Exploring the early universe through the CMB sky

Meerburg, P.D.

P. DANIEL MEERBURG

EXPLORING THE EARLY UNIVERSE THROUGH THE CMB SKY

P. DANIEL MEERBURG

EXPLORING THE EARLY UNIVERSE THROUGH THE COSMIC MICROWAVE BACKGROUND SKY

Exploring the Early Universe Through the CMB Sky

Cover: Esther Wertwijn.

This thesis was typeset in LaTeX by the author and printed in The Netherlands by IPSKAMP drukkers. Online version: http://dare.uva.nl/dissertations
Email author: daanmeerburg@gmail.com

This research has been supported by grants of De Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), NWO-toptalent grant 021.001.040. and the van Gogh Collaboration Grant VGP 63-254, the Nederlandse Onderzoekschool Voor Astronomie (NOVA), and the Leids Kerkhoven-Bosscha Fonds (LKBF). Printing of this thesis was partly funded by the Kattendijke Drucker Stichting.

Exploring the Early Universe Through the CMB Sky

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties
ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op dinsdag 27 september 2011, te 14:00 uur

door

Pieter Daniël Meerburg

geboren te Amsterdam

Promotiecommissie

Promotor:	Prof. dr. R. A. M. J. Wijers
Co-Promotor:	Dr. J. P. van der Schaar
Overige leden:	Prof. dr. M. B. M. van der Klis
	Prof. dr. J. de Boer
	Prof. dr. B. D. Wandelt
	Prof. dr. A. Achúcarro
	Prof. dr. K. Skenderis
	Dr. P. S. Corasaniti

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

The research reported in this thesis was carried out at the Astronomical Institute *"Anton Pannekoek"*, and the Institute for Theoretical Physics Amsterdam *"ITFA"*, at the University of Amsterdam in the Netherlands.

"To Whom it May Concern"

Contents

Chapter 1

Introduction

It is far better to grasp the Universe
as it really is than to persist in delusion,
however satisfying and reassuring.
Carl Sagan

1.1 Plain Vanilla Cosmology: the 6-Parameter ΛCDM Model

1.1.1 A little bit of History

The Universe is one big laboratory. In fact, as far as we know, it has no boundaries. Surprisingly then it seems that most of the (observable) Universe's large-scale history and future can be described using a *total of 6 parameters*. This remarkable conclusion is drawn from high precision measurements in the last two decades, while theoretically it had been predicted long before. Rapid theoretical advances in our understanding of the Universe were made after Einstein's proposal of General Relativity (Einstein 1915, 1916). The formulated equations allowed him to relate the geometry of spacetime to the energy distribution in the Universe. To his surprise, any solution to his equations would imply the Universe was either expanding or collapsing (Einstein 1917). Assuming the Universe was stationary, he resolved his findings by adding a (cosmological) constant on the geometry side of his equations, in order to balance out the dynamics coming from the evolving energy densities. In the following years it was shown by de Sitter (de Sitter 1918a,b) that the addition of such a constant resulted in an unstable (static) solution. Eventually the Universe would resume its dynamics, and the addition of the cosmological constant did not behave according to Einstein's prejudice. It has been said that Einstein later referred to the addition of the cosmological constant as his greatest blunder, although it might have been a free interpretation of George Gamow, the first to mention this in his memoirs (Gamow 1970). Today the cosmological constant has become one of the most important parameters in modern cosmology (albeit on the energy-momentum side of the equations) and most of us would be happy to make such a 'mistake'.

The first[1] observational evidence of a dynamic Universe came from observations of redshifts

[1]Remarkably, in a recent historical note by Block (2011) it was claimed that in fact Lemaître was the first

of known galaxies. Hubble (1929) was the first to show there exists a relation between the velocity of galaxies and their derived distance with respect to the observer. We would like to note that based on current standards and known errors his findings would have been refuted. Maybe it was intuition, but it initiated the idea (Lemaître 1927, 1931) of a 'big bang' scenario, which can be considered the pillar of modern cosmology. Although the idea was controversial at first, with its greatest critic Fred Hoyle sarcastically naming it the Big Bang in 1949, the idea itself led to several other theoretical predictions. First of all, the abundance of light elements in the observable Universe could be explained if there existed a hot and dense phase in the Universe's past. Alpher et al. (1948) showed that a period of nucleosynthesis in the early Universe could very well explain the observed ratio of Hydrogen to Helium. Secondly, in two follow-up papers (Gamow 1948; Alpher & Herman 1948) they also made another important prediction. It was argued that if the Universe went through a hot dense phase and expanded, at some point in its evolution, photons became free streaming. This should have occurred right around the time ionized Hydrogen recombined with free electrons to become neutral Hydrogen. This greatly reduced the (Thomson) cross section and practically increased the photon mean free path to infinity ($r_{mfp} \gg cH^{-1} \equiv r_H$). This free streaming radiation should be observable today as a uniform afterglow with a redshifted temperature of several Kelvin (predicted to be about 5K by Alpher & Herman (1948)).

This afterglow was first observed by Penzias & Wilson (1965) and today is known as the cosmic microwave background (CMB). Till today, it represents the most convincing evidence for a hot big bang scenario. In 1990 the Cosmic Background Explorer (COBE) satellite confirmed that the CMB was isotropic and homogeneous across the entire sky (Smoot et al. 1992) and that its mean black body temperature of 2.7 degrees Kelvin (Mather et al. 1990a,b), indicating the last scattering occurred around 380,000 years after the Big Bang at a redshift of $z \sim 1100$. The Wilkinson Microwave Anisotropy Probe (WMAP; Peiris et al. (2003); Spergel et al. (2003a, 2007); Komatsu et al. (2009, 2011)) has measured this background with very high accuracy (with a sub degree resolution, fig. 1.4) and in the near future Planck should improve current measurements with even higher resolution, probing arc minute scales. Recent observations point towards a Universe that once was very hot and dense and has since evolved into a low density, cold Universe[2]. Using supernovae type Ia as standard candles we have learned that the Universe is still expanding today and that this is happening at an accelerating rate (Perlmutter et al. 1998; Riess et al. 1998, 2007). The paradigm of modern cosmology is built upon the idea of an expanding Universe, which has evolved from a singularity commonly referred to as the Big Bang. The remainder of this chapter will introduce the theoretical framework behind this paradigm. In particular, we will introduce the model of the Universe that can be described by 6 parameters and today is dominated by two unknown components; a cosmological constant Λ and cold dark matter (CDM), the ΛCDM model of the Universe.

who proposed a relation between distance and the receding motion of distant galaxies (Hubble law). This was done in a paper in 1927, translated in 1931 (Lemaître 1927, 1931), 2 years after Hubble's paper. Referring to letters between Hubble and others as well as the original French paper, Block claims that essential parts of the paper where not included in the English translation. It is not clear if this was done on purpose and who decided to leave out these parts.

[2]For a complete list of CMB experiments (past, current and future) see for example `http://lambda.gsfc.nasa.gov/links/experimental_sites.cfm`

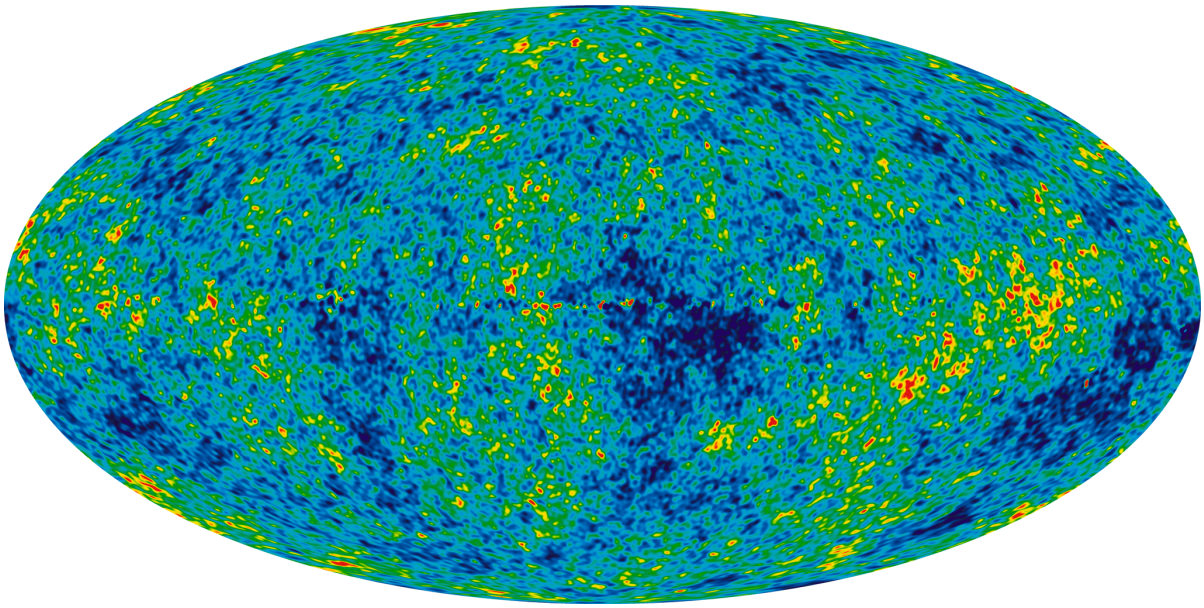

Figure 1.1: WMAP full sky image of the CMB. This Internal Linear Combination Map is a weighted linear combination of the five WMAP frequency maps (23, 33, 41, 61 and 94 GHz). The weights are computed using criteria which minimize the Galactic foreground contribution to the sky signal. The resultant map provides a low-contamination image of the CMB anisotropy. Courtesy of NASA.

1.1.2 The Metric of the Universe

Under the assumption that the Universe is isotropic and homogeneous on large scales, one can derive a simple form of the metric[3] describing an expanding Universe known as the FRLW metric (after Friedmann (1922), Robertson (1929), Lemaître (1927, 1931) & Walker (1933)), which is maximally symmetric[4]:

$$ds^2 = -dt^2 + a^2(t)\left[dr^2 + f^2(r)\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right]. \tag{1.1}$$

$a(t)$ is the overall scaling of the Universe as a function of time, while $f(r)$ is a function that has been derived to describe isotropy and homogeneity. After a coordinate transformation the metric can be rewritten in more familiar form

$$ds^2 = -dt^2 + a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right]. \tag{1.2}$$

[3]Throughout this thesis we set $c = 1$, and in specific cases we also set the reduced Planck mass $M_p = 1$.

[4]For a particularly elegant derivation see for example the TASI lectures on cosmology by Carroll & Trodden (2004).

The solution of $f(r)$ can now be connected to three different values of k, usually referred to as the curvature:

$$\begin{aligned} k = \ +1 \ &\Leftrightarrow f(r) = \sin(r) \\ k = \ 0 \ &\Leftrightarrow f(r) = r \\ k = \ -1 \ &\Leftrightarrow f(r) = \sinh(r). \end{aligned} \tag{1.3}$$

These values correspond to positively curved ($k = 1$), flat ($k = 0$) and negatively curved ($k = -1$) spatial hypersurfaces. Together with Einstein's equations

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu}, \tag{1.4}$$

we can derive the (first) Friedmann equation (Friedmann 1922, 1924):

$$H^2(t) = \frac{8\pi G}{3} \sum_i \rho_i(t) - \frac{k}{a^2}. \tag{1.5}$$

The sum runs over all components in the Universe (which may include a constant vacuum contribution) and H is the Hubble rate, describing the time evolution of the scale factor in units of the scale factor, i.e., $H = \dot{a}/a$. In other words, the Friedmann equation tells us how the presence of energy (density) in the Universe and the non-zero curvature impacts the scale evolution (time dependence) of the Universe. The Hubble rate is usually denoted as $100h$ (km/s)/Mpc, with h measured to be $0.704^{+0.013}_{-0.014}$ (Komatsu et al. 2011) today. If the Universe has been dominated by matter for most of its lifetime and has zero curvature (both turn out to be fairly good assumptions), one can obtain a quick estimate of the age of the Universe by inverting H, $t_{Universe} \sim 1/H \sim 13.7$ billion years. We can define a critical density as $\rho_{cr} = 3H^2/8\pi G \sim 2h^2 \times 10^{-29}\text{g/cm}^3$, the equivalent of about 6 protons per cubic meter. The critical density is defined to be the density in a flat Universe, with $k = 0$. Generally, all energy densities are defined relative to the critical density, i.e., $\Omega_i = \rho_i(t)/\rho_{cr}$. We can use this to rewrite the Friedmann equation at any given time as

$$1 \ = \ \sum_i \Omega_i. \tag{1.6}$$

Here we included the curvature as an energy density $\rho_k = \frac{3k}{8\pi G a^2}$. It also shows that the total measured energy density Ω_{tot} in the form of matter, radiation and possibly a cosmological constant Λ is related to the curvature k. One can distinguish three discrete geometrically isotropic and homogeneous solutions

$$\Omega_0 > 1 \ \Leftrightarrow \ k = 1, \tag{1.7}$$

$$\Omega_0 = 1 \ \Leftrightarrow \ k = 0, \tag{1.8}$$

$$\Omega_0 < 1 \ \Leftrightarrow \ k = -1. \tag{1.9}$$

relating the total energy density in Universe to its intrinsic curvature.

A second equation governing the evolution history of the Universe is derived from the space-space metric components of the Einstein equations

$$\frac{\ddot{a}}{a} + \frac{1}{2}\left(\frac{\dot{a}}{a}\right)^2 = -4\pi G \sum_i \mathcal{P}_i - \frac{k}{2a^2}, \tag{1.10}$$

with $\mathcal{P}$ the pressure of the constituents making up the Universe. Combined with the first Friedmann equation we obtain the acceleration or second Friedmann equation

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\sum_i (\rho_i + 3\mathcal{P}_i). \tag{1.11}$$

From the conservation of energy in the total Universe $\nabla_\mu T^\mu_\nu = 0$, one obtains

$$\frac{d\rho}{dt} = -3H(\rho + \mathcal{P}), \tag{1.12}$$

which tells us that the expansion of the Universe can cause changes in the energy density. It relates the energy density to the pressure. Such a relation is known as an equation of state (e.o.s.). It turns out that for all relevant components in the Universe this relation is linear $\mathcal{P} = w\rho$. Matter in the Universe (on large scales) is presureless, so $\mathcal{P}_m \sim 0$. Therefore $\rho_m \propto a^{-3}$, i.e., the energy density of matter decreases with the scale factor cubed (the comoving volume of the observable Universe). For radiation we know the equation of state equals $\mathcal{P} = \rho_r/3$, which tells us that $\rho_r \propto a^{-4}$. The number density of photons decreases with the scale volume of the Universe (a^{-3}). Meanwhile they are also redshifted as the Universe expands, adding a factor of $1/a$. For a cosmological constant, the argument can be reversed. Given that its contribution should remain constant (ρ_Λ = constant) as the Universe expands we obtain $\mathcal{P}_\Lambda = -\rho_\Lambda$, a cosmological constant Λ has *negative* equation of state. Were the Friedmann equations to hold, then the energy density needs to be positive. Therefore, a negative equation of state in a FRLW implies a negative pressure $\mathcal{P}$. This assures that the expansion does not require any energy and the total energy increases as long as the volume increases, maintaining a constant energy density.

One can measure the energy density of components that dominate the Universe today and those measurements suggest the Universe is currently dominated by a cosmological constant ($\Omega_\Lambda \sim 0.73$). The energy density of Λ is referred to as dark energy; it does not interact with light, it dominates the total energy density, and its absolute contribution to the total energy density does not change as the Universe expands. Most of the additional energy density in the Universe is not coming from baryonic matter either, but from a component that must have decoupled earlier in the history in the Universe. This can be derived from observations of structure formation as well as the CMB, which show that ordinary matter alone can not account for the observed gravitational potentials (they are too deep, given the Hubble time). Observed velocity dispersion of clusters (Zwicky 1933, 1937) and rotational curves of galaxies (Rubin et al. 1980) were in fact the first indication that some form of unobserved matter must exist. Any candidate matter is at most weakly coupled to particles from the Standard model and should predominantly interact gravitationally. Usually this matter is referred to as dark

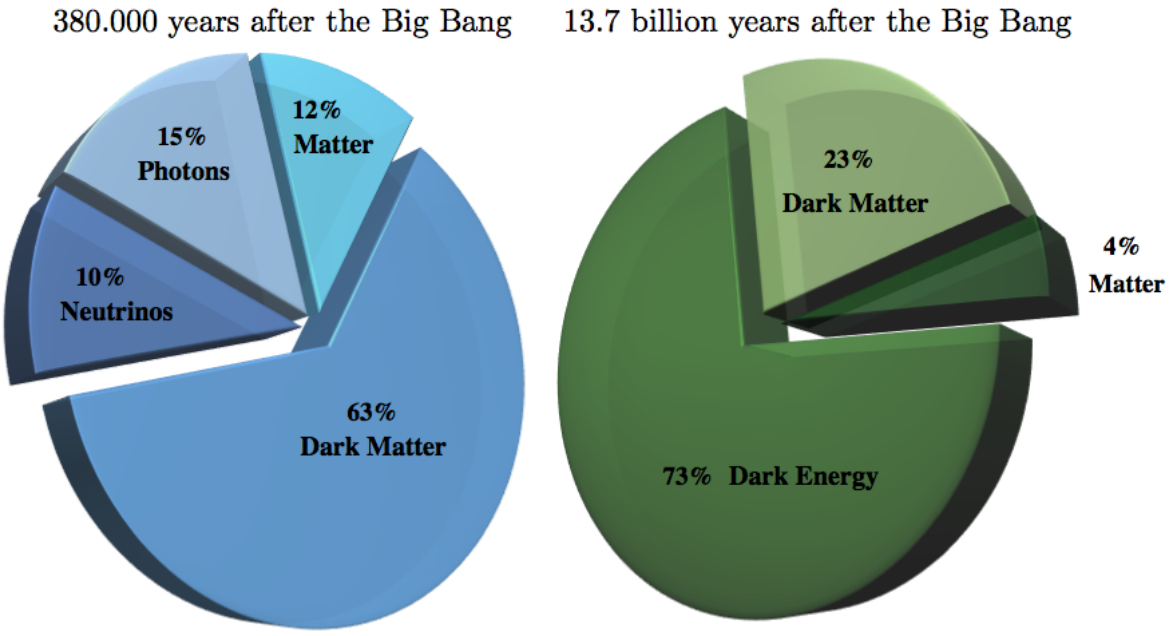

Figure 1.2: The current distributions of the energy density in the Universe (right) compared to those at the last scattering surface (left). Today, radiation has completely redshifted away, while dark energy has emerged as the dominant energy contribution. If the Universe will remain dominated by a cosmological constant in the near future, the Universe should expand at an increasing rate, diluting both matter and dark matter, until dark energy completely dominates the energy density.

matter and current measurements estimate its contribution to the total energy density in the Universe today to be about 23%. Observations also point towards a flat Universe, so the total energy density Ω_{tot} should be close to 1. The remainder of the Universe ($\sim 4.5\%$) is baryonic matter, while radiation has completely redshifted away in the current Universe. fig. 1.2 shows the relative abundance of the constituents of the Universe today compared to those immediately after last scattering.

The first three parameters of the plain vanilla ΛCDM model are Ω_m (matter), Ω_{dm} (dark matter) and Ω_Λ (cosmological constant). In their zero order form, they describe the homogeneous, isotropic dynamic evolution of the Universe. They do not describe local properties, which drive large scale structure evolution of the Universe triggered by gravitational collapse. Eventually we want to understand the origin of structure formation, which is dictated by the amplitude A_s and scale dependence n_s of local variations. These degrees of freedom are the 4th and 5th parameter of plain vanilla ΛCDM. These parameters are thought to be determined by a period of rapid expansion in the early Universe, known as inflation, which will be the topic of the next sections.

1.1.3 Cosmological Conundrums

Although Friedmann and others assumed the Universe was homogeneous and isotropic on large scales (Mpc) they did not give an explanation of why this should be the case. It was simply an assumption that made properties calculable. Both the observations of the CMB and Large Scale Structure (LSS) indicate that this assumption is correct. The Universe appears similar at any location, and in each direction. However, with the discovery of the CMB this striking property led to a problem. The issue concerns the causal horizon at the moment of CMB formation. Although the CMB is isotropic and homogeneous on all scales, knowing the age of the Universe around CMB formation, the CMB was only causally connected over a distance of about 380,000 light years. This roughly corresponds to about 1 degree on the current sky. Therefore, we need to understand how the Universe appears to be in thermal equilibrium on all (observable) scales. This issue is known as the horizon problem (fig. 1.3) and constitutes one of several problems.

The measured flatness ($k \sim 0$) of the Universe also requires an explanation. We can again rewrite the Friedmann equation

$$\Omega_{tot,0} - 1 \quad = \quad \frac{k}{(aH)^2}, \tag{1.13}$$

where on the left hand side we have the total energy density measured today ($\Omega_{tot,0} \sim 1.01 \pm 0.02$). Using the e.o.s., we can rewrite any energy density as $\rho(a) \propto 1/a^{3(1+w)}$ and derive a differential equation of the total energy density as

$$\frac{d\Omega}{da} \quad = \quad (1+3w)\frac{\Omega(\Omega-1)}{a}. \tag{1.14}$$

The solution of this differential equation has three fixed points, $\Omega = 0, 1$, and ∞. The behavior near these fixed points depends on the sign of $1+3w$. For much of the its history, the Universe has been dominated by matter and radiation with $w > -1/3$. Consequently for a very long time we have been close to the repelling solution of the $\Omega = 1$ fixed point. This leads us to conclude that because the Universe is so close to flat today, it must have been even flatter in the past. This is known as the flatness problem. Note that today we are dominated by a cosmological constant, for which $1+3w < 0$, forcing the Universe to become closer to flat over time.

In addition to the flatness and horizon problem, the Standard Model of particle physics predicts the production of magnetic monopoles in the early Universe ('t Hooft 1974). These are unobserved and this is referred to as the monopole problem (Zel'dovich & Khlapov 1978; Preskill 1979). It must be noted that all these problems are not strict inconsistencies in the standard cosmological models. In fact, it is possible to remove all these problems by incredible fine-tuning of the initial conditions, but for most of us this seems rather unsatisfactory. In the early 1980's, Alan Guth proposed a physical mechanism that could solve all these problems at once, and which removes the need for fine-tuning the initial conditions.

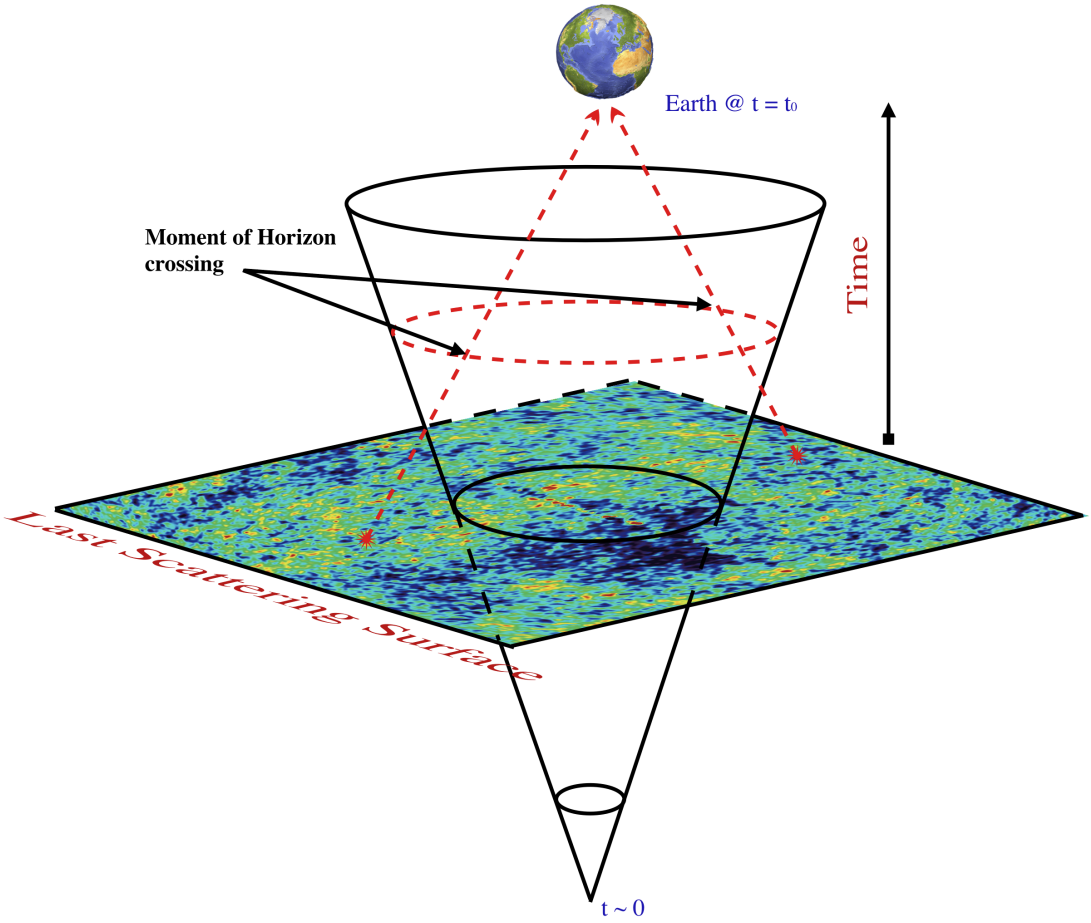

Figure 1.3: Some of the modes we can distinguish in the CMB today only entered our horizon very recently (dashed line). Surprisingly, this causally disconnected region (the CMB) has temperatures that differ from each other by only 1 part in 100000.

1.1.4 Cosmological Inflation

Consider the radial null geodesics in a flat FRWL metric

$$ds^2 = dt^2 - a(t)^2 dr^2 \equiv 0 \rightarrow dr = \pm \frac{dt}{a(t)} \equiv d\tau. \tag{1.15}$$

One can define the comoving horizon τ as the maximum distance a light ray can have travel between $t = 0$ and time t

$$\tau \equiv \int_0^t \frac{dt'}{a(t')} = \int_0^a \frac{da'}{a'} \frac{1}{a'H(a')}. \tag{1.16}$$

Note that this is equivalent to conformal time[5], which is usually denoted with η. The comoving horizon is the logarithmic integral over the comoving Hubble radius[6] $1/aH$. In both a matter (MD) and a radiation dominated (RD) universe, $1/aH$ grows with time, and consequently, the comoving horizon grows with time, i.e., $\tau \propto a$ (RD) and $\tau \propto a^{1/2}$ (MD). The comoving horizon grows monotonically with time in an expanding universe, which implies that comoving scales entering the horizon today have been far outside the horizon around last scattering (the horizon problem). By construction, particles separated by τ can never have been in causal contact (light could not bridge this distance within a time interval since $t = 0$), while particles separated by a distance $> 1/aH$ are not in casual contact now. It is possible that τ is much larger than the comoving Hubble horizon today, but at some earlier time it was not, i.e., particles could have been in causal contact in the past, while today they appear to be outside each other's (causal) horizon. If somehow the comoving Hubble radius was much larger in the past, τ could have received most of its contribution from early times. We know that the comoving Hubble radius grows within a matter or radiation dominated Universe, therefore within those era's most of τ contribution comes from late times when $1/aH$ is largest. Consequently, if τ received most of its contribution from early times, not only does $1/aH$ need to be large, it also requires $1/aH$ to decrease rapidly after, in order to avoid receiving even larger contributions at late times. The comoving Hubble radius should have decreased, transforming previously causally disconnected regions into causally connected regions. A decreasing comoving Hubble radius requires an increase of aH

$$\frac{d}{dt}[aH] \quad = \quad \frac{d}{dt}\left[\frac{da}{dt}\right] = \frac{d^2a}{dt^2} > 0. \tag{1.17}$$

Therefore, to solve the horizon problem the Universe needs to grow with an accelerating rate; in other words the Universe needs to inflate! It is not hard to show that this translates into the condition $\mathcal{P} < -\rho/3$, which is satisfied when $1 + 3w < 0$. As long as $w > -1/3$ the comoving Hubble radius grows with time, and we can not solve the horizon problem. Guth (1981) showed that if the Universe was dominated by something that had an e.o.s. with $w < -1/3$, for example a cosmological constant, you could ignite a period of accelerated expansion. This allows for a causally connected region in the Universe at early times to cover a causally disconnected region (volume) at later times. A (scalar) field configuration with $w < -1/3$ requires small kinetic energy compared to its potential energy. In the following years Linde (1982) (and independently Albrecht & Steinhardt (1982a)) quickly realized that this requires a very flat potential of the inflaton candidate, the scalar particle making up the field driving inflation. As such, the inflaton would slowly roll down its potential, causing the Hubble rate to become

[5] A spacetime is conformally flat if in some coordinates its metric is $g_{\mu\nu} = \Omega^2(x)\eta_{\mu\nu}$, where $\eta_{\mu\nu}$ is the flat Minkowski metric, with signature $-+++$ and $\Omega^2(x) \neq 0$ some function. These spacetimes can be mapped to the flat spacetime by some conformal transformation. Clearly in the metric of eq. (1.15), $\Omega^2 = a^2$. If a field theory is conformally invariant, this transformation reduces the minimally coupled action to that of a field in the flat Minkowski spacetime. In effect, a conformal field in a conformally flat spacetime is totally decoupled from gravity.

[6] As opposed to the Hubble radius $1/H$ which is a fixed quantity at time t. Allowing a particle to travel with the speed of light from one end to the other spanned by the Hubble radius, does not take into account that the Universe grows as time evolves and the particle makes its way through the universe. The comoving distance does. Therefore, this is the correct property to consider when discussing causality in the an expanding universe.

practically constant, while any causal region quickly expands beyond the horizon and creates a super horizon region that appears causally connected at late times. Linde also realized that if such a period of inflation would last long enough it could solve the flatness problem (recall eq. (1.14)). In addition, monopoles could have become so diffuse they remain unobserved.

One can quantify the relative expansion of the Universe during a period of inflation using N

$$N = \int H dt. \tag{1.18}$$

The background temperature of the CMB decreases as $1/a$ and today $T \sim 2.7$ K $\sim 2.4 \times 10^{-4}$ eV, while the energy scale of inflation corresponds to $T \sim 10^{15}$ GeV. Therefore we can derive that $a_e \simeq 10^{-28}$ around the beginning of inflation. Consequently to solve the horizon and flatness problem $N \sim 55$; the Universe needs to grow by a factor of 55 e-folds (which depends on the scale of inflation), where we assumed that after inflation the Universe grows around 10 e-folds. On the other hand, we know that the Universe at some point returned to its usual expansion, in which matter and radiation became the dominant components in the Universe (up to $z \sim 1$). Therefore inflation needs to end and the energy density of the dominant component needs to decay into other degrees of freedom. The inflaton needs to roll down its potential until it reaches the lowest energy state (originally referred to as the 'true' vacuum, as opposed to the excited 'false' vacuum (Guth 1981)) after which the inflaton degrees of freedom are transferred through inflaton decay into matter and radiation. Such a transition is usually referred to as the process of reheating (Albrecht & Steinhardt 1982b). You could consider this transition of inflatons into matter and radiation the start of a 'classic' Big Bang beginning of the Universe's evolution.

In the years following it was realized that inflation could also produce the seeds of structure formation (Mukhanov & Chibisov 1981; Starobinsky 1982; Albrecht & Steinhardt 1982a; Guth & Pi 1982; Hawking 1982a,b; Bardeen et al. 1983). It was argued that in a realistic model of inflation, not all regions would inflate equally, since inflation is driven by local fluctuations in the inflaton field. The fluctuations in the scale factor growth δa during inflation lead to density fluctuations after reheating, as $\delta a/a \propto \delta\rho/\rho$ when a given momentum scale crosses the horizon. Quantum fluctuations in the vacuum are therefore the source of fluctuations in the energy density fields of radiation and matter. Once the inflaton has decayed these fluctuations will be the source of temperature anisotropies in the CMB and density differences in the observed LSS. This is a remarkable prediction of inflation, unforeseen at the time Guth first formulated his theory. Interestingly, these predictions exactly matched those made years earlier by Harrison (1970) and Zel'dovich (1972). They argued an initial scale invariant spectrum of scalar perturbations was required to account for the observed cosmological fluctuations, which is the reason such a spectrum of fluctuations is ofter referred to as a Harrison-Zel'dovich spectrum (also see Peebles & Yu (1970) and Sunyaev (1978)). The associated power spectrum of primordial quantum fluctuations at horizon crossing is relevant for all the late time density distributions, in particular the CMB anisotropy spectrum. The 4th and 5th parameter of ΛCDM define this primordial power spectrum via $P(k) \propto A_s k^{n_s-1}$, where k is the comoving momentum $k = ap$. A non-zero $n_s - 1$ represents the first order deviation from scale invariance. For a constant Hubble rate during inflation, $n_s = 1$ and all scales have the same power A_s. Since inflation has to end, $\dot{H} \propto \epsilon$, where ϵ is a very small number that defines the slope of

the inflaton potential. In the next paragraph we will derive the primordial power spectrum for the simplest model of inflation, which its governed by the slow-roll parameters. This type of inflation is known as single-field slow-roll inflation.

1.1.5 Slow-Roll Inflation

If inflation is driven by a single canonical scalar field ϕ, one can write down the minimally coupled action for the inflaton field

$$S = \frac{1}{2}\int d^4x\sqrt{-g}\left[g_{\mu\nu}\phi_{,\mu}\phi_{,\nu} - 2V(\phi)\right]. \tag{1.19}$$

Here the potential is unspecified, but we will later see that the slow-roll condition requires $V'(\phi) \ll 1$ and $V''(\phi) \ll 1$, where $'$ denote derivatives with respect to the homogeneous part of the field ϕ, $\phi^{(0)}$. Let us first consider the zero order free field. We can derive the equation of motion by varying the action

$$\ddot{\phi}^{(0)} + 3H(t)\dot{\phi}^{(0)} + V' = 0. \tag{1.20}$$

In addition, by deriving the energy-momentum tensor from (1.19) through $T_{\alpha\beta} \propto \delta S/\delta g^{\alpha\beta}$ and realizing $T_0{}^0 = -\rho$ we obtain an expression for the energy density of the scalar field

$$\rho(\phi) \sim \frac{1}{2}\left(\dot{\phi}^{(0)}\right)^2 + V(\phi). \tag{1.21}$$

Inserting this back into the Friedman equation (eq. (1.5)) we obtain

$$H^2 = \frac{8\pi G}{3}\left(\frac{1}{2}\left(\dot{\phi}^{(0)}\right)^2 + V(\phi)\right). \tag{1.22}$$

The requirement that the $\mathcal{P} > \rho$ implies $(\dot{\phi}^{(0)})^2 \ll V(\phi)$ (ϕ should practically be constant during inflation) and therefore

$$H^2 = \frac{8\pi G}{3}V(\phi), \tag{1.23}$$

$$3H\dot{\phi}^{(0)} \simeq -V'(\phi). \tag{1.24}$$

We can use (1.23) to equate the derivative of the potential with respect to the zero order field $\phi^{(0)}$

$$V'(\phi) = \frac{6}{8\pi G}\frac{H\dot{H}}{\dot{\phi}^{(0)}}. \tag{1.25}$$

The condition $(\dot{\phi}^{(0)})^2 \ll V(\phi)$ translates to

$$H^2 \gg \frac{V'^2}{V}. \tag{1.26}$$

The change in the Hubble rate should be small during inflation and we can define a slow-roll parameter ϵ

$$\epsilon = -\dot{H}/H^2, \tag{1.27}$$

which can be rewritten in terms of potential derivatives using eq. (1.23) and eq. (1.25)

$$\epsilon = \frac{1}{16\pi G}\left(\frac{V'}{V}\right)^2 \ll 1. \tag{1.28}$$

Additionally, the slow-roll requirement puts a constraint on the second derivative of the potential $\ddot{\phi}^{(0)} \ll V(\phi) \rightarrow V'' \ll H^2$. This requirement is necessary in order for of the potential to *remain* flat for at least 55 e-folds. The second slow roll parameter η (not to be confused with conformal time) is given by

$$\eta = \frac{1}{8\pi G}\frac{V''}{V} \ll 1. \tag{1.29}$$

The slow-roll parameters govern slow roll inflation; they are small for as long as inflation lasts. Inflation ends as soon as the slow-roll conditions are violated, and kinetic potential of the inflaton is transferred to other degrees of freedom (reheating), populating the Universe with the particles we observe today.

We can rewrite equation (1.18) as

$$N = \int H dt = \int \frac{H}{\dot{\phi}} d\phi = \frac{1}{\sqrt{2}M_p^2}\int_{\phi_i}^{\phi_e} \frac{d\phi}{\sqrt{\epsilon}}, \tag{1.30}$$

where we used $M_p^2 = (8\pi G)^{-1}$, with M_p the reduced planck mass. The integral runs from the beginning of inflation to the end of inflation. It shows that the number of e-folds is directly proportional to the inverse of the square root of slow-roll parameter, emphasizing that the number of e-folds increases as the potential gets flatter or as inflation lasts longer.

1.1.6 The Primordial Power Spectrum

So far we have only considered the zero order component $\phi^{(0)}$ of the inflaton field. The next step is to consider perturbations that allow for small spatial variations in the field

$$\phi(\vec{x}, t) = \phi^{(0)}(t) + \delta\phi(\vec{x}, t). \tag{1.31}$$

Perturbations in the inflaton field will induce perturbations in the metric. One can introduce a gauge invariant[7] variable

$$\zeta = \psi - \frac{H}{\dot{\phi}^{(0)}}\delta\phi, \tag{1.32}$$

[7]Note that only *gauge invariant* quantities have an inherent physical meaning. Gauge dependent quantities have a physical meaning only to the extent that, in a particular gauge, they can be *identified* with a gauge independent quantity either exactly or approximately.

where ψ represents the scalar perturbations in the metric. One can fix a gauge such that the perturbations in the inflaton field, $\delta\phi$, are completely decoupled from those in the metric. This gauge is known as the comoving gauge, and is a good choice when computing primordial quantities as we are free to set the perturbations to the field to zero, since they will not effect the equation of motion of the metric perturbation. ζ can be related to Bardeen's curvature perturbation $\Phi = -\psi$ in the matter dominated era (CMB formation and observation) by $\zeta = -\frac{5}{3}\Phi$. At late times, ζ can be related to an observable $\Delta T/T = \frac{1}{5}\zeta$. This identification allows one to compare the theoretically predicted power spectrum of inflaton fluctuations with the observed late time spectrum of temperature fluctuations.

An elegant approach to deriving the perturbed equations of motion is the so called ADM formalism (after R. Arnowitt, S. Deser, and W. Misner (1964)). The formalism is also known as spacetime splitting. The ADM metric is given by

$$ds^2 = -N^2dt^2 + h_{ij}(dx^i + N^i dt)(dx^j + N^j dt). \tag{1.33}$$

In this representation h_{ij} is the three-dimensional metric on slices of constant time t. The lapse function N and shift vector N^i relate the different coordinates from one slice to another. They satisfy algebraic constraints that can be substituted back into the action to derive the action for the dynamical variable h_{ij}. From here on we will drop the $^{(0)}$ superscript from the homogeneous part of the field ϕ.

Considering only scalar perturbations, the spatial part of the metric in the comoving gauge is given by

$$h_{ij} = a^2 e^{2\zeta}\delta_{ij}. \tag{1.34}$$

If we expand the exponential to lowest order in ζ we obtain $h_{ij} = a^2(1+2\zeta)\delta_{ij}$. An advantage of this gauge is that ζ will remain constant outside of the horizon, so we only need to compute the dynamics up to horizon crossing (when the power spectrum, or higher order spectra, are frozen).

The indices on the shift vector can be raised and lowered using the metric (eq. (1.34)). Rewriting the scalar field action (eq. 1.19) using this metric we obtain

$$\begin{aligned} S &= \frac{1}{2}\int d^4x\sqrt{h}\Big(NR^{(3)} - 2NV(\phi) - N(K^2 - K_{ij}K^{ij}) + \\ &\quad N^{-1}(\frac{\partial\phi}{\partial t} - N^i\phi_{,i})^2 - Nh^{ij}\phi_{,i}\phi_{,j}\Big), \end{aligned} \tag{1.35}$$

with K_{ij} the *extrinsic* curvature

$$K_{ij} = \frac{1}{2}N^{-1}\left(h'_{ij} - (D_iN_j + D_jN_i)\right), \tag{1.36}$$

and $K = K^i_{\ i} = h^{ij}K_{ij}$. The D_i represent the covariant derivative on the three-dimensional subspace with metric h_{ij}. In the comoving gauge $\delta\phi$ is zero, i.e., ϕ does not depend on its spatial position, thus all derivatives w.r.t. spatial component vanish. The expression above is valid in *any* gauge.

We aim to determine the equations of motion for N and N^i and solve these equations in the special case of a comoving gauge. Inserting the solutions back into the action, should give us the action for the scalar perturbation parameter ζ, which in turn should give us the equation of motion for ζ.

The equations of motion can be obtained if we vary the action with respect to either N or N^i. First, varying the action with respect to N gives the Hamiltonian constraint

$$\delta\mathcal{L} = R^{(3)} - 2V(\phi) - (K^2 - K_{ij}K^{ij}) - N^{-2}\left(\frac{\partial\phi}{\partial t}\right)^2 = 0. \tag{1.37}$$

Varying the action with respect to the shift vector and neglecting terms of $O((\delta N_i)^2)$ and higher we derive an additional constraint

$$D_i\left[K^i{}_j - \delta^i{}_j K\right] = 0. \tag{1.38}$$

We can solve these differential equations to sufficient order in N and N_i (Maldacena 2003; Chen et al. 2007), insert these back into the action and obtain a perturbed action in terms of the curvature ζ.

Generally it suffices to compute N only up to order $n-2$ to obtain the n^{th} order action of ζ (see for example Chen et al. (2007)). We are interested in the perturbed action up to *second-order* (as required for the power spectrum, the 2-point correlation function) in ζ. Therefore we have to insert the first order expressions for N and N_i in terms of ζ (not given) into the action of eq. (1.35), with $\delta\phi = 0$. By construction, the three dimensional Ricci scalar $R^{(3)}$, is independent of N and N^i and given by

$$R^{(3)} = a^{-2}e^{-2\zeta}\left(-4\delta^2\zeta - 2(\delta\zeta)^2\right). \tag{1.39}$$

∂ represents the partial derivative in the spatial direction. This term is valid up to all orders in ζ. Furthermore $\sqrt{h} = a^3e^{3\zeta}$, and $h_{ik}h^{jm} = \delta_i{}^j$. Setting $M_p = 1$ we can write $a = e^\rho$, with ρ the familiar energy density of the Universe (here $H = \dot{\rho}$). Consequently we find $\dot{h}_{ij} = (2\dot{\rho} + 2\dot{\zeta})h_{ij}$ and $K^{ij} = h^{ik}h^{jm}K_{km}$. Using these expressions we evaluate $K_{ij}K^{ij}$ and K^2 in terms of the ζ, which leads to the following perturbed action

$$\begin{aligned} S = \frac{1}{2}\int d^4x \Bigg[& e^{\rho+\zeta}(1+\frac{\dot{\zeta}}{\dot{\rho}})(-4\Delta\zeta - 2(\partial\zeta)^2 - 2Ve^{2\rho+2\zeta}) \\ & + e^{3\rho+3\zeta}\frac{1}{1+\frac{\dot{\zeta}}{\dot{\rho}}}(-6(\dot{\rho}+\dot{\zeta})^2 + \left(\frac{\partial\phi}{\partial t}\right)^2\Bigg]. \end{aligned} \tag{1.40}$$

We expand each exponential up to appropriate powers of ζ (being maximally quadratic). Rewriting this action using the background equation of motion (e.o.m.) yields

$$S^{(2)} = \frac{1}{2}\int dtd^3x\frac{\dot{\phi}^2}{\rho'^2}\left(e^{3\rho}\dot{\zeta}^2 - e^{\rho}(\partial\zeta)^2\right) = \int dtd^3x\left(a^3\epsilon\dot{\zeta}^2 - a\epsilon(\partial\zeta)^2\right), \tag{1.41}$$

where the superscript $^{(2)}$ denotes second order in ζ.

In units of conformal time ($dt = ad\eta$) the action reads

$$S^{(2)} = \int d^3x d\eta \epsilon a^2 \left(\dot{\zeta}^2 - (\partial\zeta)^2\right). \tag{1.42}$$

Note that overdots now represent derivatives with respect to conformal time. Varying this action and Fourier transforming to decompose the perturbations into momentum modes gives us the equation of motion for ζ

$$\ddot{\zeta}_{\vec{k}} + \frac{1}{a^2\epsilon}\frac{d}{d\eta}(a^2\epsilon)\dot{\zeta}_{\vec{k}} + k^2\zeta_{\vec{k}} = 0. \tag{1.43}$$

We can slightly simplify this equation if we define the following transformations

$$v_{\vec{k}} \equiv z\zeta_{\vec{k}}, \qquad z \equiv a\sqrt{\epsilon}, \tag{1.44}$$

and the e.o.m. for v becomes

$$\ddot{v}_{\vec{k}} + \left(k^2 - \frac{\ddot{z}}{z}\right) v_{\vec{k}} = 0. \tag{1.45}$$

To leading order $\ddot{z}/z = 2a^2H^2(1 + \mathcal{O}(\epsilon))$ and the e.o.m. further reduces to

$$\ddot{v}_{\vec{k}} + \left(k^2 - 2a^2H^2\right) v_{\vec{k}} = 0. \tag{1.46}$$

This tells us that one can distinguish *three* different regimes. A wave with wavenumber k has physical momentum $p = a(\eta)k$, while the comoving wavelength $\lambda_c \sim k^{-1}$ (give or take a factor of 2π) and the physical wavelength $\lambda_p = a(\eta)\lambda_c$. We can thus write

$$k|\eta| \sim \frac{1}{\lambda_c}\frac{1}{aH} = \frac{H^{-1}}{\lambda_p}. \tag{1.47}$$

This gives us a clear picture of the physical interpretation of $k|\eta|$.

The first regime corresponds to $k \gg aH$ or $k|\eta| \gg 1$, for which eq. (1.46) reduces to a harmonic oscillator equation. The modes that oscillate are the modes that are much shorter than the Hubble horizon. These modes are known as *subhorizon* modes.

The second extreme is when $k \ll aH$ or $k|\eta| \ll 1$. The solution to the differential equation is then proportional to $1/|\eta|$. For these *superhorizon* modes, the physical wavelength λ_p stretches far beyond the Hubble horizon ($1/H$). All modes with comoving wavenumber k are subhorizon at early times, i.e., the modes are all well within the horizon for large $|\eta|$. As time evolves, however, (in de Sitter space) the absolute value of η decreases and consequently each mode will become superhorizon at late times. We can identify this moment of *horizon crossing* as the moment at which the physical wavelength equals the Hubble horizon, i.e., at $k|\eta_k| = 1$ (which is equivalent to $k = aH$ at the end of inflation). This signifies the transition between an oscillating and a comoving mode and corresponds to the third regime.

Together with the appropriate initial conditions we can now deduce the solution to the e.o.m.

$$v_{\vec{k}}(\eta) = \frac{e^{-ik\eta}}{\sqrt{4k}}\left(1 - \frac{i}{k\eta}\right), \tag{1.48}$$

which immediately relates to a solution for the mode function ζ

$$\zeta_{\vec{k}}(\eta) \quad = \quad \zeta_k(\eta) = \frac{iHe^{-ik\eta}}{\sqrt{4\epsilon k^3}}\left(1+ik\eta\right). \tag{1.49}$$

This is the classical solution, of the classical equation of motion. Before we can determine the primordial power spectrum, we need to quantize the theory by writing the operator ζ in terms of creation and annihilation operators

$$\hat{\zeta}_{\vec{k}} \quad = \quad \zeta_{\vec{k}}(\eta)\hat{a}^{\dagger}_{\vec{k}} + \zeta^{*}_{\vec{k}}(\eta)\hat{a}_{-\vec{k}}, \tag{1.50}$$

with the usual commutation relations between the creation and annihilation operators $\hat{a}^{\dagger}_{\vec{k}}$ and $\hat{a}_{-\vec{k}}$. The power spectrum is given by the 2-point correlation function

$$\langle\zeta_{\vec{k_1}}\zeta_{\vec{k_2}}\rangle = \delta^3(\vec{k_1}+\vec{k_2})|\zeta_k^{classic}(\eta)|^2 \equiv (2\pi)^2 P(k)\delta^3(\vec{k_1}+\vec{k_2}) \tag{1.51}$$

Therefore the power spectrum at horizon crossing is given by

$$P_{\zeta}(k) = \frac{1}{(2\pi)^2}|\zeta_k(\eta)|^2 = \left.\frac{1}{8\pi^2 k^3 M_P^2}\frac{H^2}{\epsilon}\right|_{k=aH}. \tag{1.52}$$

In app. A we explicitly show that ζ is constant outside the horizon.

The scale dependence of the primordial power spectrum is defined as

$$n_s - 1 \quad = \quad \frac{d\log k^3 P_{\zeta}(k)}{d\log k}. \tag{1.53}$$

For single field slow-roll we find that at horizon crossing $n_s - 1 = 6\epsilon + 2\eta$. Therefore, a slow-roll potential will result in an almost scale invariant spectrum. All scale dependence is due to a small deviation from pure de Sitter (for which H is a constant). The primordial power spectrum is only a function of scale (k^{n_s-1}) and of has an overall amplitude (A_s), hence the fact that these are the 4^{th} and the 5^{th} parameters in the plain vanilla ΛCDM model. In order to relate observation with theory (more on that in the next paragraph), we have to identify a pivot scale k_*

$$P_{\zeta}(k) \quad = \quad \frac{1}{8\pi^2 M_P^2 k^3}\frac{H_*^2}{\epsilon_*}\left(\frac{k_*}{aH_*}\right)^{n_s-1}. \tag{1.54}$$

The choice of k_* determines the normalization of the power spectrum, and thereby the value of the constraint amplitude A_s, the WMAP collaboration uses $k_* = 0.002\text{Mpc}^{-1}$.

Similarly, one can derive a power spectrum for the tensor perturbations h_{ij}. The resulting spectrum is given by

$$P_h(k) \quad = \quad \frac{2H_*^2}{\pi^2 M_P^2 k^3}\left(\frac{k_*}{aH_*}\right)^{n_t}, \tag{1.55}$$

where $n_t = -2\epsilon$. For single field slow-roll we expect the tensor spectrum to be suppressed as $P_{\zeta}/P_h \sim 16\epsilon$. Measuring this ratio, known as r, is sometimes called the holy grail of cosmology.

The reason is that a measurement of r puts a constraint on the (energy) scale of inflation. The constraint on the potential as a function of ϵ becomes

$$V \simeq 10^{67}\epsilon(\mathrm{GeV})^4. \tag{1.56}$$

In the next paragraph we will discuss how primordial conditions seed the temperature anisotropy spectrum of the CMB.

1.1.7 Connecting Theory with Observation

The final step in this section is to relate primordial fluctuations to those at late times. So far we have explained how the components in the Universe ((dark) matter, radiation and a cosmological constant) govern the evolution of the Universe. This evolution does not depend on local variations and can simply be derived from the assumption that Universe is homogeneous and isotropic. To understand the level of inhomogeneity and anisotropy in the Universe requires us to understand the evolution of small perturbations in the initial conditions. Inflation has been proposed as a source of small anisotropies; it provides the conditions to produce an almost scale invariant spectrum of inflaton fluctuations. These inflaton fluctuations $\delta\phi$ are the source of fluctuations in the metric ψ, which couple to the content of the Universe $\delta\rho_m$, and eventually the temperature fluctuations ΔT

$$\delta\phi \rightarrow \psi \rightarrow \delta\rho \rightarrow \Delta T. \tag{1.57}$$

The equations controlling the evolution of these perturbations are the Einstein equations, which relate perturbations in the metric to local variations in the energy density. A set of Boltzmann equations tell us how various components interact. For example, in the early Universe the photons are tightly coupled to electrons through Compton scattering, while protons and electrons are coupled through Coulomb scattering. Neutrinos and dark matter in their turn are only coupled to the other species through the metric (the Einstein equations). The Boltzmann and Einstein equations together form a set of coupled differential equations to be solved numerically[8]. Deriving this set of differential equations and solving them in terms of the late time perturbation spectra, is beyond the scope of this thesis. Here we will simply show the results. Fortunately, most of the physics can be contained in a single (set of) function(s). This function is known as transfer function and relates the primordial perturbation spectrum to the late time observed spectrum. For example, for radiation one can derive the photon transfer function, while for matter one can derive the matter transfer function containing much of the physics relevant between the observer today and the moment of horizon crossing (CMB formation). Before we write down the relation between the primordial power spectrum and the late time spectra, let us first introduce the observed power spectrum of temperature fluctuations, the CMB power spectrum.

[8]There is much excellent literature, explaining in detail the derivation of these equations. Liddle & Lyth (2000); Dodelson (2003); Mukhanov (2005) and Durrer (2008) all have written excellent textbooks on modern cosmology in the last decade. In certain limits, there exist analytical solutions to the set of coupled differential equations. See, e.g., Hu & Sugiyama (1994).

We are interested in the n-point correlation statistics of temperature fluctuations in the CMB

$$\langle \Delta T(\hat{n}_1)\Delta T(\hat{n}_2)\Delta T(\hat{n}_3)....\Delta T(\hat{n}_n)\rangle, \tag{1.58}$$

a statistic describing the clustering patterns of temperature fluctuations in the CMB sky. The brackets "$\langle$" and "$\rangle$" denote the ensemble average over many universes and the $\hat{n}$ are unit vectors in the sky. The real space correlation function is not particularly useful, because data points in the correlation function at different angular scales are generally correlated. For example, the 2-point correlation function at 1 degree is correlated with that at 2 degrees. A more practical way to handle statistical ensembles, is to decompose the fluctuations into a sum over functions that form an orthonormal basis on the sphere. The transformation is known as a spherical harmonic transformation, and the space is referred to as multipole space. The spherical harmonic functions $Y_{lm}(\hat{n})$ form a set of orthonormal functions

$$\int d\Omega Y_{lm}(\hat{n})Y^*_{l'm'}(\hat{n}) \quad = \quad \delta_{ll'}\delta_{mm'}. \tag{1.59}$$

To work in multipole space we need to harmonically transform the temperature fluctuation $\Delta T(\hat{n})$

$$\Delta T(\hat{n}) \quad = \quad \sum_l \sum_{m=-l}^{m=l} a_{lm}Y_{lm}(\hat{n}). \tag{1.60}$$

The a_{lm} are complex coefficients for all $m \neq 0$. However, the reality condition of the $T(\hat{n})$ tells us that $a^*_{lm} = (-1)^{-m}a_{l-m}$, and therefore the number of independent modes is reduced to $2l+1$. Computing the n-point correlation function now becomes equivalent to computing the angular n-point harmonic spectrum

$$\langle a_{l_1m_1}a_{l_2m_2}...a_{l_nm_n}\rangle. \tag{1.61}$$

Since we can only observe one Universe, we are only able to observe one single realization and we will not be able to take the ensemble average. The resulting measure is very noisy and we should find a way to average this measure to reduce the noise. Fortunately, isotropy of the CMB sky translates into statistical isotropy or rotational invariance of the a_{lm}. This allows us to average over the azimuthal direction m, with an appropriate weight (Hu 2001). The advantage of spherical harmonic space now becomes apparent, as the (Wigner) rotation operator D working on the spherical harmonic functions Y_{lm} is a linear operator in this representation, only transforming the azimuthal direction m

$$DY_{lm}(\hat{n}) = \sum_{m'} D^{(l)}_{mm'}Y_{lm'}(\hat{n}). \tag{1.62}$$

After applying the requirement of rotational invariance

$$\langle a_{l_1m_1}a^*_{l'_2m'_2}\rangle \quad = \quad \sum_{m_1m'_2} \langle a_{l_1m'_1}a^*_{l_2m'_2}\rangle D^{(l_1)}_{m'_1m_1}D^{(l_2)*}_{m'_2m_2}, \tag{1.63}$$

we derive a rotational invariant representation of the angular power spectrum, the C_l, of which the optimal estimator is given by

$$C_l = \frac{1}{2l+1} \sum_{m=-l}^{l} a_{lm} a_{lm}^*. \tag{1.64}$$

This is known as the CMB power spectrum. The power spectrum highlights characteristic structures on the sky at a given angular scale. Note that non-uniform noise renders the a_{lm} inhomogeneous and anisotropic. For example, the scanning pattern of the CMB with the WMAP satellite is not uniform, causing the observed map to have non-uniform noise. Such noise needs to be removed before one can extract the true C_l from the data.

The observed power spectrum C_l is matched with the theoretical power spectrum. The variance of the C_l, or the covariance matrix $\langle C_l C_{l'} \rangle - \langle C_l \rangle \langle C_{l'} \rangle$ is used to search for the best fit of the data on theoretically computed spectra. The covariance matrix tells us how two different variables in the model change together. The best fit is determined as a probability through the likelihood function: Given the data and given a cosmological model, what is the most likely value of the parameters given a prior estimate of the parameters?

$$P(\Theta_i|D) \quad = \quad \frac{P(D|M(\Theta_i))P(\Theta_i)}{P(M)}. \tag{1.65}$$

Here D represent the data, Θ_i are a set of parameters describing a model M. $P(\Theta_i)$ describes the parameter priors , while $P(D|M(\Theta_i)$, the probability of the measured data given the model M with parameters Θ_i, is known as the Likelihood function $\mathcal{L}$. For the CMB this translated to (Verde et al. 2003)

$$P(\Theta_i|\hat{\mathcal{C}}_l) \quad = \quad \mathcal{L}(\hat{\mathcal{C}}_l|\mathcal{C}_l^{\rm th}(\Theta_i))P(\Theta_i), \tag{1.66}$$

where $\hat{\mathcal{C}}_l$ is the best estimator of the true[9] $\mathcal{C}_l$, while $\mathcal{C}_l^{\rm th}$ is the model predicted value of the power spectrum, given the parameters Θ_i. The likelihood function $\mathcal{L}$ will depend on the characteristics of the experiment, the skycut (e.g. due to the presence of foregrounds) and on the statistical properties of the data (e.g. Gaussianity and isotropy). For the WMAP data the likelihood can be computed using a code provided by the WMAP collaboration, of which the details are beyond the scope of this thesis.

We are interested in assessing the posterior probability of several (model) parameter values given the data, $P(\Theta_i|D)$. If we know little about the model parameters and their probability, we would be required to compute the probability, and thus the likelihood, of *every point* in parameter space. Even if we were fortunate enough to have priors on all parameters, it would still take a long time to compute the multi-dimensional likelihood, if we have a large number of parameters in our model. To address this problem in Cosmology (e.g., at least 6 parameters for ΛCDM), one usually runs a so-called Markov Chain Monte Carlo (MCMC) to scan the multi-dimensional likelihood. This process stochastically determines the posterior distribution of the parameters given the data. It does so by randomly sampling the likelihood values in parameter

[9]Here $\mathcal{C}_l = l(l+1)C_l/2\pi$.

space rather than on a grid, as follows: Using prior probabilities on a set of parameters, we pick a random parameter vector Θ_i^j (where the index i labels the parameters within the vector and other index, j, labels successive parameter vectors ('samples') in the series). For $j = 1$, we simply pick a sample and compute its likelihood, $\mathcal{L}_1$. Now we pick a next candidate sample by choosing another random parameter vector from the prior probability distribution and evaluating its likelihood, $\mathcal{L}_2$. We now calculate the likelihood ratio $\mathcal{L}_2/\mathcal{L}_1$ and determine based on this whether to keep the new candidate sample for the posterior distribution. The selection could be quite simple (if $\mathcal{L}_2 > \mathcal{L}_2$ keep it, otherwise don't), but generally it is also probabilistic: even if the new likelihood $\mathcal{L}_2$ is smaller than $\mathcal{L}_1$ there is small chance to accept this new vector. This generally pushes the set of samples towards higher likelihood, but without the risk of accepting no further samples if we hit a very high $\mathcal{L}$ early on, thereby avoiding recovering only part of the posterior distribution. If we accept the sample, we have found Θ_i^2. To address a probability to the vector sample Θ_i^2 we count (multiplicity m) how often subsequent proposal samples are rejected based on the rejection criterion. Eventually, we form a chain of samples, each with multiplicity m, reflecting the probability of that particular set of parameter values, as apposed to all other samples. Because this method is probabilistic it allows the chain to find the posterior distribution much more efficient than a grid search. To establish whether we have used a long enough chain to have converged to a static posterior distribution of a given parameter, we can compare parallel chains to each other for statistical agreement. One method of doing so is the so-called Gelman & Rubin (1992) criterion, which will be discussed when we use the MCMC method in ch. 6.

To run an MCMC one has to compare theoretical power spectra, with the data. The theoretical spectra can be computed using a numerical code (e.g. CAMB (Lewis et al. 2000) or CMBFAST (Seljak & Zaldarriaga 1996)) that computes the transfer functions for a given cosmological model, with set parameter values. The last and 6^{th} parameter in plain vanilla ΛCDM is a parameter that is of importance to the transfer function: the reionization optical depth τ_{rei} has an effect on the observed free streaming photons between last scattering surface and the observer. After recombination, electrons have recombined with protons to form neutral hydrogen, the Universe becomes transparent and photons can free stream over a distance of at least a Hubble radius. However, before the average baryon density has diffused sufficiently stellar objects are formed, due to the gravitational collapse of over-densities, . These first stars (population III) reionize the surrounding hydrogen. Current estimates are that the first stars started reionizing the inter galactic medium (IGM) around a redshift of 10. The ionization was presumably completed around $z = 6$ (Komatsu et al. 2011). Note that the IGM, despite being completely ionized, is still optically thin since the number density of free electrons is low. The effect of a reionized IGM is to cause Thomson scattering of photons on electrons. This causes small scale primordial anisotropies to be erased, since it affects only those scales that are subhorizon when reionization starts. It turns out that $\tau_{rei} \sim 0.1$ so about 10% of the total number of CMB photons observed has been affected.

The theoretical C_l can be computed from

$$C_l = \frac{18}{25\pi} \int_0^\infty k^2 dk P_\zeta(k) \Delta_l^2(k). \tag{1.67}$$

Here $\Delta_l(k)$ is the photon transfer function. For large scales (superhorizon after CMB formation,

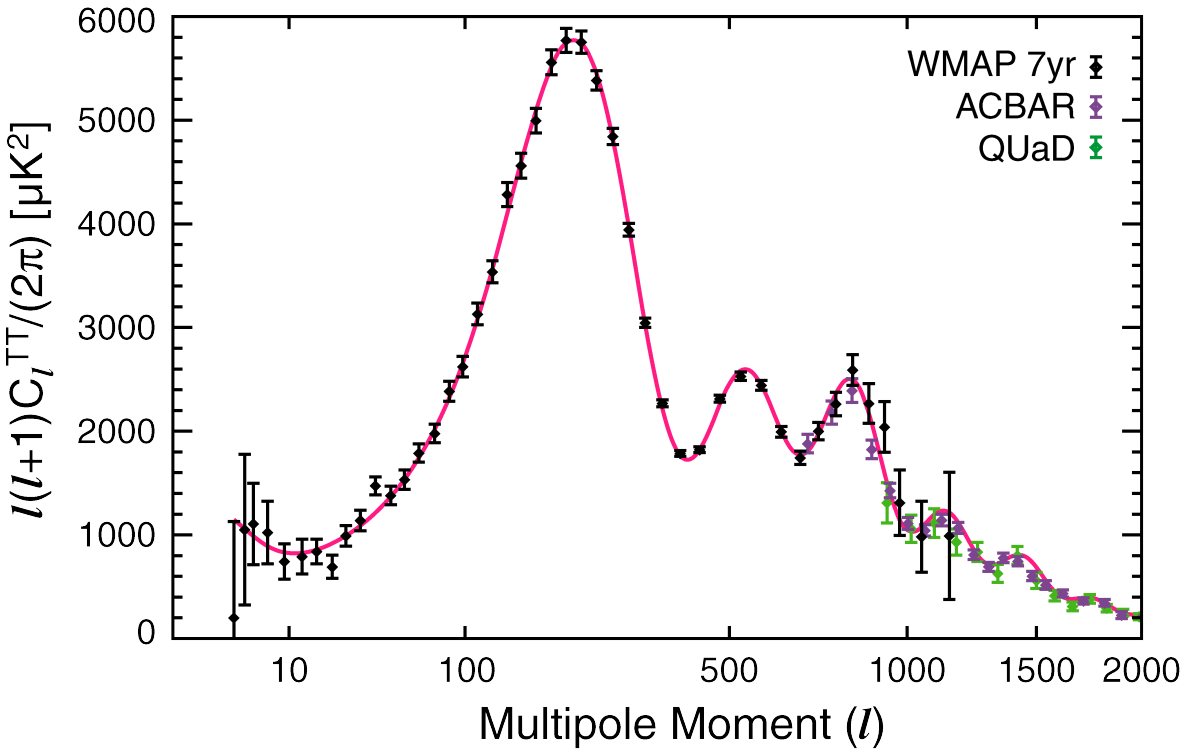

Figure 1.4: The binned WMAP 7-year temperature power spectrum (Larson et al. 2010), along with the temperature power spectra from the ACBAR (Reichardt et al. 2009) and QUaD (Brown et al. 2009) experiments. The ACBAR and QUaD are only shown at $l \geq 690$, where the errors in the WMAP power spectrum are dominated by (systematic) noise. The solid line shows the best-fitting 6-parameter flat ΛCDM model to the WMAP data alone.

i.e., $l \ll 200 \sim 1^0$) the transfer function is not affected by the Boltzmann equations, since these scale modes are not in causal contact, and can certainly not be affected by scattering processes. These scales are dictated by geometry only and the resulting transfer function $\Delta_l(k) \simeq j_l(k\Delta\eta_*)$. Here j_l is a spherical Bessel function and $\Delta\eta_*$ is the look back radius to the last scattering surface, i.e., $\Delta\eta_* = |\eta_* - \eta_0|$ with η_* the conformal time at the last scattering surface. For these scales the integral above can be computed analytically. One finds $C_l \propto 1/l(l+1)$, which is the reason that one usually plots $l(l+1)C_l$ as this approximately renders the low l spectrum constant.

This concludes the discussion of the plain vanilla ΛCDM model. In fig. 1.4 the best fit power spectrum to the WMAP7 data is shown (Larson et al. 2010) in addition to small scale experimental data from ACBAR (Reichardt et al. 2009) and QUad (Brown et al. 2009). Table 1.1 shows a result of the 6 parameter values Ω_b, Ω_c, Ω_Λ, n_s, A_s (or $\Delta^2_{\mathcal{R}}$) and τ_{rei}. Constraints from WMAP data can be combined with an independent measurement of the Hubble constant H_0 through Cepheid variables (Riess et al. 2009). Using the Sloan digital sky survey (SDSS) one can measure the baryonic acoustic oscillations in the matter power spectrum. This measurement of the apparent size of the BAO leads to a measurement of the Hubble constant and

the angular diameter distance (Reid et al. 2009) to the last scattering surface. The addition of these data sets to WMAP data leads to improved constraints (tighter bounds) on the parameters, as shown in the 3^{th} and 4^{th} column of table 1.1. Besides WMAP and Cobe, ground and balloon based (BOOMerANG; (Mauskopf 2000; Melchiorri 2000), MAXIMA; (Hanany 2000)) experiments, have measured CMB small scale anisotropy, confirming the Universe has evolved from a hot dense plasma (baryon-photon fluid) with nearly scale invariant initial conditions and currently dominated by a cosmological constant component. The constraints from WMAP and the small angular scale (ground based) CMB experiments ACT (Dunkley et al. 2010) and QUAD (Brown et al. 2009) are shown in the last column of table 1.1.

Parameter	WMAP	WMAP+H_0	WMAP+BAO	WMAP + CMB
Ω_b	0.0449 ± 0.0028	0.0462 ± 0.0018	0.0438 ± 0.0022	$0.0443^{+0.0025}_{-0.0026}$
Ω_c	0.222 ± 0.026	0.233 ± 0.016	0.210 ± 0.020	0.218 ± 0.024
Ω_Λ	0.734 ± 0.029	0.720 ± 0.018	0.746 ± 0.022	$0.738^{+0.027}_{-0.026}$
n_s	0.963 ± 0.014	$0.96^{+0.013}_{-0.012}$	0.967 ± 0.013	$0.962^{+0.014}_{-0.013}$
$\Delta^2_{\mathcal{R}}$	$(2.43 \pm 0.11) \times 10^{-9}$	$(2.466^{+0.092}_{-0.093}) \times 10^{-9}$	$(2.39 \pm 0.10) \times 10^{-9}$	$(2.43 \pm 0.11) \times 10^{-9}$
τ	0.088 ± 0.015	0.086 ± 0.014	0.090 ± 0.015	$0.088^{+0.015}_{-0.014}$

Table 1.1: The best fit values of the 6-parameter ΛCDM model, derived from WMAP, WMAP+H_0, WMAP+BAO and WMAP + (ACT & QUAD).

Before we continue, let us emphasize the following. Even though this model only counts 6 parameters, there is a lot more physics involved than represented by only those 6 parameters. For example, the parameters describing ΛCDM do not include the abundance of light elements or the number of neutrino species[10]. Nor does it explain why neutrinos have such a small mass. These parameters are derived from other observations. In addition, the evolution of the Universe is derived on the assumption that general relativity is the correct theory of gravity (on large scales). These are all premises within (a 6 parameter) ΛCDM and the predictions made in ΛCDM on the observed distribution of temperature fluctuations in the CMB does depend on those premises and several observed input parameters. GR has thus far withstand several assays, on all scales (see Reyes et al. (2010) and Will (2001) for a review), and the number of neutrino species can independently be derived from particle physics experiments. Preferably, in Bayesian terminology we should consider all models (including those with alternative theories of gravity for example), marginalize over those models to see what model is the most likely (see, e.g., Parkinson & Liddle (2010)). In the rest of this thesis we will assume that the premises of ΛCDM are valid and that several required input parameters are sufficiently accurate. We will use the 6-parameter ΛCDM as a set framework to which we can add additional degrees of freedom (such as non-Gaussianities) that extend the theory.

1.2 Primordial Conditions

The 6-parameter ΛCDM is an incredibly elegant picture of the Universe. It describes how the Universe evolves and allows us to derive the initial conditions through measurements of the

[10]Some of these premises can be constrained using the CMB, but in a 6-parameter ΛCDM they are given.

CMB anisotropy. However, as explained the initial conditions are determined by a limited set of parameters. For the scalar degrees of freedom, only two parameters are relevant (n_s and $A_s \equiv \Delta^2_R$). If we allow for non-vanishing (observationally and theoretically) tensor degrees of freedom, we have two additional parameters. From an observational point of view this makes things very transparent (albeit challenging). There is a large number of theoretical models that can produce the observed level of scale dependence in the scalar power spectrum. Consequently, if all there is to measure from the CMB is the power and weak scale dependence of the 2-point correlation function, it becomes rather difficult to discriminate between different theoretical proposals, given that these proposals do not make very different predictions. In particular, it could be of tremendous value if we could extract additional information from the CMB that could elucidate the exact distribution of initial fluctuations, further probing the physics of inflation. For example, the scale dependence of the initial power spectrum is a first order expansion of the primordial power spectrum. We could investigate second order contributions, that measure the change in scale dependence as a function of scale, known as the running r_s. If inflation is driven by a single field slowly rolling down its potential, its running will be negligible. On the other hand, running could prove an excellent discriminator between more exotic inflationary scenarios.

Running could also present itself in features on top of a nearly scale invariant spectrum. Effectively the running would either be localized (a sharp feature) or oscillating. In this thesis we will discuss such a running of the power spectrum in ch. 6.

In addition to higher order contributions within the power spectrum, we could also consider higher order correlation statistics. Instead of only looking at the 2-point correlation function, correlating two points in the CMB sky, we could consider the 3-point correlation function (and beyond), which describes the correlation between different points in the sky (forming a triangle). In spherical harmonics the 3-point correlator becomes the bispectrum and the 4-point correlation the trispectrum. Again, for single-field slow-roll inflation such higher order spectra will be negligible (Maldacena 2003; Acquaviva et al. 2003) as they turn out to be slow-roll suppressed. Conceptually, this is not hard to understand. Single-field slow-roll is a free field, only weakly coupled to gravity and it has a very flat potential. This leads to nearly Gaussian distribution of primordial quantum fluctuations as its action consists of a canonical kinetic term that drives inflation down this potential. All scale dependence is hidden in the field slowly rolling down its potential, and therefore any deviations from Gaussianity are due to slow-roll alone. The bispectrum is a measure of the interacting theory of inflation and the interaction in single-field slow-roll are proportional to the coupling of the field. Therefore the bispectrum should be close to zero for single-field slow-roll inflation. In other words, if the initial distribution of quantum fluctuations is (nearly) Gaussian, all there is to learn about those initial conditions from the CMB is hidden in the power spectrum. Vice versa, a measurement of non-Gaussianity (through the bispectrum) would imply single-field slow-roll is excluded and instead a model including relevant interactions drives inflation. Note that even in such theories there are often ways to make non-Gaussianity unmeasurably small, but for single-field slow roll it is impossible to make non-Gaussianity large.

1.2.1 Non-Gaussianities

Before we can answer what Non-Gaussianity really means in the context of cosmological perturbations, let us briefly introduce the concept of Gaussianity. To understand Gaussianity we need to define a probability distribution of temperature fluctuations ΔT. The probability density function, or PDF, $P(\Delta T)$ tells us the probability of finding a temperature difference ΔT between pixels in the CMB sky. If the temperature anisotropy is Gaussian the PDF can be written as (Komatsu 2002, 2010):

$$P(\Delta T) = \frac{1}{(2\pi)^{N_{pix}/2}|\Theta|^{1/2}} \exp\left[-\frac{1}{2}\sum_{ij}\Delta T_i(\Theta^{-1})_{ij}\Delta T_j\right]. \tag{1.68}$$

Here $\Delta T_i = \Delta T(\hat{n}_i)$ and Θ_{ij} the 2-point correlation function $\langle\Delta T(\hat{n}_i)\Delta T(\hat{n}_j)\rangle$ (covariance matrix) of the CMB temperature anisotropy. N_{pix} is the number of pixels in the observed sky (for WMAP this is about $\sim 3\times10^6$, while Planck has about 10 times as many). Converting this to spherical harmonic space (eq. (1.60)) we obtain

$$P_G(a) = \frac{1}{(2\pi)^{N_{harm}/2}|C|^{1/2}} \exp\left[-\frac{1}{2}\sum_{lm}\sum_{l'm'}a^*_{lm}(C^{-1})_{lm,l'm'}a_{l'm'}\right]. \tag{1.69}$$

$N_{harm} = l(2l+1)$ is the number of l and m and $C_{lm,l'm'} = \langle a^*_{lm}a_{l'm'}\rangle$. Here a represents the set of $\{a_{lm}\}$, and P the probability distribution of that set. For statistically homogeneous and isotropic a_{lm} we can use eq. (1.64) and write

$$P_G(a) = \prod_{lm}\frac{\exp[-|a_{lm}|^2/2C_l]}{\sqrt{2\pi C_l}}. \tag{1.70}$$

The probability of finding an a_{lm} (in the set a) for Gaussian fluctuations is completely determined by the power spectrum C_l. For non-Gaussian fluctuations this is not the case: the PDF contains higher-order correlation functions. If non-Gaussianity is weak, and the trispectrum is small compared to the bispectrum and the power spectrum, one can expand the PDF around a Gaussian distribution (Taylor & Watts 2001; Komatsu 2010)

$$P_{NG}(a) = \left(1-\frac{1}{3!}\sum_{l_im_j}B^{m_1m_2m_3}_{l_1l_2l_3}\partial_{a_{l_1m_1}}\partial_{a_{l_2m_2}}\partial_{a_{l_3m_3}}\right)\times\frac{\exp[-\frac{1}{2}\sum_{ll'mm'}a^*_{lm}(C^{-1})_{lm,l'm'}a_{l'm'}]}{(2\pi)^{N_{harm}/2}|C|^{1/2}}. \tag{1.71}$$

Here $B^{m_1m_2m_3}_{l_1l_2l_3} \equiv \langle a_{l_1m_1}a_{l_2m_2}a_{l_3m_3}\rangle$ is known as the angular bispectrum. We can evaluate the derivatives with respect to the a_{lm} and obtain

$$P_{NG}(a) = P_G(a)\times\left(1+\frac{1}{6}\sum_{l_im_j}B^{m_1m_2m_3}_{l_1l_2l_3}[\prod_k(C^{-1}a)_{l_km_k} - 3(C^{-1})_{l_1m_1,l_2m_2}(C^{-1}a)_{l_3m_3}]\right). \tag{1.72}$$

$P_G(a)$ is the PDF for Gaussian fluctuations given by eq. (1.69). We can use this equation to determine how to optimally estimate $B_{l_1 l_2 l_3}^{m_1 m_2 m_3}$ from the data, as it should maximize this PDF (what the most likely B given the data). A common approach is to parametrize the bispectrum in terms of the angular reduced bispectrum $b_{l_1 l_2 l_3}$ and an amplitude $f_{\rm NL}$

$$B_{l_1 l_2 l_3}^{m_1 m_2 m_3} = \mathcal{G}_{l_1 l_2 l_3}^{m_1 m_2 m_3} \sum_i f_{\rm NL}^{(i)} b_{l_1 l_2 l_3}^{(i)}. \tag{1.73}$$

Here $\mathcal{G}_{l_1 l_2 l_3}$ is the Gaunt integral which constrains the 3 angular and azimuthal (l_i, m_i) paired coordinates to construct a triangle. It has the geometric solution

$$\begin{aligned} \mathcal{G}_{l_1 l_2 l_3}^{m_1 m_2 m_3} &= \int d^2\hat{n} Y_{l_1 m_1}(\hat{n}) Y_{l_2 m_2}(\hat{n}) Y_{l_3 m_3}(\hat{n}) \\ &= \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}, \end{aligned} \tag{1.74}$$

where $\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$ is the Wigner 3j symbol. When computing the bispectrum, such factors can be precomputed as they are purely geometric and do not depend on the underlying physics. Maximizing the PDF (eq. 1.72) with respect to $f_{\rm NL}$ yields constraints on the amplitude of the bispectrum, given a shape of the angular bispectrum $b_{l_1 l_2 l_3}$. The optimal estimator (Komatsu & Spergel 2001; Komatsu et al. 2003, 2005; Creminelli et al. 2007b; Yadav et al. 2007a,b) is thus obtained by solving $d \ln P_{NG}/df_{\rm NL}^{(i)} = 0$

$$f_{\rm NL}^{(i)} = \sum_j (F^{-1})_{ij} S_j. \tag{1.75}$$

with F_{ij} the Fisher matrix of $f_{\rm NL}^{(i)}$ and

$$\begin{aligned} S_i \equiv\ & \frac{1}{6} \sum_{lm} \mathcal{G}_{l_1 l_2 l_3}^{m_1 m_2 m_3} b_{l_1 l_2 l_3}^{i} \\ & \times \left[(C^{-1}a)_{l_1 m_1} (C^{-1}a)_{l_2 m_2} (C^{-1}a)_{l_3 m_3} - 3 (C^{-1})_{l_1 m_1, l_2 m_2} (C^{-1}a)_{l_3 m_3} \right]. \end{aligned} \tag{1.76}$$

The inverse of the Fisher matrix is the covariance matrix of the various $f_{\rm NL}$, describing the relation between two types of non-Gaussianity, i.e.,

$$(F^{-1})_{ij} = \langle f_{\rm NL}^{(i)} f_{\rm NL}^{(j)} \rangle - \langle f_{\rm NL}^{(i)} \rangle \langle f_{\rm NL}^{(j)} \rangle. \tag{1.77}$$

If two measured $f_{\rm NL}$ are completely independent (orthogonal), their Fisher matrix is diagonal and their covariance is zero, i.e., one is not affected by a change in the other. The 68% confidence interval of a measured $f_{\rm NL}^{(i)}$ is given by inverse of the diagonal component of the Fisher matrix, $(F^{-1})_{ii}$.

Let us define $e_l(\hat{n})$ and $d_{ll'}(\hat{n})$ as

$$e_l(\hat{n}) \equiv \sum_m (C^{-1}a) Y_{lm}(\hat{n}), \tag{1.78}$$

$$d_{ll'}(\hat{n}) \equiv \sum_{mm'} (C^{-1})_{lm,l'm'} Y_{lm}(\hat{n}) Y_{l'm'}(\hat{n}). \tag{1.79}$$

Together with the Gaunt integral, S_i can now be rewritten as

$$S_i = \frac{1}{6}\int d^2\hat{n} \sum_{l_1 l_2 l_3} b^{(i)}_{l_1 l_2 l_3} \left[e_{l_1}(\hat{n}) e_{l_2}(\hat{n}) e_{l_3}(\hat{n}) - 3 d_{l_1 l_2}(\hat{n}) e_{l_3}(\hat{n})\right]. \tag{1.80}$$

The explicit form of the Fisher matrix is given by

$$\begin{aligned} F_{ij} &= \frac{f_{sky}}{6} \sum_{lm} \sum_{l'm'} \mathcal{G}^{m_1 m_2 m_3}_{l_1 l_2 l_3} b^{(i)}_{l_1 l_2 l_3} & (1.81)\\ &\quad \times (C^{-1})_{l_1 m_1, l_1' m_1'} (C^{-1})_{l_2 m_2, l_2' m_2'} (C^{-1})_{l_1 m_3, l_1' m_3'} b^{(j)}_{l_1' l_2' l_3'} \mathcal{G}^{m_1' m_2' m_3'}_{l_1' l_2' l_3'}, & (1.82) \end{aligned}$$

where f_{sky} is the fraction of sky outside a mask applied to the full CMB sky data. A mask is typically applied to the full sky data, in order to block the galactic foreground, which is expected to contain significant non primordial non-Gaussianity. If the covariance matrix C_l is diagonal, $(C^{-1})_{lm,l'm'} = (C_l)^{-1}\delta_{ll'}\delta_{mm'}$ and the Fisher matrix reduces to

$$F_{ij} = \frac{f_{sky}}{6} \sum_l I^2_{l_1 l_2 l_3} \frac{b^{(i)}_{l_1 l_2 l_3} b^{(j)}_{l_1 l_2 l_3}}{C_{l_1} C_{l_2} C_{l_3}}, \tag{1.83}$$

with

$$I^2_{l_1 l_2 l_3} = \sum_m (\mathcal{G}^{m_1 m_2 m_3}_{l_1 l_2 l_3})^2 = \frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^2. \tag{1.84}$$

If non-Gaussianities are weak and the theoretically predicted bispectra are independent, the Fisher matrix is diagonal and our best estimator becomes

$$f^i_{\rm NL} = \frac{1}{f_{sky}} \sum_{2 \le l_3 \le l_2 \le l_1} \frac{S_i}{\frac{1}{\Delta_{l_1 l_2 l_3}} b_{l_1 l_2 l_3} C^{-1}_{l_1} C^{-1}_{l_2} C^{-1}_{l_3} b_{l_1 l_2 l_3} I^2_{l_1 l_2 l_3}}. \tag{1.85}$$

The summation over the l's now accounts for symmetry in the 3d multipole matrix through $\Delta_{l_1 l_2 l_3}$, which is equal to 1 when all l's are different, 2 if two are the same, and 6 if all are equal. S_i is data driven and depends on the type of non-Gaussianity one is trying to constrain. In app. B we explicitly derive the optimal estimator for enfolded type non-Gaussianities, inspired by a template proposed in ch. 2. This estimator is generally referred to as the KSW estimator, after Komatsu, Spergel and Wandelt who first introduced it.

1.2.2 Primordial Non-Gaussianity

For single-field slow-roll inflation the bispectrum is predicted to be small and likely unobservable (Maldacena 2003; Acquaviva et al. 2003). In order to produce large, possibly observable, non-Gaussianity one should break the conditions that apply to single-field slow-roll inflation:

- **Single Field**. Inflation is driven by only a single field. The initial conditions are set by this field alone.

- **Canonical Kinetic Energy**. The effective propagation speed of fluctuations in the inflaton field is equal to the speed of light, i.e., $c_s = c$.
- **Slow Roll**. The potential of the inflaton is flat enough for the evolution of the inflaton field to be slow compared to the Hubble time during inflation.
- **Initial State**. The initial vacuum state is given by the "Bunch Davies" vacuum.

If any of the conditions[11] above is broken, the initial distribution of quantum fluctuations can be significantly non-Gaussian. In this thesis, we will be mostly concerned with the last scenario, although we will implement this in multiple inflationary models.

Just as the power spectrum, the angular bispectrum can be related to the primordial bispectrum through

$$\begin{aligned} b_{l_1,l_2,l_3} &= \left(\frac{6}{5\pi}\right)^3 \int_{\eta_*}^{\eta_0} r^2 dr \int_{k_{min}}^{k_{max}} dk_1 dk_2 dk_3 k_1^2 k_2^2 k_3^2 F(k_1,k_2,k_3) \times \\ & j_{l_1}(k_1 r) j_{l_2}(k_2 r) j_{l_1}(k_2 r) \Delta_{l_1}(k_1) \Delta_{l_2}(k_2) \Delta_{l_3}(k_3), \end{aligned} \tag{1.86}$$

with j_l the spherical Bessel functions, Δ_l the radiation transfer functions and $F(k_1, k_2, k_3)$ the primordial shape function defined through the primordial bispectrum

$$\langle \zeta_{\vec{k}_1} \zeta_{\vec{k}_2} \zeta_{\vec{k}_2} \rangle = A \cdot (2\pi)^3 \delta \left(\sum_i \vec{k}_i \right) F(k_1, k_2, k_3), \tag{1.87}$$

Here A is defined (in the comoving gauge) as

$$A = (2\pi)^4 (\frac{3}{5} f_{\rm NL}) k^6 P_\zeta^2(k). \tag{1.88}$$

with $P_\zeta(k)$ given by eq. (1.52). The amplitude $f_{\rm NL}$ can be considered the equivalent of A_s in the primordial power spectrum. The shape F carries two *continuous* degrees of freedom since the bispectrum is constrained by two scales (or angles) that vary between k_{max} an k_{min} (see ch. 2 for a breakdown of the degrees of freedom of the primordial bispectrum). The bispectrum therefore carries significantly more information than the power spectrum alone. This makes the bispectrum an excellent measure to study different theories of inflation. If the bispectrum is scale independent, $F \propto k^{-6}$. Like the power spectrum, the bispectrum can be tilted, resulting in small deviations from this proportionality.

The (connected) bispectrum describes a triangle in comoving momentum k space (as well as a projected triangle on the CMB sky). The form of the triangle is contained within F, hence the name shape function[12]. Typically, we distinguish three different classes: local, equilateral and enfolded type non-Gaussianities. The classes represent 3 different triangles shown in fig. 1.5. In the next paragraphs we will explain how different shapes relate to various cosmological models. We will introduce the concept of factorizability, which is a condition required to be able to compute the bispectrum in reasonable times.

[11] Non-Einstein gravities is an additional possibility to produce observable levels of non-Gaussianity. We will not discuss such models here. See, e.g., Chen (2010) for a recent review.

[12] Sometimes the shape function is referred to as $S = F/(k_1 k_2 k_3)^2$, just as the power spectrum is sometimes called $P' = P/k^3$.

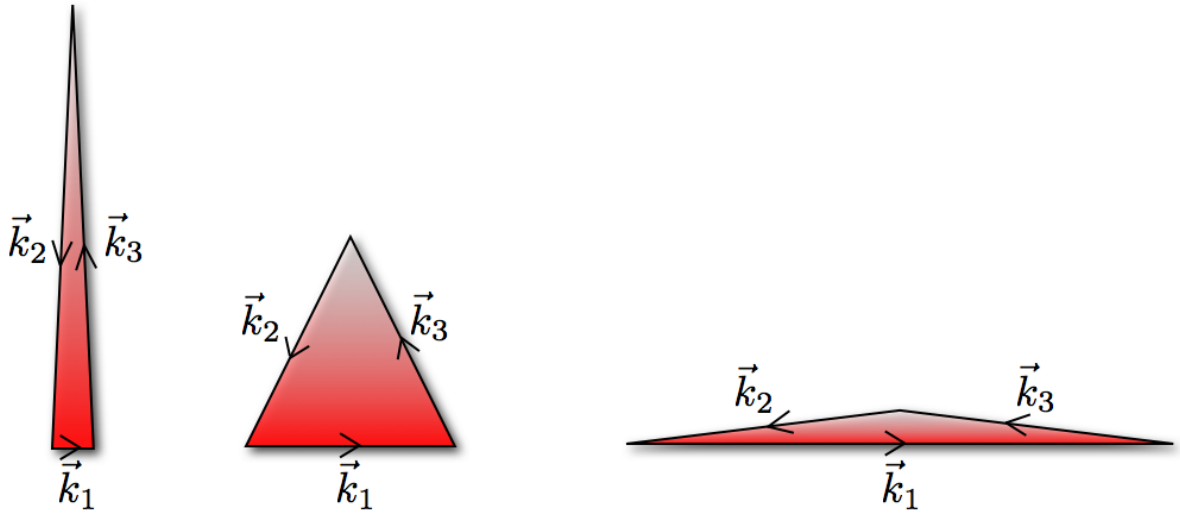

Figure 1.5: From left to right: local (squeezed), equilateral and enfolded (squashed) triangle.

1.2.3 Shapes of the Bispectrum

One of the key issues in constraining the bispectrum is the form of eq. (1.86). The computation of $b_{l_1 l_2 l_3}$ is numerically challenging. Generally, the primordial shape F can be any function of the three comoving momenta. When performing the k integration with a resolution n in each direction, it requires $\mathcal{O}(n^3)$ samples within the integration grid. Wang & Kamionkowski (2000) first showed that this issue was overcome once F became of the separable or factorizable form:

$$F(k_1, k_2, k_3) = \sum_{j=1}^{n} \prod_{i=1}^{3} f_{ij}(k_i) = f_{11}(k_1) f_{12}(k_2) f_{13}(k_3) + f_{21}(k_1) f_{22}(k_2) f_{23}(k_3) + ...$$

If F is of this form, the integrations over the 3 comoving momenta are separable and one now only has to compute $\mathcal{O}(3n)$ points (i.e., n for each of the 3 comoving momentum integrals). In terms of computing $f_{\rm NL}$, factorizability of the primordial bispectrum allows one to perform sums over l before doing integration over the sphere, since the there is no induced cross correlation between different l in the Bessel and transfer functions: in eq. (1.85) you can pull the summation over the l through the integration over r. For a full sky CMB experiment $l_{max} \propto \sqrt{N}$, with N the number of pixels in your map. The number of possible correlations scales as N^3, which reduces to $N^{5/2}$ given the triangle constraint. Another reduction by a factor of N occurs since instead of $(N^{1/2})^3$ (each l contributes $(N^{1/2})$, only $\sim N^{1/2}$ remains due to separability. Therefore the final number of computations for the bispectrum scales as $N^{3/2}$. With $N \sim 10^6$ for WMAP and $N \sim 10^7$ for Planck, these can lead to significant reductions in computational time[13].

[13]Fergusson & Shellard (2007) showed that choosing a different basis in k space allows one to circumvent factorizability in certain cases. The idea is to rewrite the primordial shape in coordinates that make it separable and then compute the corresponding Jacobi matrix. In their coordinate transformation, the original primordial shape should depend on $f(k_t)$, with $k_t = k_1 + k_2 + k_3$. DBI inflation, for example, has a dependence like this.

Local Non-Gaussianities

In general, predicted bispectra need not be of the factorizable form. The best constrained bispectrum, the local bispectrum, is in fact of the factorizable form. This bispectrum was first introduced by Komatsu & Spergel (2001) when they tried to expand the primordial gravitational potential as a Taylor expansion around a Gaussian part

$$\Phi = \Phi_G + f_{\rm NL}^{\rm local}\left(\Phi_G^2 - \langle\Phi_G\rangle^2\right), \tag{1.89}$$

where $\langle\Phi_G\rangle^2$ is subtracted to explicitely set $\langle\Phi\rangle = 0$. Computing the bispectrum to first order in the parameter $f_{\rm NL}^{\rm local}$ gives the shape function

$$F_{\rm local}(k_1, k_2, k_3) = C\left[\frac{1}{k_1^{4-n_s}k_2^{4-n_s}} + (2\text{ perm.})\right]. \tag{1.90}$$

The constant C depends on the choice of gauge. It should be considered a normalization factor that can be used to compare different bispectra. Initially the normalization was chosen such that all shapes had the same amplitude in the equilateral limit. However, as was pointed out by Fergusson & Shellard (2007) and Meerburg et al. (2009) such a normalization condition is not very convenient since some shapes have very little power in this limit. Therefore Fergusson & Shellard (2007) proposed a more intuitive normalization, namely the square root of the shape's inner product (see ch. 2 how to define such an inner product). In some literature C contains the amplitude $f_{\rm NL}$, but here we have chosen to incorporate that into A (eq. (1.88)), which is the definition of F we use throughout this thesis[14]. The local shape F describes triangles in comoving momentum space with a squeezed configuration. In other words, the amplitude of the angular bispectrum will be large when one of the comoving wavevectors (momenta) is small compared to the other two $k_i \ll k_j, k_l$. Maldacena (2003) and Acquaviva et al. (2003) showed that single field slow roll inflation is 'dominated' by a bispectrum of this form. In the squeezed limit the amplitude $f_{\rm NL}$ is $\mathcal{O}(n_s - 1)$ and therefore $f_{\rm NL}$ from the simplest model of inflation should be unobservably small. Examples of inflationary models that predict observable levels of local non-Gaussianity are multifield models (Enqvist & Vaihkonen 2004; Allen et al. 2004; Vernizzi & Wands 2006; Kim & Liddle 2006; Battefeld & Easther 2007; Battefeld & Battefeld 2007; Yokoyama et al. 2008; Huang et al. 2008) and curvaton models (Lyth et al. 2003; Bartolo et al. 2004; Sasaki et al. 2004; Malik & Lyth 2006; Assadullahi et al. 2007) while the ekpyrotic model (Koyama et al. 2007; Buchbinder et al. 2008; Lehners & Steinhardt 2008) also predicts local type non-Gaussianities. Planck is expected to constrain $f_{\rm NL}^{\rm local} \sim 5$. Currently the best constraint on $f_{\rm NL}^{\rm local}$ is derived from WMAP 7 data (Komatsu et al. 2011)

$$f_{\rm NL}^{\rm local} = 32 \pm 21 (68\% CL), \tag{1.91}$$

using the optimal estimator obtained from eq. (1.85) for local type non-Gaussianities. There are several other ways to extract information on local non-Gaussianities from the CMB data.

[14] It does not really matter what type of normalization is chosen, as long as one properly accounts for all the factors when extracting non-Gaussianities from the data and that the choice of normalization is consistent among extracted types of non-Gaussianity.

Examples include Minkowski functionals (Hikage et al 2006, 2008), needlet bispectrum analysis (Rudjord et al. 2009; Cabella et al 2009) and spherical wavelet decomposition (Aliaga et al. 2002; Curto et al 2009). In recent years, the use of large scale structure (LSS) data has shown to be an alternative observable to constrain non-Gaussianities. Early results have been very promising (Xia et al. 2011) and it is generally assumed that future LSS surveys should be more constraining than CMB measurements.

Equilateral Non-Gaussianities

Creminelli (2003) showed that if the dynamics of the inflaton is dominated by higher derivative terms (non-canonical kinetic term) the initial distribution of inflaton perturbations could become non-Gaussian. Generally a non-canonical action (Chen et al. 2007)

$$S = \frac{1}{2}\int d^4x\sqrt{-g}[M_p^2 R - 2P(X,\phi)], \tag{1.92}$$

where ϕ is the inflaton field and $X = g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi$, describes a set of inflationary models (Creminelli 2003; Silverstein & Tong 2004; Alishahiha et al. 2004; Chen 2005b,c; Khoury & Piazza 2008; Langlois et al. 2008; Kinney & Tzirakis 2008) that produce non-Gaussianities of the equilateral type, which maximizes when all comoving momenta are equal, i.e., for $k_1 \sim k_2 \sim k_3$. The theoretically predicted bispectra derived from this action are generally not of the factorizable form. In order to constrain equilateral bispectra from the data using the KSW estimator, Creminelli et al. (2006) proposed an equilateral template that is of the separable form and mimics the shape of predicted bispectra

$$F_{\rm equil}(k_1,k_2,k_3) = C\left[-F_{\rm local} - \frac{2}{(k_1k_2k_3)^{\frac{2}{3}(4-n_s)}} + \left[\frac{1}{k_1^{\frac{1}{3}(4-n_s)}k_2^{\frac{2}{3}(4-n_s)}k_3^{4-n_s}} + 5\text{ perm.}\right]\right], \tag{1.93}$$

with C again a normalization constant. The addition of non-canonical kinetic terms leads to perturbations that propagate with a reduced speed of sound. For a canonical field $c_s = c$, with c the speed of light. The predicted amplitude of equilateral non-Gaussianities depends on the speed of sound as $f_{\rm NL} \propto 1 - 1/c_s^2$ (Chen et al. 2007). Here c_s is in units of $c = 1$, such that for $c_s = c = 1$ the amplitude vanishes (i.e., for canonical models). For Dirac Born Infeld (DBI) models of inflation, where the inflaton represents the coordinate of D branes in the warped throat of a Calabi Yau manifold, the speed of sound could be as small as $c_s = \mathcal{O}(0.1)$. Therefore, non-Gaussianities produced in such a scenario could be as large as $f_{\rm NL} \sim 100$. We will investigate such models in ch. 3. Recently it was shown that in holographic models (McFadden & Skenderis 2010), non-Gaussianities are exactly of the equilateral form (McFadden & Skenderis 2011), i.e., they are equivalent to the equilateral template. The predicted amplitude $f_{\rm NL} \sim \mathcal{O}(1)$, is unobservable with current generation instruments.

The best constraint on equilateral (template) non-Gaussianities comes from WMAP 7 (Komatsu et al. 2011)

$$f_{\rm NL}^{\rm equil} = 26 \pm 140(68\%CL). \tag{1.94}$$

Enfolded Non-Gaussianities

The third triangle in fig. 1.5 represents non-Gaussianities of the enfolded type, which have their maximal contributions to the angular bispectrum when $k_i \sim k_j + k_l$. Theoretically, such bispectra were first predicted by Chen et al. (2007) due to initial state modifications. In this context, the initial state refers to the vacuum state during inflation. Usually, the vacuum state is taken to be the Bunch-Davies vacuum, which is an assumption that will be explained in sec. 1.2.4 of this chapter. Again, generally the theoretically predicted enfolded spectra are not factorizable. In ch. 2 we have made an attempt to build a separable form of the enfolded bispectrum

$$F_{\rm enf}(k_1,k_2,k_3) \;=\; C\left[F_{\rm local} + \frac{3}{(k_1k_2k_3)^{\frac{2}{3}(4-n_s)}} - \left[\frac{1}{k_1^{\frac{1}{3}(4-n_s)}k_2^{\frac{2}{3}(4-n_s)}k_3^{4-n_s}} + 5\text{ perm.}\right]\right]. \tag{1.95}$$

Non-Gaussianities of this form have not been constrained by CMB data, however Xia et al. (2011) have constrained it from LSS data and found

$$f_{\rm NL}^{\rm enf} \;=\; 183 \pm 95(68\%CL). \tag{1.96}$$

Orthogonal Non-Gaussianities and other types

The orthogonal bispectrum was first introduced by Senatore et al. (2009) as a linear combination between the enfolded and equilateral triangle: for positive values of the comoving momenta, the template has a maximum in the equilateral limit, while for negative momenta, it has a maximum in the collinear or enfolded limit. It is known as orthogonal, since it can be shown to be orthogonal to equilateral and local type bispectra. Such a bispectrum arises under very special conditions in an effective field theory description of inflation. Although Senatore et al. (2009) claimed this to be a new type of non-Gaussianity, Chen et al. (2007) already made mention of this type of bispectrum. In fact, a simple parameter transformation shows that both results are completely equivalent (see ch. 3).

In that same paper Senatore et al. (2009) introduced a factorizable template, that mimics this orthogonality to the equilateral template (eq. 1.93). The way they derived this template is very similar to the way we derived the enfolded template (see app. B)

$$F_{\rm ort}(k_1,k_2,k_3) \;=\; C\left[-F_{\rm local} - \frac{8}{(k_1k_2k_3)^{\frac{2}{3}(4-n_s)}} + - \left[\frac{1}{k_1^{\frac{1}{3}(4-n_s)}k_2^{\frac{2}{3}(4-n_s)}k_3^{4-n_s}} + 5\text{ perm.}\right]\right]. \tag{1.97}$$

This bispectrum has been constrained by several groups, including the WMAP 7 collaboration, which again leads to the most stringent constraint

$$f_{\rm NL}^{\rm ort} \;=\; -202 \pm 104(68\%CL). \tag{1.98}$$

The bispectral shapes introduced here, represent a large class of models. However, there exist other bispectra. Quite often these bispectra can be described by some weighted linear combination of the bispectral shapes introduced above. In some cases however, this is not possible. For example, in this thesis we will find that certain theoretical models predict bispectra that contain features or oscillations (Meerburg et al. 2009). For those bispectra, factorizability can be an issue. In ch. 5 we will make an attempt to resolve this issue by expanding these oscillating bispectra into bispectra that are factorized by construction, a method first introduced by Fergusson & Shellard (2007).

1.2.4 On Initial State Modifications

In the final paragraph of this section, we will put forward our motivation to study possible deviations from a Bunch Davies vacuum during inflation. The main part of the research reported in this thesis has been to compute effects from such a modification on the initial conditions, that is the primordial probability distributions of the inflaton perturbations. Before we present our motivation, let us briefly recapitulate what we have discussed so far.

Currently, the best model of the Universe is ΛCDM. This model contains (at least) 6 parameters and these parameters can be measured using cosmological observations such as the CMB and the LSS. Two of these 6 parameters represent the initial conditions. In particular, they describe the distribution of curvature fluctuations at the end of inflation through the power spectrum, the Fourier equivalent of the 2-point correlation function. These fluctuations are induced by quantum fluctuations that become classical as they are stretched beyond the causal horizon during inflation. Those initial fluctuations in the curvature of spacetime couple to the contents of the Universe; the power spectrum of inflaton perturbations seeds the power spectrum of temperature fluctuations observed in the CMB, and the power spectrum of density fluctuations in the LSS. It is precisely this coupling that allows us to probe the primordial conditions through observations of the CMB power spectrum. Although plain vanilla ΛCDM is an elegant picture of the Universe, we have argued that it might not completely describe the initial distribution of fluctuations. The reason to be interested in these initial conditions[15] is that it could possibly teach us the underlying physics of inflation. Since the typical energy scale of inflation is expected to be around the GUT scale (10^{16} GeV), this represents an indirect window on new physics beyond the Standard Model. In addition, the distribution of small fluctuations in the metric is sourced by quantum fluctuations at very high energy scales. Hypothetically, this could provide an opportunity to probe even higher energy scales, possibly close to the Planck scale (10^{19} GeV).

If the initial distribution of inflaton perturbations is perfectly Gaussian, the CMB temperature fluctuations can be fully described by the power spectrum. In order to extract additional information from the CMB through the power spectrum, one might consider a model that contains more than just two parameters in the primordial power spectrum. In addition to the tilt, one could search for a change of tilt as a function of scale, known as the running, or one could

[15]To avoid confusion, let us emphasize that here with initial conditions we refer to the distribution of fluctuations set by inflation and determined after each scale has crossed the horizon due to the cosmological expansion. These initial conditions are set *by* inflation, while these on their turn are determined by the initial conditions *of* inflation, which is the main topic of this thesis.

consider features on top of an almost scale invariant primordial spectrum. We will discuss the presence of oscillations on top of an almost scale invariant power spectrum in ch. 6. If instead the initial distribution of fluctuations is not perfectly Gaussian, there is more information in the CMB than represented by the power spectrum alone. Additional knowledge on the physics of inflation could be contained in the higher order statistics of density fluctuations, such as the bispectrum, the Fourier transform of the 3-point correlations function. What the bispectrum does is equivalent to what direct measurement of large scale structure does for dark matter and what the LHC does for particle physics. It probes the dynamics of the field sourcing inflation. Measurements of these interactions through the bispectrum (and higher order correlations statistics) allows us to probe the Lagrangian of the field and hence the effective description of physics at inflationary energy scales (Komatsu et al. 2009).

In this thesis we investigate not only the Lagrangian of the field that drives inflation, but also the exact vacuum state it thrived from. Before we can explain why we are interested in probing the initial vacuum state, let us introduce the concept of the vacuum in quantum physics. In quantum mechanics we define the vacuum as the state that is zero after applying the annihilation operator on that state, i.e., if

$$\hat{a}_{\vec{k}}|\Psi\rangle \quad = \quad 0. \tag{1.99}$$

then $|\Psi\rangle$ is the vacuum state. For a free falling observer in a flat, Minkowski spacetime, a non-interacting quantum field can naturally be decomposed into time-independent quanta, or particles. These quanta correspond to (Fourier) modes with a fixed time-independent frequency ω_k. Consequently the particle number operator $N_k = \hat{a}^*_{\vec{k}}\hat{a}_{\vec{k}}$, which counts the number of quanta in a given state, is a conserved quantity. We can define a state $|\Psi\rangle$ with zero quanta, $\langle\Psi|N_k|\Psi\rangle = 0$ and this state will remain empty at all times. In general spacetimes however, the number of field quanta is not a conserved quantity. For example in de Sitter, a curved spacetime, we can pick a state as in eq. (1.99), but as time evolves, this state will become excited and will quickly fill with particles due to the time-dependent de Sitter background. A related phenomenon is Hawking radiation near the horizon of Black Holes (Hawking 1974) as well as Unruh radiation (Fulling 1973; Davies 1975; Unruh 1976) for accelerated observers in Minkowski space.

The spacetime during inflation is close to de Sitter, so we expect N_k not to be a conserved quantity and we can not unambiguously define a time-independent empty or vacuum state. The general solution to the equations of motion in an inflating background is a superposition of a decaying and growing modes (as opposed to a plane wave in Minkowski space)

$$\zeta_{\vec{k}}(\eta) \quad = \quad \frac{iH}{\sqrt{4\epsilon k^3}}\left[P_+\left(1+ik\eta\right)e^{-ik\eta} + P_-\left(1-ik\eta\right)e^{ik\eta}\right]. \tag{1.100}$$

It should be apparent from the equation above that we can recover a vacuum similar to that of Minkowski space if we consider very early times ($\eta \to -\infty$). Then the second term in this expression vanishes and we recover a solution with constant ω_k. This choice of vacuum is known as the Bunch Davies (BD) vacuum (Bunch & Davies 1978) and corresponds to a solution $P_- = 0$ (and $P_+ = 1$), resulting in eq. (1.49). This is commonly referred to as the adiabatic regime as $\dot{\omega}/\omega^2 \to 0$ in this limit. We consider only those modes that are all deep inside the horizon

and are small compared to the curvature radius, and as such effectively all relevant modes experience Minkowski spacetime. The standard procedure to compute primordial correlation functions, such as the power spectrum and the bispectrum, is to assume the empty state is the BD vacuum state.

However, there is a serious concern with choosing the BD state as the empty state in an inflationary background. Although the effective theory of inflation is assumed to be only weakly interacting, as soon as we consider earlier times, interactions are expected to become more relevant. The BD state is defined in the assumption that the solutions to the equation of motion are still given by eq. (1.100) when we take the limit $\eta \to -\infty$. Once interactions become important, the mode functions will no longer correspond to plane waves. *This is our theoretical motivation to study initial state modifications*: the presence of interactions in the past almost certainly affects the vacuum state.

Besides this qualitative motivation, we are also driven by pragmatism; can we constrain deviations from a BD state using cosmological observations? *This is our phenomenological motivation.* If the data are constraining enough, we might conclude that the vacuum state is BD after all, and we have to reexamine our theoretical motivation, or we conclude that deviations are simply too small for detection, which constrains the parameters describing inflation.

The origin and the magnitude of a deviation from a BD vacuum has been discussed in several papers (Martin et al. 2000; Martin & Brandenberger 2001; Danielsson 2002; Easther et al. 2002; Schalm et al. 2004; Chen 2010c; Chen & Dent 2011) . Usually these modifications are referred to as "Trans Planckian Effects", given that physics near Planck scale energies is responsible for a vacuum modification. In order to constrain initial state modification phenomenologically, we need need to consider a particular representation of such a modification. An initial state modification is can be modeled using a (diagonal) Bogolyubov transformation (which is roughly represented by P_- and P_+). Suppose we have two sets of mode functions u_k and v_k, both solutions to the equations of motion. We can then rewrite v_k in terms of u_k

$$v_k^* \quad = \quad \alpha_k u_k^*(\eta) + \beta_k u_k(\eta), \tag{1.101}$$

since the modes u_k form a basis. The commutation relation between the annihilation and creation operators puts a normalization constraint on the mode functions, which tells us that

$$|\alpha|^2 - |\beta|^2 = 1. \tag{1.102}$$

If we expand our modes ζ_k in terms of the new mode functions u_k and u_k^*, with (new) annihilation $\hat{A}_{\vec{k}}$ and creation $\hat{A}^\dagger_{\vec{k}}$ operator one can show

$$\hat{A}_{\vec{k}} = \alpha_k \hat{a}_{\vec{k}} + \beta_k^* \hat{a}^\dagger_{\vec{k}}, \;\; \hat{A}^\dagger_{\vec{k}} = \alpha_k^* \hat{a}^\dagger_{\vec{k}} + \beta_k \hat{a}_{\vec{k}}. \tag{1.103}$$

The relation (1.103) and the complex coefficients α_k and β_k are known as a Bogolyubov transformation and the Bogolyubov coefficients respectively. β_k will be prescribed by a given theoretical framework. Generally, we can distinguish two different scenarios. The first scenario is known as the "New Physics Hypersurface" (NPH). In the NPH scenario one traces every momentum mode k back to some large physical cut-off scale and imposes the standard flat space vacuum state (the BD state) mode by mode, resulting in a prediction for β_k that is independent of k,

which only gives rise to a small departure from scale invariance after taking into account the slow-roll evolution of the Hubble parameter.

In the second scenario one fixes an initial time where one calculates corrections to the usual BD initial condition using boundary effective field theory. The result is a Bogolyubov parameter β_k depending on k, resulting in an explicit breaking of scale-invariance in the 2-point power spectrum. This modification is known as the Boundary Effective Field Theory, or BEFT, proposal of initial state modifications.

The difference between these two proposals is that in BEFT one breaks scale invariance explicitly, as each mode will cross the *fixed time* with a different scale size, therefore all observable scales will be modified at the same time but differently. In the NPH scenario however, the modification is done at a *fixed scale*, so each scale is modified similarly, albeit at different times. The breaking of scale invariance is due to the dynamics of the geometry during inflation, causing fixed scales to be modified differently as different times correspond to slightly different geometries. Both these scenarios will be topic of discussion in the following chapters. A constraint on β from cosmological observables, in particular the CMB, is one of the key subjects of the research presented in this thesis.

1.3 Guide to this Thesis

To conclude the introduction, let us briefly present the content of each chapter. In the next 3 chapters (2, 3, and 4) we will investigate the consequences of modifying the initial state during inflation on the production of a non-Gaussian distribution of fluctuations. We will compute the primordial bispectrum in the context of the NPH scenario, inserting a fixed cut-off scale. The Bogolyubov parameter will only be weakly scale dependent and the resulting bispectrum will therefore practically be scale invariant. The bispectrum due to initial state modifications has been investigated by Porrati (2004b);BEFT and Holman & Tolley (2008);NPH. Recently, Ashoorioon & Shiu (2011) and Chen & Dent (2011) have contributed some new insights. In ch. 2 we will consider such a modification primarily in the context of single field slow-roll inflation and slow-roll inflation including one higher order derivative term. We find that adding such a term, boosts the amplitude of non-Gaussianity significantly. This boost is expected, as such a term introduces additional interactions in the inflaton field. As explained, a deviation from the BD vacuum state introduces a non-zero particle density and the interactions between those particles is responsible for the increased amplitude. This finding motivates the work of ch. 3. In this chapter we extend our analysis to a general class of inflationary models which have a non-canonical kinetic term. These include DBI models of inflation, which are thought to be a possible candidate inflationary model derived from string theory. In order to constrain possible deviations from a Bunch Davies state, we compare our predicted bispectra with the latest constraints from WMAP. We also consider complex Bogolyubov coefficients instead of only real ones, as was assumed in ch. 2. We show that the predicted bispectra are different from those currently constrained. Consequently, the derived constraints are rather weak. Therefore we propose a new template (the enfolded template) and we investigate the improvement this new template might have on future constraints.

The bispectra derived in ch. 2 and 3 are independent of the form of the Bogolyubov trans-

formation. In ch. 4 we consider a specific prediction for β first proposed by Danielsson (2002). From the primordial bispectrum, we derive constraints on several important parameters, and show that under certain conditions such modifications can be ruled out by the data.

One of the important results in this thesis is that bispectral shapes F coming from initial state modifications typically contain (a large number of) oscillations. As a result, comparing observed bispectra (which are smooth and usually positive definite) with the predicted bispectra is a challenge. As explained, one of the key issues in constraining oscillatory bispectra directly from the data is that these need to be of the separable form, at least for the KSW estimator of non-Gaussianities. In order to achieve separability we will investigate the method of mode expansion (Fergusson & Shellard 2009; Fergusson et al. 2010a) for oscillatory bispectra in ch. 5. We show that such an expansion could lead to a much more efficient extraction of such bispectra from the data.

Besides oscillations in the primordial bispectrum, it was long known before that initial state modifications also predict oscillations in the primordial power spectrum (Easther et al. 2002; Greene et al. 2004). The predicted amplitude of these corrections to the power spectrum depends on the model (BEFT vs. NPH), but is typically small. We complete this thesis by looking at constraints on oscillations in the power spectrum using WMAP 7 data in ch. 6. We investigate the predictions from BEFT and from a recently proposed and more general NPH scenario (Jackson & Schalm 2010, 2011), which both give distinguishable predictions for the primordial power spectrum of curvature fluctuations. We compute the likelihood of such oscillations and derive preliminary constraints on some of the parameters describing the initial conditions. This chapter represents work in progress and we hope to finalize this work in the coming months.

We end this thesis with a brief epilogue, summarizing the aim of and the results presented in the thesis. We shortly describe some (new) insights not included when the papers were written and suggest possible directions in which to continue this research.

Chapter 2

Signatures of Initial State Modifications on Bispectrum Statistics

P. Daniel Meerburg, Jan Pieter van der Schaar & Pier-Stefano Corasaniti
JCAP **0905** (2009) 018

Abstract: Modifications of the initial-state of the inflaton field can induce a departure from Gaussianity and leave a testable imprint on the higher order correlations of the CMB and large scale structures in the Universe. We focus on the bispectrum statistics of the primordial curvature perturbation and its projection on the CMB. For a canonical single-field action the three-point correlator enhancement is localized, maximizing in the collinear limit, corresponding to enfolded or squashed triangles in comoving momentum space. We show that the available local and equilateral template are very insensitive to this localized enhancement and do not generate noteworthy constraints on initial-state modifications. On the other hand, when considering the addition of a dimension 8 higher order derivative term, we find a dominant rapidly oscillating contribution, which had previously been overlooked and whose significantly enhanced amplitude is independent of the triangle under consideration. Nevertheless, the oscillatory nature of (the sign of) the correlation function implies the signal is nearly orthogonal to currently available observational templates, strongly reducing the sensitivity to the enhancement. Constraints on departures from the standard Bunch-Davies vacuum state can be derived, but also depend on the next-to-leading terms. We emphasize that the construction and application of especially adapted templates could lead to CMB bispectrum constraints on modified initial states already competing with those derived from the power spectrum.

2.1 Introduction

It has been known for some time (Kofman 1991; Falk et al. 1993; Komatsu 2002) that the potential presence of non-Gaussian signatures in the CMB is a powerful probe of the physics of inflation and beyond. Computations of the primordial bispectra (Arkani-Hamed et al. 2004; Senatore 2005; Moss & Xiong 2007; Bernardeau & Brunier 2007; Barnaby & Cline 2007a; Lyth et al. 2003; Bartolo et al. 2004; Sasaki et al. 2004; Malik & Lyth 2006; Assadullahi et al.

2007; Huang et al. 2008; Creminelli 2003; Chen et al. 2007, 2008; Bartolo et al. 2005; Lyth & Rodriguez 2005a,b; Bernardeau et al. 2006b; Moss & Graham 2007; Enqvist & Vaihkonen 2004; Allen et al. 2004; Vernizzi & Wands 2006; Kim & Liddle 2006; Battefeld & Easther 2007; Battefeld & Battefeld 2007; Barnaby & Cline 2007b; Yokoyama et al. 2008; Huang et al. 2008; Koyama et al. 2007; Buchbinder et al. 2008; Lehners & Steinhardt 2008; Koyama et al. 2007; Buchbinder et al. 2008; Lehners & Steinhardt 2008; Chen 2005a; Seery & Lidsey 2005; Chen et al. 2007) (and later trispectra (Komatsu 2002; Byrnes et al. 2006; Huang & Shiu 2006; Seery & Lidsey 2007; Arroja & Koyama 2008)) have shown that different models of inflation can produce rather unique features, which would allow, when detected, to discriminate between them. For the bispectrum, the distinction between models relies on two features 1) the overall amplitude of the non-Gaussian signal and 2) the detailed dependence on the comoving momenta. Obviously, when the overall amplitude of the signal is low, the second feature will be much harder to observe. Observational limitations due to foreground contamination (Creminelli & Zaldarriaga 2004; Serra & Cooray 2008) and cosmic variance limit the detection of non-Gaussianity in the CMB temperature and polarization spectrum (Komatsu & Spergel 2001; Babich & Zaldarriaga 2004b). For that reason one can already conclude that non-Gaussianity should be observably absent if a single, slowly rolling, scalar field is responsible for inflation (Acquaviva et al. 2003; Maldacena 2003).

Even if a model predicts a detectable non-Gaussian amplitude, it will remain a challenge to measure the actual momentum dependence, since the inferred constraints on the level of non-Gaussianity (Komatsu et al. 2003; Creminelli et al. 2006, 2007a; Yadav & Wandelt 2007c; Komatsu et al. 2009) are based on a sum over all modes of a pre-assumed momentum dependence[1]. Such dependencies are known as 'local' (or 'squeezed') and 'equilateral' template, which correspond to particular shapes that maximize in some 'extreme' triangle configuration in momentum space. The possibility to distinguish between different theoretical models producing a sizable non-Gaussian amplitude relies on the fact that in the models considered so far the produced non-Gaussianities are well approximated by *one* of these templates. For example, it has been shown (Seery & Lidsey 2005; Khoury & Piazza 2008) that non-canonical kinetic terms and higher derivative contributions to the inflaton potential can produce significant levels of non-Gaussianity of the equilateral type if the speed of sound in these models is much smaller than the speed of light, which can be realized in certain brane inflation scenarios (Silverstein & Tong 2004; Chen 2005a). Local shape non-Gaussianities were the first type to be considered (Bond & Salopek 2007; Pyne & Carroll 1996; Komatsu & Spergel 2001) and are a direct consequence of the nonlinear relation between the inflaton fluctuations and the curvature perturbations that couple to matter and radiation. Acquaviva et al. (2003) and Maldacena (2003) showed that the amplitude of local type non-Gaussianities in single-field slow-roll infla-

[1]Other methods such as Minkowski Functionals (see (Schmalzing et al. 1995) for theory and (Komatsu et al. 2009) and references therein for observational results) and a Wavelet approach ((Hobson et al. 1998) and subsequent papers) exist which typically do not rely on a pre-assumed momentum dependence. Here however we refer to the approach initiated by Komatsu & Spergel (2001) and further developed by Komatsu et al. (2003); Babich et al. (2004a); Liguori et al. (2006); Kogo & Komatsu (2006); Creminelli et al. (2007b); Yadav et al. (2007a); Liguori et al. (2007); Yadav et al. (2007b), which seems to give the most consistent and stringent constraints (Komatsu et al. 2009) so far.

tion is proportional to the slow-roll parameter ϵ [2], which is very small by construction. The amplitude of local type non-Gaussianities should therefore be undetectably small if single-field slow-roll is responsible for inflation. In contrast, large local non-Gaussianities can be generated in curvaton models (Lyth et al. 2003; Bartolo et al. 2004; Sasaki et al. 2004; Malik & Lyth 2006; Assadullahi et al. 2007; Huang 2008), where the curvature perturbation ζ can evolve outside the horizon, or inflationary models with multiple scalar fields. Models of new ekpyrosis (Koyama et al. 2007; Buchbinder et al. 2008; Lehners & Steinhardt 2008) and the recently proposed contracting models with an increasing speed of sound (Khoury & Piazza 2008), in which a bouncing universe is replacing inflation, also yield large non-Gaussianities of the local type.

In this chapter we will focus on non-Gaussian features arising from an arbitrary initial-state modification. This type of non-Gaussianity has been discussed earlier by Martin et al. (2000); Gangui et al. (2002); Porrati (2004a,b); Chen et al. (2007); Holman & Tolley (2008), but here we provide a more detailed analysis, in particular on detectability. In the language of boundary effective field theory (Schalm et al. 2004) one can generally divide the contributions into two parts; non-Gaussianities coming directly from the initial-state boundary (which are absent when considering Gaussian initial-state modifications), and 'bulk' non-Gaussianities generated by the presence of (interacting) particles in the modified initial-state. In the boundary effective field theory formalism it has been shown that the leading non-Gaussian initial-state modification is of the local type (Porrati 2004a,b). However, the 'bulk' non-Gaussianities generated by the non-zero Bogolyubov coefficient seem to have a unique momentum dependence, which is very different from that of the local and equilateral types (Seery & Lidsey 2005; Chen et al. 2007; Holman & Tolley 2008). For example, for a canonical single-field inflaton action, in momentum space the non-Gaussian signal produced by a, possibly Gaussian, modification of the initial-state maximizes for triangles where two momentum vectors are collinear, i.e., when the magnitude of one of the comoving momenta equals the sum of the other two: $k_p = k_q + k_r$ with $p \neq q \neq r$ (a squashed, flattened or enfolded triangle), and are known as collinear or enfolded type non-Gaussianities.

Thus far only the local and equilateral type of non-Gaussianities have been constrained by the data (Creminelli et al. 2007a; Komatsu et al. 2009), although recently a strong case has been made for a more general set-up (Fergusson & Shellard 2009). There are essentially two reasons for this. First of all, the realization that initial-state modifications give rise to a unique non-Gaussian shape, that might even be detectable due to subtle enhancements, is rather recent and its theoretical motivation might be considered less compelling. Putting aside plausible theoretical concerns associated with modifications of the vacuum state and instead taking a phenomenological point of view, deviations from the standard Bunch-Davies state are tightly constrained (Spergel et al. 2007) because of their unique oscillatory signatures in the 2-point power spectrum (see Greene et al. (2004) and references therein). As emphasized by Holman & Tolley (2008) the bispectrum (and possibly even higher n-point functions) might be as good, or even better, in constraining initial-state modifications. Clearly, with the expected future improvements in detecting primordial non-Gaussian signals, it is worthwhile to look for the

[2] This is the semi-classical, tree level, result. The effects of quantum loop corrections have been studied by Weinberg (2005, 2006); Seery (2007); Riotto & Sloth (2008) and Cogollo et al. (2008).

presence of enfolded type or, as we will argue in this chapter, oscillatory non-Gaussianities in the CMB data to constrain initial-state modifications. The second more pragmatic reason why enfolded type non-Gaussianities have not been compared to the data yet is that in analyzing the data computational limitations demand that the momentum dependence is *factorizable*. Generic 3-point correlators are not factorizable, so one resorts to constructing a factorizable template that approximates the actual theoretical bispectrum, maximizing in the appropriate 'extreme' triangle. It is this factorizable template that is then compared to the data. Such templates have been constructed for the local (Komatsu 2002) and equilateral shapes (Creminelli et al. 2006), but has not yet been constructed for the type of non-Gaussianities predicted by initial-state modifications, which are typically expected to extremize in an enfolded (collinear, squashed or flattened) triangle. The goal of this chapter is two-fold: to present a detailed analysis on the detectability of non-Gaussianities produced by initial-state modifications using currently available templates, both with and without higher derivative corrections, and secondly to determine how much improvement can theoretically be gained by using more optimal templates.

The chapter is organized as as follows. In sec. 2.2 we review the standard analytical tools to study non-Gaussianity, in particular the computation of the 3-point correlation function in momentum space and its relation to different triangular shapes. In sec. 2.3 we will present a detailed analysis of the 3-dimensional bispectrum from initial-state modifications in the single-field slow-roll inflationary scenario. In sec. 2.4 we analyze the case of modified initial-state non-Gaussianities in the presence of a dimension 8 higher order derivative term in the Lagrangian. In sec. 2.6 we discuss the results of the CMB bispectrum computation and finally we present our conclusions in sec. 2.7.

2.2 Three-Dimensional Bispectrum Preliminaries

In this section we will briefly review the standard tools for analyzing non-Gaussianities as first described by Babich et al. (2004a). In the next sections we will apply these tools to the case of initial-state modifications. Let us start considering the primordial spectrum of curvature perturbations generated by the inflaton. In three-dimensional comoving momentum space a generic three-point correlator of the curvature perturbation $\zeta_{\vec{k}}$ is a function of the three comoving momenta $\vec{k}_1$, $\vec{k}_2$ and $\vec{k}_3$, which in 3 dimensions corresponds to a total of 9 parameters. Translational invariance forces the three-point function to conserve momentum

$$\langle \zeta_{\vec{k}_1} \zeta_{\vec{k}_2} \zeta_{\vec{k}_2} \rangle \;=\; A \cdot (2\pi)^3 \delta \left(\sum_i \vec{k}_i \right) F(\vec{k}_1, \vec{k}_2, \vec{k}_3), \tag{2.1}$$

which fixes one of the momenta, reducing the number of free parameters from 9 to 6. Rotational invariance allows one to pick a 2-dimensional plane defined by the remaining two momenta and adjust the axes such that one of the momenta is along one of the axes of the plane. This fixes another $2 + 1 = 3$ parameters, leaving only 3 variables to parametrize the three-point correlator. These can be identified with two angles and the overall scale of the triangle formed by $\vec{k}_1$, $\vec{k}_2$, $\vec{k}_3$. Since the primordial power spectrum is approximately scale invariant, we expect the correlator to be a homogeneous function F of degree -6 in comoving momentum space, i.e.,

$F(\lambda\vec{k}_1, \lambda\vec{k}_2, \lambda\vec{k}_3) = \lambda^{-6} F(\vec{k}_1, \vec{k}_2, \vec{k}_3)$. So (approximate) scale invariance fixes the dependency of the three-point correlator on the scale of the triangle, further reducing the number of free parameters to the 2 angles. Instead of writing the function F in terms of these angles, it is most convenient to consider the two independent ratios given by the magnitudes of the comoving momenta $x_2 \equiv k_2/k_1$ and $x_3 \equiv k_3/k_1$. In order to determine the relevant x_2, x_3 domain, one assumes $k_1 \geq k_2 \geq k_1$, giving $x_2 \leq 1$ and $x_3 \leq 1$, and then uses the triangle constraint to find that $1 - x_2 \leq x_3 \leq 1$, identifying the top-right triangle in x_2, x_3 space (see, e.g., 2.3). Since the distributions are symmetric in x_2 an x_3, one could further reduce the domain by half only considering $x_3 \geq x_2$. Hitherto, unlike the power spectrum, which only depends on the 'reciprocal distance' between two-points, the bispectrum F depends on two variables, typically represented by the ratios of the magnitudes of the comoving momenta $F = F(x_2, x_3)$.

To measure the overall amplitude A in eq. (2.1), one assumes a particular theoretical template shape function $F(x_2, x_3)$, sums over all triangles and then normalizes appropriately, taking into account the variance of a given mode in Fourier space. This procedure leads to the following estimator of the non-Gaussian amplitude A

$$\hat{A} \;=\; \frac{\sum_{\vec{k}_i} \zeta_{\vec{k}_1}\zeta_{\vec{k}_2}\zeta_{\vec{k}_2} F(\vec{k}_1, \vec{k}_2, \vec{k}_3)/\left(\sigma^2_{k_1}\sigma^2_{k_2}\sigma^2_{k_3}\right)}{\sum_{\vec{k}_i} F^2(\vec{k}_1, \vec{k}_2, \vec{k}_3)/\left(\sigma^2_{k_1}\sigma^2_{k_2}\sigma^2_{k_3}\right)}. \tag{2.2}$$

Here the σ_k represent the variances of the different modes and the sum runs over all triangles in momentum space. The above estimator naturally defines a scalar product between two distributions F_X and F_Y as (Babich et al. 2004a)

$$F_X \cdot F_Y \;=\; \sum_{\vec{k}_i} \frac{F_X(\vec{k}_1, \vec{k}_2, \vec{k}_3) F_Y(\vec{k}_1, \vec{k}_2, \vec{k}_3)}{\sigma^2_{k_1}\sigma^2_{k_2}\sigma^2_{k_3}}. \tag{2.3}$$

This scalar product allows us to quantitatively verify how well a particular template distribution, say F_X, can be used to constrain a theoretical signal described by the distribution F_Y. In terms of the (reduced set of) parameters x_2, x_3 the sum over triangles can be written as an integral with an appropriate measure equal to $x_2^4 x_3^4$

$$F_X \cdot F_Y \;\propto\; \int dx_2 dx_3 F_X(x_2, x_3) F_Y(x_2, x_3) x_2^4 x_3^4. \tag{2.4}$$

To derive optimal constraints using a template F_X one would like the scalar product, or the overlap, to be as large as possible. Using the scalar product one can construct a normalization independent 'cosine' between two distributions

$$\mathrm{Cos}(F_X, F_Y) \;\equiv\; \frac{F_X \cdot F_Y}{(F_X \cdot F_X)^{1/2}(F_Y \cdot F_Y)^{1/2}}\,, \tag{2.5}$$

which is close to 1 for shapes that are very similar and considerably smaller than 1 for shapes that are very distinct. It follows that optimal constraints can be obtained only if the (factorizable) templates, which are used to analyze the data, have a cosine close to 1 with the theoretically predicted non-Gaussian signal. Nevertheless, for non-optimal templates F_X one

can still derive constraints on a theoretically predicted non-Gaussian signal F_Y provided one introduces the so-called fudge factor Δ_F, defined as (Babich et al. 2004a)

$$\Delta_F \;=\; \frac{F_Y \cdot F_X}{(F_X \cdot F_X)} = \mathrm{Cos}(F_X, F_Y) \left(\frac{F_Y \cdot F_Y}{F_X \cdot F_X} \right)^{1/2} . \tag{2.6}$$

The fudge factor allows to deduce the relevant constraints for different theoretical predictions F_Y using the results inferred from the data analysis of a particular template distribution F_X. In such a case the constraint on the amplitude of the type F_X will be degraded by a factor $1/\Delta_F$, thus the smaller the scalar product between F_Y and F_X the weaker the constraints on the F_Y type non-Gaussianities using the F_X template. Looking at eq. (2.6) it should be clear that optimal constraints can be achieved by maximizing the cosine between the template and the theoretical prediction. The other contribution to the fudge factor has to do with some conventional choice of normalization for the template and the theoretical distribution involved and can be adapted accordingly. We will apply these techniques to obtain constraints on non-Gaussianities predicted by modified initial-states using the latest results on local and equilateral type non-Gaussianities, and to derive what can (theoretically) be gained by analyzing the data with an more optimal (enfolded) template.

Let us briefly discuss the normalization conventions for the non-Gaussian amplitudes, which are important for a correct interpretation and comparison of the results obtained for different distributions. To compare the local and equilateral template one typically equates the distributions in the equilateral triangle $k_1 = k_2 = k_3$ (Babich et al. 2004a). We will follow this convention, which allows us to directly use the constraints from the CMB for the local and equilateral non-Gaussian amplitudes. To be explicit, for the local template distribution the standard definition of the $f_{\mathrm{NL}}^{\mathrm{local}}$ parameter, starting from the general three-point function in eq. (2.1), is related to the amplitude A of the three-point function of curvature perturbations in the following way[3]

$$A \;=\; (2\pi)^4 \left(-\frac{3}{5} f_{\mathrm{NL}}^{\mathrm{local}} \right) \frac{\Delta_\Phi^2}{k_1^6} , \tag{2.7}$$

where $\Delta_\Phi = \frac{1}{8\pi^2} \frac{H^2}{\epsilon M_p^2}$ is the amplitude of the two-point power spectrum, which has been observed to be approximately equal to 10^{-10}, M_p is the reduced Planck mass and $\epsilon = \frac{1}{2} M_p^2 \left(\frac{V'}{V} \right)^2$ is the first slow-roll parameter. In the above expression for the amplitude A we included the overall k_1 scaling dependence, implying that the local shape F^{local} can be identified as the following function of the reduced number of variables x_2, x_3

$$F^{\mathrm{local}}(x_2, x_3) \;=\; 2 \left(\frac{1}{x_2^3} + \frac{1}{x_3^3} + \frac{1}{x_2^3 x_3^3} \right) , \tag{2.8}$$

For non-Gaussianities of the equilateral type it was shown by Creminelli et al. (2006); Babich et al. (2004a) that these are well approximated by the following shape function

$$F^{\mathrm{equil}}(x_2, x_3) \;=\; 6 \left[-\frac{1}{x_2^3} - \frac{1}{x_3^3} - \frac{1}{x_2^3 x_3^3} - \frac{2}{x_2^2 x_3^2} + \left(\frac{1}{x_2^2 x_3^3} + 5 \text{ perm} \right) \right] , \tag{2.9}$$

[3] Our sign convention for the non-Gaussian amplitude follows (Babich et al. 2004a; Maldacena 2003), which is different from that used by Komatsu et al. (2005).

where the normalization has been fixed such that the local and equilateral template shape functions both equal 6 in the equilateral limit $x_2 = x_3 = 1$. Comparing to the local template definition of the non-Gaussian amplitude $f_{\rm NL}$, this then suggests a similar definition of $f_{\rm NL}^{\rm equil}$

$$A \;=\; (2\pi)^4 \left(-\frac{3}{5} f_{\rm NL}^{\rm equil}\right) \frac{\Delta_\Phi^2}{k_1^6}. \tag{2.10}$$

A crucial property of the local and equilateral template is that they are factorized in their comoving momentum dependence. This allows for a drastic (and necessary) reduction in the computational time needed to compare the template distributions to the CMB data, yielding constraints on the parameters $f_{\rm NL}^{\rm local}$ and $f_{\rm NL}^{\rm equil}$. The analysis of the WMAP-5 year data for local and equilateral non-Gaussianities gives the following limits, (Komatsu et al. 2009)

$$\begin{aligned} -9 \;&<\; f_{\rm NL}^{\rm local} < 111 \\ -151 \;&<\; f_{\rm NL}^{\rm equil} < 253\,, \end{aligned} \tag{2.11}$$

which we will use in sec. 2.6. It is worth stressing that the non-Gaussian amplitude $f_{\rm NL}$ is not uniquely defined, it depends on a specific choice for the shape function F_X, which is equivalent to fixing the integrated norm $|F_X| \equiv \sqrt{F_X \cdot F_X}$. It is the combination $f_{\rm NL}^X |F_X|$ that is independent of a particular normalization scheme and which measures the (integrated) non-Gaussian amplitude. Obviously any choice will do, as long as one properly takes into account the corresponding norm $|F_X|$ when for example deducing constraints on the non-Gaussian amplitude $f_{\rm NL}^X$ from the equilateral and local template results.

2.3 Modified Initial-State Non-Gaussianities

Theoretically predicted three-point functions, evaluated in the regular Bunch-Davies vacuum state, describe a non-Gaussian signal of either local or equilateral type, depending on whether higher derivative corrections play a significant role in the inflationary evolution. If this is the case, as in DBI models of inflation (Silverstein & Tong 2004; Alishahiha et al. 2004), then the dominant contribution is of the equilateral type and can be large enough to be detectable in the near future. The existence of different shapes can be nicely understood in terms of the nonlinear origin of the non-Gaussian signal. For the local shape it is the nonlinear relation between the inflaton and the curvature perturbation on super-horizon scales that produces the maximal effect, whereas in the DBI case nonlinear effects in the inflaton sector are most relevant and maximize when all momenta cross the horizon.

Chen et al. (2007) and Holman & Tolley (2008) showed non-Gaussian effects can also be generated by dropping the assumption that the vacuum state is Bunch-Davies. To fundamentally address the vacuum state ambiguity one would first need a full understanding of physics at the highest energy scales, where the description in terms of a free inflaton field is expected to break down, as well as the physics preceding inflation. Nevertheless departures from the free Bunch-Davies state can be studied on a phenomenological basis and it seems worthwhile to use observations to constrain the possibilities. The two-point power spectrum already provides strong constraints on the initial-state, which has to be close to Bunch-Davies (Spergel et al.

2007). Interestingly though, according to (Holman & Tolley 2008), three-and higher n-point functions might be very constraining as well, mainly due to subtle enhancement effects, which increase the non-Gaussian amplitude in collinear or enfolded triangles. In this section we will focus on the simplest case, with a three-point correlation function derived in the general context of slow-roll inflation, but evaluated in a vacuum state different from standard Bunch-Davies, as parametrized by an undetermined Bogolyubov parameter β_k. The leading non-Gaussian contribution due to the appearance of a negative frequency mode is essentially obtained by swapping the sign of one of the comoving momenta in the slow-roll inflation result (Maldacena 2003). In app. C we confirm the result first derived by Holman & Tolley (2008) that the correction to the three-point correlation function is given by

$$\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle^{\rm nBD} = (2\pi)^3\delta^{(3)}\left(\sum\vec{k}_i\right)\frac{1}{M_p^2}\frac{4}{\prod(2k_i^3)}\frac{H^6}{\dot{\phi}^2}\sum_j\frac{3k_1^2k_2^2k_3^2}{k_j^2\tilde{k}_j}\mathcal{R}e(\beta_{k_j})\left(\cos(\tilde{k}_j\eta_0)-1\right) \tag{2.12}$$

In the above expression $\tilde{k}_j = \sum_i k_i - 2k_j$ and η_0 represents the initial conformal time which has to be introduced to ensure that the non-Gaussian effects can consistently be calculated using an effective field theory description valid below some physical cutoff momentum scale M (Holman & Tolley 2008), i.e. non-Gaussianities generated by initial-state modifications are sensitive to the details of the ultraviolet complete theory and the physical cutoff scale M is introduced to parameterize our ignorance. This suggests that the initial time η_0 should be a function of the comoving momentum k, allowing the combination $|k\,\eta_0(k)| = M/H \gg 1$ to be a large fixed number independent of k. This prescription treats all comoving momenta equivalently, tracing back different comoving momenta from the time their physical momentum equals the cutoff scale M, preserving scale invariance. Instead considering η_0 to be some fixed initial time would immediately result in a breaking of scale invariance because different comoving momenta would receive contributions from a different range of physical momentum scales in the time integral involving the interaction Hamiltonian (see app. C).

This is very reminiscent of the distinction between the New Physics Hypersurface (NPH) (Danielsson 2002; Easther et al. 2002) and Boundary Effective Field Theory (BEFT) (Schalm et al. 2004; Greene et al. 2004) proposals to model initial-state modifications. In the latter case one fixes an initial time where one calculates corrections to the usual Bunch-Davies initial condition using boundary effective field theory. The result is a Bogolyubov parameter β_k depending on k, resulting in an explicit breaking of scale-invariance in the two-point power spectrum. In the NPH scenario one traces every momentum mode back to some large physical cutoff scale M and imposes the standard flat space vacuum state (corresponding to positive frequency modes only), mode by mode, resulting in a prediction for β_k that is independent of k, which only gives rise to a small departure from scale-invariance after taking into account the slow-roll evolution of the Hubble parameter. Note that the NPH vacuum state proposal is not grounded (yet) in some effective field theory scheme that can systematically be applied to calculate quantum state corrections, as opposed to the BEFT approach. Both proposals have their problems and, as emphasized before, should at this stage and for our purposes here be considered as purely phenomenological models distinguished primarily by their consequences for scale-invariance.

For the bispectrum, considering an initial time η_0 independent of k (BEFT) or a k-independent

combination $|k\eta_0(k)| = M/H$ (NPH) immediately results in a breaking of scale-invariance, even independent of any specific k-dependent prediction for β_k. Because the tools to analyze non-Gaussian shapes introduced in the previous section crucially rely on scale-invariance of the bispectrum, we will assume that the initial time $\eta_0(k)$ depends on the comoving momentum such that $|k\eta_0| = M/H$ is k-independent, in the spirit of the scale-invariant NPH scenario. It would be interesting to relax this assumption and apply more general techniques, for instance those recently developed by (Fergusson & Shellard 2009), to analyze the scale-dependent non-Gaussianities that arise in a BEFT approach of initial-state modifications.

Looking at the dependence on comoving momenta, we can see that the three-point correlator eq. (2.12) maximizes in the collinear or enfolded triangle defined by $\tilde{k}_1/k_1 = \pi/|k_1\eta_0| \sim 0$, and with an amplitude proportional to $k_1\eta_0$. Here k_1 is assumed to be the largest comoving momentum in the triangle, whose overall power-law dependence manifests the expected scale invariance of the three-point function. Similarly to the local and equilateral shapes, this enfolded type of non-Gaussianity can be associated to a dominant source of nonlinearities, in this specific case this is related to the unavoidable presence of interacting particles in the modified initial-state at sub-horizon scales. Based on the $|k_1\eta_0| = M/H \gg 1$ enhancement of the non-Gaussian signal a rough order of magnitude estimate of the observational constraints on modified initial-state non-Gaussianities was given by Holman & Tolley (2008). However, their estimate was inferred by considering the maximum signal in the enfolded limit, and directly compared to existing bounds on the local non-Gaussian amplitude. In contrast, a full analysis of the sensitivity of current non-Gaussian constraints on departures from the Bunch-Davies vacuum must involve integrating over all triangles and crucially relies on the scalar product between the theoretical template prediction and the different observational template distributions. We will address this issue throughout the rest of this chapter.

In order to proceed and calculate the scalar product, cosine and fudge factor, we need to determine the dominant contribution to the shape function and identify the corresponding non-Gaussian amplitude. Starting from eq. (2.12) we identify the relevant comoving momentum dependent part as

$$\begin{aligned} F^{\rm modin}(k_1,k_2,k_3) &= \frac{1}{k_1k_2k_3}\left\{\frac{1-\cos[\eta_0(k_1+k_2-k_3)]}{k_3^2(k_1+k_2-k_3)}+\frac{1-\cos[\eta_0(k_1-k_2+k_3)]}{k_2^2(k_1-k_2+k_3)}+\right. \\ &\left.\frac{1-\cos[\eta_0(-k_1+k_2+k_3)]}{k_1^2(-k_1+k_2+k_3)}\right\}, \end{aligned} \tag{2.13}$$

which by having scaled out the standard k_1^{-6} dependence leads to the corresponding definition for the amplitude A,

$$A = (2\pi)^4\,(3\epsilon|\beta|)\,\frac{\Delta_\Phi^2}{k_1^6},$$

where we have replaced $\mathcal{R}e(\beta_k)$ with the absolute value $|\beta|$. By comparing eq. (2.14) to the amplitudes of the local and equilateral templates, a standard definition of $f_{\rm NL}^{\rm enf}$ suggests that $f_{\rm NL}^{\rm enf} = 5\epsilon|\beta|$. Consequently, without any enhancement from a large fudge factor this non-Gaussian amplitude is obviously undetectable, since it is suppressed by both the slow-roll

parameter ϵ and the Bogolyubov parameter $|\beta|$. In terms of the reduced variables x_2, x_3 we have

$$F^{\rm modin}(k_1\eta_0, x_2, x_3) = \frac{1}{x_2 x_3}\left\{\frac{1-\cos[k_1\eta_0(1+x_2-x_3)]}{x_3^2(1+x_2-x_3)}+\frac{1-\cos[k_1\eta_0(1-x_2+x_3)]}{x_2^2(1-x_2+x_3)}+\frac{1-\cos[k_1\eta_0(-1+x_2+x_3)]}{(-1+x_2+x_3)}\right\}. \quad (2.14)$$

We now explicitly see the dependence of the shape function on $|k_1\eta_0| = \frac{(|k_1/a_0|)}{H} = M/H$, namely the ratio of the physical cutoff scale to the Hubble parameter, as was explained earlier. Again, not fixing the combination $|k_1\eta_0|$ to be k_1-independent results in an obvious breaking of scale-invariance and would not allow us to use the introduced tools for comparison with the available templates. The cutoff scale should be significantly larger than the Hubble scale and we will typically consider it to be somewhere in between $10^2 - 10^3$. This implies the shape function is rapidly oscillating, which complicates the evaluation of the integrals to determine the cosines and fudge factors with the available templates. When possible the integrals were evaluated analytically in the limit $|k_1\eta_0| \gg 1$. Let us compute the squared norm of the modified initial-state shape function given by

$$\left|F^{\rm modin}(k_1\eta_0, x_2, x_3)\right|^2 = \int_0^1 dx_2 \int_{1-x}^1 dx_3 \left[F^{\rm modin}(k_1\eta_0, x_2, x_3)\right]^2 x_2^4 x_3^4 = \frac{\pi}{60}|k_1\eta_0| + \frac{5}{4}\log|k_1\eta_0| + 6.05\,, \quad (2.15)$$

where the integral in x_2, x_3 space is over the (triangle) domain $0 \leq x_2 \leq 1$, $1 - x_2 \leq x_3 \leq 1$. From eq. (2.15) we can derive some important conclusions about the detectability of this non-Gaussian signal. In an ideal situation the data analysis would be performed using the theoretical template eq. (2.13) to directly infer on the non-Gaussian amplitude. As previously discussed this is the product of the normalization A times the norm of the shape function, $A|F^{modin}|$. Hence, the best one can do by using an observational template perfectly aligned with the theoretical prediction is a leading enhancement factor of order $\sqrt{|k_1\eta_0|}$. However, such an enhancement is lost when the data analysis is performed using a local template, due to the integrated nature of the non-Gaussian analysis. Evaluating the scalar product as defined by eq. (2.4) between the initial-state modification template and the local one we find

$$F^{\rm modin}(k_1\eta_0, x_2, x_3)\cdot F^{\rm local}(x_2, x_3) = \int_0^1 dx_2 \int_{1-x}^1 dx_3 F^{\rm modin}(k_1\eta_0, x_2, x_3) F^{\rm local}(x_2, x_3)\, x_2^4 x_3^4 = 3\log|k_1\eta_0| + 18.96. \quad (2.16)$$

We can already conclude that when using the local template to probe modified initial state non-Gaussianities the enhancement factor is further reduced to become only logarithmic in $|k_1\eta_0|$, instead of the $\sqrt{|k_1\eta_0|}$ enhancement that can be achieved in the optimal case. It is worth remarking that the constant parts in the results, for both the norm and the scalar product, depend on how some singular integrals, those independent of $k_1\eta_0$, are being cutoff. The singular integrals always blow up in the local (squeezed) limit, corresponding to one of the

momenta being much smaller than the other two. Fortunately, a natural cutoff is given by the fact that only a finite range of modes contribute to the CMB. Specifically, the ratio between the smallest and the largest observable scales on the CMB is roughly equal to 10^{-3}, which is used to regularize the local type integrals. Throughout this chapter we will always quote results using this cutoff if necessary. For the squared norm of the local shape, which clearly blows up in the squeezed limit, the need for this cutoff is most apparent. To be more explicit, using the 10^{-3} cutoff the result for the squared norm of the local shape equals $\left|F^{\rm local}(x_2,x_3)\right|^2 = 176.5$. From the scalar product and the local and modified initial-state norms we can infer the cosine factor, which reads as

$$\begin{aligned}\mathrm{Cos}\left[F^{\rm modin},F^{\rm local}\right] &= \frac{F^{\rm modin}(k_1\eta_0,x_2,x_3)\cdot F^{\rm local}(x_2,x_3)}{|F^{\rm modin}(k_1\eta_0,x_2,x_3)|\,|F^{\rm local}(x_2,x_3)|}\\ &= 7.53\cdot 10^{-2}\,\frac{(18.96+3\log|k_1\eta_0|)}{\sqrt{6.05+\frac{\pi}{60}|k_1\eta_0|+\frac{5}{4}\log|k_1\eta_0|}}\,. \end{aligned} \tag{2.17}$$

The fudge factor necessary to transform the limits on local type non-Gaussianities into constraints on modified initial-state non-Gaussianities is given by

$$\begin{aligned}\Delta_F\left[F^{\rm modin},F^{\rm local}\right] &= \frac{F^{\rm modin}(k_1\eta_0,x_2,x_3)\cdot F^{\rm local}(x_2,x_3)}{\left|F^{\rm local}(x_2,x_3)\right|^2}\\ &= 5.67\cdot 10^{-3}\,(18.96+3\log|k_1\eta_0|)\,. \end{aligned} \tag{2.18}$$

As we already concluded from the scalar product alone, the fudge factor is logarithmically dependent on $k_1\eta_0$. In addition, the coefficient is also relatively small, implying that over a realistic range range of $|k_1\eta_0|$ values, the fudge factor can essentially be considered constant.

Similarly, we calculate the scalar product between the equilateral template and the modified initial-state distribution, which a priori can be expected to depend on $k_1\eta_0$ as well. Surprisingly, the leading $k_1\eta_0$ dependent terms cancel and the only contribution comes from a, cutoff independent, constant number for the scalar product

$$\begin{aligned}F^{\rm modin}(k_1\eta_0,x_2,x_3)\cdot F^{\rm equil}(x_2,x_3) &= \\ \int_0^1 dx_2\int_{1-x}^1 dx_3 F^{\rm modin}(k_1\eta_0,x_2,x_3)F^{\rm equil}(x_2,x_3)\,x_2^4x_3^4 &= 6.5\,. \end{aligned} \tag{2.19}$$

We can therefore conclude that all enhancement is lost when using the equilateral template to probe modified initial-state non-Gaussianities. As for the local template this will imply a constant fudge factor, even though the theoretical non-Gaussian distribution is linearly enhanced in enfolded triangles. The squared norm of the equilateral shape functions is also cutoff independent (i.e., finite), and the numerical integration gives $\left|F^{\rm equil}(x_2,x_3)\right|^2 = 7.9$. Combining this with the squared norm of the modified initial-state shape function this leads to the following expression for the normalization independent cosine

$$\begin{aligned}\mathrm{Cos}\left[F^{\rm modin},F^{\rm equil}\right] &= \frac{F^{\rm modin}(k_1\eta_0,x_2,x_3)\cdot F^{\rm equil}(x_2,x_3)}{|F^{\rm modin}(k_1\eta_0,x_2,x_3)|\,|F^{\rm equil}(x_2,x_3)|}\\ &= \frac{2.31}{\sqrt{6.05+\frac{\pi}{60}|k_1\eta_0|+\frac{5}{4}\log|k_1\eta_0|}}\,. \end{aligned} \tag{2.20}$$

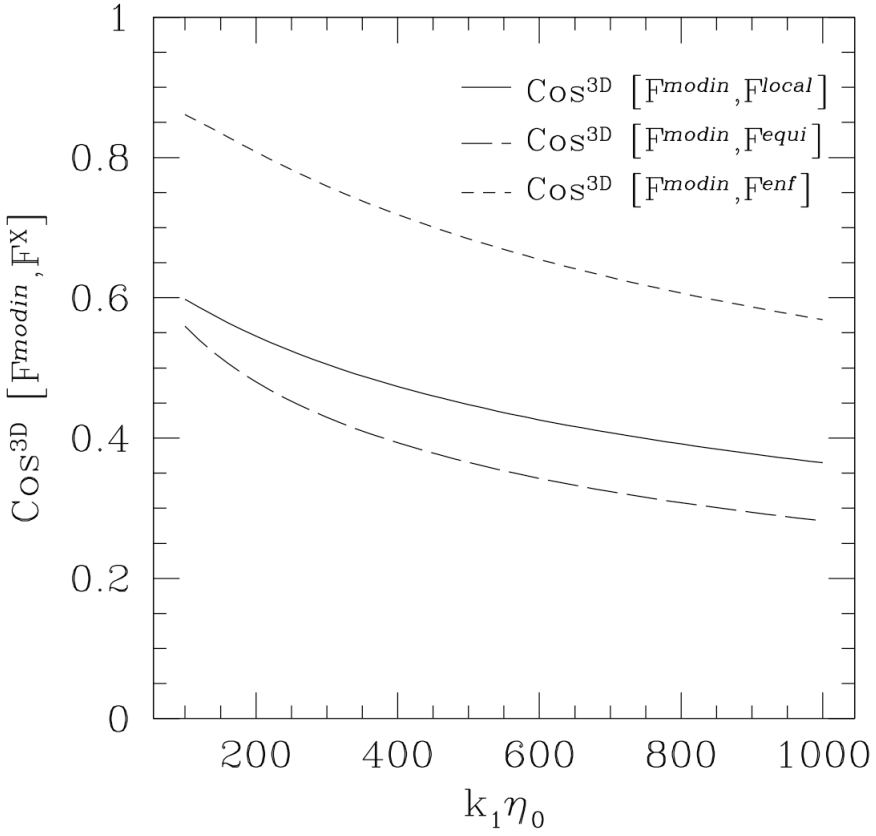

Figure 2.1: Cosines factors between the initial-state modification shape and the local (solid line), equilateral (long dashed line) and the enfolded template proposal (short dashed line) as functions of $|k_1\eta_0|$ in 3-D.

Even though the cosine is a function of $k_1\eta_0$, while the scalar product is not, the fudge factor will also be independent and equals

$$\begin{aligned}\Delta_F\left[F^{\rm modin},F^{\rm equil}\right] &= \frac{F^{\rm modin}(k_1\eta_0,x_2,x_3)\cdot F^{\rm equil}(x_2,x_3)}{\left|F^{\rm equil}(x_2,x_3)\right|^2}\\ &= 0.82\,. \end{aligned} \tag{2.21}$$

The constancy of the fudge factor explicitly confirms that all enhancement due to the large $|k_1\eta_0|$ parameter is lost. In fig. 2.1 and 2.2 we plot the cosine and fudge factors between the initial-state modification and the local (solid line) and equilateral (long dashed line) templates as function of $k_1\eta_0$. From the plot of the cosine factor we infer that indeed the local and equilateral templates poorly overlap with the modified initial-state distribution as the $\cos[F^{\rm modin},F^X]<0.6$. We conclude that although the non-Gaussian amplitude of initial-state modifications is linearly enhanced in enfolded triangles, the measured local and equilateral

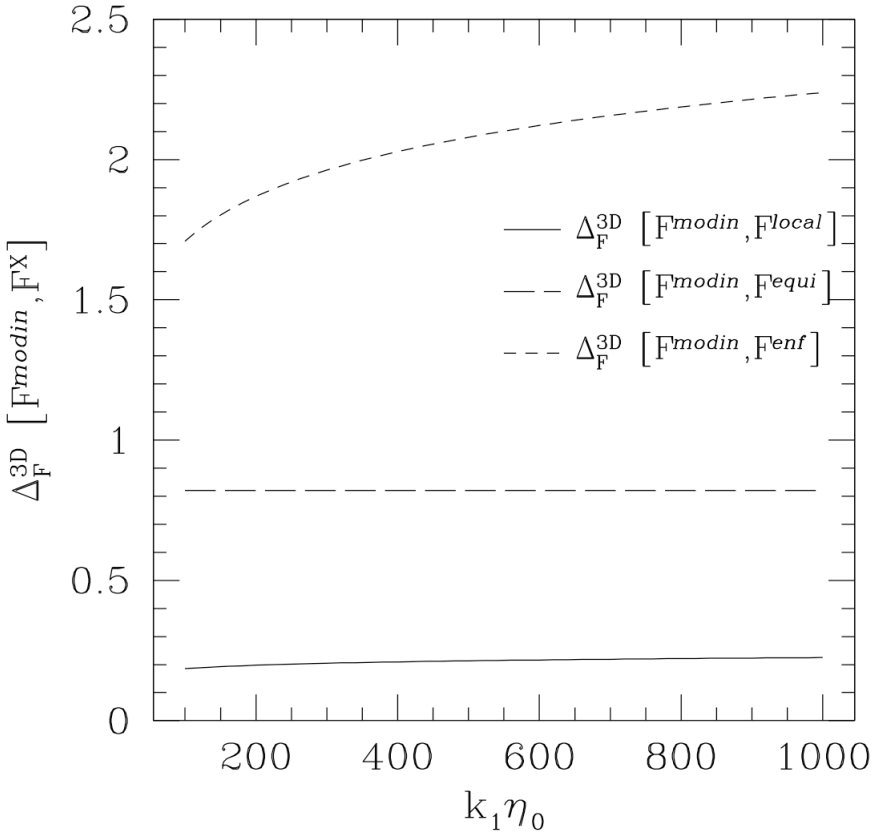

Figure 2.2: Fudge factors as functions of $|k_1\eta_0|$ for the templates.

templates are completely insensitive to this localized enhancement, thus spoiling any chance of obtaining a stringent bound on departures from the standard Bunch-Davies vacuum state.

Consequently, probing standard slow-roll modified initial-state non-Gaussianities is impossible unless a new template distribution is introduced which, unlike the local and equilateral templates, is sensitive to the localized enfolded enhancement. As pointed out in the previous discussion, using a perfect template will lead to a signal enhancement of $\sqrt{|k_1\eta_0|}$. In sec. 2.5 we will describe a first proposal for such an improved, factorized, enfolded template. In the next section we will focus on the combined effect of a specific higher derivative correction and an initial-state modification on the bispectrum.

2.4 Adding Higher Derivative Corrections

As was shown in the previous section, the enhancement effect of an initial-state modification in the bispectrum, assuming standard slow-roll inflation, is impossible to probe using the currently

available local or equilateral templates. What we would like to study is whether the same conclusion holds after adding higher derivative corrections, which according to (Holman & Tolley 2008) could be even more sensitive to initial-state modifications. A priori, one might expect similar conclusions, that even though there is a strong enhancement effect in the enfolded triangle limit, its measure in the space of all triangles versus the local or equilateral template will again be too small to allow detection. We consider the addition of a dimension 8 higher derivative term to the scalar field lagrangian of the following form:

$$\Delta\mathcal{L}_{\rm HD} = \sqrt{-g}\frac{\lambda}{8M^4}\left((\nabla\phi)^2\right)^2 , \tag{2.22}$$

where the scale M corresponds to the high energy cutoff scale and 'natural' corrections would correspond to a coupling $\lambda \sim 1$. This higher derivative correction is the same as the one discussed by Holman & Tolley (2008) and was first studied by Creminelli (2003). We provide a detailed derivation of the corresponding bispectrum in app. C. Creminelli (2003) showed that assuming the standard Bunch-Davies vacuum, this leads to non-Gaussianities of the equilateral type with an amplitude $f_{\rm NL}^{\rm equil} \propto \left(\frac{M_{Pl}^2H^2}{M^4}\right)\lambda\epsilon$ which, at best, can be of order 1 (in order not to spoil the higher derivative expansion). In the interaction Hamiltonian for the relevant perturbation variable ζ it leads to an additional term of the form

$$\Delta H_I = -\frac{\lambda H}{2M^4}\int d^3x a(\eta)\left(\frac{\dot\phi}{H}\right)^3\zeta'\left(\zeta'^2 - (\partial_i\zeta)^2\right). \tag{2.23}$$

Holman & Tolley (2008) first showed that the associated bispectrum correction due to an initial-state modification is a complicated function of the comoving momenta. Most importantly, compared to the result obtained for the standard slow-roll computation, after integrating over conformal time one now finds terms proportional to η_0 and η_0^2, in addition to contributions independent of η_0. The different powers of η_0 can, as before, be combined with one of the comoving momenta k_1 to give the large number $|k_1\eta_0| = M/H \gg 1$. Consequently, the amplitude of the three-point function is expected to be dominated by contributions proportional to η_0^2. Collecting the leading contributions and neglecting terms that are not (locally) enhanced at η_0^2 order, we obtain

$$\begin{aligned}\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle_{\rm nBD}^{\rm HD} \approx\ & (2\pi)^3\delta^{(3)}\left(\sum\vec{k}_i\right)\frac{\lambda}{M^4}\frac{1}{\prod(2k_i^3)}\frac{H^8}{\dot\phi^2}\sum_j 2\mathcal{R}e(\beta_{k_j}) \\ & \times\left[\left(\frac{1-\cos(\tilde k_j\eta_0)}{\tilde k_j^2} - \eta_0\frac{\sin(\tilde k_j\eta_0)}{\tilde k_j}\right)\mathcal{P}(k_j,k_{j+1},k_{j+2}) \right. \qquad (2.24)\\ & \left. +\eta_0^2\left(\cos(\tilde k_j\eta_0)\,\mathcal{Q}(k_j,k_{j+1},k_{j+2})\right)\right], \qquad (2.25)\end{aligned}$$

where

$$\begin{aligned}\mathcal{P}(k_j,k_{j+1},k_{j+2}) &= -4k_{j+1}k_{j+2}(k_{j+1}+k_{j+2})(k_{j+1}^2+k_{j+2}^2+k_{j+1}k_{j+2})\\ &\quad +2\times\prod_i k_i(k_t^2-4k_{j+1}k_{j+2}).\\ \mathcal{Q}(k_j,k_{j+1},k_{j+2}) &= \prod_i k_i(k_t^2-4k_{j+1}k_{j+2}).\end{aligned}$$

In the above expression $k_t=\sum_i k_i$ represents the sum of all (absolute values of) comoving momenta and $\tilde{k}_j$ is defined as before $\tilde{k}_j=k_t-2k_j$, i.e., the sum of comoving momenta with one of the signs reversed. The terms in eq. (2.24) are enhanced to an order η_0^2 only in the collinear limit, while being suppressed for all other triangular configurations. In contrast the term in eq. (2.25) is enhanced by η_0^2 over the full triangle domain and is therefore expected to be the dominant contribution. This was apparently not noticed by Holman & Tolley (2008), maybe because the collinear limit was assumed from the start. As a result the shape of the higher derivative bispectrum with a modified initial-state is not of the expected enfolded type. To be explicit let us rewrite the dominant overall enhanced contribution in terms of the two variables $x_2\equiv\frac{k_2}{k_1}$ and $x_3\equiv\frac{k_3}{k_1}$, where we scaled out the usual k_1^{-6} and absorbed the enhancement factor $(k_1\eta_0)^2$ into the non-Gaussian amplitude $f_{\rm NL}$, producing the following shape function

$$\begin{aligned}F^{\rm HD-dom}(k_1\eta_0,x_2,x_3) &= \frac{1}{x_2^2x_3^2}\left\{\cos(k_1\eta_0(x_2+x_3-1))\left[(1+x_2+x_3)^2-4x_2x_3\right]\right.\\ &\quad +\cos(k_1\eta_0(-x_2+x_3+1))\left[(1+x_2+x_3)^2-4x_3\right])\\ &\quad \left.+\cos(k_1\eta_0(x_2-x_3+1))\left[(1+x_2+x_3)^2-4x_2\right]\right\}. \qquad (2.26)\end{aligned}$$

As before, scale-invariance of the bispectrum therefore requires the combination $|k_1\eta_0|=M/H\gg 1$ to be k_1 independent. Note that the cosines appearing in this shape function imply that the three-point function is constantly changing sign. The norm of the full higher derivative modified initial-state distribution is well approximated using only the contribution described by the shape function $F^{\rm HD-dom}$. We find, after averaging over the cosine, that $|F^{\rm HDmodin}|^2\sim|F^{\rm HD-dom}|^2\approx 23.3$[4]. Since it is the normalization independent combination $f_{\rm NL}\,|F|$ that is actually being measured, we conclude that a perfect template would be sensitive to the full $(k_1\eta_0)^2$ enhancement factor. Because this number could be as large as 10^6 it indicates that the higher derivative terms are extremely sensitive to initial-state modifications, potentially leading to strong constraints on departures from the standard Bunch-Davies vacuum state. This derives from the fact that the nonlinear higher derivative interaction plays an important role at sub-horizon scales. Sub-horizon particle occupation numbers as a consequence of the modified initial-state allow for the generation of a significant non-Gaussian signal due to the crucial presence of the higher derivative interactions at that stage. This is different from the standard slow-roll situation where the required (gravitational) nonlinearities are far less important at sub-horizon scales.

[4]Remember that the overall $(k_1\eta_0)^2$ enhancement factor was absorbed into the non-Gaussian amplitude $f_{\rm NL}$, explaining why it does not show up in the norm.

Unfortunately though, due to the oscillating sign nature of the dominant contribution eq. (2.25) the currently available local and equilateral templates are extremely insensitive to this term, i.e., the scalar products between $F^{\rm HD-dom}$ and the equilateral and local templates are suppressed because of cancellations inside the scalar product integral. A quick inspection of the scalar product integral reveals it could scale as $1/(k_1\eta_0)$ times some oscillating function of $k_1\eta_0$, which would reduce the overall $(k_1\eta_0)^2$ level of enhancement by at least one power. It is for this reason that we have kept the locally enhanced terms of eq. (2.24), since these could give rise to contributions in the scalar product of similar order in $k_1\eta_0$. The sine term in eq. (2.24) is overall enhanced with one power of $k_1\eta_0$ and, based on the results in the previous section, the localized enhancement due to the single $\tilde{k}_j$ in the denominator is expected to disappear after calculating the scalar product with the local or equilateral template, neglecting possible logarithmic terms. The cosine term is locally enhanced by a factor $(k_1\eta_0)^2$ due to the squared $\tilde{k}_j$ dependence in the denominator. One of those enhancement factors is again expected to be lost after performing the scalar product integral with the local or equilateral templates.

We anticipate a linear $k_1\eta_0$ scaling at best (neglecting possible logarithmic terms) and we should keep track of all terms in eq. (2.24) and eq. (2.25) when evaluating the scalar product with the local or equilateral template. Unlike the previous section we were unable to perform the relevant integrals analytically and instead relied on a numerical approach, fitting the scalar product integral results for a large sample of $k_1\eta_0$ values to estimate the $k_1\eta_0$ dependence. The relevant shape function is identified in exactly the same way as the dominant contribution $F^{\rm HD-dom}$, except this time no overall factors of $k_1\eta_0$ are absorbed into the definition of the non-Gaussian amplitude

$$A \;=\; (2\pi)^4 \left(\frac{M_{Pl}^2 H^2}{M^4}\right) \frac{1}{2}\lambda\epsilon\,|\beta|\,\frac{\Delta_\Phi^2}{k_1^6}.$$

As before we assume that the Bogolyubov parameter $\mathcal{R}e(\beta_{k_j}) \sim |\beta|$. Having fixed the non-Gaussian amplitude and shape function ambiguities we find that the leading order behavior of the scalar product with the local template distribution is well approximated by

$$F^{\rm local}\cdot F^{\rm HDmodin} \approx (k_1\eta_0)\,(-72 + 10\log|k_1\eta_0|)\;, \tag{2.27}$$

confirming the general expectation on the order of magnitude of the result. We should point out that the relative minus sign between the different terms in eq. (2.24) is the source of the relative minus sign in the final result eq. (2.27). Rather unfortunately, the different coefficients conspire in such a way that the scalar product has a minimum and then crosses through zero in the domain of interest $100 \leq |k_1\eta_0| \leq 1000$. This implies significantly smaller fudge factors, for a small range of $|k_1\eta_0|$ values, than would be expected on the basis of scaling alone. Another consequence of this, confirmed by the numerical results, is that the unavoidable oscillatory contributions, that we neglected when fitting the numerical data to produce eq. (2.27), are bound to give rise to relatively large corrections in the $|k_1\eta_0|$ domain of interest. This makes the full structure of the scalar product rather complicated. Although the general trend is nicely described by a linear plus logarithmic scaling with $|k_1\eta_0|$, in the $|k_1\eta_0|$ domain of interest the actual value of the scalar product fluctuates and can deviate from the expected order of magnitude for some values of $|k_1\eta_0|$. Dividing the scalar product by the norm of the local

distribution (which was already computed in the previous section) one obtains the fudge factor

$$\begin{aligned}\Delta_F\left[F^{\rm HDmodin}, F^{\rm local}\right] &= \frac{F^{\rm HDmodin}(k_1\eta_0, x_2, x_3)\cdot F^{\rm local}(x_2, x_3)}{\left|F^{\rm local}(x_2, x_3)\right|^2} \\ &\approx 5.7\cdot 10^{-3}|k_1\eta_0|\left(-72+10\log|k_1\eta_0|\right)\,. \end{aligned} \tag{2.28}$$

We conclude that a $|k_1\eta_0|\log|k_1\eta_0|$ enhancement remains, which is almost one power of $k_1\eta_0$ less as compared to the optimal scenario. For the fudge factor, as was true for the scalar product, the same cautionary remarks apply. The above result describes the average trend and the detailed numerical results show that fluctuations can have a significant effect on the actual value of the fudge factor in the $|k_1\eta_0|$ domain of interest. As we will see, the results for the 2D fudge factor with the local template exhibit a similar complicated behavior as a function of $k_1\eta_0$, although the actual numbers for the fudge factor, due to the larger coefficients, are roughly one order of magnitude larger. For the equilateral template the final scaling result is the same, although somewhat surprisingly, it is the dominant contribution in eq. (2.25) that is solely responsible for the final result. As one can easily check, both terms in eq. (2.24) are in fact maximizing exactly at the line $x_2+x_3-1=0$, whereas the equilateral template is exactly vanishing at the line $x_2+x_3-1=0$. The result is a suppressed contribution to the scalar product which is negligible in the limit of large $|k_1\eta_0|$ as compared to the other contribution. Incidentally this observable enhancement is the same as reported by Holman & Tolley (2008), but the underlying reason is very different. It is a consequence of using a template that is far from optimal and it should be possible to achieve significantly higher sensitivity by constructing a more suitable template to analyze the data. In particular, the non-Gaussian signal described by eq. (2.24) and eq. (2.25) is not of the enfolded type and has strong oscillatory features, which might allow for a clear distinction from other non-Gaussian sources.

The generic appearance of at least a single factor of $|k_1\eta_0| = M/H$ in the fudge factor with respect to the local (or equilateral) template implies an enhancement possibly as large as 10^3, ignoring the fluctuations of the fudge factor as a function of $k_1\eta_0$. At the start of this section we mentioned that in the standard Bunch-Davies vacuum the higher derivative term would give rise to a maximal $f_{\rm NL}^{\rm equil}$ of order 1. Compared to the original higher derivative non-Gaussian amplitude, the modified initial-state amplitude eq. (2.27) introduces an additional suppression with the Bogolyubov parameter. On the other hand the fudge factor introduces a linear $|k_1\eta_0| \equiv M/H$ factor enhancing the original higher derivative non-Gaussian amplitude by $\beta\,(M/H)$ when probed with the local or equilateral template. The CMB two-point power spectrum constrains the Bogolyubov parameter already at the 10^{-2} level, so at best this would allow for a local or equilateral non-Gaussian amplitude of order 10 due to initial-state modifications, assuming $M/H \sim 10^3$. This might be detectable in the future, although there are many other sources for a local or equilateral non-Gaussian signal at that level. In sec. 2.6 we will confirm the same level of enhancement by computing the projected 2D fudge factor and use the most recent WMAP constraints on local type non-Gaussianities to derive an order of magnitude constraint on the Bogolyubov parameter.

2.5 An Enfolded Template Proposal

In the absence of higher derivative corrections we have shown that a non-Gaussian signal due to a modified initial-state, which maximizes in collinear triangles, cannot be probed using the available local and equilateral templates. Both templates are not sensitive enough to the localized enfolded enhancement to give rise to a significant (preferably power law) dependence of the fudge factor for large $|k_1\eta_0| = M/H$. Instead the local and equilateral fudge factors are at best logarithmically dependent on $|k_1\eta_0|$. To see if one can improve on that situation one would like to introduce a more suitable template, one better aligned with the theoretical prediction of modified initial-state non-Gaussianities.

The distinguishing feature of such a template should be that it maximizes in the enfolded (or squashed) triangle limit, as opposed to the local and equilateral templates which maximize in squeezed and equilateral triangles respectively. Whereas the squeezed triangle is obtained by taking one of the comoving momenta to zero $k_i \to 0$, and equilateral corresponds to all momenta equal, enfolded triangles imply two collinear momenta, and therefore $k_i = k_j + k_m$ with $i \neq j \neq m$ and k_j and k_m representing the two collinear momenta. Clearly the squeezed, equilateral and squashed triangle limits exhaust all possibilities, which are shown in fig. 1.5, and nicely correspond to the three different classes of theoretical non-Gaussian predictions: local, equilateral and enfolded. So besides its potential theoretical relevance, also from the point of view of completeness it might be worthwhile to develop a third factorized template shape that would maximize in the enfolded triangle limit. This would introduce a third non-Gaussian observable $f_{\rm NL}^{\rm enf}$ measuring the enfolded amplitude.

The reason why one cannot directly compare theoretical predictions to the CMB data and needs especially designed templates lies in the computational complexity of reconstructing the non-Gaussian amplitude from the two-dimensional CMB temperature data. The projection of a 3-point correlator to a 3-point function in spherical harmonic space involves the Wigner $3j$ symbol (to construct the angular averaged bispectrum) and a complicated integral over transfer and Bessel functions. This is computationally very challenging, scaling as $N^{5/2}$; $N^{1/2}$ for every multipole l and N for the averaging over m, where N equals the total number of pixels in the CMB map. In the last few years different suggestions have been made to accomplish a reduction of computational time (Komatsu et al. 2003; Babich et al. 2004a; Liguori et al. 2006; Kogo & Komatsu 2006; Creminelli et al. 2007b; Yadav et al. 2007a; Liguori et al. 2007; Yadav et al. 2007b). A significant reduction in the number of calculations can be achieved if the three-point function is factorizable in its momentum dependence, i.e., $F(k_1, k_2, k_3) \to f_a(k_1)f_b(k_2)f_c(k_3)$, leading to a reduction from $N^{5/2}$ to $N^{3/2}$. As it turns out, local type non-Gaussianities are indeed described by a factorized shape function $F(k_1, k_2, k_3)$, whereas the theoretical predictions for non-Gaussianities of the equilateral and enfolded type are not factorizable. This makes the direct comparison of equilateral and enfolded type non-Gaussianities to the two-dimensional CMB data extremely difficult for the time being, although recently some progress has been made to allow for a more direct comparison of arbitrary signals (Fergusson & Shellard 2007, 2009; Smith & Zaldarriaga 2006). Up to now one instead relies on factorized template approximations to the theoretical signals. Not so long ago an equilateral template has been successfully identified and compared to the data (Komatsu et al. 2009; Creminelli et al. 2006, 2007a,b), but an enfolded observational template has not yet been

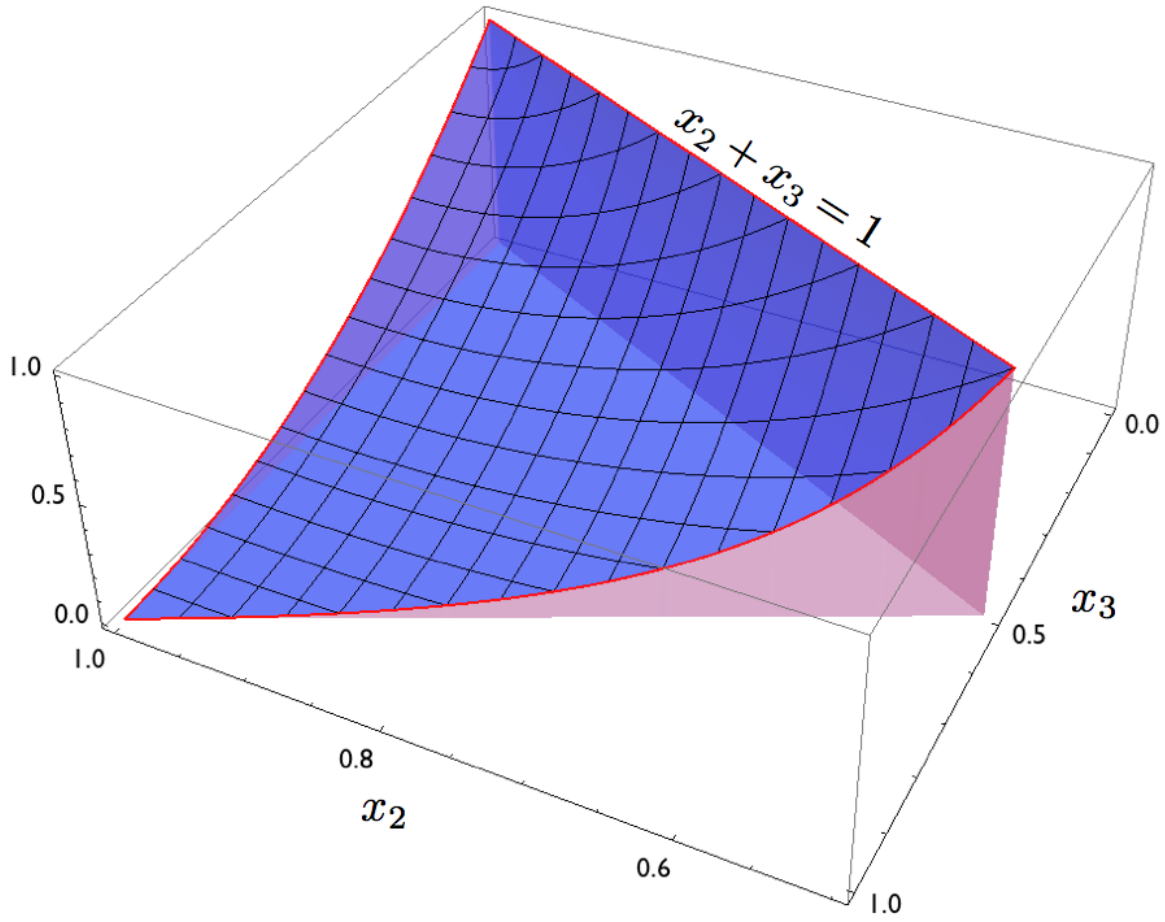

Figure 2.3: The enfolded template shape $F(x_2, x_3) \times x_2^2 x_3^2$.

constructed. Below we will construct a first proposal for a factorized enfolded template and analyze how much better it will be able to constrain the modified initial-state non-Gaussianities discussed in sec. 3.

Looking at eq. (2.13) it is clear that the three-point correlation function due to initial-state modifications is not factorizable. As explained this non-Gaussian shape function is the result of adding a minus sign to one of the comoving momenta to first order in the Bogolyubov parameter β_k (Seery & Lidsey 2005; Holman & Tolley 2008). For instance a term behaving as $1/(k_1 + k_2 + k_3)$ would change to $1/(-k_1 + k_2 + k_3)$ plus permutations. This suggests an approach where one starts with the factorized equilateral template eq.(2.9) and just replaces $k_i \to -k_i$, symmetrizing over all the indices. Applying this idea produces the shape function $F(k_1, k_2, k_3) = -F_{\rm equil}(k_1, k_2, k_3)$, which does not yet resemble the desired enfolded distribution nor does it add additional information, i.e. it is simply the equilateral shape multiplied by a minus sign. Fortunately though, it requires only a small modification to come up with a factorized shape function that seems to be a reasonable candidate for an enfolded template.

Starting from the equilateral shape function, replacing $k \to -k$, introducing $x_2 = k_2/k_1$ and $x_3 = k_3/k_1$ and plotting the obtained distribution $F(x_2, x_3)$ times the appropriate measure factor[5] $x_2^2\, x_3^2$, it becomes apparent that a term proportional to $1/k_1^2 k_2^2 k_3^2$ acts as a kind of constant 'normalization' of the template. Additional $1/k_1^2 k_2^2 k_3^2$ terms therefore simply *lift* or *lower* the whole graph. By adjusting the number of such terms, so $F^{\rm enf} = -F^{\rm equil} + c/k_1^2 k_2^2 k_3^2$, it is possible to lift the obtained shape in such a way that it resembles an enfolded type distribution, maximizing on the line $k_2 + k_3 \simeq k_1$, corresponding to enfolded triangles. We find that the best choice requires adding only one such term to the $-F_{\rm equil}$ distribution, i.e., $c = 1$ (see app. B). Consequently our proposal for the factorized enfolded template, as a function of x_2, x_3, becomes

$$F^{\rm enf}(x_2, x_3) \;=\; 6\left[\frac{1}{x_2^3} + 2\ \text{perm} + \frac{3}{x_2^2 x_3^2} - \left(\frac{1}{x_2^2 x_3^3} + 5\ \text{perm}\right)\right]. \tag{2.29}$$

We have plotted the template shape function in fig. 2.3. In app. B we explain in what sense $c = 1$ corresponds to the optimal choice and the details are presented on how to translate this template into an observable using the 'fast best estimator' approach developed by (Komatsu et al. 2003, 2005; Creminelli et al. 2007b; Yadav et al. 2007a,b).

We should determine how well the proposed enfolded template overlaps with the theoretical modified initial-state three-point function of eq. (2.13). To quantify this we will perform the same analysis as in section 3, calculating the scalar product, cosine and fudge factor, now using the enfolded template. The squared norm of the modified initial-state shape function is given by eq. (2.15) and a numerical integration gives $|F^{\rm enf}(x_2, x_3)|^2 = 4.34$. After computing the scalar product, this leads to the following expression for the cosine as a function of $|k_1\eta_0|$

$$\begin{aligned} \text{Cos}\left[F^{\rm modin}, F^{\rm enf}\right] \;&=\; \frac{F^{\rm modin}(k_1\eta_0, x_2, x_3)\cdot F^{\rm enf}(x_2, x_3)}{|F^{\rm modin}(k_1\eta_0, x_2, x_3)|\,|F^{\rm enf}(x_2, x_3)|} \\ &=\; 4.80\cdot 10^{-1}\,\frac{(2.80 + \log|k_1\eta_0|)}{\sqrt{6.05 + \frac{\pi}{60}|k_1\eta_0| + \frac{5}{4}\log|k_1\eta_0|}}\,. \end{aligned} \tag{2.30}$$

We have plotted this function in fig. 2.1, including the cosine functions between the modified initial-state shape distribution and the local and equilateral templates. As should be clear from the plot the cosine between the enfolded template and the theoretical distribution is closer to one, but there is certainly room for improvement. As the parameter $|k_1\eta_0|$ grows, the enfolded template will depart more from the theoretically predicted modified initial-state shape. Nevertheless, the enfolded template has significantly higher overlap with the theoretical distribution than the local or equilateral templates, at least in the comoving (3-D) momentum space. Since the cosine is the relevant quantity that determines the relative improvement, comparing to the plots (see also table 1) for the local and equilateral cosines one concludes that a rough $35-45$ percent level of improvement should theoretically be achievable using the enfolded template. This is certainly not enough to derive interesting constraints for the theoretically predicted modified initial-state non-Gaussianities, as can be seen more directly by turning our

[5] This is the relevant quantity because of the measure in the scalar product of eq. (2.4)

Shape F_Y	3-D Cos ($F_X = F_{\rm loc}$)	3-D Cos ($F_X = F_{\rm eq.}$)	3-D Cos ($F_X = F_{\rm enf}$)
Local	1	0.41	0.68
Equil.	0.41	1	0.49
Enf.	0.68	0.49	1
HD	0.45	0.99	0.59
Mod	0.6 – 0.3	0.6 – 0.4	0.9 – 0.6

Table 2.1: The 3D Cosine

Shape F_Y	3-D Fudge ($F_X = F_{\rm loc}$)	3-D Fudge ($F_X = F_{\rm eq.}$)	3-D Fudge ($F_X = F_{\rm enf}$)
Local	1	1.94	4.29
Equil.	0.09	1	0.66
Enf.	0.11	0.36	1
HD	0.10	1.07	0.86
Mod	$0.11 + 0.017 \log \lvert k_1\eta_0 \rvert$	0.82	$0.65 + 0.23 \log \lvert k_1\eta_0 \rvert$

Table 2.2: The 3D Fudge Factors. Note that HD are the higher derivative contributions from (Babich et al. 2004a). We added these for completeness and to show consistency with previous results (Babich et al. 2004a).

attention to the fudge factor, which explicitly identifies the level of $k_1\eta_0$ enhancement. From the expression of the cosine it is straightforward to read off the fudge factor

$$\Delta_F \left[F^{\rm modin}, F^{\rm enf} \right] = 0.65 + 0.23 \log |k_1\eta_0| \tag{2.31}$$

which disappointingly implies that the enhancement, or the sensitivity, that would be achieved using the proposed enfolded template is still only logarithmic in $|k_1\eta_0|$, far removed from the maximally attainable level of $\sqrt{|k_1\eta_0|}$ enhancement. Comparing to the local and equilateral fudge factors the coefficients are bigger which means the enfolded fudge factor will be larger and changes considerably over the natural range of $|k_1\eta_0|$ $(10^2 - 10^3)$. This is however still far removed from the (power law) level of enhancement that one would need to derive interesting constraints on the Bogolyubov parameter from the bispectrum data.

In tables 2.1 and 2.2 we have collected the cosine and fudge factor results for the different templates with respect to each other and some important theoretical predictions (equilateral higher derivative and enfolded modified initial-state). Note that the modified initial-state entries typically depend on the (large) parameter $k_1\eta_0$, which for the cosines has been denoted by the range of possible values, whereas for the fudge factors we have chosen to explicitly write the function. Independent from the original theoretical motivation, one could argue that the enfolded template nicely completes a general analysis of non-Gaussian signals. Table 2.1 then shows how much complementary information each template would provide. Compare this to a decomposition of a general vector into a set of basis vectors. Ideally, one would prefer to come up with a set of orthogonal basis shapes. Instead, the local, equilateral and enfolded template are far from orthogonal, but each does provide complementary information that can be precisely quantified in terms of the different cosine values[6] listed in the table. Decomposing

[6]Note that since we are dealing with a function space, obviously a complete decomposition would formally require an infinite set of basis functions on the relevant triangle domain.

a general three-point signal in these template shapes might therefore still be useful, even though the enfolded template by itself is unable to probe modified initial-state non-Gaussianities.

2.6 Two-Dimensional Bispectrum Results

Differently from the 3-D case the CMB temperature anisotropies are a 2-D projection of the linearly evolved primordial curvature perturbation field, hence the 2-D bispectrum is the result of the convolution of the shape function with the photon transfer function projected on the sky. In the following we give a brief review of the basic formalism, then we will discuss the results of the CMB bispectrum computation.

Let us consider the standard spherical harmonic decomposition of the CMB temperature fluctuation along the direction $\hat{n}$ of the sky,

$$\frac{\Delta T}{T}(\hat{n}) = \sum_{l,m} a_l^m Y_l^m(\hat{n}), \tag{2.32}$$

the multipole coefficients a_l^m contain all statistical information about the temperature anisotropy field, and are the starting point to construct the various correlator functions. The angular bispectrum in multipole space is given by

$$B_{l_1 l_2 l_3}^{m_1 m_2 m_3} \equiv < a_{l_1}^{m_1} a_{l_2}^{m_2} a_{l_3}^{m_3} >, \tag{2.33}$$

and assuming rotational invariance, the angle-averaged bispectrum reads as

$$B_{l_1 l_2 l_3} = \sum_{m_1, m_2, m_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} < a_{l_1}^{m_1} a_{l_2}^{m_2} a_{l_3}^{m_3} > . \tag{2.34}$$

Substituting the expression of the multipole coefficients in terms of the photon transfer function and the primordial curvature perturbation, eq. (2.34) becomes

$$B_{l_1 l_2 l_3} = \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} b_{l_1 l_2 l_3}, \tag{2.35}$$

with $b_{l_1 l_2 l_3}$ the reduced bispectrum given by,

$$b_{l_1 l_2 l_3} = \left(\frac{2}{\pi}\right)^3 \int dx dk_1 dk_2 dk_3 (x k_1 k_2 k_3)^2 j_{l_1}(k_1 x) j_{l_2}(k_2 x) j_{l_3}(k_3 x) \times F(k_1, k_2, k_3) \Delta_{l_1}(k_1) \Delta_{l_2}(k_2) \Delta_{l_3}(k_3), \tag{2.36}$$

where $\Delta_{l_i}(k_i)$ is the photon transfer function and $j_{l_i}(k_i x)$ is the Bessel function (for a detailed derivation see (Komatsu & Spergel 2001)). As discussed by Babich et al. (2004a), the evaluation of the reduced bispectrum is computationally challenging, on the other hand in the flat-sky approximation the computation is simplified, since for example the integral over the Bessel function does not explicitly appear. In such a case one has:

$$b_{l_1 l_2 l_3} = \frac{(\tau_0 - \tau_R)^2}{(2\pi)^2} \int dk_1^z dk_2^z dk_3^z \delta(k_1^z + k_2^z + k_3^z) F(k_1', k_2', k_3') \tilde{\Delta}_{l_1}(k_1^z) \tilde{\Delta}_{l_1}(k_2^z) \tilde{\Delta}_{l_1}(k_3^z), \tag{2.37}$$

where τ_0 and τ_R are the conformal time today and at decoupling respectively, k_i^z is the component of the wave-vector orthogonal to the plane tangent to the last scattering surface, and $k' = \sqrt{(k^z)^2 + l^2/(\tau_0 - \tau_R)^2}$; the photon transfer function along the orthogonal direction is given by

$$\tilde{\Delta}_l(k^z) = \int \frac{d\tau}{(\tau_0 - \tau)^2} S(\sqrt{(k^z)^2 + l^2/(\tau_0 - \tau)^2}) e^{ik^z\tau}, \tag{2.38}$$

where $S(...)$ is the CMB source function. The presence of the delta-function ensures that the projected modes form closed triangles. Hereafter we assume a vanilla LCDM cosmology with model parameter values corresponding to the WMAP 5-years best-fit model. All cosmologically relevant quantities such as the source functions have been computed with the publicly available CMBFAST code (Seljak & Zaldarriaga 1996). We then have evaluated the reduced bispectrum for different shapes using eq. (2.37).

Following (Babich et al. 2004a) we introduce a scalar product

$$B_X \cdot B_Y = \sum_{l_1,l_2,l_3} \frac{B^X_{l_1 l_2 l_3} B^Y_{l_1 l_2 l_3}}{f_{l_1 l_2 l_3} C_{l_1} C_{l_2} C_{l_3}}, \tag{2.39}$$

where $f_{l_1 l_2 l_3}$ is a combinatorial factor which is 1 if all three multipoles are different, 2 if two of them are equal and 6 if all of them are equal; C_l is the angular CMB power spectrum which includes the experimental noise evaluated in the Gaussian approximation assuming WMAP experimental characteristics. The cosine reads as

$$\mathrm{Cos}^{2D}[B_X, B_Y] = \frac{B_X \cdot B_Y}{\sqrt{B_X \cdot B_X}\sqrt{B_Y \cdot B_Y}}, \tag{2.40}$$

and the fudge factor as

$$\Delta_F^{2D}[B_X, B_Y] = \frac{B_X \cdot B_{\mathrm{local}}}{B_{\mathrm{local}} \cdot B_{\mathrm{local}}}. \tag{2.41}$$

In principle we may expect that the non-Gaussian signal given by different triangular shapes on the CMB differs from that obtained in 3-D, since triangles of different shapes in 3-D can be projected into the same 2-D configuration. If this is the case then the values of the cosine factors should be shifted upwards. Evaluating the cosine and fudge factors between local and equilateral shapes we find $\mathrm{Cos}^{2\mathrm{D}}[F^{\mathrm{local}}, F^{\mathrm{equil}}] = 0.62$ and $\Delta_F^{2\mathrm{D}}[F^{\mathrm{local}}, F^{\mathrm{equil}}] = 0.13$ respectively, which is consistent with the results presented by Babich et al. (2004a). The cosine and fudge factors between the initial-state modification shape function and the local, equilateral and enfolded templates are shown in fig. 2.4 and 2.5 as a function of the $k_1\eta_0$ parameter respectively.

We may notice a trend similar to that inferred from the 3-D evaluation. In particular, the cosine decreases as a function of $k_1\eta_0$ for all three templates, whereas the fudge factors are constant for the local and equilateral case, and increasing for the enfolded template. The enfolded template has the largest overlap with the initial state modification shape, although not significantly better then the local one. Overall the cosine values are slightly larger than what we have found in the 3-D calculation. This is because different triangular configuration in 3-D can be degenerate in the 2-D, hence the projection tends to systematically increase the overlapping between different templates.

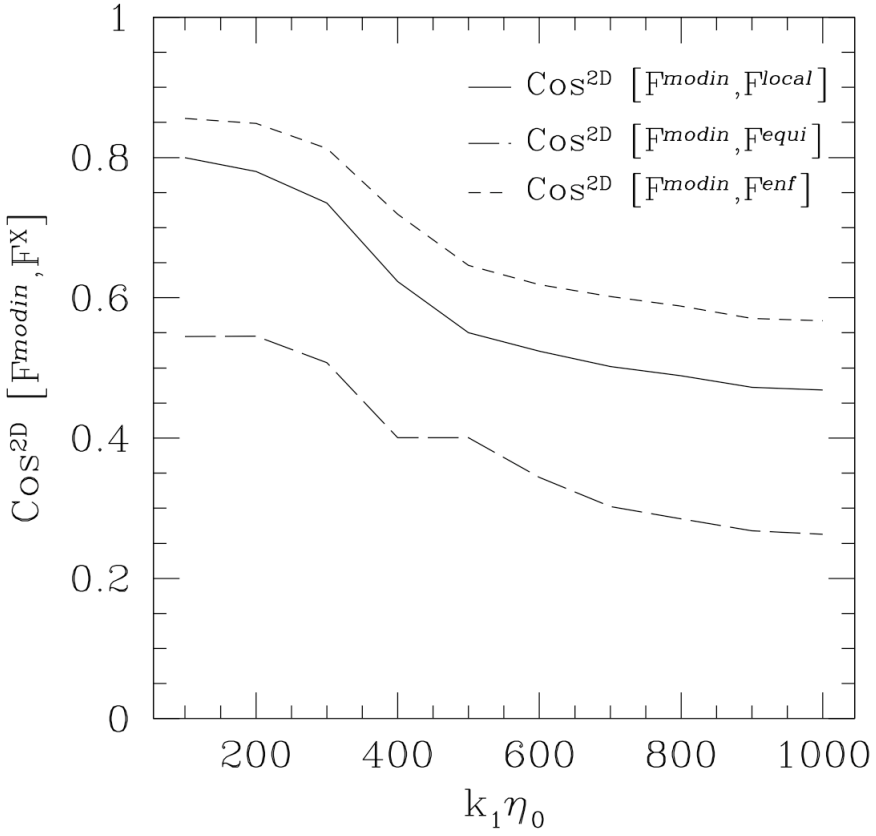

Figure 2.4: 2-D cosine factors between the initial state modification template and the local (solid line), equilateral (long dashed line) and enfolded (short dashed line).

We have also computed the reduced CMB bispectrum for the initial-state modification in the presence of higher order derivative terms as given by the shape function eq. (2.24) and eq. (2.25). We confirm the enhancement effect discussed in section 2.4, as an example calculating the fudge factor with the local template for $k_1\eta_0 = 100$ we find $|\Delta_F^{2D}| \approx 100$, whereas for $k_1\eta_0 = 10^3$ we obtain $|\Delta_F^{2D}| \approx 6000$. This clearly shows that the non-Gaussianity induced by initial state modifications is enhanced by the presence of higher derivative terms leading to potentially large detectable signals. The specific functional dependence of the fudge factor on $k_1\eta_0$ in the range of interests $(10^2 - 10^3)$ is far from trivial due to the interplay of the different terms in eq. (2.24) and eq. (2.25). These contain oscillatory factors that leads to a modulated oscillations of the fudge factor dependence on $k_1\eta_0$. Besides, the same term causes an oscillatory dependence of the reduced bispectrum as function of the multipoles. These oscillations are responsible for cancellations in the evaluation of the cosine factor, hence leading to a very small overlap with the other (non-oscillatory) templates. For example evaluating the cosine factor with the local shape for $k_1\eta_0$ in the range $10^2 - 10^3$ we find $|\mathrm{Cos}_F^{2D}| \approx 0.01$, which implies that current

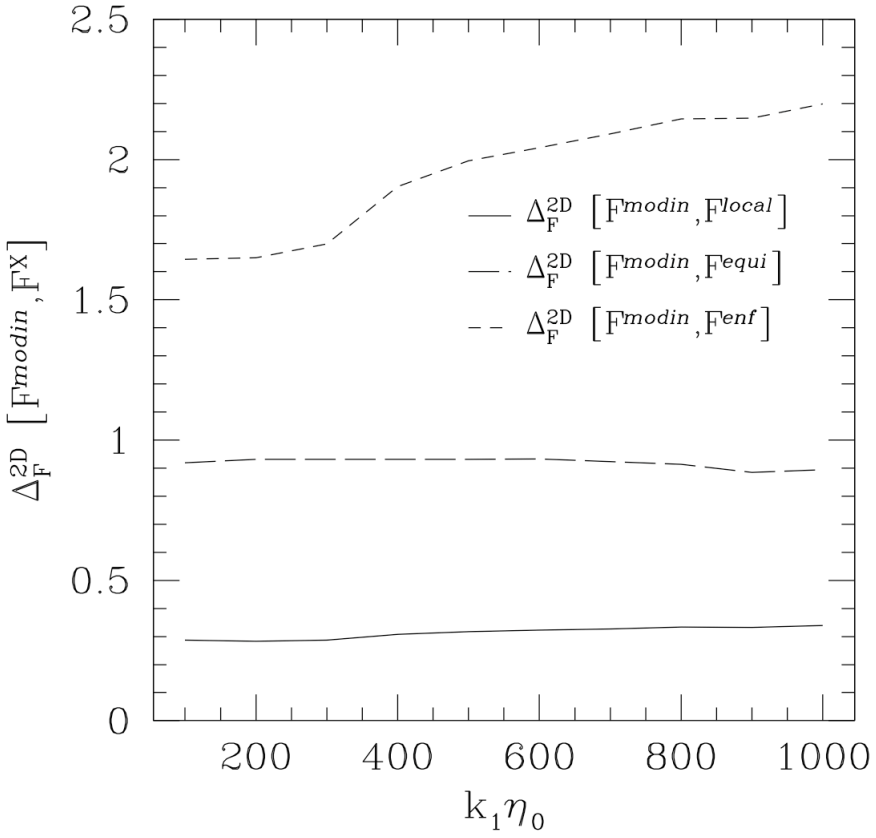

Figure 2.5: 2-D fudge factors between the initial state modification template and the local (solid line), equilateral (long dashed line) and enfolded (short dashed line).

observational templates are not apt to detect such a non-Gaussian signal. The determination of an optimal observational template that can account for the feature produced by this type of non-Gaussianity is therefore necessary and we leave this search to future work. Nevertheless using the above estimates of the fudge factor, we can use current constraints on local non-Gaussianity to infer limits on $f_{\rm NL}^{\rm HDmodin}$ and thus on the Bogolyubov parameter. As an example the prediction for a local non-Gaussian contribution due to a modified initial-state in the presence of a higher derivative operator reads

$$|\Delta f_{\rm NL}^{\rm local}| = \frac{5}{6}\epsilon|\beta|\left(\frac{M_{Pl}^2H^2}{M^4}\right)|\Delta_F^{\rm 2D}|\,, \tag{2.42}$$

where we assumed that the coefficient λ in eq.. (2.27) equals one. Using the observed amplitude of the power spectrum, the slow-roll parameter ϵ can be replaced with $\frac{10^{10}}{8\pi^2}\frac{H^2}{M_{Pl}^2}$. Assuming $M/H \sim 10^2$ this gives rise to the following constraint on the Bogolyubov parameter, using the

latest WMAP 5-years central value for $f_{\rm NL}^{\rm local}$ (51) and the result for the fudge factor $|\Delta_F^{\rm 2D}| \approx 100$

$$|\beta| \leq 0.5 \tag{2.43}$$

This is a constraint comparable to constraints coming from the power spectrum[7]. Note that the 95 CL lower limit of $f_{\rm NL}^{\rm local}$, tightens this constraint by a factor of 10. Most proposals of initial state modifications expect the Bogolyubov parameter $|\beta|$ to be a function of H/M. For instance in the New Physics Hypersurface scenario $|\beta|$ is predicted to be linear in H/M (Danielsson 2002; Easther et al. 2002) . Using the results above for $H/M \sim 10^{-2}$ this predicts at best an order 1 contribution to the local non-Gaussian signal. We should stress that an ideal template could improve the limits on the Bogolyubov parameter by another factor of M/H.

2.7 Conclusion

We have analyzed inflationary three-point correlators as a result of a small departure from the standard Bunch-Davies vacuum. In the simplest scenario where we avoided higher derivative interactions we confirmed that the initial-state modification causes the three-point correlator to maximize in the collinear or enfolded triangle limit, corresponding to a uniquely different shape as compared to local and equilateral non-Gaussian signals. Since the maximal signal scales linearly with the cutoff scale M, the non-Gaussian amplitude in enfolded triangles can be quite large, perhaps allowing for detection or providing interesting constraints on departures from the Bunch-Davies vacuum. However, by computing the scalar products, and consequently the cosine and fudge factors, between the theoretical prediction for the three-point function and the existing local and equilateral observational templates, we concluded that essentially all enhancement is lost due to the inefficiency of the available observational templates combined with the localized nature of the enhancement. Although (Holman & Tolley 2008) reached a similar conclusion, their argument was very different, relying on the projection to the two-dimensional CMB sphere. Instead, we have shown that the currently available method of comparing theoretical three-point functions to CMB bispectra, involving observational templates and the necessary integration over all triangles, already removes most sensitivity to localized enhancements in enfolded triangles, even before projecting to the two-dimensional CMB sphere.

The situation can in principle be improved by constructing a suitable enfolded template (better) adapted to the theoretical prediction. Moreover, from a general non-Gaussian analysis point of view, the introduction of an enfolded template might be interesting in itself, potentially providing complementary information in addition to the local and equilateral templates. The enfolded template proposed here was unfortunately only a marginal improvement over the local and equilateral template, still being insensitive to the localized enhancement. It would certainly be worthwhile to look for a more optimal enfolded template that could approach the theoretically maximum level of sensitivity to the localized enhancement, corresponding to a $\sqrt{|k_1\eta_0|} = \sqrt{M/H}$ dependence of the corresponding fudge factor.

After adding a specific higher derivative term we surprisingly found that the localized nature of the enhancement is substituted by an overall enhancement of $(k_1\eta_0)^2$ that can be absorbed

[7]The derivation of the constraint has been updated after publication, as we discovered a small mistake. We rectified this mistake in a subsequent publication (Meerburg & van der Schaar 2011).

directly into the non-Gaussian amplitude $f_{\rm NL}$. Even though (sub-leading) terms exist that are displaying a localized form of enhancement in the collinear limit, it turns out that the leading contribution is enhanced over the full comoving momentum triangle domain. In other words, the non-Gaussian signal is not of the enfolded type in this particular case and no enhancement sensitivity can be lost by the integration over all triangles. This dominant term to the three-point function is rapidly oscillating, causing the sign of the bispectrum to oscillate as well. This oscillating sign feature implies that the leading contribution to the scalar product with the currently available templates is severely suppressed and are not sensitive to the full $(k_1\eta_0)^2$ enhancement. Instead, subleading order $|k_1\eta_0|$ terms also contribute to the scalar product with the local or equilateral template. The end result is that the local and equilateral templates only probe a linear $k_1\eta_0$ enhancement. The details of the 2D and 3D fudge factor are complicated for the values of $|k_1\eta_0| = M/H$ of interest, but the numerically determined 2D fudge factor for $M/H = 10^2$ was used to put a constraint on the Bogolyubov parameter of order $\mathcal{O}(10^{-1})$, close to the bound derived from the two-point power spectrum. The main message however should be that improved templates, sensitive to the oscillatory nature of the dominant contribution to the bispectrum, would considerably tighten these constraints (as well as improved constraints on various $f_{\rm NL}$). The oscillatory nature of the signal in momentum space suggests that specific, perhaps observable, features could appear in 3D and 2D position space. In any case it would be worthwhile to generalize the range of available non-Gaussian shapes that can be compared to the data, including oscillatory signals, which we hope to report on in the future. Theoretically at least, for an optimal template, this would lead to a limit on vacuum modifications orders of magnitude stronger than the bound obtained from the two-point power spectrum, which would be quite remarkable.

One important general conclusion supported by our results is that higher derivative corrections, which on general grounds are always expected to be present, are extremely sensitive to departures from the standard Bunch-Davies vacuum state. Throughout this chapter we assumed that the combination $|k_1\eta_0|$ is independent of the actual comoving momenta involved and equal to M/H, in the spirit of the New Physics Hypersurface approach to vacuum state modifications. The reason for this was scale invariance of the bispectrum, which we relied on to allow for comparison with the available (scale-invariant) template shapes. Fixing η_0 instead, as one would do in a Boundary Effective Field Theory approach to vacuum state modifications, immediately results in a scale-dependent bispectrum. It would be interesting to study such scale-dependent scenarios, requiring more general analysis tools (Fergusson & Shellard 2009), and determine to what extent (future) analysis of 3D large scale structure or 2D CMB data can constrain bispectrum departures from scale-invariance. As reported, the bispectrum or three-point function is extremely sensitive to initial-state modifications in the presence of a higher derivative operator, and there is no reason to think this could not similarly be true for all higher n-point functions. A more general perturbative analysis, including higher n-point functions, might lead to a hierarchy of (theoretical) constraints on vacuum state modifications, perhaps pointing to the standard Bunch-Davies state as the only consistent possibility in practice. We hope that ongoing future work in this direction can further help us understand and identify the phenomenological and theoretical constraints on the vacuum state ambiguity.

Acknowledgments

We thank Benjamin Wandelt, Andrew Tolley and Paolo Creminelli for very useful discussions. PDM was supported by the Netherlands Organization for Scientific Research (NWO), NWO-toptalent grant 021.001.040. The research of JPvdS is financially supported by Foundation of Fundamental Research on Matter (FOM) grant 06PR2510. PDM and JPvdS were supported in part by a van Gogh Collaboration Grant VGP 63-254 from the Netherlands Organisation for Scientific Research (NWO) and PSC was supported by the van Gogh program grant N-18150RG of the "Ministère des Affaires Etrangères et Européennes".

Chapter 3

Bispectrum Signatures of a Modified Vacuum in Single-Field Inflation with a Small Speed of Sound

P. Daniel Meerburg, Jan Pieter van der Schaar & Mark Jackson
JCAP **1002** (2010) 001

Abstract: Deviations from the Bunch-Davies vacuum during an inflationary period can leave a testable imprint on the higher-order correlations of the CMB and large scale structures in the Universe. The effect is particularly pronounced if the statistical non-Gaussianity is inherently large, such as in models of inflation with a small speed of sound, e.g. DBI. First reviewing the motivations for a modified vacuum, we calculate the non-Gaussianity for a general action with a small speed of sound. The shape of its bispectrum is found to most resemble the 'orthogonal' or 'local' templates depending on the phase of the Bogolyubov parameter. In particular, for DBI models of inflation the bispectrum can have a profound 'local' template feature, in contrast to previous results. Determining the projection into the observational templates allows us to derive constraints on the absolute value of the Bogolyubov parameter. In the small sound speed limit, the derived constraints are generally stronger than the constraint obtainable from the power spectrum. The bound on the absolute value of the Bogolyubov parameter ranges from the 10^{-6} to the 10^{-3} level for $H/\Lambda = 10^{-3}$, depending on the specific details of the model, the sound speed and the phase of the Bogolyubov parameter.

3.1 Introduction

Recent years have witnessed a steady increase in the amount of precision data from the cosmic microwave background and large scale structure (Spergel et al. 2003a, 2007; Komatsu et al. 2009). This opportunity has motivated the study of both the power spectrum (or two-point correlation) of the primordial density perturbations, essentially probing the free theory of the inflationary Lagrangian, as well as higher-order statistics, probing the interacting theory. The leading statistical probe of interactions is the bispectrum (or three-point correlation, indicating

non-Gaussian statistics) and there is currently a large collection of work on the non-Gaussian predictions for specific models of inflation (Creminelli 2003; Chen et al. 2007; Allen et al. 2004; Battefeld & Easther 2007; Huang et al. 2008; Langlois et al. 2008; Bartolo et al. 2004; Sasaki et al. 2004; Koyama et al. 2007; Buchbinder et al. 2008; Lehners & Steinhardt 2008; Arkani-Hamed et al. 2004; Moss & Xiong 2007; Bernardeau & Brunier 2007; Barnaby & Cline 2007a). Interestingly, any detection of non-Gaussianity will immediately rule out all models of single-field slow-roll inflation. While this would be a monumental discovery in itself, it would also be the starting point for further constraining the inflationary Lagrangian. Indeed, if the non-Gaussian amplitude is high enough to be observable it opens up a new, and very distinctive, window on the details of inflation (Komatsu et al. 2009).

In a previous article by two of us (PDM, JPvdS) (Meerburg et al. 2009) we computed the primordial bispectrum for a single-field slow-roll inflationary model, with the assumption that the initial state is not the Bunch-Davies (BD) vacuum (Bunch & Davies 1978). This assumption is well-motivated because inflation is by construction an effective theory and therefore there is no 'knowledge' of anything beyond a specified cutoff scale in energy. In expanding backgrounds this energy-cutoff is tantamount to a (possibly momentum-dependent) time-cutoff, allowing one to instead parameterize ignorance as an initial state from which to evolve the dynamics. As with all effective theories, the details underlying the construction of such a state are irrelevant; the only concern is to use this state to compute the consequences on the bispectrum, and if possible compare theoretical predictions with the data to place constraints on such a deviation.

As first reported by Holman & Tolley (2008), our results (Meerburg et al. 2009) confirmed that small initial state modifications can induce surprisingly large corrections to the primordial bispectrum, especially if one includes higher-derivative operators (see also (Collins & Holman 2009)). The intuitive explanation is essentially that any deviation from BD introduces particles on sub-horizon scales at the initial stages of inflation, and these particles then have ample time to produce a large non-Gaussian signal in the presence of (irrelevant) interactions. Unfortunately, this yielded only a rather weak bound of order $|\beta| < 10^{-2}$ on the deviations from the BD state in terms of the Bogolyubov β-parameter. This is because although the *magnitude* of such induced non-Gaussianity was large, its *shape* in comoving momentum space was found to substantially depart from standardized observational templates, providing poor overlap with existing constraints. This motivates the need for better templates and possible systematic analysis of oscillatory features in the bispectrum, for which we hope to report some progress in the future.

In this article we generalize the class of models for which initial state modifications are considered. We opine that initial state modifications are generally natural for these models. In particular, we will argue that for both branches in DBI inflation, UV (Alishahiha et al. 2004) and IR (Chen 2005b,c), small deviations can be expected or at least are worth considering when computing the full bispectrum. Our starting point will be the most general single-field action, for which the bispectrum has been previously computed by Seery & Lidsey (2005) and Chen et al. (2007). Specific single-field models (slow-roll, DBI, kinetic) are all limiting cases of this action, and we will be interested in the limit where the speed of sound c_s (that is, the speed at which inflaton perturbations evolve) is significantly reduced relative to c, the speed at which metric perturbations evolve. Whereas bispectrum components in standard slow-roll

models are typically proportional to ϵ and η (Maldacena 2003), in the limit of small c_s one finds new contributions to the bispectrum that are proportional to $1/c_s^2$. Assuming the BD-vacuum these potentially large corrections are typically of the equilateral type and have been previously constrained (Creminelli et al. 2007a; Senatore et al. 2009). When one allows for (initially Gaussian) vacuum state modifications, there are yet-additional components to the bispectrum which are also expected to be enhanced in the small speed of sound limit. As a consequence, we will be able to put relatively strong constraints on $|\beta|$. Importantly, we will also study the effect of an arbitrary phase in the Bogolyubov parameter, which has so far been neglected. Somewhat surprisingly, we will find that specific choices for the phase can have dramatic effects on the shape of the leading non-Gaussian contribution. As a consequence the phase of the complex Bogolyubov parameter β strongly affects the observational constraints on the absolute value $|\beta|$ that can be derived. An important restriction in our analysis is that we will have to assume approximate scale-invariance throughout and will not discuss (big or small) departures from scale invariance and how these can be observationally constrained (Khoury & Piazza 2008; LoVerde 2008; Sefusatti et al. 2009).

This article is organized as follows. In sec. 3.2 we will briefly summarize the small-c_s models of inflation and explain how initial state modifications are quite natural in this context. In sec. 3.3 we begin from the most general single-field inflationary action and then compute the corrections to the bispectrum. In sec. 3.4 we identify the leading-order and sub-leading components and then constrain modifications to the initial state. Constraints can only be set using already measured bispectra by comparing the theoretical components to observed templates. In order to achieve this, we compute the leakage via projection of these components into measured components, for which we make use of a formalism first proposed by Babich et al. (2004a) and optimized by Fergusson & Shellard (2009). In sec. 3.4.3 we use the two derived constraints on different $f_{\rm NL}$ to derive constraints on DBI models with three free parameters. In sec. 3.5 we summarize and conclude.

3.2 A Case for Vacuum State Modifications in Small Sound Speed Models

It has now become clear that vacuum state modifications cannot be excluded on purely theoretical grounds, even though they are constrained by bounds on backreaction (Greene et al. 2005; Anderson et al. 2005) and observations of the primordial power spectrum (Easther et al. 2005a; Greene et al. 2004; Easther et al. 2005b; Spergel et al. 2007; Hamann et al. 2008). The theoretical challenge is to relate potential vacuum state modifications to ultraviolet physics at the string- or Planck-scale. In terms of an effective field theory description of inflation such a relation would appear natural, since the details of the vacuum (or equivalently initial) state of inflation usually depends on physics far beyond the cutoff in the effective field theory description. This raises the hope of being able to probe string- or Planck-scale physics through the signatures of possible modifications to the vacuum state.

Unfortunately, so far the available proposals have been phenomenological in nature and can be separated into two classes. In the original work of Danielsson (Danielsson 2002) scale-invariance is preserved by proposing that the initial time η_0 (beyond which an effective field

theory theory breaks down) is identified as the time where the physical momentum equals some high energy cutoff, defining a New Physics Hypersurface (NPH). This immediately implies that the initial "cutoff" time will be different for different comoving momenta, i.e., η_0 will be a function of k allowing for the preservation of scale-invariance. A specific, but rather ad-hoc, prediction for the Bogolyubov parameter is then obtained by assuming a local empty state at the initial time $\eta_0(k)$ (Danielsson 2002; Easther et al. 2002).

In a different approach known as Boundary Effective Field Theory (BEFT) (Schalm et al. 2004), scale-invariance is explicitly broken by identifying an initial time surface η_0 where the usual BD boundary conditions are corrected by including leading corrections. As one would expect from effective field theory, the corrections to the BD state grow as a function of comoving momentum k, with the largest comoving momenta corresponding to physical momenta closest to the cutoff scale on the initial time surface. The BEFT approach has the advantage of allowing one to calculate corrections to the BD state systematically from first principles. However, the breaking of scale-invariance and the dependence upon the initial time η_0 are serious drawbacks.

In both scenarios the magnitude of the (leading) corrections, parameterized by the absolute value of the Bogolyubov parameter, is proportional to the inflationary Hubble parameter H divided by a physical cutoff scale Λ. The appearance of the Hubble scale is natural for various reasons and can in particular be related to the fact that an initial vacuum state can only be defined in the adiabatic regime, i.e., for physical momenta much bigger than the Hubble scale. This immediately suggests that any modification should be proportional to H/Λ, ensuring that the breakdown of a small (perturbative) initial state correction coincides with the breakdown of adiabatic behavior.

This brings us to our first general observation regarding inflationary models with a small speed of sound. As a direct consequence of the small speed of sound the adiabatic regime breaks down long before the Hubble scale is reached. Instead, the relevant scale in this case is $H^\star \equiv H/c_s \gg H$, whose inverse corresponds to the radius of the sound horizon. Consequently, a vacuum state modification is expected to be proportional to $H^\star$ in any small speed of sound scenario. Therefore irrespective of a particular motivation or method to describe vacuum state modifications in models with a small speed of sound, its magnitude is anticipated to be enhanced by a factor of $1/c_s$ as compared to standard slow-roll inflationary models (for the same Hubble parameter). Taking into account the constraint on c_s from bispectrum data, which roughly corresponds to $c_s \geq 10^{-2}$ (Senatore et al. 2009), this could add two orders of magnitude to the observable range of H/Λ as compared to standard slow-roll inflation.

The identification of the sound horizon as the relevant scale, signaling the breakdown of the adiabatic approximation, also has a more fundamental consequence for the initial vacuum state. Generically, the speed of sound slowly evolves during inflation and the models can be separated in two classes: ones where the speed of sound decreases (as in UV DBI models) or models where the speed of sound increases (as in IR DBI models) towards the end of inflation. In the second case, as was first pointed out by Kinney & Tzirakis (2008), when $\epsilon > 1 - s$ (where ϵ is the usual slow-roll parameter and $s \equiv \dot{c}_s/H\,c_s$ measures the rate of change in the speed of sound) the fluctuation modes evolve from superhorizon to subhorizon, which excludes the possibility of an adiabatic initial state such as the Bunch-Davies state. Of course, this also rules out any phenomenological approaches to vacuum state modifications that require an (approximately)

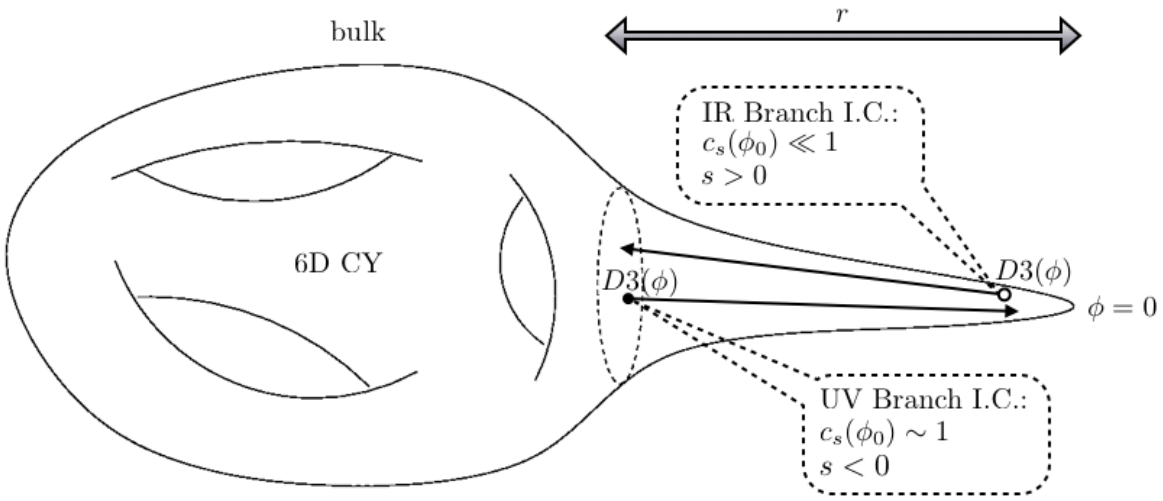

Figure 3.1: DBI inflation; IR and UV branches resulting in different evolution for c_s, leading to different motivations for changes in the Initial Conditions (I.C.)

adiabatic regime, such as the NPH and BEFT proposals. Instead the (different) initial vacuum state in this case should depend crucially on the physics before the inflationary phase. In IR models of DBI inflation this would correspond to the region near the tip of the throat where (strongly coupled) stringy physics is expected to become important.

In the opposite scenario, where the sound speed is decreasing towards the end of inflation, there is also an additional reason why the vacuum could be different from the standard BD vacuum (Kinney & Tzirakis 2008). In these models the speed of sound at the time of horizon crossing of the scales observable today should be large enough (roughly $c_s > 10^{-2}$) to avoid violating current bounds on non-Gaussianity. Depending on the values of the "slow-roll" parameter ϵ and the sound speed rate of change s, this requirement can imply that there are not enough e-folds left for the largest observable scales to allow for an approximately adiabatic vacuum choice. In other words, some modes in this case are not in the adiabatic regime when inflation began and are probing, just as in the previous case, the phase before inflation started. In the specific example of UV models of DBI inflation (whose simplest versions are in fact already ruled out), these modes would be probing the bulk region outside of the throat, which is governed by stringy physics (see fig. 3.1).

Finally, considering in particular DBI models of inflation, it is important to realize that the DBI action is obtained as the effective theory of open strings after heavy closed strings have been integrated out. As was pointed out by Burgess et al. (2003) integrating out heavy (closed string) states during an inflationary stage could give rise to modified initial conditions in the effective theory of the light (open string) degrees of freedom.

The arguments above provide a relatively strong case for considering departures from the standard BD vacuum in models with a small speed of sound, in addition to the generic effective field theory motivation. It should therefore be worthwhile to look for bispectrum signatures of

modified initial states in the limit of a small sound speed, especially as compared to canonical slow-roll inflation. In our analysis we will consider initial vacuum state modifications parameterized by an arbitrary complex Bogolyubov parameter, so no particular model for vacuum state modifications is assumed. As first reported by Holman & Tolley (2008), and corroborated by (Meerburg et al. 2009; Collins & Holman 2009), the bispectrum is significantly enhanced due to the presence of particles in the modified initial vacuum state. The important requirement of scale-invariance in order to do our analysis, in addition to the need for a (comoving dependent) cutoff scale to regulate the integrals after a change in the initial vacuum state, implies that the enhancement factor is expected to be proportional to a power of the ratio $\frac{\Lambda}{H^\star}$. This explains why the bispectrum can give rise to relatively strong constraints as compared to the power spectrum. Even though the appearance of $H^\star \equiv H/c_s$ implies that the enhancement is reduced (for fixed H) in the small sound speed limit, the specific powers of the enhancement factor that will appear, together with the expected overall enhancement of the non-Gaussian signal in the small sound speed limit, will result in a bispectrum whose magnitude is significantly larger as compared to the $c_s = 1$ case.

3.3 Corrections to the Bispectrum

3.3.1 General Single-Field Models of Inflation

We will be working in the context of a general (minimally coupled) 4d action

$$S = \frac{1}{2}\int d^4x\sqrt{-g}[M_p^2 R - 2P(X,\phi)], \tag{3.1}$$

where ϕ is the inflaton field and $X = g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi$. From this action we can obtain expressions for the usual slow-roll parameters, as well as a few parameters relevant for the three-point function first introduced by Chen et al. (2007) and Seery & Lidsey (2007), which are proportional to higher-order derivatives of $P(X,\phi)$:

$$\begin{aligned}
\epsilon &= \frac{XP_{,X}}{M_p^2H^2}, \\
c_s^2 &= \frac{P_{,X}}{P_{,X}+2XP_{,XX}} = \frac{M_p^2H^2\epsilon}{\Sigma}, \\
\Sigma &= XP_{,X}+2X^2P_{,XX}, \\
\lambda &= X^2P_{,XX}+\frac{2}{3}X^3P_{,XXX} = \frac{1}{3}\left(X\frac{\partial\Sigma}{\partial X}-\Sigma\right).
\end{aligned} \tag{3.2}$$

In a recent paper by Senatore et al. (2009) the description of single-field inflationary fluctuations was generalized further by focusing on the (broken) time-translation symmetry and writing down the most general effective Lagrangian of the corresponding Goldstone boson degree of freedom. In their description contributions to the three-point function are organized differently and known results derived in the standard formalism are only reproduced after appropriately translating the relevant variables. In particular they introduce a parameter $\tilde{c}_3$ relevant for the

(shape of the) three-point function, which is related to the above set of parameters as

$$\tilde{c}_3 = \frac{3}{2}c_s^2\left[\frac{2\lambda}{\Sigma(1-c_s^2)} - 1\right]. \tag{3.3}$$

As pointed out by Senatore et al. (2009), whereas the generic expectation for the shape of the three-point function in single-field models with a small speed of sound is equilateral, i.e., a maximal signal in equilateral triangles, at special values of $\tilde{c}_3$ (around -5) the non-Gaussian signal peaks in enfolded (or collinear) triangles instead. In fact, the orthogonal template put forward by Senatore et al. (2009) to analyze the CMB data is closely related to an earlier proposal by Meerburg et al. (2009) designed with initial state modifications in mind, which also tends to peak in enfolded triangles. It would be interesting to find a specific Lagrangian function $P(X,\phi)$ that can lead to these special values of $\tilde{c}_3$.This might be impossible. Subsequently we will base our three-point function calculations on the more "traditional" general single-field action described by a function $P(X,\phi)$, which can always be mapped to the corresponding Goldstone boson action using (3.3).

A specific example of much interest is the string-inspired model of DBI inflation (Alishahiha et al. 2004; Chen 2005b),

$$P(X,\phi) = -f(\phi)^{-1}\sqrt{1-2Xf(\phi)} + f(\phi)^{-1} - V(\phi).$$

In this case,

$$\begin{aligned}\Sigma &= \frac{H^2M_p^2\,\epsilon}{c_s^2},\\ \lambda &= \frac{H^2M_p^2\,\epsilon}{2c_s^4}(1-c_s^2).\end{aligned}$$

Substituting this into (3.3) one finds the corresponding $\tilde{c}_3 = \frac{3}{2}(1-c_s^2)$, which equals 3/2 in the small sound speed limit. This will cause the leading contribution in the three-point function to vanish for DBI models, which will be of importance when comparing the theoretical non-Gaussian amplitudes and shapes to the observational constraints.

In order to use the action given by (3.1) to calculate our quantity of interest, the three-point function, we need to determine the corresponding interaction Hamiltonian for the gauge-invariant perturbation variable. After first perturbing in the inflaton field $\phi = \phi_0 + \delta\phi$, we then convert this quantity to the gauge-invariant fluctuation $\zeta = -(H/\dot{\phi}_0)\delta\phi$. Expanding the action to third order in these perturbations, the effective interaction Hamiltonian associated with the cubic action is (Chen 2005c)

$$\begin{aligned}H_I(\eta) &= -\int d^3x\left[-a\left(\Sigma(1-\frac{1}{c_s^2})+2\lambda\right)\frac{(\dot{\zeta}_x)^3}{H^3} + \frac{a^2\epsilon}{c_s^4}(\epsilon-3+3c_s^2)\zeta_x\dot{\zeta}_x^2\right.\\ &\quad\left.+\frac{a^2\epsilon}{c_s^2}(\epsilon-2s+1-c_s^2)\zeta_x(\partial\zeta_x)^2 - 2\frac{\epsilon}{c_s^2}a\dot{\zeta}_x(\partial\zeta_x)(\partial\chi)\right],\end{aligned} \tag{3.4}$$

where overdots represent derivatives with respect to conformal time. This interaction Hamiltonian will serve as our starting point for the perturbative calculation of the three-point function in a modified initial state.

3.3.2 The Three-Point Function

We are interested in calculating the three-point correlation function $\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle$, where ζ_k is the gauge-invariant primordial density perturbation of comoving momentum k. To first order in the interaction Hamiltonian H_I this correlation can be written as an integral over the free-field correlator $\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}H_I(\eta)\rangle$,

$$\begin{aligned}\langle\psi(\eta)|\zeta_{k_1}(\eta)\zeta_{k_2}(\eta)\zeta_{k_3}(\eta)|\psi(\eta)\rangle &= \langle\psi(\eta_0)|\zeta_{k_1}(\eta)\zeta_{k_2}(\eta)\zeta_{k_3}(\eta)|\psi(\eta_0)\rangle \\ &\quad -i\int_{\eta_0}^{\eta}d\eta'\langle\psi(\eta_0)|[\zeta_{k_1}(\eta)\zeta_{k_2}(\eta)\zeta_{k_3}(\eta),H_I(\eta')]|\psi(\eta_0)\rangle+\mathcal{O}(H_I^2),\end{aligned} \tag{3.5}$$

where the time-evolved initial state is given by

$$|\psi(\eta)\rangle = Te^{-i\int_{\eta_0}^{\eta}H_I(\eta')d\eta'}|\psi(\eta_0)\rangle.$$

The first term on the right in eq. (3.5) will be zero if the initial state $|\psi(\eta_0)\rangle$ is Gaussian, as we will assume. The interaction Hamiltonian will encode all higher-order contributions in the effective action, so it will be sufficient to consider interactions to lowest order in H_I. Recall that if the initial state is assumed to be BD at the onset of inflation, the 3-point correlator at tree level is usually computed from $\eta_0 \to -\infty$. In our case we assume there is some part of the physics of which we are ignorant, ending up in the initial state at the start of inflation. Consequently we integrate from some initial time η_0 representing the initial time corresponding to the cutoff scale in physical momenta $p = k/a$. To stress this again, η_0 will therefore be k-dependent and lead to a scale-invariant spectrum, which we require for our analysis. The presence of such a cutoff must be explicit because initial state modifications would otherwise have infinite time to interact, and the 3-point correlator would then be infinitely large - and hence excluded by observation! In our case it is the time at which we chose to set the initial state to be Gaussian-distributed parameterized by an *a priori* arbitrary complex Bogolyubov parameter. Relating the initial time to the moment the physical momentum of a mode reaches the cutoff scale implies that all modes have equal time to evolve until they exit the horizon, which allows for the preservation of scale-invariance in the bispectrum.

One can expand the second line of the three-point correlator (3.5) into a product of Wightman functions $G_k^>$. These are defined as

$$\langle\zeta_{k_1}(\eta)\zeta_{k_2}(\eta')\rangle = (2\pi)^3\delta^{(3)}(\vec{k}_1+\vec{k}_2)G_{k_1}^>(\eta,\eta'). \tag{3.6}$$

The Wightman functions can be found by solving the classical equations of motion of the inflaton field action eq. (3.1) minimally coupled to gravity,

$$u_k(\eta) = \frac{iH}{\sqrt{4\epsilon c_s k^3}}(1+ikc_s\eta)e^{-ikc_s\eta}.$$

The quantized form of the dimensionless curvature perturbation is given by

$$\zeta_k(\eta) = u_k(\eta)a_k + u_{-k}^*(\eta)a_{-k}^\dagger,$$

where a_k and $a_k^\dagger$ are the annihilation and creation operators respectively and the mode-functions $u_k(\eta)$ are dimensionless by introducing the appropriate powers of the reduced Planck mass M_p. Using this form of $\zeta_k(\eta)$ one obtains the Wightman function

$$G_k^>(\eta,\eta') = \frac{H^2}{4M_p^2\epsilon c_s k^3}(1+ikc_s\eta)(1-ikc_s\eta')e^{-ikc_s(\eta-\eta')}.$$

We compute the late-time cases of interest to us,

$$G_k^>(0,\eta) = \frac{H^2}{4M_p^2\epsilon c_s k^3}(1-ikc_s\eta)e^{ikc_s\eta},$$

and

$$\partial_\eta G_k^>(0,\eta) = -\frac{H}{4M_p^2\epsilon k}\frac{c_s}{a(\eta)}e^{ikc_s\eta}, \tag{3.7}$$

with $a(\eta) = -1/\eta H$ during inflation in the usual assumption $\dot{H} \simeq 0$. To explicitly compute the correlation (3.5), a useful identity is

$$\begin{aligned}\langle\zeta_{k_1}(\eta)\zeta_{k_2}(\eta)\zeta_{k_3}(\eta)\rangle &\sim -i\int_{\eta_0}^{\eta} d\eta'\langle\psi(\eta_0)|[\zeta_{k_1}(\eta)\zeta_{k_2}(\eta)\zeta_{k_3}(\eta), H_I(\eta')]|\psi(\eta_0)\rangle \\ &= -2\mathcal{R}e\left[\int_{\eta_0}^{\eta} id\eta'\langle\psi(\eta_0)|\zeta_{k_1}(\eta)\zeta_{k_2}(\eta)\zeta_{k_3}(\eta)H_I(\eta')|\psi(\eta_0)\rangle\right], \end{aligned}\tag{3.8}$$

which removes the commutator and assures that we are dealing with a true observable. Next we can express $\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle$ in terms of these Wightman functions. For illustrational purposes let us just consider the first term on the right in eq. (3.4) in the absence of initial state modifications. The three-point function at late times $\eta \to 0$ is then

$$\begin{aligned}\langle\zeta_{k_1}(0)\zeta_{k_2}(0)\zeta_{k_3}(0)\rangle_{\rm BD} &= -12(2\pi)^3\delta^{(3)}(\sum\vec{k}_i)\left[\Sigma\left(1-\frac{1}{c_s^2}\right)+2\lambda\right]\frac{1}{H^3}\times \\ &\quad \mathcal{R}e\left[i\int_{\eta_0}^{0} d\eta'\ a\partial_{\eta'}G_{k_1}^>(0,\eta')\partial_{\eta'}G_{k_2}^>(0,\eta')\partial_{\eta'}G_{k_3}^>(0,\eta')\right].\end{aligned}$$

The factor 12 comes from 2 times the real part and the 6 possible permutations, and we have added the subscript to indicate that this is the result for the Bunch-Davies vacuum. From here on we will omit the $\eta = 0$ dependence of the correlation function, since we are only computing it at late times. Using the Wightman functions (3.7) we obtain

$$\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle_{\rm BD} = \frac{12}{8}(2\pi)^3\delta^{(3)}(\sum\vec{k}_i)\left[\Sigma\left(1-\frac{1}{c_s^2}\right)+2\lambda\right]\frac{c_s^3}{8M_p^6e^3}\mathcal{R}e\left[i\int_{\eta_0}^{0}\frac{d\eta'}{a^2}\frac{1}{k_1k_2k_3}e^{ik_tc_s\eta'}\right],$$

where $k_t = k_1+k_2+k_3$. In the limit $\eta_0 \to -\infty$, the (analytically continued) integral becomes

$$\frac{H^2}{k_1k_2k_3}\int_{-\infty}^{0} d\eta'\ \eta'^2e^{ik_tc_s\eta'} = -\frac{H^2}{k_t^3k_1k_2k_3}\frac{2}{ic_s^3}.$$

Substituting this into eq. (3.9) we find

$$\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle_{BD} = -3(2\pi)^3\delta^{(3)}(\sum\vec{k}_i)\left[\Sigma\left(1-\frac{1}{c_s^2}\right)+2\lambda\right]\frac{H^2}{8M_p^6\epsilon^3}\frac{1}{k_1k_2k_3k_t^3}. \quad (3.9)$$

To relate this to the familiar result found by Chen et al. (2007) one uses the fact that $\Sigma = \epsilon M_p^2 H^2/c_s^2$, leading to the following expression:

$$\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle_{\rm BD} = -\frac{3}{8}(2\pi)^3\frac{H^4}{M_p^4\epsilon^2c_s^2}\left(1-\frac{1}{c_s^2}+\frac{2\lambda}{\Sigma}\right)\delta^{(3)}(\sum\vec{k}_i)\frac{1}{k_1k_2k_3k_t^3}. \quad (3.10)$$

As is well-known, the momentum dependence of this contribution to the three-point function is well-approximated by the equilateral template, i.e., its magnitude maximizes in equilateral triangles. The full three-point function obviously contains additional contributions from the other terms in the interaction Hamiltonian. As mentioned earlier, in DBI models the above contribution vanishes exactly and another leading order equilateral term dominates (in the limit of slow-roll and small sound speed). All other contributions turn out to be higher order in the slow-roll expansion (Chen et al. 2007).

3.3.3 Bispectrum Corrections due to a Modified Vacuum State

Next consider the following Bogolyubov transformation of the initial state,

$$u_k \rightarrow \alpha_k v_k + \beta_k v_k^*, \quad (3.11)$$

with $|\alpha_k|^2 - |\beta_k|^2 = 1$. We will be interested in small departures from the usual BD vacuum, so $|\beta_k|/|\alpha_k| \ll 1$, as required by power spectrum constraints. For the BD vacuum ($\alpha_k \equiv 1$), our results will of course reduce to (3.9). Now let us suppose $\beta_k \neq 0$ and compute the corresponding bispectrum to lowest order in β_k. This will result in one of the modes effectively having negative energy and will carry along a corresponding factor of β_k. As emphasized previously, we will have to introduce an initial time cutoff since we expect particle interaction before the end of inflation to dominate the bispectrum, which could become arbitrarily large if the earliest modes have infinite time to grow (before they cross the horizon). To preserve scale-invariance, we introduce a fixed physical cutoff scale and determine the initial time as $kc_s\eta_0 = \Lambda/H^*$. Of course, scale-invariance also requires that the complex Bogolyubov parameter depends at most only weakly upon the comoving momentum k. The starting point will therefore be a constant complex parameter $\beta \equiv |\beta|\exp(i\delta)$ that parameterizes the departure from the BD state. Under these assumptions, the correction to the three-point correlation function eq.(3.10) is

$$\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle_{\rm nBD} = \frac{4}{8}(2\pi)^3\delta^{(3)}(\sum\vec{k}_i)\left[\Sigma\left(1-\frac{1}{c_s^2}\right)+2\lambda\right]\frac{H^6}{\dot{\phi}_0^6}c_s^3\times$$
$$\mathcal{R}e\left[\sum_j^3 i\beta_{k_j}\int_{\eta_0}^0\frac{d\eta'}{a^2}\frac{1}{k_1k_2k_3}e^{i\tilde{k}_jc_s\eta'}\right], \quad (3.12)$$

where $\tilde{k}_j = k_t - 2k_j$. The integral is evaluated to be equal to

$$\frac{H^2}{k_1 k_2 k_3} \sum_j \int_{\eta_0}^0 d\eta' \; \eta'^2 e^{i\tilde{k}_j c_s \eta'} = \frac{H^2}{k_1 k_2 k_3} \sum_j \left[i \frac{2}{\tilde{k}_j^3 c_s^3} + i \frac{e^{i\tilde{k}_j c_s \eta_0}}{\tilde{k}_j c_s} \left(\eta_0^2 + i \frac{2\eta_0}{\tilde{k}_j c_s} - \frac{2}{\tilde{k}_j^2 c_s^2} \right) \right]. \quad (3.13)$$

The first term is similar to our previous result, with the substitution of $\tilde{k}_j \leftrightarrow k_t$. The remaining terms are a result of the presence of a cutoff. The final result is

$$\begin{aligned} \langle \zeta_{k_1} \zeta_{k_2} \zeta_{k_3} \rangle_{\mathrm{nBD(1)}} &= -\frac{1}{4}(2\pi)^3 \frac{H^4 |\beta|}{\epsilon^2 c_s^4} \left(c_s^2 - 1 + \frac{2\lambda c_s^2}{\Sigma} \right) \delta^{(3)}(\sum \vec{k}_i) \frac{1}{k_1 k_2 k_3} \times \quad (3.14) \\ & \sum_j \left[\frac{c_s^2 \eta_0^2}{2} \frac{\cos(\tilde{k}_j c_s \eta_0 + \delta)}{\tilde{k}_j} - c_s \eta_0 \frac{\sin(\tilde{k}_j c_s \eta_0 + \delta)}{\tilde{k}_j^2} + \frac{\cos\delta - \cos(\tilde{k}_j c_s \eta_0 + \delta)}{\tilde{k}_j^3} \right]. \end{aligned}$$

Let us pause for a moment to interpret this expression. First, we notice that all terms contain a phase shift δ as a result of the phase in the Bogolyubov parameter. As will become clear this phase significantly influences the projection onto the observational templates and therefore the constraints on the vacuum state modification that can be derived.

Secondly, we notice that in the limit $\tilde{k}_j \to 0$ (and unlike previous results (Meerburg et al. 2009), ch. 2) all terms scale as $1/\tilde{k}_j$, which might cause some concern when considering the limit $\tilde{k}_j \to 0$. We certainly expect the result to be finite for any phase δ and any combination of the three comoving momenta and fortunately explicit computation of the limit $\tilde{k}_j \to 0$ in the app. C confirms this expectation. The first maximum occurs for $\tilde{k}_j \propto 1/c_s\eta_0$ and since $kc_s\eta_0 \gg 1$, where k would be the largest momentum scale under consideration, this effectively corresponds to the enfolded limit $\tilde{k}_j/k \sim 0$. Also note that this localized form of enhancement would scale as $(kc_s\eta_0)^3$, which is expected to be very large. Projection onto observational templates generically results in the loss of at most one factor of $kc_s\eta_0$, which would then provide us with a quadratic enhancement in the large number $kc_s\eta_0$. A detailed analysis of the projections and constraints will be discussed in sec. 4.

After this rather explicit introduction and discussion of the corrected three-point function due to the first term in the interaction Hamiltonian we would now like to present the complete (leading order) result for the corrected three-point function, for which the detailed calculations are presented in app. C. Using the expression for the power spectrum of general single-field inflation

$$P_\zeta = \frac{1}{8\pi^2} \frac{H^2}{c_s \epsilon}, \quad (3.15)$$

we can write the leading order behavior, to first order in β, as follows

$$\begin{aligned}\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle &= (2\pi)^7|\beta]P_\zeta^2\delta(\sum\vec{k}_i)\frac{1}{k_1^3k_2^3k_3^3}\times \\ &\left[\sum_j\left(\frac{1}{c_s^2}-1+\frac{2\lambda}{\Sigma}\right)(k_1c_s\eta_0)^3\mathcal{B}_{k_j}^{(1)}+(k_1c_s\eta_0)^2\left(\left(\frac{1}{c_s^2}-1\right)\left(\mathcal{B}_{k_j}^{(2)}+\mathcal{B}_{k_j}^{(3)}\right)+\right.\right. \\ &+\frac{\epsilon}{c_s^2}\left(\mathcal{B}_{k_j}^{(2)}+\mathcal{B}_{k_j}^{(3)}-\mathcal{B}_{k_j}^{(4)}\right)-\frac{2s}{c_s^2}\mathcal{B}_{k_j}^{(3)}\Big)+ \\ &\left.k_1c_s\eta_0\left(\frac{-2s+1-c_s^2}{c_s^2}\mathcal{B}_{k_j}^{(5)}+\frac{\epsilon}{c_s^2}\left(\mathcal{B}_{k_j}^{(5)}-\mathcal{B}_{k_j}^{(6)}\right)\right)+...\right],\end{aligned} \tag{3.16}$$

where the ... represent terms further suppressed slow-roll and in the cutoff η_0. Here the different functions $\mathcal{B}$ equal

$$\mathcal{B}_{k_j}^{(1)} = \frac{k_2^2k_3^2}{k_1}\left[\frac{1}{2}S_1\left(1-\frac{2}{(\tilde{k}_jc_s\eta_0)^2}\right)-\frac{S_2}{\tilde{k}_jc_s\eta_0}+\frac{\cos\delta}{(\tilde{k}_jc_s\eta_0)^3}\right], \tag{3.17}$$

$$\begin{aligned}\mathcal{B}_{k_j}^{(2)} &= \frac{3}{4}k_2^2k_3^2\left[\left(\frac{\tilde{k}_j-k_j}{k_j^2}+\frac{\tilde{k}_j+k_{j+1}}{k_{j+1}^2}+\frac{\tilde{k}_j+k_{j+2}}{k_{j+2}^2}\right)\left(\frac{\cos\delta}{(\tilde{k}_jc_s\eta_0)^2}-\frac{S_1}{\tilde{k}_jc_s\eta_0}\right)\right. \\ &\left.+\left(\frac{1}{k_j}-\frac{1}{k_{j+1}}-\frac{1}{k_{j+2}}\right)S_2\right],\end{aligned} \tag{3.18}$$

$$\begin{aligned}\mathcal{B}_{k_j}^{(3)} &= \frac{1}{8}\left[\frac{k_2k_3}{k_1}(2k_j^2-k_{j+1}^2-k_{j+2}^2)S_2-\right. \\ &\left.\frac{1}{k_1^2}\left(k_1k_2k_3+\tilde{k}_j(k_jk_{j+1}+k_jk_{j+2}-k_{j+1}k_{j+2})\right)\left(\frac{\cos\delta}{(\tilde{k}_jc_s\eta_0)^2}-\frac{S_1}{\tilde{k}_jc_s\eta_0}\right)\right],\end{aligned} \tag{3.19}$$

$$\begin{aligned}\mathcal{B}_{k_j}^{(4)} &= \frac{1}{4}\frac{1}{k_1^2}\left(k_j^4(k_{j+1}+k_{j+2})+k_{j+1}^4(k_{j+2}-k_j)+k_{j+2}^4(k_{j+1}-k_j)+\right. \\ &+k_j^2k_{j+2}^2(k_j-2k_{j+1}-k_{j+2})+k_{j+1}^2k_{j+2}^2(2k_j-k_{j+1}-k_{j+2})+ \\ &\left.k_j^2k_{j+1}^2(k_j-k_{j+1}-2k_{j+2})\right)\left(\frac{\cos\delta}{(\tilde{k}_jc_s\eta_0)^2}-\frac{S_1}{\tilde{k}_jc_s\eta_0}-S_2\right),\end{aligned} \tag{3.20}$$

$$\mathcal{B}_{k_j}^{(5)} = \frac{3}{8}\frac{(k_1^2+k_2^2+k_3^2)}{k_1}(k_jk_{j+1}+k_jk_{j+2}-k_{j+1}k_{j+2})S_1, \tag{3.21}$$

$$\mathcal{B}_{k_j}^{(6)} = \frac{1}{4}\frac{1}{k_1}\left(\frac{1}{2}(k_1^4+k_2^4+k_3^4)-(k_1^2k_2^2+k_2^2k_3^2+k_3^2k_1^2)\right)S_1, \tag{3.22}$$

and

$$\begin{aligned}S_1(\tilde{k}_jc_s\eta_0,\delta) &= \frac{\cos(\tilde{k}_jc_s\eta_0+\delta)}{\tilde{k}_jc_s\eta_0}, \\ S_2(\tilde{k}_jc_s\eta_0,\delta) &= \frac{\sin(\tilde{k}_jc_s\eta_0+\delta)}{\tilde{k}_jc_s\eta_0}.\end{aligned}$$

We have organized the bispectrum into pieces with different (decreasing) powers of $k_1 c_s \eta_0$, i.e., the ratio between the Hubble radius and the largest physical momentum scale at the cutoff time η_0, i.e., $k_1 c_s \eta_0 = (k_1/a_0)/(H/c_s)$ which can be as large as 10^3. For our purposes only the first two terms will be relevant as these are both enhanced in $k_1 c_s \eta_0$. From here on we will refer to these terms as the leading (proportional to $(k_1 c_s \eta_0)^3$) and the subleading (proportional to $(k_1 c_s \eta_0)^2$) term, respectively.

On first impression it seems that the momentum-dependence, or the shape, of the different terms in the above bispectrum strongly deviate from the constrained local, equilateral and orthogonal templates. Most importantly, all terms rapidly oscillate. Furthermore, as mentioned already, the leading-order term maximizes in the collinear limit. In the previous chapter we proposed a template for collinear triangles (Meerburg et al. 2009), which unfortunately is only a marginal improvement as compared to the local and equilateral template. Ideally, one would like to introduce a template that can (also) incorporate oscillations with an *a priori* undetermined frequency, introducing an additional parameter. It remains to be seen whether oscillations in the momentum dependence of the bispectrum are observable at all, as the finite resolution of the data might make determination of this additional frequency parameter very difficult. We will not be concerned with this problem here since we will only be interested in projecting onto bispectrum templates that have already been constrained. We hope to report on the opportunities for the detection of oscillations in the primordial bispectrum in future work.

3.4 Observational Constraints

3.4.1 Preliminaries

In order to use our results to place constraints on $|\beta|$ we need to relate them to existing templates. This is done by computing the dot product between our bispectrum result and the local, equilateral and orthogonal templates, which are standardized and whose magnitude has been constrained in the literature (Spergel et al. 2007; Senatore et al. 2009) (at the time, these were the best available constraints). To facilitate this, let us first rewrite our result in terms of the dimensionless variables $x_2 = k_2/k_1$ and $x_3 = k_3/k_1$, scaling out a factor of k_1^{-6}, indicating that the bispectrum is scale-invariant, where the comoving momentum magnitude k_1 is identified as the largest momentum. Because of the triangle constraint the shape function then only depends on a finite (triangle) domain in x_2 and x_3 space ($0 \leq x_2 \leq 1$ and $1 - x_2 \leq x_3 \leq 1$). Using these dimensionless variables, the leading shape function is then expressed as

$$\begin{aligned} F_{\rm leading}(x_2, x_3, p, \delta) &= \frac{1}{x_1 x_2}\left[\frac{1}{2}\frac{\cos(p(x_1+x_2-1)+\delta)}{p(x_1+x_2-1)} - \frac{\sin(p(x_1+x_2-1)+\delta)}{p^2(x_1+x_2-1)^2}\right. \\ &\left. + \frac{\cos\delta - \cos(p(x_1+x_2-1)+\delta)}{p^3(x_1+x_2-1)^3} + \text{cyc. perm.}\right], \end{aligned} \tag{3.23}$$

where $p \equiv k_1 c_s \eta_0$ is the enhancement factor we encountered several times before. Having identified the shape of the leading contribution we automatically also fixed a particular expression for the non-Gaussian amplitude $f_{\rm NL}$. It is important to note that the above shape function implies that a factor of p^3 has been incorporated in the non-Gaussian amplitude $f_{\rm NL}$.

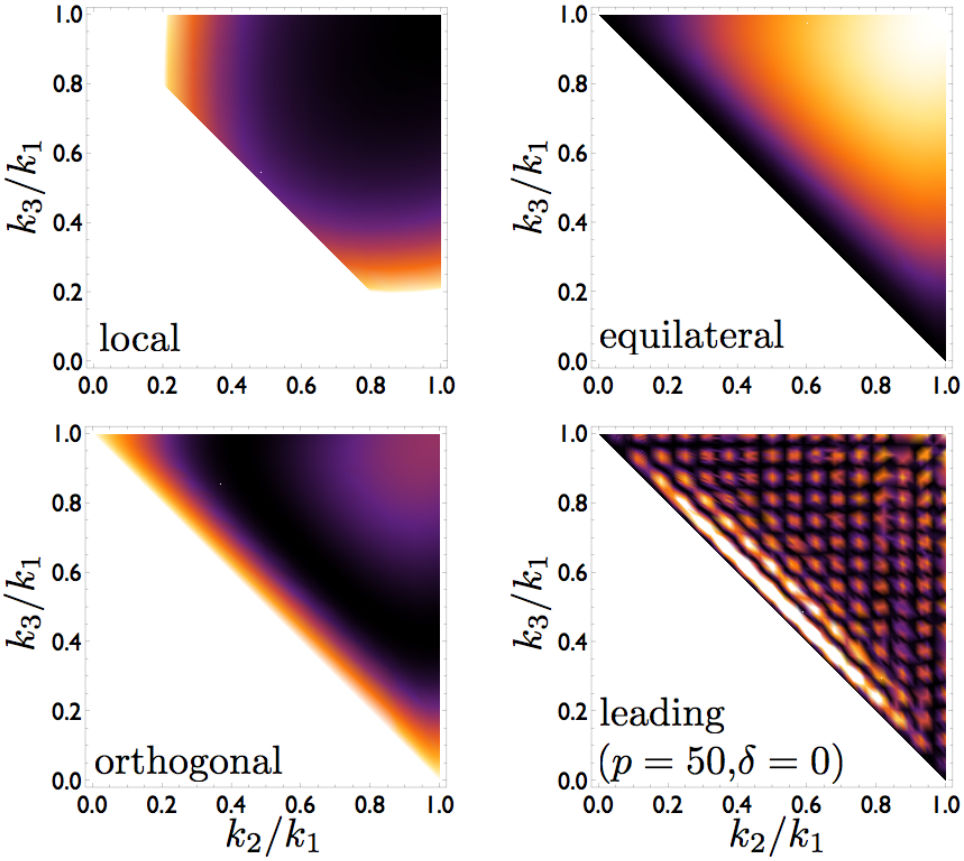

Figure 3.2: Here we plotted the projected 'density' of the different non-Gaussian shapes. As can be seen from the color (shading), large density corresponds to lighter color (shading). It is clear from inspection that the equilateral template has zero density on the collinear line (spanned by $x_2 + x_3 = 1$), while the leading order term from initial state modifications maximizes in the vicinity of this line, resulting in virtually zero overlap (which is already small due to the oscillatory nature of the leading term).

There are three different observational templates that we will be projecting our signal onto. These are the *squeezed* or *local*, *equilateral* and *orthogonal* templates. In terms of x_2 and x_3, their shape functions are expressed as follows:

$$F_{\rm local}(1, x_2, x_3) = 2\left(\frac{1}{x_2^3} + \frac{1}{x_3^3} + \frac{1}{x_2^3 x_3^3}\right), \tag{3.24}$$

$$F_{\rm eq}(1, x_2, x_3) = 6\left(-\frac{F_{\rm local}}{2} - \frac{2}{x_2^2 x_3^2} + \frac{1}{x_2^2 x_3^3} + \frac{1}{x_2 x_3^3} + \frac{1}{x_2 x_3^2} + \text{perm.}\right), \tag{3.25}$$

$$F_{\rm ort}(1, x_2, x_3) = 3\left(F_{\rm eq} - \frac{4}{x_2^2 x_3^2}\right). \tag{3.26}$$

Each of these templates has a different physical origin. The squeezed template was introduced by Spergel and Komatsu (Komatsu & Spergel 2001) to parameterize non-Gaussianities result-

ing from a quadratic correction in real space (i.e., a local term), corresponding to the leading nonlinear coupling between the inflaton field and the curvature perturbation at superhorizon scales. The equilateral template was introduced by Creminelli et al. (2006) to parameterize nonlinear higher-derivative corrections in the inflaton field, whose nonlinear magnitude is fixed at horizon crossing. As mentioned, the orthogonal template was recently introduced by Senatore et al. (2009) to allow for an optimal comparison with exotic non-canonical single-field scenarios that are, rather surprisingly, not of the equilateral type. A closely related template was also put forward in earlier work (Meerburg et al. 2009).

In order to make comparisons between these different shapes, Babich et al. (2004a); Fergusson & Shellard (2007) introduced dot products, cosines (the normalized dot product) and projection factors (to be introduced later) were defined to be

$$\begin{aligned} F_X \star F_Y &\equiv \int_{\Delta_k} dk_1 dk_2 dk_3 \frac{k_1^4 k_2^4 k_3^4}{k_t} F_X(k_1,k_2,k_3) F_Y(k_1,k_2,k_3), \\ \cos(F_X, F_Y) &\equiv \frac{F_X \star F_Y}{(F_X \star F_X)^{1/2}(F_Y \star F_Y)^{1/2}}. \end{aligned} \tag{3.27}$$

In the case where both F_Y and F_X are scale-invariant we can remove the integral over one variable, (i.e., k_1). In a previous definition of the dot (or star) product (Babich et al. 2004a) the factor $1/k_t$ was omitted. It was argued by Fergusson & Shellard (2009) that the addition of this extra 'weight' reduces the difference between the three-dimensional dot product for the bispectra and the two-dimensional dot product for the respective multipole equivalents. Since one actually measures the three-point function in two-dimensional multipole space, it is the two-dimensional dot product that one is interested in computing to derive constraints. Since the two-dimensional multipole bispectrum and its corresponding dot product are computationally challenging (and analytically ill-understood), it would indeed be a great improvement if a three-dimensional dot product could be introduced that guarantees reproduction of the two-dimensional dot product in multipole space. This is what the above dot product (including the $1/k_t$) is supposed to do, removing the need for a fully fledged two-dimensional analysis.

We will take this opportunity to clarify why the assumption of scale-invariance produces a difference with the old dot product which is (approximately) only a factor of 1/2. Note that the weight function $1/k_t$ is given by $1/(1+x_2+x_3)$ (after scaling out k_1), and the integration limits are given by $0 \leq x_2 \leq 1$ and $1-x_2 \leq x_3 \leq 1$. Consequently $1+x_2 \leq 1+x_2+x_3 \leq 2+x_2$ or $1 \leq 1+x_2+x_3 \leq 3$, resulting on average in $k_t = 1+x_2+x_3 \approx 2$, explaining the factor 1/2 difference. When computing the projection factor, a ratio of dot products that we are ultimately interested in, this factor 1/2 drops out and one ends up with exactly the same result using the former definition of the projection factor. We should emphasize that this is only true under the assumption of scale-invariance. Once this assumption is dropped one cannot scale out the k_1 explicitly and one has to perform the full three-dimensional integration in k-space. Since the dot product without the weight $1/k_t$ can be computed analytically we will actually be working with the standard definition of the three-dimensional dot product (see app. D eq. (D.1)) and assume the results will (approximately) equal the two-dimensional multipole space dot product.

Since we expect our bispectra to have poor overlap with the templates, and the dot product between the theoretically predicted bispectra is very hard to compute analytically, we do not

consider the cosine between the spectra. However, to get an indication how well (or how poor) a template does as compared to each other, we can define the ratio between two cosines

$$\mathcal{R}(X,Y) \equiv \left|\frac{\cos(F_X,F)}{\cos(F_Y,F)}\right| = \left|\frac{F_X\cdot F}{F_Y\cdot F}\right| \frac{\sqrt{F_Y\cdot F_Y}}{\sqrt{F_X\cdot F_X}}, \tag{3.28}$$

where F represents the theoretically predicted bispectrum, and F_X and F_Y represent different bispectrum templates. If $\mathcal{R}(X,Y)>1$, F_X is a better template and if $\mathcal{R}(X,Y)<1$ the opposite is true.

Fig. 2 shows the density of each shape which gives a qualitative indication of where the shapes peak as a function of the comoving momenta. We can immediately see that the orthogonal template naively appears to 'best' match the qualitative features in our signal (at least for $\delta=0$), an intuition that will be (partially) supported by calculations of the dot product. As an instructive example, let us consider the cross-correlations with the leading shape function eq.(3.23), which can be analytically determined,

$$\begin{aligned} F_{\rm local}\cdot F_{\rm leading} &= -\frac{3}{4}\frac{\cos\delta}{p}+\frac{\sin\delta}{p^2}+\mathcal{O}(p^{-3}),\\ F_{\rm eq}\cdot F_{\rm leading} &= 2\frac{\sin\delta}{p^2}+\mathcal{O}(p^{-3}),\\ F_{\rm ort}\cdot F_{\rm leading} &= \frac{1}{2}\frac{\cos\delta}{p}+3\frac{\sin\delta}{p^2}+\mathcal{O}(p^{-3}). \end{aligned}$$

Consequently both the dot product with the local as well as the orthogonal template peak for $\delta=0$, while the dot product with the equilateral template is $\pi/2$ out of phase. Note that we extracted three powers of p from the amplitude, so the dot products should be multiplied with p^3 to determine the strength of the enhancement that remains. The reason for this reduction in the level of enhancement is simply that the original enhancement was localized in the (leading) three-point function and integrating over the full domain will smear out this enhancement, typically reducing it with one power of p. Clearly, the dot product with the equilateral template is additionally suppressed in p as compared to the local and orthogonal template (for $\delta=0$ in fact all enhancement is lost). Visual inspection of the leading-order term in the density profile of fig. 3.2 indicates that the small dot product with the equilateral template is explained by the fact that this term peaks in the collinear limit, where the equilateral template actually vanishes. As mentioned before, this is precisely where the orthogonal template peaks, and therefore it is not a surprise that it would have a larger dot product. A significant feature of the orthogonal template is that it is not positive-definite within the triangle domain. As such it shares a feature with the oscillating contributions, possibly resulting in an even larger dot product. For comparison we find $\mathcal{R}(F_{\rm local},F_{\rm ort})\sim 0.42$ for $\delta=0$ which quantitatively confirms the orthogonal template has the best overlap with the theoretical spectrum[1].

We conclude that for generic values of the phase δ ($\delta\neq\pi/2$), the local and orthogonal templates would provide the best constraints if the shape function eq.(3.23) would be a good approximation to the full three-point function. As one should expect, the constraints resulting

[1] As a reference, the overlap between the local and equilateral shape is $\cos(F_{\rm local},F_{\rm eq})\sim 0.41$.

from these expressions will also depend on the tightness of the existing observational constraints on the corresponding $f_{\rm NL}$, which will favor the local template.

For completeness we should also consider, in addition to the term (3.23) which is a unique consequence of introducing a vacuum state modification, the non-Gaussian signature as a consequence of non-standard c_s (i.e., $\mathcal{O}(\beta^0)$). Chen et al. (2007) showed that the leading terms are given by

$$\mathcal{B}_\lambda = \frac{1}{k_1^3 k_2^3 k_3^3}\left(\frac{1}{c_s^2} - 1 - \frac{2\lambda}{\Sigma}\right)\frac{3k_1^2 k_2^2 k_3^2}{2k_t^3}, \tag{3.29}$$

$$\mathcal{B}_c = \frac{1}{k_1^3 k_2^3 k_3^3}\left(\frac{1}{c_s^2} - 1\right)\left(-\frac{1}{k_t}\sum_{i>j} k_i^2 k_j^2 + \frac{1}{2k_t^2}\sum_{i\neq j} k_i^3 k_j^3 + \frac{1}{8}\sum_i k_i^3\right), \tag{3.30}$$

where we neglected the corrections due to time-variation of various parameters as well as corrections to the solutions of the field equations of motion. Both of these contributions have a k-dependence that strongly resembles the equilateral template,

$$\cos(F_{\rm eq}, \mathcal{B}_\lambda) \approx \cos(F_{\rm eq}, \mathcal{B}_c) \approx 0.95. \tag{3.31}$$

This contribution is therefore expected to add a term to the equilateral non-Gaussian amplitude, in addition to the modified vacuum state effect. It will of course be interesting to see how these contributions compare.

3.4.2 General Inflationary Models with a Small Speed of Sound

We can use the dot product, i.e., the correlator, to define projections, which measures the relative "leakage" of one shape function into one of the available observational templates

$$\Delta(F_X, F_Y) \equiv \frac{F_Y \cdot F_X}{F_X \cdot F_X}. \tag{3.32}$$

Using the normalizations (self-dot products) of the different templates that are given in app. D, it is straightforward to compute the projection factors for the different templates with the leading shape behavior of the corrected three-point function due to a modified vacuum state. For large p this is well-described by the shape function that was discussed in the previous paragraph. The results are

$$\Delta(F_{\rm leading}, F_{\rm local}) \approx \frac{1}{176.5}\left[-\frac{3}{4}\frac{\cos\delta}{p} + \frac{\sin\delta}{p^2}\right], \tag{3.33}$$

$$\Delta(F_{\rm leading}, F_{\rm eq}) \approx \frac{2}{7.9}\frac{\sin\delta}{p^2}, \tag{3.34}$$

$$\Delta(F_{\rm leading}, F_{\rm ort}) \approx \frac{3}{13.8}\left[\frac{1}{6}\frac{\cos\delta}{p} + \frac{\sin\delta}{p^2}\right], \tag{3.35}$$

$$\Delta(\mathcal{B}_\lambda, F_{\rm eq}) \approx 0.01, \tag{3.36}$$

$$\Delta(\mathcal{B}_c, F_{\rm eq}) \approx -0.05. \tag{3.37}$$

From the corresponding amplitude of the three-point function, including the p^3 enhancement factor, one then derives the following expressions for the local, equilateral (including contributions due to a small speed of sound) and orthogonal non-Gaussian amplitudes:

$$f_{\rm NL}^{\rm local} \simeq \frac{1}{176.5}\left(-\frac{5}{4}\cos\delta+\frac{5}{3}p^{-1}\sin\delta\right)\left(\frac{1}{c_s^2}-1+\frac{2\lambda}{\Sigma}\right)|\beta|p^2, \tag{3.38}$$

$$f_{\rm NL}^{\rm eq} \simeq -\frac{5}{3}\left[\left(\frac{1}{c_s^2}-1\right)\left(\frac{-2}{7.9}|\beta|p\sin\delta+0.04\right)+\left(\frac{-2}{7.9}|\beta|p\sin\delta-0.01\right)\frac{2\lambda}{\Sigma}\right], \tag{3.39}$$

$$f_{\rm NL}^{\rm ort} \simeq \frac{1}{13.8}\left(\frac{5}{6}\cos\delta+5p^{-1}\sin\delta\right)\left(\frac{1}{c_s^2}-1+\frac{2\lambda}{\Sigma}\right)|\beta|p^2. \tag{3.40}$$

Again, in deriving these results we have assumed, based on the arguments presented earlier, that the three-dimensional dot products are good approximations of the two-dimensional dot products of the multipole bispectra. Also note that the two terms in the equilateral non-Gaussian amplitude that are not proportional to $|\beta|$ are a direct consequence of a small speed of sound, which produces an enhanced equilateral non-Gaussian component which cannot be neglected for $\delta = 0$. The local and orthogonal contributions are, however, uniquely due to the modified vacuum state. We are now in a position to use these predictions to put constraints on the parameter space, consisting of c_s, $|\beta|$, δ and p. The existing constraints on $f_{\rm NL}^{\rm local}$ (Smith et al. 2009), $f_{\rm NL}^{\rm equil}$ (Spergel et al. 2007; Senatore et al. 2009) and $f_{\rm NL}^{\rm ort}$ (Senatore et al. 2009) are given by

$$\begin{aligned} -4 &\leq f_{\rm NL}^{\rm local} \leq 80, \\ -125 &\leq f_{\rm NL}^{\rm equil} \leq 435, \\ -369 &\leq f_{\rm NL}^{\rm ort} \leq 71. \end{aligned}$$

Based on these results let us now consider two extremes: $\delta = 0$ (real β) and $\delta = \pi/2$ (imaginary β).

For $\delta = 0$, eq. (3.34) vanishes, which implies the linearly enhanced equilateral contributions proportional to $|\beta|$ vanish. In that case we can actually use the equilateral constraints to provide a bound on the speed of sound, similar to that found by Senatore et al. (2009), but as we will see this is not even required to derive an interesting constraint on $|\beta|$. One should also realize that the $\delta = 0$ case describes the generic order of magnitude expectation, in the sense that as long as $\delta \neq \pi/2$ the local and orthogonal contributions remain p^2 enhanced, which is the most important property as far as the constraints are concerned. The constraints from the orthogonal template and the local template are very similar. Since $|\beta|$ is positive-definite this implies only the positive constraint remains (the negative constraint can be used for $\delta = \pi$). Note that the projection factor between the local and the leading shape function has a negative sign, so the final constraint comes from the (negative) lower bound on local non-Gaussianities $f_{\rm NL}^{\rm local}$, which is given by

$$|\beta| \lesssim \frac{6\times10^2}{p^2}\left(\frac{1}{c_s^2}-1+\frac{2\lambda}{\Sigma}\right)^{-1}, \quad (\delta = 0). \tag{3.41}$$

To make this more concrete, let us assume that the ratio $2\lambda/\Sigma \ll 1/c_s^2$ in the small sound speed limit. The constraint then reads $|\beta| < 6\times10^2 c_s^2/p^2$ and after realizing that $p \equiv k_1 c_s \eta_0 = \Lambda c_s/H$

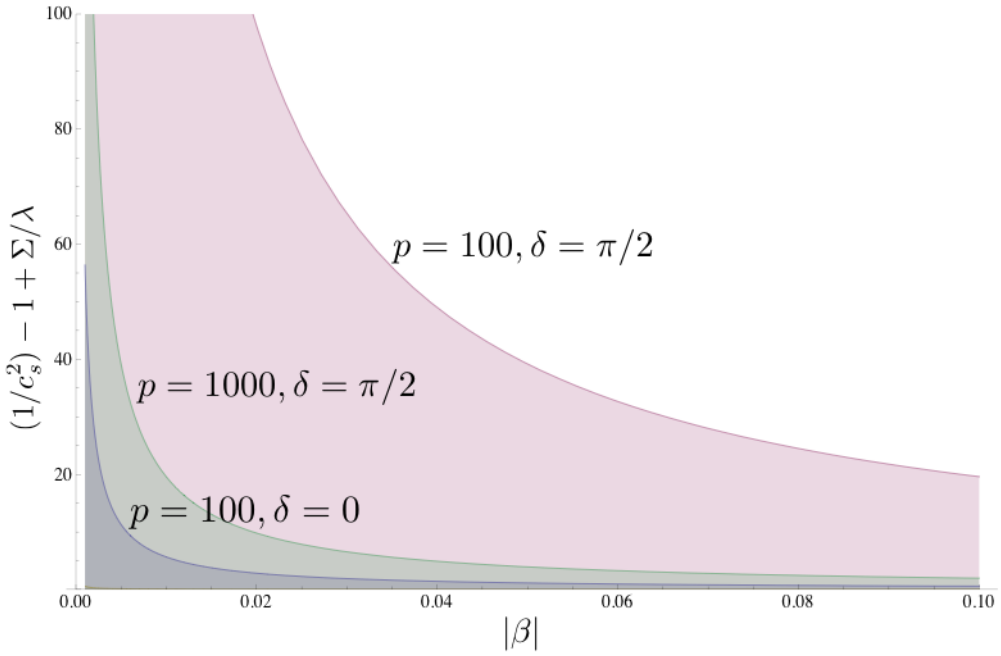

Figure 3.3: Constraining contours for different values of $p = c_s k_1 \eta_0$ and two values of δ. The constraints for intermediate values of δ should lie somewhere in between. For values $\pi/2 < \delta < 3\pi/2$ one has to consider the opposite (in sign) constraints on the various $f_{\rm NL}$. The constraint for $p = 1000$ and $\delta = 0$ is too tight to be visible in this graph. For $\delta = \pi/2$ and $p = 1000$ the constraint is still weaker than the constraint for $\delta = 0$ and $p = 100$.

this implies the interesting, sound speed-independent, constraint $|\beta| < 6 \times 10^2 H^2/\Lambda^2$. This could easily give rise to a constraint on $|\beta|$ at the level of 10^{-4}, independent of any specific model for $|\beta|$. In particular it seems to exclude Bogolyubov parameters linear in H/Λ (Danielsson 2002) [2]. Therefore in the small speed of sound limit (and $2\lambda/\Sigma \ll 1/c_s^2$), we obtain an impressive constraint on $|\beta|$. As long as the phase $\delta \neq \pi/2$ this conclusion is unaffected. For values $\delta > \pi$ the sign in the leakage factors is altered and the best constraint is obtained from orthogonal non-Gaussianities. Due to the larger number bounds on this $f_{\rm NL}$, this results in a constraint that is one order of magnitude below the best constraint.

Now let us briefly consider the special case $\delta = \pi/2$. For the local and orthogonal contributions this means the leading p behavior is one order of magnitude less, which implies that the equilateral contribution is now of the same order. Looking at the equilateral contribution we conclude that we can neglect the equilateral contributions from the zeroth order bispectrum (i.e., without initial state modification) as long as $|\beta| > 0.1/p$. If the Bogolyubov parameter is smaller than that we can actually use the equilateral contribution to put a constraint on c_s.

[2] Strictly speaking, starting with the assumption that the Bobolyubov parameter is proportional to $\frac{H^\star}{\Lambda}$, which introduces another factor of c_s, one would obtain a lower bound on the ratio of $\frac{H}{\Lambda}$ depending on the speed of sound which turns out to be in conflict with constraints from the power spectrum in the limit of a small sound speed.

This time, the strongest constraint is derived from the orthogonal template and reads

$$|\beta| \lesssim \frac{2\times 10^2}{p}\left(\frac{1}{c_s^2}-1+\frac{2\lambda}{\Sigma}\right)^{-1}, \quad (\delta=\pi/2). \tag{3.42}$$

Obviously, this constraint is much weaker and can be written as $|\beta| < 2\times 10^2 H c_s/\Lambda$, in the limit of a small speed of sound and $2\lambda/\Sigma \ll 1/c_s^2$. Since the enhancement in p is reduced in this case one should worry about subleading contributions since they might be comparable in magnitude. Upon checking this one finds that for $\delta=\pi/2$ one should indeed take into account both the subleading and the leading order contribution. In the next section we will derive the projection factors for the subleading terms, since these are also the relevant terms in DBI inflation for which $(1/c_s^2-1+2\Sigma/\lambda)$ vanishes identically.

3.4.3 DBI Models of Inflation

For DBI models (IR or UV) the leading-order term of the vacuum state corrected three-point function vanishes. Consequently we will need to compute the dot product with the next-to-leading order terms. The next-to-leading contribution goes as (from eq.(3.43) and eq.(3.43))

$$\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle_{sl} \sim \frac{1}{k_1^3k_2^3k_3^3}|\beta|\left(\frac{1}{c_s^2}-1\right)(k_1c_s\eta_0)^2\sum_j(\mathcal{B}^{(2)}_{k_j}+\mathcal{B}^{(3)}_{k_j}),$$

where we have 2 distinguishable contributions,

$$\begin{aligned}
\mathcal{B}^{(2)}_{k_j} &= \frac{3}{4}k_2^2k_3^2\left[\left(\frac{\tilde{k}_j-k_j}{k_j^2}+\frac{\tilde{k}_j+k_{j+1}}{k_{j+1}^2}+\frac{\tilde{k}_j+k_{j+2}}{k_{j+2}^2}\right)\left(\frac{\cos\delta}{(\tilde{k}_jc_s\eta_0)^2}-\frac{S_1}{\tilde{k}_jc_s\eta_0}\right)\right.\\
&\quad\left.+\left(\frac{1}{k_j}-\frac{1}{k_{j+1}}-\frac{1}{k_{j+2}}\right)S_2\right],\\
\mathcal{B}^{(3)}_{k_j} &= \frac{1}{8}\left[\frac{k_2k_3}{k_1}(2k_j^2-k_{j+1}^2-k_{j+2}^2)S_2-\right.\\
&\quad\left.\frac{1}{k_1^2}\left(k_1k_2k_3+\tilde{k}_j(k_jk_{j+1}+k_jk_{j+2}-k_{j+1}k_{j+2})\right)\left(\frac{\cos\delta}{(\tilde{k}_jc_s\eta_0)^2}-\frac{S_1}{\tilde{k}_jc_s\eta_0}\right)\right].
\end{aligned}$$

As was the case for the leading contribution, both terms are finite in the collinear limit (as shown explicitly in app. C). Both contributions are of the same order of magnitude over the full triangle domain, so both have to be included in our analysis. One might argue that the cosine in both of these functions is suppressed and can be neglected. However, for an analytical treatment of the projection factors it turns out to be convenient to work with the full expression because of its finite collinear limit. Moreover, as we will see shortly, the phase δ significantly influences the final results for the projection factors.

As before, let us rewrite these contributions in terms of the dimensionless variables x_2 and

x_3. We will identify the shape of the first sub-leading contribution as follows

$$F_{sl(1)} = -\frac{3}{4}\frac{1}{x_2x_3}\left[\left(1-\frac{1}{x_2}-\frac{1}{x_3}\right)\frac{\sin(p(x_2+x_3-1)+\delta)}{p(x_2+x_3-1)}\right. \\ \left. +\left(x_2+x_3-2+\frac{1-2x_2+x_3}{x_2^2}+\frac{1+x_2-2x_3}{x_3^2}\right)\frac{\cos\delta-\cos(p(x_2+x_3-1)+\delta)}{p^2(x_2+x_3-1)^2}\right], \tag{3.43}$$

plus cyclic permutations. The shape of the second sub-leading term equals

$$F_{sl(2)} = -\frac{1}{8}\frac{1}{x_2^3x_3^2}\left[x_2x_3\left(1+x_2^2+x_3^2\right)\frac{\sin(p(x_2+x_3-1)+\delta)}{p(x_2+x_3-1)}\right. \\ \left. -\left((1+x_2^2+x_3^2)(x_2x_3+(x_2+x_3-1)(x_2+x_3-x_2x_3)\right)\frac{\cos\delta-\cos(p(x_2+x_3-1)+\delta)}{p^2(x_2+x_3-1)^2}\right], \tag{3.44}$$

plus cyclic permutations. These identifications of the shape function imply we have extracted a factor of p^2 from the non-Gaussian amplitude.

Before computing the dot products let us have a crude first look at the shape functions. They appear quite similar to the leading order shape, but somewhat surprisingly, as soon as δ deviates from zero the sub-leading contributions change qualitatively. On the left in fig. (3.4) we have plotted eq. (3.44) for $p = 100$ and $\delta = 0$, corresponding to a rapidly-oscillating function that maximizes in the collinear limit. On the right however, we plotted the same shape function with the phase $\delta = 3\pi/2$, which suddenly appears to resemble the local shape. Interestingly, the appearance of local type features was also noted in the different context of multi-field DBI models (assuming the standard Bunch-Davies vacuum state) (Renaux-Petel 2009). Based on this observation one would expect the dot product with the local shape to increase as a function of δ (as opposed to what happens for the leading shape function). Of course, the collinear feature is still present, but simply swamped by a strong local feature that appears for $\delta = 3\pi/2$. In the previous analysis of the leading term we also found that the dot product with the orthogonal template was π out of phase as compared to the dot product with the local template and that it was comparable in magnitude. We would expect to find something similar in this sub-leading DBI case.

Because of the clear presence of local non-Gaussian features for $\delta > 0$, which are actually divergent in the $x_2 = 0$ or $x_3 = 0$ limit, we have to introduce a cutoff when computing dot products, the same cutoff used to normalize the local template. This cutoff is necessary to reflect that we cannot observe infinite wavenumbers in the CMB. It is given by $k_{\rm min}/k_{\rm max} \sim 10^{-3}$, the largest observable physical scale divided by the smallest observable physical scale. So we will require that $x_2, x_3 > \epsilon$, with the cutoff ϵ equal to $\epsilon = 10^{-3}$.

We find that there is substantial overlap with both the local and orthogonal templates. For the dot product with the local template we obtain

$$(F_{sl(1)}+F_{sl(2)})\cdot F_{\rm local} = \frac{1}{2p}(9+8\log\epsilon)\sin\delta+\frac{1}{p^2\epsilon}\left[2\cos(\delta+2p\epsilon)-2(1-\pi p\epsilon)\cos\delta\right. \\ \left. +4p\epsilon\left(\mathrm{ci}(2p\epsilon)\sin\delta+\mathrm{si}(2p\epsilon)\cos\delta\right)\right]+\mathcal{O}(p^{-2},\epsilon). \tag{3.45}$$

Here ci = cosintegral and si = sinintegral. Note that since $\epsilon \sim 10^{-3}$ we should be careful identifying the second term as p^{-2} suppressed. For small values of p the first term dominates.

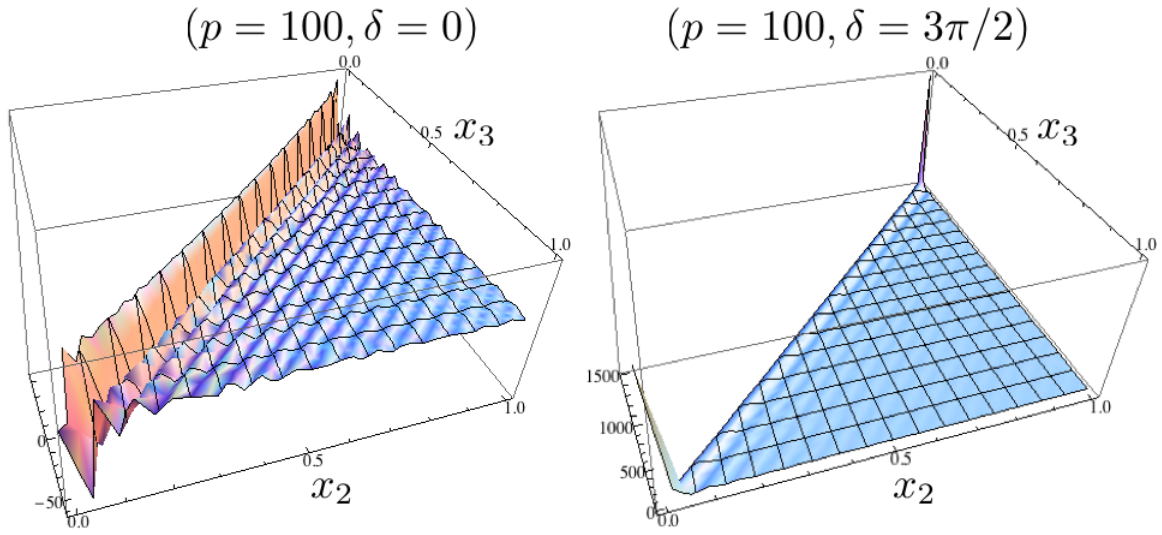

Figure 3.4: The shape of the first subleading contribution (eq.(3.44)) for different values of δ. For $\delta = 0$ the shape of this function is very similar to the leading order term. Once $\delta \neq 0$ the shape picks up a significant local feature which leads us to expect a significant overlap with the local non-Gaussian template. Since the oscillations are still present, we also expect a comparable overlap with the orthonormal template. As for the second subleading contribution (eq. (3.45), not shown) we find similar behavior, but opposite in sign, resulting in a small but non-negligible cancellation.

First of all, remember that a factor of p^2 was scaled into the non-Gaussian amplitude $f_{\rm NL}$, explaining the appearance of the leading $1/p$ suppression. Considering the case $\delta = \pi/2$, $\epsilon = 10^{-3}$, we find that $(F_{sl(1)} + F_{sl(2)}) \cdot F_{\rm local} \sim -31/p$. Properly comparing this to the dot product of the local template with the leading order shape function (which requires scaling in one factor of p, to match the normalizations of the amplitude), we find that that the overlap between the subleading term and the local shape is almost 42 times better, which is rather surprising.

In fig. 3.5 we have plotted the dot product for two values of p as a function of δ. It clearly shows that even though the dot product is smaller for $\delta = 0$ it is still relatively large compared to the dot product between the leading term and the local template. There exist values for δ for which the dot product is exactly zero. For those values of the phase the local template would not be useful to derive constraints on vacuum state modifications.

We were unable to obtain an analytical expression for the dot product with the equilateral template and the subleading bispectrum. Instead we performed a numerical analysis, which is discussed below. Next, we compute the dot product with the equilateral and orthogonal templates. Semi-analytically we obtain from best fitting:

$$(F_{sl(1)} + F_{sl(2)}) \cdot F_{\rm ort} \approx 5.3\frac{\sin\delta}{p} + 11.3\frac{\log p}{p^2}\cos(\delta - 0.2 - 2.4\sqrt{p}). \tag{3.46}$$

We compute that $\mathcal{R}(F_{\rm local}, F_{\rm ort}) \sim 1.71$ for $\delta = \pi/2$, indicating that, as expected, the local

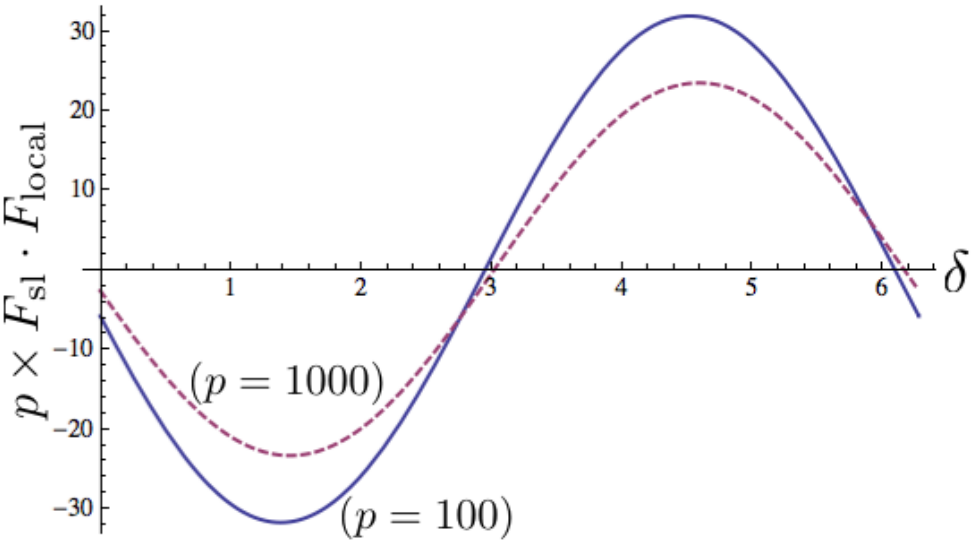

Figure 3.5: The dot product between the subleading ($F_{\rm sl} = F_{\rm sl(1)} + F_{\rm sl(2)}$) and local template (multiplied with p) for two values of p. Although the dot product is much smaller for $\delta = 0$ it is non-zero for both p. However, there exist values of the phase for which the dot product (and consequently the leakage or fudge factor) is exactly zero. For such δ the local template would not be suitable (or at least limited) for detecting any deviation from a BD state (in a DBI model of inflation) and we will have to use the much less capable equilateral template.

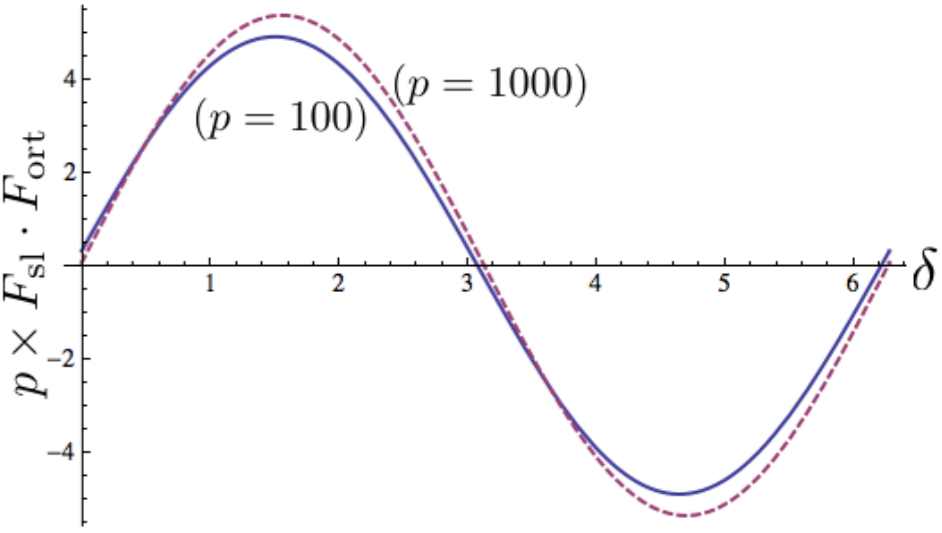

Figure 3.6: The dot product between the subleading and the orthogonal template (multiplied with p) for two values of p. This dot product is almost completely dominated by the sines and therefore almost zero for $\delta = 0$

shape has the best resemblance with the subleading bispectrum.

As in the previous section we use these to obtain the projection factors and then translate these into expressions for the various $f_{\rm NL}$ (which introduces a factor of p^2 that was absorbed in the non-Gaussian amplitude):

$$f_{\rm NL}^{\rm local} \simeq \frac{1}{176.5}\frac{5}{6}\Big[(9-32\log 10)\sin\delta+\frac{10^3}{p}\left(2\cos(\delta+2\times10^3 p)-2(1+\pi\times10^{-3}p)\cos\delta\right.$$
$$\left.+4\times10^{-3}p(\mathrm{ci}(2\times10^{-3}p)\sin\delta+\mathrm{si}(2\times10^{-3}p)\cos\delta)\right)\Big]\left(\frac{1}{c_s^2}-1\right)|\beta|p \qquad (3.47)$$

and

$$f_{\rm NL}^{\rm ort} \simeq \frac{1}{13.8}\frac{5}{3}\left(5.3\sin\delta+11.3\frac{\log p}{p}\cos(\delta-0.2-2.4\sqrt{p})\right)\left(\frac{1}{c_s^2}-1\right)|\beta|p. \qquad (3.48)$$

Once more we distinguish two values of δ. For $\delta=\pi/2$ and in the small sound speed limit, using that $p\equiv\Lambda\, c_s/H$, the best constraint comes from local non-Gaussinaties and reads

$$|\beta| \lesssim \frac{14H\,c_s}{\Lambda},\ \ (\delta=\pi/2), \qquad (3.49)$$

which is substantially tighter than the constraint from leading order contributions for the same value of the phase δ (eq. (3.42)). For all the values in between (i.e., $0\leq\delta\leq\pi/2$) the constraints will in general be somewhat weaker. Once the sign of the leakage factor changes, the constraints loose about an order of magnitude, which still exceeds the constraints from the leading term. From eq. (3.47) we find the following constraint for $\delta=0$

$$|\beta| \lesssim \frac{72H\,c_s}{\Lambda},\ \ (\delta=0). \qquad (3.50)$$

For $\delta\sim 2.9$ eq. (3.49) is zero and therefore not useful to place any constraints on $|\beta|$. Similar to the dot products with the leading order terms the dot product with the equilateral template is approximately $\pi/2$ out of phase with the orthogonal and local ones. Obtaining an analytical solution for the dot product between the equilateral template and the subleading non-Gaussianities turned out to be technically involved, instead we computed these numerically for different values of p and δ. These are shown in fig. 3.7. What we find is similar to what we found for the leading order term: it is indeed out of phase with the other two dot products, nor is it enhanced in p (only log-scale). These observations imply that although in principle one could use the equilateral template to constrain $|\beta|$ in the vicinity of those values for which the other two templates would not be able to place a constraint at all, due to the poor overlap with equilateral non-Gaussianities any constraint on $|\beta|$ would be very weak. It must be noted that the range of values for δ for which the constraints are weak is limited and it appears, due to small corrections, that the zero points of the the dot products between local and subleading and orthogonal and subleading do not exactly coincide. As a consequence one might still be able to obtain slightly better results using the orthogonal constraints. Since the equilateral template is only log-scale enhanced by the cutoff, given the constraint on $|\beta|$ from the power spectrum (roughly $|\beta|\lesssim 10^{-2}$), the signal in the equilateral template will sometimes be dominated by the

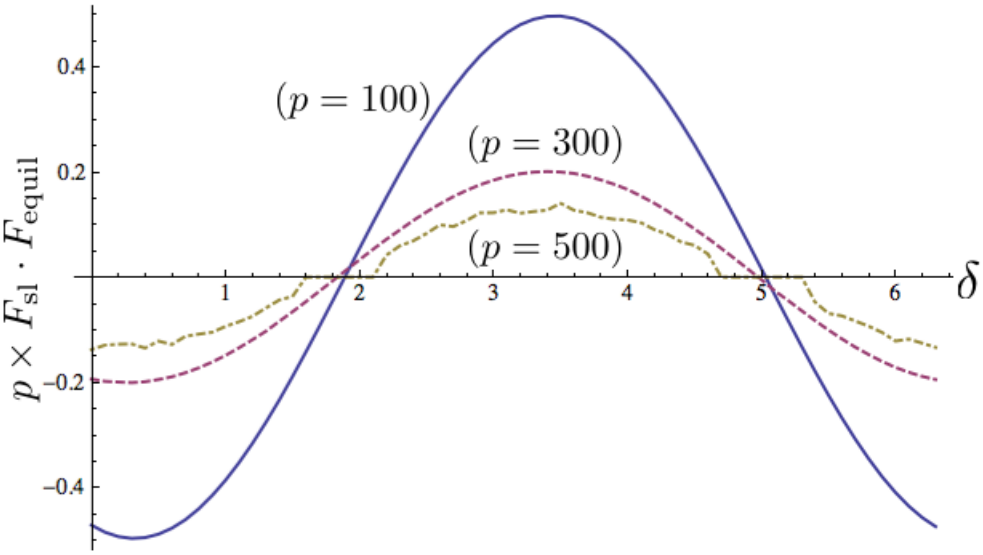

Figure 3.7: Numerical results for the dot product (multiplied by p) between the equilateral template and the subleading terms. The dot product gets smaller for larger values of p, indicating that the constraints using this template are not improved for larger values of p (recall that one p was absorbed into the definition of the associated $f_{\rm NL}$). In addition, the dot product is already small to begin with, indicating that the equilateral template might be useful to constrain DBI inflation itself, but is rather poor in constraining DBI inflation with modified initial states. For values of $p = 500$ and beyond the numerical resolution is not sufficient to reproduce the expected (cosine) oscillations. This holds especially for values that are close to zero. There was no need for a cutoff (ϵ) in these calculations as the dot product was not divergent (as expected) for small values of x_2 and x_3.

'regular' DBI contribution, which would allow us to put a constraint on c_s thereby breaking the degeneracy between c_s and $|\beta|$. Unfortunately, this is only true for a small range in δ, effectively only for those values of δ for which the equilateral contribution from vacuum modification is zero in any case. From best fitting our numerical results, we have been able to deduce that

$$(F_{sl(1)} + F_{sl(2)}) \cdot F_{\rm equil} \quad \sim \quad -10.8\cos(\delta - 0.3) \times \frac{\log p}{p^2}. \tag{3.51}$$

Consequently, $f_{\rm NL}^{\rm equil}$ can be written as

$$f_{\rm NL}^{\rm equil} \quad \simeq \quad -\frac{5}{3}(0.05 + 10.8\cos(\delta - 0.3) \times |\beta| \log p)\left(\frac{1}{c_s^2} - 1\right). \tag{3.52}$$

For DBI models of inflation the speed of sound c_s^{-1} can easily be $\mathcal{O}(10)$. In such a scenario for generic values of the phase δ the constraints become

$$|\beta| \quad \lesssim \quad 1.4\frac{\Lambda}{H} \;\; (\delta = \pi/2), \tag{3.53}$$

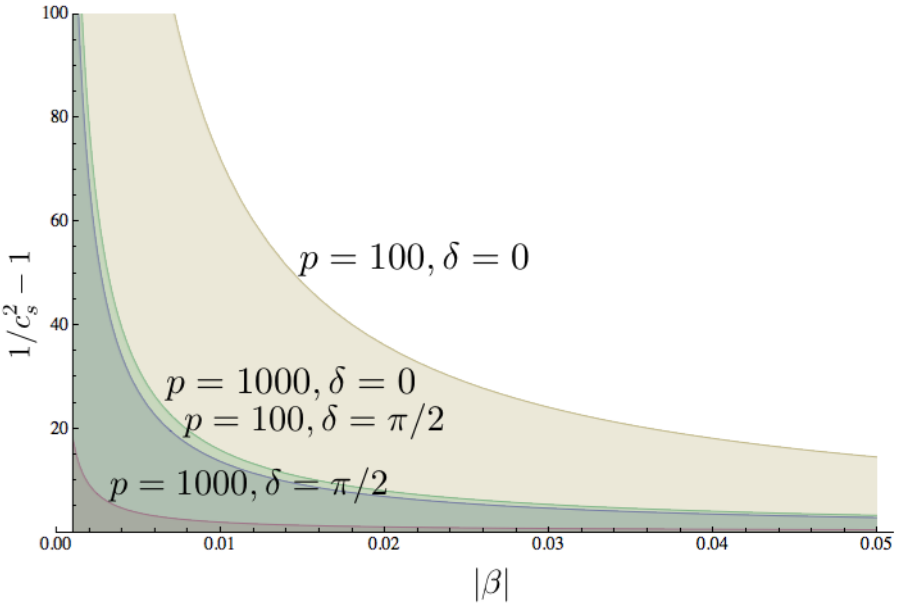

Figure 3.8: Constraints on $|\beta|$ as function of $1/c_s^2 - 1$ derived from the dot product and the 3σ levels of local non-Gaussianities for different values of p and δ. These constraints are better than the constraints coming from measurements of the power spectrum. However, this requires several assumptions due to the appearance of 4 independent variables (p, c_s, $|\beta|$ and δ). p could be measured independently if one could be able to measure the frequency of the oscillations in the bispectrum. c_s could be determined if one could deduce what type of DBI model initiated inflation. As such, one could break down some of the degeneracy.

and

$$|\beta| \lesssim 7.2\frac{\Lambda}{H} \ (\delta = \pi/2), \tag{3.54}$$

which implies that for realistic values of the cutoff $|\beta| \lesssim 10^{-3}$. This improves the constraints from the power spectrum by one order of magnitude. As noted however, there exists a limited range in δ for which the constraints are weaker due to the poor overlap with existing templates. Unlike the constraint coming from the bispectrum, the constraints on $|\beta|$ obtained from the power spectrum do not depend on the cutoff scale or the phase of the Bogolyubov parameter. As such, the bound on $|\beta|$ for these values of the phase is best constraint by the power spectrum.

3.5 Conclusions

We have analyzed three-point correlators in single-field inflation with a small speed of sound after introducing a small departure from the standard Bunch-Davies vacuum. Our aim was to derive the best possible constraints on slightly modified vacua in the specific context of small sound speed models like DBI inflation. In previous work (ch. 2, Meerburg et al. (2009)) we concluded that the bispectrum is an excellent probe of vacuum state modifications due to

enhancement of the non-Gaussian signal by a factor proportional to (powers of) the ratio of the cutoff scale and the Hubble parameter $p \equiv \Lambda/Hc_s$, which could easily be as large as 10^3. We have shown that in the small sound speed limit the powers of p that appear can be as large as p^3, which adds one factor of p as compared to the $c_s = 1$ model. In addition the level of non-Gaussianity in small sound speed models is already large to begin with. Therefore it is not surprising that the constraints that can be derived in such models are in general stronger. The computation of the contributions to the local, equilateral and orthogonal template confirmed this expectation. In general single field models with a small speed of sound the localized enhancement is proportional to p^3, and after projecting onto the local and orthogonal template one factor of p is lost, ending up with p^2 enhancement. The inclusion of an arbitrary phase in our analysis in the Bogolyubov parameter turns out to be able to reduce the enhancement with one factor of p for very specific values of the phase. For general single field models with a small speed of sound this implies that for generic phases the constraint on the absolute value of the Bogolyubov parameter equals $|\beta| \lesssim 6 \times 10^2 H^2/\Lambda^2$, which is actually independent of the speed of sound. This strong constraint is in fact excluding Bogolyubov parameters linear in H/Λ in the small sound speed limit. To be precise, assuming $|\beta| \propto \frac{H^\star}{\Lambda}$ one obtains a (c_s dependent) *lower* bound on the ratio $\frac{H}{\Lambda}$, which in the limit of a small sound speed would violate the power spectrum constraint. For the specific value of the phase $\delta = \pi/2$, the subleading terms are of the same order and have to be included in the analysis. Similarly, it is the subleading terms that govern the constraint for DBI models, for which the leading term identically vanishes.

In the context of DBI models we performed the same analysis, with a reduced enhancement of p^2 in general. After projecting onto the different templates and enforcing the observational bounds this resulted in a constraint on the absolute value of the Bogolyubov parameter $|\beta| \lesssim 14H/(\Lambda c_s)$, which is linear in H/Λ and explicitly depends on the speed of sound. This subleading contribution is the one that should also be taken into account for the specific value of the phase $\delta = \pi/2$ in the general single field case, since this bound is much stronger than the bound from the second order terms in the leading expression. For general single field inflation with a small speed of sound this results in a very tight bound on allowed values of $|\beta|$ ranging from 10^{-5} to 10^{-3}. For DBI models of inflation the constraints are only determined by the subleading term as the leading term vanishes. Again, the inclusion of the phase of the Bogolyubov parameter has the effect that for the specific value of $\delta = 0$ the constraint weakens slightly, equal to $|\beta| \lesssim 72H/\Lambda c_s$, but this time remains at the same order in p. Overall we conclude that for generic values of δ and realistic values of the cutoff this still implies $|\beta| \lesssim 10^{-3} - 10^{-2}$. For very specific values of the phase, the bounds could vanish altogether. For these special phases one would have to rely on the power spectrum to set a constraint on $|\beta|$.

Our main message is that for single field models with a small speed of sound, including DBI, the constraints on vacuum state modifications from the bispectrum for generic values of the phase are already better than the power spectrum constraints, relying on the available templates. The main reason for this is the huge (localized) enhancement that appears after introducing a modified initial state and a corresponding cutoff that regularizes the relevant integrals for the bispectrum calculation. Improved templates, which would have to be sensitive to the oscillatory nature of most of the contributions to the bispectrum, are guaranteed to tighten these constraints significantly. As was previously noted (Meerburg et al. 2009), the

oscillatory nature of the signal in momentum space suggests that specific, perhaps observable, features could appear in three- and two-dimensional position space. In any case it would be worthwhile to generalize the range of available non-Gaussian shapes that can be compared to the data, including oscillatory signals, on which we hope to report in the future.

Our results confirm that higher-derivative corrections are excellent probes for departures from the standard Bunch-Davies vacuum state. Throughout this chapter we assumed that the combination $|k_1\eta_0|$ is independent of the actual comoving momenta involved and equal to Λ/H, in the spirit of the New Physics Hypersurface approach to vacuum state modifications. The reason for this was scale-invariance of the bispectrum, which we relied on to allow for comparison with the available (scale-invariant) template shapes. Fixing η_0 instead, as one would do in a Boundary Effective Field Theory approach to vacuum state modifications, immediately results in a scale-dependent bispectrum. It would be interesting to study such scale-dependent scenarios and determine to what extent (future) analysis of 3d large scale structure or 2d CMB data can constrain bispectrum departures from scale-invariance (Sefusatti et al. 2009).

As reported, the bispectrum or three-point function is extremely sensitive to initial vacuum state modifications in the presence of a higher-derivative operator, and there is no reason to think this could not similarly be true for all higher n-point functions. A more general perturbative analysis, including higher n-point functions (Huang & Shiu 2006; Seery & Lidsey 2007; Seery et al. 2009; Gao & Hu 2009; Arroja et al. 2009; Lehners & Renaux-Petel 2009; Gao et al. 2009; Chen et al. 2009), might lead to a hierarchy of (theoretical) constraints on vacuum state modifications, perhaps ruling out departures from the standard Bunch-Davies state altogether. Indeed, for the four-point function (trispectrum) the effect of a vacuum state modification was recently studied by Chen et al. (2009), again showing a strong sensitivity. We hope that ongoing future work in this direction can help us further understand and identify the phenomenological and theoretical constraints the inflationary vacuum state has to satisfy.

Acknowledgments

This research was supported in part by a VIDI and a VICI Innovative Research Incentive Grant from the Netherlands Organization for Scientific Research (NWO). PDM was supported by the Netherlands Organization for Scientific Research (NWO), NWO-toptalent grant 021.001.040. The research of JPvdS is financially supported by Foundation of Fundamental Research on Matter (FOM) grant 06PR2510.

Chapter 4

Minimal Cut-Off Vacuum State Constraints from CMB Bispectrum Statistics

P. Daniel Meerburg & Jan Pieter van der Schaar
Phys. Rev. D **83** (2011) 043520

Abstract: In this short note we translate the best available observational bounds on the CMB bispectrum amplitudes into constraints on a specific scale-invariant New Physics Hypersurface (NPH) model of vacuum state modifications, as first proposed by Danielsson (2002), in general models of single-field inflation. As compared to the power spectrum the bispectrum constraints are less ambiguous and provide an interesting upper bound on the cutoff scale in general models of single-field inflation with a small speed of sound. This upper bound is incompatible with the power spectrum constraint for most of the parameter domain, leaving very little room for minimal cutoff vacuum state modifications in general single-field models with a small speed of sound.

4.1 Introduction

Bounds on higher-order non-Gaussian contributions to the observed and approximately Gaussian temperature anisotropy map of the Cosmic Microwave Background are rapidly entering a phase where they can be used to constrain and sometimes even rule out certain (exotic) models of inflation. For instance, bounds on the amplitude of the CMB three-point function, or bispectrum, have already narrowed down the parameter range of string inspired Dirac-Born Infeld models of inflation.

In a separate development it was recently shown that modifications to the inflationary vacuum state can also result in significantly enhanced non-Gaussian signals (Chen et al. 2007; Holman & Tolley 2008; Meerburg et al. 2009, 2010a). In particular the bispectrum can reveal enhanced features that essentially arise due to interactions between excited quanta in the modified vacuum. The power spectrum is also sensitive to deviations from the Bunch-Davies vacuum, but there the corrections are suppressed by the Bogolyubov parameter, whose magnitude is small and typically related to the inflationary Hubble parameter divided by the scale of

new physics, which is already constrained to be smaller than[1] 10^{-1} (Martin & Ringeval 2004, 2005; Easther et al. 2005a; Pahud et al. 2009; Flauger et al. 2010). For the bispectrum one instead encounters an additional enhancement factor, independent of the Bogolyubov parameter, that is in fact proportional to the scale where effective field theory is supposed to break down. The appearance of this enhancement factor can be traced back to the New Physics Hypersurface (NPH) used to define the modified vacuum state. The higher the cutoff scale representing new physics, the more time the fields have before they cross the horizon and freeze, and the more time the interactions effectively have to generate a significant non-Gaussian contribution. This also implies that the (decoupling) limit of taking the cutoff scale to infinity in a modified vacuum is in general ill-defined, producing an infinite bispectrum, whereas the power spectrum yields the standard result in this decoupling limit (assuming the Bogolyubov parameter is suppressed by one over the cutoff scale).

In this work the most recent bispectrum constraints will be applied to a particular and arguably the simplest model of vacuum state corrections, which was first considered by Easther et al. (2001, 2003), but constructed and applied more generally by Danielsson (2002). We will derive to what extent this vacuum state model can be ruled out by the bispectrum data alone in two particular scenarios; a slow-roll single field inflationary model with a higher derivative operator representing the interactions and a non-canonical single field model with a small speed of sound. The bispectrum prediction for a general (NPH) vacuum state modification in these single-field models was computed by Meerburg et al. (2009, 2010a). As an important and hopefully illuminating illustration we will slightly extend and improve the analysis, and apply it to the specific vacuum state proposal of Danielsson.

This chapter is organized as follows. We will start with a short review and small generalization of the Danielsson vacuum state proposal, followed by a summary of the results for the bispectrum in the context of a vacuum state modification in general single-field models of inflation. We then explain how these results can be turned into an *upper bound* on the scale of new physics and to what extent this vacuum state proposal is ruled out or not. We end with our conclusions and prospects for future improvement of the bounds.

4.2 A Minimal Model of Vacuum State Modifications

Let us begin with a short review of Danielsson's original proposal for a modified vacuum state, under the natural assumption that a high energy cutoff scale Λ exists beyond which effective field theory breaks down (Danielsson 2002). Solving for the mode functions in an inflating (pure de Sitter) background one writes down solutions for the field and conjugate momentum operators in terms of creation and annihilation operators in the standard way. Now identify a (conformal) time η_0 serving as the initial time where one would like to define the vacuum state. This initial time η_0 is related to the cutoff Λ by demanding that the physical momentum at η_0 equals the cutoff scale, i.e., $|k\eta_0| = \frac{\Lambda}{H}$. This ensures that the description is always within the effective field theory regime. Note that since we are imposing a cutoff on the physical momentum the initial time η_0 is necessarily a function of the comoving momentum k. With respect to this initial time η_0 the creation operators (and subsequently the annihilation

[1] See ch. 6.

operators) can then be expressed as

$$a_k(\eta) = u_k(\eta)\, a_k(\eta_0) + v_k(\eta) a^{\dagger}_{-k}(\eta_0)\,, \tag{4.1}$$

describing nothing else but the mixing of creation and annihilation operators as time progresses. To ensure the time-independent commutation relations for the creation- and annihilation operators requires that $|u_k|^2 - |v_k|^2 = 1$. Let us now define a natural candidate for a vacuum state at some initial conformal time η_0

$$a_k(\eta_0)\,|0,\eta_0\rangle = 0\,. \tag{4.2}$$

Obviously this choice requires that $v_k(\eta)$ vanishes at $\eta = \eta_0$ and as a direct consequence this candidate vacuum corresponds to a minimal uncertainty state at $\eta = \eta_0$ (Danielsson 2002). This choice of vacuum can be understood as selecting the 'local' empty state at the time the physical momentum equals the high-energy cutoff Λ. This time will be different for different comoving momenta and defines a so-called "New Physics Hypersurface". It can therefore be viewed as the vacuum state that is closest to the Bunch-Davies state in the presence of a high-energy cutoff and can in that sense be regarded as a minimal modification.

To understand the relation of this choice of vacuum to the standard Bunch-Davies state, consider the field operator which should be proportional to the sum of the creation and annihilation operators. In terms of mode-function solutions to the equations of motion $f_k(\eta)$ this implies that

$$f_k(\eta) = N_k\left(u_k(\eta) + v_k^{\star}(\eta)\right)\,, \tag{4.3}$$

where the overall (real) normalization N_k is fixed by the Klein-Gordon normalization condition. Similarly, the expression for the canonical momentum operator defines a function $g_k(\eta)$, which is proportional to the difference between $u_k(\eta)$ and $v_k(\eta)$.

$$g_k(\eta) = \tilde{N}_k\left(u_k(\eta) - v_k^{\star}(\eta)\right)\,. \tag{4.4}$$

Using these relations one can derive the following expression for the function $v_k(\eta)$ in terms of f_k and g_k

$$v_k^{\star}(\eta) = \frac{1}{2}\left[N_k^{-1}\, f_k(\eta) - \tilde{N}_k^{-1}\, g_k(\eta)\right]\,. \tag{4.5}$$

Defining a vacuum state $a_k(\eta_0)\,|0,\eta_0\rangle = 0$ is now explicitly seen to be equivalent to picking a mode-function solution f_k (and its canonically conjugate function g_k) such that $v_k(\eta)$ vanishes at some $\eta = \eta_0$. Bunch-Davies mode-functions have the property that $v_k(\eta)$ only vanishes in the limit $\eta \to -\infty$, i.e., the Bunch-Davies state in this class of vacua corresponds to the minimal uncertainty vacuum state in the infinite past.

Now let us instead consider the general solution, which can be constructed from the Bunch-Davies mode-function and its complex conjugate

$$f_k(\eta) = A_k\, f_k^{BD}(\eta) + B_k\, f_k^{BD\star}(\eta)\,, \tag{4.6}$$

where A_k and B_k are complex coefficients satisfying $|A_k|^2 - |B_k|^2 = 1$. Similarly, the momentum mode-function $g_k(\eta)$ can be expressed in terms of the Bunch-Davies momentum mode-functions

$$g_k(\eta) = A_k\, g_k^{BD}(\eta) - B_k\, g_k^{BD\star}(\eta)\,. \tag{4.7}$$

To find the natural candidate vacua at some fixed time η_0 one demands that $v_k(\eta_0) = 0$. Using (4.5) this gives the following expression for the Bogolyubov rotation parameter $\beta_k \equiv \frac{B_k}{A_k}$ in terms of the Bunch-Davies field and momentum mode-functions

$$\beta_k = \frac{\frac{N_k}{\tilde{N}_k} g_k^{BD}(\eta_0) - f_k^{BD}(\eta_0)}{\frac{N_k}{\tilde{N}_k} g_k^{BD\star}(\eta_0) + f_k^{BD\star}(\eta_0)} \,. \tag{4.8}$$

In terms of the (Bunch-Davies) functions ${u_k}^{BD}(\eta)$ and ${v_k}^{BD}(\eta)$ that appear in the expression for the annihilation and creation operators (4.1) the result simply reads

$$\beta_k = -\frac{v_k^{\star BD}(\eta_0)}{u_k^{\star BD}(\eta_0)} \,. \tag{4.9}$$

A couple of comments are in order. Since the (late time) power spectrum of fluctuations is proportional to $|f_k|^2$, the Bogolyubov rotation parameter β_k denotes the leading order correction to the power spectrum due to a vacuum state modification.Scale invariant vacuum states of course require that β_k is independent of the comoving momentum k, which will be guaranteed by the relation $k\eta_0 = -\frac{\Lambda}{H}$, i.e., η_0 is k-dependent. Also note that this minimal cutoff vacuum construction has so far been completely general, not depending on any details of the inflationary Lagrangian. Whatever inflationary model one is interested in, to construct these states and determine the Bogolyubov parameter β_k one simply plugs in the relevant (mode-) functions in the inflationary case of interest.

The best known (and originally discussed) example is that of the massless scalar field, that is associated to standard slow-roll inflation. The Bunch-Davies field and momentum mode-function in that case read

$$f_k^{BD}(\eta) = \frac{1}{\sqrt{2k}} e^{-ik\eta} \left(1 - \frac{i}{k\eta}\right) \quad , \quad g_k^{BD}(\eta) = \sqrt{\frac{k}{2}} e^{-ik\eta} \,. \tag{4.10}$$

Using the expression (4.8) or (4.9) one then derives the following expression for the Bogolyubov parameter β_k

$$\beta_k = \frac{i}{2k\eta_0 + i} e^{-2ik\eta_0} \,. \tag{4.11}$$

Since $|k\eta_0| = \frac{\Lambda}{H} \gg 1$ this Bogolyubov parameter is indeed scale invariant and approximates to $\beta_k \approx \frac{H}{2\Lambda} e^{i(\frac{3}{2}\pi - \frac{2\Lambda}{H})}$.

As we have emphasized, the construction of these minimal cutoff states is completely general and can just as well be applied in the context of small speed of sound models of inflation, which includes DBI inflation as a special class. One might suggest that in non-canonical models of inflation the introduction of a cutoff scale is not required, since an infinite number of higher order corrections have been re-summed leading to the non-canonical kinetic term. However, these corrections are only a subset of all the possible higher-dimensional operators that are expected to contribute as the string- (or cutoff) scale is approached, motivating the general introduction of a cutoff scale[2]. So the main, and for our purposes only, difference between small

[2]More specifically, Kinney & Tzirakis (2008) showed that in DBI models of inflation under certain conditions the standard procedure for defining the Bunch-Davies vacuum can fail.

speed of sound and slow-roll models of inflation is the introduction of a reduced (and assumed constant) speed of sound $c_s < 1$ of the inflaton fluctuations. For the Bunch-Davies functions u_k^{BD} and v_k^{BD} one can check that the only effect is the appearance of a single factor of the speed of sound c_s in the $k\eta_0$ terms as compared to the standard slow-roll case. The end-result for the Bogolyubov parameter therefore reads

$$\beta_k = \frac{i}{2kc_s\eta_0 + i} e^{-2ikc_s\eta_0} \,, \tag{4.12}$$

which translates into the scale-invariant approximate expression $\beta_k \approx \frac{H}{2c_s\Lambda} e^{i(\frac{3}{2}\pi - \frac{2c_s\Lambda}{H})}$. One observes that the ratio H/Λ governing the magnitude of the corrections in standard slow-roll is modified to $H/c_s\Lambda$ in these models, which could be orders of magnitude bigger. Note that the length scale $1/H^* \equiv c_s/H < 1/H$ corresponds to the sound horizon in small speed of sound models, corresponding to the length scale at which the behavior of the mode-functions turns non-adiabatic. The appearance of the factor $kc_s\eta_0$ is a general feature in these models, which was for instance also observed for the enhancement factors in the three-point function that appear whenever one introduces an arbitrary modified initial state at some high-energy cutoff. We conclude that the same is true for the Bogolyubov parameter in cutoff modified initial state proposals: the magnitude (and phase) depends on the combination $kc_s\eta_0$ which results in larger absolute values of the Bogolyubov parameter for the same value of the cutoff and inflationary Hubble parameter, as compared to standard slow-roll models.

This ends our short summary of minimal cutoff initial states. The results for the three-point function, as reported by Meerburg et al. (2009, 2010a), were model-independent in the sense that they only relied on the presence of a physical momentum cutoff. The Bogolyubov parameter was left as a free parameter and general bounds were derived on its magnitude for different inflationary Lagrangians. In this work we would like to specifically constrain the minimal cutoff modified vacuum state models for which we will need the expressions for the Bogolyubov parameters derived above.

4.3 General Bispectrum Predictions

As already referred to, in our previous work we calculated the results for the three-point function under the assumption of an arbitrary scale-invariant initial state modification in the NPH framework. The presence of a physical high-energy momentum cutoff was assumed, which identified an initial time for each comoving momentum mode where the initial state was defined by introducing a k-independent, but undetermined, Bogolyubov parameter. The presence of such a cutoff alone already has major consequences for the three-point function, also known as the bispectrum in Fourier space. Whenever the Bogolyubov parameter is non-zero, particles are injected at the initial (cutoff) time, which allows potential (self-) interactions to generate a large bispectrum amplitude at the time of horizon crossing. In the calculations these effects show up as large enhancement factors that depend on powers of $\Lambda/H \gg 1$. Depending on the details of the inflationary model under consideration the power of the enhancement factor could be as large as three, resulting in a huge bispectrum amplitude. The only reason why these large non-Gaussian signals have not yet been detected or ruled out is their specific (os-

cillatory) shape, which is orthogonal to any of the non-Gaussian templates used in analyzing the CMB data so far. Proposals for better adapted templates and improved methods to detect or constrain these oscillatory non-Gaussian signals have recently been put forward by Chen (2010c) and Meerburg (2010b), but have not yet been applied to the available data, forcing us to concentrate on the standard non-Gaussian shapes that have been constrained. Projecting the large bispectrum onto any one of the available templates still allows one to derive reasonable constraints in general. In fact, in some cases the constraints are already stronger than those available from the power spectrum. Before discussing the results of the projections to the observational templates it should be pointed out that in principle the results only apply to the three-dimensional bispectrum. However, in general the changes resulting from the reduction to the two-dimensional sphere are minimal and can (partially) be taken into account by introducing a weight function in the three-dimensional analysis (Fergusson & Shellard 2009).

For models with a small speed of sound ($c_s \ll 1$) the projections to the local and orthogonal templates (Senatore et al. 2009), including the phase, were computed analytically (Meerburg et al. 2010a). To derive these results the assumption was made that the modifications to the BD state were strictly oriented in the k_1 direction, introducing a constant $k_1\eta_0$ parameter, breaking the symmetry between the different momenta in the triangle. Properly maintaining the symmetry would instead require introducing three constant parameters $k_i\eta_0$, where k_i is the direction in which one has perturbed the BD vacuum state via a Bogolyubov transformation. Consequently, one can factor out the constant $k_i\eta_0$ for each perturbed state β_{k_i}. As expected, the results of this more complete analysis that preserves the symmetry between triangle momenta differ only slightly. Moreover, we have ameliorated our (numerical) integration methods and used the improved definition of the inner product between shapes as proposed by Fergusson & Shellard (2009), which introduces a weight function to simulate projection effects onto the 2D CMB sky. This increases the correlation between smooth and oscillating shapes as oscillations are slightly damped by the weight function.

General single field models are defined by a single (Lagrangian) function $P(X,\phi)$, where $X = g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi$. From this function $P(X,\phi)$ expressions can be derived for the slow-roll parameters, the speed of sound and two other variables Σ and λ, whose ratio $\frac{2\lambda}{\Sigma}$ will appear in the three-point function (Chen et al. 2007; Seery & Lidsey 2005). Introducing a modified NPH vacuum state the final enhancement factor in the projections turns out to be at best quadratic in $p \equiv kc_s\eta_0 = \Lambda c_s/H \gg 1$. The results, for an undetermined Bogolyubov parameter $\beta = |\beta|e^{i\delta}$, in the limit that $\frac{1}{c_s^2} \gg \frac{2\lambda}{\Sigma}$ and $p \gg 1$ is

$$\Delta f_{\rm NL}^{\rm local} \simeq 4\times 10^{-3}\,\frac{p^2|\beta|}{c_s^2}\,\cos(\delta), \tag{4.13}$$

$$\Delta f_{\rm NL}^{\rm ort} \simeq -6\times 10^{-2}\,\frac{p^2|\beta|}{c_s^2}\,\cos(\delta)\,. \tag{4.14}$$

The additional terms that have been neglected are subleading in p and are out of phase in the sense that the phase parameter δ is shifted with $\pi/2$. Also note that these results do not reduce to the standard slow-roll results in the $c_s = 1$ limit. Subdominant terms that have been neglected in the $c_s \ll 1$ limit will turn in to the dominant contributions to the bispectrum in the $c_s = 1$ limit, explaining why the $c_s = 1$ limit of the above result does not reproduce the

slow-roll bispectrum. It is also important to stress that the DBI models are a special class of small sound speed models for which $2\lambda/\Sigma = 1 - 1/c_s^2$ and as a result the leading contributions to the bispectrum cancel exactly (Seery & Lidsey 2005). The remaining contributions to the projections of the DBI modified NPH initial state bispectrum are only linearly enhanced in the small c_s limit and the details of the projections in the DBI case can only be evaluated semi-analytically (Meerburg et al. 2010a). This leads to DBI constraints that are far less interesting as compared to general small sound speed models. We will therefore concentrate on the general models and not discuss the DBI results in detail.

In the case of standard canonical single-field slow-roll inflation with a specific dimension 8 higher-derivative term (Creminelli 2003) the enhancement factor appearing in the bispectrum amplitude was found to be quadratic in Λ/H (Meerburg et al. 2010a). However, in the projection to the available observational templates $f_{\rm NL}^O = \Delta(O,T)\, f_{\rm NL}^T$, where $\Delta(T,O)$ is the projection factor, one looses one power of the enhancement, leaving a linear enhancement in Λ/H as far as the projections are concerned. In addition the amplitude of the enhancement is further reduced by slow-roll (via ϵ) and the dimensionful coupling of the higher-derivative term. In these results the effects of a (cutoff dependent) phase in the Bogolyubov parameter was neglected and for our purposes here we would like to remedy this situation. To incorporate the phase we had to rely on numerical methods to determine the projection factors. For the projection onto local non-Gaussianities we obtain (in the limit of large $p \equiv \Lambda/H$)

$$\Delta f_{\rm NL}^{\rm local} \;\simeq\; \frac{5}{6}\, p\, \lambda\, \epsilon\, |\beta| \left(\frac{M_{pl}^2 H^2}{\Lambda^4}\right) \cos(\delta), \tag{4.15}$$

which, not surprisingly, is very similar to eq. (6.11) in Meerburg et al. (2009), except for the appearance of a phase dependence.

It is important to realize that the inclusion of the perturbative effect of the dimension 8 higher-derivative operator is only consistent when $\dot\phi/\Lambda^2 \lesssim 1$ (Creminelli 2003), which translates into $\epsilon \geq \Lambda^4/M_{pl}^2 H^2$. The power spectrum observations tell us that $\epsilon = \frac{10^{10}}{8\pi^2}\frac{H^2}{M_{pl}^2}$, which implies that $H/\Lambda < 10^{-2}$ to ensure a consistent higher-derivative expansion. As a consequence the non-Gaussianities due to the presence of the dimension 8 higher-derivative term reach a maximal magnitude for $H/\Lambda \sim 10^{-2}$, yielding an equilateral amplitude of at most $f_{\rm NL}^{\rm equil} \sim \mathcal{O}(1)$. On the other hand, modifying the initial state leads to an additional enhancement factor that is proportional to the ratio between the UV cutoff Λ and the scale of inflation H. The required consistency of the higher-derivative expansion, being inversely proportional to the cutoff scale Λ, reduces this enhancement and as a consequence weakens the constraints on β. In fact, as we will see in the next section, for the minimal cutoff vacuum state proposal assumed in this chapter the constraint becomes practically meaningless. This should not come as much of a surprise, since in order to boost non-Gaussianities from higher-derivative operators the cutoff scale needs to be relatively close to the inflationary scale, whereas the non-Gaussian amplitude as a consequence of initial state modifications grows as the the hierarchy between the cutoff and the Hubble scale increases. For completeness, let us state that the correlation between the orthogonal template and the standard single-field higher-derivative bispectrum, being very similar, yields comparable constraints.

After this very short review and slight improvement of what is known about bispectrum projections due to a generic NPH initial state, let us now derive the best available constraints on the parameters of minimal cutoff initial states.

4.4 Bispectrum Constraints on Minimal Cut-Off Vacuum States

Having summarized the results for minimal cutoff initial states and the general projected non-Gaussian amplitudes due to arbitrary cutoff (or NPH) modified vacua, we would now like to apply the most recent observational bounds to derive constraints on the parameters of the minimal cutoff initial state proposal.

To proceed we will make use of the following recent (2 σ) constraints on $f_{\rm NL}^{\rm local}$ and $f_{\rm NL}^{\rm ort}$ (Senatore et al. 2009; Komatsu et al. 2011)

$$\begin{aligned} -10 \quad &\leq f_{\rm NL}^{\rm local} \leq \quad 74, \\ -410 \quad &\leq f_{\rm NL}^{\rm ort} \leq \quad 6. \end{aligned}$$

The observational constraints on the equilateral amplitude $f_{\rm NL}^{\rm equil}$ can be ignored, because the projection to the equilateral template turns out to be orders of magnitude smaller. One should keep in mind that the power spectrum constraint on $|\beta|$ is roughly 10^{-1}, which is based on the lack of evidence for oscillatory behavior with a larger amplitude in the power spectrum (Martin & Ringeval 2004, 2005; Easther et al. 2005a; Pahud et al. 2009; Flauger et al. 2010).

A notable feature of the results for the projections is that they oscillate as a function of the phase δ of the Bogolyubov parameter (4.13–4.14). Different values for the phase can therefore result in constraints that deviate considerably. Indeed, for special values of the phase the projection vanishes and the constraints disappear altogether[3]. In minimal cutoff models the phase of the Bogolyubov parameter is actually a function of the cutoff scale Λ (4.11–4.12) and as a consequence for special values of Λ the projected amplitudes would be significantly reduced, as can be seen in fig. (4.1). Since the precise value of the cutoff Λ is unknown, the constraints will be derived under the reasonable assumption that a typical value of Λ is expected to be closer to an (absolute value) maximum in the projection than to the special point where the projection vanishes. To implement this we will derive the constraints based on an expression that is related to the absolute value average over a single oscillatory domain in the parameter Λ.

Let us first analyze general single field models with a small speed of sound. The minimal cutoff vacuum state proposal predicts a Bogolyubov parameter with an absolute value approximately equal to $|\beta_k| \approx \frac{H}{2c_s\Lambda}$ and a phase $\delta = i(\frac{3}{2}\pi - \frac{2c_s\Lambda}{H})$. In the limit $p \equiv \Lambda\, c_s/H \gg 1$ we derive, to leading order in p,

$$\Delta f_{\rm NL}^{\rm local} \simeq -2 \times 10^{-3}\, \frac{p}{c_s^2}\, \sin(2p), \tag{4.16}$$

$$\Delta f_{\rm NL}^{\rm ort} \simeq 3 \times 10^{-2}\, \frac{p}{c_s^2}\, \sin(2p). \tag{4.17}$$

[3]To be more precise, for those values one can no longer neglect the subleading terms in p, which are in effect $\pi/2$ out of phase and therefore maximize exactly when the leading terms minimize. Nevertheless, the fact that they are subleading in p implies that the constraints for those values of p become practically meaningless.

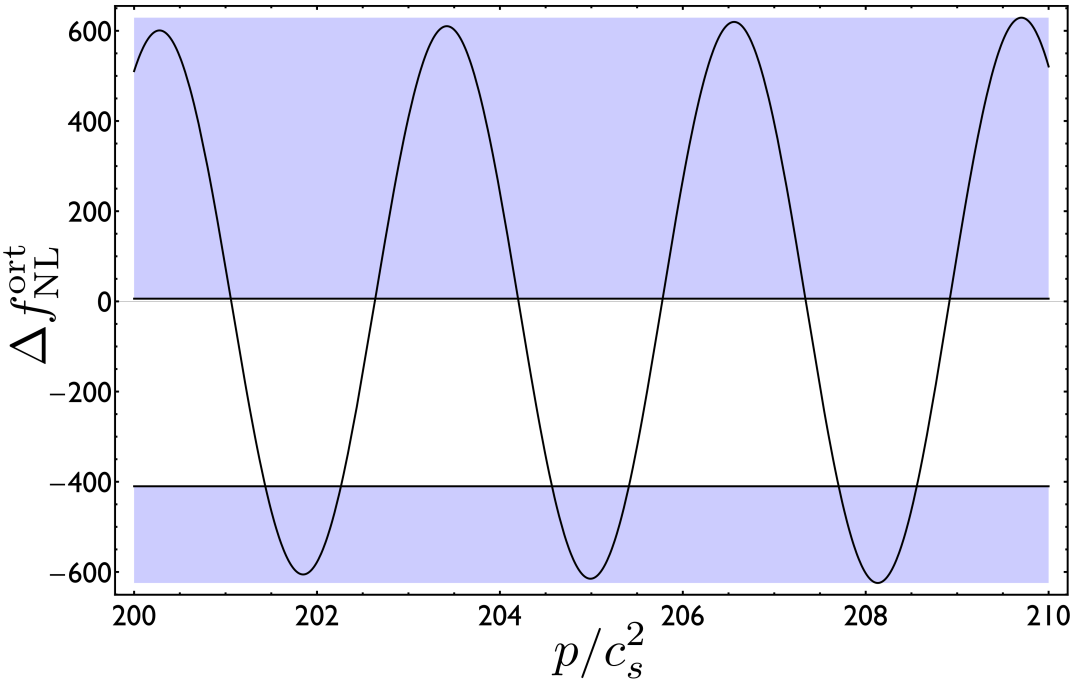

Figure 4.1: Contribution of initial state modifications to orthogonal non-Gaussianities for various values of p/c_s^2. The shaded upper and lower parts in the plot correspond to the observationally excluded regions.

The absolute value of the Bogolyubov coefficient has removed one enhancement factor, resulting in a linear p dependence of the prefactor in front of the oscillatory $\sin(2p)$ term. As an example, the maximum contribution to the orthogonal non-Gaussian template from initial state modifications has been plotted in fig. 4.1. As already alluded to, the oscillations in the projection factor force us to work with the average over a single oscillatory domain in order to derive a constraint on the ratio Λ/H. Simply replacing the $\sin(2p)$ with the average $2/\pi$ the derived bounds would only apply to about 60 percent of the p domain. If instead demanding that at least 90 percent of the domain is included within the constraints, we need to replace $\sin(2p)$ with ~ 0.16.

The ambiguous sign due to the oscillatory behavior also prompts us to consider the bound on the (largest) absolute value of the non-Gaussian amplitude, which for the local amplitude is $|f_{\rm NL}^{\rm local}| \leq 74$ and for the orthogonal type $|f_{\rm NL}^{\rm ort}| \leq 410$. Using these largest absolute values for the local and orthogonal amplitude and insisting on a 90 percent coverage, an upper bound on the parameter $\frac{\Lambda}{H}$ can be derived from local and orthogonal type non-Gaussianities

$$\text{local}: \quad \frac{\Lambda}{H} \leq 2.5 \times 10^5 \, c_s \quad , \quad \text{orthogonal}: \quad \frac{\Lambda}{H} \leq 8.5 \times 10^4 \, c_s. \tag{4.18}$$

Clearly the strongest result is obtained from the orthogonal bound, becoming stronger as c_s is smaller. In fact, the speed of sound c_s cannot be much smaller than 10^{-2} because that would

produce a leading order equilateral non-Gaussian result in conflict with observations. Applying that minimal value of c_s the upper bound on $\frac{\Lambda}{H}$ roughly equals

$$\frac{\Lambda}{H} \leq 10^3\,, \tag{4.19}$$

Again we would like to stress that this is an *upper* bound. The cutoff scale should be close to the inflationary Hubble scale to avoid too large an enhancement of the (projected) non-Gaussian signal, with the exception of special higher values of Λ for which the phase δ is close to an integer number times π, corresponding to 10 percent of the parameter domain. This constraint is weaker than initially anticipated (Meerburg et al. 2010a), mainly because the phase in this case is a function of Λ, producing oscillations in the non-Gaussian signal, forcing us to introduce a factor related to the average of the absolute non-Gaussian signal and the most conservative (upper) bound on the absolute value of the observed constraints.

Incorporating the, admittedly crude, *lower* bound on $\frac{\Lambda\, c_s}{H} > 10^2$ from analysis of the power spectrum, for a general speed of sound one arrives at

$$\frac{10^2}{c_s} \leq \frac{\Lambda}{H} \leq 8.5 \times 10^4\, c_s. \tag{4.20}$$

There is a small window remaining for $c_s \sim 10^{-1}$, but for small speed of sound models with $c_s \sim 10^{-2}$ one can conclude that minimal initial state modifications are practically ruled out, except for those special values of the phase at which the contribution to the orthogonal non-Gaussian template (nearly) vanishes, corresponding to about 10 percent of the parameter domain.

To briefly illustrate the impact of the oscillatory behavior in this analysis, consider neglecting the oscillating behavior, choosing $\sin 2p = 1$ for local non-Gaussianities and $\sin 2p = -1$ for orthogonal non-Gaussianities. Again this would yield a strongest *upper* bound from the orthogonal constraints that is roughly equal to $\Lambda/H \leq 2 \times 10^2\, c_s$, which for $c_s \sim 10^{-2}$ can essentially be ruled out from the non-Gaussian data alone and would in general be inconsistent with the power spectrum data that shows no evidence of an oscillatory effect with an amplitude that large.

As already commented on, for DBI models the constraints weaken considerably, because the enhancement of the non-Gaussian signal is reduced by a factor of p. This results in no enhancement at all in the projection to the local (or orthogonal) template, removing all dependence on the cutoff scale. As a a consequence the observational constraints do not result in (interesting) bounds on the cutoff scale Λ, which parameterizes the initial state modification.

Finally, let us briefly consider standard slow-roll inflation ($c_s = 1$), including a dimension 8 higher derivative term that, together with the initial state modification, is responsible for a large p^2 enhanced non-Gaussian signal. Inserting $|\beta|$ and the phase into eq. (4.14) one obtains

$$f_{\rm NL}^{\rm local} \;\simeq\; -\frac{5}{12}\sin(2p). \tag{4.21}$$

All enhancement in $p \equiv \Lambda/H$ is lost and one is left with a maximum contribution to the local non-Gaussian template that is of (less than) order 1, which is significantly below set constraints on local type non-Gaussianities. As a consequence minimal initial state modifications

in this particular canonical single field inflationary model are not constrained by the local non-Gaussian projection. In order to become sensitive to the intrinsically large non-Gaussian signal set by the ratio Λ/H, templates better adapted to the theoretical non-Gaussian shape should be designed and compared to the available data.

4.5 Conclusions

To summarize, we have applied the currently best available constraints on local, orthogonal and equilateral shape non-Gaussianities to derive bounds on the cutoff scale parameterizing a (slightly generalized) minimal model of initial state modifications. The results strongly depend on the model of inflation under consideration. Interesting bounds can be derived in the context of general, non-DBI, single field small speed of sound models of inflation. In that case, non-Gaussian constraints from the orthogonal template alone lead to the following *upper* bound on the cutoff scale

$$\frac{\Lambda}{H} \leq 8.5 \times 10^4 \, c_s \, . \tag{4.22}$$

Combined with results from the equilateral template and the power spectrum leaves only a very small window ($c_s \sim 10^{-1}$, $\Lambda/H = 10^3$), and minimal initial state modifications in general single field models with a small speed of sound can almost be ruled out. To be more precise, the above upper bound is valid for 90 percent of the cutoff parameter domain, excluding small isolated regions covering in total 10 percent of the parameter domain where the projection to the orthogonal template nearly vanishes as a consequence of the oscillating nature of the projection.

For DBI and a canonical single field model with a dimension 8 higher derivative operator the analysis does not lead to interesting non-Gaussian constraints. The reason is the complete lack of sensitivity of the available templates to the large non-Gaussian signal of minimal initial state modifications. Consequently, more suitable templates have to be developed to efficiently probe (minimal) initial state modifications in these models (Chen 2010c; Meerburg 2010b). Work in this direction is ongoing and we hope to report on more effective non-Gaussian templates and observational strategies in the near future.

An important property that we have emphasized is that (minimal) initial state modifications lead to a non-Gaussian signal that increases as the separation between the cutoff scale and the inflationary Hubble parameter increases. As a consequence non-Gaussian constraints can only give rise to an upper bound on the cutoff scale parameterizing initial state modifications arising due to new physics. Together with lower bounds on the cutoff scale from other considerations (including the power spectrum data or generic effective field theory arguments) this carries a strong potential to rule out these types of initial state modifications in the future as the non-Gaussian analysis and corresponding constraints improve.

At the same time the interesting feature of an upper bound on the cutoff scale poses a theoretical conundrum, since it is obviously in conflict with the idea of decoupling. Perhaps this should be interpreted as signaling a fundamental flaw of these type of initial state modifications that might be responsible for inconsistencies in the perturbative expansion of the quantum field theory, conceivably ruling out these vacuum states. We hope to come back to this important issue in future work.

Acknowledgments

We would like to thank Pier Stefano Corasaniti for very beneficial and fruitful discussions. PDM and JPvdS were supported in part by a van Gogh collaboration grant VGP 63-254 from the Netherlands Organization for Scientific Research (NWO). PDM is supported by the Netherlands Organization for Scientific Research (NWO) through a NWO-toptalent grant 021.001.040. The research of JPvdS is financially supported by Foundation of Fundamental Research on Matter (FOM) grant 06PR2510.

Chapter 5

Oscillations in the Primordial Bispectrum: Mode Expansion

P. Daniel Meerburg
Phys. Rev. D **82** (2010) 063517

Abstract: We consider the presence of oscillations in the primordial bispectrum, inspired by three different cosmological models; features in the primordial potential, resonant type non-Gaussianities and deviation from the standard Bunch Davies vacuum. In order to put constraints on their bispectra, a logical first step is to put these into factorized form which can be achieved via the recently proposed method of polynomial basis expansion on the tetrahedral domain. We investigate the viability of such an expansion for the oscillatory bispectra and find that one needs an increasing number of orthonormal mode functions to achieve significant correlation between the expansion and the original spectrum as a function of their frequency. To reduce the number of modes required, we propose a basis consisting of Fourier functions orthonormalized on the tetrahedral domain. We show that the use of Fourier mode functions instead of polynomial mode functions can lead to the necessary factorizability with the use of only 1/5 of the total number of modes required to reconstruct the bispectra with polynomial mode functions. Moreover, from an observational perspective, the expansion has unique signatures depending on the orientation of the oscillation due to a resonance effect between the mode functions and the original spectrum. This effect opens the possibility to extract information about both the frequency of the bispectrum as well as its shape while considering only a limited number of modes. The resonance effect is independent of the phase of the reconstructed bispectrum suggesting Fourier mode extraction could be an efficient way to detect oscillatory bispectra in the data.

5.1 Introduction

In recent years it has become evident that determining the precise physics of inflation requires the observation of higher order correlation functions beyond the power spectrum (Komatsu

et al. 2009). These correlation functions can be obtained from the Cosmic Microwave Background (CMB) (Komatsu et al. 2009, 2011; Senatore et al. 2009; Smidt et al. 2010b; Smith et al. 2009) and Large Scale Structure (LSS) (Xia et al. 2010; Verde 2010), but recently (Cooray 2006; Pillepich et al. 2007; Cooray et al. 2008), it has been shown that in principle 21-cm observations of the early universe can also be used to measure n-point statistics. Because higher order correlation functions introduce more free parameters they can be used to constrain more complex models of inflation, since an increased set of parameters will allow for a unique fitting of the model to the observed data (Komatsu et al. 2009). However, both due to computational and observational limitations, only the bispectrum has been reasonably investigated. For the detection of higher order correlations we will have to wait for more advanced data sets, such as Planck and improved analysis methods, although preliminary attempts have been made (Smidt et al. 2010a,b; Regan et al. 2010). Even the detection of the bispectrum is not optimal, as a bispectrum would at least be a continuous three parameter observable but thus far only constraints have been set on limiting cases, in which 2 of the parameters are fixed and the third one is measured for a predetermined triangular configuration. The limiting cases (shapes) are known as the local, equilateral and orthogonal (and in the context of limiting triangular configurations; enfolded) non-Gaussian features. Precisely these features have been chosen, as it has been shown theoretically that most models of inflation produce non-Gaussianities that fall in one of these three classes (for recent reviews see (Chen 2010; Bartolo et al. 2010; Komatsu 2010)).

When constraining non-Gaussianities using the bispectrum, it has been a prerequisite that the comoving momentum dependence should be factorizable; the bispectrum should be separable into a product of functions of one variable, each variable being one of the three comoving momenta making up the connected correlation triangle. Foremost, this requirement is set because of computational limitations that would render the analysis intractable if a given primordial bispectrum is not of the factorized form. The number integrals and sums one has to perform when computing an unfactorized bispectrum scale with the number of pixels as $N^{5/2}$, while for factorizable shapes this reduces by one factor of N (Wang & Kamionkowski 2000). Although one integral can be computed fairly quickly the number of pixels ($\mathcal{O}(10^6)$ for WMAP and $\mathcal{O}(10^7)$ for Planck) is large and one factor of N can make all the difference. The constrained bispectra, local, equilateral and orthogonal, have thus far been factorized templates. In case of equilateral (Creminelli et al. 2006) and orthogonal (Senatore et al. 2009) these have been constructed via approximation of a predicted signal, in the local case, the template is a direct representation of the theory (Gangui et al. 1994; Komatsu & Spergel 2001; Maldacena 2003). For a particular type of bispectrum to be constrained, it is necessary to construct a factorized template that 'matches' the bispectrum. Until recently, there was no given prescription how to factorize a given theoretical bispectrum. Fergusson & Shellard (2009); Fergusson et al. (2010a) showed that factorizability can be achieved in both comoving momentum and multipole space by expanding the bispectrum in mode functions that are orthogonal on the domain of the bispectrum dictated by triangle constraints. The purpose of this factorization is to be able to quickly compute the full CMB bispectrum ($B_{l_1 l_2 l_3}$) and generate CMB maps with a arbitrary primordial statistics (up to the trispectrum (Regan et al. 2010)) which are used to determine the variance of the (statistical) estimator. In the same paper, it was also shown that one can

efficiently extract information about non-Gaussianity in the observed CMB by measuring the weight of each mode in the data and comparing this to theoretical predictions[1].

In this chapter we investigate how well this mode expansion works for a class of bispectra that contain (a large number of) oscillations. The reason to be interested in such features is that a number of theoretical models (Bean et al. 2008; Chen et al. 2007, 2008; Meerburg et al. 2009, 2010a; Flauger et al. 2010) predict oscillations in the bispectrum and in order to be able to constrain such models, a plausible first step is to factorize these bispectra. As it is, oscillations can be considered as an extra, distinguishable, degree of freedom within the bispectrum which could result in narrowing down the number of potential scenarios of inflation.

We introduce three different cosmological scenarios in which oscillations in the bispectrum can appear. We will briefly discuss the theory behind these models and show to what extend these would be distinguishable from one another in the data in sec. 5.2. Two out of three bispectra can have significant correlation and it could be difficult to discriminate between such models in future surveys. We will discuss the method of polynomial expansion in order to rewrite the primordial bispectra in factorized/separable form in sec. 5.3. As expected, the number of modes required in the expansion grows along with the frequency of the theoretical spectra. In sec. 5.3.1 we show how fast polynomial expansion would yield a reasonable reconstruction of the given bispectra predicted by the three cosmological scenarios. Subsequently we will investigate another set of modes that can lead to a separable expansion of the theoretical bispectrum in sec. 5.3.2. These modes are based on the sine and cosine and the resulting set of orthonormal functions can be considered a Fourier-type basis on the tetrahedral domain. After detailing the construction of this set of orthonormal mode functions, we will compare the number of modes required to achieve comparable correlation with the polynomial mode expansion. We find out that this number is reduced significantly and as such Fourier expansion can be considered a reasonable alternative to expand oscillatory spectra. For larger frequencies both Fourier and polynomial mode expansion become inefficient. Fortunately, for various oscillatory signals only a limited number of modes contribute significantly in the reconstruction of the original spectrum. This has several consequences for the viability of Fourier mode expansion as well as possible observational advantages compared to polynomial modes, which will be discussed in sec. 5.4. In this class of models, just as the frequency, the phase can be considered a free parameter of the theory. In a polynomial mode expansion, different phases can result in significantly different expansions. In a Fourier mode expansion the phase is taken care of much more naturally. Effectively the phase can be absorbed into the weights of the expansion, and as such have minimal effect on the overall expansion. Consequently, we will see that the norm of the mode expansion coefficients will be very similar for each phase, making Fourier expansion much more elegant and suitable for this type of spectra. We conclude this chapter in sec. 5.5.

[1]During the finalization of this paper, the same group published a paper (Fergusson et al. 2010a) in which many non-factorizable non-Gaussian shapes have been constrained using the WMAP 5 year data and the method of mode expansion.

5.2 Oscillations in Primordial Bispectra

In this section we will briefly discuss 3 distinct possibilities that can produce non-Gaussianities that have an oscillatory component. Two of these examples have an exact solution, while a third has only been solved numerically and we will use an approximate form. In the following paragraphs we will describe the physics behind these models and quote their theoretically predicted primordial bispectra. In addition we investigate how well these bispectra can be distinguished from one another by computing their correlation, which will be defined shortly. Since all these bispectra have poor overlap with existing spectra, there exists substantial room for improvement, which we could achieve by approximating these shapes via mode expansion. This will be the topic of the next section.

For completeness, let us introduce (standard) notation. The primordial bispectrum is given by

$$\langle \zeta_{\vec{k}_1} \zeta_{\vec{k}_2} \zeta_{\vec{k}_3} \rangle = (2\pi)^7 f_{\rm NL} \Delta^2 \delta^K \left(\sum_{i=1}^{3} k_i \right) F(k_1, k_2, k_3), \tag{5.1}$$

where ζ is the gauge invariant curvature perturbation ($\zeta = -H\delta\phi/\dot{\phi}_0$) which is constant after horizon exit, Δ is the amplitude of the primordial power spectrum (i.e., for single field slow-roll $\Delta = H^2/8\pi\epsilon$, where H is the Hubble rate at the end of inflation and ϵ the slow-roll parameter) and $F(k_1, k_2, k_3)$ is the shape of the bispectrum. We will also make use of $S \equiv k_1^2 k_2^2 k_3^2 F$. In the following we will discuss the shapes of the bispectra and quote theoretically predicted ranges of their associated $f_{\rm NL}$. We would like to refer to the literature for a detailed examination of the theoretically predicted values of $f_{\rm NL}$(Chen et al. 2008; Meerburg et al. 2009, 2010a; Flauger et al. 2010) in various theoretical contexts.

5.2.1 Features in the Potential

Sharp features in the potential can temporarily break slow-roll and produce large non-Gaussianities (Chen et al. 2007, 2008). As long as the system relaxes within several Hubble times, inflation can still lead to a significant amount of e-folds to solve the standard cosmological problems. The motivation for these type of features is two-fold. First, there are hints of glitches in the primordial power spectrum that could be cross-checked using the bispectrum (Covi et al. 2006). A second motivation is theoretical in nature. In certain brane inflation models the effective 4-dimensional potential displays sharp features (see Chen (2010) and references therein).

One of the possible sharp features is a step in the potential, which can be parameterized as

$$V(\phi) = \frac{1}{2} m^2 \phi^2 \left[1 + c \tanh\left(\frac{\phi - \phi_s}{d} \right) \right], \tag{5.2}$$

where c, d and ϕ_s respectively determines the height, width and location of the feature.

The resulting bispectrum can only be computed numerically. Chen et al. (2008) have proposed an approximate analytic form

$$F_{Feat} \simeq \frac{\sin(k_t/k_* + \delta)}{k_1^2 k_2^2 k_3^2}. \tag{5.3}$$

The approximation can in principle be improved (Chen 2010) by multiplying by an 'envelope' function, but such improvement would not gain us any more useful insight required for the analysis in this chapter and we will therefore omit it. Here k_* is related the location of the feature in the potential ϕ_s. Evidence for features in the power spectrum around $l \sim 30$ have been put forward by Covi et al. (2006). It was shown that the inclusion of features in the primordial potential could improve the χ^2 best-fit. Such a feature would approximately correspond to $k_* = 30/\eta_0 \sim 0.002\text{Mpc}^{-1}$. This relation also indicates that the smaller the scale at which the feature appears the larger the associated wavelength. Roughly the wavelength corresponds to the location of the feature, e.g. for a feature at $l = 30$ the wavelength $\delta l \sim 30$ (Bean et al. 2008). Here we do not necessarily relate to an observed feature at a specific value in multipole space since features that lead to non-vanishing bispectra can still be present with minimal consequences for the observable power spectrum. The quantities we will compute in the remainder of this chapter are mostly integrals that run over the domain of comoving momentum space between $k_{min} \leq k \leq k_{max}$. It is therefore convenient to choose our reference scale $k_{max} \sim 10^{-1}\text{Mpc}^{-1}$, the smallest observable scale in the data, in order to be able to compare the frequencies in the various models. We then define $x_1 = k_1/k_{max}$, $x_2 = k_2/k_{max}$, $x_3 = k_3/k_{max}$, $x_t = k_t/k_{max}$ and rewrite the shape of this bispectrum as

$$F_{Feat} = k_{max}^{-6} \frac{sin(\omega_f x_t + \delta)}{x_1^2 x_2^2 x_3^2}, \tag{5.4}$$

with $\omega_f = k_{max}/k_*$. For a feature at $k_* = 0.002\text{Mpc}^{-1}$ we therefore find $\omega_f \sim 50$. Note that $\omega = 50$ can be considered an upper limit in allowable frequencies due to features in the potential. For features at smaller scales the frequency will be smaller. This bispectrum with a frequency of $\omega_f = 50$ is shown in the bottom of fig. 5.1.

The amplitude of this type of non-Gaussianity is governed by the width and the depth of the feature in the potential

$$f_{\text{NL}}^{feat} \sim \frac{7c^{1/2}}{d\epsilon}, \tag{5.5}$$

which for a feature at $l \sim 30$ would imply $f_{\text{NL}}^{feat} \sim \mathcal{O}(10)$ (Covi et al. 2006).

5.2.2 Resonant non-Gaussianity

This type of non-Gaussianity is a result of a periodic feature in the inflaton potential as apposed to a sharp feature explored in the previous example. These features will cause oscillations in the coupling(s) of the interaction terms of the inflaton field. Resonance occurs when an oscillatory mode well within the horizon grows during inflation until its frequency hits the same frequency as those of the couplings. So as long as $\omega > H$ resonance will occur at some point within the inflationary history of the mode. This resonance can result in a large contribution to the three point correlation function (Chen et al. 2008).

In a general scenario, with an oscillatory potential we obtain an expression for the bispectrum

of the form (Chen et al. 2008; Flauger et al. 2010)

$$F_{res} = \frac{1}{k_1^2 k_2^2 k_3^2}\left(\sin(C\ln(k_t/k_*)) + C^{-1}\cos(C\ln(k_t/k_*))\sum_{i\neq j}\frac{k_i}{k_j}\right) \tag{5.6}$$

Here C is related to the frequency as $C = \omega/H$ with H the Hubble rate during inflation (which is approximately constant) and k_* introduces a phase. One can also compute the general expected amplitude of non-Gaussianity which is related to the frequency as

$$f_{\rm NL}^{res} \sim \frac{\sqrt{\pi}}{2\sqrt{8}}\frac{\omega^{1/2}\dot{\eta}_A}{H^{3/2}}. \tag{5.7}$$

Here η_A represents the amplitude of the oscillatory component of the couplings.

Physically such features might be realized in terms of brane inflation (Bean et al. 2008) where the periodic feature comes from a duality cascade in the warped throat, as well as axion-monodromy inflation where the periodic feature is a result of instanton effects (Flauger et al. 2010; Flauger & Pajer 2011). As an example let us consider the latter. Axion inflation is well embedded in string theory and represents a favorable candidate for inflation if the observed tensor modes are relatively large ($r \sim 0.07$). Such a scenario implies inflation occurred at energies close to the GUT scale and would indicate that we require the knowledge of the UV completion.

The axion potential is given by

$$V(\phi) = V_0(\phi) + \Lambda^4\cos\phi/f. \tag{5.8}$$

The parameter f represents the axion decay parameter. The range of f which would generate observable non-Gaussianities and is still consistent with observations of the power spectrum is given by $10^{-4} \lesssim f \lesssim 6\times10^{-3}$ (Flauger et al. 2010; Flauger & Pajer 2011) . The lower bound is set by the requirement that the period of the oscillation should be larger than $\Delta l \sim 1$ for $l \equiv 200$. For a linear zero order potential the resulting bispectrum is then given by

$$F_{res} = \frac{k_{max}^{-6}}{x_1^2 x_2^2 x_3^2}\left(\sin(\omega_r\ln x_t + \gamma_1) + \omega_r^{-1}\cos(\omega_r\ln x_t + \gamma_1)\sum_{i\neq j}\frac{x_i}{x_j}\right) \tag{5.9}$$

with $\omega_r = (f\phi_*)^{-1}$ and $\gamma_1 = \omega_r\ln k_{max}/k_*$. Here k_* is pivot scale ($k_* = 0.002{\rm Mpc}^{-1}$), and ϕ_* is the value of the inflaton field when the pivot scale exits the horizon and is of order 10 (M_p). Given the range for $f\phi_*$ the frequency of the oscillations in the bispectrum lie within $20 \lesssim \omega_r \lesssim 10^3$. A plot of this shape is shown in the top right fig. 5.1.

The amplitude of the axion bispectrum (for a linear potential) is given by

$$f_{\rm NL}^{res} = \frac{3\sqrt{2\pi}b}{8(f\phi_*)^{3/2}}. \tag{5.10}$$

The amplitude is therefore proportional to a power of the frequency. For a linear potential $b = \Lambda^4/(\mu^3 f)$, where $\mu \sim 6\times10^{-4}$ is fixed by COBE normalization. From observations of the power spectrum one can constrain $bf < 10^{-4}$ (Flauger et al. 2010) and therefore

$$f_{\rm NL}^{res} \sim 10^{-3}\omega_r^{5/2}, \tag{5.11}$$

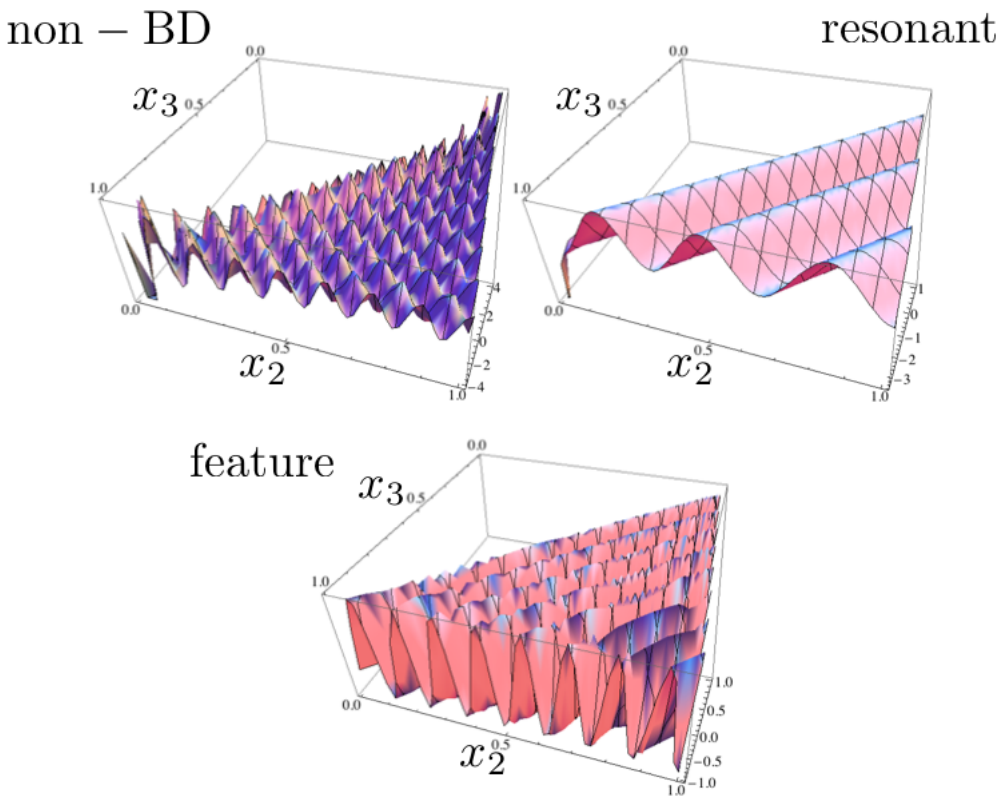

Figure 5.1: 3 examples of oscillating bispectra. We have set $\omega_v = \omega_f = \omega_r = 50$. The pivot scale in eq. (5.9) is set to $k_* = 0.002\text{Mpc}^{-1}$ and $x_1 = 1 \rightarrow k_1 = k_{max}$. In addition $\delta = 0$ in both the modified initial state and feature scenario. The non-BD bispectrum contains the most features and, not clear from this image, the number of features (effective frequency) increases rapidly for smaller values of x_1 making this bispectrum particularly hard to reconstruct using mode expansion.

allowing $\mathcal{O}(1) \leq f_{\text{NL}}^{res} \leq \mathcal{O}(10^4)$.

5.2.3 Initial State Modifications

Since inflation is an effective field theory in a curved background, choosing an appropriate vacuum state is by no means evident[2]. In general the initial or vacuum state is chosen to be equivalent to the free field vacuum state in flat Minkowski space, know as the Bunch Davies (BD) vacuum. Although it seems that possible corrections to this assumption are constrained to be small (from general observation of the power spectrum (Easther et al. 2005b; Flauger et al. 2010; Pahud et al. 2009) and backreaction constraints (Schalm et al. 2005; Holman & Tolley 2008)), it has been shown that small corrections in the BD state can result in rather large

[2]For a in depth discussion on deviations from a BD vacuum see for example Meerburg (2010b); Greene et al. (2004) and Chen (2010) sec. 6.4

non-Gaussian effects (Chen et al. 2007; Holman & Tolley 2008; Meerburg et al. 2009, 2010a). Using the currently available bounds on non-Gaussianity from CMB data, deviations from a pure Bunch Davies state have been constraint even further, although these constraints strongly depend on the inflationary model. However, there exist significant room for improvement as non-Gaussianities from these modifications are highly oscillatory and therefore the derived constraints are relatively poor since they depend on the correlation with measured smooth bispectra.

A number of different scenarios have been considered in which initial state modifications were investigated. Here we will not discuss all of these, although the results can differ significantly (Meerburg et al. 2010a). Such differences make it difficult to make robust predictions, it seems inevitable however that once you introduce a effective field theory cutoff, oscillations appear in both the power and bispectrum. We will consider one example that represents a large class of models with a non-canonical effective field theory action, which already drives large non-Gaussianities to start with. This particular class has a speed of sound $c_s < 1$, such that perturbations in the medium propagate slow compared to the growth of the causal horizon. The leading order shape of the resulting bispectrum is given by Meerburg et al. (2009)

$$F_{BD} = \frac{c_s\eta_0}{k_1k_2k_3}\sum_j\left(\frac{1}{2}\frac{\cos(\tilde{k}_jc_s\eta_0+\delta)}{\tilde{k}_jc_s\eta_0}-\frac{\sin(\tilde{k}_jc_s\eta_0+\delta)}{(\tilde{k}_jc_s\eta_0)^2}+\frac{\cos\delta-\cos(\tilde{k}_jc_s\eta_0+\delta)}{(\tilde{k}_jc_s\eta_0)^3}\right) \tag{5.12}$$

Here $\tilde{k}_j = k_t - 2k_j$. Meerburg et al. (2009) assumed that there exists a fixed physical cutoff hyper-surface η_0 that is scale dependent such that the overall momentum dependence of the bispectrum becomes scale invariant. Such a choice is known as the New Physics Hypersurface (NPH), as apposed to Boundary Effective Field Theory (BEFT) approach in which the cutoff is time dependent (Greene et al. 2004). The subtlety is that the cutoff appears due to the presence of a non-BD state in each direction in comoving momentum space. Consequently, $\eta_0(k_i)$ will depend on the k_i direction the BD vacuum has been perturbed in. This direction is set by the direction in which k_i picks up a minus sign due to the Bunch Davies vacuum perturbation as explained by Meerburg et al. (2009). One could allow for scale invariance breaking and consider BEFT, however there are some suggestions (Sefusatti et al. 2009) that such large scale invariance should have been observed already. We can rewrite the bispectrum as

$$\begin{aligned} F_{nBD} = \frac{\omega_v k_{max}^{-6}}{x_1x_2x_3}\sum_j\frac{1}{x_j^3}\Bigg(&\frac{1}{2}\frac{\cos\left(\omega_v\frac{x_{j+1}+x_{j+2}}{x_j}+\gamma_2\right)}{\omega_v\left(\frac{x_{j+1}+x_{j+2}}{x_j}-1\right)}-\frac{\sin\omega_v\left(\omega_v\frac{x_{j+1}+x_{j+2}}{x_j}+\gamma_2\right)}{\omega_v^2\left(\frac{x_{j+1}+x_{j+2}}{x_j}-1\right)^2}\\ &\frac{\cos\delta-\cos\left(\omega_v\frac{x_{j+1}+x_{j+2}}{x_j}+\gamma_2\right)}{\omega_v^3\left(\frac{x_{j+1}+x_{j+2}}{x_j}-1\right)^3}\Bigg), \end{aligned} \tag{5.13}$$

where $\gamma_2 = \delta - \omega_v$ and $\omega_v = k\eta_0c_s = (k/a_0)/(H/c_s)$ or the ratio between the largest physical momentum scale and the Hubble radius at time η_0 which can be as large as 10^3 (Meerburg

et al. 2009, 2010a; Holman & Tolley 2008). Note that from this expression it seem that $x_j = x_{j+1} + x_{j+2}$ represents a singular line (the enfolded limit). However, one can show that all infinities are cancelled against each other and the the expression is finite and vanishing[3] When computing quantities numerically, such as the correlator in sec. 5.3, these apparent singularities can be hard to handle and we need to be aware of these. We have plotted this shape in the top left fig. 5.1.

The amplitude of the non-BD bispectrum is a function of the frequency and the Bogoliubov parameter quantifying the deformation away from the BD state.The way this bispectrum was computed, considered a Bogolyubov correction of linear order β and small speed of sound c_s. In this particular scenario, $f_{\rm NL}$ is roughly given by

$$f_{\rm NL}^{nBD} \sim \frac{1}{c_s^2}\omega_v^3\beta. \tag{5.14}$$

From backreaction and power spectrum constraints $\beta \lesssim 10^{-2}$, which could still allow observable levels of non-Gaussianity.

5.2.4 Distinguishability

Although the presented theoretical bispectra have different characteristics, we would like to get an indication how well these could be discriminated. For instance, it seems obvious that the similarity between the feature bispectrum and the resonant bispectrum could lead to significant confusion when actually traced in the data. In order to do so, we want to measure the distinguishability of these shapes, which is usually quantified using the amount of overlap or correlation between two shapes. One can define a inner product between two shapes

$$\begin{aligned} F_X \star F_Y &\equiv \int_{\Delta_k} dk_1 dk_2 dk_3 k_1^4 k_2^4 k_3^4 w_k F_X F_Y \\ &= \int_{\Delta_k} dk_1 dk_2 dk_3 w_k S_X S_Y. \end{aligned} \tag{5.15}$$

The correlation between two shapes F_X and F_Y is then defined as

$$\mathrm{C}(F_X, F_Y) \equiv \frac{F_X \star F_Y}{(F_X \star F_X)^{1/2}(F_Y \star F_Y)^{1/2}}. \tag{5.16}$$

Here w_k is a weight function, which was chosen as $w_k = 1/k_t$ by Fergusson & Shellard (2009) to increase resemblance with the Fisher matrix (correlation) found in multipole space. The integral

[3] This limit is on the enfolded line $x_j = x_{j+1} + x_{j+2}$ within the sum. Outside the sum, this expression is non-zero but finite. For example $x_1 \to x_2 + x_3$ gives:

$$\frac{1}{8x_2^4x_3^4\omega^2}\left[(x_2+x_3)^3 - x_3\left((x_2+x_3)^2 - 2x_2^2\omega^2\right)\cos\left(\frac{2x_2\omega}{x_2+x_3}\right) - x_2\left(\left((x_2+x_3)^2 - 2x_3^2\omega^2\right)\cos\left(\frac{2x_3\omega}{x_2+x_3}\right) + 2x_3(x_2+x_3)\omega\left(\sin\left(\frac{2x_2\omega}{x_2+x_3}\right) + \sin\left(\frac{2x_3\omega}{x_2+x_3}\right)\right)\right)\right].$$

.

runs over the 'tetrahedral' domain, which is bounded by the following triangle constraints

$$k_a \leq k_b + k_c \text{ for } k_a \geq k_b, k_c$$
$$k_a, k_b, k_c \leq k_{max},$$

where $a, b, c = \{1, 2, 3\}$, $a \neq b \neq c$.

Before we compute the correlation between the shapes, let us perform a quick qualitative analysis in order to get an indication of what to expect. First of all, note that the shape coming from initial state modifications (eq. (5.13)) is clearly different from the other two. While for features (eq. (5.4) and eq. (5.9)) the argument in the oscillating functions explicitly depends on the sum all three comoving momenta, the argument in eq. (5.13) depends on the ratio of momenta. Consequently we can expect a rather small overlap. This becomes even more apparent once we adapt a new set of variables

$$\begin{aligned} k = k_t/2, \qquad & k_1 = k(1-\beta) \\ k_2 = \frac{1}{2}k(1+\alpha+\beta) \qquad & k_3 = \frac{1}{2}k(1-\alpha+\beta) \\ dk_1 dk_2 dk_3 \quad &= \quad k^2 dk d\alpha d\beta, \end{aligned}$$

proposed by Fergusson & Shellard (2007). As a consequence the argument in eq. (5.13) will depend on the two variables α and β, while the arguments in eq. (5.4) and (5.9) will only depend on k. In that sense, we can say that oscillations in these shapes are in *orthogonal* directions.

In addition, for both the feature and resonant bispectrum the frequency is fixed along one direction. That is, the frequency does not change (feature) or only slightly changes (resonant) when you run through a fixed direction in comoving momentum space. For the non-BD bispectrum however the argument in the oscillating function has a component that scales as $1/x_j$. Consequently for $x_j \to 0$ the effective frequency $\omega_{eff} \to \infty$. Naturally, x_j is cutoff from below (as $k_{min}/k_{max} \sim 10^{-4}$), however even with a cutoff the range in effective frequencies is large along a direction. This effect is present at all frequencies, and it turns out it will determine the efficiency of mode expansion for this bispectrum discussed in the next section.

We have numerically calculated the correlator as defined in eq. (5.16) between both feature bispectra and non-BD spectrum. We found the correlation to be maximal for low values of both frequencies (of order 1 percent around $\omega = 10$), indicating that there is no evidence for a particular resonant frequency; the largest correlation occurs due to the fact that there are less oscillations, thereby decreasing the chance for (almost perfect) cancelations in the integral. As expected, we can safely conclude that these shapes are distinguishable/orthogonal.

For the two bispectra of eq. (5.4) and (5.9) we can expect a larger correlation. The appearance of a log in eq. (5.9) is the only major difference between the two bispectra. In the new coordinate set, the bispectrum of (5.4) does not depend on the α or β. Let us try to make a simple analytical approximation of the relevant correlator before we compute the correlation numerically. The first term in (5.9) dominates the second for large values of $(f\phi_*)^{-1}$. Therefore for simplicity we neglect the second term. As a consequence both terms now depend only on k. In the computation of the correlator the integration over α and β drops out and to get an

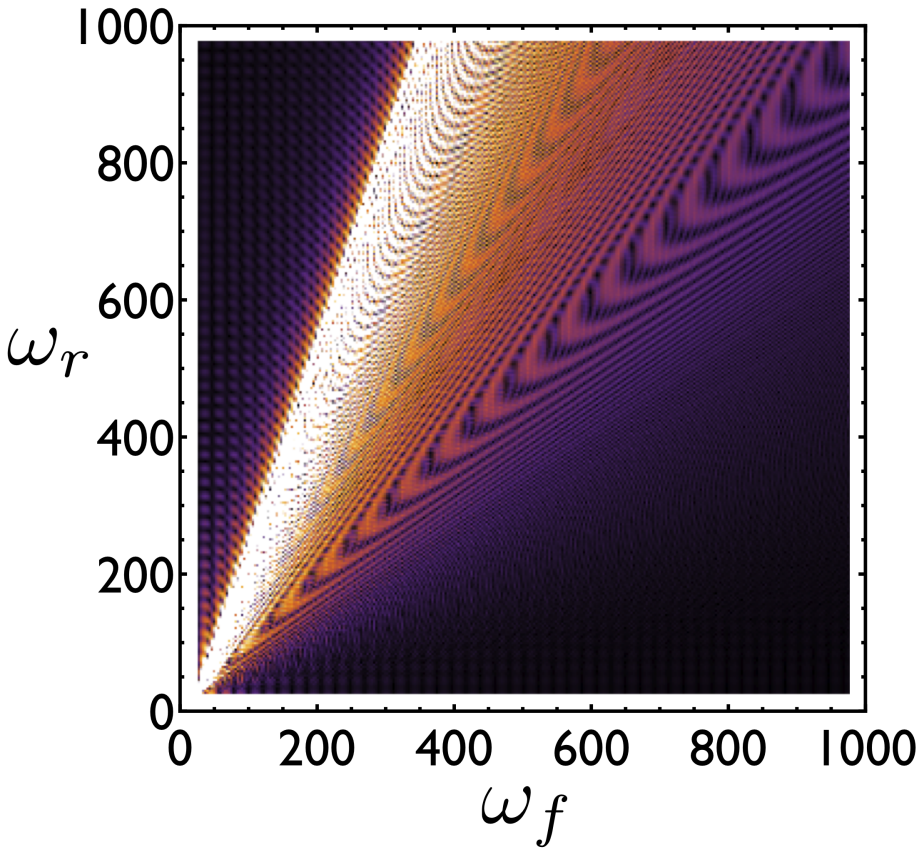

Figure 5.2: The correlation between the bispectral shapes of (5.4) and (5.9) for various values of the frequency. The light areas correspond to correlations of order $\mathcal{O}(1)$, while the dark shaded areas correspond to correlations close to 0. The correlation was computed with $\delta = \gamma_1 = 0$.

indication of the resonance we only need to investigate the following integral:

$$\int_0^3 x_t dx_t \sin(\omega_f x_t + \delta) \sin(\omega_r \log x_t + \gamma_1). \tag{5.17}$$

where we assumed that at most $k_t = 3k_{max} \rightarrow x_t = 3$. This integral can be done analytically and results in a sum of Γ functions (we have set $\delta = \gamma_1 = 0$). The interpretation of the result is rather complicated as all terms are divergent and there are no terms that can be easily neglected. However, one can plot the result and find that there is a clear resonance 'area' around $\omega_r \simeq 20\omega_f$. We have confirmed this resonance as a function of frequency when considering the full expression and allowing both phases to be non-zero. We have plotted (fig. 5.2) the correlation for a range of frequencies ($10 < \omega < 1000$) and a phase $\delta = \gamma_1 = 0$. The largest values obtained from this numerical computation are of order 0.6, or 60 percent correlation (we have used discrete steps of $\delta\omega = 10$), and we expect there to exist correlation of $\mathcal{O}(1)$ for some specific values of ω). As such it will be hard to discriminate between these two

models solely using observations of the bispectrum (as one could simply confuse frequencies). However, as mentioned before, axion inflation for example predict a large scalar to tensor ratio. Measurement of r could break the degeneracy between a sharp feature in the potential versus axion inflation. In addition, one does not expect $\omega_r < 10$ since it would not produce observational $f_{\rm NL}^{res}$, while for the feature bispectrum the natural frequency is no larger than $\omega_f \sim 50$. If one would be able to extract a frequency from the data, a large frequency would favor a resonant model while a low frequency could indicate a sharp feature.

5.3 Mode Expansion

5.3.1 Power Modes

The discussed primordial bispectra have very little in common with the constrained local, equilateral and orthogonal bispectra. Typically, to constrain any type of non-Gaussianity one computes the correlator (eq. (5.16)) and derive the so-called 'fudge' factor which indication how much 'signal' leaks into an existing template With the use of the fudge factor one is able to deduce a bound on the amplitude of the unconstrained bispectrum. The reason why certain templates have been constrained and some others have not, is two-fold. First and foremost, until now most models produced non-Gaussianities that can roughly be placed in one of the constrained types. For this reason, it was not immediate to search for any other type, simply because there were no models that indicated bispectra with completely orthogonal characteristics. Of course, optimally, one would simply look for the full bispectrum as a function of the multipole numbers instead of constraining the amplitude in particular bispectral configuration, but the low S/N and computational limitations have so-far restrained us to the former.

The second reason not to look for more 'exotic' bispectra is that for a fast estimator, the bispectrum one would like to constrain needs to be factorizable and scale invariant. That is, it is useful if the bispectrum can be we written as sum of products of functions, where each function only depends on one direction in multipole or comoving momentum space. It has been shown that such factorizability reduced the number of computations one has to make in order to constrain the amplitude of the bispectrum by a factor l^2, where l is the number of observable multipoles of the experiment (leaving only l^3 computations).

The constrained non-Gaussian amplitudes (in the form of $f_{\rm NL}^i$, where i labels the comoving momentum type, local, equilateral or orthogonal) are all based on templates that are factorized in the manner explained above. For instance, although DBI inflation does not produce a factorized bispectrum, it is well approximated by the equilateral template (Creminelli et al. 2006), that is factorized by construction. The same is true for both the local and orthogonal template, as well as the enfolded template (Meerburg et al. 2009). However, the method for constructing such factorized approximations of existing theoretical bispectra is rather ad-hoc. Until recently there was no procedure no construct a factorized bispectrum using a consistent prescription.

Fergusson et al. (2010a) proposed a method for constructing factorized approximations to theoretical bispectra using polynomial expansion. The approach is fairly straightforward; one defines a set of orthonormal 3 dimensional functions (where orthonormal is defined using a

correlator of the form[4] eq. (5.16), and the weight function can be adjusted) which are a-priori factorized and from there one computes the corresponding weight factors (α_n) via the inner product between a number of polynomial modes (R_n) up until a sufficient overlap between the polynomial expansion and the original bispectrum is established, i.e., until N such that

$$S(x_1, x_2, x_3) \simeq \sum_{n=0}^{N} \alpha_n R_n(x_1, x_2, x_3). \tag{5.18}$$

Without discussing the details of constructing such polynomial modes (see (Fergusson et al. 2010a) for a detailed description), here we want to try and investigate how well this would work in case of oscillatory bispectra of (5.4), (5.9) and (5.13).

Before we do so, let us make a few notes. First of all, recall that the objective of the expansion is to factorize a given theoretical bispectrum. However, as you can see from eq. (5.4), this particular bispectrum, albeit a best-fit approximation[5], is already of the factorized form. One can still try to expand this in terms of power law polynomials, as described here, since polynomial modes will in general behave better numerically. The other two examples of primordial bispectra are not factorizable in terms of oscillating functions using simple identities. Consequently, the polynomial expansion seems to be a good first effort in order to set up an approximately factorized form.

Secondly, were we able to expand these into a factorized form, and subsequently projected to multipole space and applied to the data, we might still miss the entire signal, simply because one of the free parameters is the frequency of the oscillations. For a non-BD bispectrum and the axion inflation model, the range of possible frequencies spans (at least) 2 orders of magnitude. Therefore, if we would fix the frequency, searching for a signal with a constructed factorized template would probably not be the best approach. Fortunately, we will later see that if you would measure mode functions in the data, instead of a fixed template, one could in principle extract information about a variety of oscillating signals. Let us emphasize that even if we would not be able to reconstruct a factorized form of a given spectrum with a small number of modes, it is still very well possible we could observe the same spectra by measuring a small number of mode functions in the data (effectively the frequency (and the phase) remain a free parameter during mode extraction).

Feature Bispectrum

First we consider the bispectrum coming from a feature in the potential (eq. (5.4)). Out of the given examples it has the simplest form (excluding the envelope). We choose $\delta = 0$ for

[4]For the construction of these polynomial modes we set $w = 1$. Once computing the correlator between the original and the reconstructed spectrum one can take $w = 1/k_t$ in order to see how much of an effect projection onto multipole space can have. We find that it reduces the correlation by 5 to 10% in both polynomial expansion and Fourier expansion. As such, it should not effect the conclusions we draw in this chapter where all correlation shown are based on $w = 1$. In order to build modes that are optimized for multipole expansion you should start by considering a weight function $1/k_t$. This is beyond the scope of this paper.

[5]The proposed envelop function has the form $(k_1+k_2+k_3)^n e^{(k_1+k_2+k_3)/k_* m}$,, where m and n are fitted to the numerical results. The envelope function is therefore also factorizable. Again, we did not consider this envelop since it is smooth compared to the oscillatory part of the bispectrum. However, such an envelope could be of significant influence in predicting the correlation in multipole space (Fergusson & Shellard 2009).

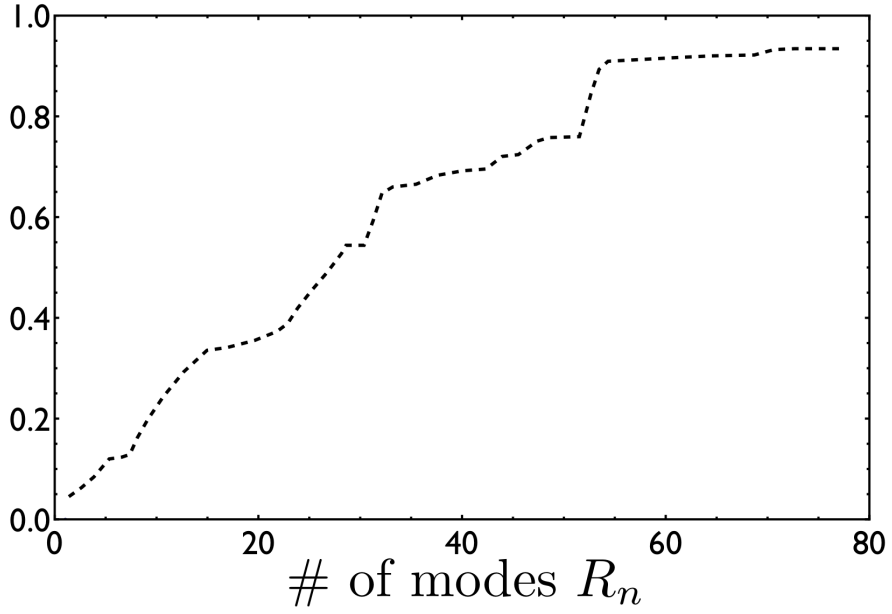

Figure 5.3: Example of the increasing correlation (eq. (5.16)) between the approximation of eq. (5.4) and the original spectrum. Here $\omega_f = 9$ and we find that it requires over 80 modes to achieve perfect correlation.

ω_f	1	2	3	4	5	6	7	8	9*
# of modes R_n	1	5	8	12	18	43	55	69	82

Table 5.1: As the frequency is increased it requires a rapidly growing number of modes to get over 98% correlation with the original spectrum.

simplicity, and since the phase can always be scaled out it will not affect the results[6]. In table 5.1 we have computed the number of modes necessary to get a correlation of at least 98% with the original spectrum for several values of ω_f. As expected, as the frequency is increased, one has to expand the bispectrum with a (rapidly) growing number of modes. For $\omega_f = 9$ we get a 93% correlation with 82 modes. On itself, it actually quite remarkable that one is able to reproduce the spectrum with a limited number of modes. Recall that the possible feature at $l \sim 30$ would result in a (decaying) oscillation with $\omega_f \sim 50$, which would be hard to fit this way. On the other hand, as we argued earlier, a frequency of $\omega_f = 50$ can be considered an upper limit, as features at higher multipole number would result in longer wavelengths. We have plotted an example of how the correlation between the original spectrum and the expansion increases as a function of the number of mode functions in the expansion in fig. 5.3.

[6]One would also have to consider $\cos \omega_f x_t$ but we found no difference when expanding between the cosine and sine in terms of the required number of modes.

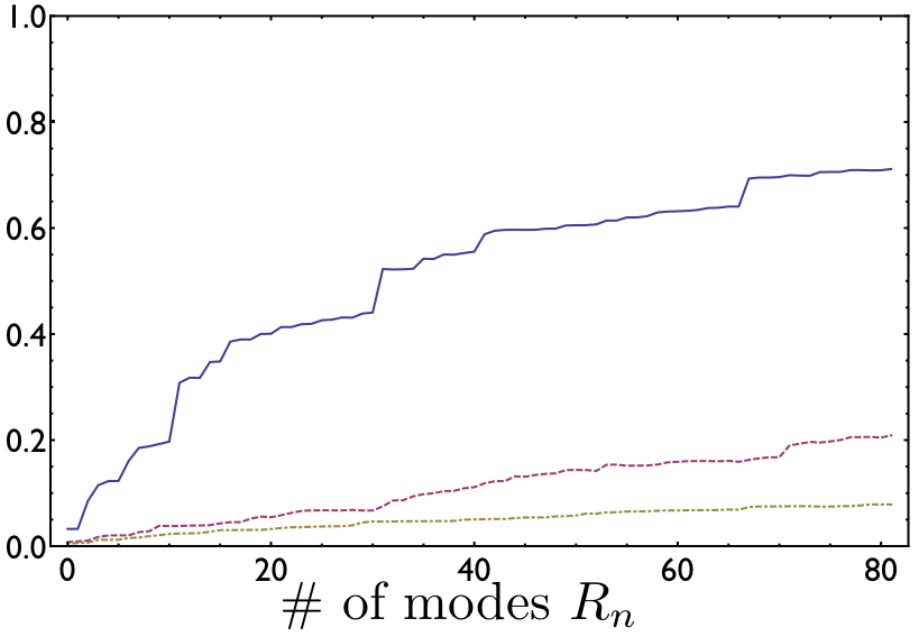

Figure 5.4: The correlation between $\sin(\omega_r \ln x_t)$ for 3 different frequencies. From top to bottom $\omega_r = 20, 40$ and 60. Beyond frequencies of 60 polynomial expansion would require many modes to achieve significant correlation with the original spectrum.

Resonant Bispectrum

Next, let us consider the resonant bispectrum. It is quite similar to the feature bispectrum, but theoretically we expect much larger frequencies ($20 \leq \omega_r \leq 10^3$). We have computed (fig. 5.4) the correlation between the expansion and the original spectrum, chosen to be $\sin(\omega_r \ln x_t)$ since for the same reason as before the phase will barely affect the number of modes required to reconstruct the spectrum. As expected, the convergence of the correlation towards one (perfect overlap) proceeds slowly. For the lowest frequency we considered ($\omega_r = 20$), the correlation reaches 71% after 82 modes. For $\omega_r = 60$ the largest correlation we can achieve is 7% after 82 modes. Recall that the amplitude of the resonant bispectrum is proportional to its frequency. The maximum correlation between existing templates and the axion spectrum is of order 1% (Flauger & Pajer 2011) (although for small frequencies this can be 10% for the equilateral template) and possibly measuring these modes in the data would therefore still allow for a constraint on axion inflation that is 10-100 times[7] better than what we have now. In general, increasing the frequency above $\omega_r \sim 60$, large correlation becomes hard to achieve with a limited number of modes.

[7]This would require adding the modes once extracted from the data.

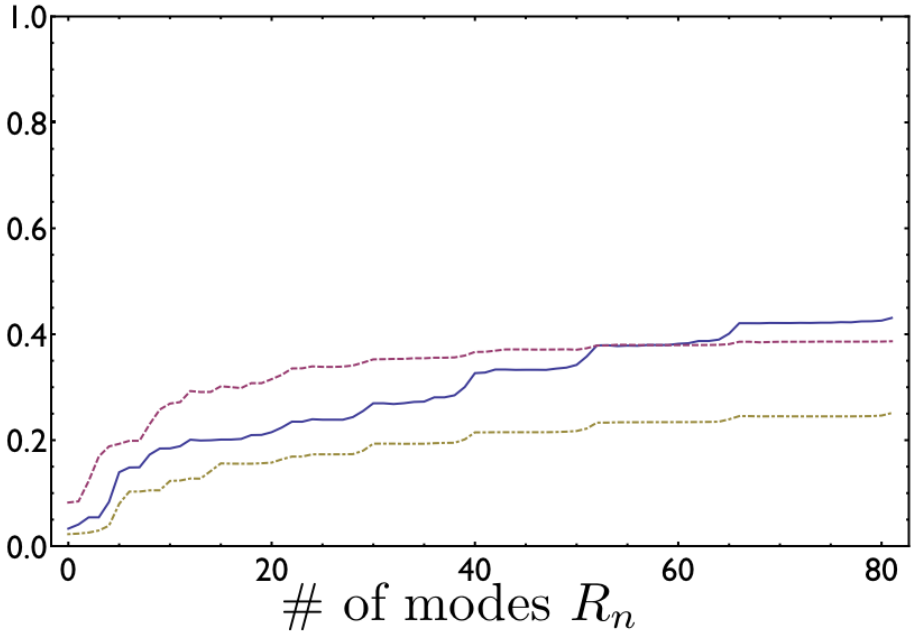

Figure 5.5: The correlation between the non-BD bispectrum of eq. (5.13) and its polynomial expansion as a function of mode number for 3 different frequencies $\omega_v = 20$ (solid), 40 (dashed) and 60 (dot-dashed). We have set $\gamma_2 = 0$ but have found very little difference for non-zero γ_2.

Non-BD Bispectrum

For the non-BD spectrum of eq. (5.13) we find that computing the correlator numerically requires a very high resolution, because this shape contains terms that are singular and the spectrum as a whole is only finite due to the exact cancellation between all the specific terms. To avoid these problems one has to stay away from the line(s) $x_{j+1} + x_{j+2} = x_j$, which can be done by adding a small ϵ in the vicinity of this line in the integral that defines the dot product (eq. (5.15)). The results are shown in fig. 5.5. Even for low frequency ($\omega_v = 20$) we can not achieve a large correlation with 82 modes. On the other hand, increasing the frequency does not really affect the ability to reach similar correlation. Overall, we find that the non-BD bispectrum is the most difficult to reconstruct due to the appearance of terms that diverge inside the argument, since for $x_j \to 0$ the frequency of the signal becomes extremely large at some of the edges of the tetrahedral domain. The observation that we can still reach some correlation is because there are also areas on the tetrahedral domain where the effective frequency is relatively small. These areas remain even if ω_v increases (al though they should become smaller and smaller) Consequently we find that achievable correlation with 82 modes is small but does not decrease significantly when you go to higher frequencies. The correlation with smooth spectra is typically of order $\ll 1\%$ (Meerburg et al. 2009) which means that an accumulated correlation of only a few percent could drastically improve the constraints we can put on $|\beta|$ as the amplitude scales proportionally to ω_v^3.

$n = 0 \to 000$	$n = 5 \to 012$	$n = 10 \to 004$	$n = 14 \to 014$	$n = 19 \to 114$	$n = 24 \to 133$
$n = 1 \to 001$	$n = 6 \to 003$	$n = 9 \to 013$	$n = 15 \to 005$	$n = 20 \to 024$	$n = 25 \to 124$
$n = 2 \to 011$	$n = 7 \to 112$	$n = 11 \to 112$	$n = 16 \to 222$	$n = 21 \to 015$	$n = 26 \to 034$
$n = 3 \to 002$	$n = 8 \to 022$	$n = 12 \to 113$	$n = 17 \to 123$	$n = 22 \to 006$	$n = 27 \to 115$
$n = 4 \to 111$	$n = 9 \to 013$	$n = 13 \to 023$	$n = 18 \to 033$	$n = 23 \to 223$	$n = 28 \to 025$

Table 5.2: The association of mode numbers for the first 28 modes. For example $n = 10$ corresponds to the mode for which one direction is of maximally 4th order and the other two are 0 order, i.e., for polynomial modes $R_{10} \propto f(1, x_1, x_1^2, x_1^3, x_1^4) + f(1, x_2, x_2^2, x_2^3, x_2^4) + f(1, x_3, x_3^2, x_3^3, x_3^4)$.

Polynomial expansion seems to work reasonably well for low frequencies of the various bispectra. For larger frequencies, to reconstruct the original spectrum the polynomial expansion requires an increasing number of modes. Given the large allowable frequencies for the resonant and non-BD bispectra, polynomial expansion might not be the most effective way of expanding. In the next section we will explore another type of expansion which uses a Fourier basis. We will investigate if such a basis would require less modes to achieve similar correlation.

5.3.2 Fourier Modes

The polynomial expansion introduced by Fergusson et al. (2010a) is based on power modes, i.e., the expansion is in increasing order of x^n. This is not necessarily optimal for describing oscillatory functions. There are two possible alternatives; the first one would be to expand the argument into a sum of functions, that each depend on one direction only. A such, one can again use trigonometric identities to expand the cosine and sine into factorized forms (be that oscillatory functions). The second option could be to use a Fourier expansion instead of a polynomial expansion. This would only be useful if for large frequencies you would need a small(er) number of modes. Before we get into Fourier mode expansion let us briefly discuss the alternative of expanding the argument in the oscillatory function.

This option would only suffice if the approximation requires 2 modes maximally. If it requires more modes, you will get product of two or three different directions in momentum space, and as a result you will not be able to expand the cosine and the sine. Let us consider the axion model. The argument is given by $\omega \log x_t$. Using the mode expansion, one finds that one can achieve $> 99\%$ correlation after just two polynomial modes; zero order and first order. Not surprisingly this is almost equivalent to a Taylor expansion to first order of $\log k_t/k_*$ around the point $k_t \sim 1.4 \sim \sqrt{2}$. Consequently, there are no cross-terms, and one can expand the cosine and sine into factorizable function of the three comoving momenta, just like you could expand the feature spectrum into oscillating functions. As it turns out however, although there is a 99% correlation between the arguments after expansion, the full bispectrum is very sensitive to small deviations in the argument, especially for large frequency. Consequently, the correlation between the full bispectrum and the approximated bispectrum decreases as a function of the frequency; from $\sim 90\%$ for $\omega_r = 1$ to $\sim 50\%$ for $\omega_r = 20$. Although this is equivalent to what can be achieved with the polynomial expansion using 7 modes, the problem is that we can not improve it in any way. Since this will only work for a first order expansion, we can never reach beyond 50% correlation, unlike the polynomial expansion, where we can simply include more

modes. Note that for non-BD model this method will not work as the argument is already a product of two directions in comoving momentum space, i.e., $(k_{j+1} + k_{j+2})/k_j$.

The second option is to consider a Fourier expansion, where we try and fit terms such as $\sin \omega f(x, y, z)$ to a sum of Fourier modes that all depend on one direction only. Such factorization would still lead to the l^2 reduction in computation, since the integrals in k space can now be performed individually[8]. We consider $\exp[i2\pi nx]$ as our basis function (as apposed to x^n) and constructed a orthogonal set of three dimensional mode function similar to (Fergusson et al. 2010a). The first few one dimensional functions are given by

$$\begin{aligned}
f_0(x_1) &= \sqrt{2} \\
f_1(x_1) &= 0.22 + 0.23i + 1.45e^{2i\pi x_1} \\
f_2(x_1) &= -0.0087 + 0.041i(0.088 + 0.62i)e^{2i\pi x_1} - \\
&\quad (0.31 + 1.12i)e^{4i\pi x_1} \\
f_3(x_1) &= (-4.9 + 2.2i)10^{-3} - (0.15 - 0.11i)e^{2i\pi x_1} + \\
&\quad (0.68 - 0.65i)e^{4i\pi x_1} - (0.59 - 0.65i)e^{6i\pi x_1} \\
f_4(x_1) &= (-6.5 - 3.8i)10^{-4} - (0.042 + 0.017i)e^{2i\pi x_1} + \\
&\quad (0.44 + 0.12i)e^{4i\pi x_1} - (1. + 0.2i)e^{6i\pi x_1} + \\
&\quad (0.63 + 0.1i)e^{8i\pi x_1} \\
f_5(x_1) &= (-1.1 - 11.1i)10^{-5} - (0.002 + 0.011i)e^{2i\pi x_1} + \\
&\quad (0.051 + 0.16i)e^{4i\pi x_1} - (0.25 + 0.63i)e^{6i\pi x_1} + \\
&\quad (0.4 + 0.89i)e^{8i\pi x_1} - (0.21 + 0.41i)e^{10i\pi x_1} \\
&\ldots,
\end{aligned}$$

The functions are shown up to $n = 10$ in fig. 5.6. From these one can construct the three dimensional basis functions via a product of each mode and symmetrization of three comoving momentum arguments; x_1, x_2 and x_3

$$\mathcal{Z}_{prs}(x_1, x_2, x_3) \quad \propto \quad [f_p(x_1) f_r(x_2) f_s(x_3) + 5 \text{ perm}]. \tag{5.19}$$

One has to introduce a counting scheme to re-numerate the three labels $\{p, r, s\}$ to n. We have chosen equal slicing counting (Fergusson et al. 2010a), of which the first 28 modes (n) and their association ($\{p, r, s\}$) are shown in table 5.2. After the construction of these modes, one has to apply additional Gramm Schmidt orthogonalization to $\mathcal{Z}_n$ to increase orthonormality of different mode functions. We refer to the three dimensional orthonormalized modes as $\mathcal{F}_n$ and the corresponding weights as $\tilde{\alpha}_n$.

$$S(x_1, x_2, x_3) \quad \simeq \quad \sum_{n=0}^{N} \mathcal{R}e\left(\tilde{\alpha}_n \mathcal{F}_n(x_1, x_2, x_3)\right). \tag{5.20}$$

If S would have been complex, one should add $i\mathcal{I}m\left(\tilde{\alpha}_n \mathcal{F}_n\right)$ in order to take this into account. The coefficients $\tilde{\alpha}_n$ can be computed by taking the inner product (eq. (5.15)) between the

[8] Fergusson et al. (2010a) briefly discussed Fourier mode expansion in sec. E as a possible orthonormal basis, however no results are shown.

original shape function (bispectrum) and the various mode functions $\mathcal{F}_n$, i.e.,

$$\tilde{\alpha}_n = \int_{\Delta_{x_i}} dx_1 dx_2 dx_3 S(x_1, x_2, x_3) \mathcal{F}_n^*(x_1, x_2, x_3). \tag{5.21}$$

Feature Bispectrum

For the feature bispectrum we do not necessarily have to consider the Fourier expansion[9] since that spectrum can be rewritten into a product of Fourier modes simply by using trigonometric identities, e.g.

$$\begin{aligned} \sin \omega_f x_t = & \cos \omega_f x_3 \left(\sin \omega_f x_1 \cos \omega_f x_2 + \cos \omega_f x_1 \sin \omega_f x_2 \right) + \\ & \sin \omega_f x_3 \left(\cos \omega_f x_1 \cos \omega_f x_2 - \sin \omega_f x_1 \sin \omega_f x_2 \right). \end{aligned} \tag{5.22}$$

The other two bispectra are not of the same form, since their arguments are non-linear functions, i.e., $\ln x_t$ for resonant non-Gaussianities and $(x_{j+1} + x_{j+2})/x_j$ for for non-BD modifications and these can be made of the form above by expanding, using the constructed Fourier modes $\mathcal{F}_n$. Given the form of the first argument you expect only a limited number of modes to significantly contribute, for example those modes that have equal mode number in the directions x_1, x_2 and x_3 (you should think about this expansion as a series around the point x_t, see table 5.2). For the second argument you expect more modes to matter, since the arguments depend on all three directions independently. Consequently the weights $\tilde{\alpha}_n$ are expected to be close to zero for many n when expanding resonant non-Gaussianities, while for the non-BD scenario they should all matter to some extend (and obviously more modes will be important for large ω_v).

Resonant Bispectrum

We have computed the correlation for the axion bispectrum with the Fourier expansion for frequency ranges of $\omega_r = 20 - 80$ up to 82 modes (fig. 5.7). As expected, we see that there are only a few modes that give significant contribution to the correlation, while most modes give only very little contribution and are not important for the expansion. We will discuss this fact in the context of CMB data mode extraction in the sec. 5.4. Given that the allowed range of frequencies $20 \lesssim \omega_r \lesssim 10^3$ this expansion is actually reasonable for the lower frequencies and the number of modes necessary to establish similar correlation as the polynomial expansion is reduced by a factor 5.

Non-BD Bispectrum

As for the polynomial basis expansion, the presence of a large number of features in the non-BD bispectrum does not allow for a fast reconstruction of the spectrum. In fact, expansion in the Fourier basis requires even more modes compared to the polynomial basis, reaching only $\sim 20\%$

[9] We will later show that we would also find it when we would search for resonant bispectra in the data, as the weights peak at almost the same mode numbers.

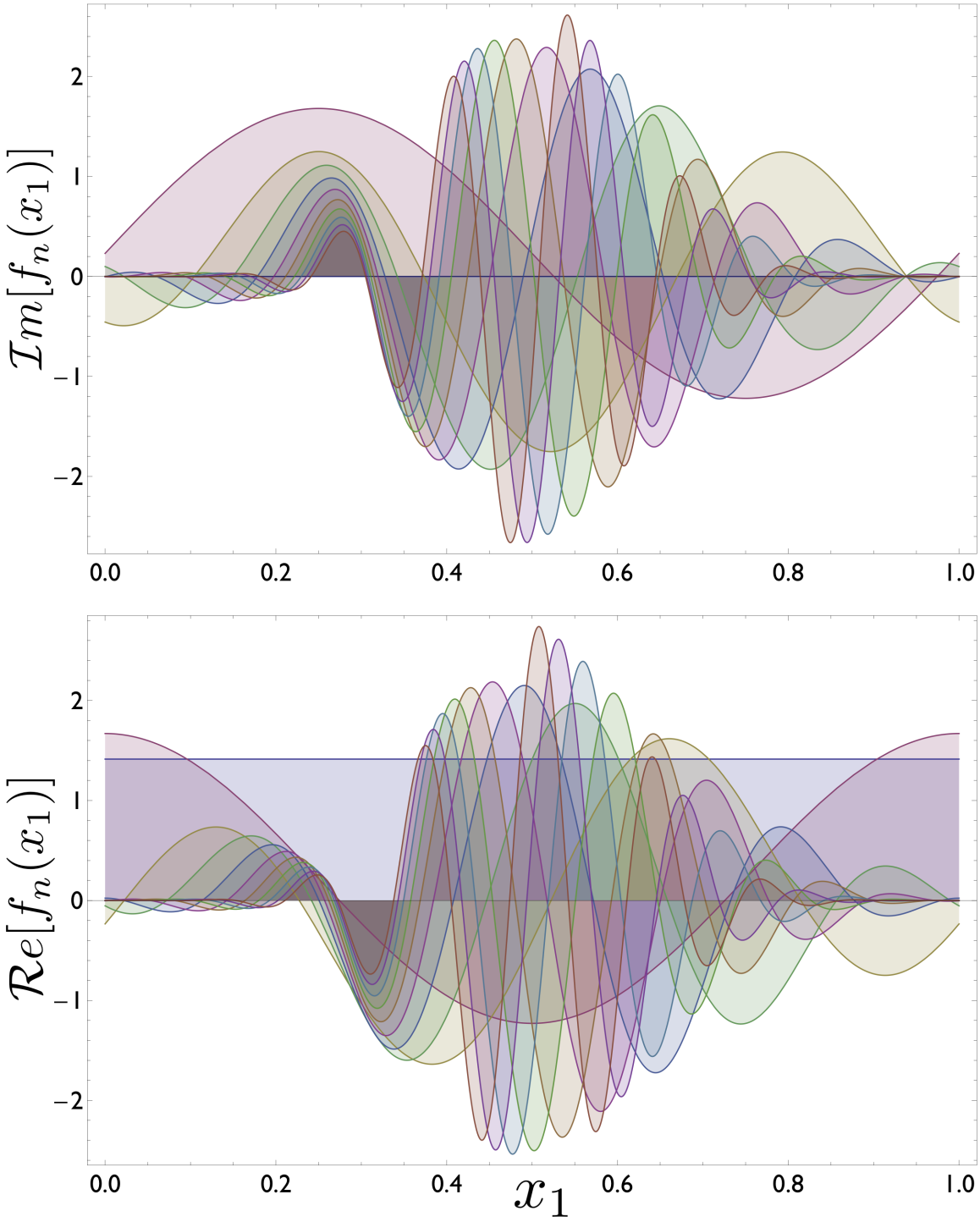

Figure 5.6: The one dimensional orthonormal Fourier functions $f_n(x_1)$ within the tetrahedral domain for the first 11 modes.

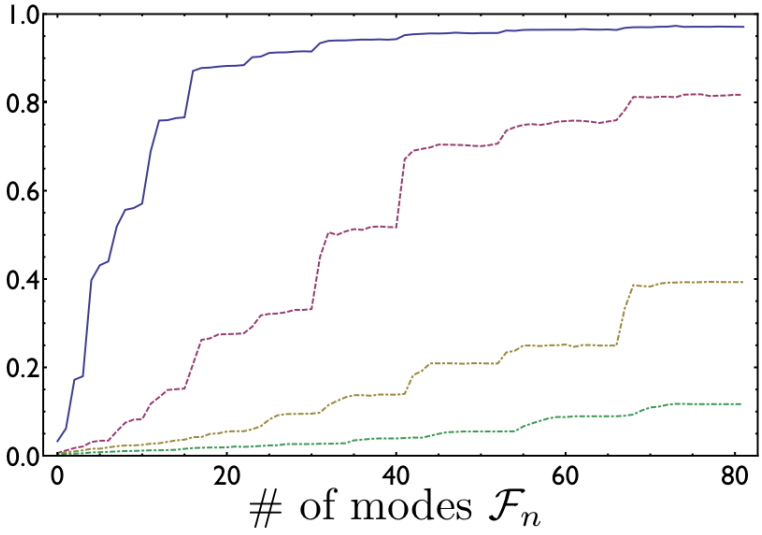

Figure 5.7: Correlation between $\sin(\omega_r \ln x_t)$ for frequencies $\omega_r = 20, 40, 60$ and 80. Compared with the polynomial mode expansion we reach similar correlation using about 5 times less modes. Also note that the increase of correlation is somewhat discrete, indicating that we might need only a fraction of these modes to reconstruct the original spectrum. We will discuss this observation in the next section.

correlation after 82 modes with $\omega_v = 20$. We also find that $\omega_f = 40$ actually reaches a slightly larger correlation, although this seems mostly due to a relatively large correlation with the zero order ($n = 0$) mode. Most likely this is caused by the fastest oscillating part of the spectrum which, in combination with numerics, could add constant power. We did observe something similar in fig. 5.5 for polynomial modes where the zero mode causes the correlation of the non-BD bispectrum reconstruction with $\omega_v = 40$ to be better initially compared to bispectrum expansion with $\omega_f = 20$.

In most realistic scenarios $\omega_v > 100$ (otherwise your effective field theory approach breaks down) and therefore both polynomial expansion and Fourier expansion fail to reconstruct this bispectrum effectively. The possible explanation why Fourier expansion is even worse than polynomial expansion for this type of bispectrum, seems to be related to the rapid change in frequency in a fixed direction. Fourier expansion is optimized for scale invariant frequencies. The polynomial expansion is simply optimized in reproducing as many different shapes as possible, explaining the observation that it is able to slowly increase correlation with the addition of modes while Fourier expansion seems to converge. Given the large enhancement of the amplitude $f_{\rm NL}^{nBD}$ (which scales as ω_v^3), one might still be able to extract some information from that data even with such small correlations.

Another possibility is that once non-BD bispectrum is projected onto multipole space one might establish a larger correlation with fewer (multipole) modes. The projection has the tendency to wash out small features (hence the weight of $1/k_t$ in the correlator). We hope to report on this in the future.

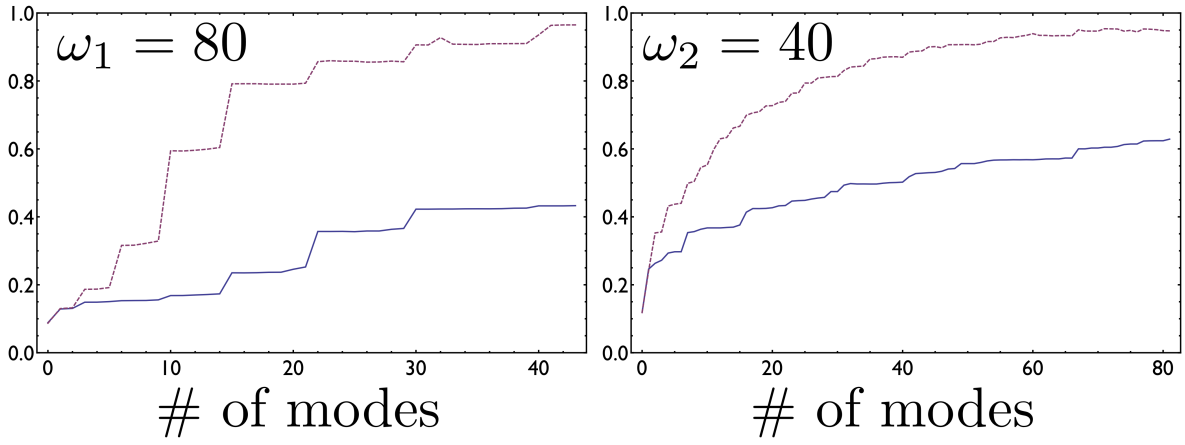

Figure 5.8: The correlation as a function of mode number for two out of three toy-spectra F_1(left) and F_2(right) in eq. (5.23). In both cases Fourier expansion (dashed) leads to faster convergence compared to polynomial expansion (solid).

Toy Spectra

To investigate the power of the Fourier expansion for oscillatory bispectra we have also tried to fit three toy-model shapes moving in different direction through comoving momentum space

$$\begin{aligned} F_1 &= \frac{1}{k_1^2 k_2^2 k_3^2}\left(\sin\frac{\omega_1}{x_1+1}+\sin\frac{\omega_1}{x_2+1}+\sin\frac{\omega_1}{x_3+1}\right), \\ F_2 &= \frac{1}{k_1^2 k_2^2 k_3^2}\sin\omega_2 x_1 x_2 x_3, \\ F_3 &= \frac{1}{k_1^2 k_2^2 k_3^2}\left(\sin\frac{\omega_3 x_t}{x_1+1}+\sin\frac{\omega_3 x_t}{x_2+1}+\sin\frac{\omega_3 x_t}{x_3+1}\right). \end{aligned} \tag{5.23}$$

We find again that for such a shapes the correlation increases about 5 times faster compared to polynomial mode expansion with the same frequency. The correlation as a function of mode numbers for F_1 and F_2 are shown fig. 5.8. We will discuss the weights of these models in the next section. Note that F_1 is already of the factorized form, however here we simply aim at showing the effectiveness of Fourier expansion. We want to emphasize that these spectra are not based on any physical model, but simply show that in general oscillatory spectra are better fitted using a Fourier basis.

Smooth Spectra

Although the Fourier expansion seems to work well for resonant non-Gaussianities and the toy-spectra, compared to polynomial expansion we confirm that Fourier expansion is not as effective: it is easier to gain fast convergence with a limited number of modes for most oscillating bispectra, but it is difficult to get correlation beyond 0.97 for smooth bispectra. This is

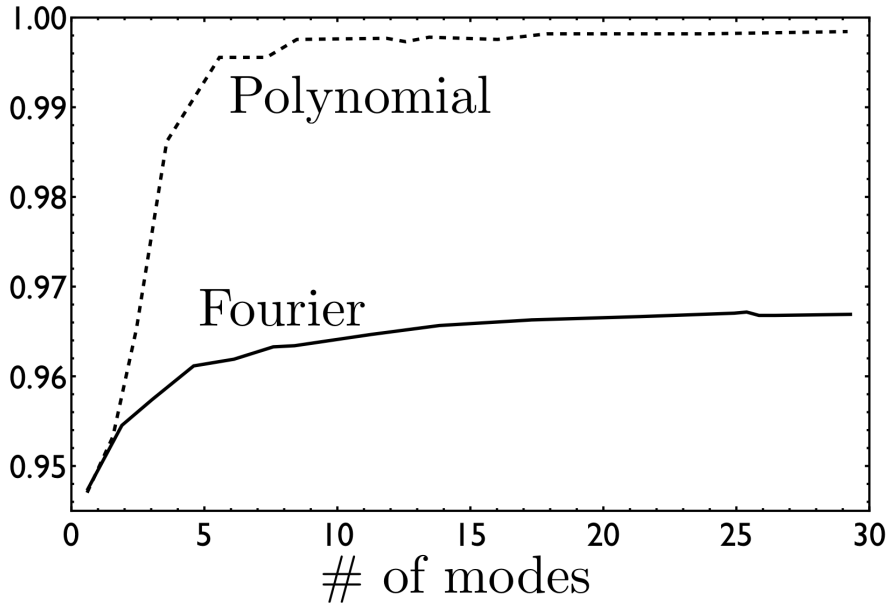

Figure 5.9: The correlation between the DBI bispectrum and both polynomial and Fourier expansion as a function of the maximum number of modes. As expected, both expansions start out equally (the zero mode of the real part of the Fourier expansion is equivalent to the zero mode of the power law expansion), but while the power law reaches a correlation of $> 99\%$ after just 5 modes, the Fourier remains stuck at 97%.

probably due to overshooting at the boundaries as discussed by Fergusson et al. (2010a). We explicitly show this in fig. 5.9 where we compare expansion of the 'smooth' DBI inflation bispectrum (which is very similar to equilateral), using Fourier modes and polynomial modes.

We conclude that Fourier expansion is a viable alternative for polynomial expansion in the case of oscillatory bispectra with relatively large frequencies. Using the Fourier expansion we can achieve factorizabilty of various oscillating bispectra with significantly less modes compared to polynomial expansion. For frequencies $\omega \gg 50$ polynomial and Fourier expansion are both unable to reconstruct the original spectrum with a small number of modes. In order to reconstruct models with such large frequencies, one should look for alternative methods. However, constraining these models with only limited number of modes seems to be a practical possibility. This will be topic of the next section.

5.4 Discussion

Even though the expansion of the oscillatory primordial bispectra becomes unavailing for really large frequencies, there are a number of interesting observations which could make constraining and expanding oscillating bispectra much more viable than presently argued. First of all, as predicted, the expansion in mode functions of the resonant bispectrum has a very discrete character; basically if you consider fig. 5.7 only few modes actually contribute significantly to

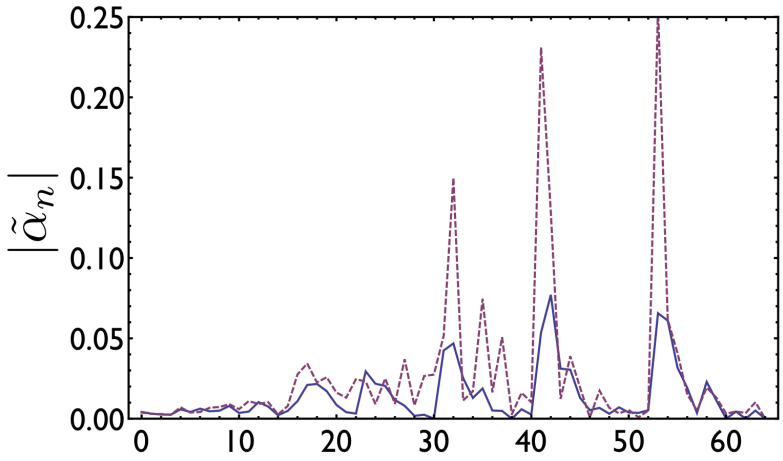

Figure 5.10: The weights derived for 65 modes for both $cos\omega_r \ln k_t$ (dashed) with $\omega_r = 50$ and $\sin \omega_f k_t$ (solid) with $\omega_f = 20$ showing that these both peak for similar mode numbers. Although distinguishing between these would be quite hard, it seems that for the feature bispectrum the values of the weights $\tilde{\alpha}_n$ are peaked sharper.

the convergence of the correlation. In fig. 5.11 we show the various weights ($|\tilde{\alpha}_n|$) as a function of mode number (as well as for F_1 and F_3 (not shown)). We can trace back the corresponding mode numbers in table 5.2. For instance there is a clear peak at $n = 16$, which correspond to all directions being maximally of quadratic order, and $n = 41$ with all directions being maximally of cubic order. Other peaks (e.g. $n = 21$, 32 and 53) correspond to the modes in which two out of three directions have one and two maximal orders less than the third, i.e., in mode number $n = 23$ two directions are maximally quadratic and the third is maximally cubic. As we already argued the location of these peaks makes sense, since the resonant model is a function of k_t (or x_t), which is the sum of the three comoving momenta. Effectively this shape is orientated in the k_t direction. One could only try to expand the spectrum only in those modes, which could significantly reduce the number of modes necessary. Since the important modes seem to be related to the direction of propagation of the oscillation, we find that this conclusion is independent of the phase. In other words, only the value of the weights will differ, not the mode numbers that are relevant for the expansion. This can be explained as follows. Consider a very simple example of an oscillating mode $\mathcal{R}e[e^{i(x+\delta)}]$. If we would expand this into polynomial mode functions, $\{1, x, x^2, ..., x^n\}$ we would find that α_n would change as a function of n if we vary δ. This makes perfect sense, since we know the polynomial expansion of these functions exactly, as they are the Taylor series of the sine and cosine. If we would expand in Fourier modes $\{1, e^{ix}, ...e^{inx}\}$, the expansion is obviously much simpler. However, more importantly, the complex phase will not affect the quantity $\alpha\alpha*$. Let us consider the mode with the largest value $\alpha\alpha*$ the resonance peak. The location of this resonance peak will be unaltered by a change of phase. For the weights of a polynomial mode expansion this is not true, as the introduction of a non-zero phase will cause this example to shift from a cosine to a sine, thereby transferring power from odd to even modes. This would cause peaks in α_n to

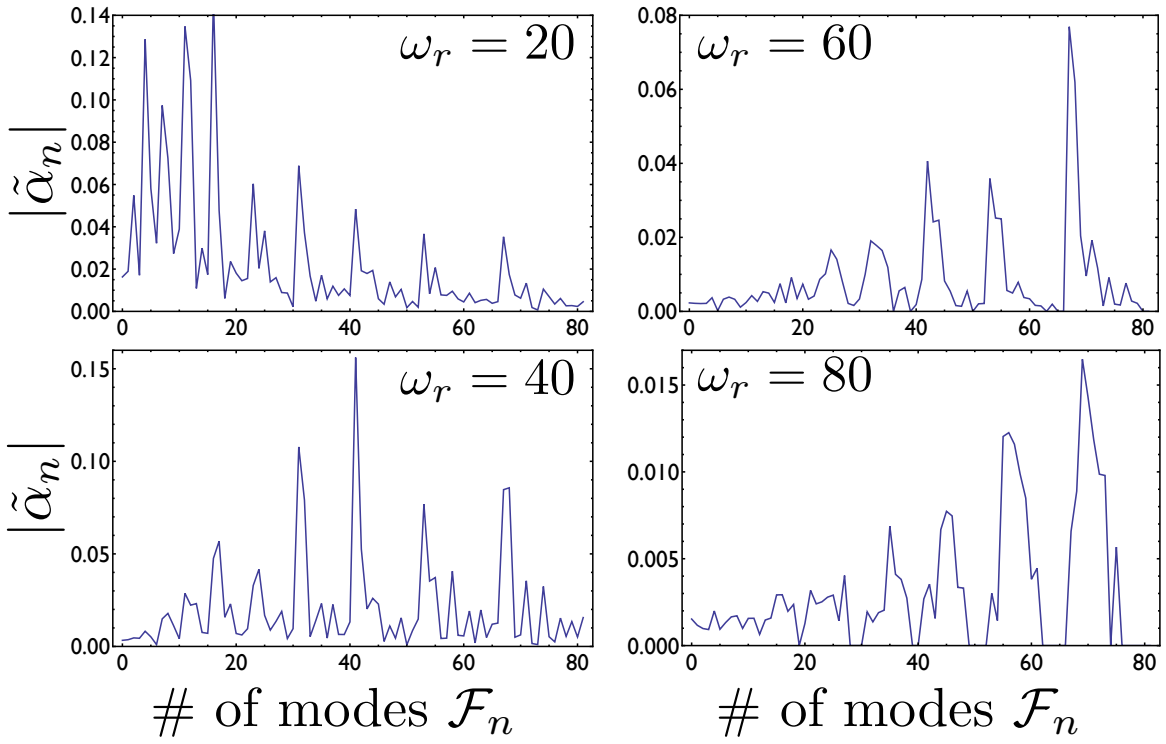

Figure 5.11: The weights $\sqrt{\alpha_n \alpha_n^*}$ for a resonant bispectrum as a function of mode number for various frequencies. It is clear that only limited number of modes are valuable in the reconstruction of the original spectrum via mode expansion. From an observational viewpoint this is very convenient as it would require the measurement of only a limited number of Fourier modes to learn about oscillations in the primordial bispectrum.

shift from n to $n+1$.

Secondly, from an observational viewpointf, given the discreteness of the correlation it is (obviously) not necessary to constrain all mode functions in the CMB data to get an indication of there is an oscillatory three point signal and what the possible frequency of this signal might be. For resonant non-Gaussianities we only need to consider those modes that have a significant weight $\tilde{\alpha}$, and the measured value of the weights would be a direct measure of the frequency. If one could extract the multipole projected Fourier modes that are responsible for most of the weight, this could in principle provide signatures of primordial bispectra with frequencies much larger than $\omega_r = 80$. Measuring modes up to e.g. $n = 100$ would not only provide information about the frequency of the signal, but could also hint on the type of primordial bispectrum. The distinction between the feature bispectrum (5.4) and the resonant bispectrum (5.9) would be more difficult, since the values of the weights peak at similar mode numbers although we have found that expanding the feature bispectrum in the constructed Fourier basis (instead of the simple trigonometric expansion discussed in sec. 5.3.2) could still be used to discriminate between the two signals (see fig. 5.10).

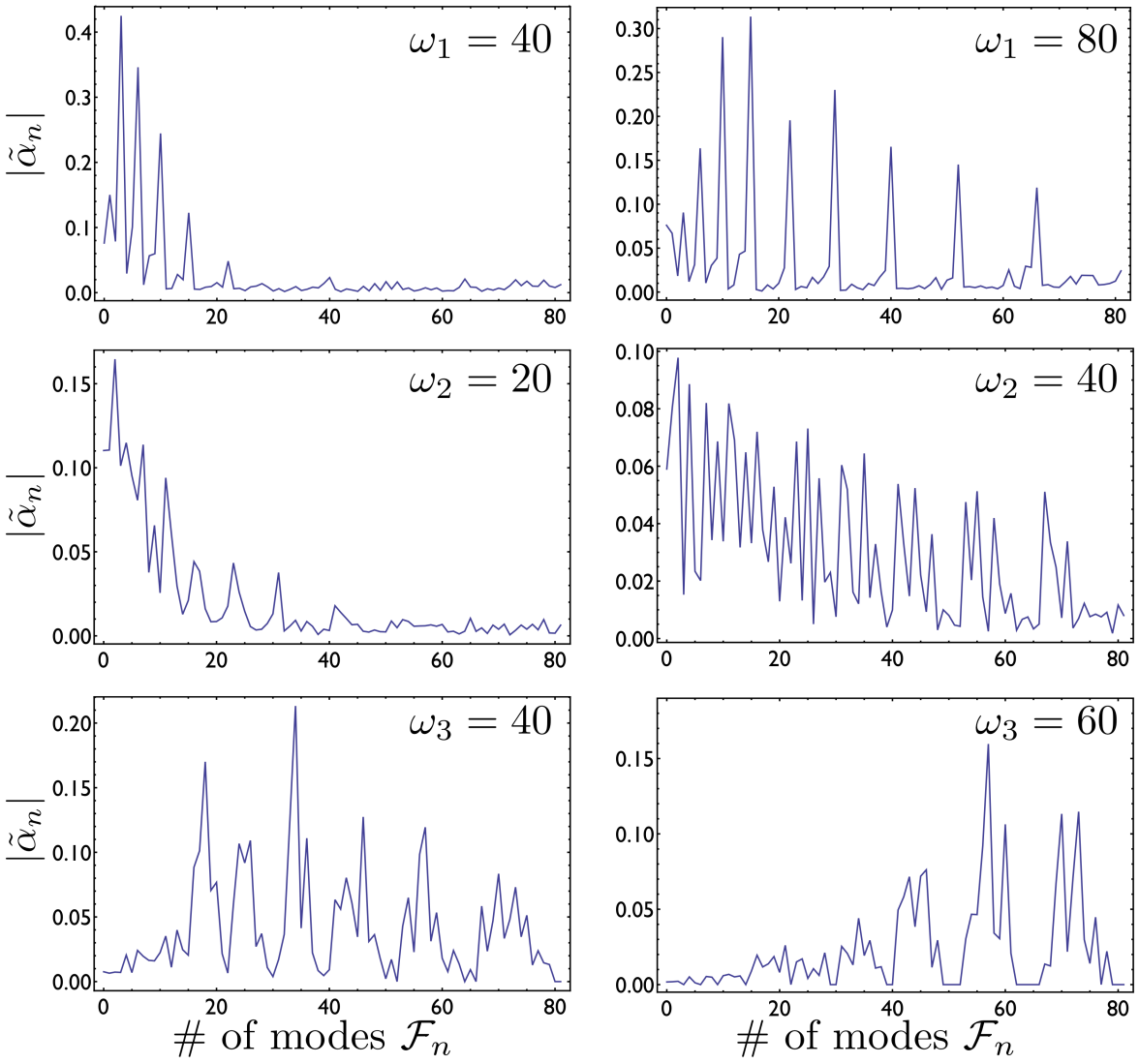

Figure 5.12: The weights $\sqrt{\alpha_n \alpha_n^*}$ for the three toy-spectra as a function of mode number for various frequencies. For the top spectrum the most relevant modes are those that have a maximal frequency of $\sim 2\pi n$ for one comoving vector and are constant for the other two. For the second toy-spectrum (middle) the symmetry of the argument makes all weights relevant (decreasing as a function of mode number). The third example (bottom) has modes dominating of which two comoving momentum vectors have non-zero (equal) frequency and the third one is constant. For example, the lower-right bottom example has $|\tilde{\alpha}_{57}|$ dominating, which corresponds to $\{p, r, s\} = \{0, 5, 5\}$.

To emphasize the ability to extract information on the primordial shape solely from the modes that are important, we have investigated three toy-model shapes of eq. (5.23). We have computed the Fourier weights for two different frequencies in fig. 5.12. As expected, F_1 has weights that peak when only one comoving momentum in k space is non-zero, i.e., it peaks at the modes where one momentum oscillates and the other two momenta are constant (see 5.2). The obvious reason is that each term in F_1 depends on one comoving momentum variable only, implying that there should be no cross terms in the expansion. For F_2 we find that many more modes are relevant, which makes perfect sense given that the argument in the sine depends on all three vectors in comoving momentum space. For F_3 however the argument effectively only depends on two comoving momenta, therefore the relevant mode functions (the ones with the largest $|\tilde{\alpha}|$) are the ones that have similar frequency in two momentum vectors and are constant in the third.

In this chapter we have only discussed mode functions in momentum space. We have to construct similar Fourier modes in multipole space or project these modes forward using the transfer function (Fergusson et al. 2010a), use these to expand a late-time oscillatory bispectrum to see if we get similar results in terms of mode number sensitivity. One expects that after projection the transfer function has caused some smoothing of the signal, which could render a Fourier basis less effective. On the other hand, intuitively it seems perfectly reasonable that a Fourier basis should be much more efficient in reconstructing oscillatory bispectra from the data. In addition, the effects of the transfer function on the correlation in l space can be examined by choosing the $1/k_t$ weight w_k in the primordial correlation function. We have found that our results were only marginally affected when including this weight factor and therefore we expect that Fourier mode expansion should be equally efficient in multipole space. To make sure that this is actually true, we should compute the projection of several oscillatory bispectra and construct a orthonormal Fourier basis in multipole space. We hope to report on this in the near future.

5.5 Conclusions

We have investigated the viability of mode expansion for bispectra that contain oscillations. The motivation for investigating such features and their mode expansion, is that recently it has been shown that several scenarios or mechanisms can produce such features not only in the power spectrum, but also in the bispectrum. The appearance of oscillations in the bispectrum makes comparison with existing bispectral constraints, based on smooth bispectra, very inefficient and there exists substantial room for improvement. In order to constrain oscillatory bispectra from the data, a logical first step is to factorize the bispectrum in order to efficiently compute its multipole counterpart. Polynomial expansion has been proposed to achieve factorization of a given theoretical bispectrum and we have investigated this for three different models. As expected, the larger the frequency of the primordial bispectrum, the more modes it requires to establish a reasonable approximation of the original spectrum. In the case of a feature in the primordial potential polynomial mode expansion might still be useful, at least for features at high multipoles (resulting in rather small frequencies in comoving momentum space). In fact, during the finalization of this paper, Fergusson et al. (2010a) have considered a feature

bispectrum and extracted 31 polynomial modes in the data, which allowed them to investigate late time bispectra with a maximal frequency of $\omega_f = 5 - 10$ (in comoving momentum space). They did not find 3σ evidence for non-zero non-Gaussianity. The other two example bispectra typically have a lot more oscillations within the tetrahedral domain, resulting in many modes necessary to realize an acceptable correlation. Fortunately, both the resonant and non-BD bispectrum have an amplitude that scales with the frequency. Therefore, a small improvement in correlation could lead to a significant improvement in the ability to constrain the model by measuring these modes in the data and reconstructing the primordial signal.

Complementarily, we have proposed a different basis expansion, based on Fourier functions instead of polynomials. This still leads to the necessary computational reduction one is after and therefore is a perfectly valid alternative. Such expansion is more relevant for resonant and non-BD scenario, since the feature bispectrum can already be transformed into Fourier modes analytically, using identities. We have shown that Fourier modes are much more efficient for the resonant bispectrum, reducing the number of modes necessary to establish the same correlation as polynomial modes by at least a factor of 5. For the non-BD bispectrum both Fourier expansion and polynomial expansion are difficult. Correlation increases fast with the addition of modes, but quickly converges to a fixed value, where the fixed value decreases a function of frequency. We believe that this is due to the exact form of the bispectrum, which has many small features near the edges of the tetrahedral domain. One might hope that some of these very small features are washed out when you compute the multipole equivalent, although that would be very time consuming since the non-BD shape is not of the factorized form. We hope to investigate this in a future attempt. In addition we have investigated three toy-spectra, not based on any particular model, which have a different oscillating orientation compared to the three theoretical models. Expanding these in Fourier modes show similar improvement compared to polynomial expansion as the resonant bispectrum. In general, we therefore believe that Fourier expansion is much more effective in the expansion of oscillatory spectra compared to polynomial basis expansion.

From an observational viewpoint, it seems that for resonant inflation only a limited number of modes contribute significantly in reproducing the original bispectrum. This allows us to consider only those modes that contribute substantially. This holds independent of the phase and frequency of the signal and is due to the specific form of this bispectrum, which oscillates (primarily) in the k_t direction. Because the modes that are important for the reconstruction of the original bispectrum are independent of the frequency, this also implies that when one would observe these modes in the data one could in fact find evidence for much larger frequencies than discussed here, simply because for larger frequencies these modes will also matter but their respective weight will be smaller. Despite the fact that we could not optimally expand the non-BD bispectrum using Fourier modes, we did look into the three toy-sepctra. We found that other modes are important. Moreover, the modes that are important directly represent the orientation of the oscillating spectrum and could therefore discriminate between different bispectra quite effectively. If this conclusion holds after forward projection into multipole space, measuring a number of Fourier mode functions in the CMB data would present an efficient way of deducing whether oscillations are present in the data and could give both an indication of the frequency and the shape of the primordial bispectrum.

Acknowledgments The author would like to thank Jan Pieter van der Schaar, Pier Stefano Corasaniti, Ralph Wijers, Licia Verde, Ben Wandelt, James Fergusson, Xingang Chen and Michele Luguori for very helpful discussions and comments on the manuscript. He would also like to thank the hospitality of DAMTP, Cambridge, where this paper was finalized. The author was supported by the Netherlands Organization for Scientific Research (NWO), NWO-toptalent grant 021.001.040.

Chapter 6

WMAP 7 Constraints on Oscillations in the Power Spectrum

P. Daniel Meerburg
In preperation

Abstract: We use the WMAP 7 data to place constraints on oscillations supplementing an almost scale-invariant primordial power spectrum. Such oscillations are predicted by a variety of models, some of which amount to assuming there is some non-trivial choice of the vacuum state at the onset of inflation. In this chapter we will explore data-driven constraints on two distinct models of initial state modifications. In both models the frequency, phase and amplitude are degrees of freedom of the theory for which the theoretical bounds on these parameters are rather weak: both the amplitude and frequency have allowed values ranging over several orders of magnitude. This requires many computationally expensive evaluations of the model CMB spectra and their goodness-of-fit, even in a Markov Chain Monte Carlo (MCMC), normally the most efficient fitting method for such a problem. To search more efficiently we first run a densely spaced grid, with only 3 varying parameters; the frequency, the amplitude and the baryon density. We obtain the optimal frequency and run an MCMC at the best fit frequency, randomly varying all other relevant parameters. To reduce the computational time of each power spectrum computation, we adjust both comoving momentum integration and spline interpolation (in l) as a function of frequency and amplitude of the primordial power spectrum. Applying this to the WMAP 7 data allows us to improve existing constraints on the presence of oscillations. We confirm earlier findings that certain frequencies can improve the fitting over a model without oscillations. For those frequencies we compute the posterior probability, allowing us to put some constraints on the primordial parameter space of both models.

6.1 Introduction

The observed statistical distribution of temperature fluctuations in the cosmic microwave background (CMB) is believed to be largely determined by the physics in the very early Universe. These CMB fluctuations were sourced by quantum fluctuations during an epoch of accelerated

expansion early on in the history of the Universe, known as inflation. They then induce curvature perturbations in the geometry of spacetime, which are preserved after horizon crossing during inflation. Once they re-enter the horizon at some later time, they couple to radiation and matter, becoming responsible for the observed statistical distribution of the CMB temperature fluctuations and the large scale structure (LSS) of matter in the Universe, respectively. Within the 6-parameter ΛCDM these initial conditions are described by only 2 parameters: the amplitude of the primordial power spectrum of curvature perturbations, and the tilt n_s describing the scale-dependence of the power spectrum to first order. By investigating the CMB power spectrum, one is therefore able to probe high energy physics in the early Universe. Given the large number of models describing the physics of the early Universe (see e.g., Chen (2010) for a recent overview), these two parameters (within 6-parameter ΛCDM) are not sufficient to be able to discriminate between various proposed models. Additional degrees of freedom, derived from the statistical analysis of the late time distribution of temperature (CMB) or density (LSS) fluctuations, could potentially be used to break the degeneracy between various models. Possible extensions to the 6-parameter model include tensor degrees of freedom (gravitational waves), higher order corrections to the scalar *and* tensor power spectra, and deviations from Gaussianity measured through high order correlation functions. The ultimate goal, of course, is to formulate a theoretical prediction of what these degrees of freedom should be, and constrain these from the observational data, just as we constrain the 2 parameters describing the tilt and the amplitude of the primordial power spectrum of scalar perturbations. The additional constrained parameters test the uniqueness of a proposed model and thus contribute to the understanding of the physics governing the early, inflationary Universe.

In this chapter we will consider modifications to the primordial power spectrum. In particular, we will search for evidence of oscillations in the almost scale-invariant spectrum. The motivation to search for these modifications is provided by a rapidly increasing number of theoretical models expecting such features in the primordial correlation statistics: e.g., in the power spectrum (Martin & Brandenberger 2001; Easther et al. 2001, 2003; Danielsson 2002; Easther et al. 2005b; Schalm et al. 2004, 2005; Greene et al. 2004; Achucarro et al. 2011; Cai et al. 2010; Flauger et al. 2010; Chen 2010c; Jackson & Schalm 2010; Chen 2011; Jackson & Schalm 2011) and bispectrum (Chen et al. 2007, 2008; Meerburg et al. 2009; Flauger & Pajer 2011; Meerburg et al. 2010a; Chen 2010c). In this chapter we will only consider oscillations in the power spectrum. Fergusson et al. (2010a) have attempted to constrain features in the bispectrum and Meerburg (2010b) have proposed a method to effectively search for features in the CMB bispectrum; recently LSS data has been proposed to search for these type of bispectra (Cyr-Racine & Schmidt 2011). The frequency, amplitude, and possibly the phase and first order scale dependence of these features are determined by the detailed physics of inflation. Detecting or constraining these parameters would help us determine the precise physics of inflation.

This is not the first time features in the power spectrum will be explored. Notably Easther et al. (2005a,b), Martin & Ringeval (2004, 2005), Covi et al. (2006), Hamann et al (2007), Pahud et al. (2009), and Flauger et al. (2010) have all investigated possible oscillations in the primordial power spectrum. Although there has been no clear detection, the data certainly allow small oscillations, as can be seen from fig. 6.1. Previous analyses were done on WMAP

1-, 3-, and 5-year data (as well as SDSS data). In this chapter we will aim at extending and improving the analysis using the latest WMAP release, WMAP 7 (Komatsu et al. 2011). One of the key hurdles in searching for oscillations in the data is the frequency of the oscillations. The frequency of the oscillations in the primordial power spectrum is a free parameter which spans several orders of magnitude for many of the proposed models. When probing the joint likelihood of our cosmological model, the large range in frequencies results in a number of issues (Groeneboom & Elgaroy 2008).

Foremost, it requires a high resolution in 'sample space', i.e., looking for the best fit parameter to the data requires a large number of computations of the CMB power spectrum from the primordial power spectrum (using, e.g., CAMB (Lewis et al. 2000) or CMBfast (Seljak & Zaldarriaga 1996)). This necessitates an efficient computing scheme. Unfortunately, the computation of the late time power spectrum from the primordial one involves a convolution of the transfer function $\Delta_l(k)$ with the primordial power spectrum $P(k)$. The appearance of rapid oscillations in both the transfer functions as well as the primordial spectrum requires smaller and more frequent steps in the integration in comoving momentum space k, increasing the required computational time for each run significantly.

Secondly, the late time spectrum must to be computed for each angular scale l. Usually, given the smoothness of the primordial spectrum, it suffices to do a spline interpolation on a number of C_l. This reduces the computational time, since the computation of the transfer functions[1] is the most time-consuming part. With the addition of oscillations on top of the smooth primordial spectrum, the number of C_l necessary to obtain an accurate fit of the interpolated C_l will depend on the frequency of the primordial oscillations. As this frequency increases, at some point *all* l will need to be considered in order to resolve the superimposed oscillations. Computing all l requires us to compute all transfer functions, which again increases the computational time. Lastly, there will often be a number of frequencies that tend to improve the fit within the large range we explore, rather than just a single one. For example, if a frequency ω is a good fit to the data, there is a fair chance 2ω will be a good fit as well. This is a major issue, as a multidimensional parameter space is most effectively searched through a Markov Chain Monte Carlo (MCMC), which is a random process. Therefore, when the frequency is not fixed, an MCMC approach to the best fit is not efficient as the likelihood is spiked and the random nature of the MCMC chain will lead to frequent tunneling of one local maximum to another within the multidimensional joint likelihood. Flauger et al. (2010) showed a way to circumvent this issue by first taking a large number of samples fixed on a grid. Obviously a grid does not allow us to vary all parameters within our cosmological model, as the number of samples grows quickly with the dimensionality of parameter space. Instead, the priority lies in varying the initial conditions, i.e., the oscillatory component of the primordial power spectrum. Once the best fit has been determined (within the prior frequency range, and the grid resolution), one can run MCMCs with a fixed best fit frequency determined through the grid. The joint likelihood is expected to no longer contain local maxima and an MCMC should converge quickly. This eventually allows us to put constraints on the amplitude and perhaps the first-order scale dependence of the amplitude of the oscillatory feature.

[1] These can generally not be precomputed, as they depend on all relevant parameters of the theory, which are varied when searching for the best fit.

Although there are a large number of different features and oscillations predicted by a variety of models, in this chapter we will focus on only two, distinct, theoretically motivated modified primordial power spectra. We will introduce these modified power spectra in section 6.2. In section 6.3 we explain how to optimize the search by making the numerical computation of the power spectrum frequency-dependent. To find a best fit frequency before we apply the MCMC, we use grid sampling. We report our findings in section 6.4. Once we have established a best fit value of the frequency we run an MCMC with that best fit value for two models. The results are discussed in section 6.5 and we compare these to theoretically derived constraints on primordial parameter space. We conclude in section 6.6.

6.2 Models

In this chapter we will consider 2 models. Both of these models can be related to possible deviations from a Bunch-Davies (BD) vacuum. Usually when computing primordial correlation statistics we consider very early times when all relevant modes are deep inside the horizon. In this adiabatic regime, we can define a vacuum: the BD vacuum. However, since inflation is an effective theory, for very small wavelength modes (large k) at early times one would expect interactions to become important. Strictly speaking this prevents the identification of the empty state with the BD vacuum. Therefore, defining a BD state for every mode is no longer possible as interactions will affect the mode function. In this chapter we are primarily motivated by pragmatism and try to constrain possible deviations from a BD state by looking at cosmological data.

The first of the models discussed in this chapter is known as the Boundary Effective Field Theory (BEFT) of initial-state modifications (Schalm et al. 2004; Greene et al. 2004). In this proposal one fixes an initial *time* where one calculates corrections to the usual BD initial condition using BEFT. The result is an explicitly scale-dependent Bogolyubov parameter $\beta_{\mathbf{k}}$, describing the vacuum modification, resulting in a strongly scale dependent power spectrum.

The second model is known as the New Physics Hypersurface (NPH) of initial-state modifications (Danielsson 2002; Easther et al. 2002). In the NPH scenario one traces every momentum mode back to some large *physical* scale of new physics M and imposes the standard flat space vacuum state (corresponding to positive frequency modes only), mode by mode, resulting in a k independent $\beta_{\mathbf{k}}$, thus it only gives rise to a small departure from scale-invariance after taking into account the slow-roll evolution of the Hubble parameter. A recent effort by Jackson and Schalm Jackson & Schalm (2011) to compute the effects of integrating out a massive scalar field onto the low energy effective theory of inflation seems to put the NPH proposal on a more fundamental footing. In this work, ignorance of the ultraviolet completion of the effective theory is directly implemented in the (effective) action of inflation. It was shown that integrating out an arbitrary massive scalar field can affect the vacuum, resulting in similar corrections to the inflationary power spectrum as those in the NPH scenario. We will use the form of the spectrum as predicted by Jackson & Schalm (2011).

The phenomenological form of the power spectrum for the two models are (Greene et al. 2004)

$$P(k) \quad = \quad P_0(k)\left[1 + \beta k A_0 \sin\left(2A_1 k + \phi\right)\right] \tag{6.1}$$

for the BEFT scenario, while for the new Jackson scenario (Jackson & Schalm 2011) (from here on NPH)

$$P(k) \quad = \quad P_0^* \left[a_0 + a_1 \ln k/k_* + (a_2 + a_3 \ln k/k_*) \sin \left(a_4 \ln k/k_* + \zeta \right) \right]. \tag{6.2}$$

In eq. (6.1) β is a parameter determined by UV physics, and is typically expected to be $\mathcal{O}(1)$, while A_0 and A_2 are parameters determined by the BEFT. $P_0(k)$ is the power spectrum from canonical slow-roll inflation in a BD state. This contains a possible tilt n_s. In eq. (6.2), P_0^* is the scale independent power spectrum, and a possible tilt is not included. The reason for this difference is that the oscillatory correction in the NPH scenario can have a tilt a_3 that differs from the tilt $a_1 = n_s - 1$. We constrain each parameter independently and apply theoretical dependencies between those parameters afterwards. For example, A_1 in the BEFT model is related to the amplitude A_0. The reason not to implement this directly is because separating these parameters results in a much 'cleaner' sampling, because the frequency causes the likelihood to be highly irregular. We therefore prefer to vary the frequency separately. We can apply the theoretical constraint relating the amplitude to the frequency once the posterior distribution has been determined. We refer to papers by Greene et al. (2004) for the BEFT and Jackson & Schalm (2011) for the NPH power spectrum, respectively. For the BEFT we are interested in A_0, A_1 and ϕ, while for the NPH model we will consider a_2, a_3, a_4 and ζ. a_0 and a_1 are considered the zero-order contribution to the power spectrum and are constrained through n_s and P_*.

6.3 Code Optimization

We have modified the CAMB code (Lewis et al. 2000) in order to efficiently search for oscillations. The modification is built upon work done by Martin & Ringeval (2004, 2005). The late time power spectrum basically is a convolution of the transfer function $\Delta_l(k)$ and the primordial power spectrum $P(k)$

$$C_l \quad \propto \quad \int_0^\infty k^2 dk P(k) \Delta_l^2(k). \tag{6.3}$$

The transfer function contains all the physics that governs the evolution of the Universe and describes an interplay between radiation pressure and gravitational collapse. This causes the transfer function to be highly oscillatory. Consequently, to gain sufficient accuracy one has to numerically integrate over k-space with an adequate number of steps. The transfer functions are not only a function of comoving momentum, but also depend on the angular scale l. To increase the speed of the code, one usually does not compute the complete transfer matrix, but rather limited number of l, and spline interpolates between them.

If the primordial power spectrum is smooth both the integral over comoving momentum and the spline interpolation in l can be done with limited samples. Once we allow for (rapid) oscillations in the primordial power spectrum, the number of samples needs to be increased.

Numerically, the most time consuming application is the number of transfer functions we will have to compute. Therefore the fewer transfer functions we compute, the faster we can determine the angular power spectrum C_l. Obviously, if the number of oscillations in our

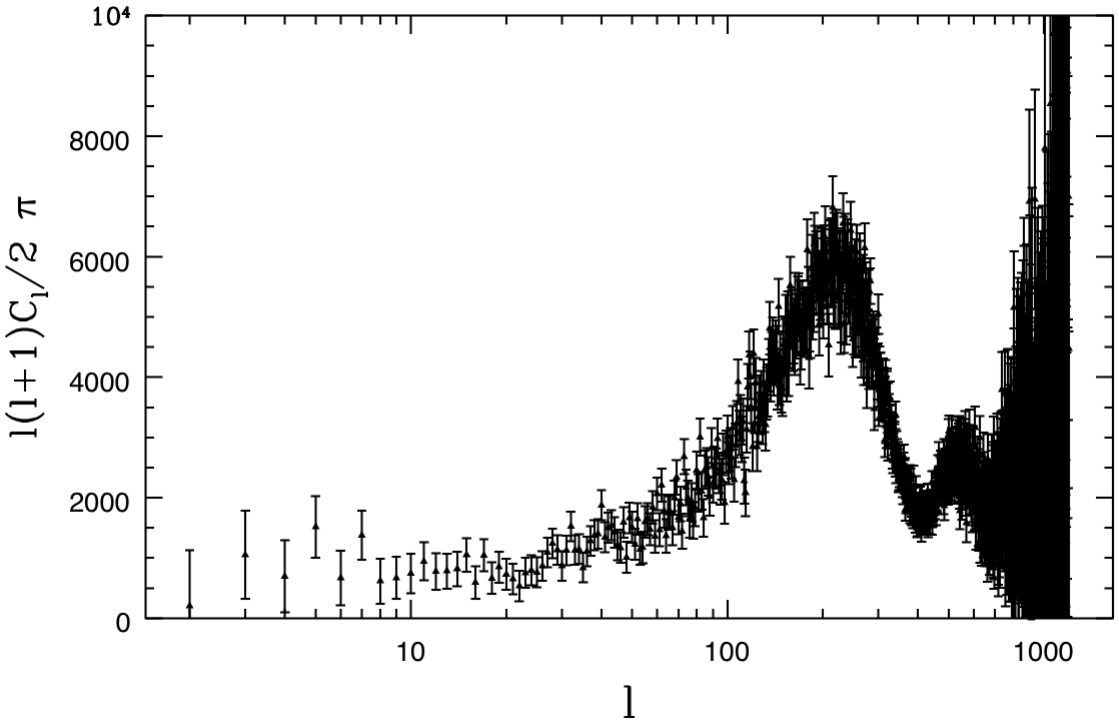

Figure 6.1: WMAP 7 data with added errors from measurement and cosmic variance. The error bars are derived from the diagonal terms in the Fisher matrix. The multipole moments are slightly coupled (inducing off-diagonal elements), so a correct treatment of errors requires use of the entire Fisher matrix, which is done when calling the likelihood WMAP code.

primordial spectrum is large, we will need many transfer functions to resolve those using a spline interpolation. If the wavelength of the primordial signal is $\delta l \sim 1$, no spline interpolation will resolve the primordial signal.

6.3.1 Frequency Dependent Power Spectrum Computation

We have modified CAMB and COSMOmc in order to optimally compute the power spectrum given a primordial frequency. Fixing all parameters to the best fit WMAP 7 values (Komatsu et al. 2011), except for the frequency in the primordial power spectrum, we investigated the convergence in the late time power spectrum when we increased the sampling in k- and l-space. As a null hypothesis, with 'optimal' accuracy, we took the primordial spectrum without oscillations. By plotting the frequency against the sample increase required to retain the same accuracy, we obtained an estimate of the optimal sampling for every frequency. As an example, in fig. 6.2 we show the increase in l-sampling required to obtain an accurate computation of the power spectrum as a function of the BEFT frequency A_1. Only for relatively low frequencies $A_1 < 2000$ and $a_4 < 60$ can we use splines in l-space to determine the CMB power spectrum. At

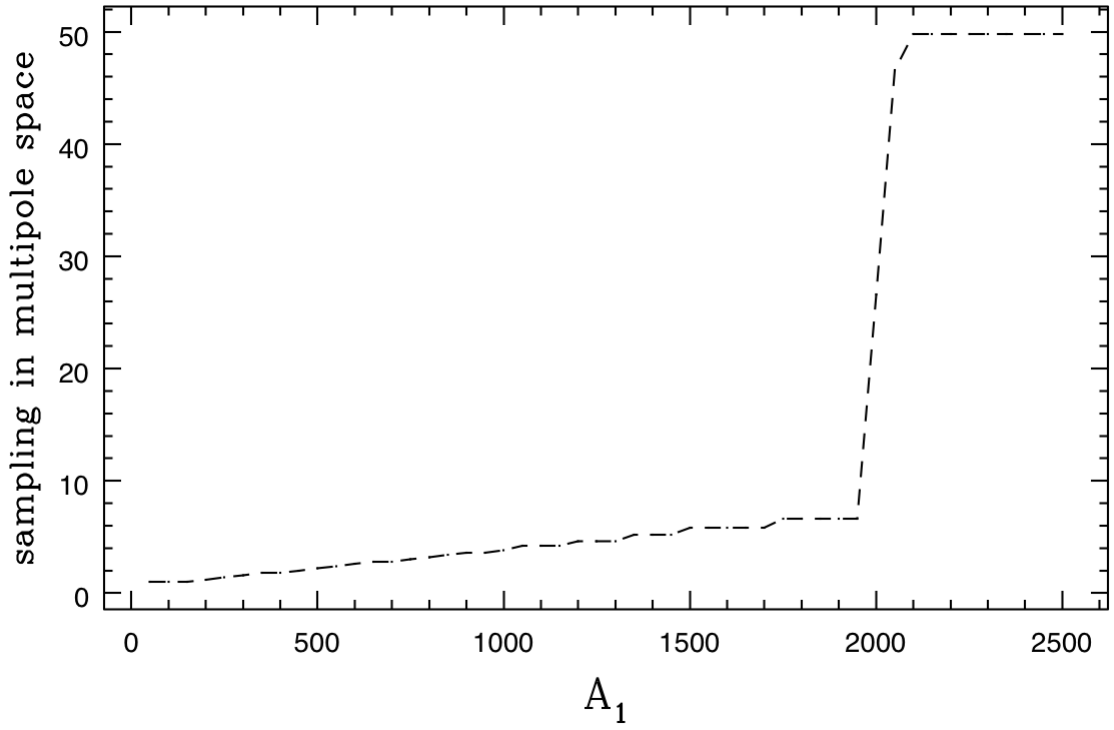

Figure 6.2: Sampling increase for splining in multipole space as a function of A_1 for the BEFT model. We fixed $A_0 = 10$, its largest possible value. As the frequency A_1 increases, it requires an increasing number of neighboring l in the interpolation to maintain the same accuracy (which was set by WMAP best fit standards). Around $A_1 \sim 2000$ the sampling factor increases beyond 50, which implies that all l's are required (i.e., splining is not sufficient and one needs to compute the full transfer matrix).

higher frequencies we need to compute the full transfer matrix in order to resolve oscillations[2]. For the sampling in k-space the optimal sampling was derived when computing all transfer functions. We find that over the range of frequencies we will probe in this chapter we need to approximately double the resolution in k-space when doing the integration.

6.4 Grid Sampling

Given the number of free parameters describing the initial conditions (at least A_s, n_s and the frequency and amplitude of the oscillating correction), many computations of the power spectrum are required to obtain the best fit (distribution). A commonly used approach is to apply a Markov Chain Monte Carlo (MCMC), randomly pick values of a set of parameters, and based on the likelihood of the computed run, reject or accept this point to be part of our

[2]The sampling is from a set of predetermined l (e.g., $l = \{1, 2, 3, ...100, 150, ...1000, 1100, ...\}$). The number of points used in the interpolation is increased within this set of predetermined l until this set is unable to achieve the same accuracy as one would obtain by computing all l. This set is based on efficient computation of a smooth primordial spectrum.

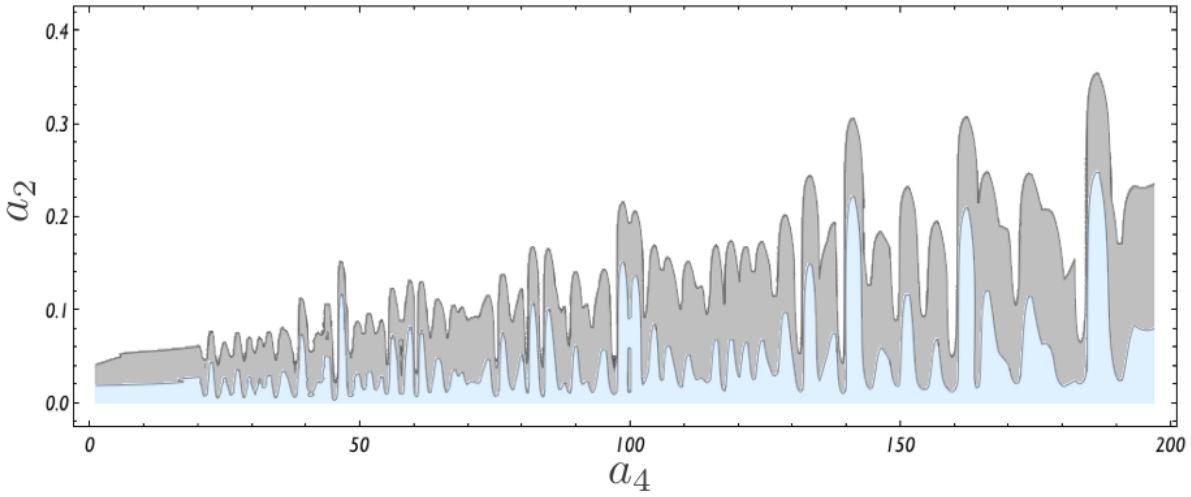

Figure 6.3: (NPH) The 68% and 95% confidence levels for amplitude a_2 versus the frequency a_4 for the low frequency grid $1 \leq a_4 \leq 200$ marginalized over the baryon density. There are many local peaks in the marginalized likelihood. The two most likely grid points are ($a_4 = 46.5$, $a_2 = 0.056$) and ($a_4 = 98.41$, $a_2 = 0.147$). Towards higher frequencies, larger amplitudes are allowed by the data but do not necessarily represent significant improvements of the overall fit.

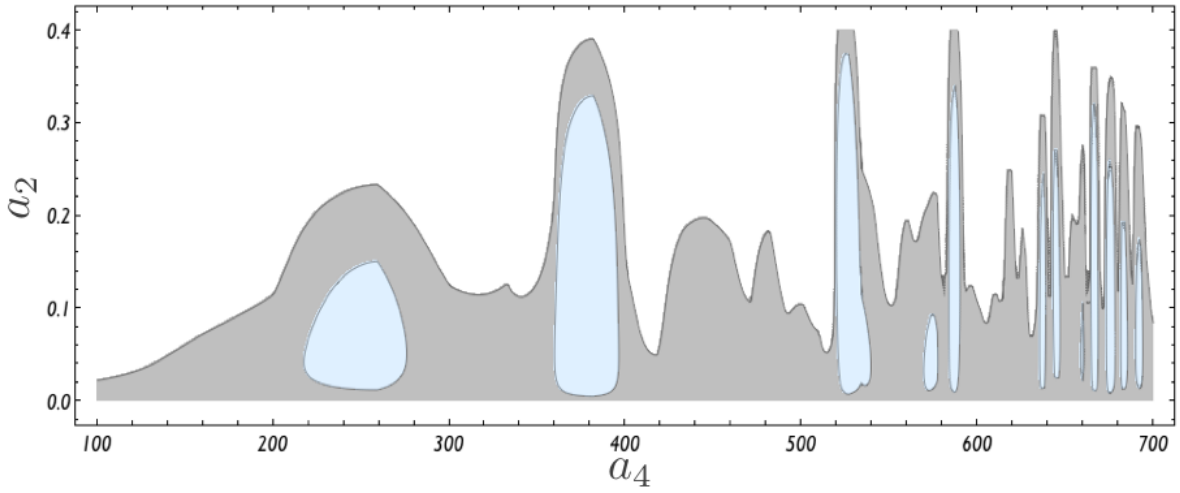

Figure 6.4: (NPH) The 68% and 95% confidence levels for a part of the high frequency grid. Note that the absence of a match between the low and high frequency regimes is artificial. In the high frequency grid, there are only 6 samples for $100 \leq a_4 \leq 400$. Additionally, a 100% likelihood was assumed within one grid. For very high frequencies, the wavelength of the oscillations is too small to be resolved at large angular scale ($l < 200$). Primordial oscillations are resolved only in the second peak and beyond. Very high frequencies can produce glitches in the large scale wing of the first peak in the angular power spectrum. For some specific frequencies and amplitudes these glitches can match outliers in the wing of the first peak of the observed angular power spectrum resulting in $\Delta\chi^2 \simeq 12$ for $a_4 \sim 998$.

parameter probability estimate. This approach is highly efficient, as we only compute a limited number of points in the multidimensional likelihood to determine the posterior distribution. Once we add oscillations to the primordial power spectrum an MCMC method becomes less efficient, as the likelihood is expected to become irregular in the coordinate of the oscillation (the frequency) and there are many local maxima in the likelihood function. At some point, we will leave this local maximum and end up in another. Therefore constraints on the initial conditions are hard to recover since the MCMC will constantly move to different local maxima in likelihood space[3]. Flauger et al. (2010) proposed a different approach. Instead of an MCMC over all parameter space, they first considered a grid over a limited number of parameters. The parameters to vary in the grid should be those that determine the primordial conditions (e.g., the frequency). They also identified the baryon density Ω_b to be degenerate with the amplitude and frequency of the oscillations, as it influences the height of the first peak. Finally they showed that oscillations in axion monodromy models of inflation are a better fit primarily to the first peak in the angular power spectrum, and hence could mimic some of the effects produced by the baryon energy density Ω_b. The grid therefore samples at least 3 parameters (amplitude, frequency and baryon density). The other parameters are fixed to the WMAP 7 (Komatsu et al. 2011) best-fit values (see table 1.1). As such the null hypothesis (no oscillations) is the fit to beat. The advantage of the grid is that one probes the likelihood completely, although with fixed conditions for most of the ΛCDM parameters. With the assumption that these other parameters are not (strongly) correlated with the parameters that we vary, we should be able to determine the absolute maximum in the likelihood rather than a local one. Once this local maximum has been determined, we can perform an MCMC with the frequency set to the best fit value, and allow all other parameters to vary. This should probe the likelihood of, e.g., the amplitude more efficiently.

We ran three grids, one for the BEFT model and two for the NPH model. While in this chapter we have opted to avoid details about the exact theoretical prediction of the parameters in each model, we decided to consider two grids for the NPH model focusing on two frequency regimes a_4. This is motivated by the fact that the NPH model is not expected to have a significant number of oscillations, as the theoretically predicted frequency is slow-roll suppressed, i.e., $a_4 \propto \epsilon$. We have therefore investigated low frequencies ($1 \leq a_4 \leq 200$) with 200 log-spaced samples, and a high frequency regime ($100 \leq a_4 \leq 1000$) separated into 500 logarithmically spaced samples. The amplitude a_2 runs from 0 to 0.4 in 120 and 200 equidistant steps, respectively. For the NPH grid $a_3 = \zeta = 0$. For the low frequency grid we set $0.021 \leq \Omega_b h^2 \leq 0.026$ in 10 equidistant steps, resulting in a total of 240.000 grid points, while for the high frequency regime we considered $0.02 \leq \Omega_b h^2 \leq 0.027$ in 16 equidistant steps, with a total of 1.6 million grid points.

In figures 6.3 and 6.4 we show the confidence contours for grids obtained for the NPH model of the frequency versus the amplitude in the low and high frequency regime, marginalized over the baryon energy density Ω_b. Peaks are areas in which the fit is best for non zero values of the amplitude of the modification, while valleys represent frequencies which are not a good fit to the data and the best fit is no modification. For example one can consider a likely frequency

[3]Constraints on other parameters unrelated to the initial conditions can be recovered with sufficiently large samples.

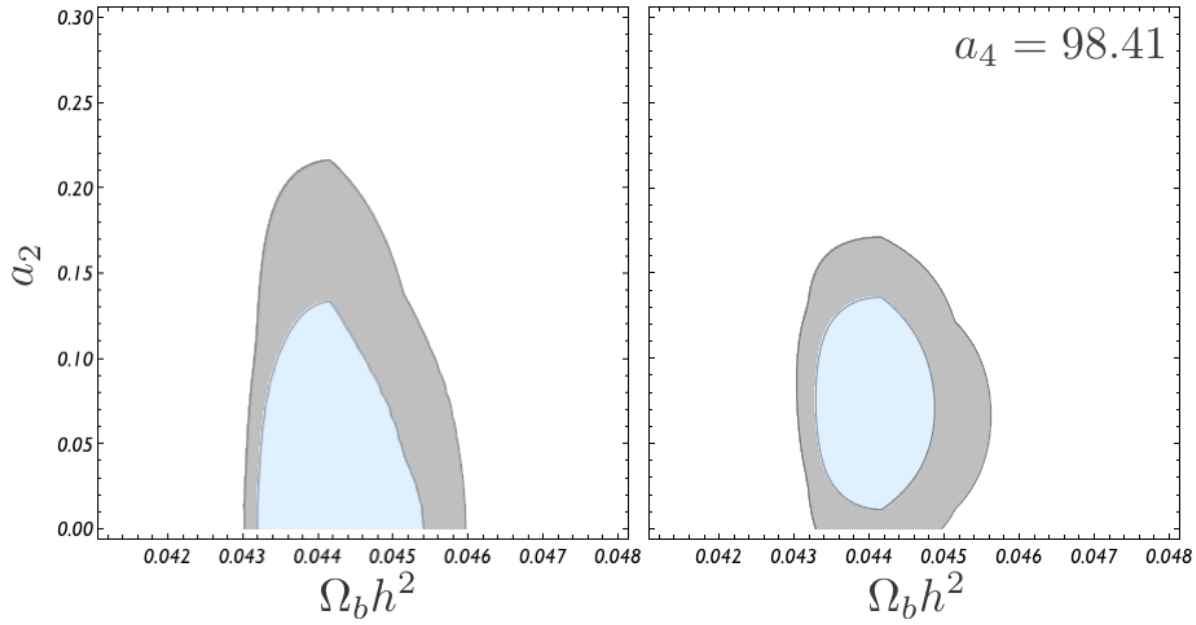

Figure 6.5: (NPH) The 68% and 95% confidence levels of $\Omega_b h^2$ versus a_2. The left shows the joint likelihood after marginalizing over the frequency a_4 (low frequency grid), while the right shows the joint likelihood for the best fit frequency $a_4 = 98.41$.

(peak) and plot the probability of the amplitude for that frequency to find that the most likely value for the amplitude is non zero. In fig. 6.5 we show the joint likelihood contour plot for the amplitude a_2 and the baryon density $\Omega_b h^2$ marginalized over the frequency a_4. We find that $a_2 = 0$ is the most likely value in the low frequency grid. We also derived the joint likelihood for one of the best-fit frequencies ($a_4 = 98.41$) to show that the best-fit point has a non-zero amplitude a_2 with almost 95% CL.

For the high frequency regime we show the effect of the highest frequencies on the marginalized amplitude a_2 in fig. 6.6. Note that the confidence levels are determined assuming that the grid contains all possible values the frequency could have. In other words, one would hope that for either extremely large or small frequencies the likelihood of the fit would go to zero. The problem is that it does not, and therefore we can only determine the confidence levels within a prior determined parameter domain. We have partly motivated this domain on theoretical arguments. Observationally, data is the limiting factor.

It should be obvious that many different frequencies represent good fits. As we had foreseen, this complicates running a large MCMC for *all* relevant cosmological parameters. The best fit point we find for the NPH model ($\Delta\chi^2 \sim 12$) centers around very high frequencies, with $a_4 = 980$ and $a_2 = 0.39$. In fact for this frequency, the best fit amplitude probably lies outside the domain of $0 \leq a_2 \leq 0.4$. This implies a relatively large number of primordial oscillations. The angular power spectrum for such high frequencies has most of its oscillations damped, since this frequency is only resolved at scales beyond the first peak. Martin & Ringeval (2004) showed that for these scales the amplitude will be suppressed, similar to how the overall power

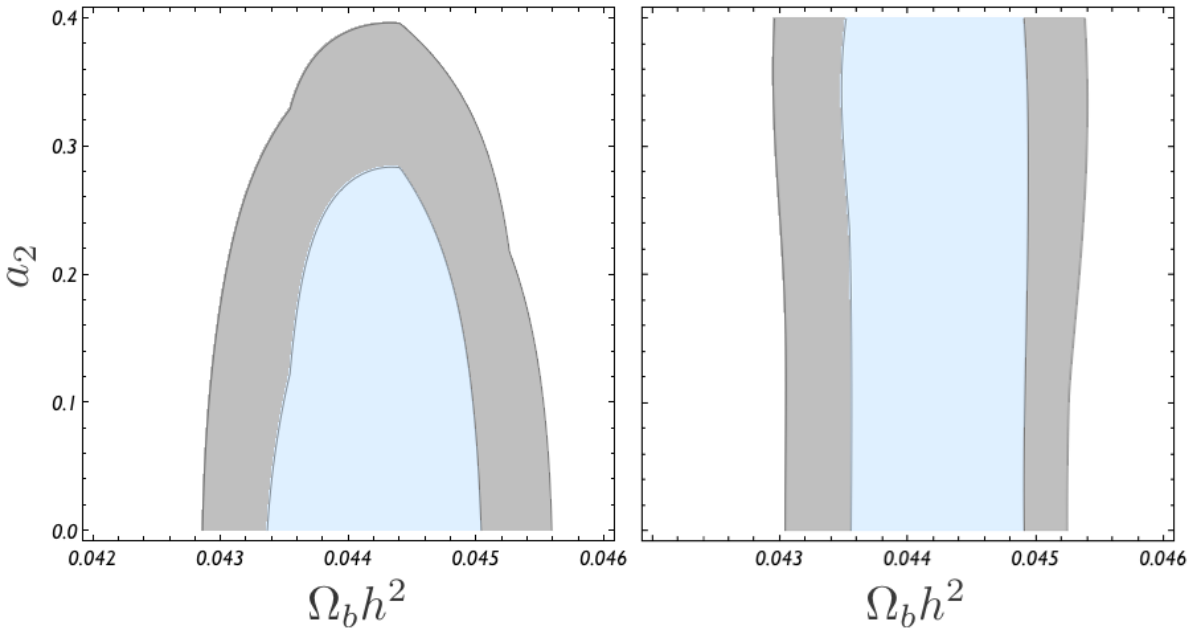

Figure 6.6: (NPH) The 68% and 95% confidence levels of $\Omega_b h^2$ versus a_2. The left shows the joint likelihood after marginalizing over the frequency a_4 for the high frequency grid up to $a_4 \leq 700$ (63 frequencies). If one includes all frequencies in the marginalization the constraint on a_2 disappears (right).

is damped. Given the large measurement error at these scales, we would not expect the fit to improve that much. It turns out that at large angular scales, the barely resolved high frequency causes glitches in the large scale wing of the first peak. The WMAP data contains outliers in the large scale (small l) wing of the first peak (attempts have been made to understand this in terms of a feature in the primordial spectrum, see e.g. Covi et al. (2006) and recently a principal component analysis has been performed to search for such a feature by Dvorkin & Hu (2011)); these are fitted for very specific unresolved frequencies of the primordial power spectrum. These glitches are expected as the angular power spectrum will no longer be able to sample all oscillations appearing in the primordial power spectrum. We have not been able to get rid of these glitches by increasing the overall accuracy of the numerics of the code.

In the low frequency regime we obtain two frequencies that give an improvement of $\Delta\chi^2 \sim 6$ compared to no oscillations. In the next section we will discuss the results from an MCMC run for one of these frequencies, with a varying phase and amplitude. For the NPH model, the results can be summarized as follows: for frequencies $1 < a_4 < 200$, $a_2 < 0.13$ at 68% and $a_2 < 0.21$ at 95% confidence. For frequencies up to $a_4 = 700$, $a_2 < 0.29$ at 68% and $a_2 < 0.39$ at 95% while for higher frequencies the amplitude is not constrained within the bounds of the grid. The most likely value within the low frequency grid is $a_4 = 46.5$ with an amplitude of $a_2 \sim 0.056$.

For the BEFT model, there is a constraint set by BEFT validity on the value of A_0. A BEFT approach to initial state modifications is only valid if the physical momentum is smaller

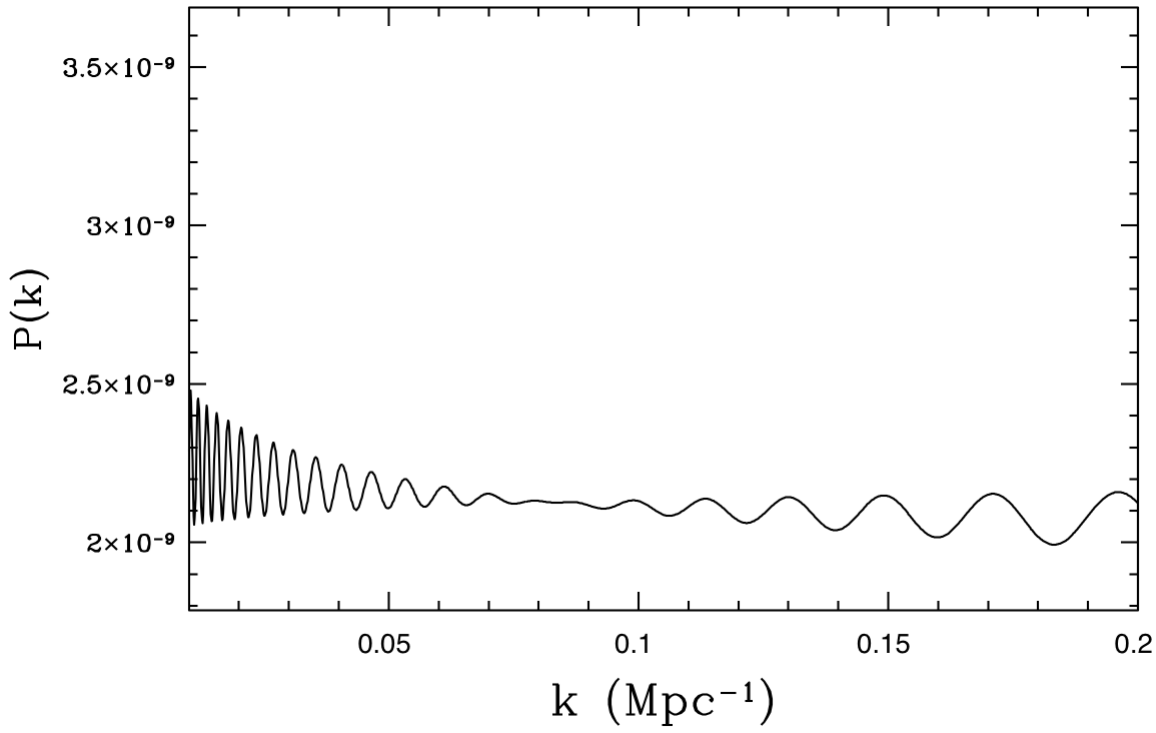

Figure 6.7: The primordial spectrum from the NPH model with $a_4 \sim 46$, $a_2 = 0.14$, $a_3 = -0.04$ and $\phi = 1.5$.

than the cutoff scale at the boundary, i.e., $k/a_0 < M$. The maximum value of k is set by the smallest observable scale in the CMB, $k_{max} = \mathcal{O}(0.1)$ and therefore we deduce $A_0 \leq 10$ from eq. (6.1). The number of samples in A_0 is set to 120 equidistant values between 1 and 10. For BEFT the frequency is not constrained by slow-roll and is proportional to M/H; the scale of new physics divided by the Hubble scale. Consequently the effective frequency can be quite high. We sample 700 logarithmically spaced steps between 10^3 and 10^4 making up a total of 1,344,000 grid points.

The likelihood confidence contours for the BEFT model are different from those of the NPH model (fig. 6.10). Most importantly, the resulting contour does not have a vanishing amplitude, which was assumed as a prior based on the theoretical form of A_0. The best fit point in the grid is given by $A_1 = 7708$ with $A_0 = 10$ corresponding to an improvement of $\Delta\chi^2 \simeq 13$.

6.5 MCMC and Model Constraints

We analyzed the WMAP data using Monte Carlo Markov sampling with a fixed frequency derived from the grids in the previous section for both BEFT and NPH model. We set a relatively weak Gelman and Rubin criterion of $R - 1 < 0.1$ (Gelman & Rubin 1992) due to time constraints. We plan to set this to a more conservative constraint of $R - 1 < 0.01$ which

is common in most literature. The Gelman-Rubin diagnostic R relies on parallel chains to test whether they all converge to the same posterior distribution by considering the variance of the parameters in each chain compared to the variance of the same parameters over all parallel chains. Convergence is diagnosed once the chains have 'forgotten' their initial values, and the output from all chains has become indistinguishable ($R - 1 = 0$). For the NPH model, we ran 8 parallel MCMC with $a_4 \sim 45.9$, close to the best fit point in the low frequency grid. In addition to a_2 we allowed both a_3 and ζ to vary. We assumed flat priors with $-0.6 \leq a_2 \leq 0.6$, $-0.5 \leq a_3 \leq 0.5$ and $0 \leq \zeta \leq \pi$. We found that with the limited number of samples in each chain achievable within the computing time (max. $60,000$), we could barely reach the relatively weak Gelman-Rubin criterion for the parameters τ, z and the amplitude of the primordial power spectrum P_* using 50% of the samples in the chain. After we ran 8 chains, we derived a proposal covariance matrix. We used this covariance matrix to speed-up the convergence. Obviously, if the proposal is derived from chains that have not converged sufficiently, we will not recover the true posterior distribution. As such, using a proposal covariance matrix is not without risks. With the addition of a covariance proposal we obtained $R - 1 < 0.003$ over 4 chains. For

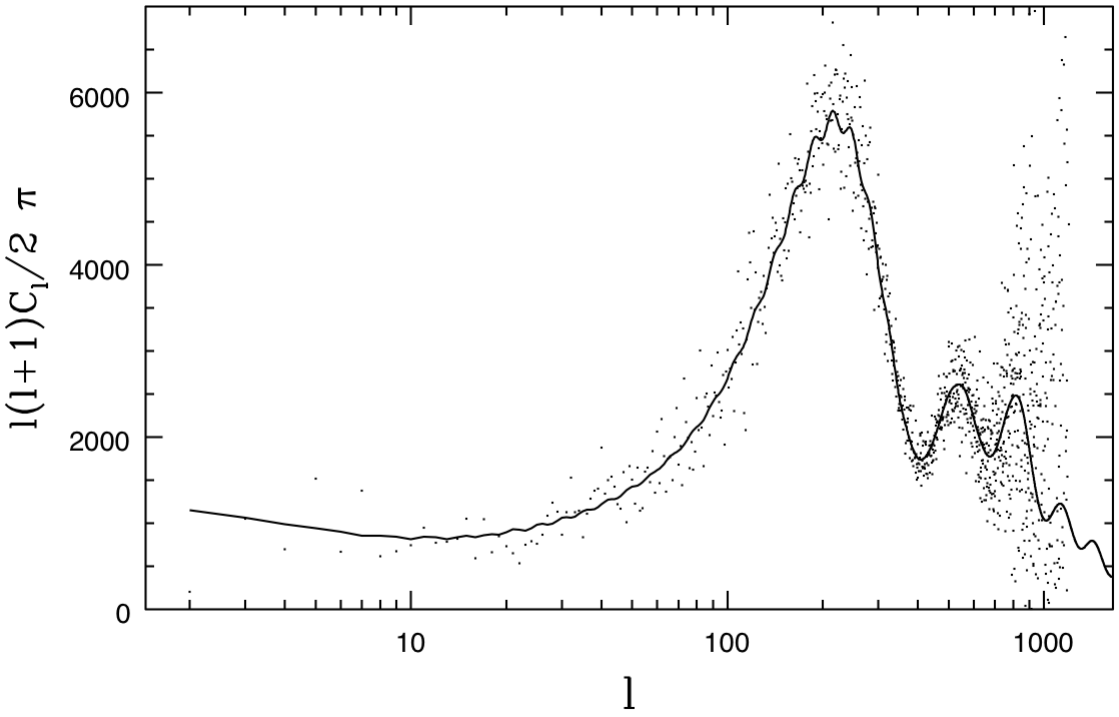

Figure 6.8: Best fit for the NPH model from the MCMC chain, with a fixed frequency $a_4 \sim 46$. The improvement of the fit compared to no oscillations is $\Delta\chi^2 \sim 12$. From the grid we found that an equally good fit is for a frequency $a_4 \sim 98$, about double this frequency. We find that the best fit amplitude is rather large, $a_2 \sim 0.14$, with the 68% level still allowing zero amplitude (fig. 6.11). The amplitude is similar to the best fit amplitude of the axion monodromy model derived by Flauger et al. (2010). However, this best fit was at a translated frequency of about 150.

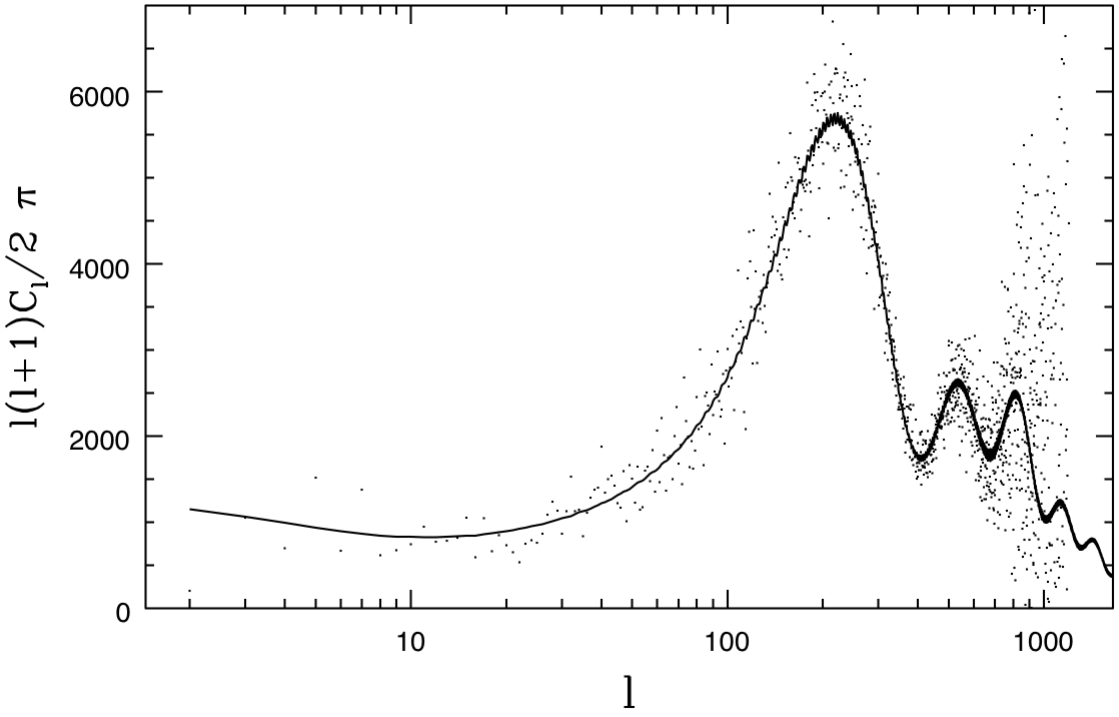

Figure 6.9: The best-fit angular power spectrum derived from the grid for the BEFT model with $A_0 = 10$ and $A_1 = 7708$, resulting in $\Delta\chi^2 \simeq 13$ relative to no oscillations.

a slightly faster convergence of our MCMC we fixed the tilt n_s to its best fit value, without oscillations. The result for some of the marginalized and joint likelihoods is shown in fig. 6.11 for the NPH model. Note the high correlation between a_2, the amplitude of the oscillation, and a_3, the scale dependence of the oscillatory correction.

From the MCMC for the NPH scenario we deduce marginalized best fit values $a_2 = 0.14^{+0.17}_{-0.16}$, $a_4 = 0.04 \pm 0.04$ and $\zeta = 1.5^{+0.4}_{-1.5}$ with a fixed frequency of $a_4 = 45.9$ and a best fit point with $\Delta\chi^2 \simeq 12$. To relate these to constraints on the NPH model one must investigate the relation between a_2, a_3 and a_4 as derived by Jackson & Schalm (2011). The exact form of these parameters is presented in eq. (31) of their paper[4] We derive a relation between a_2, a_3 and a_4

$$\frac{a_2}{a_3} \simeq -\frac{1}{a_4}\frac{M}{H}\left[5/2 + \ln M/H\right]^{-1}. \tag{6.4}$$

Here we took the lower limit $\Lambda = M$, where we consider $M/H > 10^2$. The observational limits were derived for $a_4 \sim 46$, which puts a bound on $M/H > 10^3$ for $\epsilon < 0.01$ (slow-roll). This results in a theoretical constraint $a_2/a_3 < -2.3$. We can derive a similar constraint for the upper limit $\Lambda = \frac{1}{2}(H + M^2/H)$, or $\Lambda/H \sim 1/2(M/H)^2$. It turns out that a_4 in this limit can

[4]We derived these constraints in the assumption $\epsilon_1/\epsilon_2 \sim 1$ and $H_* \sim H$ due to weak scale dependence (Jackson & Schalm 2011).

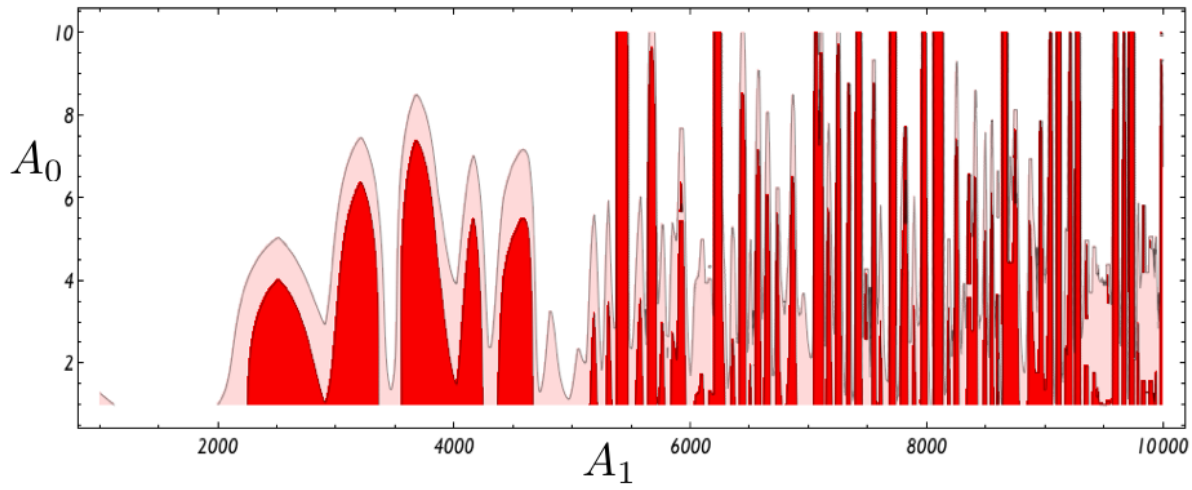

Figure 6.10: (BEFT) The 68% and 95% confidence levels for amplitude A_0 versus the frequency A_1 marginalized over the baryon density. Again, there are many local maxima in the likelihood. The cutoff in amplitude A_0 is set by theoretical constraints. If we would allow larger amplitudes ($A_0 > 10$), the likelihood contours would likely move towards higher A. The most likely grid point is ($A_1 = 7708$, $A_1 = 10$).

only have a negative sign. For a sine, this means the amplitude picks up a minus sign. We find $M/H > 10^3$ and

$$\frac{a_2}{a_3} \simeq \frac{1}{a_4}\frac{M}{H}(2 - \ln M/2H)\left[5/2 + \ln\frac{1}{2}(M/H)^2\right]^{-1} \tag{6.5}$$

which together results in $a_2/a_3 > 5.86$.

In fig. 6.12 we show the joint likelihood together with the theoretically allowed values of a_2 and a_3 for a given frequency a_4. If this frequency is a valid signal, a minimal improvement of the confidence contours could exclude this oscillation to be due to a NPH altogether.

For the BEFT model, the bound on the frequency is not that stringent and we simply chose the best fit frequency of $A_1 = 7708$ obtained from the grid. Note that this high frequency requires us to compute all l in order to resolve the primordial oscillation. We varied $0.1 \leq A_0 \leq 10$ and $0 \leq \phi \leq 2\pi$.

The 8 parallel MCMC for the BEFT model resulted in an improvement of $\Delta\chi^2 \sim 16$, with all parameters satisfying the Gelman-Rubin diagnostic, except A_0 and ϕ. The reason why the A_0 variable does not satisfy the criterion is due to the fact that we inserted a cutoff in the prior of $A_1 = 10$. As we have seen from the grid samples, the best-fit value of the amplitude is probably larger than this, and we are therefore probing only the tail of the distribution. We did run 6 parallel chains with an estimated covariance matrix from the first chains, but it did not lead to satisfying results. Although the Gelman-Rubin diagnostic was reduced for both A_0 and ϕ to $R - 1 < 0.1$, the recovered distributions were clearly derived from a wrong covariance matrix, as the MCMC posterior distribution did not trace the mean likelihood. Again, we would like to emphasize that limited time did not allow us to run the chains for too long. We expect to

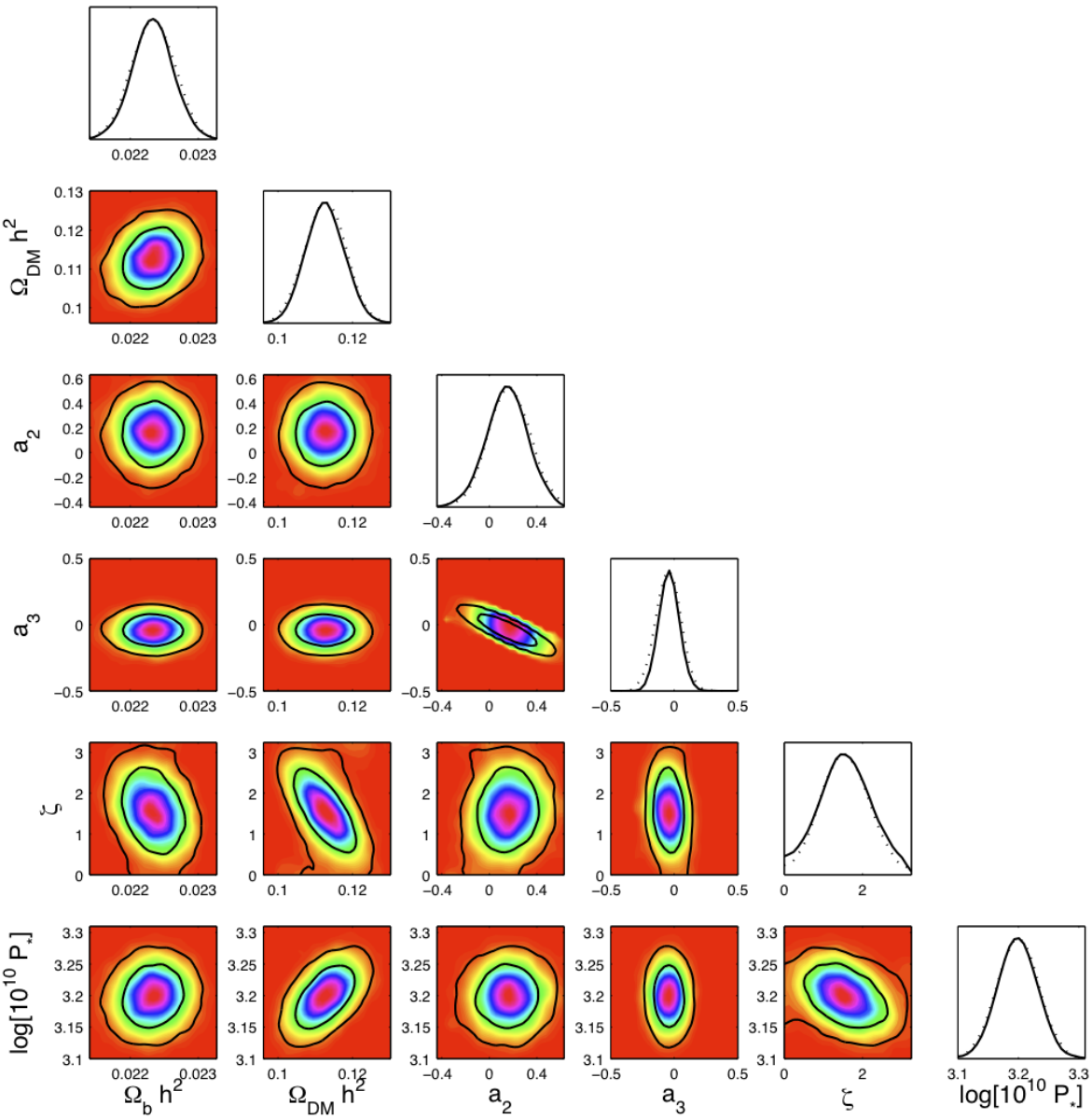

Figure 6.11: Marginalized and joint likelihoods for several parameters in the new NPH scenario. The parameters a_2 and a_3 show strong correlation. The tilt a_3 is also correlated with τ and Ω_Λ (not shown).

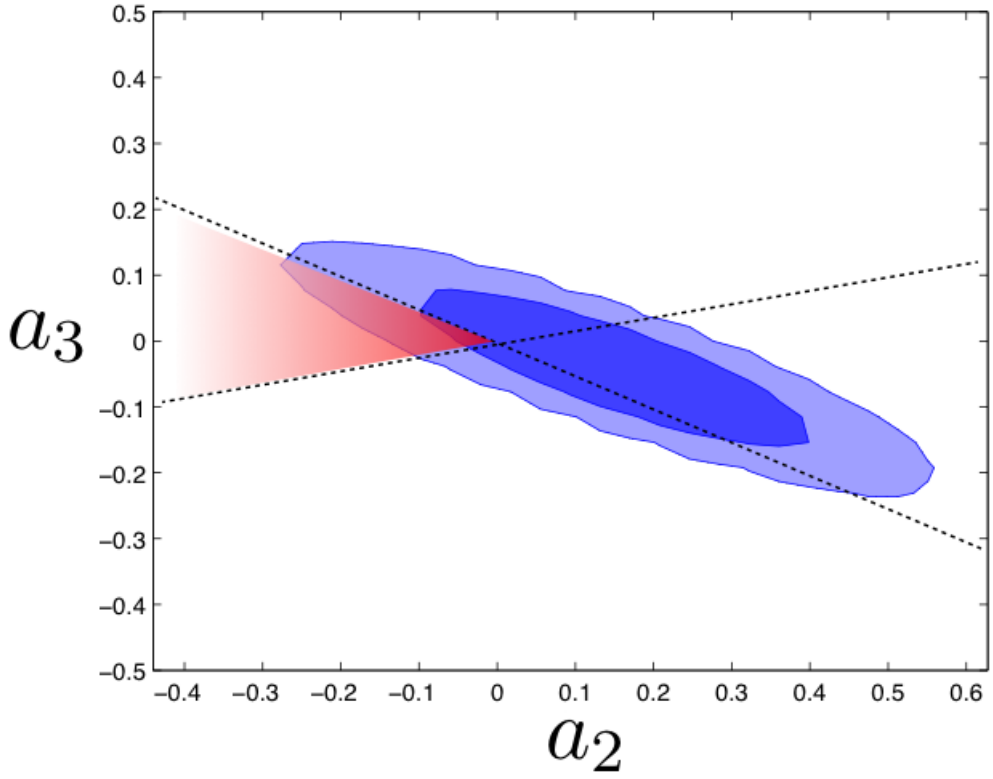

Figure 6.12: Joint likelihood for a_2 and a_3 together with the theoretical constraints relating a_2 to a_3 showing 68% and 95% confidence levels. The region (shaded, red) allowed within the NPH model constitutes a small part of the observationally allowed region. It shows that a marginal improvement of the confidence contours could exclude NPH as the source of this oscillation.

be able to derive a more accurate covariance matrix using longer chains. We will investigate longer chains in future work.

For a comparison with theoretical bounds, we derive a relation between A_0 and A_1 as $A_1 \simeq A_0 \times (M/H)$ (see Greene et al. (2004) table 1). For a fixed frequency of $A_1 \sim 7708$ and $A_0 > 3.2$ at 68% confidence, we conclude that $M/H \lesssim 3 \times 10^3$ at 68% confidence. We expect that the best fit amplitude lies closer to 0.1 and consequently $M/H \lesssim 10^3$. Note that we assumed that the coupling constant is $\beta = 1$. However, β is $\mathcal{O}(1)$ which allows for small increase of the total amplitude, βA_0, which will probably allow for a better fit. If we are in the tail of the best fit amplitude, it could significantly constrain the confidence level of A_0. We are currently running MCMC to investigate this and will to report our findings in the near future.

6.6 Conclusions

We used the latest WMAP data to constrain oscillations on top of an almost scale invariant primordial power spectrum. We argued that the primary difficulty in constraining these models

with the data is related to the irregular likelihood function of the frequency of these oscillations. There are many equally good fits and these fits are at discrete frequencies. For all other parameters one tries to constrain in a ΛCDM model of the Universe, such degeneracy usually does not exist and it is quite sufficient to scan parameter space using an MCMC. In order to avoid the random jumps from one maximum in the likelihood to another, it is preferable to keep the frequency fixed. However, the frequency is one of the key parameters being constrained, and it would make no sense to fix it prior to analyzing the data. In order to estimate the frequency before the MCMC we applied a grid search for the best fit. This grid quickly becomes incalculable for a large number of parameters (which is the primary reason to run an MCMC instead) and we only varied the parameters characterizing the modification of the power spectrum as well as the baryon density, which previously had been identified to be correlated with the amplitude of the oscillatory correction. Once the best fit had been established we performed an MCMC with the frequency fixed to its best fit value.

We found that the addition of oscillations on top of the smooth power spectrum can improve the overall fit. An improvement up to $\Delta\chi^2 \sim 16$ was found once we applied an MCMC with a predetermined best fit value of the frequency.

In the NPH scenario very high frequencies should lead to unobservable amplitudes. For completeness we analysed such high frequencies. We found that for some specific frequencies we could obtain an improvement in the goodness of fit up to $\Delta\chi^2 \sim 12$ compared to no oscillations. Analyzing the angular power spectrum with these extreme primordial frequencies revealed that these frequencies are in fact unresolved at large angular scales (small l), where the wavelength of the oscillating primordial power spectrum is smaller than the angular distance between subsequent l's. This leads to small glitches in the slope of the first peak of the angular power spectrum, which fit some of the observed outliers in the slope of the first peak in the observed power spectrum. As such, high frequencies, although unresolved, could account for some of the large scale effects we observe in the data, and improve the overall fit. We conclude, however, that these oscillations cannot be caused by NPH modifications and must be due to a different model, possibly an axion-monodromy inflation type model (Flauger et al. 2010).

Although in the grid for the low frequency domain $1 \leq a_4 \leq 200$ of the NPH model we could only found one frequency with an improvement of $\Delta\chi^2 \simeq 6$, we ran an MCMC near this frequency to obtain an improvement of $\Delta\chi^2 \simeq 12$ corresponding to a best-fit amplitude of $a_2 = 0.16$. This is close to the best-fit amplitude found by Flauger et al. (2010) with an improvement of $\Delta\chi^2 \simeq 11$ (WMAP-5). Their improvement appeared at a (translated) frequency of $a_4 \sim 150$. It is interesting to note that there appear to be subsequent improvements at low frequencies close to 50, 100 (WMAP-7) and 150 (WMAP-5) which could be an additional hint we might be looking at an oscillation as opposed to noise, because the improvement appears at equidistant intervals in frequency.

Using an MCMC we have been able to put constraints on several primordial parameters. In particular, we derived constraints on the amplitude a_2, the tilt a_3 and the phase ζ of the oscillatory correction in the NPH model. We used these observational constraints to test the NPH model by implementing the relation between various parameters as predicted by the NPH model. This shows that the signal found at this frequency could originate from NPH modifications to the primordial power spectrum. However, an improvement of the confidence

levels would exclude this possibility.

For the BEFT model we ran one grid and we found that even before we ran an MCMC we could achieve an improvement of $\Delta\chi^2 \simeq 13$. This improvement corresponded to a frequency $A_1 = 7708$ and an amplitude $A_0 = 10$. We ran an MCMC around this frequency and found that the theoretical bound on A_0 does not probe the observationally best fit point. This is signaled by a sharp increase of the marginalized probability near $A_0 = 10$. We found a best fit improvement of $\Delta\chi^2 \sim 16$ compared to no oscillations with a phase $\phi = 1.8\pi$ and an amplitude of $a_2 = 9.9$. The best fit values of A_0 and A_1 are theoretically allowed by a BEFT modification of the primordial power spectrum. We will later investigate the improvement after we drop the theoretical bound. It seems plausible that this could improve the fit as the marginalized probability for A_0, given the sharp rise of the probability near the cutoff ($A_0 = 10$). The amplitude of the oscillatory correction is also proportional to a coupling constant β which we set to 1. Once we relax this assumption, larger amplitudes are theoretically allowed through βA_0. We expect this could significantly tighten the confidence levels of A_0, possibly near its maximum.

It is hard to assess the significance of the improved fits. We could think of two possibilities: investigate simulated CMB data without oscillations with realistic beam and noise, and investigate simulated CMB data *with* oscillations and realistic beam and noise. For example, if we can find oscillations in simulated data without oscillations, this would suggest oscillations can well be mimicked by noise. Complementarily, generating realistic CMB data including primordial oscillations, can we recover these oscillations by analyzing the data? In particular, it would be interesting to determine whether there exists a threshold amplitude for oscillations to be recovered. Preferably we should generate a large number of maps in order to see if we can assess a probability of recovering oscillations from the data. We will report our findings in future work.

For now, we conclude that a primordial power spectrum with no oscillations is consistent with the data for most amplitude and frequency of the primordial signal. For some frequencies, a non-zero amplitude of the oscillatory corrections to the power spectrum seems to be preferred by the data. These signals can be investigated, and used to constrain primordial parameter space, possibly signifying some of the detailed physics driving inflation.

Acknowledgments The author would like to thank Jan Pieter van der Schaar, Ralph Wijers and Mark Jackson for very helpful discussions and comments on the manuscript. He would particularly like to thank Christophe Ringeval who provided some of the essential code which was used throughout this paper. The author was supported by the Netherlands Organization for Scientific Research (NWO), NWO-toptalent grant 021.001.040.

Epilogue

We end this thesis with a summary of our main conclusions. In addition we will briefly address some of the advances that have been made in the time since the publication of the papers presented in this thesis. We end this section with a perpective on some the directions we would like to investigate in the near future.

The aim of our research was to constrain possible deviations from a Bunch Davies (BD) vacuum at the onset of inflation. We have motivated the presence of BD modifications on the grounds that the assumption of a free quantum field description in the very early Universe is bound to be wrong. We considered the effects modifications of the vacuum state might have on the primordial distribution of quantum fluctuations. In particular we have computed its effects on the primordial bispectrum. We have shown that the bispectrum can be enhanced significantly by a small correction to the BD assumption. In short, the reason why such an enhancement is expected is due to the fact the bispectrum is a measure of the interaction of the field(s) of inflation. Although interactions are weak, deviations from a BD state introduce a non-zero particle density, which boosts the amplitude of the bispectrum. It is generally expected (Holman & Tolley 2008) that higher order correlation functions result in further enhanced corrections, although this still has not been shown explicitly for correlation spectra beyond the trispectrum (Chen et al. 2009).

We also found that the bispectra predicted by initial state modifications usually contain oscillations. The presence of these oscillations has made it very difficult to obtain stringent constraints on the Bogolyubov parameter β, a parameter used to quantify BD modifications. In ch. 5 we considered a recently developed method for reconstructing primordial bispectra into an orthogonal set of factorizable basis functions. In the introduction we explained that factorizability is required to use the KSW estimator of the bispectrum. We showed that the proposed method of factorization requires a decomposition of the original bispectrum into many terms, each of which is of the factorizable form. We also showed that the number of terms required can be reduced by looking at a different set of factorized basis functions. We hope to apply this method to WMAP data in the near future.

It has long been predicted that the power spectrum also contains small oscillating features if there is a small deviation from BD state. The frequency of these oscillations is generally related to the frequency of features in the bispectrum. Ideally, we would like to use both the power spectrum and the bispectrum to constrain deviations from a BD vacuum state, particularly because the frequency of the oscillations in both spectra are typically related. There have been

several attempts to constrain oscillations in the primordial power spectrum, and in ch. 6 we extended this search. Compared to previous work we considered two different models with a large range of frequencies and we used the latest (7 year) WMAP data release. We first ran a densely spaced grid with most parameters fixed, before we did a full ΛCDM parameter search through a Monte Carlo Markov Chain (MCMC). This allowed us to determine a best fit frequency. After establishing the best fit frequency we ran MCMCs, varying all other relevant parameters. This gave bounds on the amplitude and phase of the oscillatory corrections. We compared these *observational* constraints to *theoretically* predicted modifications to the power spectrum. We found that there are several discrete frequencies that can improve the fit compared to no oscillations, and that these frequencies can be used to constrain primordial parameter space.

Obviously, observations have improved since our first paper. This puts slightly tighter constraints on derived parameters, in particular on β. We have incorporated the tightest constraint to date in ch. 4. We did not include bounds on $f_{\rm NL}^{enf}$, which has been constrained through LSS data analysis (Xia et al. 2011). A recent paper has looked into constraining oscillating bispectra using LSS data (Cyr-Racine & Schmidt 2011) and the future looks very promising for work in that direction. Although their work only represents a proof of concept, future extraction of constraints on such bispectra from LSS data could greatly improve existing bounds on initial state modifications.

In the near future (2013), Planck should tighten constraints on non-Gaussianities, in particular local type non-Gaussianity, with an expected 68% confidence level of $\Delta f_{\rm NL} \sim 5$. This could potentially rule out single-field slow-roll entirely. In addition, the accuracy of the data obtained by Planck could potentially allow us to detect oscillations in the bispectrum directly.

We aim to continue investigating possible hints of physics beyond the standard model using cosmological observations. First we want to finalize the last chapter of this thesis into a paper by improving some of the analysis. Secondly, we hope to combine observations from the power spectrum with observations from the bispectrum, which should constitute a sound complementary search. One of the things we are currently still working on is a simple test to constrain features in the primordial bispectrum. A full-fledged search is relatively time consuming because of the larger number of allowed frequencies, and we hope to construct a fast estimator to detect the potential presence of oscillations, independent of frequency, that also has enough signal to noise to be detected with current data (WMAP, ACT, SDSS). In addition, instead of looking at the various correlation spectra, we could consider the real space correlation function. This is very similar to a Fourier transform of the angular spectra and therefore oscillations should be more easily detectable as they should end up as sharps peaks in the correlation spectrum. We have explained that this measure is rather noisy, but if the amplitude of the primordial signal is large, we should be able to detect it. Lastly, as explained in ch. 6, we would hope to assess the significance of the 'detected' features in the power spectrum.

This combined effort, both theoretical and observational, should shed more light on the physics of inflation in the foreseeable future.

Appendix A

Constant ζ

In this appendix we will demonstrate ζ is constant after horizon crossing. This is relevant because it means that the distribution of fluctuations will not change after scales have exited the horizon. The fluctuations will eventually reenter the horizon at some later time, after which they start to evolve again.

We should expand the action to first order in derivatives of the fields (i.e., terms like $\partial\zeta D_i N^i$ should be neglected, while $\partial\phi$ and $\partial\rho$ are both zero-order and can thus be multiplied with all other derivatives), while maintaining all orders in the fields. We will assume that $N = 1 + \delta N$, where the expansion in derivatives of δN starts with a first order term. In addition, we assume that $D_i N_j$ is of first order in the derivatives. We can expand the Hamiltonian constraint (equation (1.37)) to first order in derivatives of the fields, while we can neglect $R^{(3)}$ as eq. (1.39) reveals it is of second-order in the field derivatives. We left with

$$2N^2V = -N^2(K_{ij}K^{ij} - K^2) - \left(\frac{d\phi}{dt}\right)^2. \tag{A.1}$$

This can be rewritten as

$$\begin{aligned} 2(1+\delta N)^2V &= 2V + 4V\delta N \\ &= 6H^2 + 12H\dot{\zeta} - 4HD_iN^i - \left(\frac{d\phi}{dt}\right)^2, \end{aligned}$$

or

$$V + \frac{1}{2}\left(\frac{d\phi}{dt}\right)^2 - 3H^2 + 2V\delta N = 2H(3\partial\zeta - D_iN^i). \tag{A.2}$$

The first three terms on the LHS are the Friedmann equation for the scalar field, so

$$2V\delta N = 2H(3\zeta' - D_iN^i). \tag{A.3}$$

We can evaluate the action using equation (A.3). On a solution of the Hamiltonian constraint

(eq. 1.37) the action reduces to

$$\begin{aligned} S &= \int d^4x\sqrt{h}(R^{(3)} - 2V) = \int d^4x\sqrt{h}(-2V - 2V\delta N) \\ &= \int d^4x a^3 e^{3\zeta}(-2V - 6H\dot{\zeta} + 2HD_iN^i) \\ &= \int d^4x a^3 e^{3\zeta}(-6H + \left(\frac{d\phi}{dt}\right)^2 - 6H\dot{\zeta} + 2HD_iN^i) \\ &= \int d^4x a^3 e^{3\zeta}(-6H + \left(\frac{d\phi}{dt}\right)^2 - 6H\dot{\zeta}) \\ &= \int d^4x \frac{d}{dt}\left(a^3 e^{3\zeta} H\right). \end{aligned} \tag{A.4}$$

The action can be written as a integral over a total time derivative. The equation of motion for ζ deduced from the action solely depend on the boundary conditions. Within the action formalism, this implies a constant ζ is a solution. This completes the proof that for single field slow-roll ζ is constant after horizon crossing. The physical reasoning being that outside the horizon we can neglect all spatial derivatives (the longer the wavelengths get, the smaller the spatial derivatives become). Since the expansion of the fields in powers of time derivatives starts beyond second-order, we conclude that a constant ζ is a superhorizon solution of the equation of motion to all orders in powers of ζ.

Appendix B

Building a Template

We have shown in ch. 2 that non-Gaussianities due to vacuum modifications for a canonical single inflaton action enhances in the collinear limit, corresponding to enfolded or squashed triangles. A template for these type of non-Gaussianities did not exist and we have proposed one in ch. 2. We have seen that the precise shape of this template proposal does not overlap perfectly, here we will translate the template into a KSW estimator to measure the associated amplitude $f_{\rm NL}^{\rm enf}$.

The estimators for different shapes of non-Gaussianity can be written as follows (Komatsu et al. 2009)

$$\begin{aligned} f_{\rm NL}^{\rm local} &= (F^{-1})_{11}S_1 + (F^{-1})_{12}S_2 + (F^{-1})_{13}S_3 + (F^{-1})_{14}S_4 \\ f_{\rm NL}^{\rm equil} &= (F^{-1})_{22}S_2 + (F^{-1})_{21}S_1 + (F^{-1})_{23}S_3 + (F^{-1})_{24}S_4 \\ f_{\rm NL}^{\rm enf} &= (F^{-1})_{33}S_3 + (F^{-1})_{31}S_1 + (F^{-1})_{32}S_2 + (F^{-1})_{34}S_4 \\ b_{src} &= (F^{-1})_{44}S_4 + (F^{-1})_{41}S_1 + (F^{-1})_{42}S_2 + (F^{-1})_{43}S_3. \end{aligned}$$

Here F_{ij} represents the Fisher matrix and is inversely proportional to the covariance, the overlap, between two bispectra. In case one has a Gaussian likelihood and only one parameter to fit, the Fisher matrix is equal to the inverse variance; $F_{ij} = 1/\sigma_\alpha^2$, with α the fitting parameter. Here it is given by

$$F_{ij} \equiv \sum_{2\leq l_1\leq l_2\leq l_3} \frac{B^{(i)}_{l_1l_2l_3} B^{(j)}_{l_1l_2l_3}}{\tilde{C}_{l_1}\tilde{C}_{l_2}\tilde{C}_{l_3}},$$

which is practically equal to eq. (2.39). Once more, the $B^{(i)}_{l_1l_2l_3}$ are the theoretical bispectra of the various non-Gaussian shapes. $\tilde{C}_l$ represents the total angular power spectrum, which contains both the CMB signal and the noise, i.e., $\tilde{C}_l = C_l^{cmb} b_l^2 + N_l$. The b_l is the beam transfer function, which is instrument dependent. If the beam is Gaussian it has the form $b_l \propto \exp[-l^2\sigma_b^2]$, where $\sigma_b = 1/2\sqrt{2\ln 2} \simeq 0.425{*}\text{FWHM}$[1]

[1]For a Gaussian beam we at full width half maximum we solve $e^{-\frac{1}{2\sigma^2}l^2} = 1/2$.

If the fisher matrix is close to diagonal. we can neglect possible leakage of one type of non-Gausssianity into another. Komatsu et al. (2009) showed that this was true at least for the first three bispectra. From the derived cosine in ch. 2 we can assume this should hold for our template proposal. In case of a diagonal Fisher we write

$$f_{\rm NL}^{\rm local} = S_1/F_{11}, \quad f_{\rm NL}^{\rm equil} = S_2/F_{22}, \quad f_{\rm NL}^{\rm enf} = S_3/F_{33}, \quad b_{src} = S_4/F_{44}.$$

We can now directly compute the pre-factor $\propto F_{ij}^{-1}$, without first inverting the full Fisher matrix.

In this paper we have proposed a template for the enfolded or squashed triangles, that should measure $f_{\rm NL}^{\rm enf}$

$$F(k_1,k_2,k_3) \;=\; 6\Delta_\Phi^2\left[\frac{1}{k_1^3k_2^3}+2\text{ perm}+\frac{3}{k_1^2k_2^2k_3^2}-\left(\frac{1}{k_1k_2^2k_3^3}+5\text{ perm}\right)\right]. \qquad \text{(B.1)}$$

Note that if we took a slightly different template, e.g., choosing a 4 instead of a 3 in the equation above, the following derivation of the estimator is completely analogous, and one simply needs to replace this 3 with a 4 from now on.

To quantify to what extent the enfolded template gives complementary information once applied to the data, as compared to local and equilateral templates, we should consider the scalar product between the different templates. As pointed out in the main text, the starting point for deriving the most optimal enfolded template shape function is $F^{\rm enf} = -F^{\rm equil} + c/k_1^2k_2^2k_3^2$, in terms of a general parameter c. We will plot the cosine of this template distribution with the equilateral template as a function of c, with c running from 0 to 4. The result is shown in fig. B.1. If we demand the template to have a definite (positive) sign, one should really only consider $c \geq 1$. In that case it should be clear that the optimal value, i.e. the smallest cosine equal to $\mathrm{Cos}(F^{\rm enf}, F^{\rm equil}) = 0.49$, is achieved for $c = 1$, as claimed in sec. 5 of ch. 2. For the cosine with the local template we find that it is more or less independent of c, as exemplified by the fact that $\mathrm{Cos}(F^{\rm enf}, F^{\rm local})_{c=3}/\mathrm{Cos}(F^{\rm enf}, F^{\rm local})_{c=1} \simeq 1.06$ and growing ever slower for larger c. For the cosine with the local template we find, for $c = 1$, that $\mathrm{Cos}(F^{\rm enf}, F^{\rm local}) = 0.67$. This is quite large, which we should have expected since we can imagine the local template to be a special case of the factorized enfolded distribution, for which only the endpoints of the line $x_2 + x_3 = 1$ are maximal, versus the whole line for the enfolded template. For completeness let us also mention the cosine between the local and equilateral template $\mathrm{Cos}(F^{\rm equil}, F^{\rm local}) = 0.41$.

Any deviations from scale invariance can simply be taken into account by replacing the power of n with $n - (n/3)(n_s - 1)$, with n_s the spectral index (see ch. 1). If one forgets about the pre-factors and divides out the k_1 dependence, the shape can plotted as a function of k_2/k_1 and k_3/k_1. This is shown in fig. 2.3.

Indeed the template shape maximizes when $k_1 = k_2 + k_3$. However it does not blow up at this limit, which is the case when the denominator is proportional to $k_1 - k_2 - k_3$. Such behavior can (possibly) not be mimicked, using factorizable templates. Therefore such a denominator would not be allowed, since a function with such a denominator cannot be split up into functions of individual comoving momenta k_1, k_2 and k_3.

There exists another approach (Smith & Zaldarriaga 2006), in which such a denominator is written as follows

$$\frac{1}{(k_1-k_2-k_3)^2} = \int_0^\infty te^{-t(k_1-k_2-k_3)}dt.$$

Now one has an integral over exponentials, which are factorizable. For eq. (2.13) this would imply

$$F^{\rm modin}(k_1,k_2,k_3) \quad \propto \quad \frac{1}{k_1k_2k_3}\sum_j \int_0^{\eta_0} \frac{\sin(\tilde{k}_j t)}{k_j^2} dt. \tag{B.2}$$

Obviously this introduces another integration, increasing computational time one is trying to win by factorizing. In (Smith & Zaldarriaga 2006) it was shown that the double integration can be done rather quickly for an equilateral template. In our case, this may not be possible because of the large number of oscillations in the cosine, which we expect to require a large number of quadrature points, when one replaces the integral over t by a weighted sum. For now, we will focus on the enfolded template shape of eq. (B.1) and leave the investigation of the method above for future work.

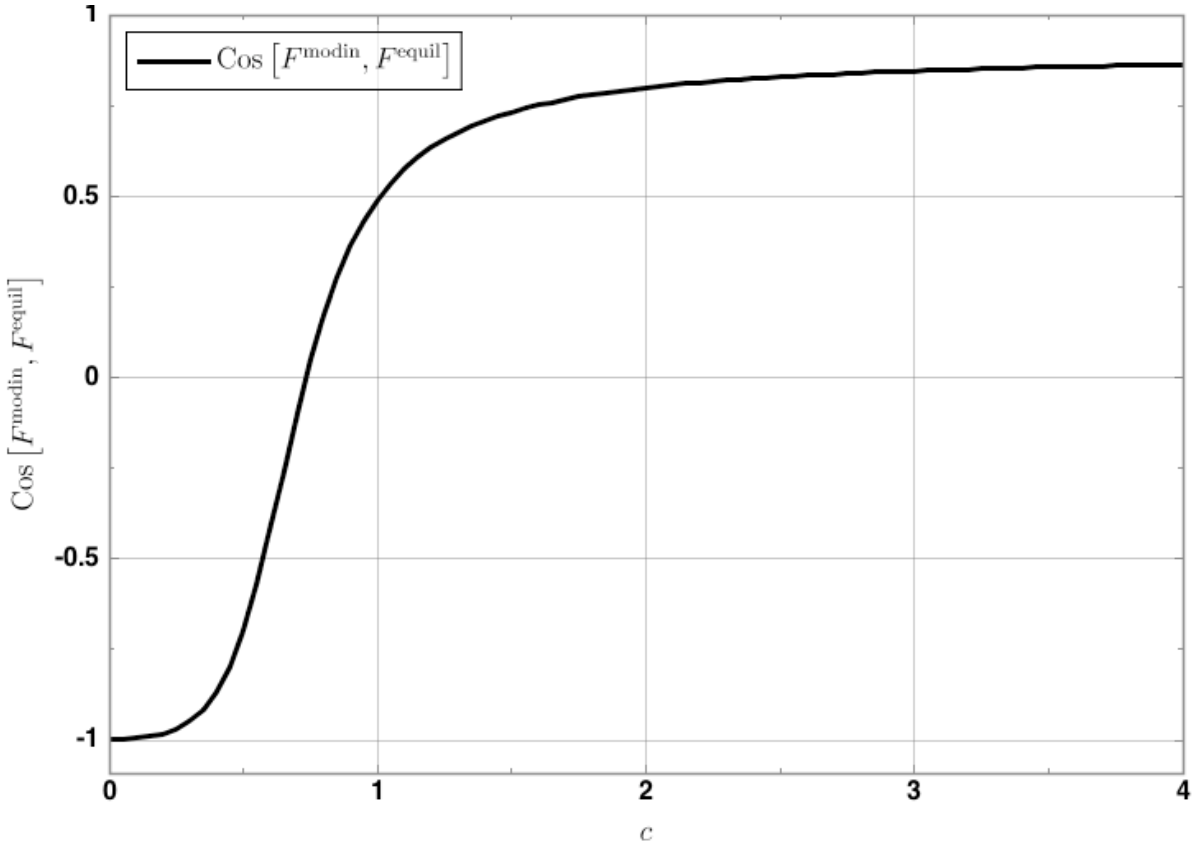

Figure B.1: The plot shows how the Cos($F^{\rm enf}, F^{\rm equil}$) changes as a function of c. The smaller the value of the cosine, the more distinct the two shapes are. As expected for $c = 0$ the cosine is -1, i.e., $F^{\rm enf} = -F^{\rm equil}$. Note that for $0 < c < 1$ the value of the cosine crosses zero. However these values of c can not be used, because the sign of the three-point function should be definite. For $0 < c < 1$ this is not the case and the reason for getting a smaller value for the cosine is due to cancellations between positive and negative parts of the shape function.

One can define the following maps

$$\alpha_l(r) = \frac{2}{\pi}\int k^2 dk \Delta_l(k) j_l(kr) \tag{B.3}$$

$$\beta_l(r) = \frac{2}{\pi}\int k^2 dk P_\Phi \Delta_l(k) j_l(kr) \tag{B.4}$$

$$\gamma_l(r) = \frac{2}{\pi}\int k^2 dk P_\Phi^{1/3} \Delta_l(k) j_l(kr) \tag{B.5}$$

$$\delta_l(r) = \frac{2}{\pi}\int k^2 dk P_\Phi^{2/3} \Delta_l(k) j_l(kr) \tag{B.6}$$

Here $\Delta_l(k)$ is the photon transfer function, introduced in sec. 6 of ch. 2, which is used to compute the (theoretical) angular power spectrum $C_l = (2/\pi)\int k^2 dk P_\Phi(k)\Delta_l^2$. $P_\phi(k)$ is the primordial power spectrum $P_\Phi(k) \propto k^{n_s-1}/k^3$. It can be seen that all maps have a different primordial power spectrum dependence (and consequently different dimension). These maps are required to set up an estimator that has the same k dependence as the template (B.1).

Using eq. (B.3) through (B.6) one can construct 4 'filtered' maps (recall that $\Delta T(\hat{n}) = \sum_{lm} a_{lm} Y_{lm}(\hat{n})$)

$$A(\hat{n}, r) = \sum_{l=2}^{l_{max}} \sum_{m=-l}^{l} \alpha_l(r) \frac{b_l}{\tilde{C}_l} a_{lm} Y_{lm}(\hat{n}), \tag{B.7}$$

$$B(\hat{n}, r) = \sum_{l=2}^{l_{max}} \sum_{m=-l}^{l} \beta_l(r) \frac{b_l}{\tilde{C}_l} a_{lm} Y_{lm}(\hat{n}), \tag{B.8}$$

$$C(\hat{n}, r) = \sum_{l=2}^{l_{max}} \sum_{m=-l}^{l} \gamma_l(r) \frac{b_l}{\tilde{C}_l} a_{lm} Y_{lm}(\hat{n}), \tag{B.9}$$

$$D(\hat{n}, r) = \sum_{l=2}^{l_{max}} \sum_{m=-l}^{l} \delta_l(r) \frac{b_l}{\tilde{C}_l} a_{lm} Y_{lm}(\hat{n}). \tag{B.10}$$

The sum over l runs from 2 to l_{max}, since the monopole and the dipole are hard/impossible to measure and WMAP (or any other instrument for that matter) can only measure up to a certain l_{max} based on the instrument's technical limitations. Now we need to set up a bispectrum that has the 'same' comoving momentum dependence as the template. For reasons that will become clear later, it is convenient to define the bispectrum related to local shape ($f_{\rm NL}^{\rm local}$). The local shape is local in real space and its template is exact, that is, the theoretical shape is equivalent to a factorized template

$$F(k_1, k_2, k_3) = f_{\rm NL}^{\rm local} \Delta_\Phi^2 2 \left(\frac{1}{k_1^3 k_2^3} + \frac{1}{k_2^3 k_3^3} + \frac{1}{k_3^3 k_1^3} \right). \tag{B.11}$$

It can be seen from the template that the shape is proportional to a product of power spectra ($n_s = 1$). We can thus use cyclic product of the angular maps (B.3) and (B.4). The (angular averaged) bispectrum can be written as

$$B_{l_1 l_2 l_3}^{\rm local} = 2 I_{l_1 l_2 l_3} \int_0^\infty r^2 dr \left[\alpha_{l_1}(r) \beta_{l_2}(r) \beta_{l_3}(r) + \text{cycl. perm} \right]. \tag{B.12}$$

One integrates over the comoving distance. The sampling rate (in r space) depends on the behavior of the transfer function Δ_l. In addition, $I_{l_1 l_2 l_3}$ is known the Gaunt factor introduced in ch. 1. It is given by (eq. (2.35))

$$I_{l_1 l_2 l_3} = \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}. \tag{B.13}$$

The first term in the enfolded shape is equivalent to the local shape, so we can use $B^{\rm local}_{l_1 l_2 l_3}$ to express (partly) $B^{\rm enf}_{l_1 l_2 l_3}$. The rest is obtained via carefully combining products of the angular maps (B.3) through (B.6)

$$B^{\rm enf}_{l_1 l_2 l_3} = 3B^{\rm local}_{l_1 l_2 l_3} + 6I_{l_1 l_2 l_3} \int r^2 dr \left[-\left(\beta_{l_1}(r)\gamma_{l_2}(r)\delta_{l_3}(r) + \text{cycl. perm}\right) + 3\delta_{l_1}(r)\delta_{l_2}(r)\delta_{l_3}(r)\right]$$

We can easily set up the skewness (KSW) estimator as explained by Yadav et al. (2007a,b); Komatsu et al. (2009)

$$S^{\rm enf} = S_{\rm prim} + S^{\rm linear}_{\rm prim},$$

where the first term is simply the term that represents the shape of the bispectrum (the cubic term, cubic in the filtered maps), while the second, the linear term (linear in the filtered maps), is added to minimize the effect caused by the inhomogeneous noise that breaks rotational invariance ($\hat{C}_l$ will have off-diagonal terms). The linear term should be constructed such that it minimizes the variance of the estimator. If the linear term is constructed as follows, this can indeed by achieved. One first has to derive the filtered maps A, B, C and D of eq. (B.7), (B.8), (B.9) and (B.10) that can be used to set up the cubic statistic estimator, $S_{\rm prim}$. Subsequently one takes the Monte Carlo average, $\langle S_{\rm prim}\rangle_{\rm MC}$. Now let us suppose that $S_{\rm prim}$ is constructed out of 3 filtered maps A, B and C. One can apply Wick's theorem to rewrite the average of a cubic product $\langle ABC\rangle_{\rm MC} = \langle A\rangle_{\rm MC}\langle BC\rangle_{\rm MC} + \langle B\rangle_{\rm MC}\langle AC\rangle_{\rm MC} + \langle C\rangle_{\rm MC}\langle AB\rangle_{\rm MC}$. Next, remove the MC average from the single maps and replace the maps within brackets with simulated maps. Our linear estimator becomes: $A\langle B_{\rm sim}C_{\rm sim}\rangle_{\rm MC} + B\langle A_{\rm sim}C_{\rm sim}\rangle_{\rm MC} + C\langle A_{\rm sim}B_{\rm sim}\rangle_{\rm MC}$. If we apply this trick and apply weighting functions we get

$$\begin{aligned} S^{\rm local} \equiv\ & 4\pi \int r^2 dr \int \frac{d^2\hat{n}}{w_3} \big(A(\hat{n},r)B^2(\hat{n},r) \\ & -2B(\hat{n},r)\langle A_{\rm sim}(\hat{n},r)B_{\rm sim}(\hat{n},r)\rangle_{\rm MC} - A(\hat{n},r)\langle B^2_{\rm sim}(\hat{n},r)\rangle_{\rm MC}\big), \end{aligned} \tag{B.14}$$

for the local estimator and

$$\begin{aligned} S^{\rm enf} \equiv\ & 3S^{\rm local} \\ & +24\pi \int r^2 dr \int \frac{d^2\hat{n}}{w_3} \left[-B(\hat{n},r)C(\hat{n},r)D(\hat{n},r) + B(\hat{n},r)\langle C_{\rm sim}(\hat{n},r)D_{\rm sim}(\hat{n},r)\rangle_{\rm MC} \right. \\ & +C(\hat{n},r)\langle B_{\rm sim}(\hat{n},r)D_{\rm sim}(\hat{n},r)\rangle_{\rm MC} + D(\hat{n},r)\langle B_{\rm sim}(\hat{n},r)C_{\rm sim}(\hat{n},r)\rangle_{\rm MC}) \\ & \left. + \big(D^3(\hat{n},r) - 4D(\hat{n},r)\langle D^2_{\rm sim}(\hat{n},r)\rangle_{\rm MC}\big)\right] \end{aligned} \tag{B.15}$$

for the enfolded estimator. Here w_3 is sum of the weighting functions cubed

$$w_3 = \int d^2\hat{n} W^3(\hat{n}).$$

The cube is a result of the fact that one has 3 a_{lm}'s in each cubic product of the filtered maps in the skewness estimators. In real space a mask is simply a multiplication (i.e., one can multiply each $\Delta T(\hat{n})$ with either zero or one). However, this becomes a convolution in Fourier space. Consequently we have an integral over the solid angle $d^2\hat{n}$. If there is uniform weighting (that is, each pixel is masked or not) $W(\hat{n}) = M(\hat{n})$, the mask function and the integral becomes

$$w_3 = 4\pi f_{sky},$$

with f_{sky} the sky cut (in fact it is the opposite, it represents the percentage of sky that remains after masking). In the latest WMAP analysis they have not used a uniform weighting, but a "combination signal-plus-noise weight", which turns out to be optimal for the analysis of equilateral shaped bispectra, while the local shape is barely affected by simply using the uniform weight. It should be checked to what extent uniform weighting changes the estimates of the enfolded shape, compared to the more advanced combined weighting used by WMAP team (Komatsu et al. 2009).

Appendix C

Primordial Bispectrum Computation

C.1 Bispectrum from Single-Field Slow-Roll

In this appendix we will demonstrate how we derived the bispectra discussed in ch. 2 and 3. First, we derive the bispectrum for single-field slow-roll starting in non-BD state. We repeat this exercise in addition of a higher order derivative term and we end this appendix with the computation of the bispectrum for inflationary models with a non-canonical action.

We start with eq. (3.11) in Holman & Tolley (2008), from which we can extract the interaction Hamiltonian from the canonical effective action minimally coupled to gravity

$$H_I = -\frac{H}{M_p^2}\int d^3x a(\eta)^3 \left(\frac{\dot{\phi}}{H}\right)^4 \zeta'^2 \partial^{-2}\zeta'. \tag{C.1}$$

As carefully explained in Holman & Tolley (2008), the three-point correlation function $\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle$ can be written (to first order in the interaction Hamiltonian H_I) in an integral over the free field correlator $\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}H_I(\eta)\rangle$, where the free field correlator can be expanded in a product of two point correlators (i.e., Wightman functions) via Wick's theorem. Consequently it is straightforward to show that the three-point correlation function in the case we consider the interaction Hamiltonian of eq. (C.1) is given by

$$\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle = -i(2\pi)^3\delta^{(3)}\left(\sum\vec{k}_i\right)\frac{H}{M_p^2}\left(\frac{\dot{\phi}}{H}\right)^4 \times \tag{C.2}$$

$$\int_{\eta)}^{0} d\eta a^3(\eta)\frac{1}{k_3^2}\partial_\eta G^{>}_{k_1}(0,\eta)\partial_\eta G^{>}_{k_2}(0,\eta)\partial_\eta G^{>}_{k_3}(0,\eta) + \text{perm} + \text{c.c}, \tag{C.3}$$

as can be found in Holman & Tolley (2008). Here the Wightman functions $G^{>}_k$ are defined as

$$\langle\zeta_{k_1}\zeta_{k_2}\rangle = (2\pi)^3\delta^{(3)}(\vec{k}_1+\vec{k}_2)G^{>}_{k_1}(\eta,\eta'). \tag{C.4}$$

The Wightman functions can be found by solving the equations of motion of the inflaton field minimally coupled to gravity

$$G^{>}_k(\eta,\eta') = \frac{H^2}{\dot{\phi}^2}\frac{H^2}{2k^3}(1+ik\eta)(1-ik\eta')e^{-ik(\eta-\eta')}. \tag{C.5}$$

Consequently we compute

$$G_k^{>}(0,\eta) = \frac{H^2}{\dot{\phi}^2}\frac{H^2}{2k^3}(1-ik\eta)e^{ik\eta}, \tag{C.6}$$

and

$$\partial_\eta G_k^{>}(0,\eta) = -\frac{H^2}{\dot{\phi}^2}\frac{H}{2k}\frac{1}{a(\eta)}e^{ik\eta}, \tag{C.7}$$

with $a(\eta) = 1/\eta H$ during inflation in the assumption $\dot{H} \simeq 0$.

Next we can express $\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle$ in terms of these Wightman functions. However, we are interested in the case where we are not in the BD vacuum state. To first order in the Bogolyubov parameter β_k, what happens is that one of the solutions to the equation of motion 'picks up' a minus sign in k. It is easy to incorporate this minus sign to find

$$\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle^{\rm nBD} = -i(2\pi)^3\delta^{(3)}\left(\sum\vec{k}_i\right)\frac{1}{M_p^2}\frac{2}{\prod(2k_i^3)}\frac{H^6}{\dot{\phi}^2}\int_{\eta_0}^{0}d\eta\sum_j\beta_{k_j}^*\frac{3k_1^2k_2^2k_3^2}{k_j^2}e^{i\tilde{k}_j\eta} + \text{c.c.} \tag{C.8}$$

Note the factor of 6 is the result of the 6 possible permutations, while the sum is the result of implementing an initial-state modification for all of these 6 permutations. $\tilde{k}_j = k_t - 2k_j$, with $k_t = k_1 + k_2 + k_3$. The integral can be easily performed as well as adding the complex conjugate part. In steps

$$-i\times\int_{\eta_0}^{0}d\eta e^{i\tilde{k}_j\eta} + \text{c.c.} = \frac{2(\cos(\tilde{k}_j\eta_0)-1)}{\tilde{k}_j}$$

resulting in

$$\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle^{\rm nBD} = (2\pi)^3\delta^{(3)}\left(\sum\vec{k}_i\right)\frac{1}{M_p^2}\frac{4}{\prod(2k_i^3)}\frac{H^6}{\dot{\phi}^2}\sum_j\frac{3k_1^2k_2^2k_3^2}{k_j^2\tilde{k}_j}\mathcal{R}e(\beta_{k_j})\left(\cos(\tilde{k}_j\eta_0)-1\right). \tag{C.9}$$

This is the result we have used in eq. (2.12) of ch. 2. In case we assume that the enhancement occurs when $\tilde{k}_j \to 0$, we can apply this limit to the above expression to find

$$\begin{aligned}\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle^{\rm nBD} &= -(2\pi)^3\delta^{(3)}\left(\sum\vec{k}_i\right)\frac{1}{M_p^2}\frac{4}{\prod(2k_i^3)}\frac{H^6}{\dot{\phi}^2}\sum_j\frac{3k_1^2k_2^2k_3^2}{k_j^2}\mathcal{R}e(\beta_{k_j})\\ &\quad\times\frac{1}{2}\tilde{k}_j\eta_0^2 + \mathcal{O}(\tilde{k}_j^3),\end{aligned}$$

which is slightly different from the result found in Holman & Tolley (2008) (as it turned out they considered the imaginary part of β instead). Note that when $\tilde{k}_j = 0$ this whole expression actually vanishes, but it does have a maximum nearby (i.e., $k_{max} \sim \eta_0^{-1}$).

C.2 Bispectrum from Higher Order Derivative Term

Next we consider higher-order terms in the action of the form

$$\mathcal{L}_I = \sqrt{-g}\frac{\lambda}{8M^4}((\nabla\Phi)(\nabla\Phi))^2. \tag{C.10}$$

It is not hard to compute the corresponding interaction Hamiltonian up to third order in the curvature field ζ (and $\zeta \simeq -(H/\dot{\phi})\delta\phi$, where $\Phi = \phi(\eta) + \delta\phi(\eta, x)$). It can be shown that

$$H_I = -\frac{\lambda H}{2M^4}\int d^3x a(\eta)\left(\frac{\dot{\phi}}{H}\right)^3 \zeta'\left(\zeta'^2 - (\partial_i\zeta)^2\right). \tag{C.11}$$

In similar fashion we can write down the three-point correlator in terms of integrals over products of Wightman functions, while Fourier transforming to k space

$$\begin{aligned}\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle^{\rm HD} &= -i(2\pi)^3\delta^{(3)}\left(\sum\vec{k}_i\right)\frac{\lambda}{2HM^4}\left(\frac{\dot{\phi}}{H}\right)^3\int d\eta a(\eta)\\ \left(\partial_\eta G^>_{k_1}(0,\eta)\partial_\eta G^>_{k_2}(0,\eta)\partial_\eta G^>_{k_3}(0,\eta)\right. &+ \left.\vec{k}_1\cdot\vec{k}_2 G^>_{k_1}(0,\eta)G^>_{k_2}(0,\eta)\partial_\eta G^>_{k_3}(0,\eta) + \text{perm}\right),\end{aligned}$$

and its complex conjugate. Here the dot product comes from the Fourier transform of two partial spatial derivatives in the interaction Hamiltonian of eq. (C.11). This is a much 'larger' three-point correlator, so let us compute it in such a way that we do not lose track of all different components. The best approach is to first compute all the different terms in the integral and then group these in proportionality to η (i.e., $\propto \eta^0, \eta$ and η^2) . In addition we can compute the whole 'pre-factor' independently. It is not hard to show that it is given by

$$P = (2\pi)^3\delta^{(3)}\left(\sum\vec{k}_i\right)\frac{\lambda}{M^4}\frac{1}{\prod(2k_i^3)}\frac{H^8}{\dot{\phi}^2}, \tag{C.12}$$

and therefore

$$\begin{aligned}\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle^{\rm HD}_{\rm nBD} &= -iP\int_{\eta_0}^0 d\eta\sum_j\beta^*_{k_j}\left[3\eta^2k_1^2k_2^2k_3^2\right.\\ &+\vec{k}_j\cdot\vec{k}_{j+1}k_{j+2}^2(1+ik_j\eta)(1-ik_{j+1}\eta)\\ &+\vec{k}_j\cdot\vec{k}_{j+2}k_{j+1}^2(1+ik_j\eta)(1-ik_{j+2}\eta)\\ &+\left.\vec{k}_{j+1}\cdot\vec{k}_{j+2}k_j^2(1-ik_{j+1}\eta)(1-ik_{j+2}\eta)\right]e^{i\tilde{k}_j\eta},\end{aligned} \tag{C.13}$$

where the j's are cyclic in 1, 2 and 3.

Let us first rewrite the dot product using the triangle vector constraint $\vec{k}_1 + \vec{k}_2 + \vec{k}_3 = 0$. Using this constraint we can deduce that

$$\vec{k}_j\cdot\vec{k}_{j+1} = \frac{1}{2}(k_{j+2}^2 - k_j^2 - k_{j+1}^2).$$

If we write the integral as $\int_{\eta_0}^{0} d\eta \sum_j \beta^*_{k_j} S(k_1, k_2, k_3, \eta) e^{i\tilde{k}_j \eta}$ we obtain

$$\begin{aligned} S(k_1,k_2,k_3,\eta) &= +\frac{1}{2}\left[(k_{j+2}^2 - k_j^2 - k_{j+1}^2)k_{j+2}^2 + k_{j+1}^2 - k_j^2 - k_{j+2}^2)k_{j+1}^2\right. \\ &\quad \left.+(k_j^2 - k_{j+1}^2 - k_{j+2}^2)k_j^2\right] + \frac{i\eta}{2}\left[(k_{j+2}^2 - k_j^2 - k_{J+1}^2)k_{j+2}^2(k_j - k_{j+1})\right. \\ &\quad +(k_{j+1}^2 - k_j^2 - k_{j+2}^2)k_{j+1}^2(k_j - k_{j+2}) \\ &\quad \left.-(k_j^2 - k_{j+1}^2 - k_{J+2}^2)k_j^2(k_{j+1} + k_{j+2})\right] \\ &\quad +\frac{\eta^2}{2}\left[k_j k_{j+1} k_{j+2}\left((k_{j+2}(k_{j+2}^2 - k_j^2 - k_{j+1}^2) + k_{j+1}(k_{j+1}^2 - k_j^2 - k_{j+2}^2)\right.\right. \\ &\quad \left.\left.-k_j(k_j^2 - k_{j+1}^2 - k_{j+2}^2)\right) + 6k_1^2 k_2^2 k_3^2\right]. \end{aligned} \tag{C.14}$$

This can be rewritten as

$$\begin{aligned} S(k_1,k_2,k_3,\eta) &= -\frac{1}{2}k_t \prod_i \tilde{k}_i + \frac{i\eta}{2}\left[\tilde{k}_j\left(2k_{j+1}k_{j+2}(k_{j+1} + k_{j+2})^2\right.\right. \\ &\quad -k_t(\tilde{k}_j(k_{j+1}^2 + k_{j+2}^2) + (k_{j+1} + k_{j+2})(2k_{j+1}^2 + 3k_{j+1}k_{j+2} + 2k_{j+2}^2) \\ &\quad \left.\left.-\tilde{k}_j(k_{j+1} + k_{j+2})k_t\right)\right. \\ &\quad \left.-4k_{j+1}k_{j+2}(k_{j+1} + k_{j+2})(k_{j+1}^2 + k_{j+2}^2 + k_{j+1}k_{j+2})\right] \\ &\quad +\frac{\eta^2}{2}\left[\tilde{k}_j \prod_i k_i(k_t^2 - 4k_{j+1}k_{j+2}\right], \end{aligned} \tag{C.15}$$

which is very similar to eq. (3.30) in Holman & Tolley (2008) except for the term linear in η, for which we find additional terms. Since the integral runs over conformal time η, all that is left is to compute the different integrals in terms of η. Again, it is useful to pre-compute the following integrals

$$\begin{aligned} -i \times \int_{\eta_0}^{0} d\eta e^{i\tilde{k}_j\eta} + \text{c.c.} &= \frac{2(\cos(\tilde{k}_j\eta_0) - 1)}{\tilde{k}_j} \\ -i \times \int_{\eta_0}^{0} d\eta (i\eta) e^{i\tilde{k}_j\eta} + \text{c.c.} &= \frac{-2\eta_0 \sin(\tilde{k}_j\eta_0)}{\tilde{k}_j} + \frac{2(1 - \cos(\tilde{k}_j\eta_0))}{\tilde{k}_j^2} \\ -i \times \int_{\eta_0}^{0} d\eta \eta^2 e^{i\tilde{k}_j\eta} + \text{c.c.} &= \frac{2\eta_0^2 \cos(\tilde{k}_j\eta_0)}{\tilde{k}_j} - \frac{4\eta_0 \sin(\tilde{k}_j\eta_0)}{\tilde{k}_j^2} + \frac{4(1 - \cos(\tilde{k}_j\eta_0))}{\tilde{k}_j^3}. \end{aligned}$$

Since we now have terms that are proportional to $\tilde{k}_j^{-3}$ one would expect some terms to diverge in case $\tilde{k}_j \to 0$. However, this proportionality only appears in the last term of eq. (C.15), which is proportional to $\tilde{k}_j$ in itself. Consequently we lose a factor of $\tilde{k}_j$ after multiplication, just enough to make that term finite for $\tilde{k}_j \to 0$. Similar argumentation can be applied to the other terms, once we realize that the limits of $\sin(x)/x$ and $(1 - \cos(x))/x^2$ are finite in the

limit $x \to 0$. The three-point correlator therefore becomes

$$\begin{aligned}\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle^{\rm HD}_{\rm nBD} &= (2\pi)^3\delta^{(3)}\left(\sum\vec{k}_i\right)\frac{\lambda}{M^4}\frac{1}{\prod(2k_i^3)}\frac{H^8}{\dot\phi^2}\sum_j 2\mathcal{R}e(\beta_{k_j})\\ &\times\left[\frac{(1-\cos(\tilde{k}_j\eta_0))}{\tilde{k}_j}\left(-k_t\prod_i\tilde{k}_i+\mathcal{A}_1(k)+\frac{1}{\tilde{k}_j}\left(\mathcal{A}_2(k)+2\mathcal{B}(k)\right)\right)\right.\\ &\left.-\frac{\eta_0\sin(\tilde{k}_j\eta_0)}{\tilde{k}_j}\left(\tilde{k}_j\mathcal{A}_1+\mathcal{A}_2(k)+2\mathcal{B}(k)\right)+\eta_0^2\cos(\tilde{k}_j\eta_0)\mathcal{B}(k)\right], \quad \text{(C.16)}\end{aligned}$$

where

$$\begin{aligned}\mathcal{A}_1(k_j,k_{j+1},k_{j+2}) &= 2k_{j+1}k_{j+2}(k_{j+1}+k_{j+2})^2-k_t\left[\tilde{k}_j(k_{j+1}^2+k_{j+2}^2)\right.\\ &\left.+(k_{j+1}+k_{j+2})(2k_{j+1}^2+3k_{j+1}k_{j+2}+2k_{j+2}^2)\right]-\tilde{k}_j(k_{j+1}+k_{j+2})k_t^2\\ \mathcal{A}_2(k_j,k_{j+1},k_{j+2}) &= -4k_{j+1}k_{j+2}(k_{j+1}+k_{j+2})(k_{j+1}^2+k_{j+2}^2+k_{j+1}k_{j+2})\\ \mathcal{B}(k_j,k_{j+1},k_{j+2}) &= \prod_i k_i(k_t^2-4k_{j+1}k_{j+2}).\end{aligned}$$

Unlike in the simple case, and aside the comoving dependence of the denominator, this three-point correlator has terms that are proportional to $(1-\cos(\tilde{k}_j\eta_0))/\tilde{k}_j^2$ and $\sin(\tilde{k}_j\eta_0)/\tilde{k}_j$. Therefore we can take the exact limit $\tilde{k}_j = 0$, and have a non-vanishing result. The easiest way to see what happens is to look at the integrals. In the limit $\tilde{k}_j = 0$ the first integral and the last integral vanish. Consequently, we can write

$$\begin{aligned}\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle^{\rm HD}_{\tilde{\rm k}_{\rm j}=0} &= -(2\pi)^3\delta^{(3)}\left(\sum\vec{k}_i\right)\frac{\lambda}{M^4}\frac{1}{\prod(2k_i^3)}\frac{H^8}{\dot\phi^2}\sum_j 2\mathcal{R}e(\beta_{k_j})\\ &\times\frac{\eta_0^2}{2}\mathcal{A}_2(k_j,k_{j+1},k_{j+2})\\ &= -(2\pi)^3\delta^{(3)}\left(\sum\vec{k}_i\right)\frac{\lambda}{M^4}\frac{1}{\prod(2k_i^3)}\frac{H^8}{\dot\phi^2}\sum_j 2\mathcal{R}e(\beta_{k_j})\\ &\times\frac{\eta_0^2}{2}\left(-4k_{j+1}k_{j+2}(k_{j+1}+k_{j+2})(k_{j+1}^2+k_{j+2}^2+k_{j+1}k_{j+2})\right),\end{aligned}$$

which is equivalent, up to a minus sign, to what Holman & Tolley (2008) found.

C.3 Bispectrum from a Non-Canonical Action

In this section we will compute the full bispectrum coming from all terms in the interaction Hamiltonian and for general but k-independent β with an arbitrary phase, i.e., $\beta = |\beta|\exp(i\delta)$.

We have already computed the leading-order contribution. In the collinear limit $\tilde{k}_j \to 0$ this

reduces to

$$\lim_{\tilde{k}_j\to 0}\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle_{\rm nBD(1)} = \frac{1}{24}(2\pi)^3\frac{H^4|\beta|}{\epsilon^2 c_s^4}\left(c_s^2-1+\frac{2\lambda c_s^2}{\Sigma}\right)\delta^{(3)}(\sum\vec{k}_i)\times (\eta_0 c_s)^3\left[\frac{1}{k_1k_2(k_1+k_2)}+\frac{1}{k_2k_3(k_2+k_3)}+\frac{1}{k_1k_3(k_1+k_3)}\right]\sin\delta. \tag{C.17}$$

Performing a similar analysis as done for the first term on the remaining terms in the interaction Hamiltonian yields for the second term:

$$\begin{aligned}\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle_{\rm nBD(2)} &= -\frac{1}{16}(2\pi)^3\frac{H^4|\beta|}{\epsilon^2c_s^4}(\epsilon-3+3c_s^2)\delta(\sum\vec{k}_i)\times\\ &\frac{1}{k_1k_2k_3}\sum_j\left[\left(\frac{\tilde{k}_j-k_j}{k_j^2}+\frac{\tilde{k}_j+k_{j+1}}{k_{j+1}^2}+\frac{\tilde{k}_j+k_{j+2}}{k_{j+2}^2}\right)\left(\frac{\cos\delta-\cos(\tilde{k}_jc_s\eta_0+\delta)}{\tilde{k}_j^2}\right)\right.\\ &\left.+\left(\frac{1}{k_j}-\frac{1}{k_{j+1}}-\frac{1}{k_{j+2}}\right)\frac{\eta_0c_s\sin(\tilde{k}_jc_s\eta_0+\delta)}{\tilde{k}_j}\right].\end{aligned}$$

In the colinear limit $\tilde{k}_j\to 0$ we obtain

$$\begin{aligned}\lim_{\tilde{k}_j\to 0}\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle_{\rm nBD(2)} &= \frac{1}{32}(2\pi)^3\frac{H^4|\beta|}{\epsilon^2c_s^4}(\epsilon-3+3c_s^2)\delta(\sum\vec{k}_i)\\ &\eta_0c_s\frac{(k_1^2+k_1k_2+k_2^2)\left((k_1k_2(k_1+k_2)\eta_0c_s\cos\delta-2(k_1^2+k_1k_2+k_2^2)\sin\delta\right)}{k_1^3k_2^3(k_1+k_2)},\end{aligned}$$

plus cyclic permutations. There are two things we would like to mention. First of all, it is clear that the subleading term in this limit has two contributions that are out of phase. Secondly, the term proportional to $\sin\delta$ is only linearly enhanced. Therefore, for large values of the cutoff this term is negligible. Consequently, the leading and subleading terms are out of phase in their enhancement. Although this observation seems preliminary given that we are considering the collinear limit, it turns out as shown in sec. 5 of ch. 3, that this is true for the full domain.

The third term yields

$$\begin{aligned}\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle_{\rm nBD(3)} &= \frac{1}{16}(2\pi)^3\frac{\epsilon}{c_s^2}(\epsilon-2s+1-c_s^2)\frac{H^{10}}{\dot{\phi}_0^6}\frac{1}{c_s^3}\delta(\sum\vec{k}_i)\frac{(k_1^2+k_2^2+k_3^2)}{k_1^3k_2^3k_3^3}\times\\ &\mathcal{R}e\left[i\beta\sum_j\int_{\eta_0}^0 d\eta\left[\frac{1}{\eta^2}-i\tilde{k}_jc_s\frac{1}{\eta}+(k_jk_{j+1}+k_jk_{j+2}-k_{j+1}k_{j+2})c_s^2\right.\right.\\ &\left.\left.-ik_1k_2k_3c_s^3\eta\right]\exp(i\tilde{k}_jc_s\eta)\right].\end{aligned} \tag{C.18}$$

This can be integrated to find

$$\begin{aligned}\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle_{\rm nBD(3a)} &= -\frac{1}{32}(2\pi)^3\frac{H^4|\beta|}{\epsilon^2 c_s^4}(\epsilon-2s+1-c_s^2)\delta(\sum\vec{k}_i)\frac{1}{k_1^3k_2^3k_3^3}\times \\ &\sum_j\left(k_1^2+k_2^2+k_3^2\right)\Bigg[k_1k_2k_3\left(-\frac{\eta_0 c_s\sin(\tilde{k}_j c_s\eta_0+\delta)}{\tilde{k}_j}\right)+ \\ &\tilde{k}_j\left(\lim_{\eta\to 0}\frac{\sin(\tilde{k}_j c_s\eta+\delta)}{c_s\eta}-\frac{\sin(\tilde{k}_j c_s\eta_0+\delta)}{c_s\eta_0}\right)+ \\ &\Big(k_1k_2k_3+\tilde{k}_j(k_jk_{j+1}+k_jk_{j+2}-k_{j+1}k_{j+2})\Big)\frac{\cos\delta-\cos(\tilde{k}_j c_s\eta_0+\delta)}{\tilde{k}_j^2}\Bigg].\end{aligned}$$

The contribution from the third line would be infinite in the limit $\eta\to 0$ and therefore would produce a divergent result. However, the limit of the conformal time to zero is not an actual limit since this would suggest we would have infinite time to observe all scales. This limit is usually taken for convenience since it turns out that in most cases it does not add power to the 3-point function. In order to constrain this term one should take the limit $\eta\to\eta^*$ where η^* is the conformal time at CMB formation. In this way, this term is finite and negligible compared to all other contributions. We can investigate the behavior of this bispectrum in the limit $\tilde{k}_j\to 0$ to obtain

$$\begin{aligned}\lim_{\tilde{k}_j\to 0}\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle_{\rm nBD(3)} &= \frac{1}{32}(2\pi)^3\frac{H^4|\beta|}{\epsilon^2 c_s^4}(\epsilon-2s+1-c_s^2)\delta(\sum\vec{k}_i) \\ &\eta_0 c_s\frac{(k_1^2+k_1k_2+k_2^2)\left((k_1k_2(k_1+k_2)\eta_0 c_s\cos\delta-2(k_1^2+k_1k_2+k_2^2)\sin\delta\right)}{k_1^3k_2^3(k_1+k_2)},\end{aligned}$$

plus cyclic permutations, making this limit exactly similar and opposite in sign to the previous collinear limit. In effect it implies that for all values of the phase δ the collinear limit is zero for DBI models of inflation.

For the last term in the interaction Hamiltonian we obtain

$$\begin{aligned}\langle\zeta_{k_1}\zeta_{k_2}\zeta_{k_3}\rangle_{\rm nBD(4)} &= -\frac{1}{16}(2\pi)^3\frac{H^4}{\epsilon c_s^4}\delta(\sum\vec{k}_i)\frac{|\beta|}{k_1^3k_2^3k_3^3}\times \\ &\Bigg[\sum_j\frac{\cos\delta-\cos(\tilde{k}_j c_s\eta_0+\delta)}{\tilde{k}_j}\left(\frac{1}{2}(k_1^4+k_2^4+k_3^4)-(k_1^2k_2^2+k_2^2k_3^2+k_3^2k_1^2)\right) \\ &+\sum_j\left(-\frac{c_s\eta_0\sin(\tilde{k}_j c_s\eta_0+\delta)}{\tilde{k}_j}+\frac{\cos\delta-\cos(\tilde{k}_j c_s\eta_0+\delta)}{\tilde{k}_j^2}\right)\times\big(k_j^4(k_{j+1}+k_{j+2}) \\ &+k_{j+1}^4(k_{j+2}-k_j)+k_{j+2}^4(k_{j+1}-k_j)+k_j^2k_{j+2}^2(k_j-2k_{j+1}-k_{j+2}) \\ &+k_{j+1}^2k_{j+2}^2(2k_j-k_{j+1}-k_{j+2})+k_j^2k_{j+1}^2(k_j-k_{j+1}-2k_{j+2})\big)\Bigg].\end{aligned} \tag{C.19}$$

Appendix D

Template Normalization

The projection between two shapes F_X and F_Y is determined via the dot product

$$F_X \cdot F_Y \equiv \int_{\Delta_k} dx_2 dx_3 F_X(x_2, x_3) F_Y(x_2, x_3) x_2^4 x_3^4, \tag{D.1}$$

first introduced by Babich et al. (2004a). This was recently revised and replaced with the improved dot product (eq. (5.16)) by Fergusson & Shellard (2009). Although this revised 'correlator' improves the results compared to the observed 2-dimensional projected dot products, in the case of scale-invariance the differences are marginal. The limits of integration are given by the triangle domain, i.e., $\sum \vec{k}_i = 0$, which corresponds to $0 \leq x_2 \leq 1$ and $1 - x_2 \leq x_3 \leq 1$. One can then define a cosine between two different shapes to 'measure' how similar two shapes with different comoving momenta are

$$\cos(F_X, F_Y) \equiv \frac{F_X \cdot F_Y}{(F_X \cdot F_X)^{1/2} (F_Y \cdot F_Y)^{1/2}}. \tag{D.2}$$

In previous work we calculated the normalization factors $F_{\rm local} \cdot F_{\rm local} = 176.5$ and $F_{\rm eq} \cdot F_{\rm eq} = 7.9$. The normalization of the newly proposed orthogonal template $F_{\rm ort} \cdot F_{\rm ort} = 13.8$. Note that the normalization of the local template is actually infinite if one integrates over the full triangle domain. This is because most of the signal comes from scales $k \to 0$, which corresponds to perturbations with an infinite physical length scale. These are not observable in the CMB and therefore one truncates the limit of integration such that only observable scales are included, which we take to be $x_2 > 0.001$, as this is the ratio between the smallest and largest observable scales.

Samenvatting

Inleiding

We leven in een immens heelal, gevuld met gas, sterrenstelsels, sterren en planeten. Toch kan de grootschalige evolutie van het heelal beschreven worden met verassend weinig parameters. Het meest succesvolle model van het heelal tot op heden telt *slechts* 6 parameters. Dit is de conclusie die de wetenschap heeft opgemaakt bijna 100 jaar na de introductie van Einstein's zwaartekrachtswetten. Deze wetten maakten het voor het eerst mogelijk een dynamisch heelal te beschrijven en uit te rekenen. Ruwweg kunnen deze 6 parameters als volgt worden opgedeeld:

- Interacties in het heelal (1)
- Samenstelling van het heelal (3)
- Begincondities (2)

Het onderzoek in dit proefschrift richt zich op de begincondities van het heelal. De moderne kosmologie verteld ons dat het 13.7 miljard jaar geleden ontstond uit de oerknal. Als we in dit proefschrift spreken over begincondities, bedoelen we met 'begin' niet het moment dat de oerknal plaats vond. In deze samenvatting zal ik uitleggen dat de begincondities die we kunnen meten en waarmee we mogelijk onbekende fysica kunnen onderzoeken, zijn ontstaan in een periode *vlak na de oerknal*. In die periode was het heelal zeer heet, dicht en energetisch. Het interessante aan deze periode is dat de natuurkunde die een rol moet hebben gespeeld bij deze extreme omstandigheden, slecht bekend is. Er zijn wel ideeën hoe de natuur zich gedraagt onder deze buitengewone omstandigheden, maar deze ideeën zijn uiterst moeizaam te toetsen omdat we niet in staat zijn deze extreme omstandigheden na te bootsen in een lab op aarde. Aan de andere kant *biedt* de oerknal ook een mogelijkheid om de natuur onder deze omstandigheden te bestuderen. Dit is ook meteen de motivatie van het onderzoek in dit proefschrift. De uitdaging ligt in het achterhalen van de juiste begincondities en die te relateren aan de fundamentele wetten in de natuur. Kortweg hebben we in dit proefschrift twee vragen proberen te beantwoorden: Gegeven een theorie die de meest extreme toestand in het heelal omschrijft, wat is de voorspelling voor de begintoestand? en Hoe achterhaal ik de werkelijke begintoestand met behulp van een kosmologische waarneming? Zodra beide vragen zijn beantwoord ontstaat de mogelijkheid om te achterhalen wat de juiste beschrijving is van

de natuur bij de meest extreme omstandigheden in het vroege heelal. Voor dat ik kan uitleggen hoe ik heb geprobeerd deze vragen te beantwoorden, zal ik eerst een korte samenvatting geven over de onstaansgeschiedenis van het heelal.

Einstein's Zwaartekracht

Met de introductie van de moderne zwaartekracht theorieën van Einstein aan het begin van de vorige eeuw (1905-1916), werd het mogelijk om de evolutie van het heelal te beschrijven. De essentiële postulaten van deze theorie vertellen ons dat ruimte en tijd aan elkaar gerelateerd zijn (ruimtetijd), dat de vorm van ruimtetijd (de geometrie) is gerelateerd aan de inhoud (de energie) en dat er een eindige snelheid bestaat waarmee informatie zich kan verplaatsen. Met deze postulaten is het mogelijk om ons heelal te beschrijven en voorspellingen te doen over hoe het heelal er vroeger uit heeft gezien, en mogelijk over hoe het heelal er in de toekomst uit gaat zien.

De rede dat we spreken over ruimtetijd wordt gemotiveerd door Einstein's bevinding dat ruimte en tijd niet *invariant* (=onveranderd) zijn onder de bewegingstoestand van een waarnemer. Dat wil zeggen dat zowel afstanden als tijdsintervallen afhankelijk zijn van de waarnemer. Een waarnemer in rust zal bijvoorbeeld een andere tijdswaarneming hebben dan een waarnemer die met een constante snelheid beweegt. De verschillen tussen de metingen nemen toe naarmate de relatieve snelheid tussen de waarnemers groter wordt. De maximale snelheid is de licht snelheid en zodra die bereikt wordt, zijn de verschillen tussen de metingen van beide waarnemers maximaal. Als je natuurwetten formuleert zou je graag willen dat deze natuurwetten waarnemer onafhankelijk zijn. Einstein liet zien dat hoewel ruimte en tijd allebei niet invariant zijn onder de bewegingstoestand van de waarnemer, de combinatie ruimtetijd dat wel is; de verschillen tussen de aftanden en tijdsintervallen die beide waarnemers meten, heffen elkaar als het ware op. De bepaling van de invariante ruimtetijd kunnen we doen met behulp van een zogenaamde metriek; de metriek is eigenlijk een soort voorschrift voor afstanden in ruimtetijd en hangt af van de vorm van ruimtetijd. Om de Einstein's wetten te gebruiken in de kosmologie zullen we een metriek moeten vinden die we kunnen gebruiken voor ruimtetijd in het Universum. De metriek die hoort bij ons Universum, de zogenaamde FRLW metriek (vernoemd naar Friedmann, Robertson, Lemaître en Walker), is afgeleid van de waarneming dat het heelal op grote schaal isotroop en homogeen is. Dat wil zeggen, het heelal is vanuit elk punt in de ruimte hetzelfde in elke richting. Daarnaast hebben we kunnen afleiden uit de waarnemingen dat het heelal vlak is. Net zoals een oppervlak gekromd kan zijn, bijvoorbeeld de buitenkant van een voetbal, kan een hoger dimensionale ruimte ook gekromd zijn. In dit geval gaat het om de 3 ruimtelijke dimensies waarin wij ons voort bewegen. Het is wellicht moeilijk voor te stellen hoe een 3 dimensionale gekromde ruimte eruit ziet, toch kunnen we metingen verrichten die aangeven of onze ruimte gekromd is. Zo blijkt dat de som van de hoeken in een gelijkzijdige driehoek in een gekromde ruimte verschilt van die in een vlakke ruimte. Deze verschillen stellen ons in staat af te leiden of het heelal vlak is of niet, en uit de recente metingen blijkt dat het heelal niet gekromd is.

Het tweede postulaat vertelt ons hoe de geometrie, opgelegd door de metriek, afhangt van de inhoud van de ruimtetijd, de massa of energie (massa is ook gerelateerd aan energie via de

welbekende relatie $E = mc^2$). Dus, als we iets in het heelal stoppen, bijvoorbeeld 1 deeltje, hoe verandert dan het ruimtetijd interval tussen twee objecten als functie van plaats en tijd. De theorie van Einstein maakte het mogelijk om te onderzoeken hoe het heelal zich ontwikkelt in de tijd, gegeven de metriek en de inhoud. Stel je nu eens voor dat je leefde in de tijd van Einstein. Je keek omhoog naar de sterrenhemel en je probeerde te begrijpen hoe het heelal mogelijk zou kunnen veranderen in de tijd. De meest voor de hand liggende waarneming die je zou maken is dat het heelal helemaal niet verandert. Dit is dezelfde conclusie die Einstein trok (en vele voor hem) en hij probeerde een oplossing te vinden van zijn eigen theorie die dit als antwoord gaf. Tot zijn eigen verbazing vond hij deze niet; alle oplossingen van zijn vergelijkingen leidde tot een dynamisch heelal. Hij heeft geprobeerd om dit kunstmatig op te lossen, maar later werd bewezen door de Nederlandse natuurkundige Willem de Sitter dat ook dit niet werkte; het heelal kon niet anders dan evolueren. Je kunt je wel voorstellen dat dit een enorme ontdekking was. Zijn bevinding impliceerden dat het heelal er hoe dan ook in het verleden anders uit moet hebben gezien[1].

De eerste aanwijzing dat het heelal daadwerkelijk evolueerde werd gevonden door Edwin Hubble in 1929. Hij ontdekte dat verre sterrenstelsels van ons vandaan bewogen. Bovendien bepaalde hij dat deze snelheid toenaam naarmate de sterrenstelsels verder van ons vandaan stonden. Dit zou betekenen dat het heelal aan het uitdijen was. Uiteraard was deze waarneming in eerste instantie behoorlijk controversieel, maar het weerhield verschillende mensen er niet van om te filosoferen. Zo werd beredeneerd dat als het heelal momenteel aan het uitdijen was, het kleiner en compacter moet zijn geweest in het verleden. Dit leidde uiteindelijk tot het idee van een oerknal.

Microgolfachtergrondstraling

In 1948 kwamen George Gamow en 2 van zijn collega's tot de conclusie dat een oerknal een handig neveneffect zou kunnen hebben; hun berekeningen toonden aan dat vlak na de oerknal de juiste condities bestonden om precies de juiste hoeveelheden waterstof en helium te produceren die in het huidige heelal werden gemeten. Daarnaast concludeerden zij dat als het heelal was ontstaan uit een oerknal er straling moet zijn geproduceerd in het vroege heelal die vandaag de dag waarneembaar zou moeten zijn. Deze straling werd voor het eerst waargenomen in 1964 en is tot op heden *het beste bewijs voor het hete-Oerknalmodel*. Deze straling staat bekend als de microgolfachtergrondstraling en is in feite een directe foto van het vroege heelal. Door deze foto nauwkeurig te bestuderen kunnen we iets leren over de begincondities.

Hoe ontstaat de microgolfachtergrondstraling? Vlak na het ontstaan van het heelal is het heelal heel heet, dicht en energetisch. Dat betekent dat alle deeltjes in het heelal heel dicht op elkaar zitten en dat ze snel bewegen en vaak met elkaar botsen (verstrooiing). Het vroege heelal wordt in de eerste paar honderdduizend jaar gedomineerd door straling. Straling bestaat uit lichtdeeltjes, zogenaamde fotonen. Daarnaast is er ook gewone materie, die voor het grootste

[1]Heel eerlijk moet gezegd worden dat de eerste twee postulaten aan elkaar gerelateerd zijn. Immers, de metriek die ons het invariante ruimtetijd interval voorschrijft hangt af van de samenstelling van de ruimte die we beschouwen. Een FRWL metriek is afgeleid in de veronderstelling dat we kennis hebben van wat er zich in het heelal bevind (energie) en hoe het is verdeeld (homogeen en isotroop).

deel bestaat uit waterstof (75%) en Helium (25%). Een waterstofatoom bestaat uit één proton (positieve lading) en één elektron (negatieve lading). De hoge temperatuur in het vroege heelal leidt tot fotonen met een relatief hoge energie. Deze fotonen zijn in staat om de elektronen uit het waterstofatoom te stoten. Het overgebleven proton en het vrije elektron vormen geïoniseerd waterstof. Zodoende heb je dus een hete soep van fotonen en geïnoniseerde waterstof. De geïoniseerde waterstofatomen hebben een grotere kans op interactie met de fotonen dan neutraal waterstof. Het gevolg is dat een foton niet ver kan bewegen voordat het een interactie aangaat met materie; licht kan zich niet vrij bewegen. In het hete-Oerknalmodel dijt het heelal uit met de tijd. Daardoor nemen zowel de dichtheid als de effectieve temperatuur af. Langzaam zijn er steeds minder fotonen beschikbaar met een energie die nodig is om waterstof te ioniseren. De protonen en elektronen kunnen zich samenvoegen tot een neutraal waterstof atoom., een proces dat recombinatie wordt genoemd. Een neutraal waterstof heeft veel minder interactie met de lichtdeeltjes, dan geïoniseerde waterstof. Daardoor is er veel minder verstrooiing van de fotonen en kan een foton nu een veel langere weg afleggen. Recombinatie gebeurt in een relatief korte periode en de afstand die fotonen kunnen afleggen zonder te worden verstrooid is veel groter dan de afstand die ze kunnen afleggen binnen de huidige leeftijd van het heelal (13.7 miljard jaar). Het gevolg is dat vlak na recombinatie vrijwel alle fotonen plots vrij kunnen bewegen: dit is het moment waarop de achtergrondstraling wordt gevormd. Op dat moment is de temperatuur van de oersoep enkele duizenden graden. Omdat het heelal langzaam verder uitdijt neemt de temperatuur van de vrijgekomen straling af. De gemiddelde energie (temperatuur) van de fotonen komt overeen met een bepaalde frequentie. Vandaag de dag is de gemiddelde frequentie van de achtergrond straling rond een paar honderd Giga Hertz, wat overeenkomt met millimeter golflengtes. Vandaar de naam microgolfachtergrondstraling. Omdat deze straling overal in het heelal min of meer tegelijk ontstaat komt deze straling van alle kanten in het heelal. Het maakt dus niet uit in welke richting je je detector zet, je zult altijd fotonen van deze achtergrondstraling opvangen. De microgolfachtergrondstraling is de grootste bron van straling in het heelal met een totale stralingsenergie groter dat alle sterren in het heelal samen. Als je je analoge TV aanzet en je haalt de kabel uit de ontvanger, dan zie je ruis (sneeuw). Dit is voornamelijk thermische ruis afkomstig van de TV zelf. Maar een klein deel, ongeveer 1 procent, is het gevolg van de microgolfachtergrondstraling.

Inflatie

De eerste waarneming van de microgolfachtergrondstraling (Cosmic Microwave Background of CMB) in 1964 door Penzias en Wilson (nobelprijs 1978) werd al gauw gevolgd door vele nieuwe metingen, waaronder metingen met behulp van satellieten die in staat zijn om de gehele hemel in kaart te brengen. Een kaart van de CMB is te zien in fig. D.1. Een van de ontdekkingen door studie van de CMB is dat de temperatuur van de achtergrondstraling enorm homogeen is. Dat wil zeggen, de afwijking van de gemiddelde temperatuur van 2.7 Kelvin (-270.3 graden Celsius) zijn klein. Dit betekent dat er thermisch evenwicht heerste ten tijde van de creatie van de microgolfachtergrondstraling. Thermisch evenwicht ontstaat doordat de de deeltjes in een gesloten systeem geleidelijk de energie aan elkaar zullen overdragen, tot dat de gemiddelde energie per deeltje ongeveer gelijk is. Dit gebeurt bijvoorbeeld als je koude melk bij een kop

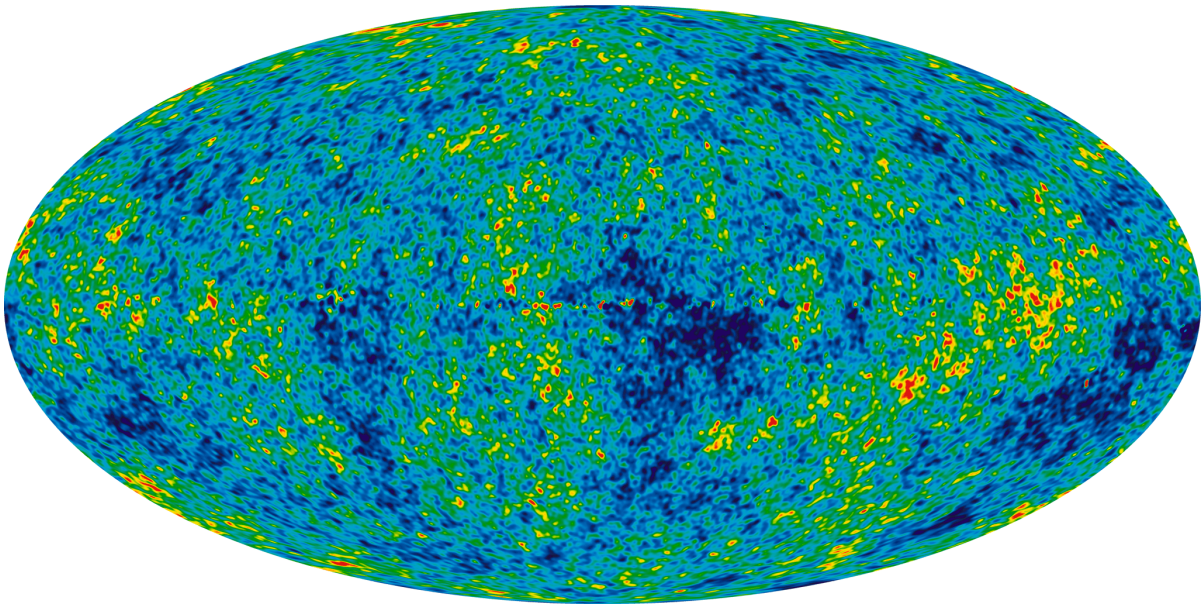

Figure D.1: Een opname van de microgolf achtergrondstraling gemaakt met behulp van de Wilkinson Microwave Anisotropy Probe (WMAP). De gemiddelde temperatuur van de straling is ongeveer 2.7 Kelvin. Daarboven op zijn kleine variaties (enkele micro Kelvin) die het gevolg zijn processen in het zeer jonge heelal. Het statistische verdeling van deze fluctuaties geeft inzicht in de natuurkunde vlak na de oerknal. Voor een kleurenversie zie fig. 1.4 in hfst. 1.

warme thee gooit. De warme water deeltjes van de thee dragen hun energie geleidelijk over aan de koude deeltjes in de melk en de uiteindelijke temperatuur van de kop thee is lager dan de temperatuur van het theewater en hoger dan die van de melk. Als je lang genoeg wacht dan komt de thee met melk uiteindelijk ook in thermisch evenwicht met de omgeving. De temperatuur zakt dan tot de temperatuur van de lucht, bijvoorbeeld de 20 graden Celsius in de huiskamer. Belangrijk is hierbij dat dit niet meteen gebeurt, maar een tijdje duurt omdat de deeltjes nu eenmaal dicht genoeg bij elkaar in de buurt moeten komen om energie te kunnen overdragen. De snelst mogelijke overdacht van deze 'informatie' over de energie van individuele deeltjes is de lichtsnelheid. Als twee punten in het heelal gedurende de leeftijd van het heelal kunnen communiceren dat zijn deze twee punten in *causaal* contact: veranderingen in de toestand van deze twee punten kunnen elkaar in principe beïnvloeden door informatie uit te wisselen met de lichtsnelheid. De CMB bestaat uit fotonen en al deze fotonen blijken dus in thermisch evenwicht te zijn. Dit betekent dat fotonen aan bijvoorbeeld de 'noordpool' in de CMB (top van fig. 1.4) op een moment in het verleden in contact moeten zijn geweest met fotonen aan de 'zuidpool'. Deze twee punten zouden ooit causaal contact moeten hebben gehad. Dit is in principe niet verontrustend, ware het niet dat de gehele hemel een enorme afstand overbrugt. De totale afstand van noord naar is 2 maal de leeftijd van het heelal in licht jaar, dus ongeveer 28 miljard licht jaar. Dit zou betekenen dat fotonen van de noord- en zuidpool nooit met elkaar in contact kunnen zijn geweest aangezien het heelal 'slechts' 13.7 miljard jaar oud is. Deze punten zijn dus niet causaal verbonden. Sterker nog, de maximale

afstand die fotonen hebben kunnen overbruggen is bepaald door de leeftijd van het heelal op het moment dat de achtergrondstraling *werd gevormd*. Dit was ongeveer 380000 jaar na de oerknal. Door dit te vermenigvuldigen met de lichtsnelheid vind je een maximale afstand die fotonen kunnen hebben afgelegd. Deze afstand kun je projecteren op de CMB map en je vindt dat dit overeenkomt met een afstand aan de huidige hemel van 1 graad. Dus fotonen die de achtergrondstraling zouden feitelijk alleen causaal verbonden kunnen zijn over de afstand van 1 graad aan de huidige hemel. We zouden op basis van deze redenering verwachten dat thermisch evenwicht alleen kan zijn ontstaan over dezelfde afstand. Deze ogenschijnlijke paradox staat bekend als het horizonprobleem.

Naast het horizonprobleem is het ook opvallend dat het heelal niet gekromd is. Uit Einstein's vergelijkingen volgt dat dit een onstabiele oplossing is: als het heelal vlak na de oerknal maar een klein beetje gekromd is, wordt die kromming alleen maar groter gedurende de evolutie van het heelal. Dit is het zogenaamde vlakheidsprobleem. In feite zijn het horizon en vlakheidsprobleem voorbeelden van 'fine-tunning' problemen; alleen als de begincondities zeer precies afgesteld zijn (fine-tunen), bijvoorbeeld door alle fotonen over enorme afstandsschalen dezelfde initiële temperatuur te geven en daarnaast een heelal creëren dat precies vlak is, kunnen we de huidige waarnemingen verklaren. Dit wordt vaak als problematisch erevaren, omdat we het liefst onze waarnemingen zouden verklaren als een onvermijdelijk gevolg van een fysisch proces.

In 1981 kwam Alan Guth met een oplossing. Hij had berekend dat als er vlak na de oerknal een periode is geweest waarin het heelal zeer snel groeide, het horizon en vlakheidsprobleem konden worden opgelost. Het heelal zou in die periode minimaal 10^{28} (een 1 gevolgd door 28 nullen) keer zo groot zijn geworden in slechts 10^{-35} seconde! De snelheid waarmee het heelal zou moeten groeien ligt hoger dan lichtsnelheid. Echter, er wordt hier geen informatie overgedragen en dus is het geen overtreding van Einstein's postulaat. Deze *exponentiele* expansie noemde Guth kosmologische inflatie. Hij liet zien dat inflatie kan optreden zodra de potentiële energie van een component het heelal domineert. Dat is nodig omdat versnelde expansie alleen optreed zodra de druk *negatief* is. Guth beargumenteerde dat inflatie kon ontstaan door een *veld*. Een veld heeft de eigenschap dat het een waarde heeft op elk punt in de ruimte (in tegenstelling tot een deeltje). Straling die de CMB vormt is een voorbeeld van een veld. Inflatie blijkt niet te realiseren met een stralingsveld en is hoogstwaarschijnlijk het gevolg van een nog onbekend scalair veld, het zogenaamde inflaton veld. De potentiële energie van het inflaton veld kan hoger zijn dan de bewegingsenergie, precies de eigenschap die nodig voor negatieve druk die lijdt tot exponentiële expansie. Het veld dat verantwoordelijk voor inflatie is niet bekent en binnen de huidige velden theorie bestaat er geen scalair veld dat inflatie zou kunnen veroorzaken. Dit betekent dat als we inflatie kunnen onderzoeken we in feite direct bezig zijn met het onderzoeken van onbekende fysica.

Inflatie voorkomt het 'fine-tunen' van de begincondities. Het vlakheidsprobleem wordt opgelost doordat elk mogelijke afwijking van vlakheid wordt gladgestreken door de enorme expansie. Een veelgebruikt voorbeeld is het oppervlak van een ballon. Een bijna lege ballon heeft duidelijk een gekromd oppervlak. Zodra je meer lucht in de ballon blaast, wordt de lokale kromming steeds kleiner. Dit is uitgebeeld in fig. D.5. Het horizonprobleem wordt ook verholpen door een korte periode van inflatie. Het causaal verbonden heelal was voor inflatie heel erg klein, maar is door inflatie vele malen groter geworden. Als de huidige afmeting van

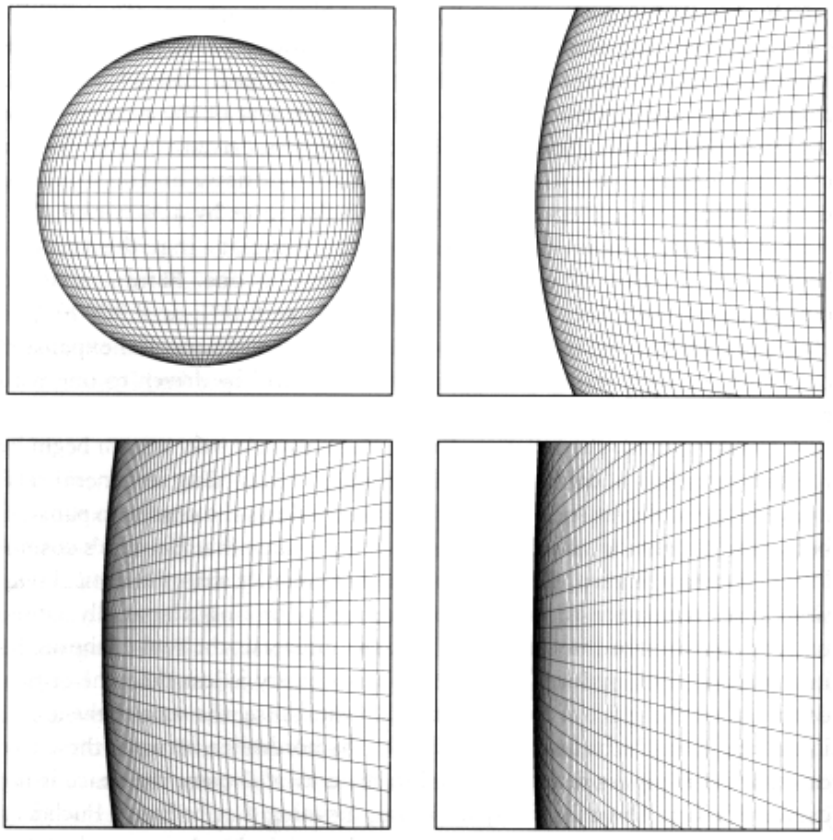

Figure D.2: Door de expansie van de het sterk gekromde oppervlak wordt de lokale kromming steeds minder zichtbaar. Zo lost inflatie het vlakheidsprobleem op.

dit causale gedeelte tenminste zo groot is als het huidige zichtbare heelal (de hemel gevuld met de CMB) dan kunnen alle gebieden binnen het zichtbare heelal in thermische evenwicht zijn. Immers, voor de expansie was dit gebied causaal verbonden en er heeft dus thermalisatie kunnen plaats vinden.

Inflatie 2.0

Een onverwacht bijverschijnsel van inflatie werd pas enkele jaren later ontdekt (1982). Zoals je in fig. 1.4 kunt zien zijn er wel degelijk hele kleine variaties in de CMB. Waar komen deze kleine fluctuaties vandaan? We hebben juist net beargumenteerd dat we inflatie nodig hebben om het heelal op niet causale schalen in thermisch evenwicht te krijgen. Toch is inflatie *ook* verantwoordelijk voor de kleine variaties in temperatuur in de achtergrondstraling. We hebben uitgelegd dat inflatie wordt veroorzaakt door een veld. Bij de hele hoge energieën die gedurende inflatie een rol speelden in het heelal krijg je ook te maken met de *quantum* eigenschappen van het veld. Een quantumveld is in staat om voor een korte tijd zijn energiedichtheid te verhogen

of te verlagen. Dit wordt een quantumfluctuatie genoemd en is het gevolg van Heisenberg's onzekerheidsprincipe in de quantummechanica. Dit principe vertelt ons dat in de quantumwereld er een systematische onzekerheid bestaat bij het bepalen van de plaats en de snelheid van een deeltje. Dit kan direct worden gekoppeld aan een onzekerheid bij het bepalen van de tijd en de energie van een quantumsysteem. Deze onzekerheid is fundamenteel, in de zin dat er geen experiment bestaat dat deze onzekerheid kan ontlopen. Hoe korter deze fluctuatie duurt (nauwkeurige tijdsbepaling), de te groter is de variatie in energie dichtheid (onnauwkeurige energiebepaling). Deze fluctuaties gebeuren in elk quantumveld, dus ook in een elektrisch of magnetisch veld. Echter, onder normale omstandigheden blijft dit een quantumfenomeen; op grote schaal (groter dan een paar atomen) middelen deze fluctuaties elkaar uit. Het bijzondere van deze fluctuaties tijdens kosmologische inflatie is juist dat deze kunnen worden vergroot. De snelle expansie van het heelal zorgt ervoor dat deze quantumfluctuaties niet meer uitmiddelen omdat ze voor die tijd al te veel zijn opgeblazen. Zodra dit gebeurt verliezen deze fluctuaties hun quantumeigenschap en zullen ze dus niet meer verdwijnen. Je kunt je voorstellen dat deze fluctuaties in energie (dichtheid) leiden tot kleine verschillen in de metriek (de geometrie), vanwege de koppeling tussen geometrie en energiedichtheid. Deze variaties in de geometrie beïnvloeden op hun beurt de fotonen, elektronen en protonen in de oersoep. Zo kan het dus dat quantumfluctuaties tijdens inflatie verantwoordelijk zijn voor de variaties in de CMB. Sterker nog, ook de fluctuaties in materie dichtheid zijn het gevolg van deze fluctuaties. En juist deze fluctuaties hebben geleid tot de vorming van sterrenstelsels, sterren en planeten. Deze worden gevormd na recombinatie wanneer ook materie onder de zwaartekracht ineen valt. De eerste sterren zijn waarschijnlijk een paar honderd miljoen jaar na het einde van inflatie ontstaan. Zonder quantum fluctuaties zou het heelal een saaie bedoeling zijn geweest.

ΛCDM

In vogelvlucht hebben we nu de belangrijkste aspecten van het moderne Oerknalmodel behandeld. In fig. D.6 zie je een overzicht. Het heelal ontstaat 13.7 miljard jaar geleden. Hoe het heelal precies is ontstaan is niet bekend en op het moment is het niet duidelijk of we ooit een waarneming kunnen doen die direct inzicht zou kunnen geven bij het beantwoorden van deze vraag. We kunnen wel speculeren, en dat wordt ook volop gedaan, maar daar zal ik nu verder niet op in gaan. Vrijwel direct na de oerknal is er een korte periode waarin het heelal exponentieel uitdijt ten gevolge van inflatie. Inflatie zorgt voor de juiste begincondities die o.a. verklaren waarom het heelal in thermisch evenwicht lijkt op niet causale afstanden en waarom het heelal een vlakke geometrie heeft. Direct na inflatie wordt de potentiële energie in het quantum veld omgezet in fotonen (straling) en materie. In feite zou je kunnen zeggen dat vanaf dat moment het oude oerknal model (voor 1980) begint. Straling en materie vormen voor enkele duizenden jaren een dichte oersoep. Nadat protonen en elektronen neutraal waterstof hebben gevormd ontstaat de achtergrondstraling. Tot de vorming van het CMB gebeurd er relatief weinig met de statische eigenschappen van de oersoep en geeft de CMB een goede representatie van het heelal vlak na het einde van inflatie. Daarmee biedt de CMB de ideale gelegenheid om inflatie te bestuderen. De verdeling van temperatuurfluctuaties in de CMB is een direct gevolg van de verdeling van kleine dichtheids verschillen vlak na het einde van in-

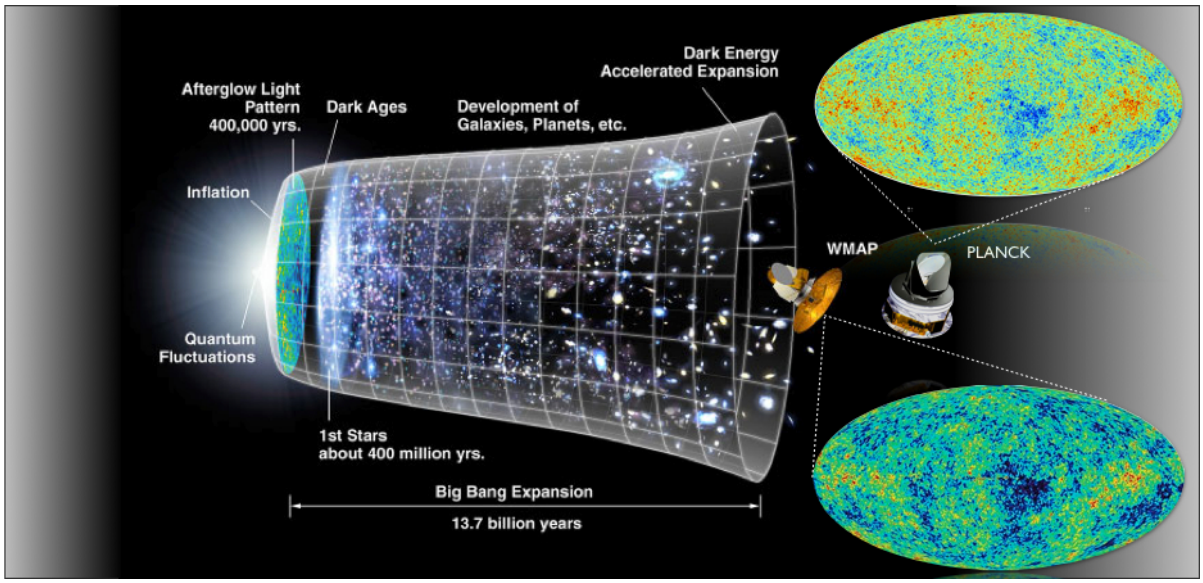

Figure D.3: De onstaansgeschiedenis van het heelal. Kort na de oerknal begint inflatie. In een zeer korte periode groeit het heelal exponentieel. Na inflatie domineren straling en gewone materie het heelal. In het begin zijn straling en materie nog heel sterk aan elkaar gekoppeld tot dat het heelal zo ver is uitgedijd dat neutraal waterstof ontstaat. Op dat moment kunnen fotonen vrij bewegen en vormen ze de achtergrondstraling, die vandaag de dag op microgolflengtes zichtbaar is met behulp van bijvoorbeeld de WMAP en de Planck satelliet. Kleine quantumfluctuaties tijdens inflatie leiden tot kleine variaties in de CMB en in de dichtheid van materie. Die kleine fluctuaties zijn de bron van structuur in het heelal; de sterrenstelsels, sterren en planeten. Het huidige heelal bestaat voornamelijk uit twee onbekende componenten, donkere energie en donkere materie. Donkere energie zorgt ervoor dar het heelal opnieuw versneld aan het uitdijen is.

flatie. Het bestuderen van die verdeling is een van de onderwerpen van dit proefschrift en geeft antwoord op de tweede vraag uit de inleiding; Hoe achterhaal ik de werkelijke begintoestand met behulp van een kosmologische waarneming? Het antwoord op deze vraag is met behulp van de statistische verdeling van temperatuurfluctuaties in de microgolfachtergrondstraling.

Na de ontkoppeling van straling en materie ontstaan de eerste sterren en uiteindelijk vult het heelal zich met stelsels van sterren en hun planetaire systemen. Vandaag de dag kunnen we de achtergrondstraling met grote nauwkeurigheid meten. Met behulp van nieuwe technieken kunnen we Hubble's metingen van de expansiesnelheid van het heelal met grote nauwkeurigheid herhalen. Een verassende waarneming is dat het er nu op lijkt dat het heelal versneld aan het uitdijen is. De component verantwoordelijk voor deze versnelde expansie wordt ook wel donkere energie genoemd omdat we net zoals in het geval van inflatie niet weten wat het precies is en het zich louter en alleen als energie manifesteert. Daarnaast blijkt dat er ook nog een onbekende soort materie moet zijn (naast protonen en elektronen). Dit kan o.a. worden afgeleid uit het aantal sterrenstelsels binnen een gebied op een gegeven afstand. Hieruit blijkt dat er massa moet zijn die we niet zien en niet kunnen terugvinden in alleen gewone materie (voornamelijk protonen en elektronen). Net als de donkere energie kunnen we deze massa niet waarnemen

en die wordt daarom ook wel donkere materie genoemd. Ditmaal manifesteert de component zich wel als materie (en niet alleen als energie) en kunnen we ook over dichtheids verschillen spreken, omdat de deeltjes die samen de donkere materie vormen, worden beïnvloed door de zwaartekracht. Uit metingen blijkt dat het huidige Heelal slechts voor 4% uit gewone materie bestaat, voor 23% uit donkere materie en voor 73% uit donkere energie (wat ook de versnelde expansie verklaart; donkere energie is de dominante vorm van energie in het nabije heelal). Vooral het gegeven dat het Heelal voor maar een heel klein deel bestaat uit gewone materie is verbazingwekkend. Alles wat wij kunnen zien, de aarde, de zon, de sterren etc, bestaat namelijk uit deze gewone materie. Daarentegen bestaat het heelal dus voor een veel groter deel uit dingen die we niet kunnen zien.

Het meest bestudeerde en best geteste model van het heelal is vernoemd naar deze dominante componenten. In het ΛCDM model van het heelal staat Λ voor donkere energie en CDM voor cold dark matter, ofwel koude donkere materie. In de simpelste 6-parameter variant van dit model zijn twee van de vrije parameters de hoeveelheid Λ en de hoeveelheid CDM. Een derde parameter meet de hoeveelheid gewone materie terwijl een vierde parameter de mate van interactie meet tussen de fotonen uit de CMB en geïoniseerd waterstof gevormd door sterren. Dit beïnvloed de waargenomen CMB en bij de analyse van de CMB moet hier rekening mee gehouden worden. De laatste 2 parameters beschrijven de begincondities, vastgelegd door inflatie. Deze parameters beschrijven de statistische verdeling van de quantumfluctuaties na afloop van inflatie die ten grondslag liggen aan de temperatuurvariaties in de CMB. De eerste vraag die ik stelde in de inleiding was: Gegeven een theorie die de meest extreme toestand in het heelal omschrijft (inflatie), wat is de voorspelling voor de begintoestand? Het antwoord op deze vraag is niet eenduidig, en zal afhangen van de details van het inflationaire model.

Mijn Onderzoek

Een groot deel van het onderzoek in dit proefschrift richt zich op het beantwoorden van deze vraag voor een specifiek model van inflatie. We hebben voornamelijk gekeken wat de effecten zijn van een kleine verandering in de *vacuümtoestand* **voor** inflatie op de verdeling van fluctuaties aan het **einde** van inflatie. De vacuümtoestand speelt een belangrijke rol in quantum velden theorie het is de laagste energietoestand van een systeem. Normaal gesproken betekent dit dat de vacuümtoestand de lege toestand is, een toestand zonder deeltjes. De motivatie om kleine aanpassingen van de vacuümtoestand te bestuderen komt voort uit het feit dat de structuur van het vacuüm bij hele hoge energieën onbekend is. Kortweg, de vacuüm toestand ten tijde van inflatie is niet perse een lege toestand. In de literatuur wordt veelal het *vrije vacuüm* gebruikt om de gevolgen van inflatie op de begincondities te berekenen. Dit vrije vacuüm is in feite dezelfde lege toestand als het vacuüm bij lage enegieën in een niet inflationaire achtergrond. Wij hebben onderzocht wat er gebeurd zodra je afwijkt van de veronderstelling dat het vacuüm bij aanvang van inflatie het vrije vacuüm is. Mochten die afwijkingen groot zijn en worden waargenomen, dan geeft dat inzicht in zowel de fysica *van* inflatie en wellicht inzicht in de fysica *voor* inflatie.

Gaussiteiten en Niet-Gaussiteiten

Zoals gezegd zijn de begincondities vlak na inflatie te beschrijven met behulp van 2 parameters die kunnen worden afgeleid uit de waargenomen fluctuaties in de CMB. De verdeling van temperatuurfluctuaties in de CMB is vrijwel Gaussisch. Dit betekent dat je met twee parameters de statistische eigenschappen van deze verdeling kan beschrijven: het gemiddelde en de variantie. De gemiddelde temperatuur is voor ons onbelangrijk omdat we geïnteresseerd zijn in temperatuurverschillen. Wat overblijft isde variantie. De variantie is een maat voor de gemiddelde afwijking van het gemiddelde. Echter, we kunnen de verdeling ook bepalen als functie van de hoek aan de hemel. Een hoek aan de hemel is in feite een (geprojecteerde) lengteschaal. Het blijkt dat de variantie van de CMB van deze schaal afhangt. Het merendeel van deze afhankelijkheid is het gevolg van fysica na inflatie, maar een klein deel is een gevolg van inflatie zelf. De amplitude P en de schaal afhankelijkheid n_s van de variantie ten gevolge van inflatie vormen samen de 2 parameters die de begincondities beschrijven. In dit proefschrift hebben we echter gekeken naar parameters die als mogelijke toevoeging dienen van het 6-parameter ΛCDM model. We zoeken naar zulke extra parameters omdat de meeste modellen van inflatie min of meer dezelfde voorspelling doen voor de waarde van P en n_s. Het nadeel hiervan is dat als je deze parameters meet, met behulp van bijvoorbeeld de CMB, je niet in staat bent om modellen van elkaar te onderscheiden. Kortom, we moeten op zoek naar een andere manier om dit te bewerkstelligen.

In hoofdstukken 2, 3 en 4 hebben we onderzocht of een afwijking van het vrije vacuüm voor inflatie kan leiden tot een afwijking van een Gaussische verdeling van fluctuaties. Niet-Gaussiteiten leveren een heel palet aan nieuwe parameters op, maar observationele beperkingen geven ons momenteel alleen de mogelijkheid om het equivalent van de schaalonafhankelijke variantie, de zogenaamde 'skewness' (scheefheid) amplitude $f_{\rm NL}$, te meten. Wat de waarnemingen ons in ieder geval vertellen is dat de afwijking van een Gaussische verdeling van temperatuurvariaties heel erg klein is. Door het uitrekenen van effecten van een afwijking in het vrije vacuüm op de skewness en deze te vergelijken met de CMB data, zijn we in staat om te achterhalen hoe groot deze mogelijke afwijking is. In hoofdstukken 2, 3, en 4 hebben we uitgerekend dat een afwijking van het vrije vacuüm grote, niet-Gaussische correcties kan opleveren, maar dat de amplitude $f_{\rm NL}$ afhangt van het inflatiemodel.

We hebben ook laten zien dat het in veel gevallen lastig is om de voorspellingen te vergelijken met de data, omdat de schaalafhankelijkheid van de voorspelde niet-Gaussiteiten enorm verschillen van de type niet-Gaussiteiten die tot nu toe gemeten zijn in de CMB. Zoals gezegd is de huidige resolutie van de data niet in staat om de schaalafhankelijkheid van de skewness te meten voornamelijk omdat het signaal zo klein is. Dit betekent dat als we de skewness amplitude $f_{\rm NL}$ meten, we een aanname moeten maken over hoe deze afhangt van de schaal. De verschillen in schaalafhankelijkheid zijn te beschrijven als verschillende vormen van de niet-Gaussiteiten. Om onze berekeningen toch te kunnen vergelijken met de metingen van $f_{\rm NL}$ hebben we geprobeerd om te kwalificeren in hoeverre twee verschillende vormen op elkaar lijken. Dit levert een zogenaamde 'overlap' factor op, die ons in staat stelt om te bepalen of er wellicht iets van het voorspelde signaal verstopt zit in de meting. Omdat deze factor heel klein is, is het moeilijk om een potentiële afwijking van het vrije vacuüm volledig uit te sluiten. Wel is gebleken dat de huidige metingen al een behoorlijke limiet stellen aan de toegestane afwijking.

In hfst. 6 hebben we niet gekeken naar een extra parameter in de vorm van de skewness van de verdeling van fluctuaties, maar hebben we onderzocht of de schaalafhankelijkheid van de variantie oscilleert. In feite kijken we dan dus naar de verandering van de schaalafhankelijkheid van de variantie. Dit zouden we met minimaal 1 extra parameter kunnen beschrijven, maar omdat we hier specifiek kijken naar oscillaties hebben we te maken met tenminste 2 extra parameters: de frequentie en de amplitude van de oscillatie. Een afwijking van het vrije vacuüm voor inflatie voorspelt dat zulke oscillaties kunnen optreden. Daarnaast zijn er in de afgelopen jaren verschillende modellen van inflatie bedacht die ook oscillaties voorspellen in de schaalafhankelijkheid van de variantie. Het is gebleken dat het traceren van deze oscillaties in de data heel moeilijk is. De oorzaak hiervan ligt opnieuw in de nauwkeurigheid van de data; de data zijn simpelweg nog niet goed genoeg om met zekerheid te zeggen dat oscillerende signalen aanwezig zijn in de variantie. We hebben wel laten zien dat er bepaalde frequenties waarvoor de theoretische voorspelling iets beter overeenkomt met de waarneming. Het is echter lastig te zeggen of dit een gevolg is van de lage kwaliteit van de meting of dat we we hier inderdaad een eerste aanwijzing hebben voor het bestaan van oscillaties. We zijn nagegaan of deze frequenties overeen zouden kunnen komen met de theoretisch voorspelde oscillaties vanuit afwijking van het vrije vacuüm en dit heeft geleid tot de conclusie dat er een zeer beperkte ruimte over blijft waarin de waargenomen variantie nog overeenstemt met de voorspelde variantie. Zodoende hebben we opnieuw limieten kunnen zetten op mogelijke correcties van het vrije vacuüm .

Een van de moeilijkheden bij het vergelijken van de voorspellingen van niet-Gaussiteiten met de data is dat net als de variantie, de schaalafhankelijkheid van de skewness kan oscilleren. Dat is ook de reden dat eerder genoemde 'overlap' factor met bestaande metingen van de skewness klein zijn. In hfst. 5 hebben we onderzocht of we hier iets aan kunnen doen door direct te kijken naar oscillerende signalen in de skewness. Een probleem daarbij is dat het vergelijken van de voorspellingen met de metingen nogal wat rekenwerk vereist, veel meer dan bijvoorbeeld het meten van de variantie. Om dit rekenwerk te versnellen heb ik in hfst. 5 een nieuwe methode onderzocht die de berekening zowel efficiënter als simpeler zou moeten maken. Ik heb later zien dat in bepaalde gevallen dit inderdaad mogelijk is, zolang je de opzet van deze methode optimaliseert voor oscillerende signalen.

Wij denken dat we een paar zeer interessante resultaten hebben geboekt in dit proefschrift. Uiteraard roept dit proefschrift ook weer vragen op. Uiteindelijk zal moeten blijken of we in staat zijn om de waarnemingen en de voorspellingen op een dusdanige manier te combineren om een duidelijke stellingname te nemen over een mogelijke afwijkingen van de vacuüm toestand ten tijde van inflatie. De komende tijd ga ik mij richten op het optimaliseren van de zoektocht naar oscillerende signalen, zowel in de variantie als in de skewness. De resolutie van de nieuwe data afkomstig van de Planck satelliet zal uitkomst bieden, hoewel de eerste resultaten pas in 2012 en 2013 te verwachten zijn. Tot die tijd zal ik ook proberen om een schatting te maken van de significantie van de gevonden signalen in hfst. 6. Hiervoor zal ik Planck en WMAP data simuleren die wel of geen oscillerend signaal bevatten. Hieruit zal moeten blijken of het mogelijk is om onderscheid te maken tussen detectie en non-detectie.

Summary

Introduction

The Universe is an immense space, filled with galaxies, stars and planets. Strikingly, the grand scale evolution of the Universe can be described by a total of 6 parameters. This is the remarkable conclusion scientists have made almost 100 years after Einstein introduced his theory of Gravity. His theory allowed us for the first time to describe a dynamically evolving Universe. Roughly, we can divide the 6 parameters as follows:

- Interactions in the Universe (1)
- Energy content of the Universe (3)
- The initial condition (2)

The research carried out in this thesis aims at the exploring the initial conditions of the Universe. Modern cosmology tells us that the Universe came into existence about 13.7 Billion years ago through a process that is commonly referred to as the Big Bang. As we will see in this summary, the initial conditions we are interested in are not set right at the beginning, but rather in a very brief period directly after. During this period the conditions are extreme: super high temperatures energies and densities dominate the Universe. Since we know very little about how nature behaves under these extreme conditions, finding a way to explore this period would pose a tremendous opportunity. Especially since these high energy conditions are impossible to reproduce in a laboratory on earth. The Big Bang might therefore offer us the only possibility to investigate the laws of nature under these extreme conditions. Our challenge is to recover information about the initial conditions using cosmological observations and relate these observations to theoretical models. We can break-down the quest presented in this thesis into two questions: Given a theory describing the early, extreme Universe, what are the predictions for the initial conditions? and How do we extract valuable information about this initial condition from cosmological data? Once these questions have been answered, we should be able to reconstruct the theory describing the very early Universe. Before I begin to explain how I tried to answer these questions, let me first given a brief introduction of the evolution history of the Universe.

Einstein's Gravity

The modern theory of Gravity introduced by Einstein (1905-1916), provided the foundation to describe the dynamical evolution of the Universe as a whole. The essential building blocks of Einstein's theory postulated that space and time are related through spacetime, that the geometry of spacetime depends on the energy inside it, and that there exists a fundamental limiting speed at which information can be exchanged (the speed of light). Using these ideas it became possible to describe the Universe's past, present and future.

The reason we consider spacetime, as opposed to space and time, was motivated by Einstein. He showed that both space and time or not *invariant* (unchanged) with respect to the motion of an observer, i.e., the measurement of bits and pieces of time and space will depend on the observer. For example, an observer at rest (velocity = 0) will measure passing time differently from an observer that is moving with respect to the observer at rest (velocity $\neq 0$). The differences between those measurements will increase as the relative velocity between both observers increases. Once this difference equals the speed of light, the measured differences will be maximal. If you formulate a law describing nature, and you assume that this law should hold throughout the entire Universe, you would wish that this law is independent of the observer, in the sense that we would like to avoid drawing different conclusions on how nature works based on the *reference frame* of the observer. Einstein showed that although time and space independently are not invariant, spacetime is. Effectively, the differences between measurements of space and time cancel each other out. In order to define the invariant spacetime distance, we make use of a *metric*. A metric tells us how to add measurements of space and time in such a way to form the *invariant* quantity known as the interval. To apply Einstein's theory to cosmology we need to define the spacetime metric of the Universe. It turns out the metric of the large scale evolution of the Universe is relatively simple and this metric was derived independently by Friedman, Robertson, Walker and Lemaître. The FRLW metric is based on the observation that the Universe is homogeneous and isotropic on large scales: the Universe appears the same in all directions on every position in the Universe. In addition, it turns out that the Universe is close to flat. Equivalent to how a surface can be curved, such as the surface of a balloon, higher dimensional space can be curved as well. When we say the Universe is flat, we infer that the 3 dimensional space in the Universe is flat. Although it is very difficult to imagine a 3 dimensionally curved space, we are able to perform experiments to test whether the geometry of the Universe is curved. For example, it turns out that the sum of the angles of an equilateral triangle in curved space differs from that in flat space. Cosmological observations that perform such a test in an ingenious way have shown that the Universe is flat.

The second idea proposed by Einstein tells us that the geometry of space, described by the metric, depends on the contents inside spacetime. When we say contents or components we typically refer to mass or energy (matter and energy are related through the well known relation $E = mc^2$ with m the mass and c the speed of light) inside the Universe, such as matter. The prescription for the invariant spacetime interval, the metric, will change as function what contents are inside the Universe. Einstein's equations allowed us to build a theory of the Universe by deriving a metric. By specifying the contents, the equations allow us to compute

the evolution of the Universe as a function of time[2]. Now imagine yourself gazing at the sky in the times of Einstein, wondering how the Universe changed as time progressed. It would come as no surprise if you concluded that the Universe would not evolve at all; the Universe has been the same for all eternity. This is exactly the same view Einstein had in mind when he first tried to solve his equations. To his surprise his own equations did not allow for a static solution, no matter what he put into his Universe, his equations led to a dynamical (time dependent) solution. He tried to solve this 'issue' artificially by slightly modifying his geometry, but the Dutch physicist de Sitter showed that this would still lead to an evolving Universe. Even though Einstein did not like it, it was a remarkable conclusion and finding, since it meant the Universe must have looked different in the past and would probably look different in the future.

The first circumstantial evidence that the Universe was truly dynamical was obtained by Edwin Hubble in 1929. He discovered that distant galaxies moved away from us, and that velocity at which this happened depended on how far these galaxies were away from us. This would imply the Universe was expanding. Obviously this observation was met with serious skepticism. However, it did provide plenty of material for discussion. This discussion lead to the insight that if the Universe was truly expanding, the Universe should have been denser and compacter in its past. Eventually, if you evolve the Universe backward in time, the Universe should become infinitely dense and compact. And so, the idea of a Big Bang was born.

Microwave Background

In 1984 George Gamow and two of his colleagues came to the conclusion that the hot, dense phase after the Big Bang could produce the observed level of Hydrogen and Helium. Although this was a very important argument in favor of a Big Bang, there existed alternative theories that could produce the observed abundance. Fortunately, they predicted the Big Bang should produce an observable. They argued that a Big Bang should leave an afterglow, much like the afterglow following the ignition of a bomb, and that this afterglow should be observable today. In 1964 this afterglow was detected for the first time and its detection should be considered the single most important evidence in favor of a Big Bang scenario. This afterglow is known as the Cosmic Microwave Background (CMB) and is believed to be a direct representation of what the Universe looked like only several hundred thousand years after the Big Bang. By studying the details of this afterglow we should be able to gain insight into the earliest moments of the evolution of our Universe.

What causes the CMB? In the seconds immediately after the Big Bang, the Universe is super compact and highly energetic. This causes particles to be closely packed and have regular interactions (scattering). The first few thousand years after the Big Bang, the Universe is dominated by radiation. Radiation exists of radiation particles, known as photons. The Universe also contains matter, which consists mostly of Hydrogen (75%) and Helium (25%). A hydrogen atom consists of two particles: a proton (positive charge) and an electron (negative

[2]Although I discussed geometry and its relation to energy separately, these are not independent. For example, the metric describing the invariant spacetime interval does depend on the distribution of energy. The FRWL metric was derived in the assumption that the Universe is homogeneous and isotropic in its energy contents.

charge). The high temperatures in the early Universe make photons very energetic. As a result these photons keep the protons and electrons from forming neutral Hydrogen. They Universe is therefore dominated by ionized Hydrogen which is a way more effective scatterer compared to neutral Hydrogen. Therefore photons can not travel for a very long distance before they interact with an ionized hydrogen atom; they are not able to move *freely*. Within the hot Big Bang scenario the Universe expands as time goes by. The expansion causes the effective temperature in the Universe to drop and photos become less and less energetic. After a few hundred thousand years photons no longer have enough energy to keep free electrons from combining with free protons to form neutral Hydrogen. A neutral Hydrogen atom has much fewer interactions with photons, resulting in less photon scattering. The formation of neutral Hydrogen happens in a relatively short period while the free path of photons increases significantly. In fact, the distance the photons can suddenly travel without interactions exceeds the distance they could travel within the age of the Universe (13.7 Billion years). The transparency of the Universe coincides with the formation of the CMB. At the time of CMB formation the effective temperature of the Universe is a few thousand degrees. As the Universe expands, the radiation of the CMB slowly cools all the way to just a few degrees in todays Universe. The temperature of the photons making up the CMB today correspond to a wavelength of a few millimeters, which is the same radiation wavelength found in your microwave, hence the name Cosmic Microwave Background. Since the free streaming of photons occurs everywhere throughout the Universe at the same time, the CMB radiation will appear from all directions across the sky. No matter in which direction you put your detector, you will always collect CMB photons. The biggest source of radiation in the Universe is the CMB, producing more radiation than all stars in Universe combined. If you take out the antenna of your (analog) TV, you will see noise (snow), most of which is due to thermal noise inside your TV equipment. However a tiny fraction, around 1%, is due to the CMB.

Inflation

After the first detection of the CMB in 1964 by Penzias and Wilson (nobelprize 1978), many new experiments have been employed to collect more data. Most of these experiments covered only a small part of the sky. To cover the entire sky requires a detecter in space. The first of such detectors was launched in 1989 (Cosmic Background Explorer, COBE) which was succeeded by the Wilkinson Microwave Anisotropy Probe (WMAP) in 2001. These instruments were capable of taking an image of the full sky. Such an image, taken with WMAP, is shown in fig. D.4. The first full sky image showed that the CMB was remarkably homogeneous, which basically means that any deviations from the average (temperature) are really small. The average temperature of the CMB is about 2.7 Kelvin (-270 degrees Celsius) and any point in the CMB deviates at most a few micro Kelvin from this average. Homogeneity of the CMB implies that the CMB photons were in thermal equilibrium at the moment the CMB was formed. Thermal equilibrium is reached when particles inside a closed system exchange (thermal) energy until the average energy per particle is about the same. For example, if you pour cold milk in your hot tea, warm water particles will slowly transfer their thermal energy to the cold milk particles, until the milk-tea reaches an equilibrium temperature that exceeds that of milk and is lower than that

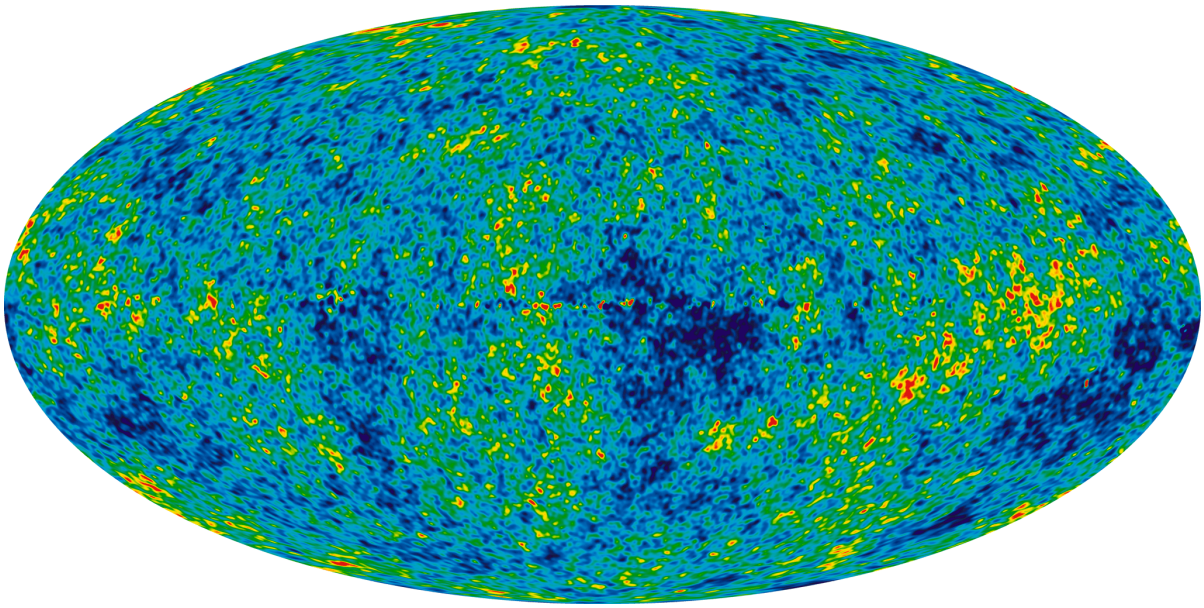

Figure D.4: An image of the Cosmic Microwave Background (CMB) taken with the camera on board the Wilkinson Microwave Anisotropy Probe (WMAP) satellite. The average temperature of the radiation making up the CMB is about 2.7K, close to absolute zero. On top of the 2.7K background there are tiny variations which are believed to be relics of the physics in the very early, highly energetic Universe. Analyzing the statistics of distribution of these variations allows us to derive information about the conditions present only a fraction of a second after the Big Bang. A color version of this image is shown in fig. 1.4, ch. 1.

of the hot tea. Eventually, the tea will also reach thermal equilibrium with its surroundings, cooling down to for example 20^O Celsius inside your living room. The reason it cools all the way down to this temperature is because there are so many 'cold' particles in your living room (and beyond); if you would have a very good thermometer, you could hypothetically measure the living room warming up due to hot particles in the tea. An important observation is that a system does not reach thermal equilibrium instantly, since it takes a while to exchange thermal energy. This is because particles need to be close enough to each other to exchange energy. The CMB is made of photons, which propagate at the speed of light. Therefore the maximum speed at which thermal information can be exchanged is the speed of light. Even though this is really fast, on cosmological distances it means that thermalizing the CMB should take a while. If two points in the Universe are able to exchange thermal energy within the life time of the Universe they are said to be causal, or causally connected. To explain a CMB in thermal equilibrium we need to know if all points in the CMB are causally connected. As it turns out, even today not all points in the Universe are causally connected. For example the 'north pole' and 'south pole' in fig. 1.4 are separated by a distance twice the age of the Universe, or around 28 billion light years. Since the Universe is only 13.7 billions years old, there is no way these point could have exchanged thermal information. However, in the assumption light from the CMB has been free streaming towards us since CMB formation, thermal equilibrium was

already reached when the CMB was formed. The Universe was only 380.000 years old at that time, which roughly corresponds to a circumference of 1 degree on the sky today. In other words, how is it possible that the CMB we measure today is in thermal equilibrium, while on causal arguments we would expect thermal equilibrium on scales of at most 1 degree. This paradox is known as the horizon problem.

Another striking observation is that the geometry of the Universe is so close to flat. Solving Einstein's equations shows that this is an unstable solution; small deviations from flatness should quickly become large as the Universe evolves. This paradox is known as the flatness problem. Both problems can be considered 'fine-tuning' problems; only by precisely fine tuning the initial conditions are we able to explain our observations. This is often considered as an inadequate understanding of the laws of nature, since we prefer to be able to (dynamically) explain these observations as unavoidable consequences of a physical process.

In 1981 Alan Guth proposed a physical solution. He argued that both the horizon and flatness problem could be solved by introducing a period of very rapid growth of the Universe. Within this period the Universe has to grow by at least a factor of of 10^{28} (a 1 followed by 28 zero's) all within a period lasting only 10^{-35} seconds! In order to solve the fine-tuning problems this period should be placed in the very early Universe. The rate at which the Universe grows actually implies parts of space are moving away from each other faster than the speed of light. However, no information is being exchanged, so there is no violation of special relativity. The resulting expansion is *exponential* and Guth called his idea cosmological inflation. He showed that inflation could be realized when the potential energy of a component in the Universe dominates the Universe. However, this requirement is necessary to make pressure become negative. This negative pressure leads to exponential expansion. When Guth formulated his theory, there was no known form of matter or radiation with the right properties resulting in negative pressure. He showed that a hypothetical *field* could meet the requirements necessary for inflation: the inflaton field. One of the characteristics of a field is that it has a value at every point in space. Radiation is an example of a field, but as said, is not suitable for inflation. The inflaton field needs to have a potential energy much larger than its kinetic energy. Since no such field exists within tested model of physics, exploring inflation would actually allow us to investigate unknown physics.

Inflation solves fine-tuning the initial conditions. The flatness problem is solved through stretching space to such extends that any curvature will be smoothed out. To understand this, consider the surface of a balloon. The surface of a balloon is curved when there is no air inside the balloon. The effect of inflation can be mimicked by letting air in the balloon. Positioning yourself at a point on the surface of the balloon you will observe all the (local) curvature disappearing as more and more air is put in the balloon. This idea is shown in fig. D.5. The horizon problem is also solved by inflation. The causally connected part of space in the Universe was very small before inflation. This part has been able to reach thermal equilibrium before inflation. The expansion of this part results in a much larger part of space that appears in thermal equilibrium today.

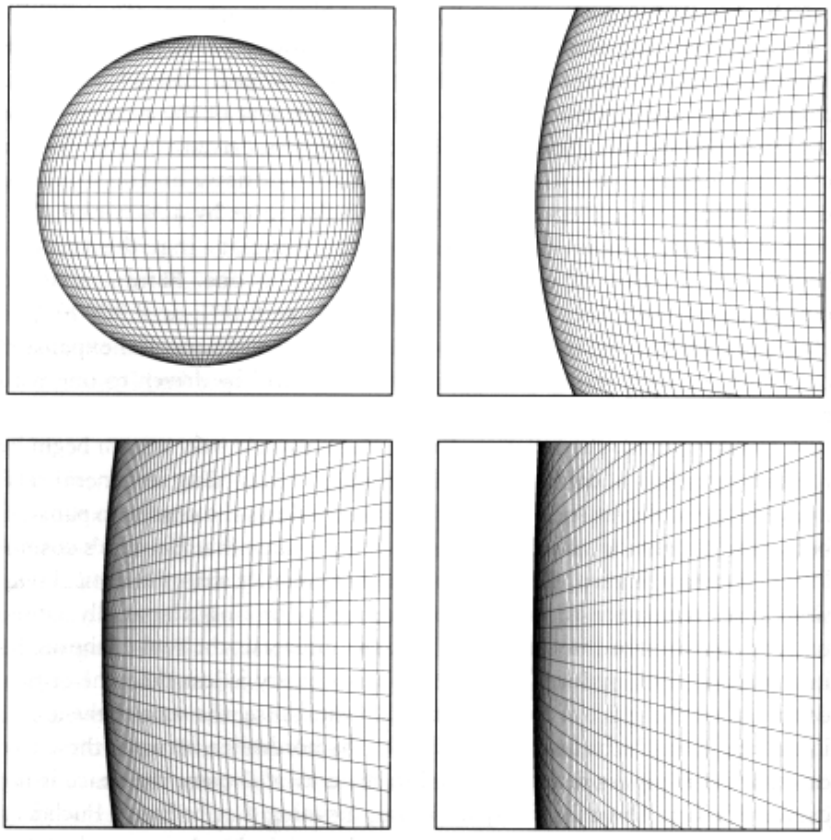

Figure D.5: The rapid expansion causes the local curvature of space to disappear, solving the flatness problem.

Inflation 2.0

An unexpected effect of inflation was discovered in 1982. From fig. 1.4 we see that the CMB *does* contain small variations. What is the origin of these small fluctuations? We just argued above inflation is required to smooth out all variations and create a patch of space that is in perfect thermal equilibrium. However, there is an elegant mechanism during inflation that is responsible for producing these tiny variations in the CMB. The inflaton field drives the exponential expansion of the Universe during a period when the Universe is extremely energetic. At these extreme energies we need to start considering quantum mechanical properties of the field. The quantum world is very different from our daily world , which is usually referred to as the *classical* world. Without explaining the details of quantum mechanics, we will note that a quantum field has a peculiar property. A quantum field is capable of temporarily lowering or rising its energy density. This is referred to as a quantum fluctuation and is a result of Heisenberg's uncertainty principle. This uncertainty is fundamental; there is no known experiment that can avoid this uncertainty. The shorter the fluctuation lasts (time), the larger the fluctuation in the energy density. These fluctuations happen all the time in the quantum

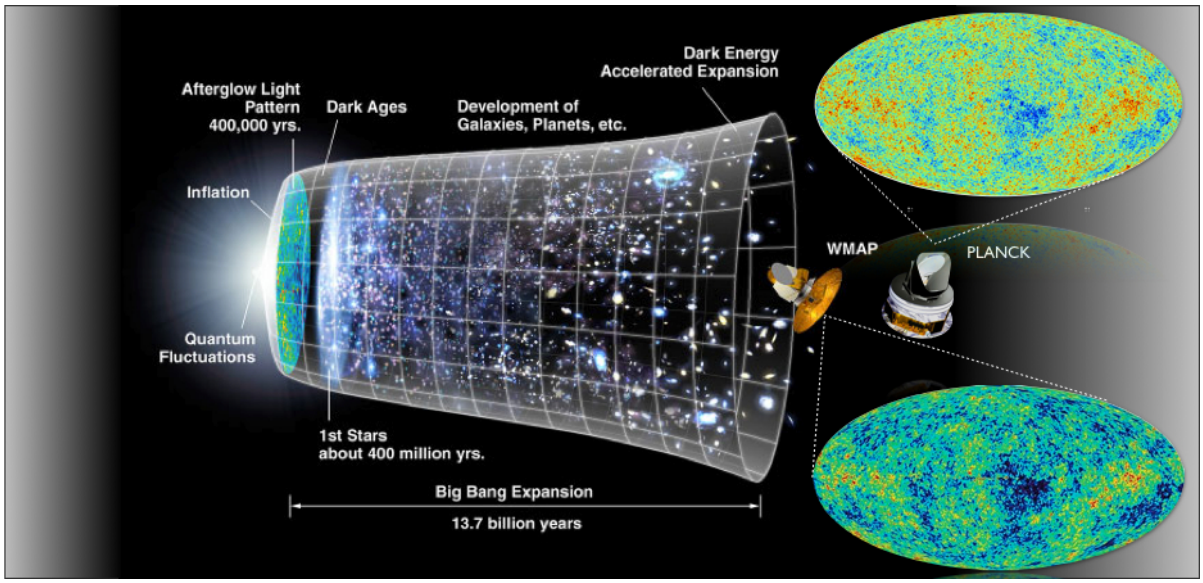

Figure D.6: The evolution history of the Universe. Directly after the Big Bang (left) inflation starts. This very brief period of exponential expansion solves the horizon and flatness problem. The potential energy of in the inflaton field is used to make matter and radiation which dominate the Universe until recently. At first, matter and radiation are tightly coupled. After about 380.000 years neutral hydrogen forms and the radiation photons become free streaming; the Universe becomes transparent and the CMB is formed. The CMB can now be detected at micro wavelengths using WMAP or the recently launched Planck satellite. Small quantum fluctuations during inflation have sourced temperature and matter fluctuations leading to the variations in the CMB and structure respectively. Currently, two unknown components dominate the Universe: dark matter and dark energy. The latter has a very similar effect as the inflaton field and as a result we have recently entered a new period of accelerated expansion.

fields, such as the electromagnetic field. The *Casimir effect* is a well-known example of this process. Usually however these fluctuations disappear, because there is an equal probability of negative and positive fluctuations, averaging out under normal circumstances. But inflation is not normal! During inflation these fluctuations can grow and become classical. As soon as they are classical, they will no longer disappear, since they lose their quantum nature; the quantum uncertainty has become a classical reality. The fluctuations in energy density couple to the geometry of spacetime through the Einstein equations. The geometry of spacetime couples to everything else (with non-zero energy) such as protons, electrons and photons. As such, quantum fluctuations are capable of sourcing temperature fluctuations in the CMB! Not only do they source the variation in the CMB, they also source small fluctuations in the matter density. These tiny fluctuations start growing rapidly as soon as the interactions between matter and radiation becomes negligible. Gravity causes the largest fluctuations to collapse and form the first stars a few hundred million years after the Big Bang. As time progresses more stars form and eventually the Universe fills up with galaxies of stars. Without a doubt, quantum fluctuations can be considered as the source of structure formation throughout the Universe, making the Universe a lot less boring than it would have been without them.

ΛCDM

In fig. D.6 we have summarized what we discussed so far, connecting the earliest moments to the Universe of today. The Universe formed approximately 13.7 billion years ago. How the Universe formed is not known, and it remains to be seen if we can actually empirically explore the creation of the Universe. That does not mean there is a lack of speculation, however we will not discuss this here. Immediately following the Big Bang, the Universe expands exponentially due to a brief period of inflation. Inflation creates a flat and homogeneous patch of Universe. The potential energy of the inflaton field is transferred into energy of particles and photons. Effectively, when this occurs the Universe enters the classical hot Big Bang evolution. Radiation and matter together form a hot dense plasma that dominates the Universe for the first few hundred thousand years. As soon as the Universe has expanded enough to cool down significantly, neutral Hydrogen forms leaving photons free streaming. This is the moment the CMB is formed. Quantum fluctuations during inflation are believed to have sourced temperature fluctuations in the CMB and density fluctuations in the distribution of matter. Very little has happened with the distribution of fluctuations between the end of inflation and the formation of the CMB, making the CMB almost a direct image of the conditions right after inflation. By studying CMB fluctuations or matter fluctuations we can explore the conditions in the early Universe. Studying the distribution of temperature fluctuations in the CMB is one of the main topics of this thesis and answer the second question phrased in the introduction: How do we extract valuable information about this initial condition from cosmological data? Answer: By using the using statistical distribution of temperature fluctuations in the cosmic microwave background.

After decoupling of matter and radiation, matter starts to clutter through gravitational forces. These clumps until they have formed compact objects, such as galaxies and stars with their planetary systems. Using present-day instruments we can measure the CMB with great accuracy. Using new techniques and modern detectors we can also repeat Hubble's expansion measurement. Surprisingly, we find that Universe is currently expanding at an accelerating rate. Measurements of the CMB confirm this conclusion. This observation implies that our Universe is dominated by component with very similar characteristics as the inflaton field. This component is called dark energy, as it is only a form of energy and does not appear to have any particle characteristics. From the CMB and a number of other independent observation we find that there must be an additional form of energy: dark matter. Very little is known about dark matter, but we can infer from its distribution in the Universe that it must consist of particles, since it tends to cluster just as ordinary matter (protons, electrons). Observations show that dark mater makes up about 23% of the total energy density in the Universe, while ordinary matter accounts for only 4% to the total observed energy. The rest (73%) is dark energy, showing dark energy dominating the Universe today, explaining the accelerated expansion. Note that in principle the contribution of dark energy will only increase as the energy density is constant as the Universe expands, while both dark and ordinary matter will slowly dilute with the expansion of the Universe. No doubt, these are remarkable findings. Our Universe is dominated by things we cannot see and only 4% of the total energy density in today's Universe accounts for all the things that we do see, such as stars and galaxies of stars.

The most studied and intensely tested model of the Universe is named of its dominating

contents, the ΛCDM model of the Universe. Λ is known as the cosmological constant, and represents dark energy, while the abbreviation CDM stands for cold dark matter. In the simplest 6-parameter ΛCDM model, 2 out of 6 parameters are the total energy density of Λ and CDM. A third parameter measures the total energy density of ordinary matter. The fourth parameter measures the interactions between CMB photons and matter in the Universe after CMB formation. This measure is relevant as it will affect the observed CMB photons. The final two parameters describe the initial conditions as they were at the end of inflation. These parameters label the statistical distribution of quantum fluctuations set by inflation, sourcing the distribution of temperature fluctuations in the CMB. The first question I expressed in the introduction was: Given a theory describing the early, extreme Universe, what are the predictions for the initial conditions? The answer to this question is to be found in the physics of inflation, which is the other main topic of this thesis.

My Research

A substantial part of the research presented in this thesis aims at answering the question above for a given model of inflation. In particular, we have investigated the possible effects of modifying the *vacuum state* **before** the onset of inflation on the statistical distribution of fluctuations at the **end** of inflation. The vacuum state plays an important role in quantum field theory. It defines the lowest energy state of a quantum system. Under normal circumstances this implies the vacuum state corresponds to the state without particles: the empty state. The reason to study modification to the vacuum state is that the structure of the vacuum at very high energies is not well known (or understood). Theoretically it seems clear that the vacuum state at the onset of inflation does not necessarily need to be the empty state. Usually, when computing statistical properties of the primordial distribution of quantum fluctuations the assumption is made the vacuum is the *free vacuum*. The free vacuum is the equivalent to the vacuum at low energies in a non-inflating background. We investigated what the consequences are once you adopt a vacuum state different from the free vacuum. If the corrections to the initial conditions are significant, we can use observational constraints to increase our understanding of the physics of inflation.

Gaussianities and Non-Gaussianities

As previously explained, the initial conditions at the end of inflation can be described using 2 parameters. The distribution of temperature fluctuations, which are sourced by the initial conditions, is close to Gaussian. For a Gaussian there are two important quantities: the mean and the variance. The mean temperature of the temperature fluctuations is not very important as we are interested in statistical properties of temperature differences, therefore we can set it to zero. All we are left with is the variance, which is a measure for the mean deviation from the mean (zero). We can determine the variance of temperature fluctuations as a projected function of an angle on the sky. An angle on the sky represents a projected length scale. It turns out that the variance of temperature fluctuations depends on this scale. Most of this dependence is a result of the physics after inflation, but a small part of this dependence is

due to inflation physics. The amplitude P and the scale dependence n_s of the variance due to inflation together present the 2 parameters that dictate the initial conditions. However, in this thesis we have looked for parameters that could complement these two. The reason to look for additional information in the CMB is given by the fact that most models of inflation give very similar predictions for the value of P and n_s. Consequently it becomes very difficult to to discriminate between various models as soon as you have measured these two parameters using the CMB. It therefore makes sense to find other ways to differentiate between various models.

In chapters 2, 3, and 4 we investigated if a deviation from a free vacuum at the onset of inflation can result in deviations from Gaussianity. Non-Gaussianities contain a plethora of new parameters, but observational limitations have restricted us to investigate just a single one. Basically, it is the equivalent of the scale independent part of the variance, the so-called skewness of the distribution known as $f_{\rm NL}$. So far, observations have shown us that deviations from Gaussianity are small. By computing the effects due to a deviation from a free vacuum state on the skewness, and comparing this with CMB observations, we were able deduce constraints on vacuum modifications. In chapters 2, 3, and 4 we found that vacuum modifications lead to a non-Gaussian distribution of primordial fluctuations, but that the enhancement heavily depends on the physics that is driving inflation.

We also showed that in many cases it can be difficult to compare theoretical predictions with observations, because the scale dependence of the non-Gaussian corrections deviate from the non-Gaussianities that have been constrained so far. As was previously explained, the resolution of the current images of the CMB is not good enough to measure the scale dependence of the skewness because the amplitude of the skewness is very small. Therefore when we want to measure the skewness in the data, we need make an assumption on its scale dependence. Different scale dependence can be represented by the *shape* of non-Gaussianities. To allow comparison between theoretical predictions and observations we tried to qualify similarities between shapes. We computed the so-called 'overlap' between our predicted shape and constrained shapes. This overlap allowed us to estimate how much of the measured signal (skewness) could be due to non-Gaussianity derived from our model. We showed that current measurements do put a serious constraint on allowed modifications. Because the overlap is really small (the predicted shapes differ a lot from constrained types), it is difficult to completely rule out deviations from a free vacuum.

It is not strictly necessary to consider non-Gaussianities if you are interested in using additional parameters to investigate the physics of inflation. In ch. 6 we investigated whether the scale dependence of the variance oscillates. Basically, we are then trying to measure the change as a function of scale of the scale dependence of the variance. We could try to parametrize this using only 1 parameter (known as the running), but since we are specifically interested in oscillations we used (at least) 2 parameters: the frequency of the oscillation and the amplitude of the oscillation. Besides predicting non-Gaussian corrections to the primordial distribution, vacuum modifications also predict an oscillating correction to the variance. Moreover, an increasing number of models, not all related to vacuum modifications, predict similar corrections to the variance. In ch. 6 we argue that measuring these oscillations is a challenge as the data is simply not good enough. We show however that there exist various discrete frequencies that do allow for a better match of the model to the data. We have not been able to assess

a significance to these findings as our result might simply be due to noise. We did compare the observationally preferred frequencies and the associated amplitudes with our theoretical predictions. We found that there is only a small range in parameter space that still allowed by current data. By this means, we have been able to obtain additional constraints on free vacuum state modifications, complementary to those derived from the skewness.

One of the obstacles faced when comparing predicted non-Gaussian from vacuum state modifications with the data is that the scale dependence oscillates. This is the reason why the discussed earlier overlap is so small; constrained type non-Gaussian shapes are relatively smooth, suppressing overlap significantly. In ch. 5 we tried to resolve this issue by investigating the possibility to constrain oscillating skewness directly from the data. One of the problems with constraining non-Gaussianities directly is that it requires relatively time-consuming computations. In fact, until recently there was only a small class of non-Gaussian shapes that could be constrained at all. In ch. 5 I looked at a new method that should reduce the computational time significantly, by transforming any predicted skewness into same form as the ones that have already been constrained. We have showed that this is indeed possible for certain oscillating non-Gaussianities. We made additional proposals to further reduce computational costs.

We believe that this thesis presents some very interesting results. It also opens up new possibilities. Our priority in the near future will be to connect all the observations in such was to optimize the constraint on vacuum state modifications. We should further develop methods that optimize the search for oscillations in the scale dependence of the variance as well as the skewness. Data wise, in the next two years Planck data will be (publicly) released opening up opportunities to search for oscillations in this data. One thing I am particularly interested in is assessing significance to the oscillatory signal(s) found thus far. This is important because it is quite possible that the true signal will be very small. This means it could be easily buried in the noise. To investigate the significance we should simulate realistic data with and without oscillating statistical distributions in the presence of noise. Analyzing this data should give us insight to what extend these signals can be mimicked or recovered from the data and how these results depend on the level of noise and the resolution of the observations.

Bibliography

A. Achucarro, J. O. Gong, S. Hardeman, G. A. Palma & S. P. Patil, "Features of Heavy Physics in the CMB Power Spectrum," JCAP **1101**, (2011) 030. arXiv:hep-th/10103693.

V. Acquaviva, N. Bartolo, S. Matarrese & A. Riotto, "Second-Order Cosmological Perturbations from Inflation," Nucl. Phys. B **667**, (2003) 119. arXiv:astro-ph/0209156.

A. M. Aliaga, E. Martínez-Gonzalez, L. Cayon, F. Argueso, J. L. Sanz, R. B. Barreiro & J. E. Gallegos, "Tests of Gaussianity," MNRAS, **336**, (2002) 22. arXiv:astro-ph/0310706.

M. Alishahiha, E. Silverstein & D. Tong, "DBI in the Sky," PRD **70**, (2004) 123505. arXiv:hep-th/0404084.

L. E. Allen, S. Gupta & D. Wands, "Non-Gaussian Perturbations from Multifield Inflation," JCAP **0601**, (2006) 006. arXiv:astro-ph/0509719.

P. R. Anderson, C. Molina-Paris & E. Mottola, "Short Distance and Initial State Effects in Inflation: Stress Tensor and Decoherence," PRD **72**, (2005) 043515. arXiv:hep-th/0504134.

A. Albrecht & P. J. Steinhardt, "Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking", PRD **48**, (1982) 1220-1223.

A. Albrecht & P. J. Steinhardt, "Reheating an Inflationary Universe", PRL **48**, (1982) 1437-1440.

R. A. Alpher, H. Bethe & G. Gamow"The Origin of Chemical Elements", PRD **73**, (1948) 803-804.

R. Alpher & R. Herman, "Evolution of the Universe", Nature **162**, (1948), 774-775.

N. Arkani-Hamed, P. Creminelli, S. Mukohyama & M. Zaldarriaga, "Ghost inflation," JCAP **0404**, (2004) 001. arXiv:hep-th/0312100.

R. Arnowitt, S. Deser & C. W. Misner, "Gravitation: an introduction to current research", Wiley, New York, 1962. Reprint: arXiv:gr-qc/0405109.

F. Arroja & K. Koyama, "Non-Gaussianity from the trispectrum in general single field inflation," PRD **77**, (2008) 083517. arXiv:hep-th/08021167.

F. Arroja, S. Mizuno, K. Koyama & T. Tanaka, "On the Full Trispectrum in Single Field DBI-inflation," PRD **80**, (2009) 043527. arXiv:hep-th/09053641.

A. Ashoorioon & G. Shiu, "A Note on Calm Excited States of Inflation," JCAP **1103**, (2011) 025.

H. Assadullahi, J. Valiviita & D. Wands, "Primordial non-Gaussianity from Two Curvaton Decays," PRD **76**, (2007) 103003. arXiv:hep-th/07080223.

D. Babich, P. Creminelli & M. Zaldarriaga, "The Shape of Non-Gaussianities," JCAP **0408**, (2004) 009. arXiv:astro-ph/0405356.

D. Babich & M. Zaldarriaga, "Primordial Bispectrum Information from CMB Polarization," PRD **70**, (2004) 083005. arXiv:astro-ph/0408455.

J. Bardeen, P. Steinhardt & M. S. Turner, "Spontaneous Creation of Almost Scale-Free Density Perturbations in an Inflationary Universe", PRD **28**, (1983) 679-693.

N. Barnaby & J. M. Cline, "NonGaussianity from Tachyonic Preheating in Hybrid Inflation," PRD **75**, (2007) 086004. arXiv:astro-ph/0611750.

N. Barnaby & J. M. Cline, "Large Non-Gaussianity from Nonlocal Inflation," JCAP **0707**, (2007) 017. [arXiv:hep-th/07043426.

N. Bartolo, S. Matarrese & A. Riotto, "On Non-Gaussianity in the Curvaton Scenario," PRD **69**, (2004) 043503. arXiv:hep-ph/0309033.

N. Bartolo, E. Komatsu, S. Matarrese & A. Riotto, "Non-Gaussianity from Inflation: Theory and Observations," Phys. Rept. **402**, (2004) 103. arXiv:astro-ph/0406398.

N. Bartolo, S. Matarrese & A. Riotto, "Primordial Non-Gaussianity from Different Cosmological Scenarios," Nucl. Phys. Proc. Suppl. **148**, (2005) 56.

N. Bartolo, S. Matarrese & A. Riotto, "Non-Gaussianity and the Cosmic Microwave Background Anisotropies," arXiv:astro-ph/10013957.

T. Battefeld & R. Easther, "Non-Gaussianities in multi-field inflation," JCAP **0703**, (2007) 020. arXiv:astro-ph/0610296.

D. Battefeld & T. Battefeld, "Non-Gaussianities in N-flation," JCAP **0705**, (2007) 012. arXiv:hep-th/0703012.

R. Bean, X. Chen, G. Hailu, S. H. Tye & J. Xu, "Duality Cascade in Brane Inflation," JCAP **0803**, (2008) 026. [arXiv:hep-th/08020491.

F. Bernardeau & T. Brunier, "Non-Gaussianities in Extended D-term Inflation," PRD **76**, (2007) 043526. arXiv:hep-th/07052501.

F. Bernardeau, T. Brunier & J. P. Uzan, "Models of Inflation with Primordial Non-Gaussianities," AIP Conf. Proc. **861**, (2006) 821. arXiv:astro-ph/0604200.

D. L. Block, "A Hubble Eclipse: Lemaĩtre and Censorship," arXiv:physics.hist-ph/11063928.

J. R. Bond & D. S. Salopek, "Nonlinear evolution of long-wavelength metric fluctuations in inflationary models," PRD **42**, (1990) 3936.

R. H. Brandenberger, "Theory of Cosmological Perturbations and Applications to Superstring Cosmology," arXiv:hep-th/0501033.

. M. L. Brown *et al.* [QUaD collaboration], "Improved Measurements of the Temperature and Polarization of the Cmb from Quad," Astrophys. J. **705**, (2009) 978.

E. I. Buchbinder, J. Khoury & B. A. Ovrut, "Non-Gaussianities in New Ekpyrotic Cosmology," PRL **100**, (2008) 171302. arXiv:hep-th/07105172.

T. S. Bunch & P. C. W. Davies, "Quantum Field Theory In De Sitter Space: Renormalization By Point Splitting," Proc. Roy. Soc. Lond. A **360**, (1978) 117.

C. P. Burgess, J. M. Cline, F. Lemieux & R. Holman, "Are Inflationary Predictions Sensitive to Very High Energy Physics?," JHEP **0302**, (2003) 048. arXiv:hep-th/0210233.

C. Dvorkin and W. Hu, "Complete WMAP Constraints on Bandlimited Inflationary Features," arXiv:astro-ph/11064016.

C. T. Byrnes, M. Sasaki & D. Wands, "The Primordial Trispectrum from Inflation," PRD **74**, (2006) 123519. arXiv:astro-ph/0611075.

P. Cabella *et al.*, "Foreground Influence on Primordial Non-Gaussianity Estimates: Needlet Analysis of WMAP 5-Year Data," MNRAS **405**, (2010) 961-968. arXiv:astro-ph/09104362.

R. G. Cai, B. Hu & H. B. Zhang, "Acoustic Signatures in the Cosmic Microwave Background Bispectrum from Primordial Magnetic Fields," JCAP **1008**, (2010) 025.

S. Carroll & M. Trodden, "TASI Lectures: Introduction to Cosmology," arXiv:astro-ph/0401547.

S. H. Chen & J. B. Dent, "A New Approach to the Vacuum of Inflationary Models," arXiv:astro-ph/10124811.

X. Chen, "Multithroat Brane Inflation," PRD **71**, (2005) 063506. arXiv:hep-th/0408084.

X. Chen, "Inflation From Warped Space," JHEP **0508**, (2005) 045. arXiv:hep-th/0501184.

X. Chen, "Running Non-Gaussianities in DBI Inflation," PRD **72**, (2005) 123518. arXiv:astro-ph/0507053.

X. Chen, M. x. Huang, S. Kachru & G. Shiu, "Observational Signatures and Non-Gaussianities of General Single Field Inflation," JCAP **0701**, (2007) 002. arXiv:hep-th/0605045.

X. Chen, R. Easther & E. A. Lim, "Large Non-Gaussianities in Single Field Inflation," JCAP **0706**, (2007) 023. arXiv:astro-ph/0611645.

X. Chen, R. Easther & E. A. Lim, "Generation and Characterization of Large Non-Gaussianities in Single Field Inflation," JCAP **0804**, (2008) 010. arXiv:astro-ph/08013295.

X. Chen, B. Hu, M. x. Huang, G. Shiu & Y. Wang, "Large Primordial Trispectra in General Single Field Inflation," JCAP **0908**, (2009) 008. arXiv:astro-ph/09053494

X. Chen, "Primordial Non-Gaussianities from Inflation Models," Adv. Astron. **2010**, (2010) 638979. arXiv:astro-ph/10021416

X. Chen, "Folded Resonant Non-Gaussianity in General Single Field Inflation," JCAP **1012**, (2010) 003.

X. Chen, "Primordial Features as Evidence for Inflation," arXiv:hep-th/11041323.

A. Cooray, "21-cm Background Anisotropies Can Discern Primordial Non-Gaussianity from Slow-Roll Inflation," PRL **97**, (2006) 261301. arXiv:astro-ph/0610257.

A. Cooray, C. Li & A. Melchiorri, "The Trispectrum of 21-cm Background Anisotropies as a Probe of Primordial Non-Gaussianity," PRD **77**, (2008) 103506. arXiv:astr-ph/08013463.

H. R. S. Cogollo, Y. Rodriguez & C. A. Valenzuela-Toledo, "On the Issue of the ζ Series Convergence and Loop Corrections in the Generation of Observable Primordial Non-Gaussianity in Slow-Roll Inflation. Part I: the Bispectrum," JCAP **0808**, (2008) 029. arXiv:astro-ph/08061546.

H. Collins & R. Holman, "Trans-Planckian Enhancements of the Primordial Non-Gaussianities," PRD **80**, (2009) 043524. arXiv:hep-ph/09054925.

L. Covi, J. Hamann, A. Melchiorri, A. Slosar & I. Sorbera, "Inflation and WMAP Three Year Data: Features have a Future!," PRD **74**, (2006) 083509. arXiv:astro-ph/0606452.

P. Creminelli, "On non-Gaussianities in Single-Field Inflation," JCAP **0310**, (2003) 003. arXiv:astro-ph/0306122.

P. Creminelli & M. Zaldarriaga, "CMB 3-point Functions Generated by Nonlinearities at Recombination," PRD **70**, (2004) 083532. arXiv:astro-ph/0405428.

P. Creminelli, A. Nicolis, L. Senatore, M. Tegmark & M. Zaldarriaga, "Limits on Non-Gaussianities From WMAP Data," JCAP **0605**, (2006) 004. arXiv:astro-ph/0509029.

P. Creminelli, L. Senatore, M. Zaldarriaga & M. Tegmark, "Limits on $f_{\rm NL}$ Parameters from WMAP 3yr Data," JCAP **0703**, (2007) 005. arXiv:astro-ph/0610600.

P. Creminelli, L. Senatore & M. Zaldarriaga, "Estimators for Local Non-Gaussianities," JCAP **0703**, (2007) 019. arXiv:astro-ph/0606001.

A. Curto, E. Martínez-Gonzalez, P. Mukherjee, R. B. Barreiro, F. K. Hansen, M. Liguori & S. Matarrese, "WMAP 5-Year Constraints on $f_{\rm NL}$ with Wavelets," MNRAS **393**, (2009) 615. arXiv:astro-ph/08070231.

F. Y. Cyr-Racine and F. Schmidt, "Oscillating Bispectra and Galaxy Clustering: a Novel Probe of Inflationary Physics with Large-Scale Structure," arXiv:astro-ph/11062806.

U. H. Danielsson, "A note on Inflation and Transplanckian Physics," PRD **66**, (2002) 023511. arXiv:hep-th/0203198.

P. C. .W. Davies, "Scalar Production in Schwarzschild and Rindler Metrics," J. Phys. A: Math. Gen. **8**, (1975).

S. Dodelson, "Modern Cosmology," Academic Press, 2003.

J. Dunkley *et al.*, "The Atacama Cosmology Telescope: Cosmological Parameters from the 2008 Power Spectra," arXiv:astro-ph/10090866.

R. Durrer, "The Cosmic Microwave Background," Cambridge University Press 2008.

R. Easther, B. R. Greene, W. H. Kinney & G. Shiu, "Inflation as a Probe of Short Distance Physics," PRD **64**, (2001) 103502. arXiv:hep-th/0104102.

R. Easther, B. R. Greene, W. H. Kinney & G. Shiu, "A Generic Estimate of Trans-Planckian Modifications to the Primordial Power Spectrum in Inflation," PRD **66**, (2002) 023518. arXiv:hep-th/0204129.

R. Easther, B. R. Greene, W. H. Kinney & G. Shiu, "Imprints of Short Distance Physics on Inflationary Cosmology," PRD **67**, (2003) 063508. arXiv:hep-th/0110226.

R. Easther, W. H. Kinney & H. Peiris, "Observing Trans-Planckian Signatures in the Cosmic Microwave Background," JCAP **0505**, (2005) 009. arXiv:astro-ph/0412613.

R. Easther, W. H. Kinney & H. Peiris, "Boundary Effective Field Theory and Trans-Planckian Perturbations: Astrophysical Implications," JCAP **0508**, (2005) 001. arXiv:astro-ph/0505426.

A. Einstein, "The Field Equations of Gravitation", Sitzungsber. Preuss. Akad. Wiss. Berline (Math. Phys.) (1915), 844-847.

A. Einstein, "The Foundation of the General Theory of Relativity", Annalaen Phys. **49**, (1916) 769-822.

A. Einstein, "Cosmological Considerations in the General Theory of Relativity", Sitzungber. Preuss. Akad. Wiss. Berlin (Math. Phys.) (1917), 142-152.

K. Enqvist & A. Vaihkonen, "Non-Gaussian Perturbations in Hybrid Inflation," JCAP **0409**, (2004) 006. arXiv:hep-ph/0405103.

T. Falk, R. Rangarajan & M. Srednicki, "The Angular Dependence of the Three-Point Correlation Function of the Cosmic Microwave Background Radiation as Predicted by Inflationary Cosmologies," Astrophys. J. **403**, (1993) L1. arXiv:astro-ph/9208001.

J. R. Fergusson & E. P. S. Shellard, "Primordial Non-Gaussianity and the CMB Bispectrum," PRD **76**, (2007) 083523. arXiv:astro-ph/0612713.

J. R. Fergusson & E. P. S. Shellard, "The Shape of Primordial Non-Gaussianity and the CMB Bispectrum," PRD **80**, (2009) 043510. arXiv:astro-ph/0812.3413.

J. R. Fergusson, M. Liguori & E. P. S. Shellard, "General CMB and Primordial Bispectrum Estimation I: Mode Expansion, Map-Making and Measures of $F_{\rm NL}$," PRD **82**, (2010) 023502.

J. R. Fergusson, M. Liguori & E. P. S. Shellard, "The CMB Bispectrum," arXiv:astro-ph/10061642

R. Flauger, L. McAllister, E. Pajer, A. Westphal & G. Xu, "Oscillations in the CMB from Axion Monodromy Inflation," JCAP **1006**, (2010) 009. arXiv:hep-th/09072916

R. Flauger & E. Pajer, "Resonant Non-Gaussianity," JCAP **1101**, (2011) 017. arXiv:hep-th/10020833.

A. Friedmann,"Uber die Krummung des Raumes," Z. fur Phys. **10**, (1922) 377-386. Reprint: "On the Curvature of Space," General Rel. and Gr. **31**, (1999) 1991-2000.

A. Friedmann, "Uber die Moglichkeit einer Welt mit konstanter negativer Krummung des Raumes," Z. fur Phys. **31**, (1924) 326-332. Reprint: "On the Possibility of a World with Constant Negative Curvature of Space," General Rel. and Gr. **31**, (1999) 2000-2008.

S. A. Fulling, "Nonuniqueness of Canonical Field Quantization in Riemannian Space-Time," PRD **7**, (1973) 28502862.

G. Gamow,"My World Line: An Informal Autobiography," Viking Adult, 1970.

G. Gamow, "The Evolution of the Universe," Nature, **162**, (1948) 680-682.

A. Gangui, F. Lucchin, S. Matarrese & S. Mollerach, "The Three Point Correlation Function Of The Cosmic Microwave Background In Inflationary Models," Astrophys. J. **430**, (1994) 447. arXiv:astro-ph/9312033.

A. Gangui, J. Martin & M. Sakellariadou, "Single Field Inflation and Non-Gaussianity," PRD **66**, (2002) 083502. arXiv:astro-ph/0205202.

X. Gao, M. Li & C. Lin, "Primordial Non-Gaussianities from the Trispectra in Multiple Field Inflationary Models," JCAP **0911**, (2009) 007. arXiv:astro-ph09061345.

X. Gao & B. Hu, "Primordial Trispectrum from Entropy Perturbations in Multifield DBI Model," JCAP **0908**, (2009) 012. arXiv:astro-ph/09031920.

A. Gelman and D. Rubin, 'General Methods for Monitoring Convergence of Iterative Simulations," Statistical Science **7**, (1992) 452.

B. Greene, K. Schalm, J. P. van der Schaar & G. Shiu, "Extracting New Physics From the CMB," *In the Proceedings of 22nd Texas Symposium on Relativistic Astrophysics at Stanford University, Stanford, California, 13-17 Dec 2004, pp 0001* arXiv:astro-ph/0503458.

B. R. Greene, K. Schalm, G. Shiu & J. P. van der Schaar, "Decoupling in an Expanding Universe: Backreaction Barely Constrains Short Distance Effects in the CMB," JCAP **0502**, (2005) 001. arXiv:hep-th/0411217.

N. E. Groeneboom & O. Elgaroy, "Detection of Transplanckian Effects in the Cosmic Microwave Background," PRD **77**, (2008) 043522. arXiv:astro-ph/07111793.

A. Guth, "The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems," PRD **23**, (1981) 347-356.

A. H. Guth & S. Y. Pi, "Fluctuations in the New Inflationary Universe," PRL **49**, (1982) 1110-1113.

E. Hubble, "A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae," Proceedings of the National Academy of Sciences of the United States of America, **15**, (1929), 168-173.

J. Hamann, L. Covi, A. Melchiorri & A. Slosar, "New Constraints on Oscillations in the Primordial Spectrum of Inflationary Perturbations," PRD **76**, (2007) 023503. arXiv:astro-ph/0701380.

J. Hamann, S. Hannestad, M. S. Sloth & Y. Y. Y. Wong, "Observing Trans-Planckian Ripples in the Primordial Power Spectrum with future Large Scale Structure Probes," JCAP **0809**, (2008) 015. arXiv:astro-ph/08074528.

S. Hanany *et al.*, "Maxima-1: a Measurement of the Cosmic Microwave Background Anisotropy on Angular Scales of 10 Arcminutes to 5 Degrees," Astrophys. J. **545** (2000) L5. arXiv:astro-ph/0005123.

S. Hannestad, T. Haugboelle, P. R. Jarnhus & M. S. Sloth, "Non-Gaussianity from Axion Monodromy Inflation," arXiv:hep-th/0912.3527.

E. R. Harrison, "Fluctuations at the Threshold of Classical Cosmology," PRD **1**, (1970) 2726-2730.

S. Hawking, "Black Hole Explosions?," Nature **248**, (1974) 30-31.

S. Hawking, "The Development of Irregularities in a Single Bubble Inflationary Universe," Phys. Lett. **115b**, (1982) 295-297.

S. Hawking, "Bubble Collisions in the very Early Universe," Phys. Rev. **26**, (1982) 2681-2693.

C. Hikage, E. Komatsu & T. Matsubara, "Primordial Non-Gaussianity and Analytical Formula for Minkowski Functionals Of the Cosmic Microwave Background and Large-Scale Structure," Astrophys. J. **653** (2006) 11. arXiv:astro-ph/0607284.

C. Hikage, T. Matsubara, P. Coles, M. Liguori, F. K. Hansen & S. Matarrese, "Limits on Primordial Non-Gaussianity from Minkowski Functionals of the WMAP Temperature Anisotropies," MNRAS **389** (2008) 1439.

G. Hinshaw *et al.* [WMAP Collaboration], "Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Data Processing, Sky Maps, & Basic Results," Astrophys. J. Suppl. **180**, (2009) 225. arXiv:astro-ph/08030732

M. Hobson, A. Jones & A. Lasenby, "Wavelet Analysis and the Detection of Non-Gaussianity in the CMB," arXiv:astro-ph/9810200.

R. Holman & A. J. Tolley, "Enhanced Non-Gaussianity from Excited Initial States", JCAP **0805**, (2008) 001. arXiv:hep-th07101302.

G. 't Hooft, "Magnetic monopoles in Unified Gauge Theories," Nucl. Phys. B **79**, (1974) 276-284.

W. Hu & N. Sugiyama, "Anisotropies in the Cosmic Microwave Background: an Analytic Approach," Astrophys. J. **444** (1995) 489. arXiv:astro-ph/9407093.

W. Hu, "Angular Trispectrum of the Cosmic Microwave Background," PRD **64** (2001) 083005. arXiv:astro-ph/0105117.

M. x. Huang & G. Shiu, "The Inflationary Trispectrum for Models with Large Non-Gaussianities," PRD **74**, (2006) 121301. arXiv:hep-th/0610235.

M. x. Huang, G. Shiu & B. Underwood, "Multifield DBI Inflation and Non-Gaussianities," PRD **77**, (2008) 023511. arXiv:hep-th07093299.

Q. G. Huang, "N-vaton," JCAP **0809**, (2008) 017. arXiv:hep-th08071567.

M. G. Jackson & K. Schalm, "Model Independent Signatures of New Physics in the Inflationary Power Spectrum," arXiv:hep-th10070185.

M. G. Jackson & K. Schalm, "Model-Independent Signatures of New Physics in Slow-Roll Inflation," arXiv:hep-th11040887.

S. A. Kim & A. R. Liddle, "Nflation: Multifield Inflationary Dynamics and Perturbations," PRD **74**, (2006) 023513. arXiv:astro-ph/0605604.

W. H. Kinney & K. Tzirakis, "Quantum Modes in DBI Inflation: Exact Solutions and Constraints from Vacuum Selection," PRD **77** (2008) 103517. arXiv:astro-ph/07122043.

J. Khoury & F. Piazza, "Rapidly-Varying Speed of Sound, Scale Invariance and Non-Gaussian Signatures," JCAP **0907** (2009) 026. arXiv:hep-th/08113633.

L. A. Kofman, "Primordial Perturbations from Inflation," Phys. Scripta **T36** (1991) 108.

E. Komatsu & D. N. Spergel, "Acoustic Signatures in the Primary Microwave Background Bispectrum," PRD **63**, (2001) 063002. arXiv:astro-ph/0005036.

E. Komatsu, "The Pursuit of Non-Gaussian Fluctuations in the Cosmic Microwave Background," arXiv:astro-ph/0206039.

E. Komatsu *et al.* [WMAP Collaboration], "First Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Tests Of Gaussianity," Astrophys. J. Suppl. **148** (2003) 119. arXiv:astro-ph/0302223.

E. Komatsu, D. N. Spergel & B. D. Wandelt, "Measuring Primordial Non-Gaussianity in the Cosmic Microwave Background," Astrophys. J. **634**, (2005) 14. arXiv:astro-ph/0305189.

N. Kogo & E. Komatsu, "Angular Trispectrum of CMB Temperature Anisotropy from Primordial Non-Gaussianity with the Full Radiation Transfer Function," PRD **73**, (2006) 083007. arXiv:astro-ph/0602099.

E. Komatsu *et al.* [WMAP Collaboration], "Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation," Astrophys. J. Suppl. **180**, (2009) 330. arXiv:astro-ph/08030547.

E. Komatsu *et al.*, "Non-Gaussianity as a Probe of the Physics of the Primordial Universe and the Astrophysics of the Low Redshift Universe," white paper arXiv:astro-ph/09024759].

E. Komatsu, "Hunting for Primordial Non-Gaussianity in the Cosmic Microwave Background," Class. Quant. Grav. **27** (2010) 124010. arXiv:astro-ph/10036097.

E. Komatsu *et al.* [WMAP Collaboration], "Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation," Astrophys. J. Suppl. **192** (2011) 18. arXiv:astro-ph/10014538.

K. Koyama, S. Mizuno, F. Vernizzi & D. Wands, "Non-Gaussianities from Ekpyrotic Collapse with Multiple Fields," JCAP **0711**, (2007) 024. arXiv:hep-th/07084321.

D. Langlois, S. Renaux-Petel, D. A. Steer & T. Tanaka, "Primordial Perturbations and Non-Gaussianities in DBI and General Multi-Field Inflation," PRD **78**, (2008) 063523. arXiv:hep-th/08060336.

D. Larson *et al.*, "Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Power Spectra and WMAP-Derived Parameters," Astrophys. J. Suppl. **192**, (2011) 16. arXiv:astro-ph/10014635.

J. L. Lehners & P. J. Steinhardt, "Non-Gaussian Density Fluctuations from Entropically Generated Curvature Perturbations in Ekpyrotic Models," PRD **77**, (2008) 063533. arXiv:hep-th/07123779.

J. L. Lehners & S. Renaux-Petel, "Multifield Cosmological Perturbations at Third Order and the Ekpyrotic Trispectrum," PhRD **80**, (2009) 063503. arXiv:hep-th/09060530.

A. G. Lemaître, "Un Universe Homogene de Masse Constante et de Rayon Croissant Rendant Compte de la Vitesse Radiale des Nëbuleuses Extra-Galactiques," Annales de la Societe

Scietifique de Bruxelles **47**, (1927) 4959. Translation: "A Homogeneous Universe of Constant Mass and Increasing Radius accounting for Radial Velocity of Extra-galactic Nebulae," MNRAS **91**, (1931) 483-490.

A. Lewis, A. Challinor & A. Lasenby, "Efficient Computation of CMB anisotropies in Closed FRW Models," Astrophys. J. **538**, (2000) 473. arXiv:astro-ph/9911177.

A. R. Liddle & D. H. Lyth, "Cosmological Inflation and Large Scale Structure," Cambridge University Press 2000.

M. Liguori, F. K. Hansen, E. Komatsu, S. Matarrese & A. Riotto, "Testing Primordial Non-Gaussianity in CMB Anisotropies," PRD **73**, (2006) 043505. arXiv:astro-ph/0509098.

M. Liguori, A. Yadav, F. K. Hansen, E. Komatsu, S. Matarrese & B. Wandelt, "Temperature and Polarization CMB Maps from Primordial non-Gaussianities of the Local Type," PRD **76**, (2007) 105016. [Erratum-ibid. D **77**, (2008) 029902] arXiv:astro-ph/07083786.

A. Linde, "A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems," Phys. Lett. B **108b**, (1982), 389-393.

M. LoVerde, A. Miller, S. Shandera and L. Verde, "Effects of Scale-Dependent Non-Gaussianity on Cosmological Structures," JCAP **0804**, (2008) 014. arXiv:astro-ph/07114126.

D. H. Lyth, C. Ungarelli & D. Wands, "The Primordial Density Perturbation in the Curvaton Scenario," PRD **67**, (2003) 023503. arXiv:astro-ph/0208055.

D. H. Lyth & Y. Rodriguez, "The Inflationary Prediction for Primordial Non-Gaussianity," PRL **95**, (2005) 121302. arXiv:astro-ph/0504045.

D. H. Lyth & Y. Rodriguez, "Non-Gaussianity from the Second-Order Cosmological Perturbation," PRD **71**, (2005) 123508. arXiv:astro-ph/0502578.

J. M. Maldacena, "Non-Gaussian Features of Primordial Fluctuations in Single Field Inflationary Models," JHEP **0305**, (2003) 013. arXiv:astro-ph/0210603.

K. A. Malik & D. H. Lyth, "A Numerical Study of Non-Gaussianity in the Curvaton Ccenario," JCAP **0609**, (2006) 008. arXiv:astro-ph/0604387.

J. Martin, A. Riazuelo & M. Sakellariadou, "Non-Vacuum Initial States for Cosmological Perturbations of Quantum-Mechanical Origin," PRD **61**, (2000) 083518. arXiv:astro-ph/9904167.

J. Martin & R. H. Brandenberger, "The Trans-Planckian Problem of Inflationary Cosmology," PRD **63**, (2001) 123501. arXiv:hep-th/0005209.

J. Martin & C. Ringeval, "Superimposed Oscillations in the WMAP Data?," PRD **69**, (2004) 083515. arXiv:astro-ph/0310382.

J. Martin & C. Ringeval, "Exploring the Superimposed Oscillations Parameter Space," JCAP **0501**, (2005) 007. arXiv:hep-ph/0405249.

J. C. Mather *et al.* , "Early Results from the Cosmic Background Explorer (Cobe),"

J. C. Mather *et al.* , "A Preliminary Measurement of the Cosmic Microwave Background Spectrum by The Cosmic Background Explorer (Cobe) Satellite," Astrophys. J. **354**, (1990) L37.

P. D. Mauskopf *et al.* [Boomerang Collaboration], "Measurement of a Peak in the Cosmic Microwave Background Power Spectrum From the North American Test Flight of Boomerang," Astrophys. J. **536**, (2000) L59. arXiv:astro-ph/9911444.

P. McFadden & K. Skenderis, "Holography for Cosmology," PRD **81**, (2010) 021301.

P. McFadden & K. Skenderis, "Holographic Non-Gaussianity," JCAP **1105**, (2011) 013.

P. D. Meerburg, J. P. van der Schaar & P. S. Corasaniti, "Signatures of Initial State Modifications on Bispectrum Statistics," JCAP **0905**, (2009) 018. arXiv:hep-th/09014044.

P. D. Meerburg, J. P. van der Schaar & M. G. Jackson, "Bispectrum Signatures of a Modified Vacuum in Single Field Inflation With a Small Speed of Sound," JCAP **1002** (2010) 001. arXiv:hep-th/09104986.

P. D. Meerburg, "Oscillations in the Primordial Bispectrum: Mode Expansion," PRD **82**, (2010) 063517. arXiv:astro-ph/10062771.

P. D. Meerburg and J. P. van der Schaar, "Minimal Cut-Off Vacuum State Constraints from CMB Bispectrum Statistics," PRD **83** (2011) 043520. arXiv:hep-th/10095660.

A. Melchiorri *et al.* [Boomerang Collaboration], "A Measurement of Omega from the North American Test Flight of Boomerang," Astrophys. J. **536**, (2000) L63. arXiv:astro-ph/9911445.

I. G. Moss & C. Xiong, "Non-Gaussianity in Fluctuations from Warm Inflation," JCAP **0704**, (2007) 007. arXiv:astro-ph/0701302.

I. G. Moss & C. M. Graham, "Testing Models of Inflation with CMB Non-Gaussianity," JCAP **0711**, (2007) 004. arXiv:astro-ph/07071647.

V. F. Mukhanov & G. V. Chibisov, "Quantum Fluctuation and Nonsingular Universe," JETP Lett. **33**, (1981) 532535.

V. F. Mukhanov, "Physical Foundations of Cosmology," Cambridge University Press 2005.

D. Parkinson & A. R. Liddle, "Application of Bayesian Model Averaging to Measurements of the Primordial Power Spectrum," PRD **82**, (2010) 103533. arXiv:astr-ph/10091394.

C. Pahud, M. Kamionkowski & A. R. Liddle, "Oscillations in the Inflaton Potential?," PhRD **79** (2009) 083503. arXiv:astro-ph/08070322.

P. J. .E. Peebles & J. T. Yu, "Primeval Adiabatic Perturbation in an Expanding Universe", Astrophys. J. **162**, (1970) 815.

H. V. Peiris *et al.*, "First year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Inflation," Astrophys. J. Suppl. **148**, (2003) 213. arXiv:astro-ph/0302225.

A. A. Penzias & R. W. Wilson, "A Measurement of Excess Antenna Temperature at 4080 Mc/s," Astrophys. J. **142**, (1965) 419-421.

S. Perlmutter *et al.*, "A Study of 42 Type Ia Supernovae and a Resulting Measurement of Ω_m and Ω_Λ," Physics Reports **307**, (1998) 325-331.

A. Pillepich, C. Porciani & S. Matarrese, "The Bispectrum of Redshifted 21-cm Fluctuations from the Dark Ages," Astrophys. J. **662**, (2007) 1. arXiv:astro-ph/0611126.

M. Porrati, "Bounds on Generic High-Energy Physics Modifications to the Primordial Power Spectrum from Back-Reaction on the Metric," Phys. Lett. B **596**, (2004) 306. arXiv:hep-th/0402038.

M. Porrati, "Effective Field Theory Approach to Cosmological Initial Conditions: Self-Consistency Bounds and Non-Gaussianities," arXiv:hep-th/0409210.

J. Preskill,"Cosmological Production of Superheavy Magnetic Monopoles," PRL **43**, (1979) 1365-1368.

T. Pyne & S. M. Carroll, "Higher-Order Gravitational Perturbations of the Cosmic Microwave Background," PRD **53**, (1996) 2920. arXiv:astro-ph/9510041.

C. L. Reichardt *et al.*, "High Resolution CMB Power Spectrum from the Complete Acbar Data Set," Astrophys. J. **694** (2009) 1200. arXiv:astro-ph/0801.1491.

B. A. Reid *et al.* [SDSS Collaboration], "Baryon Acoustic Oscillations in the Sloan Digital Sky Survey Data Release 7 Galaxy Sample," MNRAS **401** (2010) 2148.

D. M. Regan, E. P. S. Shellard & J. R. Fergusson, "General CMB and Primordial Trispectrum Estimation," arXiv:astro-ph/10042915.

S. Renaux-Petel, "Combined Local and Equilateral Non-Gaussianities from Multifield DBI inflation," JCAP **0910**, (2009) 012. arXiv:hep-th/09072476.

R. Reyes, R. Mandelbaum, U. Seljak, T. Baldauf, J. E. Gunn, L. Lombriser and R. E. Smith, "Confirmation of General Relativity on Large Scales from Weak Lensing and Galaxy Velocities," Nature **464** (2010) 256. arXiv:astro-ph/10032185.

A. Riess et al, "Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant," Astrophys. J. **116**, (1998) 1009-1038.

A. Riess et al. "Discoveries of Type Ia Supernovae at $z \geq 1$: Narrowing Constraints on the Early Behavior of Dark Energy," Astrophys. J. **656**, (2007).

A. G. Riess *et al.*, "A Redetermination of the Hubble Constant with the Hubble Space Telescope from a Differential Distance Ladder," Astrophys. J. **699**, (2009) 539. arXiv:astro-ph/09050695.

A. Riotto & M. S. Sloth, "On Resumming Inflationary Perturbations beyond One-loop," JCAP **0804**, (2008) 030. arXiv:hep-th/08011845.

H. P. Robertson, "On the Foundations of Relativistic Cosmology," Proceedings of the National Academy of Science **15**, (1929) 822829.

V. Rubin *et al.*, "Rotational Properties of 21 Sc Galaxies with a Large Range of Luminosities and Radii from NGC 4605 (R=4kpc) to UGC 2885 (R=122kpc)", Astrophys. J.˙**238**, (1980).

O. Rudjord, F. K. Hansen, X. Lan, M. Liguori, D. Marinucci & S. Matarrese, "An Estimate of the Primordial Non-Gaussianity Parameter $f_{\rm NL}$ Using the Needlet Bispectrum from WMAP," Astrophys. J. **701** (2009) 369. arXiv:astro-ph/09013154.

M. Sasaki, J. Valiviita & D. Wands, "Non-Gaussianity of the Primordial Perturbation in the Curvaton Model," PRD **74**, (2006) 103003. arXiv:astro-ph/0607627.

K. Schalm, G. Shiu & J. P. van der Schaar, "Decoupling in an Expanding Universe: Boundary RG-flow Affects Initial Conditions for Inflation," JHEP **0404**, (2004) 076. arXiv:hep-th/0401164.

K. Schalm, G. Shiu & J. P. van der Schaar, "The Cosmological Vacuum Ambiguity, Effective Actions, and Transplanckian Effects in Inflation," AIP Conf. Proc. **743**, (2005) 362. arXiv:hep-th/0412288.

J. Schmalzing, M. Kerscher & T. Buchert, "Minkowski Functionals in Cosmology," arXiv:astro-ph/9508154.

D. Seery, "One-Loop Corrections to a Scalar Field During Inflation," JCAP **0711**, (2007) 025. arXiv:astro-ph/07073377.

D. Seery & J. E. Lidsey, "Primordial Non-Gaussianities in Single Field Inflation," JCAP **0506**, (2005) 003. arXiv:astro-ph/0503692.

D. Seery & J. E. Lidsey, "Non-Gaussianity from the Inflationary Trispectrum," JCAP **0701**, (2007) 008. arXiv:astro-ph/0611034.

D. Seery, M. S. Sloth & F. Vernizzi, "Inflationary Trispectrum from Graviton Exchange," JCAP **0903**, (2009) 018. arXiv:astro-ph/08113934

E. Sefusatti, M. Liguori, A. P. S. Yadav, M. G. Jackson & E. Pajer, "Constraining Running Non-Gaussianity," JCAP **0912**, (2009) 022. arXiv:astro-ph/09060232.

U. Seljak & M. Zaldarriaga, "A line of Sight Approach to Cosmic Microwave Background Anisotropies", Astrophys. J. **469**, (1996) 437.

L. Senatore, "Tilted ghost inflation," PRD **71**, (2005) 043512. arXiv:astro-ph/0406187.

L. Senatore, K. M. Smith & M. Zaldarriaga, "Non-Gaussianities in Single Field Inflation and their Optimal Limits from the WMAP 5-year Data," arXiv:astro-ph/0905.3746.

P. Serra & A. Cooray, "Impact of Secondary Non-Gaussianities on the Search for Primordial Non-Gaussianity with CMB Maps," arXiv:astro-ph/0801.3276.

E. Silverstein & D. Tong, "Scalar Speed Limits and Cosmology: Acceleration from D-cceleration," PRD **70**, (2004) 103505. arXiv:hep-th/0310221.

W. de Sitter, "Further Remarks on the Solutions of the Field-Equations of Einstein's Theory of Gravitation", KNAW Proceedings Series B Physical Sciences **20**, (1918) 1309-1312.

W. de Sitter, "On the Curvature of Space", KNAW Proceedings Series B Physical Sciences **20**, (1918).

J. Smidt, A. Amblard, A. Cooray, A. Heavens, D. Munshi & P. Serra, "A Measurement of Cubic-Order Primordial Non-Gaussianity (g_{NL} and τ_{NL}) With WMAP 5-Year Data," arXiv:astro-ph/10015026.

J. Smidt, A. Amblard, C. T. Byrnes, A. Cooray & D. Munshi, "Cmb Constraints on Primordial Non-Gaussianity from the Bispectrum ($f_{\rm NL}$) And Trispectrum ($g_{\rm NL}$ and$\tau_{\rm NL}$) and a New Consistency Test of Single-Field Inflation," PRD **81**, (2010) 123007. arXiv:10041409.

K. M. Smith & M. Zaldarriaga, "Algorithms for Bispectra: Forecasting, Optimal Analysis, and Simulation", arXiv:astro-ph/0612571.

K. M. Smith, L. Senatore & M. Zaldarriaga, "Optimal Limits on $f_{\rm NL}^{local}$ From WMAP 5-Year Data," JCAP **0909** (2009) 006. arXiv:astro-ph/0901.2572.

G. F. Smoot *et al.*, "Structure in the Cobe Differential Microwave Radiometer First Year Maps," Astrophys. J. **396** (1992) L1.

D. N. Spergel *et al.* [WMAP Collaboration], "Wilkinson Microwave Anisotropy Probe (WMAP) Three Year Results: Implications for Cosmology," Astrophys. J. Suppl. **170**, (2007) 377. arXiv:astro-ph/0603449.

D. N. Spergel *et al.* [WMAP Collaboration], "First Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters," Astrophys. J. Suppl. **148**, (2003) 175. arXiv:astro-ph/0302209.

A. A. Starobinsky,"Dynamics of Phase Transition in the New Inflationary Scenario and Generation of Perturbations", Phys. Lett, Vol. 117b, No. 3-4, (1982), **175-178**.

R. A. Sunyaev, "Fluctuations of the Microwave Background Radiation", The large scale structure of the universe; Proceedings of the Symposium, Tallin, Estonian SSR, September 12-16, 1977. (A79-13511 03-90) Dordrecht, D. Reidel Publishing Co., 1978, p. 393-402; Discussion, p. 402-404.

A. Taylor & P. Watts, "Parameter Information from Nonlinear Cosmological Fields," MNRAS **328**, (2001) 1027. arXiv:astro-ph/0010014.

W. G. Unruh, "Notes on Black-Hole Evaporation," PRD **14**, (1976) 870892.

L. Verde, R. Jimenez, M. Kamionkowski & S. Matarrese, "Tests for Primordial Non-Gaussianity," MNRAS **325**, (2001) 412. arXiv:astro-ph/0011180.

L. Verde *et al.* [WMAP Collaboration], "First Year Wilkinson Microwave Anisotropy Probe (Wmap) Observations: Parameter Estimation Methodology," Astrophys. J. Suppl. **148**, (2003) 195. arXiv:astro-ph/0302218.

L. Verde, "Non-Gaussianity from Large-Scale Structure Surveys," arXiv:astro-ph/10015217.

F. Vernizzi & D. Wands, "Non-Gaussianities in Two-Field Inflation," JCAP **0605**, (2006) 019. arXiv:astro-ph/0603799.

A. G. Walker, "Distance in an expanding universe", MNRAS **94**, (1933) 159.

L. M. Wang & M. Kamionkowski, "The Cosmic Microwave Background Bispectrum and Inflation," PRD **61**, (2000) 063504. arXiv:astro-ph/9907431.

S. Weinberg, "Quantum Contributions to Cosmological Correlations," PRD **72**, (2005) 043514. arXiv:hep-th/0506236.

S. Weinberg, "Quantum Contributions to Cosmological Correlations. II: Can these Corrections Become Large?," PRD **74**, (2006) 023508. arXiv:hep-th/0605244.

C. M. Will, "The Confrontation Between General Relativity and Experiment," Living Rev. Rel. **4**, (2001) 4. arXiv:gr-qc/0103036.

J. Q. Xia, C. Baccigalupi, S. Matarrese, L. Verde & M. Viel, "Constraints on Primordial Non-Gaussianity from Large Scale Structure Probes," arXiv:astro-ph/11045015.

J. Q. Xia, M. Viel, C. Baccigalupi, G. De Zotti, S. Matarrese & L. Verde, "Primordial Non-Gaussianity and the NRAO VLA Sky Survey," Astrophys. J. **717** (2010) L17. arXiv:astro-ph/10033451.

A. P. S. Yadav, E. Komatsu & B. D. Wandelt, "Fast Estimator of Primordial Non-Gaussianity from Temperature and Polarization Anisotropies in the Cosmic Microwave Background," Astrophys. J. **664**, (2007) 680. arXiv:astro-ph/0701921.

A. P. S. Yadav, E. Komatsu, B. D. Wandelt, M. Liguori, F. K. Hansen & S. Matarrese, "Fast Estimator of Primordial Non-Gaussianity from Temperature and Polarization Anisotropies in the Cosmic Microwave Background II: Partial Sky Coverage and Inhomogeneous Noise," arXiv:astro-ph/07114933.

A. P. S. Yadav & B. D. Wandelt, "Detection of Primordial Non-Gaussianity (f_{NL}) in the WMAP 3-Year Data at above 99.5 % Confidence," arXiv:astro-ph/07121148.

S. Yokoyama, T. Suyama & T. Tanaka, "Primordial Non-Gaussianity in Multi-Scalar Slow-Roll Inflation," PRD **77**, (2008) 083511. arXiv:astro-ph/07053178.

Y. B. Zeldovich, "A hypothesis, unifying the structure and the entropy of the Universe," MNRAS **160**, (1972).

Y. Zeldovich & M. Y. Khlapov, "On the Concentration of Relic Monopoles in the Universe," Phys. Lett. B, **79b**, (1978) 239-241.

F. Zwicky, "Die Rotverschiebung von Extragalaktischen Nebeln," Helvetica Physica Acta 6, 1933, **110-127**.

F. Zwicky, "On the Masses of Nebulae and of Clusters of Nebulae," Astrophys. J. **217**, (1937).

Acknowledgements

The law of causality has not the significance of a statement as to the world of experience, except when observable facts ultimately appear as causes and effects.
Albert Einstein

Save the best for last! Truth be told, I have been postponing this for some time, because I was afraid of forgetting someone. In fact, I have already sent my thesis in for proof printing, without the acknowledgements. So, let's hope they accept a late addition to the manuscript.

First of all, let me say that it is true "*that the older you get, the faster time seems to pass*". I remember my master thesis taking ages, and now, 4 years later, I am *already* wrapping up my next thesis. It does not mean the time that passes is less enjoyed. On the contrary, the last 4 years have been really good. Now that this is out of the way, and I have at least included one sentence of wisdom in this thesis, let me start thanking all of you who have made a contribution in one way or another to this thesis.

If it is numbers that count, and in today's world that certainly seems the case, I want to thank Jan Pieter van der Schaar first. Dividing equal time between me, teaching and your family is a great sacrifice. JP I can not thank you enough for that. I was introduced to you for the first time in the first year of my masters, as I was interested in working on problems related to string theory and cosmology. According to Erik Verlinde you were the right person to talk to. Admittedly, my contribution to string theory is negligible to non-existent, but I do believe that we both have learned a lot about cosmology through these 6 years of collaboration. And being a non-cosmology focused institute, I think that we have done remarkably well within the community, with our first paper recently reaching top-cite on spires. However, this would all be much less satisfying, if working with you was not such a pleasure. You are just a really nice guy. I really appreciate your open and super social character. If I would associate this quality to anything I would believe that it has something to do with your 'Friese' roots. I have yet to attend a social gathering where you are not able to maintain a conversation. I think this shows an honest and healthy interest in other people (or it is just because you do not like long silences?).

Ralph Wijers, although not as dominantly present as JP in terms of science, you are a great supervisor. Sometimes I wonder when you sleep, I can sent you a mail day and night

and always receive a reply (though admittedly less frequent since you became a director!). Especially during the last months, when I worked on a paper that is more related to your own line of work, you were always there to explain to me again and again how to make a contour plot. One thing you have in common with JP is your pleasant character and your genuine interest in so many things. You are pretty much always cheerful and open for a conversation on any topic. Your walks across the hallway always remind of someone who is out for a stroll in the park. Given the work load you have it is incredible that you find time for this and I am sure it is a quality that is appreciated by everyone in our institute, and contributes to making our institute (API) one of the nicest astronomy institutes across the globe, and quite possibly the Universe.

I would especially like to thank my two collaborators Mark Jackson and Pier-Stefano Corasaniti. Mark I hope you find your way in Paris. Your stories are truly one of a kind. PS thanks for having me over many times in Paris, and despite your crazy busy life, making time to show me Paris. I hope for many more such collaborations with you.

Although I haven't always been the most involved person, I do like API a lot and all the people that make our institute such a special place. I have had loads of fun with my former roommates Gerrit, Atish, Klaas, Alexander, Kostadinos and Peter. Some special thanks goes out to people who had no personal gain in helping me, but none the less did so, some of them many times over. These include Gerrit for his companionship , Kostadinos for random questions and discussions, Martin for even more random questions, Yuri for the effort you put in getting me up and running on the LISA cluster, Diego, Evert and Eduardo usually for Unix stuff or La Palma related issues, the secretaries Minou, Milena, Eva, Nicole for booking all my trips and arranging lots of stuff in general, and Lide for all the help with financial work and grant related problems. Special thanks also goes out to Bart for sitting next to me in a confined space, that was some serious sacrifice, Bryn for being his substitute, Rens for trusting me teaching an observation course, Alex for the help getting a grant to make my PhD become reality and helping me becoming a better speaker, same is true for Sara as well as allowing me to be part of her Cosmology course, Michael for helping me set up my cosmology software on a mac, Thijs for being the only one showing up at all the parties I invited you to, Lucas, Gijs and Mihkel for being inspiring characters by being very promising scientists, Nathalie for always showing interest and just being a very nice person, let's meet again in the States, the students, most of who now have become colleagues, we had fun in La Palma did we not: Rik, Tulio, Thijs, Peter, Ricardo, Daniela, Sjors, Frank, Bram, Olga, Arjen, Rasjied, Anita and Sabastiaan especially for being our substitute goalkeeper once, Ed van de Heuvel for being the first professor that told me about the stars and also being the first professor I share an office with (what an honor), Marieke and Jaap for letting me part of the Year of Astronomy team, Montse, Peter, Arjan, Ale, Rudy, Carsten, Anastasia, Lars, Theo, Salome, Samia, Dumai, Caroline, Ji-Ling, Godelieve, Huib, Joop, Mathili, Lex, Michiel, Jouri, Richard, Saskia, Anna, Ton, Oscar, Dave, Hagues, John, Ramanpreet, Koen, Jason, Maciej, Lianne, Dominik and Ken, for those who know who I am (that guy who is hardly ever around, for the last 4 years), thank you for being such pleasant colleagues.

The first two years of my PhD I also spent half of my time at the ITFA institute. Admittedly, API and ITFA then were literally and figuratively two worlds apart. I have had only few

encounters with the most of you there, none the less I appreciated the times that I did. Special thanks go out to Jan de Boer, Balt, Sheer, Paul, Johannes, Jesper, and Sander, Michele and Reinier now all at Nikhef. The Dutch cosmology crew: Koenraad, Jurjen, Marieke, Jan, Ana, Damien, Tomislav, Ted and others, thank your for all the interesting meetings and discussions. I hope that the future will bring more cosmology to the Netherlands and given the recent GRAPPA initiative, I am sure there is a bright future ahead. Unfortunately we lost one very appreciated colleague and dear friend Sjoerd. I will always remember the many times we worked together.

I would like to thank the guy working in the MyCom shop who helped me out fixing my computer the day it lost its life, only 24 hours before I had to hand in my thesis. Unfortunately it is still broken, but at least I recovered all my precious data. Thanks!

Dear family, as far as I know most of you have been around forever and I hope it won't be any different in the future. Krijn, Marijke, Daan, Jochem, Micha, Lieke, Reini, Koos, Ingeborg,Tiny, Ted, Karuna, Joos, Trinette, Tom, Lenny, Marc, Jonathan, Lawrence, David, Kati, Daniel, Lexi, Gideon, Anna, Benjamin and Anais, thank you for making such a cool, ridiculously extended family. It feels warm and comfortable. Loesje en Opa, I miss you both very much, and I am sad you are no longer around to share this moment with me. Thank you Jack for reading part of this manuscript and giving some helpful comments, and obviously for being such a good guy. Maartje, my extra blond sister, be sure to come and visit me as often as possible in the coming years. Brazil is only a stones throw away from the US :-). Katja and Jet, if I had anything to say about it, two half sisters count way more than just 1 sister!

Mum and dad let me start by saying that I care a lot for both of you. I count myself incredibly lucky, having such great parents.

Dickie, thank you for all the times we talked about science and things that relate to it even remotely.You should realize that I am scientist because of your inspiring stories when I was just a boy. Let us continue talking about the arrow of time in the future (and possibly in the past!). And thanks a lot for reading this manuscript one more time, you are always there when I need it. That's all a son can ask for.

Mama, because you are simply the best mother in the world. The mother of our family, with your deep devotion to us all. I would not know where I would be without it.

Esther, my dearest, we will be married by the time this is printed, and we are up for an incredible adventure, extending the already amazing journey we started two and a half years ago. You have been a friend, a lover, a companion, a travelmate, a discussion partner, and illustrator. Thank you for everything, it is impossible to do you right in these acknowledgements. So, instead I will try to show my appreciation by being around you as much as I possible can every day of my life.

List of Publications

Refereed

- P. D. Meerburg, J. P. van der Schaar and P. S. Corasaniti, "Signatures of Initial State Modifications on Bispectrum Statistics," JCAP **0905**, (2009) 018.
- P. D. Meerburg, J. P. van der Schaar and M. G. Jackson, "Bispectrum signatures of a modified vacuum in single field inflation with a small speed of sound," JCAP **1002**, (2010) 001.
- P. D. Meerburg, "Oscillations in the Primordial Bispectrum: Mode Expansion," Phys. Rev. D **82**, (2010) 063517.
- T. Bagnoli *et al.*, "An Inner Gaseous Disk around the Herbig Be Star MWC 147,' Astrophysical Journal Letters **724**, (2010) L5-L8.
- P. D. Meerburg and J. P. van der Schaar, "Minimal cut-off vacuum state constraints from CMB bispectrum statistics," Phys. Rev. D **83**, (2011) 043520.

Conference Proceedings

- P. D. Meerburg, "Oscillations in the Primordial Bispectrum," J. Phys. Conf. Ser. **259**, (2010) 012049.

Unpublished

- J. Aguirre *et al.*, "Observing the Evolution of the Universe," arXiv:0903.0902 [astro-ph.CO]. Science White Paper
- S. Dodelson, R. Easther, S. Hanany *et al.*, "The Origin of the Universe as Revealed Through the Polarization of the Cosmic Microwave Background," arXiv:0902.3796 [astro-ph.CO]. Science White Paper