it is useful to define the reduced integrated bispectrum,

$$ib_R(k) = \frac{iB_R(k)}{P(k)\sigma^2_R}. \quad (4.48)$$

where $iB_R(k)$ now is the angle-averaged integrated bispectrum.

### 4.4.2.2 Position-dependent bispectrum

Building on the idea of the position-dependent power spectrum, we can divide a survey volume in subsamples and measure bispectrum in individual sub-volumes centered on $x_R$. The position-dependent bispectrum is given by (note that we have used only two
wavevector arguments below as the third wavevector of the bispectrum is \( k_3 = -k_{12} \):

\[
B(k_1, k_2)_{\triangle} = \frac{1}{V_R} \left[ \prod_{i=1}^{3} \int \frac{d^3q_i}{(2\pi)^3} W_R(q_i) e^{-ix_R \cdot q_i} \right] \delta_{k_1 - q_1} \delta_{k_2 - q_2} \delta_{-k_{12} - q_3}
\]

\( (4.49) \)

The correlation of the position-dependent bispectra with the mean overdensities of the sub-volumes is given by an integrated trispectrum:

\[
\langle B(k_1, k_2)_{\triangle} \delta_{\triangle} \rangle = \frac{1}{V^2_R} \left[ \prod_{i=1}^{4} \int \frac{d^3q_i}{(2\pi)^3} W_R(q_i) e^{-ix_R \cdot q_i} \right] \langle \delta_{k_1 - q_1} \delta_{k_2 - q_2} \delta_{-k_{12} - q_3} \delta_{-q_4} \rangle = iT(k_1, k_2)
\]

\( (4.50) \)

The above equation, on the right hand side, contains the trispectrum \( T \) defined as:

\[
\langle \delta_{q_1} \delta_{q_2} \delta_{q_3} \delta_{q_4} \rangle = (2\pi)^3 \delta_D(q_{1234}) T(q_1, q_2, q_3, q_4)
\]

\( (4.51) \)

and therefore can be written as

\[
\langle B(k_1, k_2)_{\triangle} \delta_{\triangle} \rangle = \frac{1}{V^2_R} \left[ \prod_{i=1}^{3} \int \frac{d^3q_i}{(2\pi)^3} W_R(q_i) \right] W_R(-q_{123}) T(k_1 - q_1, k_2 - q_2, -k_{12} + q_1, -q_3)
\]

\( (4.52) \)

When all modes in the bispectrum are well inside the sub-volume, \( |k_{1,2}|, |k_{12}| \gg \pi/R \), the expression is dominated by the limit of the trispectrum in which one of the wave-numbers is much smaller than the others. If we use this squeezed-limit approximation and the Fourier transform of the real space equality \( W^3_R(x) = W_R(x) \), we get

\[
iT(k_1, k_2) = \frac{1}{V^2_R} \int \frac{d^3q}{(2\pi)^3} W^2_R(q) T(k_1, k_2, -k_{12} + q, -q)
\]

\( (4.53) \)

We can then compute the angle-averaged integrated trispectrum as

\[
iT(k_1, k_2) = \int \frac{d^2k_1}{4\pi} \int \frac{d^2k_2}{4\pi} iT(k_1, k_2)
\]

\[
= \frac{1}{V^2_R} \int \frac{q_1 dq_1}{2\pi} W^2_R(q_1) \int \frac{d^2q_2}{4\pi} iT(k_1, k_2, -k_{12} + q_1, -q_2)
\]

\( (4.54) \)

where we are taking \( \hat{k}_1 \equiv \hat{z} \). The integrated trispectrum is a measure of the correlation
of the three point statistics on scales smaller than the sub-volume size $R$ with the density fluctuations on the sub-volume scale $R$. That is to say, in Fourier space, the integrated trispectrum signal is dominated by configurations of connected four-point function in which one of the momenta is small.

We do not expect the squeezed-limit approximation Eq.(4.53) to give the same result as the full expression Eq.(4.52) for arbitrary trispectra. For trispectra that depend only on the magnitudes of the momenta, the angular integrals have no additional contribution and therefore the approximation can be expected to give exact result in the $q \rightarrow 0$ limit. However, when the trispectrum also depends on the angle between momenta, the approximation may not give exact result even in the squeezed limit. For the large-scale structure trispectrum $T^{(1)}$ (see Appendix B.4; see also the expression for the $F_3$ kernel there), we find that the use of the approximation gives slightly different result for the angle-averaged trispectrum (last line of Eq.(4.54)) than if we use the large-scale structure consistency relations [218, 219]. The difference, however, is not significant to change our Fisher constraints appreciably.

4.4.3 Integrated trispectrum for a primordial trispectrum

A strong motivation to study the position-dependent statistics is to be able to constrain primordial non-Gaussianity. We will first consider statistics at the level of initial conditions (and denote the Bardeen potential by $\Phi$). After that, we will work out the corresponding expressions for the galaxy density contrast $\delta_g$.

We can write the primordial trispectrum by using symmetric kernel functions as follows [220]:

$$T_{\Phi}(k_1, k_2, k_3, k_4) = g_{NL} P_{\Phi}(k_1) P_{\Phi}(k_2) P_{\Phi}(k_3) \times N_3(k_1, k_2, k_3, k_4) + 3 \text{ cyc, }$$

(4.55)

where the kernel $N_3$ is symmetric in the first three momenta. The widely studied $g_{NL}^{\text{local}}$ model is a very useful benchmark case and corresponds to $N_3(k_1, k_2, k_3, k_4) = 6$. In the squeezed limit, where one of the momenta is much smaller than the other three, the trispectrum scales as

$$T_{\Phi}^{\text{local}}(k_1, k_2, q \rightarrow 0) = \frac{3g_{NL}^{\text{local}}}{q^3} [P_{\Phi}(k_1)P_{\Phi}(k_2) + 2 \text{ cyc.}] + O(q^0)$$

(4.56)

Notice that the quantity in the square brackets is the local bispectrum, which is non-zero (and typically normalized) in the equilateral configuration but peaks on squeezed triangles. The integrated trispectrum in this case is particularly simple:

$$iT_{\Phi}(k_1, k_2)^{g_{NL}^{\text{local}}} = 6g_{NL}^{\text{local}} \sigma_{\Phi,R}^2 [P_{\Phi}(k_1)P_{\Phi}(k_2) + 2 \text{ cyc.}]$$

(4.57)
where
\[
\sigma_{\Phi,R}^2 = \frac{1}{V_R^2} \int \frac{d^3 q}{(2\pi)^3} W_R^2(q) P_\Phi(q).
\]
Notice that for small \( q \) (modes much larger than the box size), this integral is proportional to \( \int dq/q \).

Now consider a more generic trispectrum, whose leading order behavior in the squeezed limit can be schematically written as
\[
T_\Phi(k_1, k_2, k_3, q \to 0) \propto \frac{1}{q^3} \left( \frac{q}{F(k_i)} \right)^\beta B^{\text{eff}}(k_1, k_2, k_3) \tag{4.58}
\]
where \( B^{\text{eff}} \) has the properties of a bispectrum and \( F(k_i) \) is a dimension 1 function of the momenta \( k_1, k_2, k_3 \). In other words, the squeezed trispectrum is characterized by its scaling with the long wavelength mode, \( q^{\beta-3} \), and by the weights in configurations of the remaining momenta. For a fixed configuration of the bispectrum \( B^{\text{eff}} \), all trispectra with \( \beta = 0 \) will generate the same average strength of position dependence for that configuration as the \( g^{\text{local}}_{\text{NL}} \) ansatz does.

To illustrate a bit further, suppose we restrict to cases where \( \beta \geq 0 \) and study the position-dependence of the equilateral bispectrum. The leading contribution in the squeezed limit (\( k_4 = q \to 0 \)) can be expressed in terms of the kernel:
\[
T_\Phi(k_1, k_2, -k_{12} - q, q) \approx g_{\text{NL}} P_\Phi(q) P_\Phi(k_1) P_\Phi(k_2) N_3(q, k_1, k_2, -k_{12} - q) + 2 \text{ cyc}. \tag{4.59}
\]
where we have used \( P_\Phi(q) \gg P_\Phi(k_1), P_\Phi(k_2), P_\Phi(k_3) \). Now, the integrated trispectrum becomes
\[
iT_\Phi(k_1, k_2) = g_{\text{NL}} P_\Phi(k_1) P_\Phi(k_2) \int \frac{d^3 q}{(2\pi)^3} W_R^2(q) P_\Phi(q) N_3(q, k_1, k_2, -k_{12} - q)
+ 2 \text{ cyc} \tag{4.60}
\]
It is possible to find trispectra that reduce in the squeezed limit to other bispectral shapes besides the local case. For example [220] wrote down two different examples (Eq.(D3) and Eq.(D5) of that paper) that both have squeezed limits whose leading term is
\[
T_\Phi^{\text{equl}}(k_1, k_2, k_3, q \to 0) \propto \frac{1}{q^3} \left[ P_\Phi(k_1) P_\Phi(k_2) \right.
\times \left( -6 + 4 \frac{k_1 + k_2}{k_3} + 2 \frac{k_1^2 + k_2^2}{k_3^2} - 4 \frac{k_1 k_2}{k_3^2} \right) + 2 \text{ cyc.}
\left. + O\left( \frac{1}{q^2} \right) \right] \tag{4.61}
\]
Here, the term in square brackets is the equilateral bispectrum, but notice that the
strength of coupling to the background is the same as that for the local trispectrum (i.e., $\beta = 0$).

As in the case of the integrated bispectrum, it is useful to define the reduced integrated trispectrum:

$$i_{\text{T}}(k_1, k_2) = \frac{iT(k_1, k_2)}{\frac{1}{2} \left[ P(k_1) P(k_2) + 2 \text{cyc} \right] \sigma_R^2}$$  \hfill (4.62)

such that $i_{\Phi, R}^{\text{NL}} = 18 g_{\text{NL}}^\text{local}$. The subscript $\Phi$ here is to remind that the computation was performed for primordial statistics. To make use of galaxy surveys, we need to define and compute the corresponding signals for the galaxy density contrast $\delta_{g}$. Assuming linear perturbation theory, the galaxy trispectrum generated by a primordial trispectrum (to leading order) is:

$$T_{g, \Phi}(k) = b_1^4 \alpha(k_1) \alpha(k_2) \alpha(k_3) \alpha(k_4) T_{\Phi}$$  \hfill (4.63)

where,

$$\alpha(k, z) = \frac{2 D(z)}{3 H_0^2 \Omega_m k^2 T(k)}$$  \hfill (4.64)

in which $D(z)$ is the growth function and $T(k)$ is the transfer function. The matter overdensity field $\delta$ in Fourier space is related to the Bardeen potential $\Phi$ as: $\delta(k, z) = \alpha(k, z) \Phi(k)$, and therefore $P(k, z) = \alpha^2(k, z) P_{\delta}(k)$ is the linear matter power spectrum. We will often suppress the redshift dependence when considering the overdensities at a fixed redshift, as done in Eq.(4.63).

The reduced integrated trispectrum of a large-scale structure tracer (generated by a primordial trispectrum of the form Eq.(4.59)) can therefore be written as:

$$i_{\text{T}}^{\text{NL}}(k_1, k_2) \approx \frac{g_{\text{NL}}}{b_1^2 \sigma_{\delta, R}^2} \left[ P_{\delta}(k_1) P_{\delta}(k_2) + 2 \text{cyc} \right] \frac{\alpha(k_1) P_{\delta}(k_2) P_{\delta}(k_3)}{\alpha(k_2) \alpha(k_3) + 2 \text{cyc}}$$  \hfill (4.65)

which simplifies significantly if we consider the equilateral configuration of bispectra only. That is, we will take $|k_1| \approx |k_2| \approx |k_3| = k$ and $k_i \cdot k_j \approx -k^2/2$ for $i, j = 1, 2, 3$. In this limit, the kernel reduces to a number and a simple scaling:

$$N_3(q, \Delta_k) \equiv N_3(q, k_1, k_2, k_3 | k_i \cdot k_j \approx -k^2/2) = A_{\text{equil}} (q/k)^{\beta} + \ldots$$  \hfill (4.66)
where we have used $\triangle$ to denote the equilateral configuration of bispectra. The reduced integrated trispectrum for the equilateral configuration of bispectra then simplifies to:

$$it_R^{(\triangle)}(k) \approx \frac{3g_{\text{NL}}}{b_1^2 \alpha(k) \sigma_{\delta,R}^2} \times \frac{1}{V_R^2} \int \frac{d^3q}{(2\pi)^3} W_R^2(q) \frac{P_3(q) N_3(q, \triangle_k)}{\alpha(q)}$$

(4.67)

where, for modes that are much larger than the sub-volume size, the integral on the second line scales like

$$\propto \int \frac{dq}{q} q^{(\beta+2)}.$$  

(4.68)

Both the local trispectrum and trispectra of the type considered in Eq.(4.61) have the same scaling with the long mode, $\beta = 0$. The scenarios differ only in the numerical coefficient of the equilateral configuration of short wave-length modes ($A_{\text{equil}} = 6$ for the local case, $A_{\text{equil}} = 2$ for the others), so the reduced, integrated trispectra are

$$it_R^{(\triangle)}(k) = \frac{3A_{\text{PD}(\triangle)}}{b_1^2 \alpha(k) V_R^2 \sigma_{\delta,R}^2} \int \frac{d^3q}{(2\pi)^3} W_R^2(q) \frac{P_3(q) N_3(q, \triangle_k)}{\alpha(q)}$$

(4.70)

where the amplitude of the position-dependence is $A_{\text{PD}(\triangle)} = 6g_{\text{local}}^\text{local}$ for the standard local trispectrum and $A_{\text{PD}(\triangle)} = 2g_{\text{NL}}^{\text{equil}}$ for any trispectrum that generates the equilateral template in biased sub-volumes (see Eq.(4.61)). Trispectra reducing to either bispectra in the squeezed limit can have any value of $\beta$, but $\beta = 0$ is coupling of “local” strength.

To summarize, the important features of the integrated trispectrum are the configuration of the effective bispectrum considered (which is a choice made in the analysis), and the scaling $\beta$ in the integral in Eq.(4.70), which is the information about the model and a measure of how strongly the configuration is coupled to the background. In the absence of motivation for a particular model, one could ultimately constrain $\beta$ as well as the amplitude $A_{\text{PD}(\triangle)}$. In the next section we will assume coupling of the local strength ($\beta = 0$) and quote forecast constraints on the primordial trispectrum in terms of $A_{\text{PD}(\triangle)}$. The constraints we will forecast in the next section apply equally well to any scenario with $\beta = 0$. To match to a particular trispectrum, one just needs to compute $A_{\text{PD}}$.

Note that the leading position dependence of the bispectrum (even its full shape) still does not fully characterize the trispectrum. For example, the distinction between the two trispectra in [220] that both generate equilateral bispectra in biased sub-volumes is the doubly-squeezed limit of the trispectra. One of the trispectra shifts the power spectrum when two of the momenta are small whereas the other does not. (This is
related to terms that are sub-leading in the position-dependent bispectrum.) So, a
distinction between the two can be made by correlating the square of the mean sub-
volume overdensities with the power spectra: \(<P(k)_{x_R} \delta_{x_R}^2>\). The dominant contribution
from matter trispectrum in that case, in the squeezed limit, can be obtained from the
n = 2 response function \(R_2(k)\) in [221]. We will further pursue the utility of this quantity
in distinguishing the two types of primordial trispectra in a forthcoming publication.

4.4.4 Measurement in a galaxy survey

When performing measurements with galaxy surveys, we have to account for the
contribution from non-linear gravitational evolution if we are to look for signals from a
primordial trispectrum. The galaxy trispectrum for Gaussian initial conditions, and
assuming \(\delta_g = b_1 \delta + (b_2/2) \delta^2 + (b_3/6) \delta^3\), can be written as [222]:

\[
T^{(g)} = b_1^4 T^{(1)} + \frac{b_1^2 b_2}{2} T^{(2)} + \frac{b_1^2 b_2^2}{4} T^{(3)} + \frac{b_1^3 b_3}{6} T^{(4)} \tag{4.71}
\]

The expressions for each \(T^{(i)}\) can be found in Appendix B.4 or in the text of [222].
We then obtain the angle-averaged trispectra by performing the integration Eq.(4.54)
in the equilateral limit. The reduced integrated trispectra are then:

\[
\begin{align*}
\tilde{t}_R^{(1)}(k) &= \frac{1}{b_1^2} \left[ \frac{579}{98} - \frac{32 \partial \ln P_\delta(k)}{\partial \ln k} \right] \\
\tilde{t}_R^{(2)}(k) &= \frac{b_2}{b_1^2} \left[ \frac{242}{42} P_\delta(k) \left( 242 - 28 \frac{\partial \ln P_\delta(k)}{\partial \ln k} \right) \frac{\sigma_{W_L}^2}{\sigma_L^2} \right. \\
&\quad + 7 \left( \frac{98 - 11 \partial \ln P_\delta(k)}{\partial \ln k} \right) \left( 68 + 7 \frac{\partial \ln P_\delta(k)}{\partial \ln k} \right) \frac{\sigma_{R,P}^2}{\sigma_R^2 P_\delta(k)} \right] \\
\tilde{t}_R^{(3)}(k) &= \frac{b_3}{b_1^2} \left[ 1 + P_\delta(k) \frac{\sigma_{W_L}^2}{\sigma_R^2} \right] \\
\tilde{t}_R^{(4)}(k) &= \frac{b_3}{b_1^2} \left[ 3 + P_\delta(k) \frac{\sigma_{W_L}^2}{\sigma_R^2} \right] \tag{4.72}
\end{align*}
\]

where,

\[
\begin{align*}
\sigma_{W_R}^2 &= \frac{1}{V_R^2} \int \frac{d^3 q}{(2\pi)^3} W_R^2(q), \\
\sigma_{R,P}^2 &= \frac{1}{V_R^2} \int \frac{d^3 q}{(2\pi)^3} W_R^2(q) P_\delta^2(q) \tag{4.73}
\end{align*}
\]

In Figure 4.12, we show the reduced integrated trispectra in the equilateral configuration
from perturbation theory Eqs.(4.72) and from a primordial trispectrum for the
Figure 4.12: The various reduced integrated trispectra, $i t_R^{(i)}$ (the expressions are given in Eq. (4.70) and Eq. (4.72)) for a large spherical sub-volume with radius $R = 400 \text{Mpc}/h$ at $z = 1.0$. We have taken $b_1 = 1.95$, $b_2 = 0.5$, $b_3 = 0.1$.

local-type trispectrum Eq. (4.70).

For reference, we also provide the gravity-induced integrated reduced bispectrum (in the squeezed-limit approximation) and that from the quadratic bias $b_2$ [215]:

$$ib_{\text{SPT}}(k) = \frac{1}{b_1} \left[ \frac{47}{21} - \frac{1}{3} \frac{d \ln P_\delta(k)}{d \ln k} \right]$$  \hspace{1cm} (4.74)

$$ib_{b_2}(k) = \frac{b_2}{b_1^2}$$  \hspace{1cm} (4.75)

Similarly, the integrated bispectrum from $f_{NL}^{\text{local}}$ is given by:

$$ib_R(f_{NL}^{\text{local}})(k) \approx \frac{4_{NL}^{\text{local}} \sigma_{R,P}^2}{b_1}$$  \hspace{1cm} (4.76)

Primordial non-Gaussianity introduces non-Gaussian terms in the bias parameters, as convincingly demonstrated by [44]. For models with long-short mode coupling of the local strength, these new terms grow on large scales as $1/k^2$ and so the galaxy power spectrum can itself be used as a powerful constraint on $f_{NL}$ as well as $g_{NL}$, etc. [156, 223, 157]. For SPHEREx, for example, forecasts find expected $1\sigma$ errors on $f_{NL}$ to be 0.87 from the power spectrum and 0.21 from the bispectrum [49]. Here we want to focus on understanding the position-dependent bispectrum alone, so we leave the
Figure 4.13: The galaxy number density as a function of the redshift assumed in the Fisher matrix calculations. The function approximates the large galaxy count, low-accuracy redshift sample proposed for SPHEREx ($\delta_z = 0.1$, cumulative) in Figure 10 of [49].

full treatment, including non-Gaussian corrections to the bias as well as the correlated position-dependent power spectrum, for future work.

4.4.5 Fisher forecasts

We now present the Fisher forecast method using the position-dependent method. Our method follows that of [216]. Unlike in [216], we do not perform the full integration but rather use the approximate squeezed-limit expressions for the reduced integrated trispectra and bispectra. Also, for the window function we use a spherical top-hat instead of the cubic window function. To compute the matter power spectrum in our Fisher forecasts, we use fiducial cosmological parameters from the Planck 2015 results. Specifically, we use the TT+lowP+lensing (second column of Table 4 in [20]) values: $n_s = 0.968, \sigma_8 = 0.815, \Omega_m = 0.308, \Omega_b = 0.048$.

4.4.5.1 Reduced integrated bispectrum

The Fisher matrix of the reduced integrated bispectrum is

$$F_{ibR,\alpha\beta} = \sum_{z_i} N^z_{sub} \sum_R \sum_{k \leq k_{\text{max}}} \frac{\partial ib_R(k, z_i)}{\partial p_\alpha} \frac{\partial ib_R(k, z_i)}{\partial p_\beta} \frac{1}{\Delta ib_R^2(k, z_i)}$$ (4.77)

where, the sum over $R$ is a sum over the different sizes of sub-volumes radii, $N^z_{sub} = V_{z_i}/\sum_R V_R$ is the number of sub-volumes (assume equal) of each type of sub-volumes,
\[
\Delta ib_R^2(k) = \frac{1}{N_{kR}} \left[ \sigma_{R,z}^2 + \frac{P_{\text{shot}}}{V_R} \right] \left[ P_{R,z}(k) + P_{\text{shot}} \right]^2 \frac{\sigma_{R,z}^4}{\sigma_{R,z}^4 P_{R,z}^4(k)}
\]

(4.78)

in which,

\[
N_{kR} \approx 2\pi \left( \frac{k}{k_{\min}} \right)^2
\]

is the number of independent \(k\)-modes in a sub-volume [41, 224], and

\[
P_{R,z}(k) = \frac{1}{V_R} \int \frac{d^3q}{(2\pi)^3} W_R^2(q) P_z(|k - q|)
\]

(4.79)

is the convolved power spectrum, and \(P_{\text{shot}} = 1/\bar{n}_g\) is the shot noise of the galaxy sample. We assume the survey volume and redshift information to be that of the planned SPHEREx survey. In Figure 4.13, we show the galaxy number density that we are using, which approximates the low-accuracy sample in Figure 10 of [49].

4.4.5.2 Reduced integrated trispectrum

The Fisher matrix of the reduced integrated trispectrum, similarly, is

\[
F_{itR,\alpha\beta} = \sum_{z_i} N_{\text{sub}}^{z_i} \sum_R \sum_{k \leq k_{\text{max}}} \frac{\partial it_R(k, z_i)}{\partial p_\alpha} \frac{\partial it_R(k, z_i)}{\partial p_\beta} \frac{1}{\Delta it_R^2(k, z_i)}
\]

(4.80)

where, the sum over \(R\) is a sum over the different sizes of sub-volume radii used, and

\[
\Delta it_R^2(k) = 6V_R \left[ \sigma_{R,z}^2 + \frac{P_{\text{shot}}}{V_R} \right] \left[ P_{R,z}(k) + P_{\text{shot}} \right]^3 \frac{\sigma_{R,z}^4}{\sigma_{R,z}^4 P_{R,z}^4(k)}
\]

(4.81)

in which,

\[
N_{k,\Delta} \approx 8\pi^2 \left( \frac{k}{k_{\min}} \right)^3
\]

is the number of equilateral-type triangular configuration (of size \(k\)) inside each sub-volume, and 6 is the symmetry factor for the equilateral configuration [224].

4.4.5.3 Note on correlation matrix

Note that we have assumed that there is no correlation between different \(k\) modes and different sub-volumes. To see that this is a reasonable approximation, note that the dominant contribution for the matrix element \(\langle it_R(k, x) it_R(k, x') \rangle\) separated by
Figure 4.14: The smoothed two-point correlation function $\xi_R(r)$ as a function of the comoving distance $r$ (normalized by $\xi_R(0)$), for three smoothing scales $R = 200, 100, 20$ Mpc/h. The vertical lines are $r = 2R$ lines, and are plotted to show that the correlation is small for sub-volumes separated by $r > 2R$.

$|x' - x| = r$ is given by

$$
\left\langle i t_R(k, x)i t_R(k, x') \right\rangle \approx \frac{V_R}{N_{k, \Delta}} \frac{\xi_R(r)}{\sigma_R^2} \phi_R(z) P_{R,z}(k)
$$

(4.82)

where,

$$
\xi_R(r) = \frac{1}{2\pi^2} \frac{1}{V_R^2} \int q^2 dq W_R^2(q) P_z(q) j_0(qr)
$$

(4.83)

Normalizing by the value at $r = 0$, we plot the smoothed correlation function in Figure 4.14. The normalized matrix element can be approximated (for both the integrated bispectrum and integrated trispectrum) in the zero shot noise limit by $\xi_R(r)/\sigma_R^2$. In the presence of shot noise, we expect the normalized matrix element (non-diagonal) to be smaller. For each $R$, we see that the correlation is very weak when $r > 2R$, which is the smallest distance between two sub-volumes. In addition, there should be some correlation from non-Gaussian coupling to very long wavelength modes common to neighboring sub-volumes, but the scaling of the integrands in Eq.(4.47), Eq.(4.67) shows that this should be small.

We have also assumed the correlation between different wavenumbers is small. That is

$$
\left\langle i t_R(k_1)i t_R(k_2) \right\rangle \approx \delta_D(k_1 - k_2) \Delta i t_R^2(k)
$$
(and similarly for the integrated bispectrum). This approximation breaks at smaller scales and smaller redshifts due to the non-linear evolution [225] (see in particular Figure B.1. and the discussion around it in the reference). It is also worthwhile to note other important results from [225], useful in the context of constraining primordial non-Gaussianity: (i) that the cross correlation between different \( k \) values for different sizes of sub-volumes gets weaker as these contain different long wavelength modes, (ii) that having different sized sub-volumes and different redshifts is useful in breaking the degeneracy between the primordial and late-time contribution to the integrated bispectrum. This is because, the primordial integrated bispectrum signal depends on the sub-volume size (through \( \sigma_R^2 \)) and is also inversely proportional to the growth factor \( D(z) \) whereas the late time contributions are nearly independent of these. Similarly, we see that the reduced integrated trispectrum signal (primordial) has different \( z \) and \( R \) dependence compared to the late time contribution.

Note that for a constant redshift, \( \ell_b^{(f_{\text{local}}^{\text{NL}})} \) (in the squeezed limit) and \( \ell b_2 \) are both constant and therefore degenerate. It is necessary to use more than one sub-volume sizes to break this degeneracy for a single redshift bin. However, such a strong degeneracy is not present for the integrated trispectra (see Figure 4.12). Our results considering multiple redshift bins in the range 0.1 < \( z \) < 3.0, and using the number density expected for the SPHEREx survey is presented next.

### 4.4.6 Fisher forecast results

We now present results from the Fisher analysis. We will focus on the possible constraints on the non-Gaussianity amplitudes \( f_{\text{NL}}^{\text{local}} \) and \( A_{\text{PD}} \). The fiducial values we use for the Fisher parameters are: \( f_{\text{NL}}^{\text{local}} = 0 \) and \( A_{\text{PD}} = 0 \). For the SPHEREx Fisher forecasts, our fiducial values for the bias parameters are: \( b_1 = 1.95, b_2 = 0.5 \) and \( b_3 = 0 \).

#### 4.4.6.1 \( f_{\text{NL}}^{\text{local}} \) constraint from integrated bispectrum

In Figure 4.15, we plot Fisher forecast ellipses for \( f_{\text{NL}} \) and the bias coefficients using the integrated bispectrum. We can see that constraints of order \( \sigma(f_{\text{NL}}) \approx 1 \) can be obtained using the integrated bispectrum method. This is similar to the Fisher constraint obtained in [49] using the full bispectrum. However, we note that we have not optimized the sub-volume choices (nor included the non-Gaussian corrections to the bias); it may be possible to further improve the constraint.

In addition, in Table 4.4, we also list Fisher constraint on \( f_{\text{NL}} \) by considering the luminous red galaxies (LRGs) and the quasars from the extended Baryon Oscillation Spectroscopic Survey (eBOSS). We take the survey parameters, the expected number densities, and the bias parameters from [226] (See Table 2 in the reference).
Figure 4.15: Fisher forecast ellipses for two of \((f^{\text{local}}_{\text{NL}}, b_1, b_2)\) marginalized over the other, assuming SPHEREx survey volume and other parameters given above in the figure. See Figure 4.13 for the assumed galaxy number density as a function of the redshift. The two different ellipses in each plot represent different choices for sub-volumes considered: (i) dashed green – only one type of sub-volume with radius \(R = 100 \text{ Mpc}/h\); this means that the total number of sub-volumes when dividing the whole survey is large \((N = 260924)\), (ii) solid blue – two sizes of sub-volumes with \(R = 100, 1000 \text{ Mpc}/h\) in equal numbers (except for when the volume of a redshift bin is smaller than the volume of the larger sub-volume). The Fisher constraint for these two cases are: \(\sigma(f_{\text{NL}}) = 4.1, 1.2; \sigma(b_1) = 0.02, 0.15\) and \(\sigma(b_2) = 0.04, 0.33\).

<table>
<thead>
<tr>
<th>Survey</th>
<th>(R \text{ (Mpc}/h))</th>
<th>(N_{\text{sub-volumes}})</th>
<th>(\sigma(f_{\text{NL}}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPHEREx</td>
<td>100, 1000</td>
<td>1302</td>
<td>1.20</td>
</tr>
<tr>
<td>SPHEREx</td>
<td>([1, 2, 3, 4, 5] \times 100)</td>
<td>5775</td>
<td>1.71</td>
</tr>
<tr>
<td>eBOSS LRGs</td>
<td>200, 500</td>
<td>408</td>
<td>20.5</td>
</tr>
<tr>
<td>eBOSS quasars</td>
<td>200, 500</td>
<td>1750</td>
<td>54.5</td>
</tr>
</tbody>
</table>

Table 4.4: Fisher forecast results for \(f_{\text{NL}}\). In the first row, we have considered two sub-volume sizes: one large \(R = 1000 \text{ Mpc}/h\) \((N = 255)\) and one small: \(R = 100 \text{ Mpc}/h\) \((N = 1047)\). In the second row, we have used five different sub-volume sizes: \(R = 100, 200, 300, 400, 500 \text{ Mpc}/h\); there are 1155 of each of these sub-volumes.
\[ k_{\text{max}} = 0.2 \text{(Mpc}/h)^{-1} \]
\[ N_{\text{bins}}^1 = 12 \]
\[ z_{\text{min}} = 0.1, z_{\text{max}} = 3.0 \]
\[ R = 200 \text{ Mpc}/h \]
\[ N = 32615 \]
\[ R = 200, 500 \text{ Mpc}/h \]
\[ N = 3912 \]

**Figure 4.16:** Fisher forecast ellipses for two of \( (A_{PD}, b_1, b_2) \) marginalized over the other and \( b_3 \) (which is not shown), assuming SPHEREx survey volume. The different colored ellipses represent different sets of sub-volume types: (i) dashed green – only one type of sub-volume with radius \( R = 200 \text{ Mpc}/h \), (ii) solid blue – two sizes of sub-volumes with \( R = 200, 500 \text{ Mpc}/h \) in equal numbers. The Fisher constraints for these two cases are: \( \sigma(A_{PD}) = 1.43 \times 10^6, 2.33 \times 10^6 \); \( \sigma(b_1) = 0.36, 1.4 \); \( \sigma(b_2) = 0.41, 1.6 \).

### 4.4.6.2 \( A_{PD} \) constraint from integrated trispectrum

In Figure 4.16, we plot Fisher forecast ellipses for \( A_{PD} \) and the bias coefficients using the integrated trispectrum. With the same parameters that was used for \( f_{NL} \), we obtain \( \sigma(A_{PD}) \approx 10^6 \). See Table 4.5 for a list of constraints on the non-Gaussianity parameter \( A_{PD} \) and the corresponding constraint on \( g_{NL}^{\text{local}} \) for other choices of subvolume sizes. By using only the equilateral configuration of the bispectrum, we can obtain \( \sigma(g_{NL}^{\text{local}}) \approx 3 \times 10^5 \). With the addition of other configurations, we should expect improvements in the \( g_{NL} \) constraints. Note that this is different than the case of \( f_{NL}^{\text{local}} \) using integrated bispectrum in which we use all the power spectra in the “position-dependent power spectrum”. In the equilateral configuration, the total number of triangles used is roughly given by

\[ N_{\text{equil}, \Delta} \approx 2\pi^2 \left( \frac{k_{\text{max}}}{k_{\text{min}}} \right)^4. \]
Table 4.5: Fisher forecast results for $A_{PD}$. In the last column, we have translated the constraint on $A_{PD}$ to the constraint on the local-type primordial trispectrum amplitude $g_{\text{local}}^{NL}$ using $g_{\text{local}}^{NL} = A_{PD}/6$.

<table>
<thead>
<tr>
<th>Survey</th>
<th>$R$ (Mpc/$h$)</th>
<th>$N_{\text{sub-volumes}}$</th>
<th>$\sigma(A_{PD})$</th>
<th>$\sigma(g_{\text{local}}^{NL})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPHEREx</td>
<td>[1, 2, 3, 4, 5] $\times$ 100</td>
<td>5775</td>
<td>$1.81 \times 10^6$</td>
<td>$3.01 \times 10^5$</td>
</tr>
<tr>
<td>SPHEREx</td>
<td>200, 500</td>
<td>3912</td>
<td>$2.33 \times 10^6$</td>
<td>$3.89 \times 10^5$</td>
</tr>
<tr>
<td>SPHEREx</td>
<td>100, 1000</td>
<td>1302</td>
<td>$3.96 \times 10^6$</td>
<td>$6.60 \times 10^5$</td>
</tr>
<tr>
<td>eBOSS LRGs</td>
<td>200, 500</td>
<td>408</td>
<td>$1.81 \times 10^7$</td>
<td>$3.02 \times 10^6$</td>
</tr>
<tr>
<td>eBOSS quasars</td>
<td>200, 500</td>
<td>1750</td>
<td>$6.38 \times 10^7$</td>
<td>$1.06 \times 10^7$</td>
</tr>
</tbody>
</table>

If we use all the triangles possible, then the rough count of the number of triangles is

$$N_{\text{all,} \Delta} \approx \pi^2 \left( \frac{k_{\text{max}}}{k_{\text{min}}} \right)^6.$$ 

Therefore, if we assume that the ratio of the primordial contribution to the non-primordial contributions to the integrated trispectrum do not change drastically when considering non-equilateral configurations, we can roughly estimate the approximate improvement expected in the $g_{\text{local}}^{NL}$ constraint when including all triangular configurations by taking the square root of the ratio $N_{\text{all,} \Delta}/N_{\text{equil,} \Delta}$. That is roughly one expects improvement of the order $O\left( k_{\text{max}}/k_{\text{min}} \right)$; so, it is reasonable to expect an improvement to $\sigma(g_{\text{local}}^{NL})$ by a factor of 10 than what is obtained in our Fisher forecasts by using only the equilateral configuration. In that case, $\sigma(g_{\text{local}}^{NL}) \approx 10^4$ may be possible using the position-dependent bispectrum method, which is nearly a factor of 10 better than the current Planck satellite constraints.

### 4.4.7 Conclusion

We have introduced and discussed the “position-dependent bispectrum.” We have shown that it can be used as an efficient method to measure the four-point statistics in the squeezed limit, using data from galaxy surveys. Even with a small subset (equilateral configuration of the bispectra) of all the available triangles, we find $\sigma(g_{\text{local}}^{NL}) \approx 10^5$ can be obtained using a SPHEREx-like galaxy survey. Improvements of an order of magnitude can be expected by correlating the full bispectra with the subsample overdensities. For the constraint on $f_{\text{NL}}$, we find that $\sigma(f_{\text{local}}^{NL}) \approx 1$ is possible using the position-dependent power spectrum.

One goal of constraining the position dependence of the statistics like the power spectrum and bispectrum is to bound any non-Gaussian cosmic variance that may affect the translation between properties of the observed fluctuations and the particle physics
of the primordial era. This cosmic variance arises from the coupling of modes inside our Hubble volume to the unobservable modes outside. For scenarios with mode coupling of the local strength ($\beta = 0$ for the coupling of the bispectrum to long wavelength modes), the cosmic variance uncertainty can be significant even for very low levels of observed non-Gaussianity. For example, consider a universe with a trispectrum of the sort given in Eq.(4.61), that induces a bispectrum of the equilateral type in biased sub-volumes. As plotted in [220], if our Hubble volume has values of $f_{\text{NL}}^{\text{equil}} = 10$, $g_{\text{NL}}^{\text{equil}} = 5 \times 10^3$, the value of $f_{\text{NL}}^{\text{equil}}$ in an inflationary volume with 100 extra e-folds can be between 0 and 20 at $\sim 68\%$ confidence. From Table 4.5, this value of $g_{\text{NL}}^{\text{equil}} (= A_{PD}/2)$ is three orders of magnitude below our rough estimate of what can be ruled out by a SPHERE-x like survey, and so is unlikely to be reached even by including more configurations of the bispectrum. If models that can generate a trispectrum like that in Eq.(4.61) are physically reasonable, it will be hard to conclusively tie a detection of $f_{\text{NL}}^{\text{equil}}$ to single-clock inflation.

We have made several approximations here in order to convey the basic utility of the position dependent bispectrum, and there are many ways in which our analysis can be improved. In particular, we have not included complimentary information from the non-Gaussian bias, and higher order position dependent power spectrum correlations (e.g., $\langle P(k)_{x_R} \delta^2_{x_L} \rangle$), which would further distinguish trispectra configurations. To obtain the best constraint from SPHEREx survey, we should also the extension of the position-dependent bispectrum to include more general triangular configurations and optimize the selection sub-volume sizes and numbers. We will address these issues in future work.
Chapter 5

Summary of Results and Conclusion

The main theme of the thesis has been the study of the effects of primordial non-Gaussianity in the cosmic microwave background (CMB) and the large-scale structure (LSS) of the Universe. In the process, we have discussed many important concepts including the homogeneity and isotropy of the Universe at large scales, clustering of dark matter halos, the cosmic microwave background dipole.

In this final chapter, we will summarize the important conceptual points and the results. We will end the chapter with notes on possible future work that can extend and improve upon the ideas presented in the thesis.

5.1 Summary of Results

In the first part of Chapter 3, we introduced and discussed the CMB power asymmetry. We used the Planck 2013 temperature fluctuations data and quantified the power asymmetry in the temperature fluctuations using a real-space local-variance method. Further, we discussed the power asymmetry signal due to the Doppler dipole i.e. the dipole in the CMB temperature due to our local motion with respect to the CMB. We accounted for this expected signal and computed the residual power asymmetry signal at large CMB scales. At smaller CMB scales, by accounting for the Doppler signal, we marginally improved the constraint on the power asymmetry.

In the second part of Chapter 3, we studied in detail the effect of superhorizon modes on the statistical anisotropy (focusing on the hemispherical power asymmetry) if the primordial fluctuations had non-Gaussian statistics. We showed how a probability distribution for a power asymmetry amplitude can be calculated for any kind of non-Gaussian primordial fluctuations. More importantly, we explicitly illustrated the connection between non-Gaussianity and statistical anisotropy. We showed that these two are not independent if the non-Gaussian fluctuations are generated in a larger
volume than that can be observed. Further, our work showed that a scale-dependent local-type non-Gaussianity can explain two of the CMB anomalies—the hemispherical power asymmetry and the power suppression—at large scales, without the need for any other exotic assumption about the physics in the early universe or about the presence of large inhomogeneity across our horizon. Therefore, we proposed that a more careful combined analysis of non-Gaussianity and such signals of statistical anisotropies (as the effects of non-Gaussian perturbations) could potentially tell us if these anomalies are effects of mode coupling between superhorizon modes and the large-scale modes that we observe.

In Chapter 4, we studied the effect of a two-field (one Gaussian and the other non-Gaussian) model of primordial local-type non-Gaussianity. We found the scalings of the higher moments, which can be obtained by changing the relative power of the Gaussian and non-Gaussian fields, affected the mass function and clustering of massive halos. The analytic calculations were tested using dark matter N-body simulations ran with initial conditions with varying degree of relative scaling of higher moments. We found excellent agreement and calibrated some of our results whenever necessary. Our calibrated mass function results have already been used to obtain non-Gaussianity constraints assuming two different scaling of higher moments from X-ray cluster data; see Figure 4.7.

We then studied the effect of non-Gaussianity on the bias $b = P_{hm}/P_{mm}$ as a function of scale $k$. First, we could reproduce previous results for the local model of primordial non-Gaussianity generated using a single field. Second, we found that the simulation results and the analytical calculation showed good match for our two field models after calibrating for the $\delta_c$, but with two bias parameters for the two fields involved; these however were not fitting parameters but could be obtained by cross-correlating with the matter density field of our simulations.

Further, in the two-field non-Gaussian model, we studied the stochastic bias that is potentially useful in distinguishing between single and multi field models of inflation. The stochastic bias at large scales goes to unity for Gaussian and single-source non-Gaussian initial conditions. For non-Gaussian initial conditions with more than one fields, the stochasticity gives the fraction of power from the non-Gaussian field in the initial conditions. Our simulation results verified this behavior. However, to quantitatively match the simulation results, we identified the need to include a separate bias parameter for each field, when multiple fields are used to generate the non-Gaussian initial conditions. Previous quantitative tests of the non-Gaussian halo stochasticity parameter on simulations only used a single bias parameter for the bias that had resulted in unexplainable mismatch of analytical predictions to simulation results.

As the final section of the chapter on the effects of primordial non-Gaussianity on the large-scale structure, we discussed the position-dependent bispectrum. The measurement of a trispectrum (four-point function), while interesting, is computationally challenging;
measurements of bispectrum in the large-scale structure, however, have begun. Therefore, we studied the expected constraint on the squeezed limit of a primordial trispectrum by dividing a large volume survey into subvolumes and correlating bispectrum and mean density of subvolumes.

5.2 Future Directions

On the CMB front: The work done on the effect of primordial non-Gaussianity in the observed CMB statistics (in the presence of superhorizon modes) has the potential to be extended in many useful directions. We have briefly mentioned these future directions earlier in the individual chapters. Here, we provide a more detailed description of the future directions that are useful.

1. It is useful to perform the numerical checks with more realistic CMB realizations i.e. beyond the Sachs-Wolfe approximation. The algorithms to generate these realizations exist. But the current publicly available maps do not include the monopole and dipole of the primordial gravitational potential, and therefore ignore the effect of mode coupling between superhorizon modes and the observed CMB modes. However, the algorithm can be implemented with some effort, and new CMB realizations can be generated that include the effect of superhorizon modes. These can then be used to study and test a wide range of (super) cosmic variance effects studied in this thesis and in previous works. Further, generating realistic CMB maps is necessary for CMB data analysis.

2. As we have advocated in Chapter 3, it would be useful to perform a CMB data analysis that searches for non-Gaussianity but including the signal due to CMB power asymmetry; the hope is that it may provide a tighter constraint on non-Gaussianity and/or provide evidence in support or against the non-Gaussian origin of the CMB power asymmetry anomaly.

3. It is useful to look for similar effects of monopole, dipole and higher-order modulations in the CMB polarization data for the same reason.

4. It is useful to extend the work to include other types of non-Gaussian primordial fluctuations than the one considered in this thesis i.e. beyond the simple local and scale-dependent local type bispectra. That is, we should look at the effect of other more general shapes of bispectra (particularly promising for the scale dependence of the power asymmetry) and at the effect of higher order non-Gaussianities (trispectrum for example).

On the LSS front:
1. It is useful to consider beyond the equilateral configuration of the position-dependent bispectrum studied in Section 4.4 of Chapter 4. This will improve the constraint that can be obtained from galaxy surveys on the squeezed limit of the primordial trispectrum.

2. The position-dependent power spectrum and the position-dependent bispectrum can be employed efficiently for a future large-volume survey like the SPHEREx survey to obtain constraint on local-type primordial non-Gaussianities. Therefore, a more careful work with the planned future surveys in consideration is useful.

3. On the clustering statistics, it is useful to further explore large-scale stochastic bias as a signature of non-Gaussian signal generated from multiple degrees of freedom during inflation. That is the possibility that we can jointly constrain non-Gaussianity and the presence of multiple fields during inflation through large-scale structure clustering measurements.
Appendix A

Calculations for CMB power asymmetry and non-Gaussianity

A.1 Conventions and definitions

We follow the following Fourier convention:

\[
\phi(x) = \int \frac{d^3k}{(2\pi)^3} e^{ik \cdot x} \phi(k);
\]
\[
\phi(k) = \int d^3x e^{-ik \cdot x} \phi(x).
\] (A.1)

As usual, the temperature fluctuations of the CMB are decomposed into spherical harmonics:

\[
\Delta T_T(\hat{n}) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\hat{n}),
\] (A.2)

with the spherical harmonics normalization given by

\[
\int d\Omega \hat{n} Y_{\ell m}^* (\hat{n}) Y_{\ell' m'}(\hat{n}) = \delta_{\ell \ell'} \delta_{mm'}.
\] (A.3)

Therefore, the multipole coefficients \(a_{\ell m}\) of the CMB temperature fluctuations are given by

\[
a_{\ell m} = \int d\Omega \hat{n} Y_{\ell m}^* (\hat{n}) \frac{\Delta T_T (\hat{n})}{T}.
\] (A.4)

For Sachs-Wolfe temperature fluctuations \((\Delta T_T = -\Phi/3)\) and when the Bardeen potential \(\Phi\) is a Gaussian field \((\Phi(x) = \phi(x))\), the Sachs-Wolfe \(a_{\ell m}\) is given by

\[
a_{\ell m} = -\frac{4\pi}{3} i^\ell \int \frac{d^3k}{(2\pi)^3} \phi(k) j_\ell (kx) Y_{\ell m}^*(\hat{k}).
\] (A.5)
From this, we can obtain the Sachs-Wolfe angular power spectrum $C_\ell$ defined as:

$$
\langle a_\ell m a_{\ell' m'}^* \rangle = \delta_{\ell \ell'} \delta_{mm'} C_\ell
$$

where we have defined the power spectrum $P_\phi(k)$ as

$$
\langle \phi(k)\phi(k') \rangle = (2\pi)^3 \delta_D(k + k') P_\phi(k), \quad (A.7)
$$

or equivalently, $\langle \phi(k)\phi^*(k') \rangle = (2\pi)^3 \delta_D(k - k') P_\phi(k)$, and the dimensionless power spectrum $P_\phi$ is defined as $P_\phi(k) = 2\pi^2 P_\phi(k)/k^3$.

### A.2 Derivation of modulated power spectrum

For a homogeneous and isotropic cosmology, the two-point correlation function depends only on the magnitude of the separation between the two points and is the Fourier transform of the power spectrum:

$$
\langle \Phi(x - x/2) \Phi(x + x/2) \rangle = \int \frac{d^3k}{(2\pi)^3} P_\Phi(k) e^{ik\cdot x}. \quad (A.8)
$$

A hemispherical power asymmetry (a dipole modulation) cannot be described by allowing the power spectrum to be anisotropic in the usual way, $P(k) \rightarrow P(\vec{k})$. (Since the fluctuations are real, the Fourier modes of $\vec{k}$ and $-\vec{k}$ are related, which forbids a dipole modulation. See, e.g., [227]). Instead, if the amplitude of the coincident two-point function $\langle \phi(x)\phi(x) \rangle$ varies spatially, we can likewise introduce position dependence into the power spectrum, $P_\Phi(k) \rightarrow P_\Phi(k, x)$. Both even and odd multipole modulations can now be incorporated.

This inhomogeneous power spectrum is well defined for Fourier modes with wavelengths that are much smaller than the length scale on which the power spectrum changes because these modes have unambiguous and constant wavelengths even if their amplitude varies spatially. In contrast, a plane wave with an amplitude that varies significantly within one wavelength doesn’t have a well-defined amplitude or wavelength. Therefore, $P_\Phi(k, x)$ is only defined for wavenumbers such that $P_\Phi(k, x) \simeq P_\Phi(k, x + (2\pi/k^2)x)$. On a similar note, defining

$$
\langle \Phi(x - x/2) \Phi(x + x/2) \rangle = \int \frac{d^3k}{(2\pi)^3} P_\Phi(k, x) e^{ik\cdot x} \quad (A.9)
$$

only makes sense if $P_\Phi(k, x) \simeq P_\Phi(k, x \pm x)$; otherwise, it is futile to describe the two-point function $\langle \Phi(x - x/2) \Phi(x + x/2) \rangle$ in terms of a single power spectrum.

In this appendix we show that a spatially varying power spectrum defined as in
Eq. (A.9) arises naturally in non-Gaussian scenarios where short-wavelength modes are coupled to long-wavelength modes. For example, suppose the statistics in some large volume $V_L \to \infty$ are described by the local ansatz in real space: $\Phi(\mathbf{x}) = \phi(\mathbf{x}) + f_{NL} (\phi(\mathbf{x})^2 - \langle \phi(\mathbf{x})^2 \rangle)$ where $\phi$ is a Gaussian random field. Fourier modes of the non-Gaussian field are related to those of the Gaussian field by

$$\Phi(\mathbf{k}) = \phi(\mathbf{k}) + f_{NL} \int \frac{d^3 q}{(2\pi)^3} [\phi(\mathbf{k} - \mathbf{q})\phi(\mathbf{q}) - \langle \phi(\mathbf{k} - \mathbf{q})\phi(\mathbf{q}) \rangle]. \quad (A.10)$$

The two-point correlation of the non-Gaussian field is

$$\langle \Phi(\mathbf{x} - \mathbf{x}/2) \Phi(\mathbf{x} + \mathbf{x}/2) \rangle = \int \frac{d^3 k}{(2\pi)^3} \frac{d^3 k'}{(2\pi)^3} \langle \Phi(\mathbf{k})\Phi(\mathbf{k}') \rangle e^{i(k+k') \cdot \mathbf{x}} e^{i(k-k') \cdot \mathbf{x}/2}, \quad (A.11)$$

$$= \int \frac{d^3 k}{(2\pi)^3} P_\phi(k) e^{i k \cdot \mathbf{x}}$$

$$+ f_{NL} \int \frac{d^3 k}{(2\pi)^3} \int \frac{d^3 k'}{(2\pi)^3} \langle \Phi(\mathbf{k}') \phi(\mathbf{p}) \phi(\mathbf{k} - \mathbf{p}) \rangle + \int \frac{d^3 p'}{(2\pi)^3} \langle \phi(\mathbf{k}) \phi(\mathbf{p}') \phi(\mathbf{k}' - \mathbf{p}') \rangle \right]$$

$$+ \mathcal{O}(f_{NL}^2 P_\phi^2).$$

When we consider statistics entirely within an infinite volume, the terms proportional to a single power of $f_{NL}$ vanish because $\phi$ is a Gaussian field. In that case, the power spectrum is corrected only by the last term (proportional to $f_{NL}^2$), which is small when non-Gaussianity is weak.

Now, suppose we instead consider the statistics in a single sub-volume. Each sub-volume sits on top of a single realization of modes with wavelengths the size of the sub-volume or larger. So, for example, if mode $\mathbf{p}$ corresponds to a long-wavelength mode while modes $\mathbf{k}$ and $\mathbf{k}'$ are well within the sub-volume, then

$$\langle \phi(\mathbf{k}')\phi(\mathbf{k})\phi(\mathbf{p}) \rangle |_{\text{sub-volume}} = \phi(\mathbf{p}) \langle \phi(\mathbf{k}')\phi(\mathbf{k}) \rangle |_{\text{sub-volume}}. \quad (A.12)$$

That is, $\phi(\mathbf{p})$ takes a particular value in the sub-volume, while $\phi(\mathbf{k})$ and $\phi(\mathbf{k}')$ are still randomly distributed. Then, taking Eq. (A.11) and considering a Fourier mode $\phi(\mathbf{p})$ to be stochastic only if $p = |\mathbf{p}| > k_{\text{min}}$:

$$\langle \Phi(\mathbf{x} - \mathbf{x}/2) \Phi(\mathbf{x} + \mathbf{x}/2) \rangle |_{\text{sub-volume}} = \int |\mathbf{k}| > k_{\text{min}} \frac{d^3 k}{(2\pi)^3} e^{i k \cdot \mathbf{x}} P_\phi(k).$$
\[ + 2f_{\text{NL}} \int \frac{d^3k}{(2\pi)^3} P(\phi(k)) e^{i k \cdot x} \int_{|p| < k_{\text{min}}} \frac{d^3p}{(2\pi)^3} \phi(p) [e^{i p \cdot (x + x/2)} + e^{i p \cdot (x - x/2)}] \]
\[ + O(f_{\text{NL}}^2 P_{\phi}^2), \]
\[ = \int_{|k| > k_{\text{min}}} \frac{d^3k}{(2\pi)^3} e^{i k \cdot x} P(\phi(k)) \left[ 1 + 4f_{\text{NL}} \int_{|p| < k_{\text{min}}} \frac{d^3p}{(2\pi)^3} \left[ \phi(p) \cos \left( \frac{p \cdot x}{2} \right) \right] e^{i p \cdot x} \right]. \]

(A.13)

As discussed above, \( P(\phi(k, x)) \) can be defined in terms of the two-point function
\[ \langle \Phi(x - x/2, \Phi(x + x/2) \rangle \]
only if \( P(\phi(k, x) \simeq P(\phi(k, x \pm x)) \). In Eq. (A.13), we see that the spatial variation of the power spectrum arises from the \( e^{i p \cdot x} \) factor, which implies that \( P(\phi(k, x) \) will be nearly constant on scales that are shorter than \( 1/|p| \), i.e. scales that are well within the sub-volume. Therefore, we should restrict the two-point function to separations such that \( p \cdot x \ll 1 \), in which case the cosine term in Eq. (A.13) is approximately unity. Then Eq. (A.9) implies that
\[ P_{\phi, S}(k, x) \simeq P(\phi(k) \left[ 1 + 4f_{\text{NL}} \int \frac{d^3k}{(2\pi)^3} e^{i k \cdot x} \phi(k) \right], \]
where the subscript \( \ell \) specifies that the integral is only over long wavelength modes \((k < k_{\text{min}})\).

Expanding the factor \( e^{i k \cdot x} \), we can see the effects of non-zero long wavelength modes as a multipole expansion:
\[ P_{\phi, S}(k) = P(\phi(k) \left[ 1 + f_{\text{NL}} g_{00} + f_{\text{NL}} \sum_{M = \{-1, 0, 1\}} g_{1M} Y_{1M}(\hat{n}) + \ldots \right]. \]

(A.14)

The monopole shift is not observable, so we absorb it into the coefficient to define the observed isotropic power spectrum:
\[ P_{\phi}^{\text{obs}}(k) = P(\phi(k) \left[ 1 + f_{\text{NL}} g_{00} \right] + \left( \frac{f_{\text{NL}}}{1 + f_{\text{NL}} g_{00}} \right) \sum_{M = \{-1, 0, 1\}} g_{1M} Y_{1M}(\hat{n}) + \ldots \right]. \]

(A.15)

Finally, the shift to the power spectrum also shifts the observed value of \( f_{\text{NL}} \) as defined from the local template for the bispectrum:
\[ \langle \Phi_S(k_1) \Phi_S(k_2) \Phi_S(k_3) \rangle \equiv (2\pi)^2 \delta_D(k_1 + k_2 + k_3) B^{\text{obs}}(k_1, k_2, k_3); \]

(A.16)
\[ B_{\text{local,obs}}(k_1, k_2, k_3) = 2f_{\text{NL}}[1 + f_{\text{NL}00}]P_{\phi}(k_1)P_{\phi}(k_2) + \text{sym}, \]
\[ \equiv 2f_{\text{NL}}^{\text{obs}}P_{\phi}^{\text{obs}}(k_1)P_{\phi}^{\text{obs}}(k_2) + \text{sym}. \]

Since \( P_{\phi}^{\text{obs}} = P_{\phi}[1 + f_{\text{NL}00}] \) (considering only the isotropic piece), the second and third lines imply that \( f_{\text{NL}}^{\text{obs}} = f_{\text{NL}}/[1 + f_{\text{NL}00}] \). Then, the expression for the power spectrum, including the dipole asymmetry, is
\[
P_{\phi}^{\text{obs}}(k) = P_{\phi}^{\text{obs}}(k) \left[ 1 + f_{\text{NL}}^{\text{obs}} \sum_{M = \{-1, 0, 1\}} g_{1M} Y_{1M}(\hat{n}) + \ldots \right]. \tag{A.17}
\]

The superscript “obs” on \( f_{\text{NL}} \) should also be taken to indicate that the value does not contain the contribution corresponding to the Maldacena consistency relation for single clock inflation \( [f_{\text{NL}} \propto (n_s - 1)] \), which is unobservable [120].

Related useful works that derive estimators when the primordial temperature field is modulated by an anisotropic field include [228, 229].

### A.3 Bipolar spherical harmonics

Many works quantifying the likelihood of the power asymmetry make use of bipolar spherical harmonics [230, 231]. As an aid to that analysis, we repeat the calculation of Appendix A.2, but for the \( a_{\ell m} \). In the Sachs-Wolfe approximation, the statistics of the observed CMB two-point function are simply related to the two-point correlation of the potential at the time of decoupling:
\[
\langle a_{\ell_1 m_1} a_{\ell_2 m_2}^* \rangle = \frac{1}{9} \int d\Omega_1 Y^*_{\ell_1 m_1}(\hat{n}_1) \int \frac{d^3 k_1}{(2\pi)^3} e^{i\mathbf{k}_1 \cdot \hat{n}_1 x} \int d\Omega_2 Y_{\ell_2 m_2}(\hat{n}_2) \int \frac{d^3 k_2}{(2\pi)^3} e^{-i\mathbf{k}_2 \cdot \hat{n}_2 x} \langle \Phi(\mathbf{k}_1)\Phi^*(\mathbf{k}_2) \rangle, \tag{A.18}
\]
where \( x \) is the comoving distance to the last scattering surface. When the potential is non-Gaussian according to the local ansatz, Eq.(A.10), the statistics observed in a sub-volume will depend on the realization of the long wavelength modes:
\[
\langle a_{\ell_1 m_1} a_{\ell_2 m_2}^* \rangle_{\text{sub-volume}} = \frac{1}{9} \int d\Omega_1 d\Omega_2 Y^*_{\ell_1 m_1}(\hat{n}_1) Y_{\ell_2 m_2}(\hat{n}_2) \left[ \int \frac{d^3 k_1}{(2\pi)^3} P(k_1) e^{i\mathbf{k}_1 \cdot \hat{n}_1 x} e^{-i\mathbf{k}_2 \cdot \hat{n}_2 x} \right] \tag{A.19}
\]
\[
+ 2f_{\text{NL}} \int_{|k_1| > k_{\text{min}}} \frac{d^3 k_1}{(2\pi)^3} P(k_1) \int_{|p| < k_{\text{min}}} \frac{d^3 p}{(2\pi)^3} \phi^*(p) e^{i\mathbf{k}_1 \cdot \hat{n}_1 x - i(\mathbf{k}_1 + \mathbf{p}) \cdot \hat{n}_2 x} ,
\]
\[
+ 2f_{\text{NL}} \int_{|k_2| > k_{\text{min}}} \frac{d^3 k_2}{(2\pi)^3} P(k_2) \int_{|p| < k_{\text{min}}} \frac{d^3 p}{(2\pi)^3} \phi(p) e^{i(\mathbf{p} + \mathbf{k}_2) \cdot \hat{n}_1 x - i\mathbf{k}_2 \cdot \hat{n}_2 x} ,
\]
Here $C_\ell$ is given by Eq. (3.37) as usual, and $C_{\ell_1\ell_2}^{L_1M_1}\delta_{m_1,m_2}$ are the Clebsch-Gordon coefficients.

The standard notation for the bipolar spherical harmonic expansion is

$$\langle a_{\ell_1m_1}a_{\ell_2m_2}^* \rangle = C_\ell \delta_{\ell_1\ell_2}\delta_{m_1m_2} + \sum_{LM} (-1)^m_1 C_{\ell_1\ell_2}^{L_1M_1}A_{\ell_1\ell_2}^{LM},$$

so that the local model gives

$$A_{\ell_1\ell_2}^{LM} = \frac{2f_{NL}(4\pi)(-i)^L}{4\pi(2L+1)}C_{\ell_1\ell_2}^{L_1M_1}$$

$$\times \left[ C_{\ell_1}\int_{|p|<k_{\text{min}}}(\frac{d^3p}{2\pi^3})^3 \phi^*(p)Y_{LM}(\hat{p}) + C_{\ell_2}\int_{|p|<k_{\text{min}}}(\frac{d^3p}{2\pi^3})^3 \phi(p)Y_{LM}(\hat{p}) \right].$$

### A.4 Notes on numerical realizations

In Section 3.2.3.1, we used Gaussian and non-Gaussian Sachs-Wolfe CMB maps to test some of our analytical formulas. The generation of non-Gaussian CMB maps for a more realistic CMB sky has been described in [232, 233]. Following [233], we can write the temperature multipole moments as an integration over multipole moments of the Bardeen potential as a function of comoving distance $r$:

$$a_{\ell m} = \int dr \ r^2 \Phi_{\ell m}(r) \alpha_\ell(r),$$

$$\alpha_\ell(r) = \frac{2}{\pi} \int dk \ k^2 g_\ell(k)j_\ell(kr),$$

where $g_\ell(k)$ is the transfer function of temperature in momentum space. The process of generating non-Gaussian CMB maps as detailed in [233], therefore, requires generating non-Gaussian $\Phi_{\ell m}$ at different comoving distances $r_i$, considering the covariance $\langle \Phi_{\ell_1m_1}(r_1)\Phi_{\ell_2m_2}(r_2) \rangle$, and numerically integrating over the comoving distance $r$. For a Sachs-Wolfe universe, $g_\ell(k) = -j_\ell(kr_{\text{cmb}})/3$, and [25]:

$$\alpha_\ell(r) = -\frac{2}{3\pi} \int_0^\infty dk \ k^2 j_\ell(kr_{\text{cmb}})j_\ell(kr)$$

$$= -\frac{\delta_D(r-r_{\text{cmb}})}{3r_{\text{cmb}}}.$$
Figure A.1: Example of a realization of a primordial potential $\Phi$ (Gaussian). The two maps shown above show the same realization. The one on the top has only the monopole removed (one can see the dominant dipole structure). Note that the figure shows a $\Phi$ realization and not a CMB realization for which the dipole is canceled for adiabatic perturbations. The one on the bottom has both the monopole and dipole removed and the dominant multipole is the quadrupole.

Using this, we obtain (for $\ell > 1$):

$$a_{\ell m} = -\frac{1}{3} \Phi_{\ell m}(r_{\text{cmb}}).$$ (A.24)

Therefore, we can generate the local non-Gaussian CMB Sachs-Wolfe temperature anisotropies simply using Eq.(3.36). The maps used in our study were generated using the HEALPIX software [60], with $\ell_{\text{max}} = 300$ for the input $C_\ell$s and $N_{\text{side}} = 128$ for map-making. With the 10000 Sachs-Wolfe simulated CMB maps for Gaussian and non-Gaussian potentials, we can perform a number of correlation tests in the Gaussian and non-Gaussian maps to understand what generates the power asymmetry.

In Figure A.2, we show three plots to illustrate some useful correlations among
Figure A.2: Some useful correlations among quantities in the Gaussian and non-Gaussian numerical maps. Top: The relation between the $C_\ell$ asymmetry in three orthonormal directions for 1000 of our simulated Gaussian maps (x-axis) and the corresponding local non-Gaussian maps with $f_{NL} = 500$ (y-axis). We find a correlation coefficient of 0.5. The correlation coefficient for the corresponding $A$ values (not shown) is 0.25. Middle: The correlation between the power asymmetry amplitude and $C_1$ for Gaussian maps (blue circles) and non-Gaussian maps with $f_{NL} = 500$ (red stars). The correlation for the Gaussian case is weak, while the correlation coefficient of the power asymmetry amplitude with $C_1$ in the $f_{NL} = 500$ realizations is strong. This shows that the power asymmetry due to non-Gaussianity also depends on the value of $C_1$, a measure of the background dipole anisotropy for each map. Bottom: The power asymmetry amplitude $A$ distribution from Gaussian Sachs-Wolfe CMB maps (blue) and the corresponding power asymmetry amplitude $A$ distribution from non-Gaussian maps (red histogram, the non-Gaussianity is of local type with $f_{NL} = 500$) but with $C_1 = 0$. For comparison, the green histogram shows the distribution of $A$ with $C_1$ present in the Gaussian maps.
quantities in the simulated maps. In the top panel of Figure A.2, we plot the directional power asymmetry amplitudes $A_i$ for the Gaussian maps $f_{NL} = 0$ and non-Gaussian maps with $f_{NL} = 500$. The significant correlation between $f_{NL} = 0$ and $f_{NL} = 500$ (with a correlation coefficient, $\rho = 0.5$) is expected because both maps contain approximately the same amount of power asymmetry that comes from Gaussian cosmic variance. The $f_{NL} = 500$ CMB skies contain additional power asymmetry on top of the $f_{NL} = 0$ asymmetry. We have checked that the correlation gets weaker for larger values of $f_{NL}$.

In the middle panel, we plot $C_1$ (which we have used to model the background dipole anisotropy in density fluctuations) against the dipole asymmetry amplitude $A$ measured in both Gaussian ($f_{NL} = 0$) and local non-Gaussian ($f_{NL} = 500$) models. The $C_1$ values remain almost the same with a correlation coefficient of 0.995; this indicates that the contribution from the $O(f_{NL}^2P^2)$ term to the power spectrum is small for $f_{NL} = 500$. We find that the correlation between $C_1$ and $A$ for the Gaussian maps is very weak (correlation coefficient $= 0.035$) compared to the correlation in the $f_{NL} = 500$ model (correlation coefficient $= 0.685$). This shows that the combination of the background dipole anisotropy and non-Gaussianity is responsible for the power asymmetry in the non-Gaussian maps. To further test this notion, we also generated a set of $f_{NL} = 500$ non-Gaussian maps in which $C_1 = 0$. As we can see in the bottom panel of Figure A.2, we do not find a significant increase in the power asymmetry distribution $A$ for $f_{NL} = 500$ with $C_1 = 0$, which provides further evidence that the non-Gaussian power asymmetry is generated by a mode coupling between a background dipole anisotropy with the small-scale modes.
Appendix B

Calculations for LSS clustering statistics

B.1 Integrals for $\langle \delta_R^n \rangle_c$

Here we list some of the integrals for $\langle \delta_R^n \rangle_c = \langle \delta_R^n \rangle_1 + \langle \delta_R^n \rangle_2$, where the subscripts 1 and 2 are for the $O(\tilde{f}_{NL}^{n-2})$ and $O(\tilde{f}_{NL}^n)$ terms respectively.

\[
\langle \delta_R^2 \rangle_1 = \frac{1}{q} \int \frac{d^3 k}{(2\pi)^3} \alpha(k)^2 W_R(k)^2 P_\psi(k)
\]

\[
= \frac{1}{q} \int \frac{d k}{k} \alpha(k)^2 W_R(k)^2 P_\psi(k) \tag{B.1}
\]

\[
\langle \delta_R^2 \rangle_2 = 2 \tilde{f}_{NL}^2 \int \frac{d^3 k}{(2\pi)^3} \int \frac{d^3 \vec{p}}{(2\pi)^3} \alpha(k)^2 W_R(k)^2 P_\psi(p) P_\psi(|\vec{p} - \vec{k}|)
\]

\[
= \tilde{f}_{NL}^2 \int k^2 dk \int d p \int_{-1}^{1} d \mu \left[ \alpha(k)^2 W_R(k)^2 \frac{P_\psi(p) P_\psi(|\vec{p} - \vec{k}|)}{p|\vec{p} - k|^3} \right] \tag{B.2}
\]

The $\langle \delta_R^3 \rangle$ terms are:

\[
\langle \delta_R^3 \rangle_1 = 6 \tilde{f}_{NL} \int \frac{d^3 \vec{k}_1}{(2\pi)^3} \int \frac{d^3 \vec{k}_2}{(2\pi)^3} (\alpha W)_{1,2,12} P_\psi(k_1) P_\psi(k_2)
\]

\[
= 6 \tilde{f}_{NL} \left[ \frac{1}{2} \int \frac{d k_1}{k_1} \int \frac{d k_2}{k_2} \int_{-1}^{1} d \mu [(\alpha W)_{1,2,12} P_\psi(k_1) P_\psi(k_2)] \right] \tag{B.3}
\]

where, $(\alpha W)_{1,2,12} = \alpha(k_1) W(k_1) \alpha(k_2) W(k_2) \alpha(|\vec{k}_1 - \vec{k}_2|) W(|\vec{k}_1 - \vec{k}_2|)$.

\[
\langle \delta_R^3 \rangle_2 = 8 \tilde{f}_{NL}^3 \int \frac{d^3 \vec{k}_1}{(2\pi)^3} \int \frac{d^3 \vec{k}_2}{(2\pi)^3} (\alpha W)_{1,2,12} \int \frac{d^3 \vec{p}}{(2\pi)^3} P_\psi(p) P_\psi(|\vec{k}_1 - \vec{p}|) P_\psi(|\vec{k}_2 + \vec{p}|)
\]
Figure B.1: The cumulant estimates for the various $I_{ij}(M)$ integral from the Monte-Carlo approach. See the text and specifically Eq.(B.6) and Eq.(B.5) for the definition of $I_{ij}(M)$ plotted here. Since these integrals cover a huge range of values, we are multiplying by appropriate factors as shown in the plot for clarity. The dashed lines are the corresponding results from direct numerical integrations.

\[
\frac{\tilde{f}_{NL}^3}{8\pi^2} \int k_1^2 dk_1 \int k_2^2 dk_2 \int_1^{-1} d\mu_2 \int_0^{2\pi} d\phi_2 (\alpha W)_{1,2,12} \int \frac{dp}{p} \int_1^1 d\mu \int_0^{2\pi} d\phi \left[ \frac{P_\psi(p)P_\psi(|\vec{k}_1 - \vec{p}|)P_\psi(|\vec{k}_2 + \vec{p}|)}{|\vec{k}_1 - \vec{p}|^3 |\vec{k}_1 + \vec{p}|^3} \right]
\]

where, $\phi$ is the angle between $\vec{k}_1$ and $\vec{p}$ and $\phi_2$ is the angle between $\vec{k}_2$ and $\vec{p}$.

In the above integrals, all the $k, p$ integrals go from $k_{min}$ to $k_{max}$ and the divergent pieces in the integrands are set to be greater than $k_{min}$; for example in $\langle \delta_R^2 \rangle_1$ we set $|\vec{p} - \vec{k}| \geq k_{min}$—otherwise the integral diverges. The $\langle \delta_R^2 \rangle_1$, $\langle \delta_R^2 \rangle_2$ and $\langle \delta_R^3 \rangle_1$ integrations were evaluated using Mathematica 9’s Adaptive MonteCarlo numerical integration routine. To estimate the higher order cumulants, we used the Monte-Carlo approach described in Appendix A of [173]. When applying this method to the cumulants for which numerical integration was performed, the results agree. Following [173], we
simulate a Gaussian initial curvature $\Phi_G$ and define two fields by

$$\delta_{M,\Phi_G}(\vec{k}) = W_M(k)\alpha(k) \int d^3\vec{x}e^{-ik\cdot\vec{x}}\Phi_G(\vec{x})$$

$$\delta_{M,\Phi_G}^*(\vec{k}) = W_M(k)\alpha(k) \int d^3\vec{x}e^{-ik\cdot\vec{x}}\Phi_G(\vec{x})^2$$

Then the integral estimates are given by,

$$I_{21}(M) = \langle \delta_{M,\Phi_G}(\vec{x})^2 \rangle$$

$$I_{22}(M) = \langle \delta_{M,\Phi_G}^*(\vec{x})^2 \rangle$$

$$I_{31}(M) = 3\langle \delta_{M,\Phi_G}(\vec{x})^2\delta_{M,\Phi_G}^*(\vec{x}) \rangle$$

$$I_{32}(M) = \langle \delta_{M,\Phi_G}^*(\vec{x})^3 \rangle$$

$$I_{41}(M) = 6\langle \delta_{M,\Phi_G}(\vec{x})^2\delta_{M,\Phi_G}^*(\vec{x})^2 \rangle$$

$$I_{42}(M) = \langle \delta_{M,\Phi_G}^*(\vec{x})^4 \rangle$$

$$I_{51}(M) = 10\langle \delta_{M,\Phi_G}(\vec{x})^2\delta_{M,\Phi_G}^*(\vec{x})^3 \rangle$$

$$I_{52}(M) = \langle \delta_{M,\Phi_G}^*(\vec{x})^5 \rangle$$

In Figure B.1, we plot these values obtained through four Monte-Carlo realizations in a box of size $L = 2400\text{Mpc}/h$ and same cosmological parameters as that of our N-body simulations. The error reported is 1σ variation from four realizations. To get the value of $\langle \delta_{M}^2 \rangle_j$ for each of our models parametrized by $q$ and $\tilde{f}_{NL}$, $I_{nj}(M)$ should simply be multiplied by an appropriate factor which depends on our model parameters $q$ and $\tilde{f}_{NL}$. For example,

$$\langle \delta_{M}^2 \rangle_1 = \frac{1}{q} \frac{A^2}{A^2} I_{21}(M)$$

and

$$\langle \delta_{M}^3 \rangle_1 = \left( \frac{A^2}{A^2} \right)^2 \tilde{f}_{NL} I_{31}(M).$$

Once we have these $I_{ij}$ values, we can now also study how much do the scaling of higher moments of smoothed moments deviate from the naive expectations, Eq.(4.3) and Eq.(4.4). For that, we look at the limit of small non-Gaussianity and therefore $\langle \delta_{M}^2 \rangle \approx I_{21}(M)$. Then, for the single field hierarchical case, one simply gets:

$$\mathcal{M}_{n,M}^h \approx f_{NL}^{n-2} I_{n1}/I_{21}^{n/2},$$

and for the feeder case, one gets:

$$\mathcal{M}_{n}^f \approx \left( q_{\tilde{f}_{NL}} \right)^n I_{n2}/I_{21}^{n/2}.$$
For the feeder case, we find the scaling of higher moments of the smoothed density field is only slightly different from the expectation Eq.(4.4). We get

\[ \mathcal{M}_{n,M}^f \approx 2^{n-1} (n-1)! \left( \frac{1.32 M_3}{8} \right)^{n/3}, \]

for the range of mass scale (\( \approx 10^{13} \) to \( 10^{15} \) \( h^{-1} M_\odot \)). In the same mass range, we find that the hierarchical case, similarly, satisfies a modified relation:

\[ \mathcal{M}_{n}^h \approx 2^{n-3} n! \left( \frac{1.58 M_3}{6} \right)^{n-2}, \]

but the extra factor (here 1.58 taken near \( M = 10^{14} h^{-1} M_\odot \)) is weakly \( M \) dependent (at a few percent level).
B.2 Truncation and error

For the hierarchical scaling, we will truncate the series to $N = s$ terms and call the result the $N = s$ truncation. This will produce a series with terms of order $\mathcal{M}_3^N$. For the feeder scaling, we will truncate the series at $N = s + 2$ terms and call this the $N = s + 2$ truncation; this will produce a series with terms of order $\mathcal{M}_3^{N/3}$. For both scalings, our definition of the $N$th term in the series follows the definition in [205].

The utility of the Edgeworth series formalism lies in the fact that the PDF in Eq.(4.23) is an asymptotic series. Therefore, one can estimate the error induced by truncating the PDF to $N$th order by simply looking at the next term in the series. For the error analysis, we will adopt methods similar to that of [205] and look at the maximum $\mathcal{M}_3$ values (for both scalings) that the PDF can be computed with reasonable accuracy (20 percent) for the $\nu$ range that encompasses the halo masses that we will use from our simulation outputs. We will also look at the error as a function of $\nu$ for various $\mathcal{M}_3$ values relevant for our simulations.

We find that the truncations of the PDF for the feeder scaling generate errors of magnitude $> 20\%$ for much smaller $\nu \approx 3 - 4$ when keeping similar order terms in $\mathcal{M}_n$ compared to the hierarchical case (typically $\nu \approx 5 - 6$ for $n = 5$). From Figure B.2, we can also see that the value of $\mathcal{M}_{3,\text{max}}$ is much smaller for the feeder scaling and decreases sharply as one increases $\nu$. This means that at a higher mass range or at higher redshift, our simulation results may not be well described by our analytical formula for the non-Gaussian feeder mass function. Looking at the result for $N = 14$, 

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure_B.2.png}
\caption{Left: we plot the maximum value of $\mathcal{M}_3$ for which the PDF (for various truncations $N$) produces results within 20\% error for $\nu$ specified on the x-axis to $\nu_{\text{max}} = 2.1\nu^{0.7}$ for feeder scaling of higher moments. Right: same as left but for hierarchical scaling and with $\nu_{\text{max}} = 2.2\nu^{0.7}$.}
\end{figure}
Figure B.3: Left: The error of $N = 5$ truncation for different values of $\mathcal{M}_3$ for feeder scaling. Right: same as left but for hierarchical scaling.

we see that one gains only marginally by increasing the number of terms for the feeder scaling. So, we will adopt $N = 5$ truncation for the feeder case (i.e. up to $\mathcal{M}_3^{5/3}$) to compare with the simulation results. For the hierarchical scaling, we will adopt $N = 3$ truncation (i.e up to $\mathcal{M}_3^5$).

We can also see that the error increases as one increases $\nu$ or $\mathcal{M}_3$ (see Figure B.3). So, we expect the analytic mass function to describe simulation results better when the level of non-Gaussianity is smaller and at smaller $\nu_c$ i.e. low redshift and small halo masses.

Also note that our error discussion assumes scalings: $\mathcal{M}_n^{\text{hier}} = A_n (\mathcal{M}_3 / 6)^{n-2}$ and $\mathcal{M}_n^{\text{feeder}} = B_n (\mathcal{M}_3 / 8)^{n/3}$. But from our calculated moments, we find that the scaling for the smoothed moments is modified slightly and the higher moments are larger than expected from the simple scaling assumed in the error analysis plots (see Appendix B.1). The qualitative discussion remains the same but the magnitude of $\mathcal{M}_{3,\text{max}}$ will decrease and the relative error will increase in Figures B.2 and B.3 respectively.
B.3 Calculations for bias and stochasticity

Here we calculate the large scale bias and stochasticity expressions for our two field ansatz \((4.5)\). We take as a starting point the derivation for generic non-Gaussian scenarios given in [210]. The leading contribution to the matter-halo cross spectrum, in the long wavelength limit \((k \to 0)\) is

\[
P_{mh}(k) = b_{\psi} \left( \alpha^2(k) P_{\phi,G}(k) \right) + \alpha^2(k) P_{\psi,NG}(k) \left[ b_{\psi} + \frac{1}{\alpha(k)} \left( \frac{1}{2} (b_{\psi} - 1) \delta_c + \frac{1}{2 \ln \sigma_R} \right) F_R^{(3)} \right]
\]

\[\tag{B.7}
F_R^{(3)} = \frac{1}{P_{\phi}(k) \sigma_R^2} \int \frac{d^3{\vec{p}_1}}{(2\pi)^3} \frac{d^3{\vec{p}_2}}{(2\pi)^3} \alpha_R(p_1) \alpha_R(p_2) (\Phi(\vec{k}) \Phi(\vec{p}_1) \Phi(\vec{p}_2))_c
\]

where \(\alpha_R(k) = W_R(k) \alpha(k)\). This expression agrees with the result previously derived in [208]. For our bispectrum, Eq.\((4.11)\), there are two terms in \(F_R^{(3)}\). In the \(k \to 0\) limit, the usual term (proportional to \(\tilde{f}_{NL}\)) is

\[
F_R^{(3)} = 4 \tilde{f}_{NL} q^2 \left[ 1 + \frac{1}{f_{NL}^2 q P_{\psi,G}(k) I_1(k)} \right] \int \frac{d^3{\vec{p}}}{(2\pi)^3} \frac{\alpha_R(p)^2 P_{\phi}(p)}{[1 + f_{NL}^2 q P_{\psi,G}(p) I_1(p)]}
\]

\[\approx 4 \tilde{f}_{NL} q^2 \tag{B.8}
\]

where the second line holds only if the non-Gaussian correction to the total power is negligible. In the above expression, we have used,

\[
P_{\psi,G}(k) = \frac{q P_{\phi}(k)}{1 + f_{NL}^2 q P_{\psi,G} I_1(k)} \tag{B.9}
\]

which relates the Gaussian power in the \(\psi, G\) field to the total power, \(P_\phi\).

The second term, which is usually dropped as small in single field scenarios, is

\[
F_R^{(3)} = 8 \tilde{f}_{NL}^3 \frac{P_{\phi}(k) \sigma_R^2}{P_{\phi}(k) \sigma_R^2} \int \frac{d^3{\vec{p}_1}}{(2\pi)^3} \alpha_R(p_1) \alpha_R(|\vec{p}_1 + \vec{k}|)
\]

\[
\int \frac{d^3{\vec{p}}}{(2\pi)^3} P_{\psi,G}(p) P_{\psi,G}(|\vec{p}_1 - \vec{p}|) P_{\psi,G}(|\vec{p} + \vec{k}|)
\]

Let us now try to simplify the integral by looking at the major contributions to the integral. At large scales (small \(k\)), the value of \(F_R^{(3)}\) peaks when \(p_1\) is near the halo scale \((\approx 1/R)\). Typically, when looking at large scale bias, \(k \ll p_1\). In this squeezed limit, we find that the loop bispectrum can be well approximated by:

\[
\int \frac{d^3{\vec{p}}}{(2\pi)^3} P_{\psi,G}(p) P_{\psi,G}(|\vec{p}_1 - \vec{p}|) P_{\psi,G}(|\vec{p} + \vec{k}|)
\]
Figure B.4: Numerical test of the approximation (B.10) for $F_{R,2}^{(3)}$ at $R = 8 \, \text{Mpc}/h$. The approximation is excellent at large scales $k \lesssim 0.04 \, h \, \text{Mpc}^{-1}$. Note that a different choice for $\tilde{f}_{NL}$ and normalization for $P_{\psi,G}$ (or $q$) only rescales both curves by the same factor.

This is true because the dominant term for the integral comes from when $|\vec{p} + \vec{k}|$ and $p \approx k$ are both small. This approximation breaks down as the ratio $p_1/k$ becomes smaller. However, we find that the dependence of the left hand side integral on the angular part of $\vec{p}_1$ is symmetric around the approximate value on the right hand side even when $p_1$ is only a few times larger than $k$. Since we integrate over $\vec{p}_1$ in the $F_{R}^{(3)}$ integral, we expect the following approximation to hold quite well.

$$
F_{R,2}^{(3)} \approx \frac{8 \tilde{f}_{NL}^3}{P_{\psi}(k) \sigma_R^2} \int \frac{d^3 \vec{p}_1}{(2\pi)^3} \frac{1}{\sigma_R^2} P_{\psi,G}(p_1)^2 P_{\psi,G}(p) P_{\psi,G}(|\vec{p} + \vec{k}|)
$$

Approximation (B.10) was checked numerically using the CUBA library for multidimensional integration [234]; the result is shown in Figure B.4. Further, we also tested that the derivative term in Eq.(B.7) is indeed small compared to relevant values of $(b_y - 1) \delta_c F_{R}^{(3)}$. The derivative term for feeder scaling $\frac{d}{\delta_c \rho_R} F_{R,2}^{(3)}$ was found to be of the same order as the derivative term for hierarchical scaling $\frac{d}{\delta_c \rho_R} F_{R,1}^{(3)}$.

The approximated expression for bias can also be obtained directly from a peak
background split analysis of our model. Following the peak background split derivation of [210] (section 4.1.1), as usual, we split both our Gaussian fields into short and long wavelength modes: \( \phi_G(\mathbf{x}) = \phi_{G,l}(\mathbf{x}) + \phi_{G,s}(\mathbf{x}) \) and \( \psi_G(\mathbf{x}) = \psi_{G,l}(\mathbf{x}) + \psi_{G,s}(\mathbf{x}) \). However, we also include the \( \psi(\mathbf{x})^2 \) term and the local small scale power is given by

\[
\sigma_R^2 = \sigma_R^2 \left[ 1 + 4 \frac{\sigma_{R,\psi}^2}{\sigma_R^2} \tilde{f}_{NL} \psi l(\mathbf{x}) \left( 1 + \tilde{f}_{NL} \psi l(\mathbf{x}) \right) \right]
\]

Then, allowing for separate linear bias coefficients for our two independent fields, we get:

\[
\delta_h(\mathbf{k}) = b_\phi \delta_\phi(\mathbf{k}) + b_\psi \delta_{\psi,NG} + 2 \delta_c (b_\psi - 1) \tilde{f}_{NL} \frac{\sigma_{R,\psi}^2}{\sigma_R^2} \rho \left( \psi l(\mathbf{k}) + \tilde{f}_{NL} \int \frac{d^3 \tilde{s}}{(2\pi)^3} \psi l(\mathbf{s}) \psi l(\mathbf{k} - \mathbf{s}) \right) = b_\phi \delta_\phi(\mathbf{k}) + \left( b_\psi + 2 \delta_c (b_\psi - 1) \tilde{f}_{NL} \frac{\sigma_{R,\psi}^2}{\alpha(\mathbf{k}) \sigma_R^2} \right) \delta_{\psi,NG}
\]  

(B.11)

where

\[
\delta_\phi(\mathbf{k}) = \alpha(\mathbf{k}) \phi l(\mathbf{k}) \]  

(B.12)

\[
\delta_{\psi,NG} = \alpha(\mathbf{k}) \psi l(\mathbf{k}) + \tilde{f}_{NL} \int \frac{d^3 p}{(2\pi)^3} \alpha(\mathbf{k}) \psi l(\mathbf{p}) \psi l(\mathbf{k} - \mathbf{p})
\]  

(B.13)

and therefore, the total linear matter density field is,

\[
\delta_m(\mathbf{k}) = \delta_\phi(\mathbf{k}) + \delta_{\psi,NG}
\]  

(B.14)

This gives the expression for \( P_{hm}(k) \) to be,

\[
P_{hm}(k) = b_\phi \alpha^2(k) P_\phi(k) + \left( b_\psi + 2 \delta_c (b_\psi - 1) \tilde{f}_{NL} \frac{\sigma_{R,\psi}^2}{\alpha(\mathbf{k}) \sigma_R^2} \right) \alpha^2(k) P_{\psi,NG}
\]  

(B.15)

which agrees to the bias for our two field model given by Eq.(B.7) after using the approximations for \( F_{R,1}^{(3)} \) and \( F_{R,2}^{(3)} \) from Eq.(B.8) and Eq.(B.10). Note that the factor \( \sigma_{R,\psi}^2/\sigma_R^2 \approx q \) for the case of small non-Gaussianity; in this limit one gets the simpler expression Eq.(4.36) but since we do not always stay in the limit of small non-Gaussianity in our simulations, we compute this factor given a model specified by \( q, \tilde{f}_{NL} \) in our analysis.

Now to compute the stochastic bias, defined as

\[
r^2(k) = \frac{P_{hm}^2(k)}{P_{hh}(k)P_{mm}(k)}
\]  

(B.16)
we need the halo-halo power spectrum. The leading contributions to the halo-halo power spectrum are from our trispectrum, Eq.(4.11), there are two terms in $F_R^{(4)}$. In the $k \to 0$ limit, the usual term (proportional to $\tilde{f}_{NL}^2 q^3$) is

$$F_{R,1}^{(4)} = \frac{16 \tilde{f}_{NL}^2 q^3}{[1 + q \tilde{f}_{NL}^2 I_1(k)\mathcal{P}_\psi,G(k)]} \left[ \frac{1}{\sigma_R^2} \int \frac{d^3p}{(2\pi)^3} \frac{\alpha_R(p)^2 \mathcal{P}_\Phi(p)}{[1 + q \tilde{f}_{NL}^2 I_1(p)\mathcal{P}_\psi,G(p)]} \right]^2$$  \tag{B.17}

The other term, usually much smaller than the first term in single field cases, is

$$F_{R,2}^{(4)} \approx 48 \tilde{f}_{NL}^4 \int \frac{d^3p_1}{(2\pi)^3} \frac{d^3p_2}{(2\pi)^3} \alpha_R(p_1) \alpha_R(p_2) \int \frac{d^3\tilde{p}}{(2\pi)^3} \left[ \mathcal{P}_\psi,G(p)\mathcal{P}_\psi,G(|\vec{p}_1 - \vec{p}|)\mathcal{P}_\psi,G(|\vec{k} - \vec{p}_1 + \vec{p}|)\mathcal{P}_\psi,G(|\vec{p} - \vec{p}_1 - \vec{p}_2|) \right]$$  \tag{B.18}

Similar to the case of $F_{R,2}^{(3)}$, we can approximate this integral at large scales by looking at the collapsed limit: $k \ll p_1, p_2$. In this limit, the loop trispectrum is well approximated by:

$$\int \frac{d^3\tilde{p}}{(2\pi)^3} \mathcal{P}_\psi,G(p)\mathcal{P}_\psi,G(|\vec{p}_1 - \vec{p}|)\mathcal{P}_\psi,G(|\vec{k} - \vec{p}_1 + \vec{p}|)\mathcal{P}_\psi,G(|\vec{p} - \vec{p}_1 - \vec{p}_2|) \approx \mathcal{P}_\psi,G(p_2) \int \frac{d^3\tilde{p}}{(2\pi)^3} \mathcal{P}_\psi,G(p)\mathcal{P}_\psi,G(|\vec{p}_1 - \vec{p}|)\mathcal{P}_\psi,G(|\vec{p}_1 + \vec{k} + \vec{p}|)$$  \tag{B.19}

following from the observation that the integral gets maximum contribution from when $|\vec{p}_1 - \vec{p}|$ is small. Further, when $p_2/k$ is small, the approximation breaks down as before, but since we integrate over $p_2$ to obtain $F_{R,2}^{(4)}$, the symmetry of the integrand with respect to the angular part of $p_2$ will justify the use of this approximation in $F_{R,2}^{(4)}$ even when the halo formation scale is only a few times larger than the scale $k$ at which stochastic bias is measured.xx

With this approximation $F_{R,2}^{(4)}$ becomes,

$$F_{R,2}^{(4)} \approx 48 \tilde{f}_{NL}^4 \int \frac{d^3p_1}{(2\pi)^3} \alpha_R(p_1) \int \frac{d^3\tilde{p}}{(2\pi)^3} \mathcal{P}_\psi,G(p)\mathcal{P}_\psi,G(|\vec{p}_1 - \vec{p}|)\mathcal{P}_\psi,G(|\vec{k} - \vec{p}_1 + \vec{p}|)$$  \tag{B.20}

These integrals are computationally challenging. However, as in the case of $P_{hm}(k)$, we expect the peak-background-split calculation to provide good approximation, in the large scale limit, to the halo-halo power spectrum. For this, using Eq.(B.11), one
obtains,

\[ P_{hh}(k) = b_\psi^2(\alpha^2(k)P_\psi(k)) + \left( b_\psi + 2\delta_c(b_\psi - 1) \frac{f_{\text{NL}}}{\alpha(k)} \frac{\sigma^2 R_\psi}{\sigma_R^2} \right)^2 \alpha^2(k)P_{\psi,\text{NG}} \]

(B.21)

To summarize this section, we have derived expressions Eq.(B.15) (for \( P_{hm}(k) \)) and Eq.(B.21) (for \( P_{hh}(k) \)), which were used to fit to our simulation results for large scale bias and stochastic bias (Eq.(4.37)).

### B.4 Galaxy trispectrum expressions

Here we list the galaxy trispectrum expressions used in the text, taken from [222]. See Eq.(4.71).

\[
T^{(1)} = T_a + T_b \tag{B.22}
\]

\[
T^{(2)} = 4P_1 [F_2(k_2, k_3)P_2P_3 + F_2(k_2, -k_{23})P_2P_{23}] \\
+ F_2(k_3, -k_{23})P_3P_{23} + 4P_2 [F_2(k_1, k_3)P_1P_3 \\
+ F_2(k_1, -k_{13})P_1P_{13} + F_2(k_3, -k_{13})P_3P_{13}] \\
+ 4P_3 [F_2(k_1, k_2)P_1P_2 + F_2(k_1, -k_{12})P_1P_{12} \\
+ F_2(k_2, -k_{12})P_2P_{12}] + \text{cyc.} \tag{B.23}
\]

\[
T^{(3)} = 4P_1P_2(P_{13} + P_{14}) + \text{perm.} \\
T^{(4)} = 6P_1P_2P_3 + \text{cyc.} \tag{B.24}
\]

where

\[
T_a = 4P_1P_2P_{13}F_2(k_1, -k_{13})F_2(k_2, k_{13}) + P_{14}F_2(k_1, -k_{14})F_2(k_2, k_{14})] + \text{perm.} \\
T_b = [F_3(k_1, k_2, k_3) + \text{perm.}]P_1P_2P_3 + \text{cyc.} \tag{B.25}
\]

and from [235]

\[
F_2(k_1, k_2) = 5 \frac{1}{7} + \frac{1}{2} \frac{k_1 \cdot k_2}{k_1k_2} \left( k_1 \frac{k_2}{k_1} + k_2 \frac{k_1}{k_2} \right) + \frac{2}{7} \left( k_1 \cdot k_2 \right)^2 \tag{B.26}
\]

\[
F_3(k_1, k_2, k_3) = \frac{1}{3k_1^2k_2^2k_3^2k_{123}^2} \left[ \frac{1}{21} k_1 \cdot k_2 |k_{12}|^2 + \frac{1}{14} k_2^2k_1 \cdot k_{12} \right] \\
\times \left[ 7k_2^2k_{12} \cdot k_{123} + k_3 \cdot k_{12} |k_{123}|^2 \right] \\
+ \frac{k_1 \cdot k_{23} |k_{123}|^2}{3k_1^2k_2^2k_3^2k_{23}^2} \left[ \frac{1}{21} k_2 \cdot k_3 |k_{23}|^2 + \frac{1}{14} k_3^2k_2 \cdot k_{23} \right] \\
+ \frac{k_1 \cdot k_{123} |k_{23}|^2}{18k_1^2k_2^2k_3^2} \left[ k_2 \cdot k_3 |k_{23}|^2 + 5k_3^2k_2 \cdot k_{23} \right]. \tag{B.27}
\]
By directly taking the appropriate equilateral and soft limit \(|k_4| = q \to 0\), and after angular averaging, we can get the integrated trispectrum \(iT_{R}^{(1)}(k)\). For example, for the two terms in \(T^{(1)}\), we obtain

\[
\langle T_a(k) \rangle_{\text{angle-avg}} = P^2_\delta(k)P_\delta(q) \left[ \frac{585}{147} - \frac{4}{3} \frac{\partial \ln P_\delta(k)}{\partial \ln k} \right] \\
\langle T_b(k) \rangle_{\text{angle-avg}} = P^2_\delta(k)P_\delta(q) \left[ \frac{27}{14} - \frac{4}{21} \frac{\partial \ln P_\delta(k)}{\partial \ln k} \right] \\
\Rightarrow iT^{(1)}(k) = P^2_\delta(k)P_\delta(q) \left[ \frac{579}{98} - \frac{32}{21} \frac{\partial \ln P_\delta(k)}{\partial \ln k} \right]
\]

(B.28)
Bibliography


137


[38] E. Komatsu et al., *Non-Gaussianity as a Probe of the Physics of the Primordial Universe and the Astrophysics of the Low Redshift Universe*, 0902.4759.


[48] N. Agarwal, S. Ho and S. Shandera, Constraining the initial conditions of the Universe using large scale structure, JCAP 1402 (2014) 038, [1311.2606].

[49] O. Doré et al., Cosmology with the SPHEREX All-Sky Spectral Survey, 1412.4872.


[101] D. H. Lyth, Generating $f_{NL}$ at $\ell \lesssim 60$, JCAP 1504 (2015) 039, [1405.3562].


A. Ashoorioon and T. Koivisto, *Hemispherical Asymmetry from Parity-Violating Excited Initial States*, 1507.03514.


Z. Chang and S. Wang, *Implications of primordial power spectra with statistical anisotropy on CMB temperature fluctuation and polarizations*, 1312.6575.


Euclid Theory Working Group collaboration, L. Amendola et al., Cosmology and fundamental physics with the Euclid satellite, Living Rev.Rel. 16 (2013) 6, [1206.1225].


[173] M. LoVerde and K. M. Smith, "The Non-Gaussian Halo Mass Function with $f_{NL}$, $g_{NL}$ and $\tau_{NL}$, JCAP 1108 (2011) 003, [1102.1439].


*Toward a halo mass function for precision cosmology: The Limits of universality*,


[195] F. Bernardeau, S. Colombi, E. Gaztanaga and R. Scoccimarro, *Large scale structure of the universe and cosmological perturbation theory*,


*JCAP* **1410** (2014) 077, [1312.1364].


[223] K. M. Smith, S. Ferraro and M. LoVerde, Halo clustering and $g_{NL}$-type primordial non-gaussianity, JCAP 3 (Mar., 2012) 032, [1106.0503].


Vita

Saroj Adhikari

Saroj Adhikari completed his Bachelor’s degree in physics from the University of Central Arkansas in 2010. Before moving to the US for his college education in 2006, he completed his high school education in his native country of Nepal. After college, he joined the physics PhD program at the Pennsylvania State University in 2010. After working for a short period in a high energy experiment group, he decided to switch to cosmology research, and has worked under the supervision of Prof. Sarah Shandera since 2012 for his dissertation research.

Positions

1. Postdoctoral Research Fellow, Cosmology
   University of Michigan, Ann Arbor, MI 2016 - 2019

2. Graduate Research and Teaching Assistant,
   Institute for Gravitation and the Cosmos and Department of Physics
   The Pennsylvania State University, University Park, PA 2010 - 2016

Awards

1. Frymoyer Honors Fellowship, Penn State 2015-16
2. Downsborough Graduate Fellowship in Physics, Penn State 2014-15
3. David C. Duncan Graduate Fellowship in Physics, Penn State 2013-14
4. Homer F. Braddock Scholarship, Penn State 2010-11
5. Dorothy M. Long Scholarship, University of Central Arkansas 2008-09
6. Honors Scholarship, University of Central Arkansas 2006-10