Is the Universe Normal? Constraining Scale-Dependent Primordial Non-Gaussianity

by

Adam M. Becker

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in the University of Michigan
2012

Doctoral Committee:

Assistant Professor Dragan Huterer, Chair
Professor August Evrard
Professor Timothy A. McKay
Assistant Professor Christopher John Miller
Assistant Professor Kathryn Zurek

The fact that we live at the bottom of a deep gravity well, on the surface of a gas covered planet going around a nuclear fireball 90 million miles away, and think this to be *normal* is obviously some indication of how skewed our perspective tends to be – but we have done various things over intellectual history to slowly correct some of our misapprehensions.
– Douglas Adams

To be able to see Nobody! And at that distance, too! Why, it's as much as I can do to see real people, by this light!
– *Through The Looking-Glass*

ACKNOWLEDGEMENTS

Technical acknowledgements. Nearly all of the computational work for this dissertation was done using the Python programming language, along with the NumPy and SciPy quantitative and scientific computing packages. All of the figures, with the exception of Figs. 1.3 and 1.4, were created using the matplotlib package for Python. The work that I did not do in Python mostly involved CAMB (Code for Anisotropies in the Microwave Background), maintainted by Antony Lewis, and HEALPix, maintained by NASA/JPL. I also made use of the GNU Scientific Library (GSL) in some of this work. I was supported by a grant from the NSF in the course of this work.

Personal acknowledgements. There are many people who have helped me through the process of researching and writing this dissertation, and in graduate school more generally. A few require special thanks:

- Dragan Huterer, for the major role he played in this research, in addition to being the best advisor I could possibly have hoped for.
- Kenji Kadota, for the work he did on much of this research – Chapter II, in particular, would have been impossible without him.
- Chris Byrnes, Sarah Shandera, and Amit Yadav for many helpful discussions and e-mail exchanges.

- Gus Evrard, Tim McKay, Chris Miller, and Kathryn Zurek, for agreeing to serve on my thesis committee.
- Bob Ziff, for being understanding about my choices in research.
- Cameron Gibelyou, for helping me and being generally awesome throughout.
- Daniel Jordan, for showing me that there should be one – and preferably only one – obvious way to do it.
- Marshall Weir and his family, for tea, Thanksgiving, and all the rest.
- Stefan Richter, for getting me through rough spots, and also to the gym.
- Kate Hanley and the Clarks, for their hospitality and friendship.
- Andrew McNair, for never letting me down.
- Adrienne Grant, for following me around the country and letting me steal her friends.
- Andrew Schwarzkopf, for asking the right questions and having faith in my insanity.
- My friends at the Telluride House, for giving me a home and an extended family.
- My parents, for being supportive and patient, and for answering far more questions than they ever expected to be asked.
- And Elisabeth, who knows why.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

CHAPTER I

Introduction

1.1 Constraining models of inflationary-era physics

The Friedmann-Robertson-Walker metric commonly used to describe our universe is based on the assumption that the universe looks the same everywhere, in all directions. While this is nearly true on large scales, it is manifestly untrue on small scales, as demonstrated by our existence, and more broadly the existence of galaxies and galaxy clusters. The evolution of these structures is reasonably well understood; the cosmic microwave background (CMB) gives us evidence for density perturbations on the order of one part in 10^5 at the time of recombination, and their evolution to the large density perturbations that we see today is described well by gravitational collapse. But the origin of those perturbations is far less well understood. Our best guess comes from inflation. Inflation posits that the primordial density perturbations have their origin in quantum fluctuations of the inflaton field that were "blown up" to macroscopic scale during the inflationary era in the first $\sim 10^{-33}$ seconds after the Big Bang. Inflation is a remarkably successful theory – it neatly resolves several major problems regarding the very early universe, it's passed every observational test we have thrown at it, and it has been very theoretically fruitful. If anything, though, it's been *too* fruitful – in the thirty years since it was first proposed by Guth

(Guth (1981); Albrecht and Steinhardt (1982); Linde (1983)), inflation has grown from a single theory into a large class of theories. Since we have very little empirical access to the inflationary era, these theories have proliferated with few constraints placed upon them by observation. Furthermore, there are theoretical alternatives to inflation, such as ekpyrotic models, which cannot be ruled out on the basis of current observations.

It is difficult to place observational constraints on the physics of inflation because the inflationary epoch is so early in the history of the Universe. Very few signals remain from that epoch, and there are none uncontaminated by late-time effects. Most hopes for placing constraints on inflation are pinned on seeking out properties of the primordial density perturbations that were left behind after reheating[1]. The power spectrum of the primordial perturbations has been of particular interest: its amplitude A_s, spectral index n_s, the running (scale-dependence) of the spectral index $\frac{dn_s}{d\ln k}$, and the tensor-to-scalar ratio r have all been measured or constrained, largely through measurements of the CMB. While all of these parameters can tell us about the physics of inflation, the spectral index is especially notable. Standard slow-roll inflation predicts that n_s is just below one, and the WMAP CMB data confirm this prediction (Komatsu et al. (2011)): $n_s = 0.963 \pm 0.014$. This is perhaps the greatest observational triumph of standard inflation, but the spectral index carries limited information about the physics of inflation – and there are many different types of inflation (and alternatives to inflation altogether) which predict the same value for n_s. A large number of these models are also consistent with current measurements of A_s, $\frac{dn_s}{d\ln k}$, and r, leaving us with dozens of alternatives and few prospective means of choosing among them.

[1]While there is some hope of detecting gravitational waves from inflation, it is entirely possible that these waves are far too weak to be seen with a detector smaller than the observable universe.

1.2 Non-Gaussianity

One way to mine the primordial density perturbations for more information about the physics of inflation is go beyond the power spectrum and search for *non-Gaussianity* in the distribution of the perturbations. Single-field slow-roll inflation, with a canonical kinetic term in a Bunch-Davies vacuum, predicts that the primordial distribution of density perturbations at all scales should be very nearly Gaussian – to about one part in 10^8, though this would be reduced to one part in 10^6 by secondary and late-time effects (See Maldacena (2003), among many others; for a more recent review, see Yadav and Wandelt (2010)). Specifically, the *magnitude* of the primordial fluctuations should follow a Gaussian distribution at all scales (see Figure 1.1). This follows from Wick's theorem, which guarantees that the Nth-order correlation function of the inflaton field will be equal to the Nth moment of a Gaussian distribution, given the assumptions of standard inflation (slow-roll, Bunch-Davies vacuum, canonical kinetic term, and a single inflaton field). Thus, the detection of significant non-Gaussianity would be a serious challenge to the simplest models of inflation, and would be a corresponding boon to non-standard inflationary theories.

1.2.1 Modeling non-Gaussianity

Unfortunately, searching for non-Gaussianity is not as simple as searching for a fit to a given probability distribution – "non-Gaussianity" is a wildly non-specific term. (Calling a distribution "non-Gaussian" is like calling an object "not a puppy" – many things (hats, lions, sonic screwdrivers) are not puppies.) The universe is so close to Gaussian that merely searching for deviations from Gaussianity in the distribution of the primordial perturbations isn't an enlightening line of inquiry (Figure 1.2). But sensitive estimators of non-Gaussianity can be constructed if a particular model is

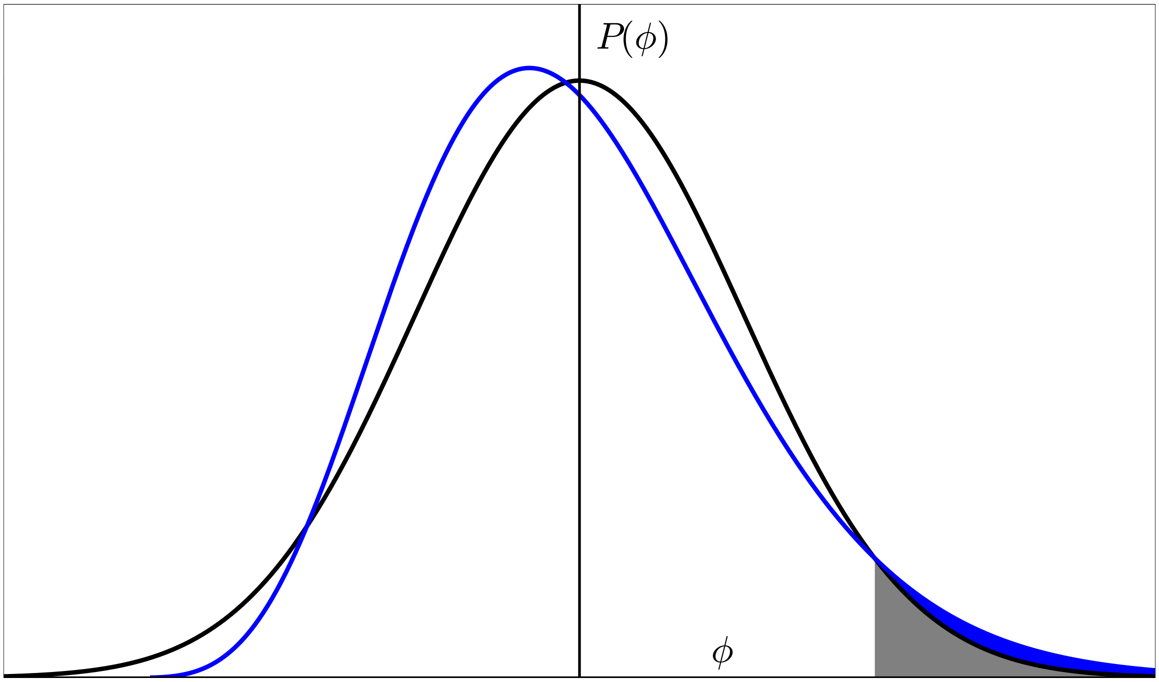


Figure 1.1: A comparison of a Gaussian distribution (black curve) with a non-Gaussian distribution of the local type (Equation 1.1; blue curve). Here, the non-Gaussianity parameter $f_{\rm NL} = 10^4$, much larger than it is in our universe; I have made it large so the difference between the two distributions is visible. (See Figure 1.2.) The slight excess in the high tail of the non-Gaussian distribution (blue filled region) is the key region for detecting non-Gaussianity in large-scale structure today.

picked. The most commonly discussed model of non-Gaussianity, known as the local or "squeezed" model, is defined via (see e.g. Komatsu and Spergel (2001)):

$$\Phi(x) = \phi_G(x) + f_{\rm NL}(\phi_G(x)^2 - \langle \phi_G(x)^2 \rangle). \tag{1.1}$$

Here, Φ denotes the primordial curvature perturbations (Bardeen's gauge-invariant potential), and $\phi_G(x)$ is a Gaussian random field. The parameter $f_{\rm NL}$ characterizes the level of non-Gaussianity – in a Gaussian universe, $f_{\rm NL} = 0$. Looking more carefully at (1.1), we will have non-Gaussianity of order unity when the second term is roughly equal to the first term; that is, when $f_{\rm NL} \sim 1/\phi_G \sim 10^5$.

The local model has been much studied, in part because it is the first two terms of the most general local form of non-Gaussianity (Babich et al. (2004)). This model is also theoretically well-motivated: various popular forms of inflation, including curvaton, multi-field, and modulated reheating models, all predict some amount of local non-Gaussianity (Yadav and Wandelt (2010)). Perhaps the most interesting thing about the local model is the remarkable result from Creminelli and Zaldarriaga (2004): the detection of $f_{\rm NL}$ much greater than unity would rule out *all* single-field inflation models, regardless of the dynamics involved.

1.2.2 Detecting non-Gaussianity

In addition to the vagueness of the term "non-Gaussianity," there is the further problem that neither the primordial curvature perturbations $\Phi(k)$ nor the primordial density fluctuations $\delta\rho/\rho$ are directly observable. To learn more about primordial non-Gaussianity, we have to turn the clock back: we must infer properties of the primordial perturbations from their "descendants," the CMB anisotropies and the large-scale structure (LSS) of the universe today. (See Figures 1.3 and 1.4.) Since the universe is so close to Gaussian – and since Gaussian distributions are fully described by their one- and two-point functions – we must look at higher-order correlation functions (and their Fourier transforms, the polyspectra) in order to test any given model of non-Gaussianity. The polyspectra generally offer a much larger number of observables, yielding a large signal-to-noise ratio even if S/N is small in each individual observable. For example, if there was non-Gaussianity of the local type in the primordial universe, then the bispectrum (the Fourier transform of the three-point function) of the CMB is directly proportional to $f_{\rm NL}$; and the number of angular-averaged terms $B_{\ell_1\ell_2\ell_3}$ in the bispectrum is proportional to $\ell^3_{\rm max}$, where $\ell_{\rm max} \sim 500$ for WMAP. In fact, the best constraints on the Gaussianity of the universe have until recently come from the bispectrum[2] of the CMB: WMAP has constrained

[2]Some models of non-Gaussianity exist that lead to modifications in the four-point function and its Fourier transform, the trispectrum; these are often characterized using the parameters $g_{\rm NL}$ and $\tau_{\rm NL}$. These are not quite as

$f_{\rm NL}$ to roughly 30 ± 20 (Komatsu et al. (2011)), corresponding to a universe that is Gaussian to about one part in $10^{3.5}$. However, Dalal et al. (2008) pointed out that a non-zero $f_{\rm NL}$ leads to a strongly scale-dependent dark matter halo bias, which can be detected in the power spectrum of large-scale structure; this technique has since emerged as a source of constraints already competitive with the CMB (Slosar et al. (2008)).[3]

Applying these methods in the context of scale-dependent models of non-Gaussianity is the focus of the rest of this work. In the rest of this chapter, I will examine the local model in greater detail, along with scale-dependent extensions of that model; in Chapter II, I'll discuss the relatively new method of constraining non-Gaussianity from the power spectrum of the LSS, following Dalal et al. (2008), in the context of scale-dependent models; in Chapter III, I'll discuss methods of constraining scale-dependent models from the CMB; in Chapter IV, I'll give actual constraints from current WMAP data on a particular scale-dependent model; finally, Chapter V comprises my overall conclusions and a summary.

1.3 Beyond the local model

Switching over to Fourier space, the local model takes the form:

$$\Phi(k) = \phi_G(k) + f_{\rm NL} \int \frac{d^3k'}{(2\pi)^3} \phi_G(k') \phi_G(k-k'). \tag{1.2}$$

For this model, the primordial curvature bispectrum takes a relatively simple form:

$$B_\phi(\vec{k}_1, \vec{k}_2, \vec{k}_3) = 2 f_{\rm NL} (2\pi)^3 \delta(\vec{k}_1 + \vec{k}_2 + \vec{k}_3)(P_\phi(k_1) P_\phi(k_2) + \text{perm.}). \tag{1.3}$$

well-studied, in part because there are few truly computationally-efficient algorithms to calculate the trispectrum. I will not be considering such models in this work.

[3]There are also other techniques involving large-scale structure, most notably the galaxy bispectrum and cluster counts – but the former is not practical to calculate, and the latter is not nearly as sensitive a probe of non-Gaussianity as the halo bias.

Here, P_ϕ is the power spectrum of the primordial curvature perturbations, and δ is the Dirac delta function, enforcing the condition that the three k-vectors must form a triangle. Assuming translational symmetry, the primordial bispectrum for any model can always be written in the form (Babich et al. (2004)):

$$B_\phi(\vec{k}_1, \vec{k}_2, \vec{k}_3) = (2\pi)^3 \delta(\vec{k}_1 + \vec{k}_2 + \vec{k}_3) F(\vec{k}_1, \vec{k}_2, \vec{k}_3), \tag{1.4}$$

where F is known as the shape function, so called because it determines which shapes of triangles in k-space are the dominant contributions to the bispectrum. Thus, we can characterize different models of primordial non-Gaussianity by looking at the shape functions that they produce. We can easily see that F for the local model is

$$F_{\text{local}}(k_1, k_2, k_3) = 2 f_{\text{NL}} (P_\phi(k_1) P_\phi(k_2) + \text{perm.}), \tag{1.5}$$

where $P_\phi(k) \propto k^{-(4-n_s)}$ is the primordial curvature power spectrum. This function is maximized for triangles with one side much shorter than the others: $k_3 << k_1 \sim k_2$ – a long thin "squeezed" isosceles triangle. (Hence the name "squeezed model" for the local model.) Other models of non-Gaussianity favor triangles of different shapes. The equilateral model, as the name suggests, has much of its power in near-equilateral triangles; this type of non-Gaussianity is seen in DBI inflation, ghost inflation, and other inflationary models with non-standard kinetic terms. Models of inflation that drop the assumption of a Bunch-Davies vacuum can give rise to non-Gaussianity with a shape function that favors "folded" triangles: $k_1 \sim k_2 \sim k_3/2$ (see e.g. Babich et al. (2004); Chen (2005)).

1.3.1 Scale-dependent non-Gaussianity

While these models each favor a different shape of triangle, the deviation from Gaussianity in each model is independent of scale.[4] But there is good theoretical

[4]Scale-independence for a particular model of non-Gaussianity does *not* imply that similar triangles of different sizes in k-space contribute equally to the primordial curvature bispectrum associated with that model. A glance

motivation to think that non-Gaussianity, if it exists, will be scale-dependent; this is a generic result of single-field inflationary models with interactions, along with most multi-field models (e.g. Salopek and Bond (1990); Luo and Schramm (1993); Wang and Kamionkowski (2000); LoVerde et al. (2008); Sefusatti et al. (2009)). We can introduce scale-dependence to the local model by promoting the parameter $f_{\rm NL}$ to a function of scale, $f_{\rm NL}(k)$. The curvature pertubations in this new model are

$$\Phi(k) = \phi(k) + f_{\rm NL}(k) \int \frac{d^3k'}{(2\pi)^3} \phi(k')\phi(k-k'). \tag{1.6}$$

This form of non-Gaussianity is expected in curvaton or modulated reheating scenarios (see e.g. Byrnes et al. (2010) and Shandera et al. (2011), where this form explicitly appears in the study of these models; see also Linde and Mukhanov (1997); Lyth and Wands (2002); and Zaldarriaga (2004), among many others). Note that this new ansatz is *not* local, which is clear when we transform back into real space:

$$\Phi(x) = \phi + f_{\rm NL}(x) * (\phi(x)^2 - \langle\phi(x)^2\rangle), \tag{1.7}$$

where $*$ represents convolution and x denotes a three-dimensional spatial coordinate. The shape function F for this model takes the form:

$$F(k_1, k_2, k_3) = 2\left(f_{\rm NL}(k_3) P_\phi(k_1) P_\phi(k_2) + 2 \text{ perm.}\right) \tag{1.8}$$

We can parametrize $f_{\rm NL}(k)$ in a way that is valid for any general form of $f_{\rm NL}(k)$ by breaking $f_{\rm NL}(k)$ into a set of piecewise-constant (in wavenumber) bins, such that $f_{\rm NL}(k)$ is equal to $f^i_{\rm NL}$ in the ith wavenumber bin (Becker et al. (2011)):

$$f^i_{\rm NL} \equiv f_{\rm NL}(k_i). \tag{1.9}$$

at (1.5) confirms this: $F_{\rm local}(\lambda k_1, \lambda k_2, \lambda k_3) = \lambda^{-(8-2n_s)} F_{\rm local}(k_1, k_2, k_3)$. This scale-dependence comes from the fact that we are looking at the *primordial curvature* bispectrum, which is related to the bispectrum of density perturbations through the Poisson equation.

In this work, we pay special attention to this parametrization of $f_{\rm NL}(k)$, as well as a simple form of non-Gaussianity analogous to the conventional parameterization of the power spectrum

$$f_{\rm NL}(k) = f^{*}_{\rm NL}\left(\frac{k}{k_{\rm piv}}\right)^{n_{f_{\rm NL}}}. \tag{1.10}$$

Here, $k_{\rm piv}$ is an arbitrary fixed parameter, leaving $f^{*}_{\rm NL}$ and $n_{f_{\rm NL}}$ as the parameters of interest in this model (Shandera et al. (2011); Becker et al. (2011)).

In the rest of this work, I will forecast and find constraints on scale-dependent non-Gaussianity of the form (1.6). Chapters II and III are focused on projecting constraints on the piecewise-constant parameters in (1.9) using LSS and the CMB, respectively. In Chapter IV, I find the constraints placed on $n_{f_{\rm NL}}$ from the WMAP7 CMB temperature data set – to the best of my knowledge, a novel result.

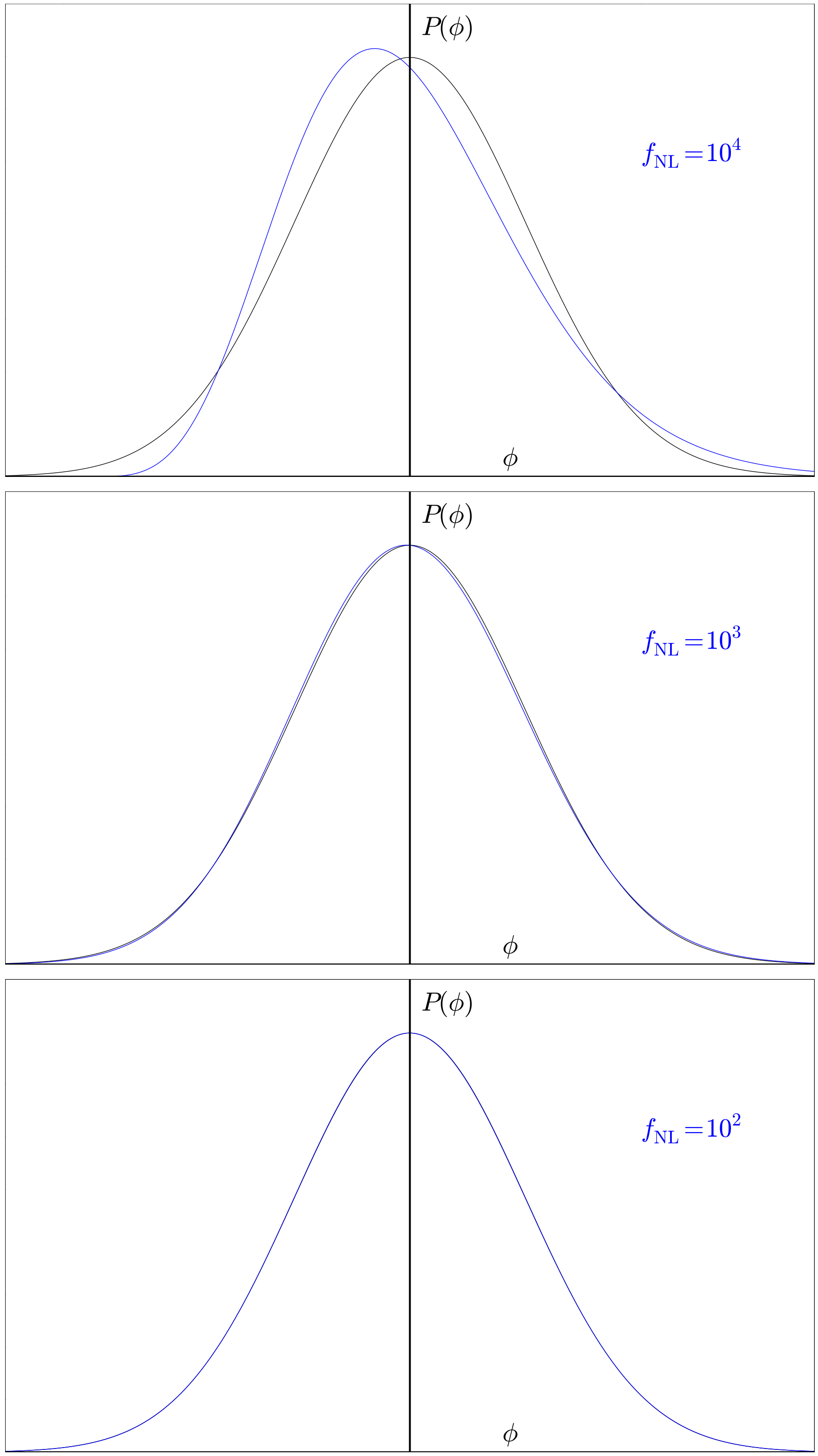


Figure 1.2: A further comparison of Gaussian and local non-Gaussian distributions. As the text in each panel indicates, the top panel has $f_{\rm NL} = 10^4$, the middle has $f_{\rm NL} = 10^3$, and the bottom panel has $f_{\rm NL} = 10^2$. For $f_{\rm NL} < 10^3$, it is quite difficult to tell the difference between the Gaussian and non-Gaussian one-point functions; thus, higher-order correlation functions and estimators are needed.

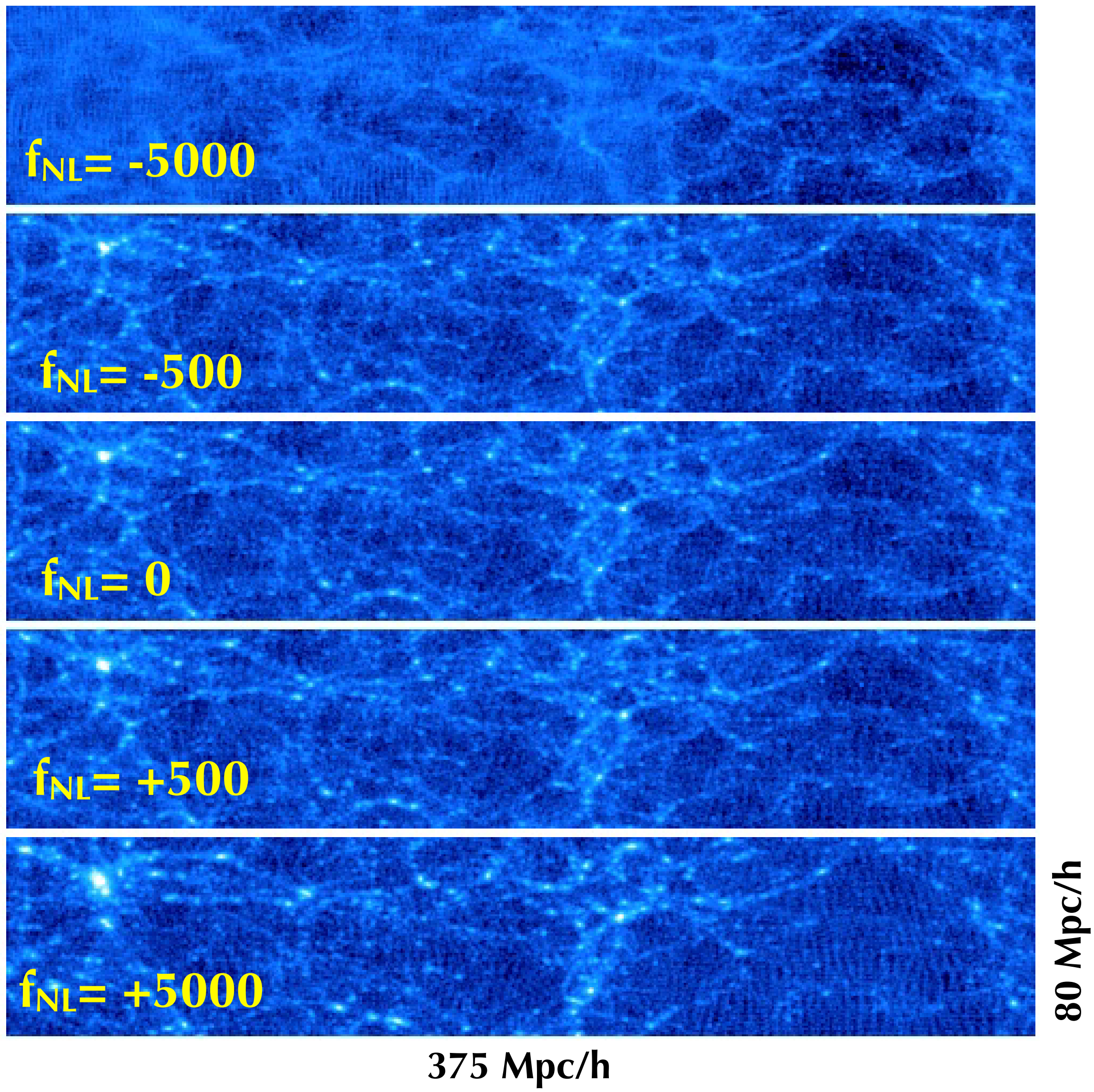


Figure 1.3: The effects of local non-Gaussianity on N-body simulations of large-scale structure (Dalal et al. (2008)). Here, we have five different simulations, each with a different value of $f_{\rm NL}$, but all with exactly the same initial conditions. Local non-Gaussianity introduces a scale-dependent bias into the halo power spectrum; see Chapter II.

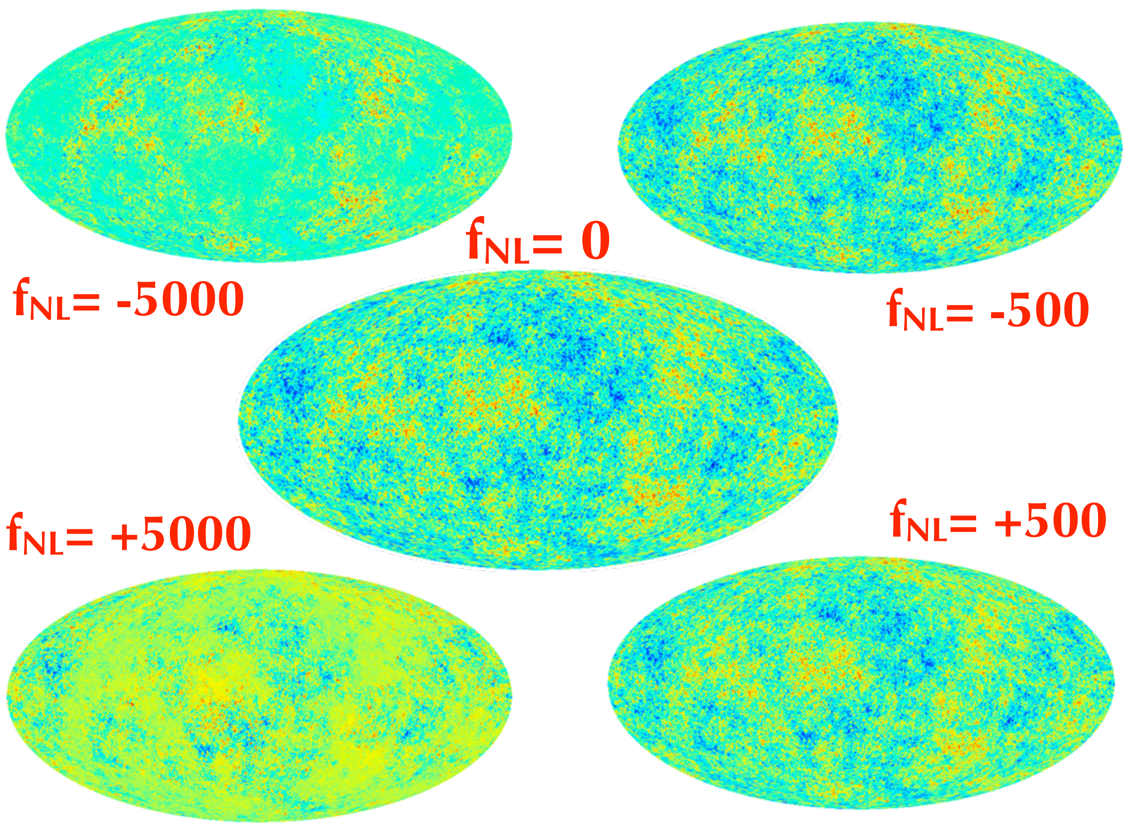


Figure 1.4: Effects of local non-Gaussianity on Monte Carlo simulations of the CMB sky, based on Elsner and Wandelt (2009).

CHAPTER II

Forecasted constraints on scale-dependent non-Gaussianity from LSS

2.1 Non-Gaussianity and bias

2.1.1 The effect of a non-vanishing bispectrum on bias

Dalal et al. (2008) found, analytically and numerically, that the bias of dark matter halos acquires strong scale dependence if $f_{\rm NL} \neq 0$:

$$b(k) = b_0 + f_{\rm NL}(b_0 - 1)\delta_c \frac{3\Omega_m H_0^2}{a\, g(a) T(k) c^2 k^2}. \tag{2.1}$$

Here, b_0 is the usual Gaussian bias (on large scales, where it is constant), $\delta_c \approx 1.686$ is the collapse threshold, a is the scale factor, Ω_m is the matter density relative to the critical density, H_0 is the Hubble constant, k is the wavenumber, $T(k)$ is the transfer function, and $g(a) = g(1)\frac{D(a)}{a}$ is the growth suppression factor . This result has been confirmed by other researchers using a variety of methods, including the peak-background split (e.g. Afshordi and Tolley (2008)), perturbation theory (e.g. McDonald (2008)), and numerical (N-body) simulations (e.g. Desjacques et al. (2009)). Astrophysical measurements of the scale dependence of the large-scale bias, using galaxy and quasar clustering as well as the cross-correlation between the galaxy density and CMB anisotropy, have recently been used to impose constraints

Portions of this chapter first appeared in:
Becker, A., Huterer, D., Kadota, K., Scale-dependent non-Gaussianity as a generalization of the local model, *Journal of Cosmology and Astroparticle Physics*, 2011, vol. 1, p. 006, doi:10.1088/1475-7516/2011/01/006

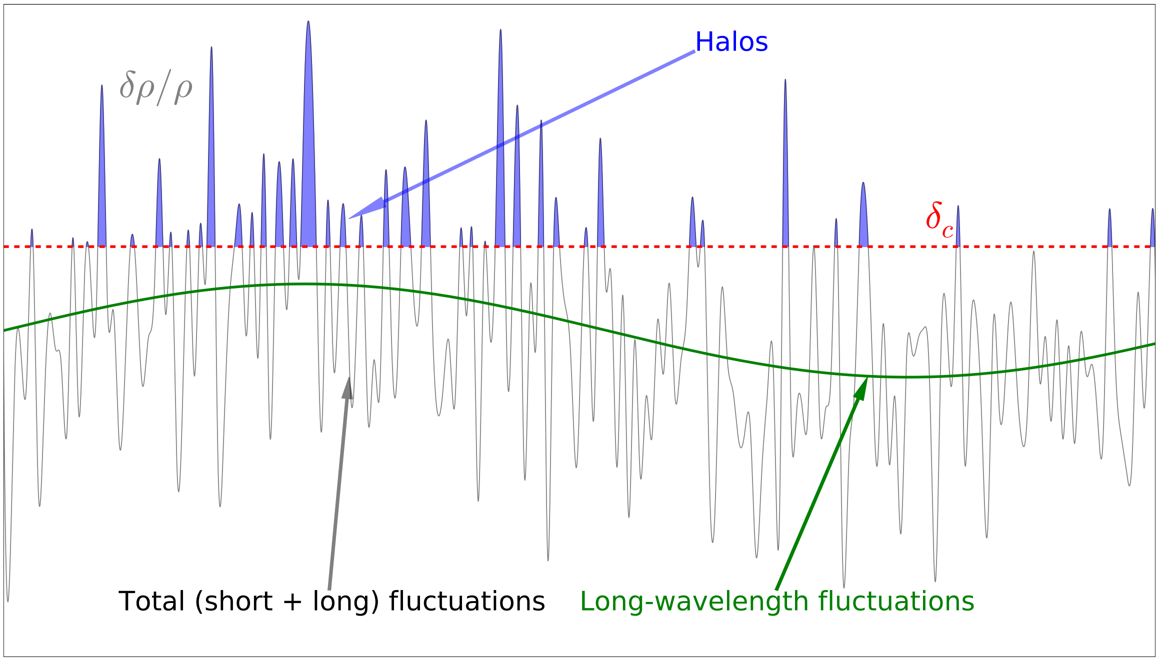


Figure 2.1: The peak-background split. Halos form when the local matter overdensity $\delta\rho/\rho$ exceeds the critical threshold for collapse, δ_c (red dashed line). In our toy model here, small-scale fluctuations are added to the large-scale fluctuations (green line) to get the overall fluctuations (black line). A given fluctuation is more likely to exceed the threshold and form a halo (blue regions) when it is sitting on top of a large-scale overdensity than when it is sitting on top of a large-scale underdensity; this is why there are more blue halos on the left than there are on the right.

on $f_{\rm NL}$ already comparable to those from the cosmic microwave background (CMB) anisotropy, giving $f_{\rm NL} = 28 \pm 23$ (1σ), with some dependence on the assumptions made in the analysis (Slosar et al. (2008)). In the future, constraints on $f_{\rm NL}$ are expected to be on the order of a few (Dalal et al. (2008); Cunha et al. (2010)). The sensitivity of the large-scale bias to other models of primordial non-Gaussianity has not yet been investigated much (though see analyses in e.g. Desjacques and Seljak (2010); Verde and Matarrese (2009)).

To get a physical picture of how halo bias is sensitive to local non-Gaussianity, first remember that halos form when the local matter overdensity $\delta\rho/\rho$ exceeds the critical threshold for collapse, δ_c. Therefore, a given small-scale fluctuation is more likely to exceed the threshold when it is sitting on top of a large-scale overdensity than when it is sitting on top of a large-scale underdensity (see Figure 2.1). This picture is called the peak-background split, and it is the primary source of the linear halo bias: $\delta_{\rm halo} = b_0\, \delta_{\rm matter}$.

Local non-Gaussianity introduces a coupling between the power in primordial *curvature* fluctuations, Φ, at small scales and large scales. Φ and $\delta\rho$ are related by the Poisson equation: $\Phi \sim \frac{\delta\rho}{k^2}$. Thus, when $f_{\rm NL} \neq 0$, the power in small-scale density fluctuations becomes tied to the power in large-scale density fluctuations, which introduces a scale-dependent term $\Delta b(k) \sim k^{-2}$ into the halo bias.

We can get a more rigorous derivation of this extra term by starting with the full Poisson equation, to find the relation between $\Phi(k)$ and the present-time (z=0) smoothed linear overdensities δ_R:

$$\delta_R(k) = \frac{2}{3}\frac{k^2 T(k)}{H_0^2 \Omega_m} \tilde{W}_R(k)\Phi(k) \equiv \mathcal{M}_R(k)\Phi(k); \tag{2.2}$$

where $T(k)$ is the matter transfer function, H_0 is the Hubble constant, Ω_m is the matter density relative to critical today, and $\tilde{W}_R(k)$ is the Fourier transform of the top-hat filter with radius R. The smoothing spatial scale R is related to the smoothing mass scale M via

$$M = \frac{4}{3}\pi R^3 \rho_{m,0}, \tag{2.3}$$

where $\rho_{m,0}$ is the matter energy density today.

One can expand the two point correlation function of dark matter halos, $\xi_h(\boldsymbol{x}_1, \boldsymbol{x}_2)$, in terms of high-order correlation functions of the underlying density field, $\xi_R^{(N)}$. In the high-threshold limit ($\nu \gg 1$), this becomes the so-called MLB formula (Grinstein and Wise (1986); Matarrese et al. (1986)):

$$\xi_h(\boldsymbol{x}_1, \boldsymbol{x}_2) = \xi_h(x_{12})$$

$$= -1 + \exp\left(\sum_{N=2}^{\infty}\sum_{j=1}^{N-1} \frac{\nu^N}{\sigma_R^N}\frac{1}{j!(N-j)!}\xi_R^{(N)}\left[\begin{array}{cc}\boldsymbol{x}_1, ..., \boldsymbol{x}_1, & \boldsymbol{x}_2, ..., \boldsymbol{x}_2 \\ j \text{ times} & (N-j) \text{ times}\end{array}\right]\right); \tag{2.4}$$

where $x_{ij} = |\boldsymbol{x}_i - \boldsymbol{x}_j|, \nu = \delta_c/\sigma_R$ is the peak height, and ${\xi_R}^{(n)}(r)$ is the n-point correlation function of the underlying matter density field smoothed with a top-hat filter of radius R. Keeping the terms up to the three-point correlation function, which is reasonable for the observationally allowed range of $f_{\rm NL}$, the expansion series gives us the halo correlation function in terms of the density field correlation functions:

$$\xi_h(x_{12}) = \frac{\nu^2}{\sigma_R^2}\xi_R^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2) + \frac{\nu^3}{\sigma_R^3}\xi_R^{(3)}(\boldsymbol{x}_1, \boldsymbol{x}_1, \boldsymbol{x}_2) + \dots \tag{2.5}$$

The power spectrum is given, to the same expansion order as Eq. (2.5), by

$$P_h(k) = \frac{\nu^2}{\sigma_R^2}P_R(k) + \frac{\nu^3}{\sigma_R^3}\int \frac{d^3q}{(2\pi)^3}B_R(k, q, |\boldsymbol{k} - \boldsymbol{q}|) + \dots \tag{2.6}$$

The first term on the right-hand side includes the familiar (Gaussian) bias $b = \nu/\sigma_R$ (in the high-peak limit for which the MLB formula is valid) for the Gaussian fluctuations. The effects of non-Gaussianity on the galaxy bias are represented by the second term, including the bispectrum B_R, which vanishes for the Gaussian fluctuations.

2.1.2 Beyond the high-peak approximation

The expression (2.1) is only correct in the high-peak, small-k limit. Desjacques et al. (2011) pointed out an additional term is required for the exact expression:

$$\Delta b(k) = 2\frac{F(k)}{\mathcal{M}_R(k)}\left((b_0 - 1)\frac{\delta_c}{D(z)} + \frac{d\ln F(k)}{d\ln \sigma_{0s}}\right) \tag{2.7}$$

where

$$F(k) \equiv \frac{1}{8\pi^2\sigma_R^2}\int dk_1 k_1^2 \mathcal{M}_R(k_1) P_\phi(k_1) \times \int_{-1}^{1} d\mu \mathcal{M}_R(k_2)\left[f_{\rm NL}(k)\frac{P_\phi(k_2)}{P_\phi(k)} + 2f_{\rm NL}(k_2)\right]. \tag{2.8}$$

The new term (second term on the left-hand side of Equation 2.7) vanishes when the fiducial $f_{\rm NL}(k) = 0$, but it remains relevant for any other constant or scale-dependent fiducial value, even for the piecewise-constant parametrization of $f_{\rm NL}(k)$ from equation (A.13). (See Appendix A.2.1 for details on this.) Desjacques et al. have found that this new term explains previously mysterious discrepancies (Shandera et al. (2011)) between the theoretical expectation for the scale-dependent bias and the results of numerical simulations.

2.2 Forecasted constraints on scale-dependent non-Gaussianity from large-scale structure

2.2.1 Fisher matrix analysis

We would like to project constraints on scale-dependent non-Gaussianity for future galaxy redshift surveys. To do this, we can calculate the Fisher information matrix for the parameters $f_{\rm NL}^j$ that describe the piecewise-constant $f_{\rm NL}(k)$. The Cramér-Rao inequality tells us that the inverse of the Fisher matrix sets a lower bound on the covariance matrix we can get on our parameters $f_{\rm NL}^i$ from our hypothetical survey. Specifically, given the Fisher matrix F_{ij}, the minimum possible marginalized and unmarginalized errors for a particular $f_{\rm NL}^i$ are $\sqrt{F_{ii}^{-1}}$ and $1/\sqrt{F_{ii}}$, respectively. Thus, the Fisher matrix allows us to forecast the extent to which scale-dependent non-Gaussianity could be constrained by future galaxy surveys. (For more on Fisher matrix analysis in general, see Appendix B.1. Details on calculating the derivative of the bias with respect to $f_{\rm NL}$ and $f_{\rm NL}(k)$, a necessary intermediate step in calculating the Fisher matrix, are in Appendix A.)

We consider measurements of the power spectrum $P_h(k)$ of dark matter halos (galaxies or clusters, for example) averaged over thin spherical shells in k-space. The variance of $P_h(k) \equiv P_h$ in each shell is (Feldman et al. (1994))

$$\sigma^2_{P_h} = \frac{2P_h^2}{V_{\text{shell}}\, V_{\text{survey}}} \left(\frac{1+nP_h}{nP_h}\right)^2 = \frac{(2\pi P_h)^2}{k^2 dk\, V_{\text{survey}}} \left(\frac{1+nP_h}{nP_h}\right)^2, \tag{2.9}$$

where $V_{\text{shell}} = 4\pi k^2 dk/(2\pi)^3$ is the volume of the shell in Fourier space (we are ignoring redshift distortion effects for simplicity here). Therefore, the Fisher matrix for measurements of $P_h(k,z)$ is the standard expression from Tegmark (1997):

$$F_{ij} = \sum_m V_m \int_{k_{\text{min}}}^{k_{\text{max}}} \frac{\partial P_h(k,z_m)}{\partial p_i} \frac{\partial P_h(k,z_m)}{\partial p_j} \frac{1}{\left[P_h(k,z_m) + \frac{1}{n}\right]^2} \frac{k^2 dk}{(2\pi)^2}, \tag{2.10}$$

where V_m is the comoving volume of the m-th redshift bin, each redshift bin is centered on z_m, and we have summed over all redshift bins. We adopt $k_{\text{min}} = 10^{-4}\, h^{-1}\,\text{Mpc}$, and we choose k_{max} as a function of z so that $\sigma(\pi/(2k_{\text{max}}), z) = 0.5$ (Seo and Eisenstein (2003)), which leads to $k_{\text{max}}(z=0) \approx 0.1h\,\text{Mpc}^{-1}$. P_h is the dark matter halo power spectrum, related to the true dark matter power spectrum P through

$$P_h(k) = b(k)^2 P(k), \tag{2.11}$$

where each quantity implicitly also depends on redshift. Finally, p_i are the parameters of interest; in our case, these are the f^i_{NL}.

2.2.2 Survey properties

We assume a future survey covering one-quarter of the sky (about 10,000 square degrees) out to $z = 1$, and find constraints for a set of 20 f^i_{NL} uniformly spaced in $\log k$ in the range $10^{-4} \leq k/(h\,\text{Mpc}^{-1}) \leq 1$, with a smoothing scale of $M_{\text{smooth}} =$

$10^{14} M_\odot$. We assume a flat universe and a fiducial model of constant non-Gaussianity at the value favored by the seven-year WMAP CMB data: $f_{\rm NL}(k) = 30 = f^i_{\rm NL}$. We include six cosmological parameters in our Fisher matrix aside from the $f^i_{\rm NL}$: Hubble's constant H_0; physical dark matter and baryon densities $\Omega_{\rm cdm}h^2$ and $\Omega_{\rm b}h^2$; equation of state of dark energy w; the log of the scalar amplitude of the matter power spectrum, $\log A_s$; and the spectral index of the matter power spectrum, n_s . Fiducial values of these parameters correspond to their best-fit WMAP7 values (Komatsu et al. (2011)). We also added the forecasted cosmological parameter constraints from the CMB experiment Planck by adding its Fisher matrix as a prior (W. Hu, private communication). Note that the CMB prior does *not* include CMB constraints on non-Gaussianity; the CMB constraints on $f_{\rm NL}(k)$ are studied separately in Chapter III. Finally, in addition to the cosmological parameters and the $f^i_{\rm NL}$, we include five Gaussian bias parameters in our Fisher matrix – one $b_0(z)$ for each redshift bin.[1] The fiducial values of these parameters are set by the Sheth-Tormen formalism (Sheth and Tormen (1999)). All of the hypothetical galaxy redshift surveys in this chapter and in Chapter III have these same assumptions, unless explicitly stated otherwise.

2.2.3 Forecasted constraints on the $f^i_{\rm NL}$

We already have the derivatives of $b(k)$ with respect to each of the $f^i_{\rm NL}$ (see Appendix A for these), so the derivative of $P_h(k)$ with respect to the $f^i_{\rm NL}$ is just

$$\frac{\partial P_h(k)}{\partial f^i_{\rm NL}} = 2\,\frac{\partial b(k)}{\partial f^i_{\rm NL}}\, b(k) P_{\rm mat}(k); \tag{2.12}$$

$P_{\rm mat}(k)$ is the ΛCDM matter power spectrum, easily obtained from a numerical code – in this case, CAMB. Since we only consider information from large scales

[1] Using six cosmological parameters along with five $b_0(z)$ and 20 $f^i_{\rm NL}$ led us into some issues with floating-point errors and numerical precision. The 31×31 Fisher matrix we obtained was rather ill-conditioned and difficult to invert reliably using 64-bit precision; we were eventually forced to move to 128-bit precision in order to accurately marginalize over the cosmological parameters and nuisance parameters.

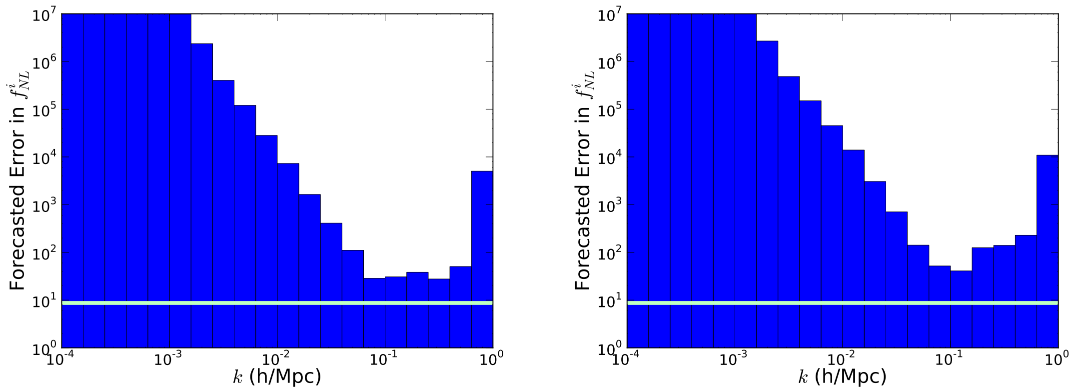
($k \leq k_{\rm max} \approx 0.1\,h\,{\rm Mpc}^{-1}$), we do not model the small amount of nonlinearity present at the high-k end of these scales. (Note that, while some of the $f_{\rm NL}^i$ have support at $k > k_{\rm max}(z=1) \approx 0.2\,h\,{\rm Mpc}^{-1}$, we only use information about those (and other) parameters coming from $k < k_{\rm max}(z)$ in each z-bin.)

The constraints vary considerably as a function of the k at which these parameters are defined. The best-constrained $f_{\rm NL}^i$ corresponds to the $10^{-0.6} < k < 10^{-0.4}$ bin, and it has an estimated unmarginalized error of $\sigma(f_{\rm NL}^{17}) = 28$; for comparison, the worst-constrained $f_{\rm NL}^i$, which corresponds to the largest scale (smallest k) bin, has an unmarginalized error well over 10^{11}. As expected, the marginalized constraints for the best-constrained parameters are a bit weaker than the unmarginalized constraints – the best-measured $f_{\rm NL}^i$ has an estimated marginalized error of 41. In general, dependence of the constraints on the value of k is determined by two competing factors: as k increases, there is a larger number of modes, each with a smaller signal (given by the smaller nongaussian bias Δb). The best-constrained k is also affected by the fact that only information out to $k = k_{\rm max} = 0.1h\,{\rm Mpc}^{-1}$ is assumed from the galaxy survey. In particular, we have checked that if we unrealistically assume information to be available at all k (instead of at $k < k_{\rm max}$) without modeling the nonlinearities, the unmarginalized constraints on $f_{\rm NL}^i$ improve monotonically with increasing k. Therefore, the raw signal-to-noise ratio in $f_{\rm NL}^i$ increases with k.

2.3 Projection and principal components

2.3.1 Constraining other $f_{\rm NL}(k)$ models

Once the Fisher matrix F has been obtained for the set of parameters $f_{\rm NL}^i$, it is quite simple to find the best possible constraints on the $f_{\rm NL}^i$ that could be obtained from a future galaxy redshift survey. By projecting this Fisher matrix into another basis, it is also possible to find the constraints on any arbitrary $f_{\rm NL}(k)$ without



(a) Unmarginalized errors (b) Marginalized errors

Figure 2.2: Forecasted unmarginalized (left panel) and marginalized (right panel) constraints on piecewise-constant parameters $f_{\rm NL}^i$ assuming a future galaxy survey covering one-quarter of the sky out to $z = 1$, with average number density of 2×10^{-4} gal/Mpc^3. For comparison, the green horizontal line is the constraint found for a constant $f_{\rm NL}$ using the same survey assumptions . While the individual parameters $f_{\rm NL}^i$ are poorly constrained as expected, their few best linear combinations – the principal components – are well measured; see the next section and text for details.

calculating a new Fisher matrix from scratch. A trivial example can be found in Appendix B.2, where we find that the estimated error on a constant $f_{\rm NL}$, assuming the same future survey as in the previous section, is $\sigma(f_{\rm NL}) = 8.7$. (Note that this forecasted constraint is on a par with the error expected from Planck, where $\sigma(f_{\rm NL}) \sim 5$.)

For another, scale-dependent example, consider a power-law form for $f_{\rm NL}(k)$ (as in Equation 1.10):

$$f_{\rm NL}(k) = f_{\rm NL}^* \left(\frac{k}{k_{\rm piv}}\right)^{n_{f_{\rm NL}}}, \tag{2.13}$$

where $k_{\rm piv}$ is an arbitrary fixed parameter, leaving $f_{\rm NL}^*$ and $n_{f_{\rm NL}}$ as the parameters of interest in this model. ($k_{\rm piv}$ is generally chosen to minimize degeneracy between $f_{\rm NL}^*$ and $n_{f_{\rm NL}}$ for the observable of interest. We have set $k_{\rm piv} = 0.20h\,{\rm Mpc}^{-1}$, close to the optimal value in our case; in CMB analysis, the optimal value is lower, around $0.08h\,{\rm Mpc}^{-1}$.) The partial derivatives of our basis of $f_{\rm NL}^i$ with respect to these parameters are:

$$\frac{\partial f_{\rm NL}^i}{\partial f_{\rm NL}^*} = \left(\frac{k}{k_{\rm piv}}\right)^{n_{f_{\rm NL}}}; \tag{2.14}$$

$$\frac{\partial f_{\rm NL}^i}{\partial n_{f_{\rm NL}}} = f_{\rm NL}^* \left(\frac{k}{k_{\rm piv}}\right)^{n_{f_{\rm NL}}} \log\left(\frac{k}{k_{\rm piv}}\right). \tag{2.15}$$

Starting in a basis of 20 $f_{\rm NL}^i$ evenly spaced in log k, we project down to a basis of $f_{\rm NL}^*$ and $n_{f_{\rm NL}}$ in order to forecast constraints on the two new parameters from a survey covering one-quarter of the sky out to $z = 1$. We are using the same limits of integration as in Section 2.2.1, along with the fiducial values $f_{\rm NL}^* = 30$ and $n_{f_{\rm NL}} = 0$. The forecasted constraints on these parameters, marginalized over each other, are $\sigma_{f_{\rm NL}^*} = 8.7$ and $\sigma_{n_{f_{\rm NL}}} = 0.85$.

2.3.2 Principal components

We now represent a general function $f_{\rm NL}(k)$ in terms of principal components (PCs). In this approach, the *data* determine which particular modes of $f_{\rm NL}(k)$ are best or worst measured. The PCs also constitute a useful form of data compression, so that one can keep only a few of the best-measured modes to make inferences about the function $f_{\rm NL}(k)$. Finally, the PCs will also enable us to measure the degree of similarity between our scale-dependent ansatz and the local and equilateral forms of non-Gaussianity.

It is rather straightforward to start from the covariance matrix for the piecewise constant parameters $f_{\rm NL}^i$ and obtain the PCs of $f_{\rm NL}(k)$. The PCs are weights in wavenumber with amplitudes that are uncorrelated by construction, and they are ordered from the best-measured ($i = 0$) to the worst-measured ($i = 19$) for the assumed fiducial survey. The construction of the PCs is described in Appendix B.3. A few of these PCs of $f_{\rm NL}(k)$ are shown in Fig. 2.3. For example, the best-measured PC has most of its weight around $k = 10^{-0.4}\,h\,{\rm Mpc}^{-1}$, which agrees with sensitivities of

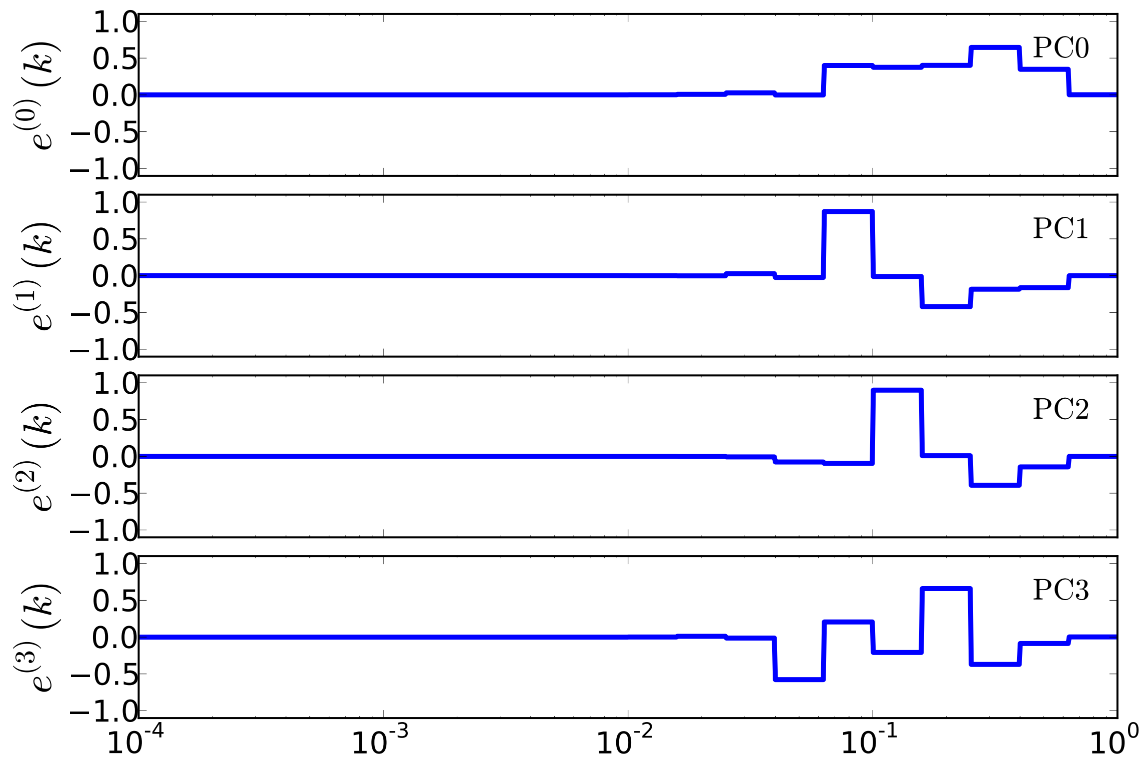


Figure 2.3: The forecasted best-measured principal components of $f_{\rm NL}(k)$. The PCs, $e^{(j)}(k)$, are eigenvectors of the Fisher matrix for the $f^i_{\rm NL}$, and are ordered from the best-measured one ($j = 0$) to the worst-measured one ($j = 19$) for the assumed fiducial survey.

piecewise-constant parameters shown in Fig. 2.2. Again, the sensitivity is not greatest at the largest value of k ($1\,h\,{\rm Mpc}^{-1}$) because we assumed cosmological information from $k \leq k_{\rm max} = 0.1\,h\,{\rm Mpc}^{-1}$. We checked that information available at a higher $k_{\rm max}$ would shift the "sweet spot" of sensitivity to higher wavenumbers in this case as well.

The forecasted error in the best-measured PC is 19.3; the error in the next-best measured PCs are 31.3 and 34.7, but the accuracy rapidly drops off from there. Thus, the first three or four PCs should be enough for any conceivable application. The forecasted error in each PC is plotted on a logarithmic scale in figure 2.4.

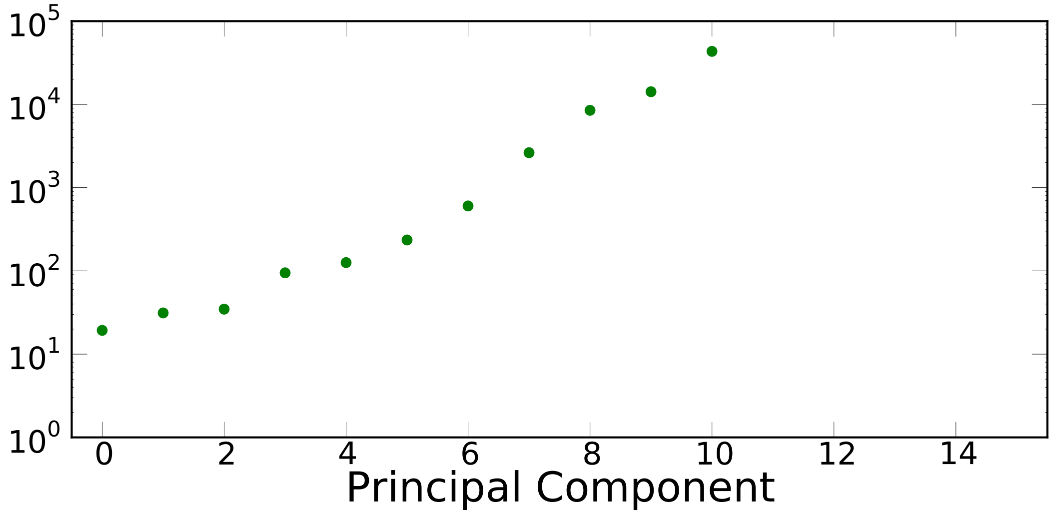


Figure 2.4: Forecasted RMS error on each principal component from LSS.

2.4 Conclusions

In this chapter, we used forecasted constraints from an intermediate-future galaxy survey to calculate errors on individual parameters $f_{\rm NL}^i$ (see Fig. 2.2). Projecting the Fisher matrix for the $f_{\rm NL}^i$ down to a different basis, we were able to project constraints on the power-law model of $f_{\rm NL}(k)$ (1.10). We further calculated the principal components of $f_{\rm NL}(k)$, and thus identified the best-measured configurations (in wavenumber) of this function (see Fig. 2.3). While the sensitivity of the survey to non-Gaussianity increases with increasing k, restricting the survey information to scales where linear perturbation theory is valid imposes a "sweet spot" in sensitivity of $k \sim 0.2h\,{\rm Mpc}^{-1}$. We will see a similar effect – but at a different scale – in the next chapter, where we forecast constraints on the $f_{\rm NL}^i$ from the CMB.

CHAPTER III

Forecasted constraints on scale-dependent non-Gaussianity from the CMB

3.1 Signatures of the generalized local model in the CMB

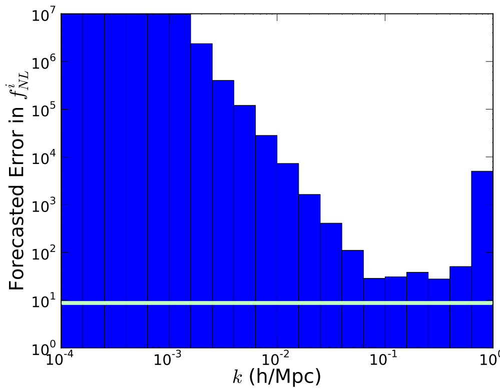
Traditionally, the best constraints on non-Gaussianity have come from the CMB. This is done almost exclusively through estimators involving the N-point correlation functions for $N > 2$ and their Fourier transforms, the polyspectra. Most emphasis has been on the $N = 3$ case, or the bispectrum of temperature fluctuations in the CMB, if only because of its relative computational simplicity. The well-known general expression for the CMB bispectrum, re-derived in Appendix C.1, is

$$\begin{aligned} B^{pqr}_{\ell_1\ell_2\ell_3} &= \left(\frac{2}{\pi}\right)^3 \sqrt{\frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4\pi}} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix} \int k_1^2 dk_1\, k_2^2 dk_2\, k_3^2 dk_3 \\ &\times \; B_\Phi(k_1,k_2,k_3) t^p_{\ell_1}(k_1) t^q_{\ell_2}(k_2) t^r_{\ell_3}(k_3) \int_0^\infty r^2 dr\; j_{\ell_1}(k_1 r) j_{\ell_2}(k_2 r) j_{\ell_3}(k_3 r). \end{aligned} \tag{3.1}$$

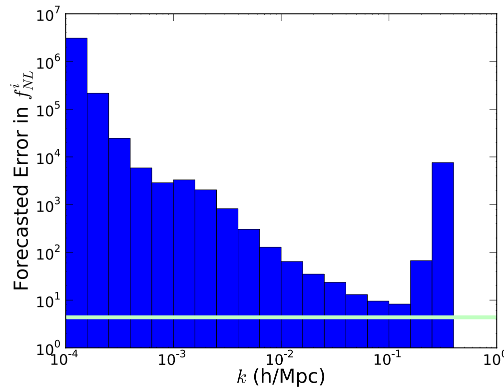
In principle, we can use this to find the Fisher matrix F_{ij} for the CMB bispectrum: (Babich and Zaldarriaga (2004); Komatsu and Spergel (2001))

$$F^{\mathrm{CMB}}_{ij} = f_{\mathrm{sky}} \sum_{lmn,pqr} \sum_{2\le\ell_1\le\ell_2\le\ell_3} \frac{1}{\Delta_{\ell_1\ell_2\ell_3}} \frac{\partial B^{lmn}_{\ell_1\ell_2\ell_3}}{\partial p_i} (\mathbf{C}^{-1}_{\ell_1\ell_2\ell_3})_{lmn,pqr} \frac{\partial B^{pqr}_{\ell_1\ell_2\ell_3}}{\partial p_j} \tag{3.2}$$

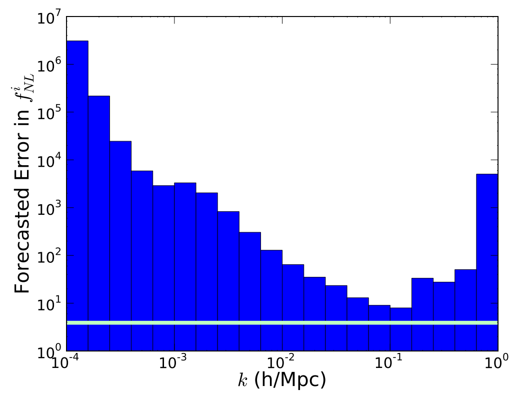
Here, $\mathbf{C}$ is the covariance of the bispectrum and $p_{i,j}$ are the parameters of interest. $\Delta_{\ell_1\ell_2\ell_3}$ is a combinatoric term – equal to 6 when $\ell_1 = \ell_2 = \ell_3$, 1 when $\ell_1 \neq \ell_2 \neq \ell_3$, and 2 otherwise (Spergel and Goldberg (1999)). The indices i, j, k and p, q, r run



(a) LSS



(b) Planck



(c) Combined

Figure 3.1: LSS (top left), CMB (top right), and combined (bottom) forecasted constraints on the piecewise constant parameters $f^i_{\rm NL}$ in the generalized local model. All constraints are unmarginalized. The LSS constraints come from the power spectrum of halos, assuming the same survey parameters as Section 2.2.2, while the CMB constraints come from the bispectrum of temperature and polarization fluctuations. See text for details. For reference, the green line is the forecasted error on a constant $f_{\rm NL}$ using the same assumptions. There are bins "missing" on the rightmost end of the Planck plot; those bins correspond to k-values too large to be probed when $\ell_{\max} = 2000$, as it is here.

independently over all eight possible ordered triplets of temperature and polarization fields (TTT, TTE...EEE). $\mathbf{C}$ can be thought of as a 6-point function, being the covariance of the 3-point function; since $f_{\rm NL} \ll 10^5$, it is reasonable to only consider the Gaussian contribution to the covariance of the bispectrum, $\mathbf{C}$. Using Wick's theorem, this is:

$$\mathbf{C}_{\ell_1\ell_2\ell_3} = C_{\ell_1} C_{\ell_2} C_{\ell_3} \tag{3.3}$$

Further details of calculating $\mathbf{C}$, $B^{pqr}_{\ell_1\ell_2\ell_3}$, and the derivatives of the bispectrum are all in Appendix C.

Equation (3.1) is a totally general result for the bispectrum of the CMB in terms of the primordial Bardeen curvature bispectrum B_Φ; we have not picked a particular model of non-Gaussianity. But (3.1) is not useful without an expression for B_Φ. For the local model (i.e. constant $f_{\rm NL}$), B_Φ is:

$$B_\Phi(k_1,k_2,k_3) = 2\Delta_\phi^2 f_{\rm NL}\left(\frac{1}{k_1^{3-(n_s-1)}k_2^{3-(n_s-1)}} + \text{perm.}\right) \tag{3.4}$$

where Δ_ϕ is the amplitude of the primordial Bardeen curvature power spectrum. Using Eqs. (3.2), (C.29), and (C.24), we have the following expression for the CMB bispectrum Fisher information in the constant $f_{\rm NL}$ case:

$$\begin{aligned} F^{\rm CMB}_{f_{\rm NL}} &= 4\Delta_\phi^4 \sum_{lmn,pqr}\sum_{2\le\ell_1\le\ell_2\le\ell_3} \frac{1}{\Delta_{\ell_1\ell_2\ell_3}}\frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4\pi}\begin{pmatrix}\ell_1 & \ell_2 & \ell_3\\ 0 & 0 & 0\end{pmatrix}^2\frac{1}{\Delta_{\ell_1\ell_2\ell_3}} \\ &\times\ (C^{-1}_{\ell_1})_{lp}(C^{-1}_{\ell_2})_{mq}(C^{-1}_{\ell_3})_{nr}\left[\int_0^\infty r^2 dr\left(\alpha^l_{\ell_1}(r)\beta^m_{\ell_2}(r)\beta^n_{\ell_3}(r)+\text{perm.}\right)\right] \\ &\times\left[\int_0^\infty r^2 dr\left(\alpha^p_{\ell_1}(r)\beta^q_{\ell_2}(r)\beta^r_{\ell_3}(r)+\text{perm.}\right)\right]. \end{aligned} \tag{3.5}$$

For the scale-dependent generalized local model, with $f_{\rm NL}(k)$ in place of $f_{\rm NL}$, things are somewhat more complicated. The Bardeen curvature bispectrum is:

$$B_\Phi(k_1,k_2,k_3) = 2\Delta_\phi^2\left(\frac{f_{\rm NL}(k_3)}{k_1^{3-(n_s-1)}k_2^{3-(n_s-1)}} + \text{perm.}\right). \tag{3.6}$$

Using the piecewise-constant parametrization of $f_{\rm NL}(k)$ together with (3.2), (C.29), and (C.25), we get an expression for the Fisher matrix of the $f^i_{\rm NL}$ that is similar to (3.5):

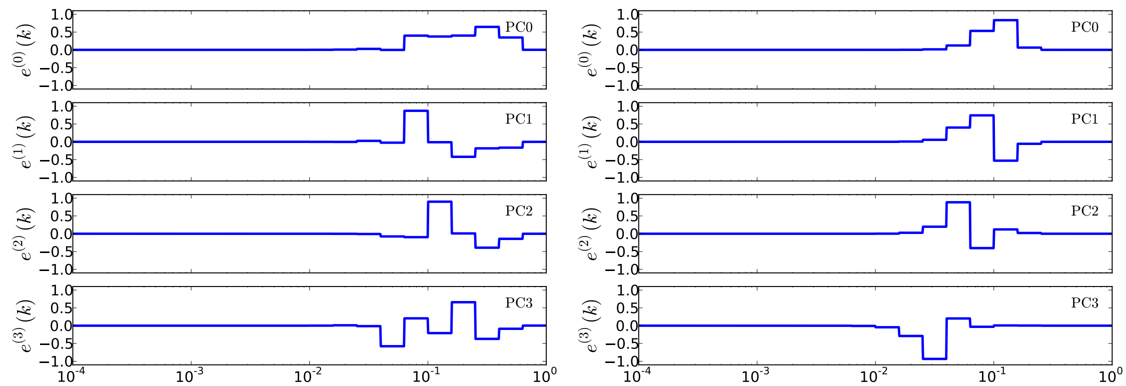
$$
\begin{aligned}
F_{ij}^{\text{CMB}} &= 4\Delta_\phi^4 \sum_{lmn,pqr} \sum_{2\le\ell_1\le\ell_2\le\ell_3}^{\ell_{\max}} \frac{1}{\Delta_{\ell_1\ell_2\ell_3}} \frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4\pi} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}^2 \\
&\times \frac{1}{\Delta_{\ell_1\ell_2\ell_3}} (C_{\ell_1}^{-1})_{ip} (C_{\ell_2}^{-1})_{jq} (C_{\ell_3}^{-1})_{kr} \left[\int_0^\infty r^2 dr \left(\alpha_{\ell_1}^{l,i}(r) \beta_{\ell_2}^m(r) \beta_{\ell_3}^n(r) + \text{perm.} \right) \right] \\
&\times \left[\int_0^\infty r^2 dr \left(\alpha_{\ell_1}^{p,j}(r) \beta_{\ell_2}^q(r) \beta_{\ell_3}^r(r) + \text{perm.} \right) \right]. \qquad (3.7)
\end{aligned}
$$

Despite appearances, calculating the full Fisher matrix F^{CMB} is relatively straightforward, and it takes roughly half an hour (on four 2.2 GHz processors) for twenty f_{NL}^i parameters with $\ell_{\max} \approx 2000$. (Some tabulation is necessary for the α and β functions, and the Wigner $3j$-symbol is not easy to calculate for large ℓ. Details on all of this are in Appendix C.) We did not include other cosmological parameters in this Fisher matrix, as the CMB bispectrum does not constrain them terribly well, nor is f_{NL} expected to be terribly degenerate with them; the CMB power spectrum places much stronger constraints on other cosmological parameters.

3.2 Results and joint constraints

3.2.1 Forecasted constraints on the f_{NL}^i

We have performed a CMB Fisher matrix analysis to forecast errors from the Planck mission: we take $\ell_{\max} = 2000$ and noise parameters from the Planck "blue book." Figure 3.1 shows the (unmarginalized) constraints on the piecewise constant parameters f_{NL}^i in the generalized local model from the LSS and Planck forecasts individually, as well as combined. Note that both types of surveys have comparable constraints at the pivot wavenumber, and the pivots also agree (though this statement is only approximate given the huge range of scales on both axes). Away from the pivot, the Planck constraints are expected to be better than those from the LSS, but both rapidly deteriorate away from the pivot $k_{\text{piv}} \approx 0.1h\,\text{Mpc}^{-1}$. Finally, the



(a) LSS PCs (b) Planck PCs

Figure 3.2: The forecasted best-measured principal components of $f_{\rm NL}(k)$ from LSS and Planck, with a fiducial $f_{\rm NL}(k) = 30$. The PCs, $e^{(j)}(k)$, are eigenvectors of the Fisher matrix for the $f^i_{\rm NL}$, and are ordered from the best-measured one ($j = 0$) to the worst-measured one ($j = 19$) for the assumed fiducial survey.

combined constraints are significantly helped by breaking of the degeneracies between the CMB and the LSS, and lead to better constraints across a wider range of scales. We will make these statements more quantitative below when we study the specific case where $f_{\rm NL}(k)$ is a pure power law in k.

Note that our Fisher matrices for the CMB – but not for the LSS – assume all cosmological parameters other than the $f^i_{\rm NL}$ are fixed (known). Adding priors from other data sets (e.g. SN Ia, the power spectrum of the CMB) constrains other cosmological parameters well enough that we would not get appreciably different results if we had those other parameters and their priors in our Fisher matrix.

3.2.2 Principal component analysis

As in Chapter II, we can represent a general function $f_{\rm NL}(k)$ in terms of principal components (PCs). Figure 3.2 shows the forecasted PCs of LSS and Planck separately, while Fig. 3.3 shows the combined PCs. Fig. 3.4 shows the forecasted 1-σ errors on the PCs for LSS, Planck, and the two combined. Typically, the lowest principal component (PC0) serves to see how well we can find the deviation of $f_{\rm NL}(k)$

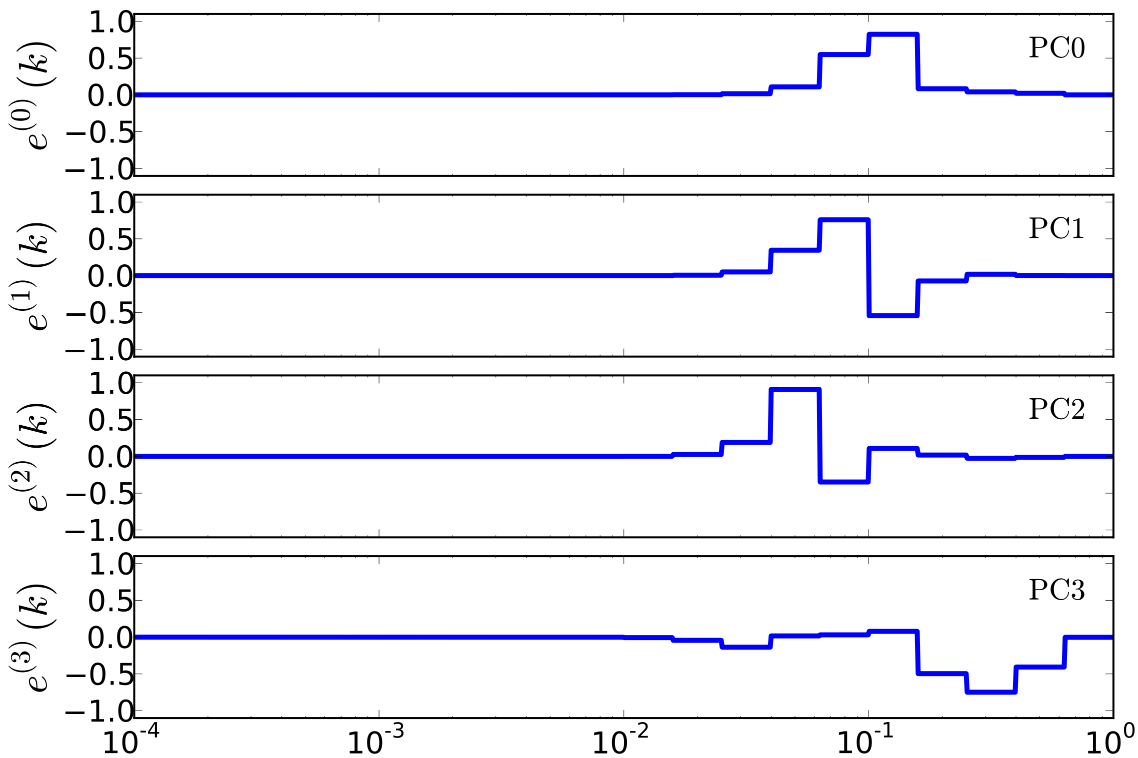


Figure 3.3: The forecasted best-measured principal components of $f_{\rm NL}(k)$ from the joint LSS + Planck data set.

at its pivot (i.e. best-determined wavenumber) from the fiducial value. The higher PCs (PC1, PC2, etc) serve to probe the k-dependence of $f_{\rm NL}$.

While the combined principal components are dominated by the contribution from the Planck PCs in this particular case, the relative strength of the LSS constraints is strongly dependent on two factors: volume of the LSS survey and, to a slightly lesser extent, fiducial (i.e. true) value of $f_{\rm NL}(k)$. We investigate these dependences further in the next section.

3.2.3 Projecting constraints on the power-law model of $f_{\rm NL}(k)$

As in section 2.3.1, we can project our Fisher matrix down to a different basis in order to study the power-law parameterization of $f_{\rm NL}(k)$ (see Equation 1.10):

$$f_{\rm NL}(k) = f^{*}_{\rm NL}\left(\frac{k}{k_{\rm piv}}\right)^{n_{f_{\rm NL}}}. \tag{3.8}$$

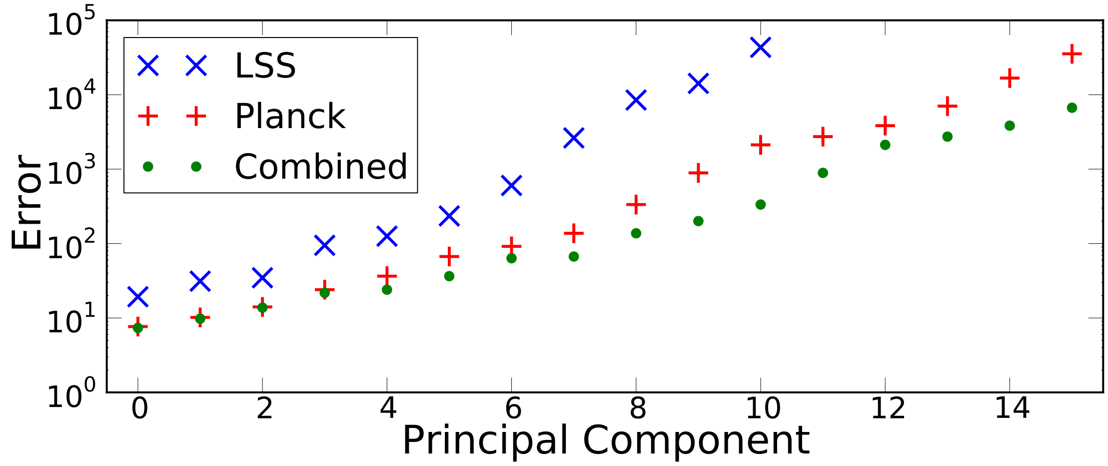


Figure 3.4: Forecasted RMS error on each principal component for LSS, Planck, and combined data sets.

We find a story similar to the one we found with the PCs; see Table 3.1. We can use the constraints on $f^*_{\rm NL}$ and $n_{f_{\rm NL}}$ to find constraints on $f_{\rm NL}(k)$ as a whole, through the usual methods of error propagation:

(3.9)

$$\sigma(f_{\rm NL}(k)) = \sqrt{\left(\frac{\partial f_{\rm NL}}{\partial f^*_{\rm NL}}\sigma(f^*_{\rm NL})\right)^2 + \left(\frac{\partial f_{\rm NL}}{\partial n_{f_{\rm NL}}}\sigma(n_{f_{\rm NL}})\right)^2 + 2\frac{\partial f_{\rm NL}}{\partial f^*_{\rm NL}}\frac{\partial f_{\rm NL}}{\partial n_{f_{\rm NL}}}C_{f^*_{\rm NL},n_{f_{\rm NL}}}},$$

where $C_{f^*_{\rm NL},n_{f_{\rm NL}}}$ is the covariance between $f^*_{\rm NL}$ and $n_{f_{\rm NL}}$. Using this relation, and given some fiducial model of $f_{\rm NL}(k)$, we can plot the forecasted constraints on $f_{\rm NL}(k)$ as a function of k. This is what we have done in Figure 3.5 for the Planck bispectrum, a

Projected errors $\sigma(f^*_{\rm NL})$ and $\sigma(n_{f_{\rm NL}})$, and the corresponding pivots				
Variable	LSS	LSS + Planck C_ℓs	Planck bispectrum	**LSS + all Planck**
$\sigma(f^*_{\rm NL})$	17	8.7	4.4	3.9
$\sigma(n_{f_{\rm NL}})$	2.0	0.85	0.29	0.22
k_{piv}	0.12	0.20	0.080	0.096

Table 3.1: Fiducial $f^*_{\rm NL} = 30$; fiducial $n_{f_{\rm NL}} = 0$. Each column's numbers are for the pivot in that column; thus the errors in the two parameters are uncorrelated in each column. LSS survey parameters are the same as in Section 2.2.2.

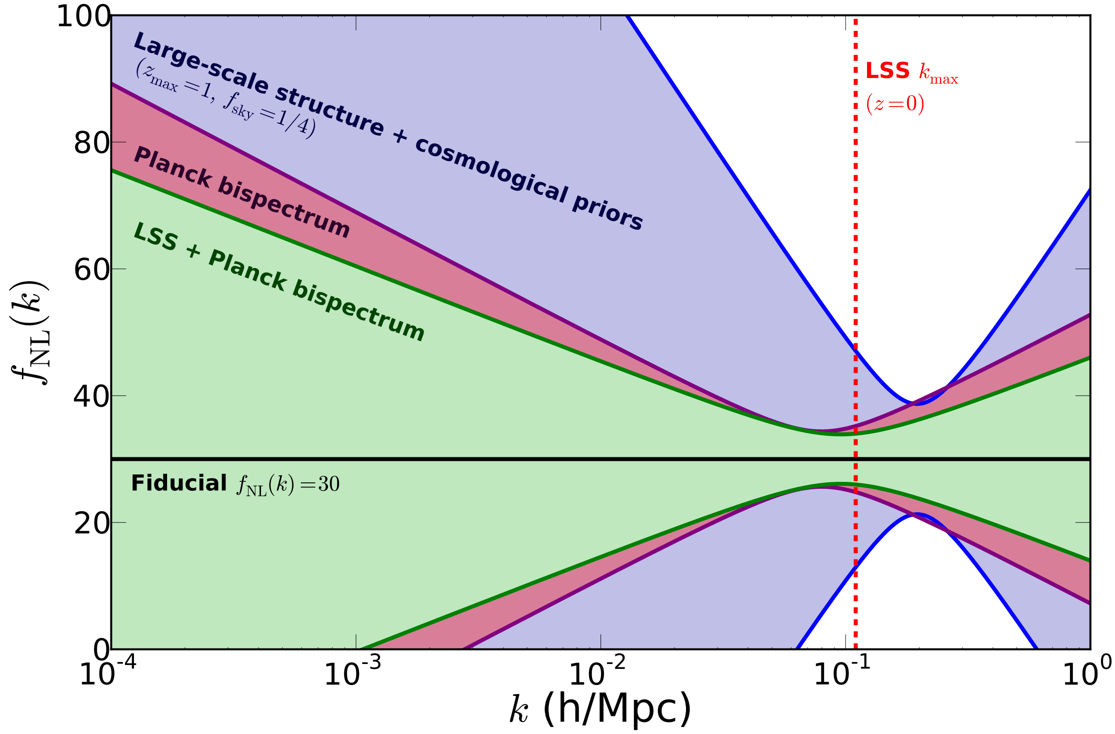


Figure 3.5: Forecasted constraints on $f_{NL}(k)$ from several different data sets, assuming the power-law model of scale-dependent non-Gaussianity: $f_{NL}(k) = f^*_{NL}(k/k_{piv})^{n_{f_{NL}}}$, projecting down from the piecewise-constant f^i_{NL} basis. The red dashed line is the maximum k for which information was kept in the LSS Fisher matrix at $z = 0$.

future large-scale structure survey, and the combination of both (along with priors on cosmological parameters from the Planck power spectrum).

The constraints on $f_{NL}(k)$ from a large-scale structure survey are quite sensitive to the survey parameters. Unlike the constraints on $f_{NL}(k)$ from the CMB bispectrum, the forecasted constraints from LSS are also sensitive to the choice made for the fiducial model of $f_{NL}(k)$, as shown in Appendix A.3. Forecasted constraints on f^*_{NL} and $n_{f_{NL}}$ for a couple of different LSS surveys, with a couple of different fiducial models, are compared to forecasted constraints from Planck in Table 3.2. (Note that all values of $n_{f_{NL}}$ are equally likely for the fiducial model where $f^*_{NL} = 0$, since $f_{NL}(k) = 0$ no matter what $n_{f_{NL}}$ is in that case.) Figures 3.6 and 3.7 are analogous

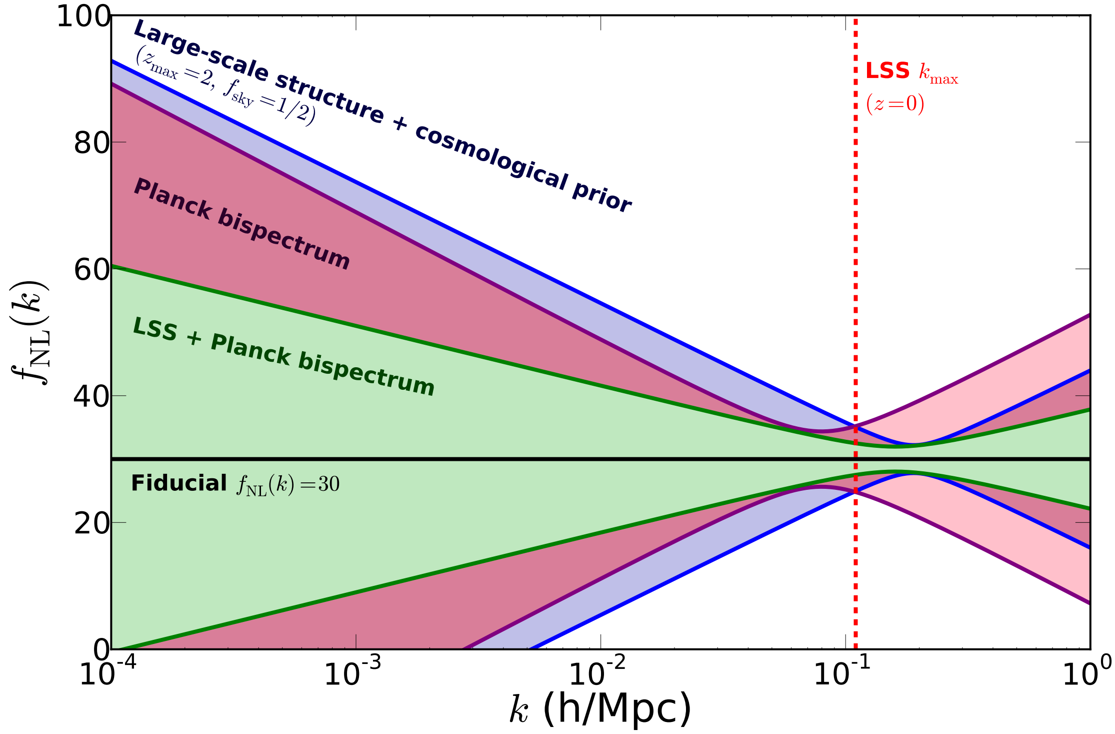


Figure 3.6: The same as Figure 3.5, but with different survey parameters for large-scale structure, similar to the planned Euclid survey.

to Figure 3.5, but for different choices of survey parameters and fiducial values of f_{NL}^{*} and $n_{f_{NL}}$, respectively.

3.3 Conclusions

In this chapter, we studied how well the generalized local model (1.6) can be probed with the combination of cosmic microwave background data and large-scale structure surveys. As in Chapter II, we started by forecasting errors on the individual parameters f_{NL}^{i} (see Fig. 3.1). We also found the best-measured linear combinations of the f_{NL}^{i} through principal component analysis (see Fig. 3.2). We also projected the Fisher matrix down to the two-parameter space for the power-law form of $f_{NL}(k)$, and then propagated the errors from those parameters to $f_{NL}(k)$ as a whole (see

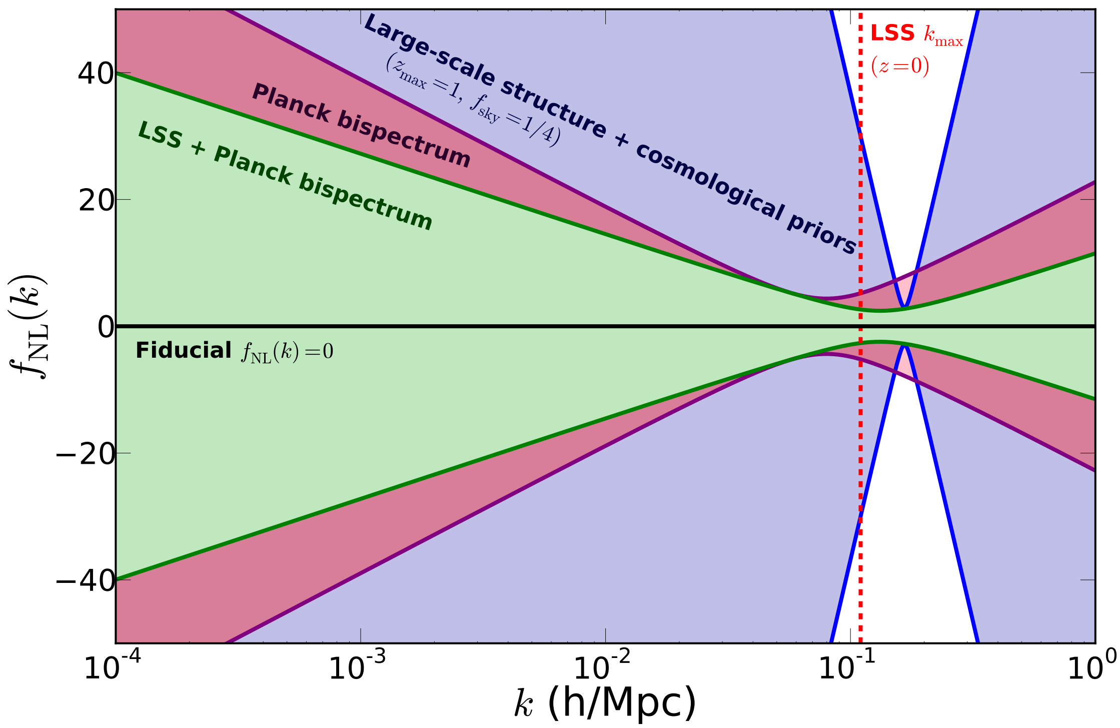


Figure 3.7: The same as Figure 3.5, but with a fiducial $f_{\rm NL}(k) = 0$.

Figures 3.5, 3.6, and 3.7.)

We found that both the bispectrum measurement from the CMB Planck survey and power spectrum measurement from an LSS survey can constrain $f_{\rm NL}(k)$ tightly in a relatively narrow range of wavenumbers around $k \simeq 0.1h\,{\rm Mpc}^{-1}$. The scale best constrained by the CMB is larger (i.e. at a smaller k) than the scale best constrained by LSS: we get complementary information about $f_{\rm NL}(k)$ from the two data sets. Constraints from the CMB and LSS should remain comparable if systematics are properly controlled for – but systematics are arguably more difficult to control for LSS surveys (witness the larger number of nuisance parameters and degeneracies in the LSS Fisher matrix). The ability of LSS to constrain $f_{\rm NL}(k)$ effectively at a wide range of scales also depends on the survey parameters and the fiducial model

Projected errors ($\sigma_{f^*_{\rm NL}}$, $\sigma_{n_{f_{\rm NL}}}$) for different surveys and different fiducial $f_{\rm NL}(k)$			
	$z_{\rm max}=1, f_{\rm sky}=1/4$	$z_{\rm max}=2, f_{\rm sky}=1/2$	Planck
Fiducial $f_{\rm NL}(k)=30$	(8.7, 0.85)	(2.2, 0.28)	(4.4, 0.29)
Fiducial $f_{\rm NL}(k)=0$	(2.9, ∞)	(0.41, ∞)	(4.4, ∞)

Table 3.2: Forecasted constraints $\sigma_{f^*_{\rm NL}}$ from different LSS surveys, assuming different fiducial models. Forecasted constraints from Planck are also shown for comparison. (All values of $n_{f_{\rm NL}}$ are equally likely in the second fiducial model, where $f^*_{\rm NL}=0$.)

of $f_{\rm NL}(k)$ chosen, as is clear from Figures 3.5 - 3.7 and Table 3.2. Nonetheless, large galaxy redshift surveys planned for the future may well be competitive with, or even better than, the constraints on the magnitude and running of $f_{\rm NL}(k)$ expected from Planck.

CHAPTER IV

Constraints on the running of local-type non-Gaussianity from WMAP 7-year data

4.1 Introduction

As mentioned elsewhere in this work (e.g. equation 1.10), a common parametrization of $f_{\rm NL}(k)$ is a simple power law:

$$f_{\rm NL}(k) = f_{\rm NL}^{*} \left(\frac{k}{k_{\rm piv}} \right)^{n_{f_{\rm NL}}} \tag{4.1}$$

Despite the relative popularity of this model, nobody has ever placed actual constraints on $n_{f_{\rm NL}}$, nor any other form of running of non-Gaussianity with scale. In this chapter, we use CMB data – specifically, the seven-year data set from the Wilkinson Microwave Anisotropy Probe (WMAP7) – to place the first-ever constraints on $n_{f_{\rm NL}}$, the running of local-type non-Gaussianity.

4.2 Estimating $n_{f_{\rm NL}}$

In order to extract information about primordial non-Gaussianity from actual CMB data, we need to have an unbiased estimator. Estimators relate the observable quantities on the CMB sky (pixels) to theoretical parameters of interest (e.g. $f_{\rm NL}$). Unfortunately, it is difficult (if not actually impossible) to construct an estimator for $n_{f_{\rm NL}}$ directly. Instead, we have adopted an alternative procedure.

We start with a fast cubic estimator for $f_{\rm NL}$ due to Komatsu, Smith, and Wandelt (Komatsu et al. (2005)) and modified it to get an estimator for $f^*_{\rm NL}$. (The details of the KSW estimator and our modification of it are in Appendix D.) We used this modified estimator to construct the likelihood as a function of both $f^*_{\rm NL}$ and $n_{f_{\rm NL}}$. Then we marginalized over $f^*_{\rm NL}$ to get the likelihood as a function of $n_{f_{\rm NL}}$ alone, which in turn gave us an estimate of $n_{f_{\rm NL}}$.

To find the likelihood, we first find a χ^2 statistic for $f^*_{\rm NL}$, given a value of $n_{f_{\rm NL}}$. The χ^2 statistic for a set of observables O_i is defined as:

$$\chi^2 \equiv \sum_i \frac{\left(O_i^{\rm obs} - O_i^{\rm theory}\right)^2}{\sigma^2_{{\rm theory},i}} \tag{4.2}$$

Taking the angular-averaged bispectrum $B_{\ell_1\ell_2\ell_3}$ as our observables, and defining $B^{\rm theory}_{\ell_1\ell_2\ell_3}(n_{f_{\rm NL}})$ as the theoretical expectation for the angular-averaged bispectrum in the case where $f^*_{\rm NL} = 1$, we have:

$$\begin{aligned}\chi^2(f^*_{\rm NL}, n_{f_{\rm NL}}) &= \sum_{\ell_1\ell_2\ell_3} \frac{\left(B^{\rm obs}_{\ell_1\ell_2\ell_3} - f^*_{\rm NL} B^{\rm theory}_{\ell_1\ell_2\ell_3}(n_{f_{\rm NL}})\right)^2}{\tilde{C}_{\ell_1}\tilde{C}_{\ell_2}\tilde{C}_{\ell_3}} \\ &= \sum_{\ell_1\ell_2\ell_3} \frac{\left(B^{\rm obs}_{\ell_1\ell_2\ell_3}\right)^2 - 2 f^*_{\rm NL} B^{\rm obs}_{\ell_1\ell_2\ell_3} B^{\rm theory}_{\ell_1\ell_2\ell_3}(n_{f_{\rm NL}}) + \left(f^*_{\rm NL} B^{\rm theory}_{\ell_1\ell_2\ell_3}(n_{f_{\rm NL}})\right)^2}{\tilde{C}_{\ell_1}\tilde{C}_{\ell_2}\tilde{C}_{\ell_3}}.\end{aligned} \tag{4.3}$$

(This works because the theoretical expectation for $B_{\ell_1\ell_2\ell_3} \propto f^*_{\rm NL}$.)

Using the skewness parameter $S(n_{f_{\rm NL}})$ from the KSW estimator (equation (D.24)), and taking advantage of the definition of the Fisher matrix $F(n_{f_{\rm NL}})$ for $f^*_{\rm NL}$ (equation (D.25)), we can rewrite χ^2 as:

$$\chi^2(f^*_{\rm NL}, n_{f_{\rm NL}}) = \left[\sum_{\ell_1\ell_2\ell_3} \frac{\left(B^{\rm obs}_{\ell_1\ell_2\ell_3}\right)^2}{\tilde{C}_{\ell_1}\tilde{C}_{\ell_2}\tilde{C}_{\ell_3}}\right] - 2 f^*_{\rm NL} S(n_{f_{\rm NL}}) + (f^*_{\rm NL})^2 F(n_{f_{\rm NL}}). \tag{4.4}$$

We can simplify this expression by introducing the following definition:

$$\chi^2_0 = \sum_{\ell_1\ell_2\ell_3} \frac{\left(B^{\rm obs}_{\ell_1\ell_2\ell_3}\right)^2}{\tilde{C}_{\ell_1}\tilde{C}_{\ell_2}\tilde{C}_{\ell_3}}. \tag{4.5}$$

χ_0^2 is the goodness-of-fit parameter for the data with respect to the $f^*_{\rm NL} = 0$ case, hence the notation. Note that the numerator of χ_0^2 is an observed quantity, and the denominator is based solely on the theoretical prediction for the power spectrum (as well as a few noise and beam parameters of WMAP). Therefore, χ_0^2 does not depend on $f^*_{\rm NL}$ or $n_{f_{\rm NL}}$ at all.

Now we can rewrite χ^2 as:

$$\chi^2(f^*_{\rm NL}, n_{f_{\rm NL}}) = \chi_0^2 - 2f^*_{\rm NL}S + (f^*_{\rm NL})^2F. \tag{4.6}$$

Completing the square, we find:

$$\chi^2(f^*_{\rm NL}, n_{f_{\rm NL}}) = F\left(f^*_{\rm NL} - \frac{S}{F}\right)^2 + \chi_0^2 - \frac{S^2}{F}. \tag{4.7}$$

Finally, we can take advantage of the definition of the modified KSW estimator itself, $\hat{f}^*_{\rm NL}(n_{f_{\rm NL}}) \equiv S/F$ (equation (D.26)):

$$\chi^2(f^*_{\rm NL}, n_{f_{\rm NL}}) = F\left(f^*_{\rm NL} - \hat{f}^*_{\rm NL}\right)^2 + \chi_0^2 - (\hat{f}^*_{\rm NL})^2F. \tag{4.8}$$

χ^2 is minimized in $f^*_{\rm NL}$ when $f^*_{\rm NL} = \hat{f}^*_{\rm NL}$:

$$\chi^2_{\rm min}(n_{f_{\rm NL}}) = \chi_0^2 - (\hat{f}^*_{\rm NL})^2F. \tag{4.9}$$

Figure 4.1 is a plot of $\chi^2_{\rm min} - \chi_0^2$ as a function of $n_{f_{\rm NL}}$.

We don't have to settle for minimizing χ^2 over $f^*_{\rm NL}$, though. We can actually find an expression for the likelihood, $\mathcal{L}(f^*_{\rm NL}, n_{f_{\rm NL}})$, and marginalize over $f^*_{\rm NL}$ to find the likelihood as a function of $n_{f_{\rm NL}}$ alone. We can get the likelihood from χ^2:

$$\mathcal{L}(n_{f_{\rm NL}}, f^*_{\rm NL}) \propto \exp\left(-\frac{\chi^2}{2}\right) = e^{-\frac{F(f^*_{\rm NL} - \hat{f}^*_{\rm NL})^2}{2}} e^{-\frac{\chi_0^2 - (\hat{f}^*_{\rm NL})^2F}{2}} \tag{4.10}$$

Figure 4.2 is a contour plot of this likelihood in the $n_{f_{\rm NL}}$ - $f^*_{\rm NL}$ plane, and figure 4.3 is a three-dimensional plot of $\mathcal{L}(n_{f_{\rm NL}}, f^*_{\rm NL})$.

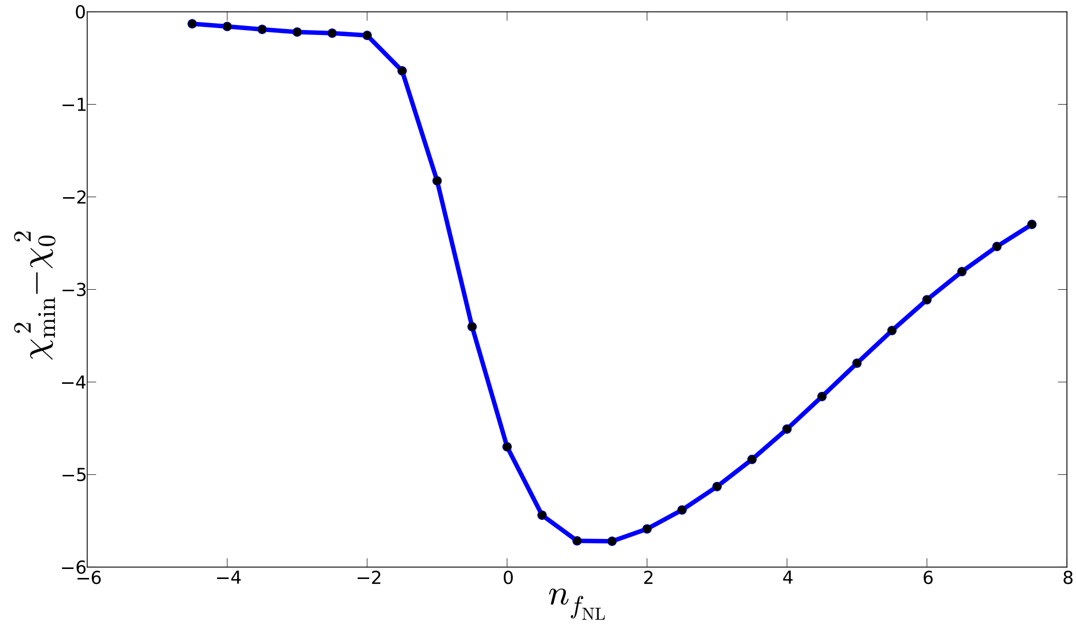


Figure 4.1: $\chi^2_{\rm min} - \chi^2_0$ as a function of $n_{f_{\rm NL}}$.

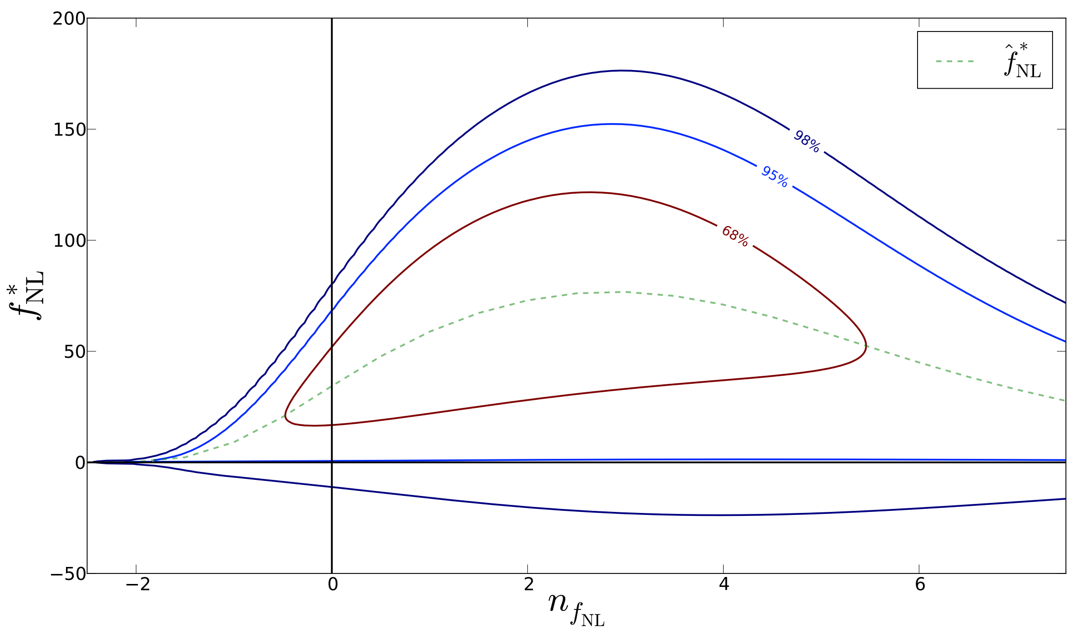


Figure 4.2: A contour plot of the likelihood in the $f^*_{\rm NL}$ - $n_{f_{\rm NL}}$ plane.

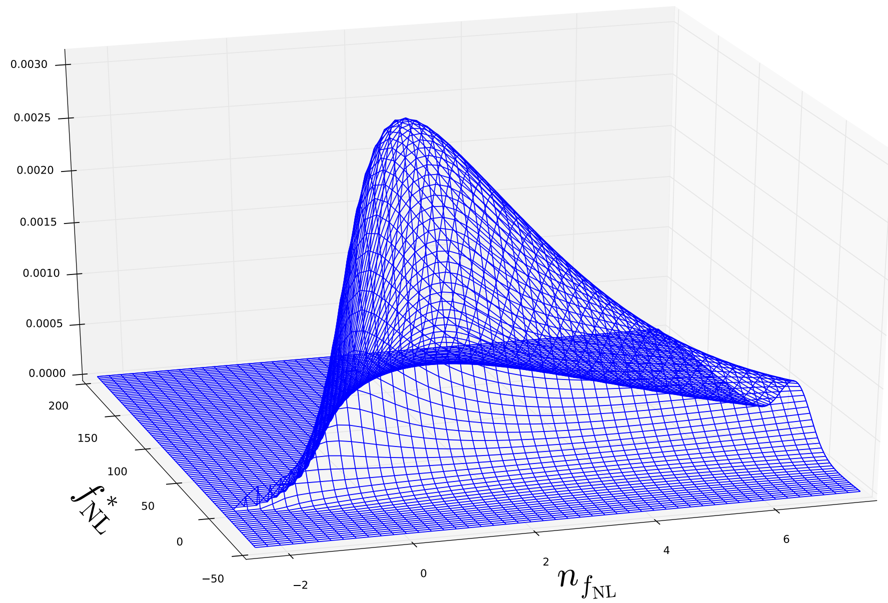


Figure 4.3: A three-dimensional plot of the likelihood, $\mathcal{L}(f^*_{\rm NL}, n_{f_{\rm NL}})$.

To marginalize over $f^*_{\rm NL}$, we integrate the likelihood:

$$\mathcal{L}(n_{f_{\rm NL}}) = \int \mathcal{L}(n_{f_{\rm NL}}, f^*_{\rm NL})\, df^*_{\rm NL} \propto \frac{1}{\sqrt{F}} e^{-\frac{\chi_0^2 - (\hat{f}^*_{\rm NL})^2 F}{2}}. \tag{4.11}$$

Remembering that χ_0^2 is constant, $e^{-\frac{\chi_0^2}{2}}$ merely contributes to the normalization, and we are left with:

$$\mathcal{L}(n_{f_{\rm NL}}) \propto \frac{1}{\sqrt{F}} e^{\frac{(\hat{f}^*_{\rm NL})^2 F}{2}}. \tag{4.12}$$

4.3 Results and conclusions

4.3.1 WMAP7 constraints on $n_{f_{\rm NL}}$

Figure 4.4 shows $\mathcal{L}$ as a function of $n_{f_{\rm NL}}$ for three different values of the pivot scale $k_{\rm piv}$. χ^2 is independent of our choice for $k_{\rm piv}$, but the likelihood itself is not, since F is inversely proportional to $k_{\rm piv}^{2n_{f_{\rm NL}}}$. This is not especially surprising, since choosing a different pivot is equivalent to choosing a different effective prior in $f^*_{\rm NL}$. The true

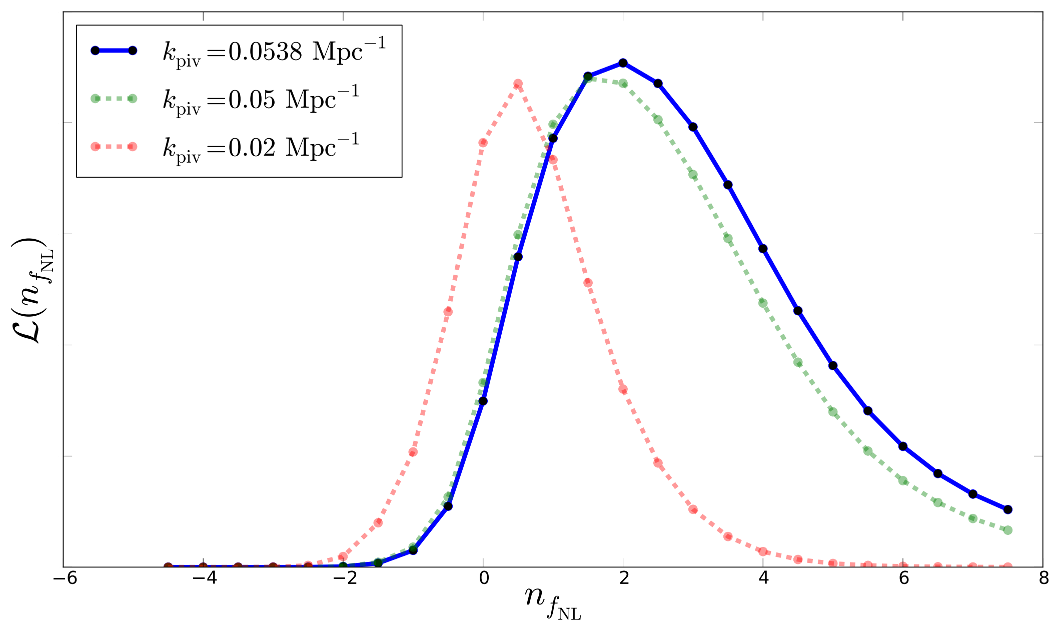


Figure 4.4: The likelihood marginalized over $f^*_{\rm NL}$ as a function of $n_{f_{\rm NL}}$ for the true pivot, along with two other pivots.

pivot scale favored by the data is the value of $k_{\rm piv}$ for which the errors in $f^*_{\rm NL}$ are uncorrelated with the errors in $n_{f_{\rm NL}}$. We find this scale by using the likelihood to calculate the covariance matrix $\mathbf{C}$ between $f^*_{\rm NL}$ and $n_{f_{\rm NL}}$:

$$\mathbf{C}_{i,j} = \langle (p_i - \bar{p}_i)(p_j - \bar{p}_j) \rangle. \tag{4.13}$$

With $\mathbf{C}$ in hand, we can find $k_{\rm piv}$ (Shandera et al. (2011)):

$$k_{\rm piv} = k_* \exp\left(- \frac{\mathbf{C}_{f^*_{\rm NL}, n_{f_{\rm NL}}}}{f^*_{\rm NL} \mathbf{C}_{n_{f_{\rm NL}}, n_{f_{\rm NL}}}} \right). \tag{4.14}$$

Here, k_* is the pivot used when evaluating $\mathbf{C}$; similarly, $f^*_{\rm NL}$ is the value used in $\mathbf{C}$. Despite the fact that k_* and $f^*_{\rm NL}$ show up in the expression, $k_{\rm piv}$ does not depend on them – the same value of $k_{\rm piv}$ will come out of (4.14) no matter what values of k_* and $f^*_{\rm NL}$ are used. We find that $k_{\rm piv}^{\rm WMAP7} \approx 0.0538\,{\rm Mpc}^{-1}$; this corresponds to the likelihood shown by the bold blue line in Figure 4.4.

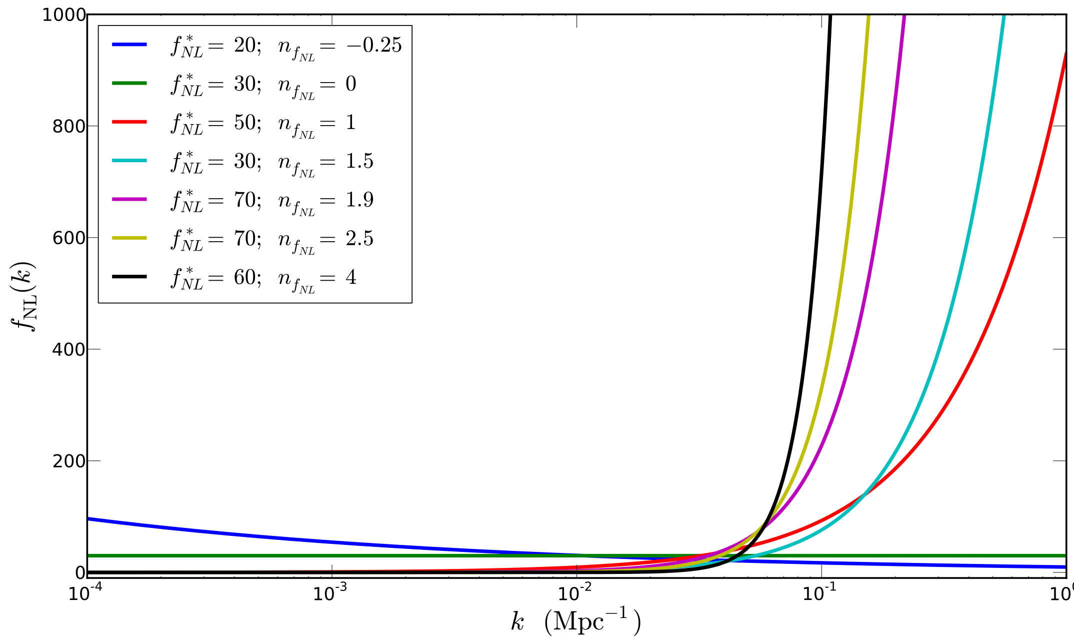


Figure 4.5: Several models of $f_{\rm NL}(k)$ with high likelihood. All of the models shown here lie within the 68% confidence region in Figure 4.2, and they all use the pivot favored by the data, $k_{\rm piv} = 0.0538$ Mpc^{-1}.

The central value for $n_{f_{\rm NL}}$ is the value which maximizes the likelihood at the correct pivot, and the uncertainty comes from the width of the likelihood (our likelihood is manifestly *not* Gaussian, so we can't just use the uncertainty from $\mathbf{C}$). Putting it all together, we have the following estimate for $n_{f_{\rm NL}}$ from the WMAP7 data, with a 68% (95%) confidence interval:

$$n_{f_{\rm NL}} = 1.9^{+2.1(+4.2)}_{-1.4(-2.1)} \tag{4.15}$$

4.3.2 Conclusions

These constraints in $n_{f_{\rm NL}}$, (4.15), are the first constraints on the scale-dependence of any form of non-Gaussianity. They are, admittedly, somewhat loose constraints – there are still a variety of power-law models for $f_{\rm NL}(k)$ that the data do not rule out (Figure 4.5). While the WMAP7 data are compatible with $n_{f_{\rm NL}} = 0$, the shape of the likelihood function does hint at a positive value for $n_{f_{\rm NL}}$. We will learn more about this hint soon with the Planck data, due out next year. For the fiducial value $f_{\rm NL} = 30$ favored by the WMAP7 data, the forecasted Planck error on $n_{f_{\rm NL}}$ in Table 3.2 is $\sigma_{n_{f_{\rm NL}}}^{\rm Planck} = 0.29$, indicating that Planck may be able to improve upon our WMAP7 constraints by nearly a full order of magnitude.

CHAPTER V

Summary and conclusions

Non-Gaussianity is a potentially powerful probe of inflationary physics in the very early universe. Single-field inflationary models with interactions, along with most multi-field models, generically produce scale-dependent non-Gaussianity. To learn more about primordial non-Gaussianity, we must infer properties of the primordial perturbations from the anisotropies in the CMB and the large-scale structure of the universe today, and we must look at higher-order correlation functions and polyspectra in order to test any given model of non-Gaussianity.

The best constraints on the Gaussianity of the universe have, until recently, come from the bispectrum of the CMB: WMAP has constrained $f_{\rm NL}$ to roughly 30 ± 20 (Komatsu et al. (2011)). But Dalal et al. (2008) pointed out that a non-zero $f_{\rm NL}$ leads to a strongly scale-dependent dark matter halo bias, which can be detected in the power spectrum of large-scale structure; this technique has since emerged as a source of constraints already competitive with the CMB (Slosar et al. (2008)).

I have focused on an extension of the local model, Equation (1.6), in which the usual local non-Gaussianity parameter $f_{\rm NL}$ is promoted to a function of scale, $f_{\rm NL}(k)$. I have paid particular attention to a piecewise-constant parametrization of $f_{\rm NL}(k)$ into a set of constants $f^i_{\rm NL}$, Equation (1.9), as well as a simple power-law model of

$f_{\rm NL}(k)$, Equation (1.10).

In Chapter II, we used forecasted constraints on the individual parameters $f^i_{\rm NL}$ from an intermediate-future galaxy survey. We also projected constraints on the power-law model of $f_{\rm NL}(k)$. We calculated the principal components of $f_{\rm NL}(k)$ to find the best-measured linear combinations of $f^i_{\rm NL}$. The sensitivity of the survey to non-Gaussianity increases with increasing k, but restricting the survey information to scales where linear perturbation theory is valid imposes a “sweet spot” in sensitivity of $k \sim 0.1h\,{\rm Mpc}^{-1}$.

In Chapter III, we studied how well the generalized local model can be probed with the combination of cosmic microwave background data and large-scale structure surveys. As in Chapter II, we started by forecasting errors on the individual parameters $f^i_{\rm NL}$. We found the principal components and forecasted errors for the power-law form of $f_{\rm NL}(k)$. We then propagated the errors from those parameters to $f_{\rm NL}(k)$ as a whole.

Constraints from the CMB and LSS should remain comparable if systematics are properly controlled for – but systematics are arguably more difficult to control for LSS surveys (witness the larger number of nuisance parameters and degeneracies in the LSS Fisher matrix). Nonetheless, large galaxy redshift surveys planned for the future may well be competitive with, or even better than, the constraints on the magnitude and running of $f_{\rm NL}(k)$ expected from Planck.

In Chapter IV, we used the WMAP7 data to obtain the first constraints on the scale-dependence of non-Gaussianity of any form. The WMAP7 data are compatible with $n_{f_{\rm NL}} = 0 : n_{f_{\rm NL}} = 1.9^{+2.1}_{-1.4}$. The Planck data, due out next year, should be able to improve on these constraints enough to tell us whether the slight hint of a positive $n_{f_{\rm NL}}$ is significant.

We are entering a very exciting era in cosmology; in the next few years, data may finally be good enough to start placing serious constraints on entire classes of inflationary models via primordial non-Gaussianity. Planck and the next generation of large-scale structure surveys will be able to constrain the non-Gaussianity of the universe down to one part in 10^5. Non-Gaussianity, if we do find it, will give us new insight into the physics at work in the first fraction of a second after the Big Bang.

APPENDICES

APPENDIX A

Finding the derivative of the halo bias with respect to $f_{\rm NL}$ and the $f_{\rm NL}^i$

If we denote the full bias of dark matter halos by $b + \Delta b$, where b represents the bias for the Gaussian fluctuations and Δb is the non-Gaussian correction, then

$$\frac{P_h}{P_R} = b^2 \left(1 + \frac{\Delta b}{b}\right)^2, \tag{A.1}$$

where P_h and P_R are the power spectra of halos and dark matter, respectively. The non-Gaussian correction to the linear peak bias to the leading order becomes

$$\frac{\Delta b}{b}(k) = \frac{\nu}{\sigma_R} \frac{1}{2P_R(k)} \int \frac{d^3q}{(2\pi)^3} B_R(k, q, |\boldsymbol{k} - \boldsymbol{q}|), \tag{A.2}$$

where B_R is the matter bispectrum on scale R. Hence, the non-Gaussian correction $\Delta b(k)$ can be expressed in terms of the primordial potential fluctuations as (Matarrese and Verde (2008)):

$$\frac{\Delta b}{b}(k) = \frac{\delta_c}{D(z)} \frac{1}{8\pi^2\sigma_R^2 \mathcal{M}_R(k)} \int_0^\infty dk_1 k_1^2 \mathcal{M}_R(k_1) \int_{-1}^1 d\mu \mathcal{M}_R(k_2) \frac{B_\phi(k_1, k_2, k)}{P_\phi(k)}. \tag{A.3}$$

We perform the integration over all triangles. The triangles' sides are k_1, k_2, and k; the cosine of the angle opposite k_2 is μ, so $k_2^2 = k_1^2 + k^2 + 2k_1 k\mu$. $\mathcal{M}_R(k)$ is the same function defined in Eq. (2.2), and the redshift dependence of the critical threshold for collapse is given as $\delta_c(z) = \delta_c/D(z)$, with $\delta_c = 1.686$.

Portions of this appendix first appeared in:
Becker, A., Huterer, D., Kadota, K., Scale-dependent non-Gaussianity as a generalization of the local model, *Journal of Cosmology and Astroparticle Physics*, 2011, vol. 1, p. 006, doi:10.1088/1475-7516/2011/01/006

A.1 Constant $f_{\rm NL}$

Eq. (A.3) leads to the famous scale-dependent bias formula in the case of a constant $f_{\rm NL}$. For this model, the bispectrum is

$$B_\phi(k_1,k_2,k_3) = 2f_{\rm NL}\,[P_\phi(k_1)P_\phi(k_2) + \text{perm.}]. \tag{A.4}$$

Through Eq. (A.3), this leads to the result

$$\begin{aligned}\frac{\Delta b}{b}(k) &= \frac{\delta_c}{D(z)}\frac{2f_{\rm NL}}{8\pi^2\sigma_R^2\mathcal{M}_R(k)}\int dk_1 k_1^2\mathcal{M}_R(k_1)P_\phi(k_1)\int d\mu\mathcal{M}_R(k_2)\left[\frac{P_\phi(k_2)}{P_\phi(k)}+2\right]\\ &\equiv \frac{2f_{\rm NL}\delta_c}{D(z)}\frac{\mathcal{F}(k)}{\mathcal{M}_R(k)},\end{aligned} \tag{A.5}$$

where

$$\mathcal{F}(k) \equiv \frac{1}{8\pi^2\sigma_R^2}\int dk_1 k_1^2\mathcal{M}_R(k_1)P_\phi(k_1)\int d\mu\mathcal{M}_R(k_2)\left[\frac{P_\phi(k_2)}{P_\phi(k)}+2\right]. \tag{A.6}$$

Note that there is a factor of 2 in Eq. (A.5) because we can exchange the order of integration of terms corresponding to k_1 and k_2.

Finally, we rewrite Eq. (A.5) by defining

$$\mathcal{F}_1(k) \equiv \frac{1}{8\pi^2\sigma_R^2\mathcal{M}_R(k)P_\phi(k)}\int dk_1 k_1^2\mathcal{M}_R(k_1)P_\phi(k_1)\int d\mu\mathcal{M}_R(k_2)P_\phi(k_2); \tag{A.7}$$

$$\mathcal{F}_2(k) \equiv \frac{2}{8\pi^2\sigma_R^2\mathcal{M}_R(k)}\int dk_1 k_1^2\mathcal{M}_R(k_1)P_\phi(k_1)\int d\mu\mathcal{M}_R(k_2). \tag{A.8}$$

Then, for constant $f_{\rm NL}$,

$$\frac{\Delta b}{b}(k) = \frac{2f_{\rm NL}\delta_c}{D(z)}\left[\mathcal{F}_1(k)+\mathcal{F}_2(k)\right], \tag{A.9}$$

and the derivative with respect to $f_{\rm NL}$ is

$$\frac{\partial}{\partial f_{\rm NL}}\left[\frac{\Delta b}{b}(k)\right] = \frac{2\delta_c}{D(z)}\left[\mathcal{F}_1(k)+\mathcal{F}_2(k)\right]. \tag{A.10}$$

A.2 Scale-dependent $f_{\rm NL}$

Now we repeat the analysis of the previous section, but we allow $f_{\rm NL}(k)$ to be an arbitrary function of scale, adopting the ansatz in Eq. (1.6). We still assume homogeneity, so $f_{\rm NL}(\vec{k}) = f_{\rm NL}(k)$. The bispectrum is given by

$$B_\phi(k_1, k_2, k_3) = 2[f_{\rm NL}(k_1)P_\phi(k_2)P_\phi(k_3) + \text{perm.}]. \tag{A.11}$$

Here, the triangle condition always holds, so that (for example) $k_1 = |\vec{k_2} + \vec{k_3}|$. Following Eq. (A.3), we get

$$\begin{aligned}\frac{\Delta b}{b}(k) &= \frac{\delta_c}{D(z)} \frac{2}{8\pi^2\sigma_R^2\mathcal{M}_R(k)} \int dk_1 k_1^2 \mathcal{M}_R(k_1)P_\phi(k_1) \\ &\times \int d\mu \mathcal{M}_R(k_2)\left[f_{\rm NL}(k)\frac{P_\phi(k_2)}{P_\phi(k)} + 2f_{\rm NL}(k_2)\right].\end{aligned} \tag{A.12}$$

This looks like Eq. (A.5) – but this time, $f_{\rm NL}(k)$ is a function, not a constant. Thus, to find the derivative of $\Delta b/b(k)$ with respect to the relevant parameters, we must parametrize $f_{\rm NL}(k)$ in a way that is valid for any general form of $f_{\rm NL}(k)$. We consider the piecewise-constant (in wavenumber) parametrization where $f_{\rm NL}(k)$ is equal to $f^i_{\rm NL}$ in the ith wavenumber bin:

$$f^i_{\rm NL} \equiv f_{\rm NL}(k_i). \tag{A.13}$$

The derivative of $\Delta b/b(k)$ with respect to these $f^i_{\rm NL}$ is:

$$\begin{aligned}\frac{\partial}{\partial f^j_{\rm NL}}\left[\frac{\Delta b}{b}(k_i)\right] &= \frac{\delta_c}{D(z)} \frac{2}{8\pi^2\sigma_R^2\mathcal{M}_R(k)} \times \\ &\left[\delta_{ij}\frac{1}{P_\phi(k)}\int dk_1 k_1^2\mathcal{M}_R(k_1)P_\phi(k_1)\int d\mu\mathcal{M}_R(k_2)P_\phi(k_2) + \right. \\ &\left. +2\int_{k_2\in k_j} dk_1 k_1^2 \mathcal{M}_R(k_1)P_\phi(k_1)\int d\mu \mathcal{M}_R(k_2)\right],\end{aligned} \tag{A.14}$$

where δ_{ij} is the Kronecker delta function. Note that the last integral over k_2 only goes over the jth wavenumber bin.

This derivative can be rewritten more concisely as

$$\frac{\partial}{\partial f_{\rm NL}^{j}}\left[\frac{\Delta b}{b}(k_i)\right] = \frac{2\delta_c}{D(z)}\left[\delta_{ij}\mathcal{F}_1(k) + \mathcal{F}_2^{j}(k)\right]. \tag{A.15}$$

The functions $\mathcal{F}_1$ and $\mathcal{F}_2$ are defined as in Eqs. (A.7) and (A.8), except that the superscript in F_2^j indicates that the integral over k_2 is to be executed only over the jth wavenumber bin.

A.2.1 The Desjacques et al. term

The new term in the bias, pointed out by Desjacques et al. (2011), is the second term of (2.7):

$$N(k) \equiv \frac{d\ln F(k)}{d\ln \sigma_R}. \tag{A.16}$$

This is not a particularly computationally friendly form. We can make it more tractable by using the chain rule:

$$N(k) = \frac{\sigma_R}{F(k)}\frac{dF}{dM}\left(\frac{d\sigma_R}{dM}\right)^{-1}. \tag{A.17}$$

Now we need to take the derivative of N with respect to the $f_{\rm NL}^{i}$, for our Fisher matrix.

$$\begin{aligned}\frac{\partial N}{\partial f_{\rm NL}^{i}} &= \sigma_R\left(\frac{d\sigma_R}{dM}\right)^{-1}\frac{\partial}{\partial f_{\rm NL}^{i}}\left[\frac{1}{F(k)}\frac{dF}{dM}\right] \\ &= \frac{\sigma_R}{F}\left(\frac{d\sigma_R}{dM}\right)^{-1}\frac{\partial}{\partial f_{\rm NL}^{i}}\left[\frac{d}{dM}\left(\frac{\partial F}{\partial f_{\rm NL}^{i}}\right) - \frac{1}{F}\frac{dF}{dM}\frac{\partial F}{\partial f_{\rm NL}^{i}}\right].\end{aligned} \tag{A.18}$$

Equations (A.17) and (A.18) are everything we need to properly account for the new term in our Fisher matrix. Note that σ_R and $\frac{d\sigma_R}{dM}$ are the only z-dependent quantities in N; since their z-dependence is linear and exactly the same, it cancels entirely, leaving N independent of z.

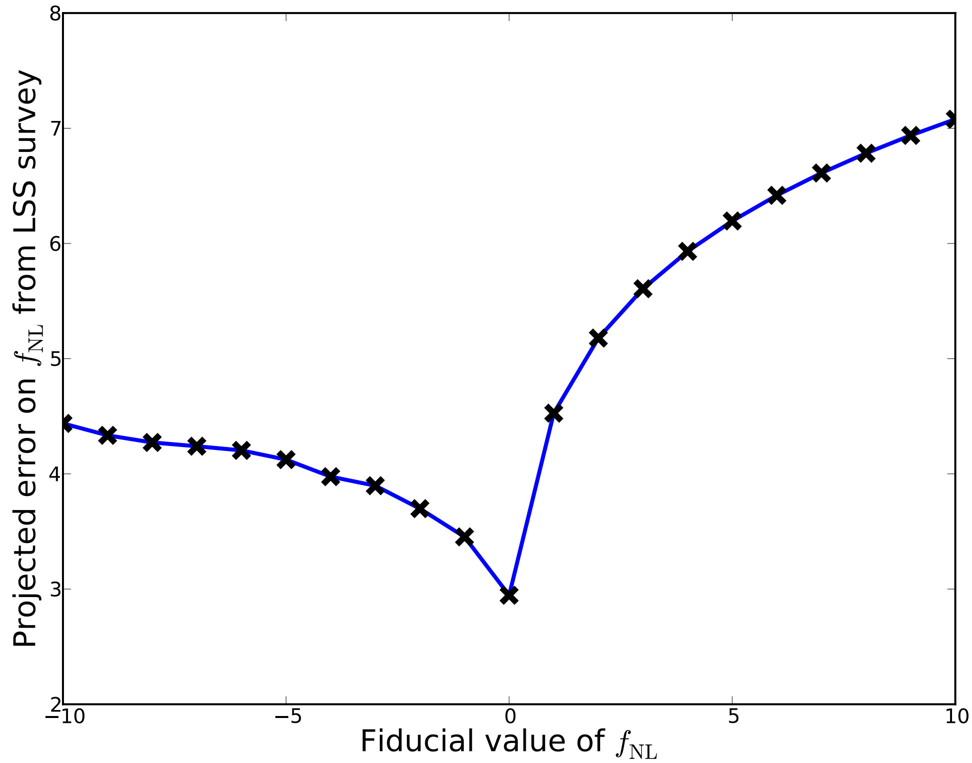


Figure A.1: How the choice of fiducial $f_{\rm NL}$ affects the forecasted constraints on constant $f_{\rm NL}$ from a future galaxy survey. See text for analytic explanation for why results are the best at a fiducial value of $f_{\rm NL} = 0$.

A.3 The effect of the fiducial value on constraints

The fiducial value of $f_{\rm NL}$ affects the Fisher matrix – and thus the forecasted constraints on $f_{\rm NL}$ itself – because the relationship between $P_h(k)$ and $f_{\rm NL}$ is non-linear. The fiducial $f_{\rm NL}$ enters the Fisher matrix through the bias, by way of $P_h = (b^2(k))P(k)$. Assuming $P_h(k) \gg 1/n$ (a reasonable assumption at large angular scales where non-Gaussianity constraints largely come from and where shot noise is negligible), we find that the Fisher matrix element corresponding to $f_{\rm NL} = $ const is

$$F^{\rm LSS} \propto \int \left(\frac{\partial b(k)}{\partial f_{\rm NL}}\right)^2 b^{-2}(k) dk = \int \left(\frac{\Delta b(k)}{f_{\rm NL}\,(b_0 + \Delta b(k))}\right)^2 dk. \tag{A.19}$$

Thus, the expression on the right-hand side will, in general, be dependent on the choice of fiducial f_{NL}. Since $|\Delta b(k)|$ blows up at small k, in that regime we have:

$$\left(\frac{\Delta b(k)}{f_{\mathrm{NL}}\left(b_0 + \Delta b(k)\right)}\right)^2 \approx \frac{1}{f_{\mathrm{NL}}^2}. \tag{A.20}$$

At large k, $\Delta b(k)$ goes to 0, taking the entire expression with it. Thus, the integral is dominated by the contribution at low k, meaning we should expect a maximal Fisher matrix element around a fiducial $f_{\mathrm{NL}} = 0$. And indeed, that is what we see in Figure A.1: the forecasted constraints on f_{NL} from a given sky survey depend on the fiducial value chosen, with the tightest constraints at $f_{\mathrm{NL}} = 0$.

APPENDIX B

Statistical methods: Fisher matrices, principal components, and all that.

B.1 Fisher information matrices: a brief introduction.

Fisher matrices are powerful tools for forecasting the constraints placed on a set of parameters from an expected future data set. It is a purely analytic method; no likelihood evaluation or parameter search of any kind is required. This makes it a particularly fast and convenient method for error forecasting. In this subsection, I will give a brief overview of the derivation and application of Fisher matrices in the abstract. More details about how I performed specific Fisher matrix calculations are provided in Chapters II and III.

B.1.1 Bayes's theorem, likelihood, and the Fisher information matrix

Any reasonable interpretation of probability admits the following truth about conditional probabilities:

$$P(A|B) = \frac{P(AB)}{P(B)} \tag{B.1}$$

In other words, the probability of A given B is equal to the probability of both A and B divided by the probability of B. Given B.1 and some other basic axioms of

Portions of this appendix first appeared in:
Becker, A., Huterer, D., Kadota, K., Scale-dependent non-Gaussianity as a generalization of the local model, *Journal of Cosmology and Astroparticle Physics*, 2011, vol. 1, p. 006, doi:10.1088/1475-7516/2011/01/006

probability, we have the following chain of reasoning concerning the probability of a hypothesis H and some data D:

$$P(H|D) = P(HD)/P(D)$$

$$P(D|H) = P(DH)/P(H)$$

$$P(H|D)P(D) = P(HD) = P(DH) = P(D|H)P(H)$$

$$\therefore \boxed{P(H|D) = P(D|H)P(H)/P(D)} \tag{B.2}$$

This is Bayes's theorem. The probability of the hypothesis given the data, $P(H|D)$, is equal to the probability of the data given the hypothesis, $P(D|H)$, multiplied by the probability of the hypothesis, $P(H)$, divided by the probability of the data, $P(D)$. $P(H)$ is known as the *prior probability*; $P(D) = \int P(D|H')P(H')dH'$ is the probability of the data marginalized over all hypotheses, and is therefore called the *marginal probability*; $P(H|D)$ is the *posterior probability*; finally, $P(D|H)$, the probability of observing the data given the truth of the the hypothesis, is known as the *likelihood*. Bayes theorem, then, can be restated:

$$\text{posterior probability} = \text{likelihood} \times \frac{\text{prior probability}}{\text{marginal probability}}$$

The marginal probability depends only on the data (and the chosen hypothesis space), not on H itself; thus, it can be viewed as an overall normalization factor. For a likelihood function sharply peaked in hypothesis space, it (nearly) doesn't matter what method you're using to assign priors to your hypotheses – the likelihood function will pick out a narrow band of hypotheses so long as we have sufficiently informative data.

But how do we quantify the notion of "sufficiently informative" for our data? Our data is sufficiently informative if the models in our model space are sensitive to the parameters our data tell us about. We already know that our models are sensitive to the parameters we're measuring if the likelihood function for those parameters is sharply peaked in our model space. So we can quantify how useful our data will be for distinguishing among different models in terms of the peak curvature of the likelihood function – and we measure a function's curvature by taking its second derivative. Thus, we arrive at the Fisher information matrix, often just called the Fisher matrix:

$$\boxed{F_{ij} = \left\langle -\frac{\partial^2 \ln \mathcal{L}}{\partial p_i \partial p_j} \right\rangle} \tag{B.3}$$

Here, $\mathcal{L}$ is the likelihood, and the p_i are the parameters of interest in the model (e.g. any cosmological parameters). The Fisher matrix gives us a quantitative measure of how well a data set can choose among available models – and thus, how much information a data set can contain about the parameters that determine our models. The brackets $\langle\rangle$ indicate an expectation value taken over realizations of the data; this enables us to find an analytic expression for the Fisher matrix. We assume that the data are distributed according to a multivariate Gaussian; in that case, the covariance matrix of the data C has all the information about the distribution of the data:

$$\mathcal{L} = \frac{1}{(2\pi)^{n/2} |\det C|^{1/2}} \exp\left[-\frac{1}{2} (d - \bar{d})_i^T C_{ij}^{-1} (d - \bar{d})_j \right], \tag{B.4}$$

where d_i are the data (with $\bar{d}_i$ the mean for each i) and C_{ij} is the covariance of the data. After some tedious but straightforward algebra, (B.3) and (B.4) combine to give an expression for the Fisher matrix:

$$\boxed{F_{ij} = \frac{1}{2} \mathrm{Tr}[C^{-1} C_{,i} C^{-1} C_{,j}] + \bar{d}_{,i}^T C^{-1} \bar{d}_{,j}} \tag{B.5}$$

where $_{,i}$ is the partial derivative with respect to p_i.

In most cases, $\bar{d}_{,i}$ and C will depend on the values chosen for the parameters p_i; in order to calculate the Fisher matrix F_{ij}, one must first choose fiducial values for these parameters. So the Fisher matrix can be used to forecast constraints on the errors in the p_i – as we are about to see – but it obviously cannot give any information about the most likely values for the p_i themselves.

B.1.2 Using Fisher matrices to estimate parameter errors

The most straightforward way to use Fisher matrices in error forecasting is through the Cramér-Rao bound, which states that an error in a cosmological parameter p_i will be greater than or equal to the corresponding Fisher matrix element:

$$\boxed{\sigma(p_i) \geq \begin{cases} \sqrt{(F^{-1})_{ii}} & \text{(marginalized error)} \\ 1/\sqrt{F_{ii}} & \text{(unmarginalized error)} \end{cases}} \tag{B.6}$$

Here, the marginalized error is the error in p_i marginalized over the uncertainties in all the other parameters in the Fisher matrix F, while the unmarginalized error is the error in p_i while holding all the other parameters perfectly fixed. The marginalized errors are generally the quantities of interest, since we are usually trying to determine the values of several parameters at once from the same set of data. Cramér-Rao only gives us a lower bound on the marginalized error – but in practice, we assume that the data will saturate the bound, allowing us to effectively forecast the best achievable errors for a given set of observations using Fisher matrices.

B.2 Calculating the error on an arbitrary parametrized $f_{\rm NL}(k)$

Projecting the constraints from an old set of parameters $f^i_{\rm NL} \equiv f_{\rm NL}(k_i)$ ($i = 1, 2, \ldots, N$) to new parameters (which we can call q; $j = 1, 2, \ldots, M$ for some M) is in principle straightforward. The Fisher matrix in the new parameters, $F^{\rm new}$, is given by

$$F^{\text{new}}_{i,j} = \sum_{k,l=1}^{N} \frac{\partial p^k}{\partial q^i}\frac{\partial p^l}{\partial q^j} F_{kl} \tag{B.7}$$

so that

$$F^{\text{new}} \equiv \mathcal{P}^T F\, \mathcal{P}, \tag{B.8}$$

where $\mathcal{P}_{ij} = \partial p^i / \partial q^j$ is the derivative matrix of old parameters with respect to new.

Let us look at a couple of examples. Projecting to the case

$$f_{\text{NL}}(k) = f_{\text{NL}} = \text{const} \tag{B.9}$$

is particularly easy, since $\mathcal{P}$ is the column vector with $\mathcal{P}_{i1} = df^i_{\text{NL}}/df_{\text{NL}} = 1$. Then F^{new}_{ij} is a 1×1 matrix that quantifies information on f_{NL}, given by

$$F^{\text{new}}_{11} = \sum_{k,l} F_{kl}. \tag{B.10}$$

The error on f_{NL} is of course given simply by $\sigma(f_{\text{NL}}) = 1/\sqrt{F^{\text{new}}_{11}}$.

Another example is given by the function

$$f_{\text{NL}}(k) = \left(\frac{k}{k_0}\right)^{n_{\text{NG}}}, \tag{B.11}$$

with two parameters, k_0 and n_{NG}. Then one can show that (labeling $k_0 \equiv q_1$ and $n_{\text{NG}} \equiv q_2$):

$$\mathcal{P}_{i1} = -\frac{n_{\text{NG}}}{k_0}\left(\frac{k_i}{k_0}\right)^{n_{\text{NG}}}; \tag{B.12}$$

$$\mathcal{P}_{i2} = \ln\left(\frac{k_i}{k_0}\right)\left(\frac{k_i}{k_0}\right)^{n_{\text{NG}}}. \tag{B.13}$$

Then, using Eq. (B.8), one can simply obtain the 2×2 Fisher matrix in k_0 and n_{NG}.

B.3 Principal components of $f_{\rm NL}(k)$

We now show how to decompose the measurement of $f_{\rm NL}(k)$ in principal components, which are essentially the eigenmodes of the covariance matrix for the aforementioned parameters $f_{\rm NL}(k_i)$. This method has been widely used in cosmology, including applications to parametrizing and describing dark energy (Huterer and Starkman (2003); Albrecht et al. (2009)). It allows us to order the best-to-worst measured weights in wavenumber of the function $f_{\rm NL}(k)$.

Let the function $f_{\rm NL}(k)$ be described in terms of piecewise constant parameters $f_{\rm NL}^i \equiv f_{\rm NL}(k_i)$, where

$$f_{\rm NL}(k) = \sum_{i=1}^{N} p_i \Theta_i(k). \tag{B.14}$$

Here, $\Theta(k) \equiv \left[H(k - k_i^{\rm lower}) - H(k - k_i^{\rm upper})\right]$ is the top-hat function of unit height over the ith wavenumber bin, and we assume a total of N bins. $k_i^{\rm lower}$ and $k_i^{\rm upper}$ are the wavenumber bin boundaries, and H is the Heaviside step function. We have effectively expanded the function around the zero value, though this is not crucial: the left-hand side could be $f_{\rm NL}(k) - f_{\rm NL}^{\rm fid}(k)$, for any fiducial $f_{\rm NL}^{\rm fid}(k)$, and the formalism still follows.

The Fisher matrix F is the inverse covariance matrix in the original piecewise-constant parameters p_i, so that $F_{ij}^{-1} = \langle p_i p_j \rangle - \langle p_i \rangle \langle p_j \rangle$. We first diagonalize the Fisher matrix F:

$$F = W^T D W, \tag{B.15}$$

where D is diagonal and W is some orthogonal matrix. The vector of uncorrelated parameters, $\mathbf{q}$, is related to the vector of original parameters $\mathbf{p}$ via

$$\mathbf{q} = W \mathbf{p}, \tag{B.16}$$

and it is easy to check that the $\mathbf{q}$ are uncorrelated; that is, $\langle \mathbf{q}\,\mathbf{q}^T \rangle = D^{-1}$. The rows of W are therefore the new parameters.

Thus, to calculate the principal components:

1. Obtain the full Fisher matrix for N parameters p_i, plus the cosmological parameters $\Omega_b h^2, \Omega_{CDM} h^2, H_0, w, \log A_s$, and n_s.
2. Marginalize over the cosmological parameters by inverting this larger Fisher matrix, taking the $N \times N$ submatrix, then inverting back to get the Fisher matrix of the p_i; we call *this* Fisher matrix F
3. Diagonalize F as in Eq. (B.15)
4. The rows of W are the principal components. More precisely, $q_a = \sum_i W_{ai} p_i$, and q_a are the PCs.

Let us now change notation slightly (to agree with the commonly used one, e.g. Huterer and Starkman (2003)), and define the shape of the a-th principal component in i-th redshift bin as $\alpha_i^{(a)}$, so that $\alpha_i^{(a)} \equiv W_{ai}$. Then we can represent the a-th principal component, $e^{(a)}(k)$, in terms of the original parameters p_i as[1]

$$e^{(a)}(k) = \sum_{i=1}^{N} \alpha_i^{(a)} p_i \, \Theta_i(k). \tag{B.17}$$

The PCs are obviously uncorrelated, and their eigenvalues λ_a, so that

$$\langle e^{(a)} e^{(b)} \rangle \equiv \sum_{i,j=1}^{N} \alpha_i^{(a)} \alpha_j^{(b)} \langle p_i p_j \rangle = \frac{\delta_{ab}}{\lambda_a}. \tag{B.18}$$

where, recall, $\lambda_a \equiv D_{aa}$.

Finally, let us calculate the coefficients $c^{(a)}$ in the expansion in principal components of an arbitrary $f_{\rm NL}(k)$

$$f_{\rm NL}(k) = \sum_{a=1}^{N} c_a e^{(a)}(k). \tag{B.19}$$

[1] This is basically the continuous version of the relation $q_a = \sum_i W_{ai} p_i$.

Let coefficients $f_{\rm NL}^i$ describe $f_{\rm NL}(k)$ in our original basis, so that $f_{\rm NL}(k) = \text{const} \equiv \sum_i f_{\rm NL}^i p_i \Theta_i(k)$, with $f_{\rm NL}^i$ being left arbitrary for now. Then, taking the expectation value of the product with $e^{(b)}$, we get

$$\langle f_{\rm NL}(k) e^{(b)} \rangle \equiv \frac{c_b}{\lambda_b} = \left\langle \left(\sum_{i=1}^{N} f_{\rm NL}^i \, p_i \right) \times \left(\sum_{j=1}^{N} \alpha_j^{(a)} p_j \right) \right\rangle \tag{B.20}$$

$$= \sum_{i,j=1}^{N} f_{\rm NL}^i \, \alpha_j^{(a)} (F^{-1})_{ij}, \tag{B.21}$$

so that

$$c_a = \lambda_a \sum_{i,j=1}^{N} f_{\rm NL}^i \, \alpha_j^{(a)} \, (F^{-1})_{ij}. \tag{B.22}$$

For example, in the simplest case of constant $f_{\rm NL}(k)$, where $f_{\rm NL}^i = \text{const} \equiv f_{\rm NL}$, the coefficients of the principal components in the expansion of $f_{\rm NL}(k)$ are

$$c_a = \lambda_a \, f_{\rm NL} \sum_{ij} \alpha_j^{(a)} \, (F^{-1})_{ij} \qquad (\text{for } f_{\rm NL}(k) \equiv f_{\rm NL} = \text{const}). \tag{B.23}$$

APPENDIX C

Calculating the CMB bispectrum Fisher matrix for local-type non-Gaussianity

C.1 Calculating the CMB bispectrum

The non-averaged bispectrum is:

$$B_{\ell_1\ell_2\ell_3,m_1m_2m_3} = \langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3} \rangle \tag{C.1}$$

where the a_{lm}s are the coefficients on the spherical harmonic decomposition of the CMB sky. The a_{lm}s can be related to the Bardeen curvature perturbations $\Phi(\mathbf{k})$ by:

$$a_{\ell m} = \int d^2\hat{\mathbf{k}}\, \frac{\Delta T(\hat{\mathbf{k}})}{T} Y^*_{\ell m}(\hat{\mathbf{k}}) = 4\pi(-i)^\ell \int \frac{d^3k}{(2\pi)^3}\, \Phi(\mathbf{k}) g_\ell(k) Y^*_{\ell m}(\hat{\mathbf{k}}) \tag{C.2}$$

Here, $g_\ell(k)$ is the CMB temperature radiation transfer function. There are several conventions used for this transfer function; $g_\ell(k)$ is related to the transfer function $T_\ell(k)$ found in (Gibelyou et al. (2010)) by:

$$g_\ell(k) = \frac{(-i)^\ell}{\sqrt{2\ell(\ell+1)}} T_\ell(k) \tag{C.3}$$

We will be using yet another convention, as both of the transfer functions above lead to messy prefactors later on. Throughout this paper, we denote the radiation transfer functions as $t_\ell(k)$, defined as:

$$t_\ell(k) = \frac{1}{(-i)^\ell} g_\ell(k) = \frac{1}{\sqrt{2\ell(\ell+1)}} T_\ell(k) \tag{C.4}$$

With these transfer functions, (C.2) becomes:

$$a_{\ell m} = \frac{4\pi}{\sqrt{2\ell(\ell+1)}}(-1)^{\ell}\int \frac{d^3k}{(2\pi)^3}\,\Phi(\mathbf{k})t_{\ell}(k)Y^*_{\ell m}(\hat{\mathbf{k}}) \tag{C.5}$$

One last word on transfer function conventions: these transfer functions connect the CMB sky to the Bardeen curvature perturbations, not the primordial curvature perturbations.

The angular-averaged bispectrum $B_{\ell_1\ell_2\ell_3}$ is related to the raw bispectrum $B_{\ell_1\ell_2\ell_3,m_1,m_2,m_3}$ of (C.1) by the relation:

$$B_{\ell_1\ell_2\ell_3} = \sum_{m_1,m_2,m_3}\begin{pmatrix}\ell_1 & \ell_2 & \ell_3\\ m_1 & m_2 & m_3\end{pmatrix}B_{\ell_1\ell_2\ell_3,m_1,m_2,m_3} \tag{C.6}$$

Here, $\begin{pmatrix}\ell_1 & \ell_2 & \ell_3\\ m_1 & m_2 & m_3\end{pmatrix}$ is the Wigner $3j$-symbol. This symbol ensures that $\ell_1+\ell_2+\ell_3$ is even, $m_1+m_2+m_3=0$, and the triangle inequality ($|\ell_i-\ell_j|\le\ell_k\le\ell_i+\ell_j$) is met for all i,j,k.[1] Substituting (C.1) and (C.5) into (C.6), we obtain the following expression for the angular-averaged bispectrum:

$$\begin{aligned} B_{\ell_1\ell_2\ell_3} = (4\pi)^3(-1)^{\ell_1+\ell_2+\ell_3}\sum_{m_1,m_2,m_3}\begin{pmatrix}\ell_1 & \ell_2 & \ell_3\\ m_1 & m_2 & m_3\end{pmatrix}\int\frac{d^3k_1}{(2\pi)^3}\frac{d^3k_2}{(2\pi)^3}\frac{d^3k_3}{(2\pi)^3}\\ \times Y^*_{\ell_1m_1}(\hat{\mathbf{k_1}})Y^*_{\ell_2m_2}(\hat{\mathbf{k_2}})Y^*_{\ell_3m_3}(\hat{\mathbf{k_3}})t_{\ell_1}(k_1)t_{\ell_2}(k_2)t_{\ell_3}(k_3)\langle\Phi(\mathbf{k_1})\Phi(\mathbf{k_2})\Phi(\mathbf{k_3})\rangle \end{aligned} \tag{C.7}$$

Using the definition of the Bardeen curvature bispectrum, B_Φ,

$$\langle\Phi(\mathbf{k_1})\Phi(\mathbf{k_2})\Phi(\mathbf{k_3})\rangle = (2\pi)^3\delta(\mathbf{k_1}+\mathbf{k_2}+\mathbf{k_3})B_\Phi(k_1,k_2,k_3), \tag{C.8}$$

we find:

$$\begin{aligned} B_{\ell_1\ell_2\ell_3} = \frac{1}{\pi^3}\sum_{m_1,m_2,m_3}\begin{pmatrix}\ell_1 & \ell_2 & \ell_3\\ m_1 & m_2 & m_3\end{pmatrix}\int d^3k_1\,d^3k_2\,d^3k_3 Y^*_{\ell_1m_1}(\hat{\mathbf{k_1}})Y^*_{\ell_2m_2}(\hat{\mathbf{k_2}})Y^*_{\ell_3m_3}(\hat{\mathbf{k_3}})\\ \times t_{\ell_1}(k_1)t_{\ell_2}(k_2)t_{\ell_3}(k_3)\delta(\mathbf{k_1}+\mathbf{k_2}+\mathbf{k_3})B_\Phi(k_1,k_2,k_3). \end{aligned} \tag{C.9}$$

[1] There are some computational difficulties that arise when evaluating the $3j$-symbol for high $l_{1,2,3}$; see Appendix C.3.2 for more on this.

(The prefactor of $(-1)^{\ell_1+\ell_2+\ell_3}$ vanished because the Wigner $3j$-symbol ensures $\ell_1 + \ell_2 + \ell_3$ is even.) Taking advantage of several identities in Wang and Kamionkowski (2000) (their (12) and (13)), the orthogonality of the spherical harmonics, and the Gaunt integral identity (Komatsu and Spergel (2001)), this becomes:

$$B_{\ell_1\ell_2\ell_3} = \left(\frac{2}{\pi}\right)^3 I_{\ell_1\ell_2\ell_3} \int k_1^2 dk_1\, k_2^2 dk_2\, k_3^2 dk_3 B_\Phi(k_1, k_2, k_3) t_{\ell_1}(k_1) t_{\ell_2}(k_2) t_{\ell_3}(k_3)$$

$$\times \int_0^\infty r^2 dr\, j_{\ell_1}(k_1 r) j_{\ell_2}(k_2 r) j_{\ell_3}(k_3 r), \tag{C.10}$$

where $I_{\ell_1\ell_2\ell_3}$ is the Gaunt integral

$$I_{\ell_1\ell_2\ell_3} = \sqrt{\frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4\pi}} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}. \tag{C.11}$$

The real-space integral is now a one-dimensional integral in the spherical coordinate r, starting at our location and ending at infinity. This real-space coordinate is the difference in the conformal time $\Delta\eta = \int_{t_e}^{t_0} \frac{dt}{a} = c(\tau_0 - \tau_e)$ between the time when the CMB was emitted and the time when we saw it. Equivalently, it is the difference between the radius of the particle horizon of the observable universe when the CMB was observed and that radius when the CMB was first emitted. Thus, nearly all of the contribution to the integral in r comes from a short period of time around the surface of last scattering, and there are no physical contributions beyond $r > r_{\max} = \eta_0 = c\tau_0 \approx 14.6$ Gpc. For our purposes, when performing this integral in Chapter III, we sampled the integral 150 times between $r_{\max}$ and $r_{\max} - 2r_*$, where $r_{\max} - r_*$ is the comoving distance to the surface of last scattering. We also sampled 50 times between $r_{\max} - 2r_*$ and 0 to capture any impact that late-time effects might have had. Increasing the sampling rate did not significantly improve our results.

C.1.1 Bispectrum and derivatives for $f_{\rm NL}$ and $f_{\rm NL}(k)$

Using (C.10) along with (3.4), we get the following expression for the angular-averaged CMB bispectrum in the constant $f_{\rm NL}$ case:

$$B_{\ell_1\ell_2\ell_3} = 2\Delta_\phi^2 f_{\rm NL}\left(\frac{2}{\pi}\right)^3 I_{\ell_1\ell_2\ell_3}\int k_1^2dk_1\, k_2^2dk_2\, k_3^2dk_3\left(\frac{1}{k_1^{3-(n_s-1)}k_2^{3-(n_s-1)}}+\text{perm.}\right)$$

$$\times\, t_{\ell_1}(k_1)t_{\ell_2}(k_2)t_{\ell_3}(k_3)\int_0^\infty r^2dr\; j_{\ell_1}(k_1r)j_{\ell_2}(k_2r)j_{\ell_3}(k_3r) \tag{C.12}$$

Following equations 33 and 34 from Yadav and Wandelt (2010) (where they are themselves following Komatsu and Spergel (2001), equations 17 and 18), we'll define a pair of functions, $\alpha_\ell(r)$ and $\beta_\ell(r)$, to help us rewrite (C.12) in a more computationally friendly way.

$$\alpha_\ell(r) \equiv \frac{2}{\pi}\int k^2 t_\ell(k)j_\ell(kr)dk \tag{C.13}$$

$$\beta_\ell(r) \equiv \frac{2}{\pi}\int k^{-(2-n_s)} t_\ell(k)j_\ell(kr)dk \tag{C.14}$$

Now (C.12) looks like this:

$$B_{\ell_1\ell_2\ell_3} = 2\Delta_\phi^2 f_{\rm NL} I_{\ell_1\ell_2\ell_3}\int_0^\infty r^2dr\,(\alpha_{\ell_1}(r)\beta_{\ell_2}(r)\beta_{\ell_3}(r)+\text{perm.}) \tag{C.15}$$

and (naturally)

$$\frac{\partial B_{\ell_1\ell_2\ell_3}}{\partial f_{\rm NL}} = \frac{1}{f_{\rm NL}}B_{\ell_1\ell_2\ell_3}. \tag{C.16}$$

For the scale-dependent $f_{\rm NL}(k)$ case, we use (3.6) to find that the angular-averaged CMB bispectrum is:

$$\frac{\partial B_{\ell_1\ell_2\ell_3}}{\partial f^i_{\rm NL}} = 2\Delta_\phi^2 I_{\ell_1\ell_2\ell_3}\int_0^\infty r^2dr\left(\alpha^i_{\ell_1}(r)\beta_{\ell_2}(r)\beta_{\ell_3}(r)+\text{perm.}\right) \tag{C.17}$$

where α^i_ℓ is:

$$\alpha^i_\ell(r) \equiv \frac{2}{\pi}\int_{k_i^{\rm lower}}^{k_i^{\rm upper}} k^2 t_\ell(k)j_\ell(kr)dk. \tag{C.18}$$

Polarization and cross-terms

The bispectrum for multiple fields is a simple extension of the single field case. By analogy with (C.1) and (C.2), the multiple-field bispectrum is

$$B^{pqr}_{\ell_1\ell_2\ell_3,m_1m_2m_3} = \langle a^p_{\ell_1 m_1} a^q_{\ell_2 m_2} a^r_{\ell_3 m_3}\rangle, \tag{C.19}$$

where

$$a^p_{\ell m} = \frac{4\pi}{\sqrt{2\ell(\ell+1)}}(-1)^\ell \int \frac{d^3k}{(2\pi)^3}\, \Phi(\mathbf{k}) t^p_\ell(k) Y^*_{\ell m}(\hat{\mathbf{k}}) \tag{C.20}$$

and $t^i_\ell(k)$ is either the temperature or polarization radiation transfer function. Using these definitions and running through equations (C.7) through (C.17) again, it's pretty clear that we can rewrite the bispectrum for multiple fields very easily if we just change (C.13), (C.14), and (C.18) slightly:

$$\alpha^p_\ell(r) \equiv \frac{2}{\pi}\int k^2 t^p_\ell(k) j_\ell(kr) dk; \tag{C.21}$$

$$\beta^p_\ell(r) \equiv \frac{2}{\pi}\int k^{-(2-n_s)} t^p_\ell(k) j_\ell(kr) dk; \tag{C.22}$$

$$\alpha^{p,i}_\ell(r) \equiv \frac{2}{\pi}\int_{k_i^{\text{lower}}}^{k_i^{\text{upper}}} k^2 t^p_\ell(k) j_\ell(kr) dk. \tag{C.23}$$

So for the constant f_{NL} case, we have

$$\frac{\partial B^{pqr}_{\ell_1\ell_2\ell_3}}{\partial f_{\text{NL}}} = 2\Delta^2_\phi I_{\ell_1\ell_2\ell_3}\int_0^\infty r^2 dr \left(\alpha^p_{\ell_1}(r)\beta^q_{\ell_2}(r)\beta^r_{\ell_3}(r) + \text{perm.}\right) \tag{C.24}$$

while for the piecewise-constant $f_{\text{NL}}(k)$ case, we have:

$$\frac{\partial B^{pqr}_{\ell_1\ell_2\ell_3}}{\partial f^i_{\text{NL}}} = 2\Delta^2_\phi I_{\ell_1\ell_2\ell_3}\int_0^\infty r^2 dr \left(\alpha^{p,i}_{\ell_1}(r)\beta^q_{\ell_2}(r)\beta^r_{\ell_3}(r) + \text{perm.}\right) \tag{C.25}$$

C.2 The covariance of the bispectrum

It is usually a good assumption to consider only the Gaussian contribution to the covariance of the bispectrum, **C**. Using Wick's theorem, one can straightforwardly show (Liguori et al. (2010); Babich and Zaldarriaga (2004); Spergel and Goldberg (1999)):

$$\mathbf{C}_{\ell_1\ell_2\ell_3} = C_{\ell_1}C_{\ell_2}C_{\ell_3} \tag{C.26}$$

where

$$C_\ell = C_\ell^{CV} + \sigma_\ell^2 W_\ell = C_\ell^{CV} + C_\ell^N \tag{C.27}$$

C_ℓ^{CV} is cosmic variance, while C_ℓ^N is the variance due to the noise and beam width in the survey. σ_ℓ^2 is the variance of the noise in the survey per pixel, and W_ℓ is a "window" term relating to the survey beam type and width (Cooray and Hu (2000); Knox (1995)).[2] For an experiment with multiple frequency channels (such as Planck or WMAP), the basic form of equation (C.27) still holds, but finding C_ℓ^N is slightly trickier (Cooray and Hu (2000)):

$$\frac{1}{C_\ell^N} = \sum_\nu \frac{1}{C_\ell^N(\nu)} = \sum_\nu \frac{1}{\sigma_\ell^2(\nu)W_\ell(\nu)}. \tag{C.28}$$

For uncorrelated Gaussian noise, $\sigma_\ell^2(\nu) = \sigma^2(\nu)$ is constant, and you can find its value for a particular experiment fairly easily; for example, the Planck beam width and noise parameters are found in the Planck mission "blue book."

We have only been dealing with temperature (TT), but it is not significantly harder to add in polarization (EE) and cross (TE) terms. The covariance matrix here is (Yadav et al. (2007); Babich and Zaldarriaga (2004))

$$(\mathbf{C}^{-1}_{\ell_1\ell_2\ell_3})_{lmn,pqr} = (C^{-1}_{\ell_1})_{lp}(C^{-1}_{\ell_2})_{mq}(C^{-1}_{\ell_3})_{nr}, \tag{C.29}$$

where

$$C_\ell = \begin{pmatrix} C_\ell^{TT} & C_\ell^{TE} \\ C_\ell^{TE} & C_\ell^{EE} \end{pmatrix}. \tag{C.30}$$

[2]Confusingly, Cooray and Hu (2000) uses w^{-1} for what we are calling σ^2.

Noise is dealt with in the same way as in (C.27) for C_ℓ^{TT} and C_ℓ^{EE} in (C.30). Assuming that the noise for T and E are uncorrelated, $\sigma_{TE}^2 = \langle \Delta T \Delta E \rangle = \langle \Delta T \rangle \langle \Delta E \rangle = 0$, and thus $C_\ell^{N,TE} = 0$ for all ℓ.

C.3 Computational details

C.3.1 ℓ sampling and binning

In evaluating equation (3.7), we do not actually use every $\ell \leq \ell_{\max}$; that would be incredibly computationally expensive. Instead, we sample and bin in ℓ. The binning in ℓ is progressive, not fixed-width: all ℓs are kept up through $\ell = 40$, at which point sampling drops off gradually until, at $\ell \gtrsim 100$, only every tenth ℓ is sampled. The "width" of the bins in ℓ are given by the equation

$$\Delta_{\ell_i} = \frac{1}{2}\left[(\ell_i - \ell_{i-1}) + (\ell_{i+1} - \ell_i)\right] = \frac{1}{2}(\ell_{i+1} - \ell_{i-1}). \tag{C.31}$$

C.3.2 Calculating the Wigner $3j$-symbol

We need to be able to calculate the Wigner $3j$-symbol for large (> 1000) values of $\ell_{1,2,3}$ in order to evaluate many of the expressions we're interested in. Unfortunately, the $3j$ function built in to the GNU Scientific Library can't properly evaluate the symbol for $\ell_{1,2,3} \gtrsim 70$. Thus, we were forced to create our own special-purpose $3j$-evaluator. Thankfully, we're only interested in the special case $m_{1,2,3} = 0$; as it turns out, in this case, the $3j$-symbol reduces to (see Wolfram Mathworld: http://mathworld.wolfram.com/Wigner3j-Symbol.html):

$$\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix} = \begin{cases} (-1)^g \sqrt{\frac{(2g-2\ell_1)!(2g-2\ell_2)!(2g-2\ell_3)!}{(2g+1)!}} \frac{g!}{(g-\ell_1)!(g-\ell_2)!(g-\ell_3)!} & \text{if } L = 2g; \\ 0 & \text{if } L = 2g+1, \end{cases} \tag{C.32}$$

where $L = \ell_1 + \ell_2 + \ell_3$. Since (C.32) involves evaluating the factorials of relatively large numbers when any of $\ell_{1,2,3}$ are large, we used Stirling's approximation to perform the factorials – but we needed the factorials to remain accurate even when the arguments were small, so we used six terms in the approximation.

APPENDIX D

The KSW estimator and the modified KSW estimator

D.1 The KSW estimator

Komatsu et al. (2005) found a fast cubic estimator for $f_{\rm NL}$ based on a full-sky CMB temperature map; Yadav et al. (2007) and Yadav et al. (2008) extended that estimator to deal with polarization, sky cuts, and inhomogeneous noise. I will refer to this estimator as the KSW estimator for convenience's sake.

We start by recalling from Appendix C several useful definitions and equations relating the primordial curvature bispectrum to that of the CMB. The angular-averaged CMB bispectrum $B_{\ell_1\ell_2\ell_3}$ is related to the shape function of the primordial curvature bispectrum F_Φ through the equation

$$B^{\rm theory}_{\ell_1\ell_2\ell_3} = \left(\frac{2}{\pi}\right)^3 I_{\ell_1\ell_2\ell_3} \int (k_1k_2k_3)^2\, dk_1\, dk_2\, dk_3\, F_\Phi(k_1,k_2,k_3)\, t_{\ell_1}(k_1)\, t_{\ell_2}(k_2)\, t_{\ell_3}(k_3) \times \int_0^\infty r^2 dr\; j_{\ell_1}(k_1 r) j_{\ell_2}(k_2 r) j_{\ell_3}(k_3 r), \tag{D.1}$$

where $I_{\ell_1\ell_2\ell_3}$ is the Gaunt integral

$$I_{\ell_1\ell_2\ell_3} = \sqrt{\frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4\pi}} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}. \tag{D.2}$$

We can reduce this to a considerably simpler form in the case of local non-Gaussianity (i.e. when $F_\Phi = F_\Phi^{\rm local}$; see equation (1.5)) :

$$B^{\rm theory}_{\ell_1\ell_2\ell_3}(f_{\rm NL}) = 2 f_{\rm NL} I_{\ell_1\ell_2\ell_3} \int_0^\infty r^2 dr\, \left(\alpha_{\ell_1}(r)\beta_{\ell_2}(r)\beta_{\ell_3}(r) + \text{perm.}\right) \tag{D.3}$$

where $\alpha_\ell(r)$ and $\beta_\ell(r)$ are defined (using a slightly different convention from Appendix C, to play nicely with the output from CAMB) as

$$\alpha_\ell(r) \equiv \frac{2}{\pi}\int k^2 t_\ell(k) j_\ell(kr) dk \tag{D.4}$$

$$\beta_\ell(r) \equiv \frac{2}{\pi}\int k^2 P_\Phi(k) t_\ell(k) j_\ell(kr) dk \tag{D.5}$$

Given a set of spherical harmonic coefficients $a_{\ell m}$ for the CMB sky, we can define a set of "filtered" maps, A and B:

$$A(\hat{\mathbf{n}}, r) \equiv \sum_{\ell,m} \alpha_\ell(r) \frac{b_\ell}{\tilde{C}_\ell} a_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}); \tag{D.6}$$

$$B(\hat{\mathbf{n}}, r) \equiv \sum_{\ell,m} \beta_\ell(r) \frac{b_\ell}{\tilde{C}_\ell} a_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}), \tag{D.7}$$

where $\tilde{C}_\ell = b_\ell^2 C_\ell + N_\ell$ is the power spectrum corrected for beam width and noise. Komatsu et al. (2005) construct a skewness parameter S from these filtered maps:

$$S \equiv \int r^2 dr \int d^2\hat{\mathbf{n}} A(\hat{\mathbf{n}}, r) B^2(\hat{\mathbf{n}}, r) \tag{D.8}$$

Equation (D.8) is the computationally friendly form of the skewness parameter, and we can skip straight to (D.15) if we just want to calculate a full-sky estimator for $f_{\rm NL}$. But to see *how* it leads us to that estimator, we have to do a little more work.

Keeping in mind that the observed CMB bispectrum is defined as

$$B^{\rm obs.}_{\ell_1\ell_2\ell_3} = \langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3} \rangle, \tag{D.9}$$

it is not hard to see that S reduces to

$$S = \sum_{\ell_1 \le \ell_2 \le \ell_3} \frac{B^{\rm obs}_{\ell_1\ell_2\ell_3} \tilde{B}^{\rm theory}_{\ell_1\ell_2\ell_3}(f_{\rm NL}=1)}{\tilde{C}_{\ell_1}\tilde{C}_{\ell_2}\tilde{C}_{\ell_3}} \tag{D.10}$$

where

$$\tilde{B}^{\rm theory}_{\ell_1\ell_2\ell_3}(f_{\rm NL}) = b_{\ell_1} b_{\ell_2} b_{\ell_3} B^{\rm theory}_{\ell_1\ell_2\ell_3}(f_{\rm NL}). \tag{D.11}$$

Performing a least-squares fit of $B^{\rm obs}_{\ell_1\ell_2\ell_3}$ to $\tilde{B}^{\rm theory}$, we find (Komatsu et al. (2005)):

$$S \approx f_{\rm NL} \sum_{\ell_1 \le \ell_2 \le \ell_3} \frac{\left(\tilde{B}^{\rm theory}_{\ell_1\ell_2\ell_3}(f_{\rm NL}=1)\right)^2}{\tilde{C}_{\ell_1}\tilde{C}_{\ell_2}\tilde{C}_{\ell_3}} \tag{D.12}$$

$\tilde{B}^{\rm theory}_{\ell_1\ell_2\ell_3}(f_{\rm NL}=1) = \partial\tilde{B}^{\rm theory}_{\ell_1\ell_2\ell_3}/\partial f_{\rm NL}$, because $B^{\rm theory}_{\ell_1\ell_2\ell_3}$ is proportional to $f_{\rm NL}$. Therefore, we can write the Fisher matrix F for $f_{\rm NL}$ as (see (3.5)):

$$F = \sum_{\ell_1 \le \ell_2 \le \ell_3} \left(\frac{\partial\tilde{B}^{\rm theory}_{\ell_1\ell_2\ell_3}}{\partial f_{\rm NL}}\right)^2 \frac{1}{\tilde{C}_{\ell_1}\tilde{C}_{\ell_2}\tilde{C}_{\ell_3}} = \sum_{\ell_1 \le \ell_2 \le \ell_3} \frac{\left(\tilde{B}^{\rm theory}_{\ell_1\ell_2\ell_3}(f_{\rm NL}=1)\right)^2}{\tilde{C}_{\ell_1}\tilde{C}_{\ell_2}\tilde{C}_{\ell_3}}. \tag{D.13}$$

This, in turn, means we can rewrite (D.12) as

$$S \sim f_{\rm NL} F. \tag{D.14}$$

Thus, the KSW estimator for $f_{\rm NL}$ is:

$$\hat{f}_{\rm NL} \equiv \frac{S}{F} \tag{D.15}$$

While this estimator works well for a full-sky map, it breaks down for a cut-sky map. To get around this, an extra term is introduced into the estimator (Yadav et al. (2008)) to account for the spurious signal introduced by the sky cut:

$$\hat{f}_{\rm NL} = \frac{S_{\rm cut}}{F} = \frac{\frac{1}{f_{\rm sky}}S + S_{\rm linear}}{F}. \tag{D.16}$$

$S_{\rm linear}$ is:

$$\begin{aligned} S_{\rm linear} = -\frac{1}{f_{\rm sky}} \int r^2 dr \int d^2\hat{\mathbf{n}} \, & \left[A(\hat{\mathbf{n}},r)\langle B^2_{\rm sim}(\hat{\mathbf{n}},r)\rangle_{MC} \right. \\ & \left. +2B(\hat{\mathbf{n}},r)\langle A_{\rm sim}(\hat{\mathbf{n}},r)B_{\rm sim}(\hat{\mathbf{n}},r)\rangle_{MC}\right]. \end{aligned} \tag{D.17}$$

The subscripted filtered maps $A_{\rm sim}$ and $B_{\rm sim}$ are generated from Monte Carlo realizations of the cut CMB sky; the brackets $\langle\rangle_{MC}$ indicate an average over all Monte Carlo maps. The Monte Carlo maps were produced using the prescription laid out in Appendix A of the WMAP5 paper (Komatsu et al. (2009)); the only difference (aside from our use of the WMAP7 data) is that we used a uniform weighting for the maps, rather than the slightly more complicated weighting given there, since it only results in a marginal improvement of the estimation of $f_{\rm NL}$. We created the Monte Carlo maps in Python; we plugged these Monte-Python maps into HEALPix, by way of HealPy, to do the forwards and backwards spherical harmonic transforms required to obtain the A and B maps.

D.2 Modifying the KSW estimator for a power-law $f_{\rm NL}(k)$

It is fairly simple to modify the KSW estimator for the case of a power-law $f_{\rm NL}(k)$ of the form

$$f_{\rm NL}(k) = f^*_{\rm NL}\left(\frac{k}{k_{\rm piv}}\right)^{n_{f_{\rm NL}}}. \tag{D.18}$$

We want an estimator for the parameter $f^*_{\rm NL}$. Note that the pivot scale, $k_{\rm piv}$, is completely degenerate with $f^*_{\rm NL}$; the choice of pivot scale is largely arbitrary, and in fact we will see that $k_{\rm piv}$ cancels entirely from some (but not all!) quantities of interest.

To get our new estimator, start with the shape function for the bispectrum associated with this $f_{\rm NL}(k)$:

$$F_\Phi = 2(f_{\rm NL}(k_1)P(k_2)P(k_3) + \text{perm.}) = 2\frac{f^*_{\rm NL}}{k_{\rm piv}^{n_{f_{\rm NL}}}}(k_1^{n_{f_{\rm NL}}}P(k_2)P(k_3) + \text{perm.}). \tag{D.19}$$

Plugging (D.19) into (D.1), and deploying the usual tricks, we get:

$$B^{\rm theory}_{\ell_1\ell_2\ell_3}(f^*_{\rm NL}, n_{f_{\rm NL}}) = 2f^*_{\rm NL}I_{\ell_1\ell_2\ell_3}\int_0^\infty r^2 dr\left(\gamma_{\ell_1}(n_{f_{\rm NL}}, r)\beta_{\ell_2}(r)\beta_{\ell_3}(r) + \text{perm.}\right). \tag{D.20}$$

Here, $\gamma_\ell(n_{f_{\rm NL}}, r)$ takes the role of $\alpha_\ell(r)$, and is similarly defined:

$$\gamma_\ell(n_{f_{\rm NL}}, r) \equiv \frac{2}{\pi}\frac{1}{k_{\rm piv}^{n_{f_{\rm NL}}}}\int k^{2+n_{f_{\rm NL}}}\, t_\ell(k) j_\ell(kr) dk. \tag{D.21}$$

We can use $\gamma_\ell(r)$ to write down a new filtered map $G(\hat{\mathbf{n}}, r)$,

$$G(n_{f_{\rm NL}}, \hat{\mathbf{n}}, r) \equiv \sum_{\ell,m} \gamma_\ell(n_{f_{\rm NL}}, r) \frac{b_\ell}{\tilde{C}_\ell} a_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}), \tag{D.22}$$

and we can use $G(n_{f_{\rm NL}}, \hat{\mathbf{n}}, r)$ to write down a new skewness parameter $S(n_{f_{\rm NL}})$.

$$S(n_{f_{\rm NL}}) \equiv \int r^2 dr \int d^2\hat{\mathbf{n}} G(n_{f_{\rm NL}}, \hat{\mathbf{n}}, r) B^2(\hat{\mathbf{n}}, r) \tag{D.23}$$

In the case where $n_{f_{\rm NL}} = 0$, $\gamma_\ell(n_{f_{\rm NL}}, r) = \alpha_\ell(r)$ and $S(n_{f_{\rm NL}})$ trivially reduces to (D.8).

The same argument that takes us from (D.8) to (D.10) applies here too, so $S(n_{f_{\rm NL}})$ must reduce to

$$S(n_{f_{\rm NL}}) = \sum_{\ell_1 \le \ell_2 \le \ell_3} \frac{B^{\rm obs}_{\ell_1\ell_2\ell_3} \left(\tilde{B}^{\rm theory}_{\ell_1\ell_2\ell_3}(f^*_{\rm NL} = 1, n_{f_{\rm NL}})\right)}{\tilde{C}_{\ell_1}\tilde{C}_{\ell_2}\tilde{C}_{\ell_3}} \tag{D.24}$$

We can write the Fisher matrix $F(n_{f_{\rm NL}})$ for $f^*_{\rm NL}$ at a given value of $n_{f_{\rm NL}}$ as:

$$F = \sum_{\ell_1 \le \ell_2 \le \ell_3} \left(\frac{\partial \tilde{B}^{\rm theory}_{\ell_1\ell_2\ell_3}(n_{f_{\rm NL}})}{\partial f^*_{\rm NL}}\right)^2 \frac{1}{\tilde{C}_{\ell_1}\tilde{C}_{\ell_2}\tilde{C}_{\ell_3}} = \sum_{\ell_1 \le \ell_2 \le \ell_3} \frac{\left(\tilde{B}^{\rm theory}_{\ell_1\ell_2\ell_3}(n_{f_{\rm NL}}, f^*_{\rm NL} = 1)\right)^2}{\tilde{C}_{\ell_1}\tilde{C}_{\ell_2}\tilde{C}_{\ell_3}}. \tag{D.25}$$

The least-squares fit (D.12) still holds, so we have the following unbiased estimator for $f^*_{\rm NL}$:

$$\hat{f}^*_{\rm NL} = \frac{S(n_{f_{\rm NL}})}{F_{f^*_{\rm NL}}(n_{f_{\rm NL}})} \tag{D.26}$$

To account for a sky cut, the same arguments used by Yadav et al. (2008) hold here, as we are still using a cubic estimator. Thus, our actual estimator for $f^*_{\rm NL}$ is

$$\hat{f}^*_{\rm NL} = \frac{S_{\rm cut}(n_{f_{\rm NL}})}{F} = \frac{\frac{1}{f_{\rm sky}} S(n_{f_{\rm NL}}) + S_{\rm linear}(n_{f_{\rm NL}})}{F}, \tag{D.27}$$

where F is the appropriately-modified Fisher matrix element for $f^*_{\rm NL}$, and $S_{\rm linear}(n_{f_{\rm NL}})$ is

$$\begin{aligned} S_{\rm linear}(n_{f_{\rm NL}}) = &-\frac{1}{f_{\rm sky}} \int r^2 dr \int d^2\hat{\mathbf{n}} \left[G(n_{f_{\rm NL}}, \hat{\mathbf{n}}, r) \langle B^2_{\rm sim}(\hat{\mathbf{n}}, r)\rangle_{MC} \right. \\ &\left. + 2B(\hat{\mathbf{n}}, r) \langle G_{\rm sim}(n_{f_{\rm NL}}, \hat{\mathbf{n}}, r) B_{\rm sim}(\hat{\mathbf{n}}, r)\rangle_{MC} \right]. \end{aligned} \tag{D.28}$$

Now we have an estimator for $f^*_{\rm NL}$ – and more importantly, we have a skewness parameter for $f^*_{\rm NL}$, which allows us to get the likelihood function for $n_{f_{\rm NL}}$ in Chapter IV.

Bibliography

Afshordi, N. and Tolley, A. J. Primordial non-gaussianity, statistics of collapsed objects, and the Integrated Sachs-Wolfe effect. *Phys. Rev. D*, 78:123507, 2008, [arXiv:0806.1046].

Albrecht, A. and Steinhardt, P. J. Cosmology for grand unified theories with radiatively induced symmetry breaking. *Phys. Rev. Letters*, 48:1220–1223, 1982.

Albrecht, A. J. et al. Findings of the Joint Dark Energy Mission Figure of Merit Science Working Group. 2009, [arXiv:0901.0721].

Babich, D., Creminelli, P. and Zaldarriaga, M. The shape of non-gaussianities. *JCAP*, 08:009, 2004, [astro-ph/0405356].

Babich, D. and Zaldarriaga, M. Primordial Bispectrum Information from CMB Polarization. *Phys. Rev. D*, 70(8):083005, 2004, [astro-ph/0408455].

Becker, A., Huterer, D. and Kadota, K. Scale-dependent non-Gaussianity as a generalization of the local model. *JCAP*, 1:6, 2011, [arXiv:1009.4189].

Byrnes, C. T., Gerstenlauer, M., Nurmi, S., Tasinato, G. and Wands, D. Scale-dependent non-Gaussianity probes inflationary physics. *JCAP*, 10:004, 2010, [arXiv:1007.4277].

Chen, X. Running Non-Gaussianities in DBI Inflation. *Phys. Rev.*, D 72:123518, 2005, [astro-ph/0507053].

Cooray, A. R. and Hu, W. Imprint of reionization on the cosmic microwave background bispectrum. *ApJ*, 534:533–550, 2000, [astro-ph/9910397].

Creminelli, P. and Zaldarriaga, M. Single field consistency relation for the 3-point function. *JCAP*, 0410:006, 2004, [astro-ph/0407059].

Cunha, C., Huterer, D. and Doré, O. Primordial non-gaussianity from the covariance of galaxy cluster counts. *Phys. Rev. D*, 82(2):023004, 2010, [arXiv:1003.2416].

Dalal, N., Dore, O., Huterer, D. and Shirokov, A. The imprints of primordial non-gaussianities on large- scale structure: scale dependent bias and abundance of virialized objects. *Phys. Rev. D*, 77:123514, 2008, [arXiv:0710.4560].

Desjacques, V., Jeong, D. and Schmidt, F. Accurate predictions for the scale-dependent galaxy bias from primordial non-gaussianity. *Phys. Rev. D*, 84:061301, 2011, [arXiv:1105.3476].

Desjacques, V. and Seljak, U. Signature of primordial non-Gaussianity of ϕ^3-type in the mass function and bias of dark matter haloes. *Phys. Rev. D*, 81:023006, 2010, [arXiv:0907.2257].

Desjacques, V., Seljak, U. and Iliev, I. T. Scale-dependent bias induced by local non-Gaussianity: a comparison to N-body simulations. *MNRAS*, 396:85–96, 2009, [arXiv:0811.2748].

Elsner, F. and Wandelt, B. D. Improved Simulation of Non-Gaussian Temperature and Polarization Cosmic Microwave Background Maps. *ApJS*, 184:264–270, 2009, [arXiv:0909.0009].

Feldman, H. A., Kaiser, N. and Peacock, J. A. Power spectrum analysis of three-dimensional redshift surveys. *ApJ*, 426:23–37, 1994, [astro-ph/9304022].

Gibelyou, C., Huterer, D. and Fang, W. Detectability of large-scale power suppression in the galaxy distribution. *Phys. Rev. D*, 82(12):123009, 2010, [arXiv:1007.0757].

Grinstein, B. and Wise, M. B. Nongaussian Fluctuations and the Correlations of Galaxies or Rich Clusters of Galaxies. *ApJ*, 310:19–22, 1986.

Guth, A. H. Inflationary universe: A possible solution to the horizon and flatness problems. *Phys. Rev. D*, 23:347–356, 1981.

Huterer, D. and Starkman, G. Parameterization of dark-energy properties: A principal- component approach. *Phys. Rev. Lett.*, 90:031301, 2003, [astro-ph/0207517].

Knox, L. Determination of inflationary observables by cosmic microwave background anisotropy experiments. *Phys. Rev. D*, 52:4307, 1995, [astro-ph/9504054].

Komatsu, E. and Spergel, D. N. Acoustic signatures in the primary microwave background bispectrum. *Phys. Rev. D*, 63(6):063002, 2001, [arXiv:astro-ph/0005036].

Komatsu, E., Spergel, D. N. and Wandelt, B. D. Measuring Primordial Non-Gaussianity in the Cosmic Microwave Background. *ApJ*, 634:14–19, 2005, [astro-ph/0305189].

Komatsu, E. et al. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations:Cosmological Interpretation. *ApJS*, 180:330–376, 2009, [arXiv:0803.0547].

Komatsu, E. et al. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *ApJS*, 192:18, 2011, [arXiv:1001.4538].

Liguori, M., Sefusatti, E., Fergusson, J. R. and Shellard, E. P. S. Primordial Non-Gaussianity and Bispectrum Measurements in the Cosmic Microwave Background and Large-Scale Structure. *Advances in Astronomy*, 2010:980523, 2010, [arXiv:1001.4707].

Linde, A. D. Chaotic inflation. *Physics Letters B*, 129:177–181, 1983.

Linde, A. D. and Mukhanov, V. F. Nongaussian isocurvature perturbations from inflation. *Phys. Rev. D*, 56:535–539, 1997, [astro-ph/9610219].

LoVerde, M., Miller, A., Shandera, S. and Verde, L. Effects of Scale-Dependent Non-Gaussianity on Cosmological Structures. *JCAP*, 04:014, 2008, [arXiv:0711.4126].

Luo, X.-c. and Schramm, D. N. Testing for the gaussian nature of cosmological density perturbations through the three-point temperature correlation function. *Phys. Rev. Lett.*, 71:1124–1127, 1993, [astro-ph/9305009].

Lyth, D. H. and Wands, D. Generating the curvature perturbation without an inflaton. *Phys. Lett.*, B 524:5–14, 2002, [hep-ph/0110002].

Maldacena, J. Non-gaussian features of primordial fluctuations in single field inflationary models. *JHEP*, 5:13, 2003, [astro-ph/0210603].

Matarrese, S., Lucchin, F. and Bonometto, S. A. A path-integral approach to large-scale matter distribution originated by non-Gaussian fluctuations. *ApJ*, 310:L21–L26, 1986.

Matarrese, S. and Verde, L. The effect of primordial non-Gaussianity on halo bias. *ApJ*, 677:L77, 2008, [arXiv:0801.4826].

McDonald, P. Primordial non-Gaussianity: large-scale structure signature in the perturbative bias model. *Phys. Rev. D*, 78:123519, 2008, [arXiv:0806.1061].

Salopek, D. and Bond, J. Nonlinear evolution of long wavelength metric fluctuations in inflationary models. *Phys. Rev. D*, 42:3936–3962, 1990.

Sefusatti, E., Liguori, M., Yadav, A. P. S., Jackson, M. G. and Pajer, E. Constraining running non-gaussianity. *JCAP*, 12:22, 2009, [arXiv:0906.0232].

Seo, H. and Eisenstein, D. J. Probing Dark Energy with Baryonic Acoustic Oscillations from Future Large Galaxy Redshift Surveys. *ApJ*, 598:720–740, 2003, [astro-ph/0307460].

Shandera, S., Dalal, N. and Huterer, D. A generalized local ansatz and its effect on halo bias. *JCAP*, 1103:017, 2011, [arXiv:1010.3722].

Sheth, R. K. and Tormen, G. Large-scale bias and the peak background split. *MNRAS*, 308:119–126, 1999, [astro-ph/9901122].

Slosar, A., Hirata, C., Seljak, U., Ho, S. and Padmanabhan, N. Constraints on local primordial non-Gaussianity from large scale structure. *JCAP*, 08:031, 2008, [arXiv:0805.3580].

Spergel, D. N. and Goldberg, D. M. Microwave background bispectrum. I. Basic formalism. *Phys. Rev. D*, 59:103001, 1999, [astro-ph/9811252].

Tegmark, M. Measuring cosmological parameters with galaxy surveys. *Phys. Rev. Lett.*, 79(20):3806–3809, 1997, [astro-ph/9706198].

Verde, L. and Matarrese, S. Detectability of the effect of Inflationary non- Gaussianity on halo bias. *ApJ*, 706:L91–L95, 2009, [arXiv:0909.3224].

Wang, L.-M. and Kamionkowski, M. The cosmic microwave background bispectrum and inflation. *Phys. Rev. D*, 61:063504, 2000, [astro-ph/9907431].

Yadav, A. P., Komatsu, E. and Wandelt, B. D. Fast Estimator of Primordial Non-Gaussianity from Temperature and Polarization Anisotropies in the Cosmic Microwave Background. *ApJ*, 664:680, 2007, [astro-ph/0701921].

Yadav, A. P. S., Komatsu, E., Wandelt, B. D., Liguori, M., Hansen, F. K. and Matarrese, S. Fast Estimator of Primordial Non-Gaussianity from Temperature and Polarization Anisotropies in the Cosmic Microwave Background. II. Partial Sky Coverage and Inhomogeneous Noise. *ApJ*, 678(2):578, 2008, [arXiv:0711.4933].

Yadav, A. P. S. and Wandelt, B. D. Primordial Non-Gaussianity in the Cosmic Microwave Background. *Advances in Astronomy*, 2010:565248, 2010, [arXiv:1006.0275].

Zaldarriaga, M. Non-Gaussianities in models with a varying inflaton decay rate. *Phys. Rev. D*, 69:043508, 2004, [astro-ph/0306006].