

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/67091> holds various files of this Leiden University dissertation.

Author: Welling, Y.M.

Title: Spectroscopy of two-field Inflation

Issue Date: 2018-11-27

Spectroscopy of Two-Field Inflation

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 27 november 2018
klokke 13.45 uur

door

Yvette M. Welling

geboren te Almelo
in 1989

Promotor: Prof. dr A. Achúcarro
Co-promotor: Dr E. Pajer (Utrecht University)

Promotie commissie: Prof. dr D. D. Baumann (University of Amsterdam)
Dr G. Palma (University of Chile, Santiago, Chile)
Dr A. Westphal (DESY, Hamburg, Germany)
Prof. dr E. R. Eliel
Prof. dr K. Schalm

Casimir PhD series, Delft-Leiden 2018-36

ISBN 978-90-8593-366-3

An electronic version of this thesis can be found at <https://openaccess.leidenuniv.nl>

This work is supported by a Leiden de Sitter Fellowship that is funded by the Netherlands Organisation for Scientific Research (NWO).

The cover shows an inflationary flower on an effective background of large scale structure. Design by MdK and YW.

Voor papa

Inhoudsopgave

1	Introduction	3
1.1	A brief history of modern cosmology	4
1.2	Inflation	10
1.3	Outline thesis	28
1.4	Large scale structure as a probe of inflation	32
I	Multi-field Inflation	49
2	Orbital Inflation	51
2.1	Introduction	51
2.2	Kinematical analysis of multi-field inflation	53
2.3	Phenomenology of Orbital Inflation	61
2.4	Exact models of Orbital Inflation	66
2.5	Summary	70
3	Orbital Inflation with ultra-light fields	73
3.1	Introduction	74
3.2	A toy model with neutrally stable orbits	75
3.3	Ultra-light Orbital inflation	80
3.4	Stability	82
3.5	Inflation with massless isocurvature perturbations	85
3.6	Phenomenology	86
4	Universality of multi-field α-attractors	89
4.1	Introduction	90
4.2	α -attractors and their supergravity implementations	92
4.3	Dynamics of multi-field α -attractors	96
4.4	Universal predictions of α -attractors	102
4.5	Universality conditions for more general α	105
4.6	Summary and Conclusions	107

4.A	Constraints on the potential	109
4.B	Full analysis of perturbations	110
5	Heavy fields and a reduced tensor-to-scalar ratio	115
5.1	Introduction	115
5.2	General setup	117
5.3	Quadratic inflation	122
5.4	Linear inflation	124
5.5	Natural inflation	125
5.6	Conclusion	127
II	Large Scale Structure	129
6	Lifting primordial non-Gaussianity above the noise	131
6.1	Introduction and summary	132
6.2	Analytical predictions for the bispectrum	134
6.3	Fisher analysis	137
6.4	Results	147
6.5	Discussion and Outlook	158
6.A	Explicit Results for the Bispectrum	162
6.B	Theoretical noise	168
6.C	Choice of binning and volume of the bins	178
6.D	Table of parameters	183
	Bibliography	184
	Publications	209
	Samenvatting	211
	Curriculum Vitae	219
	Acknowledgements	221

Introduction

Cosmology is a fascinating field of research. The advance of theory and observations in the last centuries has enabled us to reformulate philosophical musings such as “Where do we come from?” as quantitatively testable scientific questions. It is impressive that we have been able to learn about the universe far in the past. This resulted in the Hot Big Bang model: an excellent description of the universe from the moment it was merely a few minutes old. Moreover, cosmological inflation has provided us with a very compelling model of the primordial universe. It predicts the creation of seeds of structure formation from quantum perturbations at that time. Therefore, we might still find an imprint of the primordial universe in the sky. This allows us to look back to perhaps¹ the first $10^{-30}s$, a mind-blowing thought! At that time the universe was extremely dense and tiny and thus dominated by ultra high energy particle physics. Therefore, inflation connects observations on cosmological scales to particle physics not accessible at earth-based particle accelerators. In other words, inflation provides the ultimate playground for a theoretical physicist.

In the main part of this thesis we aim to improve our understanding of the signatures of new physics at the time of inflation. Additionally, in the second part we perform a statistical analysis to understand whether we can extract some of the traces thereof from the 3D map of galaxies in the near future.

Throughout this thesis we work in Planck units $\hbar = c = k_B = 1$ and the reduced Planck mass is given by $M_p = (8\pi G)^{-1/2}$.

¹Please take the time-scale of inflation with a grain of salt, it is a model dependent number. It could as well be a factor of a million off, but nevertheless it is still extremely small.

1.1 A brief history of modern cosmology

Most of modern research in cosmology is based on the Standard Model of Cosmology. For my generation of cosmologists the so-called Λ CDM model is a natural starting point, but from a historical point of view it must have been quite some journey to get here. The road was bumpy and the passage required several radically new ideas. Due to the efforts of many great scientists, Λ CDM was established within a century. In this section I aim to give an idea of the historical development of Λ CDM. This section is mainly based on [1–6].

1.1.1 The hot Big Bang model

An expanding universe

Einstein’s General Relativity (GR) [7] marked the beginning of cosmology as a predictive science. The dynamical and geometrical properties of the universe are related to its constituents by the Einstein equations. One key property follows from the premise that there are no privileged positions or directions in the universe. More precisely, the *Cosmological Principle* states that the universe is statistically homogeneous and isotropic at large scales.

Einstein believed in a static universe, and in 1917 he introduced [8] the static cosmological model as a homogeneous and isotropic solution to GR. For that purpose he famously introduced the *cosmological constant*. It turned out that this solution is unstable [9]. Around the same time De Sitter [10] published another ‘static’ cosmological model, which in fact described an exponentially expanding universe (but this was only realized later by Lemaître.) It corresponds to a universe with no matter, but only a cosmological constant. He calculated that this would result in a redshift of distant sources and he realized that this could explain the observation of Slipher [11].

In the following decades substantial observational and theoretical progress was made. On the theory side, the most general cosmological solution of a homogeneous and isotropic universe was established: today we call it the Friedmann-Lemaître-Robertson-Walker (FLRW) metric [12–16]

$$ds^2 = -dt^2 + a(t)^2 \left(\frac{dr^2}{1 - \kappa r^2} + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \right) \quad (1.1)$$

Here the constant $\kappa \in \{-1, 0, 1\}$ denotes whether the curvature of the spatial slices is negatively curved, flat or positively curved, respectively. The *scale factor* $a(t)$ probes the *physical* distance between two points on the *comoving* spatial slice at the time t . It is a measure of the relative size of the universe

compared to today, i.e. $a(t_0) = 1$. The expansion rate of the universe is measured by the *Hubble parameter* $H(t) \equiv \frac{\dot{a}(t)}{a(t)}$. The stress-energy tensor corresponding to the FLRW spacetime is that of a homogeneous and isotropic fluid with pressure p and energy density ρ . The evolution of the universe, as parameterized by the Hubble parameter, can therefore be related to its matter content by the Friedmann equations

$$H^2 \equiv \frac{\dot{a}^2}{a^2} = \frac{\rho}{3M_p^2} - \frac{\kappa^2}{a^2}, \quad (1.2a)$$

$$\dot{\rho} + 3H(\rho + p) = 0. \quad (1.2b)$$

Only when $a(t)$ is constant, or $H = 0$, this resembles the static solution of Einstein. A constant Hubble parameter H , and therefore $a \sim e^{Ht}$, corresponds to the De Sitter solution.

Lemaître [14] realized that a non-static universe, i.e. $H(t) \neq 0$, leads to *cosmological redshift* of photons. Photons moving freely in an expanding spacetime lose their energy like $E \sim a^{-1}$. In other words their wavelength stretches, and spectra emitted by distant objects redshift. He understood that the ‘apparent velocity’ of galaxies as measured by Slipher [11] and Hubble [17] was in fact mainly caused by the expansion of spacetime². He also provided the linear velocity-distance relation

$$v = Hd. \quad (1.3)$$

This is now called *Hubble’s law*³, after [20].

Big Bang nucleosynthesis

The observation that the universe is expanding led Lemaître [21] to propose the tantalizing idea that the universe must have been much smaller and denser in the past. His idea was first met with contempt by his peers; Fred Hoyle even coined the term ‘Big Bang’ to make fun of Lemaître. However, in 1949 his idea was put on firmer ground when Gamow and collaborators [22] used his idea to predict the formation of light elements during the first three minutes. They

²An earlier interpretation came from his mentor Eddington [18], based on Slipher’s measurements of 36 redshifted spiral nebulae in the 1910s. He concluded that De Sitter’s solution provided an explanation: “there is the general displacement of spectral lines to the red in distant objects due to the slowing down of atomic vibration which would be erroneously interpreted as a motion of recession.” Moreover, Lemaître reanalysed the result of De Sitter in [19] for which he found the linear velocity-distance relation.

³Apparently, in the English translation of Lemaître’s article, the relevant lines about velocity-distance relation were modified [1].

had the crucial insight that the nuclear reactions could take place because the early universe was radiation dominated and very *hot*. This was the birth of the *Hot Big Bang* model, in which *Big Bang Nucleosynthesis* (BBN) plays a key role. However, it took the detection of the Cosmic Microwave Background radiation (see below), before the Hot Big Bang scenario became accepted by the scientific community. Nowadays, the relative abundances of light elements are well-tested and - except for the Lithium abundance - in good agreement with late time observations [23–25].

Cosmic Microwave Background radiation

Soon after, also in 1948, Alpher and Hermann [26] predicted that we should see an afterglow of the Hot Big Bang. At early times, radiation was in thermal equilibrium with matter. Only after 380.000 years, when the universe cooled down to temperatures $T \sim 3000K$ such that neutral hydrogen could be formed (recombination), the number of free electrons dropped enormously. Thomson scattering, coupling photons to free electrons, became inefficient and photons started to freestream (decoupling). These photons have been travelling since then, while being redshifted with the expansion of the universe. Alpher and Hermann predicted that today we are immersed in a bath of thermal photons of temperature $T \sim 5K$. Almost twenty years later Penzias and Wilson [27] discovered a mysterious isotropic antenna noise, which turned out to be the *Cosmic Microwave Background* radiation (CMB). In 1993 the COBE mission confirmed their measurement and showed that it follows a blackbody spectrum [28] with average temperature $T \approx 2.7K$. Moreover, the COBE team found [29] small variations in the temperature of order 10^{-5} , reflecting small density inhomogeneities crucial for structure formation in the late universe. Figure 1.1 shows the most recent map of the CMB temperature variations from the *Planck* collaboration [30]

1.1.2 Dark matter and dark energy

Accepting the Hot Big Bang picture of an expanding universe, the next step was to understand its constituents and its geometry. For that purpose it is useful to define the *critical density* for which the universe is flat ($\kappa = 0$)

$$\rho_c \equiv 3M_p^2 H^2 . \quad (1.4)$$

The ratio of the total energy density to the critical density, i.e. the *density parameter* $\Omega \equiv \frac{\rho}{\rho_c}$ is a measure of the geometry of the universe. If $\Omega = 1$ the universe is flat, if $\Omega < 1$ the universe is open ($\kappa = -1$) and if $\Omega > 1$ the universe is closed ($\kappa = 1$).

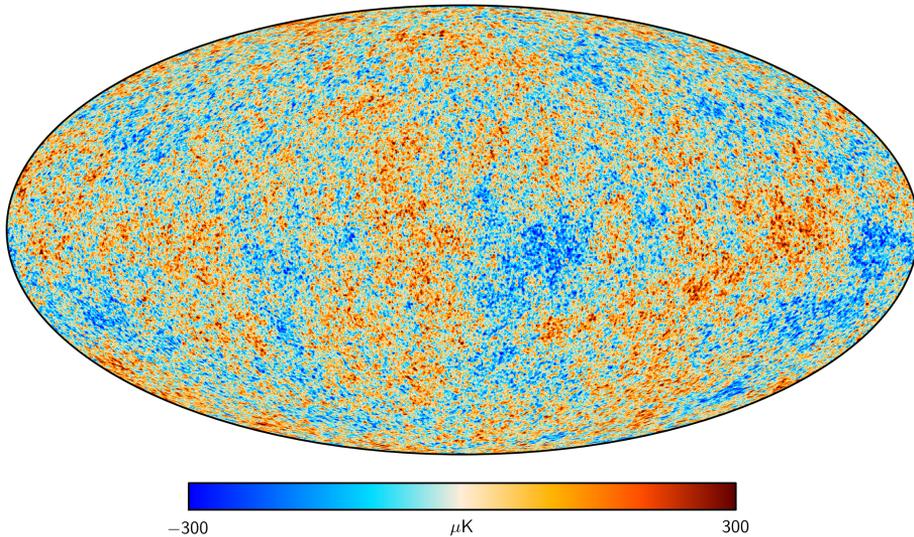


Figure 1.1: The CMB intensity map as measured by *Planck* [30]. The colours represent small temperature variations of order 10^{-5} with respect to the background $T \sim 2.73K$.

The BBN computations of Gamow and collaborators [22] provided the first theoretical prediction of the amount of baryonic matter in the universe. This was possible, because there is only a short window in time for nucleosynthesis to take place. The abundance of light elements depends on two parameters: the expansion rate of the universe (matter to radiation ratio) and the density of neutrons and protons (*baryonic matter*). The prediction of Gamow et al was consistent with the observed matter density in galaxies and intervening gases, namely a few percent of the critical density. However, this number was a bit puzzling for another reason. Observations of the rotation curves of the outer parts of galaxies, starting with the measurements of Babcock [31] and Oort [32], suggested that the amount of matter in a galaxy should be much higher, about 20-30 percent of the critical density.

Meanwhile, after the establishment of the Hot Big Bang model and with the first all-sky redshift surveys of galaxies, Peebles' picture of hierarchical structure formation by gravitational clustering [33] gained more support. This led astronomers to look for temperature variations in the CMB as a proxy for the amplitude of initial density fluctuations. However, already in the late 70s it was clear that the temperature fluctuations were too small to account for all the observed structures [34]. This led several astronomers to suggest the existence of non-baryonic matter. This idea was not new, because the ear-

lier observations of Kapteyn [35] (reinterpreted by Jeans [36]) and Zwicky [37] indicated the presence of invisible matter, or *dark matter*. Decisive evidence of dark matter came with the high precision observations of the flattening of galaxy rotation curves by Rubin [38]. The only known non-baryonic matter particle was the neutrino, but it was soon realized that its velocity dispersion is too high to support structure formation on galactic scales. Therefore, several *Cold Dark Matter* (CDM) scenarios were proposed [39–43]. The important property of CDM is that it has to move with non-relativistic velocities in the early universe. In addition, the CDM model solved the missing matter problem described above, so it killed two birds with one stone. Therefore, even though CDM is of unknown nature up until today, it became part of our modern cosmological model.

So far so good. A more coherent picture of the contents of the universe started to emerge. It was believed that the universe contains radiation and both visible and dark matter, together accounting for $\rho_m \sim 0.2\rho_c$. If no other form of energy density were present, this would imply that $\Omega \sim 0.2$ and that the universe has an open geometry. However, based on the expectation that the universe was flat (see § 1.1.3 below), Einstein’s cosmological constant was reintroduced [44–46] to have $\Omega = 1$. Evidence of a cosmological constant (now also called *dark energy*) came from supernovae data [47, 48]. It was shown that the universe is currently in accelerated expansion, exactly what happens if the cosmological energy density is dominated by dark energy. Moreover, after COBE identified the tiny temperature variations in the CMB, many more experiments followed. In particular, the WMAP satellite [49] collected data on the CMB precise enough to confirm that a flat geometry was favored.

1.1.3 Inflation

With the successes of BBN and CMB another mystery arose. Because the universe contains a minimal amount of matter, it means that Ω is non-zero today, and must have been very close to unity in the past. More specifically, it was argued that if $1 - \Omega < 0.9$ today, then $1 - \Omega < 10^{-16}$ at the time of nucleosynthesis [50]. Moreover, the detection of the CMB suggests that the photons are in thermal equilibrium across the full sky at the time of decoupling. However, in the Hot Big Bang model there are about 10^4 causally disconnected regions that have not had enough time to interact and reach thermal equilibrium. So, why is the universe so similar everywhere? It seemed that either a lot of fine-tuning was required in the initial conditions, or the presence of new physics.

These problems motivated Guth [51] to propose *cosmological inflation*⁴: a period of exponentially accelerated expansion in the very early universe. During inflation $1 - \Omega$ quickly decays, and moreover, causally connected regions become exponentially larger. In other words, the seemingly fine-tuned initial conditions would be a natural consequence of inflation. However, soon it was realized [52, 53] that his model did not allow for a smooth transition to a FLRW universe. A new type of inflationary scenario was introduced by Linde [54] and by Albrecht and Steinhardt [55], which did not suffer from the so-called ‘graceful exit problem’. Analyses of quantum effects⁵ followed [58–61] and showed that inflation also generates small density inhomogeneities. Inflation could provide the seeds of structure formation!

In light of this major success, the paradigm of inflation was embraced. Finally, Linde [62] discovered that a stage of inflation would naturally occur for a broad class of scalar potentials. This is called *chaotic inflation*, whose mechanism is discussed in more detail in § 1.2. The idea that inflation could serve as the origin of structure was so exciting that theorists tried to embed it in the cosmological model (see the discussion on dark energy § 1.1.2). Compelling evidence for inflation came with the observations of the WMAP satellite [49].

This completes the picture of Λ CDM cosmology: an expanding flat universe filled with dark energy, cold dark matter, baryonic matter, neutrinos and radiation. This is complemented by scale-free initial conditions for structure formation. A period of inflation could account for these initial seeds. After inflation ends, its energy density has to be transferred to the Standard Model particles, after which the Hot Big Bang starts.

1.1.4 What’s next?

Fast forward to today. Many more precision tests have been performed and confirmed the predictions of Λ CDM. The baseline model, parameterized by only six parameters, works surprisingly well. The current best-fit values of the density parameters of dark energy, dark matter, baryonic matter derived from the CMB are [63]

$$\Omega_{\Lambda}^0 = 0.692 \pm 0.012, \quad \Omega_c^0 h^2 = 0.1186 \pm 0.0020, \quad \Omega_b^0 h^2 = 0.02226 \pm 0.00023 .$$

⁴For more details on the development of inflation, see [3]

⁵In fact the road had been paved earlier by the quantum perturbation analysis of Mukhanov and Chibisov [56] for Starobinsky’s model [57] of the primordial universe. Also from the English translation, which can be found in [4], they already understood that a phase of cosmological expansion can create the seeds for galaxy formation.

Here the superscript 0 means that the density parameter is evaluated at the present time. The parameter h equals the value of the Hubble parameter in particular units, it is defined in Eq. 1.5. In addition, there are neutrinos and radiation, but today their contribution to the energy density is very small. Radiation did, however, dominate the energy budget at earlier times. Moreover, the value of the Hubble parameter at the present time has the best-fit value (from the CMB)

$$H_0 \equiv 100h \frac{\text{km}}{\text{s Mpc}} = 67.81 \pm 0.92 \frac{\text{km}}{\text{s Mpc}} . \quad (1.5)$$

In addition, two well-constrained assumptions of Λ CDM are the flatness of the universe and the homogeneity and isotropy of the universe.

The success of Λ CDM is fantastic in the light of observational data. However, we have no clue about the nature of dark matter, dark energy and what drives inflation. More optimistically speaking, perhaps we are on the verge of another radically new idea. For instance, some tensions between datasets are reported [64–66], which might provide us with some hints. And of course the next generation of CMB experiments [67] and large scale structure surveys [68, 69] are designed to reach unprecedented precision with the aim to discover new physics.

The remainder of the thesis is dedicated to scientific questions related to inflation. Broadly speaking, we aim to improve our understanding of which signals to look for and what we can learn about inflation from observations in the near future.

1.2 Inflation

We review some relevant elements of inflation, as it is the main topic of this thesis. We start with a recap of single-field inflation. In particular, we outline the generic mechanism of chaotic inflation in § 1.2.1 and the generation of perturbations in § 1.2.4. Two important ingredients are the slow-roll approximation and mode freezing at horizon crossing. Mode freezing connects inflation naturally to the late universe. It results in the observed ‘phase coherence’ of the CMB, which is the most compelling evidence in favor of inflation, as we will discuss in § 1.2.5.

Interestingly, the inflationary predictions are sensitive to the details of the field theory under consideration (see § 1.2.4 and § 1.3). Inflation is therefore

the crossroad where particle physicists and cosmologists meet, and precision cosmological observations might reveal a glimpse of high-energy physics. We discuss the theoretical targets in § 1.2.4 and summarize the outline of future observations that might help us to reveal the nature of inflation in § 1.2.6. Finally, we outline the scope of this thesis in § 1.3.

1.2.1 Chaotic inflation

Inflation is a period of accelerated expansion in the early universe. How do we meet the requirements from § 1.1.3 that it solves the fine-tuning problems, but also has a graceful exit to a FLRW universe? To achieve the former it turns out that we need at least 60 e-folds of inflation (see e.g. [70]). One e-fold is the time in which the universe expands by a factor of e , so after N e-folds the scale factor grows with a factor e^N . In other words: $dN = d \ln a = H dt$ (see the discussion following Eq. 1.2).

One thing that comes to mind is a slowly varying cosmological ‘constant’, or $\dot{H} \approx 0$. Linde [62] figured out that this can be realized with a simple scalar field theory, where the scalar field slowly rolls down to the minimum of its potential, providing a source of almost constant energy density. Importantly, in these models inflation is an *attractor* solution and will eventually happen for a wide range of initial conditions, which explains the name *chaotic inflation*. Let’s see how this works.

We consider the toy model of a simple scalar field ϕ minimally coupled to Einstein gravity

$$S = -\frac{1}{2} \int \sqrt{-g} [-M_p^2 R + g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi + 2V(\phi)], \quad (1.6)$$

Here R is the Ricci scalar of spacetime. The potential is a smooth function of ϕ with minimum $V(\phi_0) = 0$. Without loss of generality we can take $\phi_0 = 0$. For instance, one could think of a quadratic potential $V = \frac{1}{2} m^2 \phi^2$. The dynamics of the homogeneous scalar field is determined by its equations of motion

$$\ddot{\phi} + 3H\dot{\phi} + V_\phi = 0, \quad (1.7)$$

where we denote a derivative with respect to ϕ with a subscript $V_\phi \equiv \frac{\partial V}{\partial \phi}$, together with the Friedmann equations

$$3H^2 M_p^2 = \frac{1}{2} \dot{\phi}^2 + V, \quad (1.8)$$

$$\dot{H} = -\frac{\dot{\phi}^2}{2M_p^2}. \quad (1.9)$$

The field equation is the same as that of a damped harmonic oscillator, where the Hubble friction depends on the dynamics of the field itself. The second Friedmann equation [Eq. 1.9](#) is not an independent equation, but will be useful in [§ 1.2.3](#). Qualitatively, the friction is large when the scalar field is far from the origin, and it gets even larger when the initial velocity $\dot{\phi}$ is substantial. The friction quickly slows down the scalar field, such that it approximately follows the *gradient flow* [[54, 55](#)]

$$3H\dot{\phi} + V_\phi \approx 0 . \quad (1.10)$$

Assuming that we can neglect the contribution of the kinetic energy to the Hubble friction in [Eq. 1.8](#), the gradient flow approximation becomes (in e-folds)

$$\frac{d\phi}{dN} \approx -\frac{M_p^2 V_\phi}{V} . \quad (1.11)$$

This shows that the field displacement is small within one expansion time, as long as $V \gg M_p |V_\phi|$. In that case, the potential energy is slowly varying, and the kinetic energy is indeed negligible. This implies that the Hubble parameter is almost constant, and the scale factor grows quasi-exponentially fast for many e-folds. This is exactly as desired! Finally, we have to check the validity of the gradient flow approximation. Using [Eq. 1.11](#) we see that we can neglect the field acceleration in [Eq. 1.7](#) if in addition $V \gg M_p^2 |V_{\phi\phi}|$.

To prove the attractor behavior of chaotic inflation one has to do a more careful stability analysis. For a nice treatment see for instance [[4](#)]. In this thesis we will instead review an argument using the Hamilton-Jacobi formulation in [§ 1.2.3](#). Moreover, we assume a homogeneous scalar field from the start. For recent studies of the attractor behavior of scalar field inflation with inhomogeneous initial conditions see [[71, 72](#)], see also the review [[73](#)].

1.2.2 Slow-roll inflation

An important feature of chaotic inflation is that the scalar field quickly approaches the gradient flow. Tracing it back to the original requirement of accelerated expansion, this ensures that the Hubble parameter is slowly changing in time $-\frac{\dot{H}}{H^2} \ll 1$. Therefore, the gradient flow approximation can be formalised as the *Hubble slow-roll approximation* [[74–76](#)]. Slow-roll inflation is defined as a period of near-exponential expansion for which all Hubble slow-roll parameters are much smaller than unity

$$\epsilon \equiv -\frac{\dot{H}}{H^2}, \quad \epsilon_{n+1} \equiv \frac{\dot{\epsilon}_n}{H\epsilon_n} \quad \text{for } n \geq 2 . \quad (1.12)$$

Here it is understood that $\epsilon = \epsilon_1$. Usually, the second slow parameter is denoted by $\eta = \epsilon_2$. Notice that in the toy model of § 1.2.1 we could also have rephrased the gradient flow assumption in terms of conditions on potential slow-roll parameters $\epsilon_V \sim \frac{M_p^2}{2} \left(\frac{V_\phi}{V} \right)^2 \ll 1$ and $\eta_V \sim \frac{M_p^2}{2} \frac{V_{\phi\phi}}{V} \ll 1$. However, the Hubble slow roll approximation applies to a much more general class of inflationary models, including many models of multi-field inflation⁶. For that reason we will use the Hubble slow-roll parameters, defined in Eq. 1.12, in the remainder of this thesis.

1.2.3 Hamilton-Jacobi formulation

A particularly neat way to show linear stability of slow roll inflation is presented by Salopek and Bond [74], see also [76]. They employ the Hamilton-Jacobi formulation [74, 77, 78] of single field inflation. Moreover, this formalism can be used to construct inflationary models with exact solutions.

In the Hamilton-Jacobi formalism the scalar field itself⁷ parameterizes the inflationary trajectory, and the Hubble parameter $H(\phi)$ determines the inflationary dynamics. The Hubble parameter replaces the potential as fundamental input function and straightforwardly gives all slow-roll parameters (combining Eq. 1.12 and Eq. 1.14)

$$\epsilon \equiv 2M_p^2 \frac{H_\phi^2}{H^2}, \quad \epsilon_{n+1} \equiv -2M_p^2 \frac{\epsilon_{n,\phi} H_\phi}{\epsilon_n H} \quad \text{for } n \geq 2. \quad (1.13)$$

The potential and the Hubble parameter are related through the Hamilton-Jacobi equation, though. This allows us to prove linear stability of chaotic inflation. In addition, this formalism can be used to (locally) reconstruct the potential, given the Hubble slow-roll parameters that characterise the observations (see § 1.2.4).

⁶In a multi-dimensional field space the inflationary trajectory is in general not aligned with the gradient of the potential. Therefore, it makes no sense to use the gradient of the potential as a proxy of the size of the field velocity.

⁷The only requirement is that ϕ is monotonic. This means that, for instance, oscillations around the minimum of the potential cannot be captured within this approach.

Hamilton-Jacobi equation

An easy way to derive the Hamilton-Jacobi equation is to replace $H(t)$ with $H(\phi)$ and use the second Friedmann equation [Eq. 1.9](#) to write

$$\dot{H} = \dot{\phi} H_\phi = -\epsilon H^2 = -\frac{\dot{\phi}^2}{2M_p^2} \quad \longrightarrow \quad -2M_p^2 H_\phi = \dot{\phi}. \quad (1.14)$$

This allows us to rewrite the first Friedmann equation [Eq. 1.8](#) as the Hamilton-Jacobi equation

$$V = 3H^2 M_p^2 - 2M_p^4 H_\phi^2, \quad (1.15)$$

where all functions now explicitly depend on ϕ . Given some Hubble parameter $H(\phi)$, the potential follows straightforwardly from the Hamilton-Jacobi equation. If we take the field derivative of [Eq. 1.15](#) and use [Eq. 1.14](#) we find the field equation of motion [Eq. 1.7](#). This means that the Hamilton-Jacobi equation is equivalent to the field equation of motion, modulo some integration constant ϕ_0 . The initial value of the Hubble parameter is the same as specifying $\dot{\phi}_0$ at some ϕ_0 .

Stability of slow-roll inflation

We now review the stability argument presented by Salopek and Bond [\[74\]](#). Let's assume that we identified a solution $H(\phi)$ to [Eq. 1.15](#) for a given potential. This is possible in the regime where ϕ is monotonic. It follows directly from perturbing [Eq. 1.15](#), while keeping the potential fixed, that a small perturbation in H evolves as

$$\delta H(\phi) = \delta H(\phi_0) \exp\left(\int_{\phi_0}^{\phi} d\phi \frac{3}{2M_p^2} \frac{H}{H_\phi}\right). \quad (1.16)$$

Since $\dot{\phi}$ and $\frac{H}{H_\phi}$ have opposite sign the perturbations will decay. To understand how fast they decay rewrite the argument of the exponent in terms of e-folds by using $d\phi = \frac{\dot{\phi}}{H} dN = -2M_p^2 \frac{H_\phi}{H} dN$. This yields

$$\delta H(N) = \delta H(0) \exp(-3N). \quad (1.17)$$

We learn that different solutions $H(\phi)$ of [Eq. 1.15](#) approach each other exponentially fast. Therefore, any functional form of $H(\phi)$ provides an (possibly non-inflationary) linear attractor solution to the corresponding potential, computed from [Eq. 1.15](#), as long as it is monotonic.

To address the attractor behavior of chaotic inflation, let's make use of the slow roll expansion. The zeroth order solution (with $\epsilon = 0$) for the Hubble

parameter is given by $H_{\text{SR}}^2 = \frac{V}{3M_p^2}$. The solution $H(\phi)$ to the Hamilton-Jacobi equation is always larger than this, but to support inflation it has to obey

$$H_{\text{SR}} < H(\phi) < \sqrt{\frac{3}{2}} H_{\text{SR}} . \quad (1.18)$$

Therefore, is there a solution close to H_{SR} ? If yes, it can be found perturbatively from Eq. 1.15

$$H^{(0)} = H_{\text{SR}}, \quad H^{(1)} = H_{\text{SR}} \sqrt{1 + \frac{1}{3}\epsilon^{(0)}}, \quad (1.19)$$

etcetera. Therefore, we expect that if the slow-roll parameters computed with H_{SR} are small, then the solution to Eq. 1.15 is an inflationary solution, and moreover it is an attractor. In other words, slow-roll inflation is an attractor. Please see [76] for a more precise treatment.

Reconstruction potential

The Hamilton-Jacobi formalism can also be used to find exact inflationary solutions [75, 77, 79, 80]. To illustrate this, we reconstruct the potential which allows for slow roll parameters of the form

$$\epsilon = \frac{p}{2\Delta N + p}, \quad (1.20)$$

where $\Delta N = N_e - N$ is the number of e-folds before the end of inflation. We added a p in the denominator to ensure that $\epsilon = 1$ at the end of inflation. Notice that with this parameterization all other slow-roll parameters follow automatically. This parameterization of ϵ has to be understood as a fit to the data, i.e. it only is guaranteed to work well at the time that the observable modes cross the horizon (see § 1.2.4). Therefore, the potential we reconstruct provides a good approximation around the field values of horizon crossing only.

First, we write ϵ in terms of ϕ by integrating $d\phi/dN = -\sqrt{2\epsilon}M_p$. Here the minus sign is our choice of convention. This yields $\epsilon = \frac{2M_p^2 p^2}{\phi^2}$. From Eq. 1.13 it then follows that $H = c\phi^p$, with c some constant. Therefore, we find the following potential from the Hamilton-Jacobi equation Eq. 1.15

$$V = 3c^2 M_p^2 \phi^{2p-2} \left(\phi^2 - \frac{2p^2 M_p^2}{3} \right). \quad (1.21)$$

These are the familiar power law potentials up to a small correction, which is only there to ensure the exact behavior of the assumed slow roll parameters. One can indeed check that $\dot{\phi} = Hd\phi/dN = -2cpM_p^2\phi^{p-1}$ solves the field

equations of motion exactly.

One might worry that if we slightly change the functional form of $H(\phi)$ around horizon crossing, this could change the global structure of the potential quite a bit. However, in single field slow roll inflation this will not affect the inflationary predictions, so long the slow parameters behave smoothly. Therefore, we should not take the global behavior of the resulting potential too seriously, except possibly when the form of $H(\phi)$ is constrained by symmetries.

1.2.4 Inflation as the origin of structure

The major success of inflation is that it provides a mechanism to create the initial seeds for structure formation from quantum fluctuations [56]. The rough picture is that quantum perturbations are stretched exponentially fast during inflation. The Hubble scale $\frac{1}{aH}$ quickly drops below their comoving wavelengths, where their amplitude freezes out. At much later times, when they become sub-Hubble again, they provide the gravitational wells that triggers structure formation.

Primordial perturbations

Let us recap the main steps in the computation of primordial perturbations for single field slow roll inflation. We follow the treatments of [5, 81, 82], albeit with a different notation. The leading order behavior of quantum fluctuations is captured by a linear perturbation analysis. Therefore, we start with linear perturbations around the homogeneous inflationary background solution. Using the ADM decomposition of the metric [83] we write

$$\phi(t, \mathbf{x}) = \phi(t) + \delta\phi(t, \mathbf{x}), \quad (1.22a)$$

$$ds^2 = -\mathcal{N}dt^2 + g_{ij} (dx^i + \mathcal{N}^i dt) (dx^j + \mathcal{N}^j dt) \quad (1.22b)$$

align Here $\mathcal{N}(t, \mathbf{x})$ and $\mathcal{N}_i(t, \mathbf{x})$ are the lapse and the shift functions, respectively. They are not dynamical, but auxiliary variables solved by the Einstein equations, which become constraints in this case. The remaining perturbations in the metric can be decomposed into scalar, vector and tensor modes, depending on how they transform under rotations about the axis defined by its Fourier vector \mathbf{k} . This decomposition is quite useful because the different types of perturbations decouple at the linear level, as a consequence of translational and rotational invariance of the background [5]. The spatial metric perturbation g_{ij} together with the field perturbation can be split in two scalar, one vector and one tensor mode. However, not all of them are physical degrees of freedom, since reparametrization invariance $x^\mu \rightarrow x^\mu + \xi^\mu(t, \mathbf{x})$, removes one

scalar and one vector mode. Therefore, we are left with one scalar and one tensor mode, equivalent to one scalar and two tensor degrees of freedom. For the purpose of the computation it is useful to work in a particular gauge that fixes time and spatial reparametrizations. We choose the *comoving gauge*

$$\delta\phi = 0 \quad \text{and} \quad g_{ij} = a^2(t) \left((1 + 2\mathcal{R})\delta_{ij} + h_{ij} \right), \quad (1.23)$$

where \mathcal{R} is the *comoving curvature perturbation* which measures the curvature of the spatial hypersurfaces in this gauge. Moreover, h_{ij} is transverse and traceless and contains the two tensor degrees of freedom. The quadratic action for \mathcal{R} and h_{ij} can be found by expanding the full action Eq. 1.6 to second order, replacing \mathcal{N} and \mathcal{N}^i by their solution of the constraint equations, and performing integration by parts

$$S^{(2)} = \int d^4x a^3 M_p^2 \left[\epsilon \left(\dot{\mathcal{R}}^2 - \frac{1}{a^2} (\nabla\mathcal{R})^2 \right) + \frac{1}{8} \left(\dot{h}_{ij}^2 - \frac{1}{a^2} (\nabla h_{ij})^2 \right) \right]. \quad (1.24)$$

To perform the quantization of the curvature perturbations, it is convenient to go to conformal time $ad\tau = dt$ and change to the canonically normalized Mukhanov-Sasaki variable [84–86]

$$v \equiv z\mathcal{R}M_p \quad \text{with} \quad z \equiv a\sqrt{2\epsilon}. \quad (1.25)$$

The quantization of the theory [87] proceeds by promoting the field v to a quantum operator

$$\hat{v}(t, \mathbf{x}) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \left[v_k(\tau) \hat{a}_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}} + v_k^*(\tau) \hat{a}_{\mathbf{k}}^\dagger e^{-i\mathbf{k}\cdot\mathbf{x}} \right], \quad (1.26)$$

where the creation and annihilation operators satisfy the commutation relations

$$\left[\hat{a}_{\mathbf{k}}, \hat{a}_{\mathbf{p}}^\dagger \right] = (2\pi)^3 \delta_D(\mathbf{k} - \mathbf{p}), \quad \left[\hat{a}_{\mathbf{k}}, \hat{a}_{\mathbf{p}} \right] = \left[\hat{a}_{\mathbf{k}}^\dagger, \hat{a}_{\mathbf{p}}^\dagger \right] = 0. \quad (1.27)$$

This, together with the selection of the Bunch-Davies vacuum $\hat{a}_{\mathbf{k}}|0\rangle = 0$, fixes the initial conditions of the mode functions $v_k(\tau)$. The mode functions satisfy a modified Klein-Gordon equation in conformal time

$$v_k'' + \left(k^2 - \frac{z''}{z} \right) v_k \quad (1.28)$$

In the short wavelength limit $k^2 \gg \frac{z''}{z}$ this is the equation of a harmonic oscillator. In the long wavelength limit $k^2 \ll \frac{z''}{z}$ one of the solutions is $v_k \sim z$, which corresponds to $\mathcal{R} \sim \text{const}$. This agrees with the homogeneous equation

of motion $\partial_t(a^3\epsilon\dot{\mathcal{R}}) = 0$, derived from the quadratic action [Eq. 1.24](#). The constant solution is the dominant one, as long as $\int \frac{dt}{a^3\epsilon}$ is decaying. For slow-roll inflation this is always the case.

More precisely, at first order in slow-roll, the solution of the mode function to [Eq. 1.28](#) subject to the quantum initial conditions is given by

$$v_k(\tau) = \sqrt{\frac{\pi}{4k}} \sqrt{-k\tau} H_\nu^{(1)}(-k\tau), \quad \text{with} \quad \nu \equiv \frac{3}{2} + \epsilon - \frac{1}{2}\eta, \quad (1.29)$$

modulo a random phase. On *superhorizon scales* $-k\tau = \frac{k}{aH} \ll 1$ the mode function of the curvature perturbation asymptotes to [\[88\]](#)

$$\lim_{-k\tau \ll 1} |k^{3/2}\mathcal{R}_k| = \frac{2^\nu \Gamma(\nu)}{\sqrt{8\pi\epsilon}} \frac{H}{M_p} (-k\tau)^{3/2-\nu}. \quad (1.30)$$

Using the approximate relation [\[82\]](#) that for slow-roll inflation $\frac{H}{M_p\sqrt{\epsilon}} (-k\tau)^{3/2-\nu}$ is a constant, this allows us to evaluate [Eq. 1.30](#) at *horizon crossing* $-k\tau = \frac{k}{aH} = 1$ where $\nu \approx 3/2$. This validates the *horizon crossing formalism*, where the typical size of curvature perturbations at the end of inflation is related to the inflationary background quantities evaluated at horizon crossing.

The variance of curvature perturbations is captured by the dimensionless power spectrum $\Delta_{\mathcal{R}}^2(k)$, defined as

$$\frac{k^3}{2\pi^2} \langle \mathcal{R}(\mathbf{k})\mathcal{R}(\mathbf{k}') \rangle \equiv (2\pi)^3 \delta_D^{(3)}(\mathbf{k} + \mathbf{k}') \Delta_{\mathcal{R}}^2(k). \quad (1.31)$$

Homogeneity and isotropy imply that the correlation functions in Fourier space always come with an overall delta-function. Using [Eq. 1.30](#) the dimensionless power spectrum evaluates to

$$\Delta_{\mathcal{R}}^2(k) = \frac{H^2}{8\pi^2\epsilon M_p^2} \Big|_{k=aH}. \quad (1.32)$$

Since H and ϵ are nearly constant during slow-roll inflation, the power spectrum is almost scale-invariant. It has a small scale dependence though, because different modes cross the horizon at different times and they will all feel a different Hubble and slow-roll parameter. Therefore it is convenient to parameterize the power spectrum by its amplitude A_s and *scalar spectral index* n_s

$$\Delta_{\mathcal{R}}^2(k) = A_s(k_*) \left(\frac{k}{k_*} \right)^{n_s-1}. \quad (1.33)$$

Here k_* is the ‘pivot scale’, a reference scale at which the amplitude is evaluated.

A similar analysis applies to the tensor perturbations [57]. The transverse and traceless h_{ij} contains two tensor polarization modes. Comparing the relative normalizations of h_{ij} and \mathcal{R} in the quadratic action Eq. 1.24 we expect the power spectrum of tensor perturbations Δ_t^2 to be $2 \times 8\epsilon \times \Delta_{\mathcal{R}}^2$, which is indeed the case:

$$\Delta_t^2(k) = \frac{2H^2}{\pi^2 M_p^2} \Big|_{k=aH} . \quad (1.34)$$

Also the power spectrum of tensor perturbations can be parametrized in terms of the *tensor spectral index* n_t and the *tensor-to-scalar-ratio* r

$$\Delta_t^2(k) = A_t(k_*) \left(\frac{k}{k_*} \right)^{n_t} \quad \text{and} \quad r(k_*) \equiv \frac{A_t(k_*)}{A_s(k_*)} . \quad (1.35)$$

The tensor-to-scalar-ratio measures the ratio between the amplitude of tensor perturbations with respect to the scalar perturbations.

Theoretical Targets - First of all, a measurement of stochastic background of primordial gravitational waves on large scales would be groundbreaking, as it probes the quantum nature of gravity. Moreover, it provides another important confirmation of inflation and probes its energy scale. For instance, for the simplest models of inflation⁸ a measurement of both the scalar amplitude A_s and the tensor-to-scalar ratio r is translated into $V^{1/4} \sim \left(\frac{r}{0.009} \right)^{1/4} \times 10^{16}$ GeV [89]. It has been argued [67] that $\sigma(r) = 0.001$ is a clear theoretical benchmark, because it allows us to falsify all simple models that naturally explain the observed spectral tilt.

Moreover, using the slow-roll approximation, we can summarize the leading order predictions of the simplest models of inflation as

$$r = 16\epsilon, \quad n_s = 1 - 2\epsilon - \eta, \quad n_t = -2\epsilon . \quad (1.36)$$

Notice the *consistency relation* $r = -8n_t$ [75]. This provides us with another theoretical target: a violation of this consistency relation signals that we should break at least one of the conditions listed in footnote 8.

At this stage it is perhaps interesting to remark that, going beyond the simplest realization of inflation, the curvature perturbations could have a sound

⁸ By the simplest models of inflation we mean canonical single-field slow-roll inflationary models minimally coupled to gravity, and with Bunch-Davies initial conditions.

speed c_s smaller than unity, see e.g. [90]. The quadratic action for curvature perturbations now reads

$$S^{(2)} = \int d^4x a^3 M_p^2 \left[\epsilon \left(\frac{\dot{\mathcal{R}}^2}{c_s^2} - \frac{1}{a^2} (\nabla \mathcal{R})^2 \right) \right]. \quad (1.37)$$

Repeating similar computations as outlined above yields $r = 16\epsilon c_s$, assuming that $c_s \approx \text{const.}$ In other words, this might change our interpretation of the data. In [Chapter 5](#) we will illustrate how this effect may arise in a simple two field set-up, where the curvature perturbation interacts with a heavy degree of freedom.

Primordial non-Gaussianities

We have seen that single field inflation generates scalar and tensor perturbations. Above we only computed their variance, but in principle the full probability distribution could be computed. Primordial non-Gaussianities have been studied thoroughly since the pioneering works [81,91], see e.g. [92,93] for reviews. Since the computation of primordial perturbations is organized perturbatively, it is natural to work with correlation functions. In this thesis we consider only the two- and three-point correlation functions. The bispectrum of curvature perturbation is defined by

$$\langle \mathcal{R}(\mathbf{k}_1) \mathcal{R}(\mathbf{k}_2) \mathcal{R}(\mathbf{k}_3) \rangle \equiv (2\pi)^3 \delta_D^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B_{\mathcal{R}}(k_1, k_2, k_3). \quad (1.38)$$

A similar definition applies to the tensor perturbations and any cross-correlation between the variables. The dimensionless bispectrum and shape function are given by

$$\mathcal{I}_{\mathcal{R}}(k_1, k_2, k_3) \equiv \frac{k_1^2 k_2^2 k_3^2}{4\pi^4} B_{\mathcal{R}}(k_1, k_2, k_3), \quad (1.39a)$$

$$f_{NL}(k_1, k_2, k_3) \equiv \frac{5}{6} \frac{B_{\mathcal{R}}(k_1, k_2, k_3)}{P_{\mathcal{R}}(k_1)P_{\mathcal{R}}(k_2) + P_{\mathcal{R}}(k_1)P_{\mathcal{R}}(k_3) + P_{\mathcal{R}}(k_2)P_{\mathcal{R}}(k_3)}, \quad (1.39b)$$

respectively. The denominator of the shape function depends on the dimensionfull power spectrum $P_{\mathcal{R}}(k) \equiv \frac{2\pi^2}{k^3} \Delta_{\mathcal{R}}^2(k)$. The full scale dependence of the bispectrum for canonical single field slow-roll inflation has been computed for the first time by Maldacena [81]. It turns out that the amplitude of the signal is very small, because interaction terms are slow-roll suppressed $f_{NL} \sim \mathcal{O}(\epsilon, \eta)$.

The bispectrum contains a wealth of information. If we constrain or probe its amplitude and its three dimensional scale-dependence, we learn about the possible (self-)interactions of the inflaton. More sizeable non-Gaussianities $f_{NL} \sim 1$ might be generated if we deviate from one of the assumptions of

canonical single-field slow-roll inflation with Bunch-Davies initial conditions. However, please keep in mind that a bispectrum with amplitude of order unity is already extremely challenging to measure, since the amplitude of the power spectrum is tiny $\Delta_{\mathcal{R}}^2 \sim 10^{-9}$ (see § 1.2.5), implying that $\mathcal{I}_{\mathcal{R}} \sim 10^{-18} f_{NL}$. But this is definitely worth pushing for.

The bispectrum is a three-dimensional function, which makes it challenging to compare efficiently with data. Also, the scale-dependence cannot always be computed analytically for a given inflationary model. In order to compare with the data, it is therefore useful to have a set of templates that resemble various possible scale-dependencies of the bispectrum well. Let us highlight three well-motivated shape templates for the bispectrum:

- *The local shape.* A simple modification to the Gaussian probability distribution function (in real space) arises from a small non-linear correction to the Gaussian curvature perturbation [94]

$$\mathcal{R}(\mathbf{x}) = \mathcal{R}_g(\mathbf{x}) + \frac{3}{5} f_{NL}^{\text{loc}} (\mathcal{R}_g^2(\mathbf{x}) - \langle \mathcal{R}_g^2(\mathbf{x}) \rangle) . \quad (1.40)$$

This corresponds to a constant shape function equal to f_{NL}^{loc} . This explains why f_{NL} is defined as it is. The factor 3/5 comes from the translation to the primordial gravitational potential, see § 1.4. The dimensionless bispectrum, on the other hand, peaks in the *squeezed configuration* where one of the wavenumbers is much smaller than the others $k_1 \ll k_2, k_3$. This type of non-Gaussianity arises for instance in multi-field inflation [95], where the curvature perturbations are sourced on super-Hubble scales by the other fields. For a nice explanation and a review of other scenarios that create local non-Gaussianities see [96].

- *The equilateral shape.* The dimensionless bispectrum of equilateral non-Gaussianity peaks in the equilateral configuration where all wavenumbers are equal $k_1 = k_2 = k_3$.

$$\mathcal{I}_{\mathcal{R}}^{\text{eq}}(k_1, k_2, k_3) = \frac{6}{5} f_{NL}^{\text{eq}} \frac{81 \Delta_{\mathcal{R}}^2}{(2\pi^2)^4} \frac{k_1 k_2 k_3}{K^3}, \quad \text{with } K \equiv k_1 + k_2 + k_3 . \quad (1.41)$$

Equilateral non-Gaussianity is generated in models of inflation where the interactions are most important around the time of horizon crossing. For instance it is created in single field inflation with a small inflaton sound speed [97–99].

- *The quasi-single-field shape.* Finally, if the inflaton interacts with an isocurvaton of mass $\mu \lesssim H$ a bispectrum with a shape that interpolates

between local and equilateral is produced [100, 101]. Its scale dependence is well approximated by the following template

$$\mathcal{I}_{\mathcal{R}}^{\text{qsf}}(k_1, k_2, k_3) = \frac{6}{5} f_{NL}^{\text{qsf}} \frac{9\sqrt{3}\Delta_{\mathcal{R}}^2}{(2\pi^2)^4} \frac{k_1 k_2 k_3}{K^3} \frac{N_{\nu}(8\kappa)}{\sqrt{\kappa} N_{\nu}(8/27)}, \quad (1.42)$$

with $\kappa \equiv k_1 k_2 k_3 / K^3$ and N_{ν} the Neumann function of order $\nu \equiv \sqrt{\frac{9}{4} - \frac{\mu^2}{H^2}}$.

Theoretical targets - Which precision do we require from future observations, such that we learn something even in the absence of a detection? Concerning the amplitude of these three shapes, the theoretical benchmark for equilateral non-Gaussianity reads $\sigma(f_{NL}^{\text{eq}}) = \mathcal{O}(1)$. First of all, single-field slow-roll inflation necessarily produces an amplitude smaller than unity [97]. Second, a detection of f_{NL}^{eq} larger than unity implies that inflation becomes strongly coupled, or it signals the presence of new fields [102–104]. The theoretical benchmarks for the other two shapes are less clear. In [103] it has been argued that $\sigma(f_{NL}^{\text{loc}}) = \mathcal{O}(1)$ is of theoretical interest, because this can disfavor a particular class of inflationary models.

A very important target for future observations is to constrain the amplitude of the bispectrum in the *squeezed configuration*, where one of the wavenumbers is much smaller than the others. Maldacena [81] derived an important *consistency relation* of the single field bispectrum

$$B_{\mathcal{R}}^{\text{sq}}(q, k_s) \equiv \lim_{q \ll k_s} B_{\mathcal{R}}(q, k_s, k_s) = (1 - n_s(k_s))P(q)P(k_s). \quad (1.43)$$

This relation was proven to be valid for all single-field attractor models of inflation [105–107]. Moreover, it was shown to be non-observable [108, 109] in the late universe, up to projection effects. In terms of the shape function one gets⁹

$$f_{NL}^{\text{sq, obs}}(q, k_s) = 0 + \mathcal{O}\left((q/k_s)^2\right) \quad (1.44)$$

Therefore, any detection of a signal in the squeezed limit would rule out single field inflation!

⁹The reason is that a local observer is freely falling in a FLRW background modified by the long wavelength perturbation \mathcal{R}_{ℓ} . It turns out that the observer can only distinguish \mathcal{R}_{ℓ} from the background by its second spatial derivative (or higher) [110].

Initial conditions after inflation

After inflation ends, its energy density is transferred into a soup of particles that lead to the universe today. This process is called reheating, and its precise mechanism is unknown. After many interactions in this dense plasma it is expected that the baryons, photons and neutrinos thermalize. This is the onset of the Hot Big Bang.

The reason why we can probe the epoch of inflation, regardless of the details of reheating and other unknown physics, is due to an important theorem by Weinberg [111]. It states that the curvature perturbation \mathcal{R} remains conserved on superhorizon scales whenever perturbations are adiabatic, independent of the matter content of the universe. Only when the curvature perturbations enter the horizon, they start to evolve again. This allows us to probe the primordial universe by observing the largest scales we can access today.

Adiabatic perturbations are perturbations that mimic a local time shift of the homogeneous background. Perturbations along the inflationary background solution correspond to a local time shift of the homogeneous background. Some parts of the universe end inflation slightly ahead of time and other parts slightly behind, i.e. single field inflation produces adiabatic perturbations. After the inflaton has decayed into the hot plasma, we can relate the overdensities of the various particle species X and Y to each other

$$\delta t = \frac{\delta \rho_Y}{\dot{\rho}_Y} = \frac{\delta \rho_X}{\dot{\rho}_X} \sim \frac{\delta_X}{1 + \omega_X} = \frac{\delta_Y}{1 + \omega_Y} . \quad (1.45)$$

Here we defined the *density contrast* $\delta_X \equiv \frac{\delta \rho_X}{\rho_X}$ and the *equation of state* $\omega_X \equiv p_X/\rho_X$ for each of the particles species X . This implies the following relation between matter, cold dark matter, photon and neutrino overdensities on superhorizon scales

$$\delta_b = \delta_c = \frac{3}{4} \delta_\gamma = \frac{3}{4} \delta_\nu . \quad (1.46)$$

Therefore, as long as we understand the evolution of the density perturbations when they enter the horizon, we can relate their statistical properties to those of \mathcal{R} . It is most convenient to study perturbations on the largest scales, because i) they only entered the horizon recently, where we understand better the constituents of the universe and ii) they are well captured by linear perturbation theory and carry the cleanest information from the early universe.

1.2.5 Evidence and constraints

The statistics of the CMB temperature variations confirm a couple of non-trivial predictions of inflation. The CMB provides us with a snapshot of the universe when it was only approximately 380,000 years old. This means that we probe, among other things, superhorizon correlations of density fluctuations. This is the reason that it provides an excellent test of inflation. Moreover, the current precision reached by CMB measurements starts to constrain the subtle details of inflation. We summarize the key findings so far:

- The structure of peaks and troughs of the CMB power spectrum provides compelling evidence in favor of inflation [112]. To appreciate this idea, we give an extremely simplified explanation of CMB physics, see e.g. [113, 114] for more details. Before the CMB photons were emitted, the photons were tightly coupled to electrons through Thomson scattering. The electrons were, in turn, tightly coupled to protons by the Coulomb interaction. Therefore, the system behaved like a single photon-baryon fluid. This fluid supports ‘acoustic oscillations’ as a consequence of the competing effects of gravitational collapse and radiation pressure. The gravitational source term is set by the initial conditions provided by inflation. The special feature of inflation is that $\dot{\mathcal{R}} = 0$ on superhorizon scales, which implies that the density perturbations all start oscillating in phase (at maximal amplitude). The acoustic oscillations suddenly stop at recombination when the radiation pressure disappears and the CMB photons start free-streaming towards us. Because the photons have to climb out of the potential wells created by the matter overdensities, they provide us with a snapshot of the density field at recombination. The density perturbations of a given wavenumber are all captured in the same phase of their oscillation, because they started oscillating at the same time when they entered the horizon. On the largest scales the perturbations are still frozen. On the smaller scales some modes are captured in their minimum, some in their maximum, and this produces the peaks and troughs in the CMB temperature power spectrum. If not for this phase coherence, the peaks in the CMB temperature spectrum would be washed away.
- The CMB probes the primordial power spectrum of curvature perturbations in the range $0.0001 \text{ Mpc}^{-1} \leq k \leq 0.3 \text{ Mpc}^{-1}$. The *Planck* collaboration [115] has reported a detection of the deviation from a scale-invariant power spectrum with 5.6σ confidence level. Moreover, the amplitude is given by $\ln(10^{10}A_s) = 3.089 \pm 0.036$ at $k_* = 0.05 \text{ Mpc}^{-1}$. Joint constraints on the scalar tilt n_s and the tensor-to-scalar ratio r are shown

in Figure 1.2. They are evaluated at the pivot scale $k_* = 0.002 \text{ Mpc}^{-1}$, which corresponds to a horizon-crossing time between 50 and 60 e-folds before the end of inflation [116].

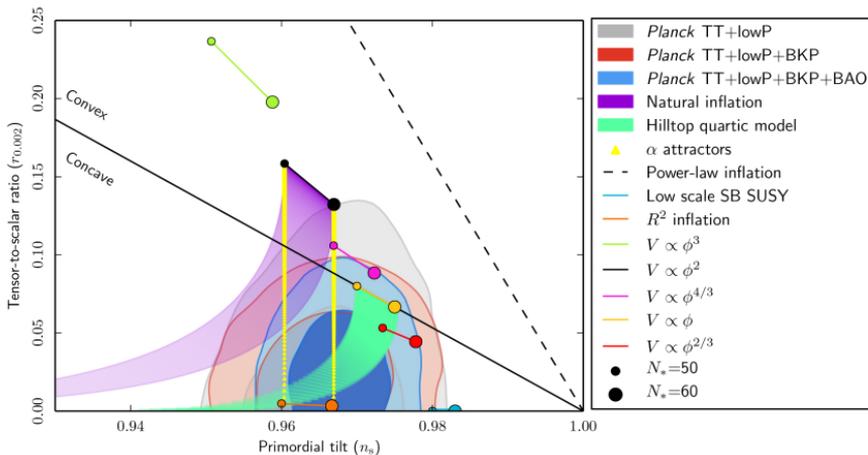


Figure 1.2: Marginalized joint 68% and 95% CL contours on the tensor-to-scalar ratio r and spectral tilt n_s at the pivot scale $k_* = 0.002 \text{ Mpc}^{-1}$ from *Planck* [115] together with other datasets. The predictions of some selected inflationary models are shown as well.

- Phase coherence puts strong constraints on the adiabaticity of the universe. *Planck* [115] has tested some of the relations in Eq. 1.46 to almost percent level accuracy. A detection would falsify single field slow-roll inflation, and is for that reason very interesting. It is possible that isocurvature (non-adiabatic) perturbations are generated during inflation, but it requires that more than one scalar degree of freedom is excited. Therefore, late-time isocurvature perturbations could give us some information about the mechanism of inflation. However, how the late-time isocurvature perturbations are related exactly to the isocurvature perturbations produced during inflation depends on the details of reheating. In fact, reheating might even wash out (part of) the primordial isocurvature perturbations [117].
- The Gaussianity of the primordial fluctuations has also been tested to a high degree. In the comparison with CMB data [118], various templates for the shape are considered, some of them are described in Eq. 1.40, Eq. 1.41 and Eq. 1.42. The amplitude is evaluated in the equilateral configuration $f_{NL}(k) \equiv f_{NL}(k, k, k)$. The current best constraints [118]

on the local and equilateral types of non-Gaussianity are

$$f_{NL}^{\text{loc}} = 0.8 \pm 5.0 \quad \text{and} \quad f_{NL}^{\text{eq}} = -4 \pm 43 \quad (68\% \text{ CL}) . \quad (1.47)$$

- The location of the first peak in the CMB spectrum depends on the geometry of the universe. This indicates that the universe is flat to very high precision $\Omega_\kappa = -0.005_{-0.017}^{+0.016}$ at 95% CL [115]. This is compatible with inflation, because it was designed to create a flat universe.

1.2.6 Future probes of inflation

What are the observations that will teach us more about inflation in the future?

Cosmic Microwave Background experiments - There is much room for improvement in CMB experiments to constrain the amplitude of primordial tensor fluctuations by observing *primordial B-mode polarization*. The CMB photons are polarized via Thomson scattering, see e.g. [119]. An electron at the last scattering surface is surrounded by a slightly non-isotropic radiation field, because of the small temperature variations. Therefore, the CMB photons will carry an overall polarization that is correlated with the temperature fluctuations. Polarization is a vector field and can be decomposed in so-called E-modes and B-modes [120–122]. This decomposition is particularly useful because *only* tensor perturbations can create a primordial B-mode signal. Therefore, constraining the amplitude of the primordial B-mode polarization is an important observational goal. It has been forecast that CMB-S4 [67] will be able to do better than $\sigma(r) = 0.001$, provided there is large increase in the number of detectors.

Concerning primordial non-Gaussianities, it has been estimated [123] that, with improved temperature and polarization measurements, the error bars can maximally decrease by a factor of two, which is unfortunately not sufficient to reach the theoretical thresholds quoted between Eq. 1.42 and Eq. 1.43. Fortunately, there is a promising opportunity to constrain the squeezed primordial bispectrum from the CMB by observing spectral distortions [124, 125]. The energy spectrum of the CMB is not a perfect blackbody, but has tiny distortions [126]. Two types of spectral distortions are sensitive to the integrated power spectrum between $1\text{Mpc}^{-1} < k < 50\text{Mpc}^{-1}$ [127] and $50\text{Mpc}^{-1} < k < 10^4\text{Mpc}^{-1}$ [126] respectively. Therefore, correlating the spectral distortion with large scale temperature fluctuations provides a measurement of the bispectrum in the ultra-squeezed limit. It has been forecast [128] that a futuristic idealized cosmic-variance-limited survey can reach error bars of order

$$\sigma(f_{NL}^{\text{sq}}) = \mathcal{O}(10^{-4}).$$

Large Scale Structure experiments - Long after the emission of the CMB photons the small initial density perturbations evolved into the large scale structures in which galaxies formed. Galaxies and neutral hydrogen trace the underlying matter distribution, and can be used to infer its statistical properties. Moreover, weak lensing in addition infers the integrated matter density along the line of sight. Since large scale structure surveys help us to reconstruct the three-dimensional distribution of matter, they contain in principle much more information than the two-dimensional CMB. The challenge, however, is to extract this information, because gravitational clustering is a non-linear process. Moreover, since galaxies are expected to form only in the densest regions, they are biased tracers of the matter distribution. In this thesis we consider the possibility of using the bispectrum as a probe of primordial non-Gaussianities (see § 1.3 and § 1.4.) Another opportunity to constrain primordial non-Gaussianities comes from galaxy bias. It turns out that a non-zero squeezed bispectrum imprints a particular scale-dependence in the galaxy power spectrum [129]. The prospects are that Euclid [69] can reach $\sigma(f_{NL}^{\text{sq}}) = \mathcal{O}(5)$ [130] and the optimized proposed SPHEREx survey can reach $\sigma(f_{NL}^{\text{sq}}) = \mathcal{O}(1)$ [130]. Moreover, future intensity mapping experiments (of neutral hydrogen) such as SKA [131], are expected to be competitive $\sigma(f_{NL}^{\text{sq}}) = \mathcal{O}(0.5)$ [132].

Other probes - With the exciting discipline of gravitational-wave cosmology [133], we can search for a stochastic background of primordial gravitational waves directly. From the CMB to ground-based GW experiments, a range of scales with more than 20 orders of magnitude [134] is scanned. This allows us to constrain the tensor tilt n_t to much better accuracy than the CMB alone [134, 135]. Moreover, it constrains the amplitude of primordial gravitational waves and therefore r . Another probe is offered by the non-detection of primordial black holes [136], which constrains the integrated primordial scalar power spectrum.

1.3 Outline thesis

This thesis is about “spectroscopy” of two-field inflation. The power spectrum and bispectrum contain a wealth of information about the primordial universe. In particular, the amplitude and scale dependence of inflationary spectra might tell us something about the spectrum of masses of other fields relevant at the time of inflation, and their couplings to the inflaton.

This thesis consists of two parts. In [Part I](#) we study the phenomenology of two-field extensions of the simplest models of inflation. In [Part II](#) we investigate the utility of the bispectrum as a probe of non-Gaussianities in large scale structure experiments.

Multi-field inflation

Inflation gives us an opportunity to probe energy scales not accessible at earth based experiments. This means that inflation may shed light on the UV completion of the Standard Model and teach us something about fundamental physics. For instance, string theory typically predicts the presence of many scalar moduli fields [137] which could be active during inflation, and leave their imprint in the data. Being agnostic about the precise UV completion of inflation and the Standard Model, it is important to identify some of its essential features. For instance, when only one light degree of freedom is excited during inflation, the effective field theory of inflation [99] provides the most general way to parametrize our ignorance about the parent theory. However, we eventually would like to interpret the EFT coefficients in terms of properties of the UV embedding, such as extra dimensions, the mass and spin of other fields, higher order kinetic terms, etcetera. Therefore, in general it is important to understand how various classes of inflationary theories affect the low energy dynamics of the inflaton and its possible coupling to other light degrees of freedom.

In [Part I](#) of this thesis we will be mainly concerned with two-field inflation as a representative of the inflationary class of multi-field models of the form

$$S = \frac{1}{2} \int d^4x \sqrt{-g} \left[M_p^2 R - G_{ab}(\phi^c) \partial_\mu \phi^a \partial^\mu \phi^b - 2V(\phi^a) \right] . \quad (1.48)$$

Here $G_{ab}(\phi^c)$ is the field metric characterizing the kinetic terms. Moreover, R is the Ricci scalar of spacetime and $V(\phi^a)$ the potential energy density of the scalar fields. The action of perturbations will now contain an extra scalar degree of freedom besides the curvature perturbation: the *isocurvature perturbation*. We study a broad mass spectrum of the isocurvature perturbation,

ranging from massless all the way to heavy with respect to the Hubble scale. Our aim is to improve our understanding about the observational viability and falsifiability of such theories when the isocurvature and inflaton are coupled to each other. Depending on the coupling strength, the transfer of power from the isocurvature perturbations to the curvature perturbations is either efficient or inefficient. In [Figure 1.3](#) we illustrate the different regimes of parameter space we cover in this thesis.

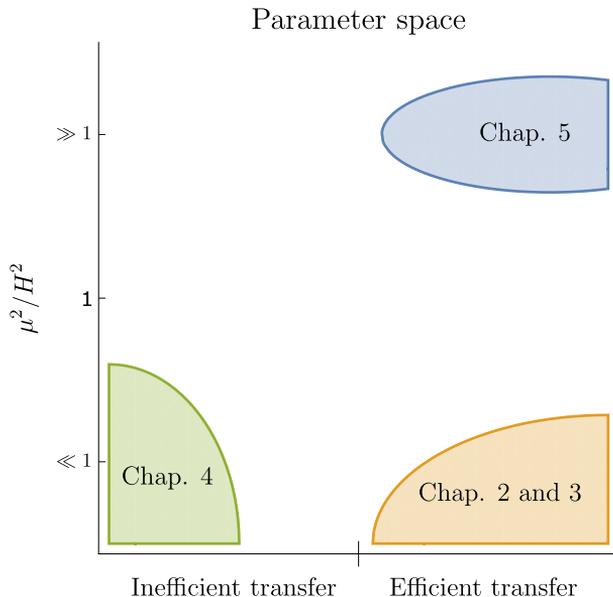


Figure 1.3: An illustration of the parameter space of two-field inflation covered in [Part I](#) of this thesis. On the vertical axis we vary the mass of the isocurvature perturbations (see [Eq. 2.13](#)) in units of the Hubble scale μ^2/H^2 . The horizontal axis represents the coupling strength between isocurvature and curvature perturbations. It is divided into two regimes. The (left) right part corresponds to (in)efficient transfer of power from the isocurvature perturbations to the curvature perturbations.

- [Chapter 2](#): the first part of this chapter serves as a review of a few selected elements of multi-field inflation. In particular, we recap the kinematical analysis of multi-field inflation and we pay special attention to the definition of the entropy mass (the effective mass of isocurvature perturbations). It turns out that the linear dynamics of two-field inflation is described by only a few kinematical and geometrical parameters. These are the field radius of curvature, the entropy mass and the Hubble slow-roll parameters.

This motivates us to introduce Orbital Inflation: a family of two-field

models of inflation where all these parameters are (approximately) constant. We discuss their phenomenology, and we highlight a few results in the regime typically not considered in quasi-single field inflation (inflation with additional fields of mass $m \sim H$). First, the larger the coupling to the isocurvature perturbations, the more they become suppressed compared to the curvature perturbations. Second, the value of the isocurvature mass determines how the predictions for n_s shift. This may help to distinguish between various entropy masses. Finally, by means of the Hamilton-Jacobi formalism we locally reconstruct potentials with exactly these properties. This allows us to numerically test our predictions. Moreover, this provides a playground for quasi-single field models of inflation, because higher order couplings can be tuned as well. As far as we know, no exact models of QSF are known.

- **Chapter 3:** we study Orbital Inflation (see [Chapter 2](#)) in the limit that the entropy mass is zero. We dub this ‘ultra-light Orbital Inflation’ because these models realize the shift symmetry described in [\[138\]](#). If the radius of curvature of the inflationary trajectory is sufficiently small, the amplitude of isocurvature perturbations and primordial non-Gaussianities are highly suppressed. These models mimic the predictions of single-field inflation, because only one degree of freedom is responsible for the observed perturbations. Inflation proceeds along an ‘angular’ isometry direction in field space at arbitrary radius and is a special case of Orbital Inflation discussed in [Chapter 2](#). We prove neutral stability of a class of exact attractor solutions.

The results in this chapter are based on joint work with Ana Achúcarro, Edmund Copeland, Oksana Iargyina, Gonzalo Palma and Dong-Gang Wang.

- **Chapter 4:** we investigate two-field cosmological α -attractors, which are characterized by a hyperbolic field metric. The important property of the single field realization of α -attractors is that, in the limit of small $\alpha < \mathcal{O}(10)$, their predictions converge to $n_s \simeq 1 - \frac{2}{N}$, $r \simeq \frac{4}{N^2}$. In the two-field case, we find that the inflationary predictions show universal behavior too, insensitive to significant modifications of the potential. In particular, when both fields are light, the multi-field effects conspire in such a way that the predictions remain unchanged with respect to the single field scenario. We emphasize the key role played by the hyperbolic field space. We also list the constraints on the potential to ensure the

validity of our results.

This chapter is based on [139]:

Universality of multi-field α -attractors, A. Achúcarro, R. Kallosh, A. Linde, D-G. Wang and Y. Welling, JCAP **1804** (2018) 07, 028, (arXiv:1711.09478 [hep-th]).

- **Chapter 5:** we illustrate the impact of heavy fields on the inflationary observables by a simple two-field embedding of a few large-field models of inflation. In our set-up the inflaton corresponds to the phase of a complex field with mildly broken $U(1)$ symmetry. This type of embedding affects the background evolution and modifies the effective sound speed of the curvature perturbation. The overall effect is that the tensor-to-scalar ratio is reduced, which improves the viability of the inflationary models under consideration.

The results in this chapter are based on [140]:

On the viability of $m^2\phi^2$ and natural inflation, A. Achúcarro, V. Atal and Y. Welling, JCAP **1507** (2015) 07, 008, (arXiv:1503.07486 [astro-ph.CO]).

Large scale structure

In [Part II](#) of this thesis we study the detectability of non-Gaussianities in near future large scale structure experiments, using the bispectrum as observable. In [§ 1.4](#) we describe how the primordial perturbations from inflation evolve into the large scale structures we see around us today. To extract information about the early universe we have to understand the relation between the distribution of galaxies and the primordial power spectrum and bispectrum. One of the ingredients is to understand the clustering of dark matter. On the largest scales this is well captured by linear perturbation theory. We review the main idea of the ‘Effective Theory of Large Scale Structure’ (EFT of LSS), a theoretical method that extends the perturbative description to quasi non-linear scales.

- **Chapter 6:** we perform a simple statistical analysis to understand whether the EFT of LSS can help us to improve the constraints on primordial non-Gaussianities in upcoming surveys such as Euclid. As a first step we focus exclusively on the matter bispectrum. Already in this simplified set-up we find that it is unlikely to reach the theoretical benchmarks quoted between [Eq. 1.42](#) and [Eq. 1.43](#). On the other hand, the EFT reduces the size of the error bars by a factor of 3 compared to standard perturbation theory (SPT) in this set-up. We put special emphasis on

the modeling of theoretical uncertainties.

The results in this chapter are based on [141]:

Lifting Primordial Non- Gaussianity Above the Noise, Y. Welling, D. van der Woude, and E. Pajer, JCAP **1608** (2016) 08, 044, (arXiv:1605.06426 [astro-ph.CO]).

1.4 Large scale structure as a probe of inflation

After inflation produced the tiny initial density inhomogeneities, they slowly evolved into the structures we see around us today. The current picture of structure formation is based on work by Zeldovich, Peebles and collaborators [33, 142–144]. It is believed that gravitational instability drives cold dark matter to evolve into large scale filaments. In this network, structures form through hierarchical clustering of small objects, such as galaxies. With our telescopes we observe light coming from these galaxies, and they trace the underlying network of dark matter. This allows us to connect the observed galaxy distribution with the dark matter distribution, which in its turn probes the initial conditions set by inflation.

However, extracting information about the primordial universe is challenging. There are several complex steps that have to be understood in order to make connection with the observational data. The first step we have to understand is the non-linear gravitational evolution of dark matter overdensities. Second, galaxies¹⁰ only form in the densest regions, and are therefore biased tracers of the dark matter distribution. Finally, there are observational complexities, such as redshift space distortions and projection effects.

Fortunately, on sufficiently large scales, gravitational evolution, biasing and redshift space distortions are controllable by perturbative methods [145–147]. Since our aim is to learn something about inflation¹¹, we are mainly interested in these large scales, where the imprint of initial conditions has not been much distorted yet. This calls for a theoretical description that captures both initial conditions and gravitational evolution by a finite number of parameters.

¹⁰Or any other tracer of dark matter.

¹¹Of course, from the mildly non-linear scales there is more to learn about cosmology than just inflation. For instance, the shape and the location of the Baryon Acoustic Oscillations (BAO) in the range $k \sim 0.05 - 0.25 h\text{Mpc}^{-1}$ provides a powerful probe of the properties of dark energy [148].

In this thesis we focus exclusively on the gravitational evolution of dark matter.

Depending on how far we can push the analytical description, large scale structure (LSS) surveys have the potential to become competitive with the CMB. The main advantage is that LSS surveys are three dimensional, and therefore they contain in principle much more information. The number of independent measurements we can extract from the CMB scales roughly as $N_{\text{CMB}} \sim (k_{\text{max}}/k_{\text{min}})^2 \sim 10^7$, whereas in LSS¹² it scales like $N_{\text{LSS}} \sim (k_{\text{max}}/k_{\text{min}})^3$. This means that if we can push k_{min} closer to the non-linear scale $k_{\text{NL}} \sim 0.2h\text{Mpc}^{-1}$ (see § 1.4.1) and map the largest volume $k_{\text{min}} \sim \frac{H}{c} \sim 10^{-4}h\text{Mpc}^{-1}$, we might even approach a factor $\mathcal{O}(10^2)$ more datapoints with respect to the CMB.

To describe the evolution of dark matter on mildly non-linear scales, we employ Eulerian perturbation theory. In this framework, dark matter is modelled as a self-gravitating effective fluid on large scales (see § 1.4.2). In its first formulation in the nineties, dark matter was assumed to take the form of a pressureless perfect fluid. This resulted in Standard Perturbation Theory (SPT), for a comprehensive review see [149]. With the increasing precision of (future) LSS surveys [68, 130, 150–152], several improvements on SPT were proposed in the following decades, see e.g. [153–165]. In particular, it was realized [159, 166–168] that SPT is limited by the assumption that dark matter can be described in terms of density and velocity perturbations only. For instance, SPT assumes a vanishing velocity dispersion. However, the velocity dispersion was shown to give percent level corrections to the predictions of SPT [166]. Therefore, this should be incorporated in the analytical description as well. More generally, we should take into account the backreaction of the short scale physics on the dynamics of the dark matter fluid.

This has led to the development of several new perturbation techniques, among which the so-called Effective Theory of Large Scale Structures (EFT of LSS) [159, 162]. The backreaction of short scale physics is captured by an effective viscous shear tensor in the fluid equations. By expanding the stress tensor in terms of the long wavelength density and velocity perturbations, the fluid equations become an effective description of dark matter, valid on large scales. The small and large scales become correlated during their joint gravitational evolution, and might also be correlated initially in the presence of primordial non-Gaussianities.

¹²This is a naive estimate of the number of independent measurements, neglecting shot-noise and cross-correlations.

In this section we review some elements of perturbation theory relevant for [Chapter 6](#). In [§ 1.4.1](#) we show the behavior of the transfer function, which captures the linear evolution of the matter density perturbation from the primordial to the late universe. This allows us to estimate the non-linear scale at which the density perturbations become order unity. In [§ 1.4.2](#) we recap how to describe dark matter as a fluid on scales larger than the non-linear scale. In particular, we pay special attention to the smoothing procedure and the appearance of an effective stress tensor. After expanding the stress tensor in terms of the long wavelength fields, we can solve the fluid equations perturbatively. In [§ 1.4.3](#) we show how to compute a non-linear correction to the power spectrum. Finally, in [§ 1.4.4](#) we discuss shortly the bispectrum, which is the topic of [Chapter 6](#).

1.4.1 Transfer Function

During most of the history of the universe the matter density perturbations were tiny $\delta_m \ll 1$. Only during the epoch of matter domination they started to grow substantially and form the non-linear structures. The linear evolution from the initial conditions from inflation to the late universe (well after matter-radiation equality) can be captured by the *transfer function*

$$\delta_{m,\mathbf{k}}(a) = \frac{2}{5\Omega_m} \frac{k^2}{a^2 H^2} T_\delta(k) \mathcal{R}_{\mathbf{k}} . \quad (1.49)$$

It is conventional to factor out some terms, such that the transfer function asymptotes to unity on large scales and only depends on k otherwise. Let's recap how the transfer function scales with k . We follow the set of lectures by Baumann [70], where everything is more rigorously derived¹³. First of all, the Poisson equation relates the evolution of the perturbed energy density to the gravitational potential as $\nabla^2 \Phi = a^2 \delta\rho / M_p^2$. Assuming we are in the epoch of matter domination or later, we can replace $\delta\rho = \bar{\rho}_m \delta_m$, because the contribution from radiation is negligible and dark energy does not cluster.

¹³In short, in [§ 1.4.1](#) we work in Newtonian gauge in conformal time $ad\tau = dt$

$$ds^2 = a^2(\tau) \left[(1 + 2\Psi)d\tau^2 - (1 - 2\Phi)\delta_{ij} dx^i dx^j \right], \quad (1.50)$$

where we assume the absence of anisotropic stress, which implies $\Psi = \Phi$. The tensor and vector modes are neglected. Here Φ corresponds to the Newtonian potential in the weak field limit. When we map the inflationary predictions to the late universe, the variables we work with have to be understood as gauge invariant variables. The gravitational potentials are identified with the Bardeen variables [169] and the curvature perturbation and density perturbation are the *comoving* curvature perturbation [86] and density perturbation Δ_m . On sub-Hubble scales $\Delta_m \approx \delta_m$.

Going to Fourier space and using the Friedmann equation [Eq. 1.8](#) we rewrite the Poisson equation as

$$\delta_{m,\mathbf{k}} = -\frac{2k^2}{3a^2\Omega_m H^2}\Phi_{\mathbf{k}}. \quad (1.51)$$

The evolution of the gravitational potential directly translates to that of δ_m . The linear dynamics of the gravitational potential turns out to be quite simple. First of all, on super Hubble scales the gravitational potential Φ is related to the curvature perturbation as follows

$$\mathcal{R} = -\frac{5+3w}{3+3w}\Phi, \quad (1.52)$$

where w is the equation of state of the background. Assuming adiabatic¹⁴ initial conditions, this implies that the gravitational potential is conserved on super Hubble scales, as long as the equation of state is constant. In the transition from radiation to matter Φ however drops by a factor of 9/10, because \mathcal{R} remains constant. The evolution of the gravitational potential on all scales is captured by the Einstein equations. It can be shown that the gravitational potential is constant throughout the matter era. If perturbations become sub-Hubble at that time we therefore have the simple relation¹⁵ $\Phi_{\mathbf{k}} = -\frac{3}{5}\mathcal{R}_{\mathbf{k}}$ which yields

$$\delta_{m,\mathbf{k}} = \frac{2}{5\Omega_m} \frac{k^2}{a^2 H^2} \mathcal{R}_{\mathbf{k}}. \quad (1.53)$$

On the other hand, if the gravitational potential enters the horizon (at $k\tau = -1$) during the radiation era it decays as τ^{-2} until matter-radiation equality. Therefore, in this case the gravitational potential receives an additional suppression $\Phi_{\mathbf{k}} \sim (k_{\text{eq}}/k)^2$. This results in the following approximate asymptotic scalings [\[114\]](#)

$$T_{\delta}(k) = 1 \quad \text{for } k \ll k_{\text{eq}}, \quad (1.54a)$$

$$T_{\delta}(k) \approx 12 \left(\frac{k_{\text{eq}}}{k}\right)^2 \ln\left(\frac{k}{8k_{\text{eq}}}\right) \quad \text{for } k \gg k_{\text{eq}}. \quad (1.54b)$$

Here the logarithmic correction reflects the logarithmic growth of matter perturbations during radiation domination. Improved theoretical fitting functions are given in [\[170, 171\]](#). Exact transfer functions can be computed numerically with CMBFAST [\[172\]](#) or CAMB [\[173\]](#). If we set initial conditions well in the

¹⁴See the discussion around [Eq. 1.46](#).

¹⁵This explains the funny normalization of f_{NL} in [Eq. 1.40](#).

epoch of matter domination, the linear dimensionless matter power spectrum is therefore given by

$$\Delta_{\delta}^2(k, \tau) = \frac{4}{25\Omega_m^2} \frac{k^4}{a^4 H^4} T_{\delta}^2(k) \Delta_{\mathcal{R}}^2(k) \quad (1.55)$$

Its exact behavior, computed using CAMB [173], is shown in Figure 1.4. The scaling of Δ_{δ}^2 agrees qualitatively with that of Eq. 1.54b. Around the scale of matter-radiation equality the power spectrum bends over from power law growth $(k/k_{eq})^4$ to logarithmic growth $\ln(k/k_{eq})$.

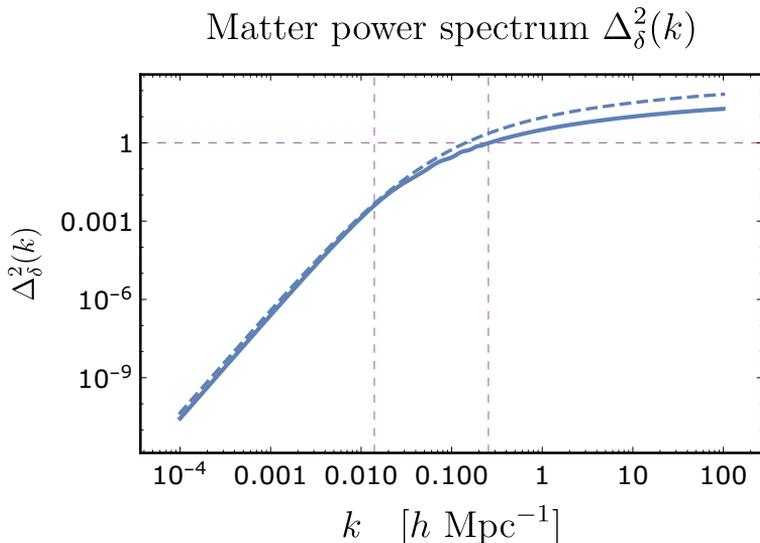


Figure 1.4: The linear dimensionless matter power spectrum today $\Delta_{\delta}^2(k, \tau_0)$ (solid line) generated with the online tool of CAMB [173]: http://lambda.gsfc.nasa.gov/toolbox/tb_camb_form.cfm. We used $\Omega_{\Lambda}^0 = 0.728$, $\Omega_b^0 = 0.046$, $\Omega_{dm}^0 = 0.226$, $\Omega_r^0 = 8.4 \cdot 10^{-5}$, $n_s = 0.967$, $h = 0.704$ and $\sigma_8 = 0.81$. We indicate the scale of matter-radiation equality $k_{eq} \approx 0.073h\Omega_m^0 h\text{Mpc}^{-1}$ [114] and the non-linear scale $k_{NL} \approx 0.25 h\text{Mpc}^{-1}$, at which $\Delta_{\delta}^2(k_{NL}, \tau_0) = 1$. Moreover, we compare with the result we get by using a transfer function that has the asymptotic limits quoted in Eq. 1.54b (dashed line). More precisely, we used $T(x \equiv \frac{k}{8k_{eq}}) = \frac{\ln(1+x)}{x} (1 + \frac{2}{3}x)^{-1}$, which we inserted into Eq. 1.55.

Non-linear scale

Note from Eq. 1.51, that, even though the amplitude of the gravitational potential is always $\Delta_{\Phi}^2(k) \sim 10^{-9} T_{\delta}^2(k)$, the amplitude of density fluctuations is much larger and grows with a^2 during matter domination, where $H^2 \sim \rho_m \sim a^{-3}$. Today the linear matter power spectrum exceeds unity at the *non-linear scale* $k_{\text{NL}} \approx 0.25 h\text{Mpc}^{-1}$. This signals the breakdown of the linear approximation, or more precisely the breakdown of perturbation theory in the density contrast δ_m . To make sensible theoretical predictions within perturbation theory, we need to consider sufficiently large scales compared to the non-linear scale. We will next see how dark matter can be modeled as an effective fluid on these scales.

1.4.2 The Dark Matter Fluid

We assume that dark matter consists of a collection of non-relativistic particles of mass m , which interact only gravitationally. They can be described by the collisionless Boltzmann equation, or Vlasov equation [174]

$$\frac{Df}{D\tau} \equiv \frac{\partial f}{\partial \tau} + \frac{\mathbf{p}}{am} \cdot \nabla f + \partial_{\tau} \mathbf{p} \cdot \frac{\partial f}{\partial \mathbf{p}} = 0. \quad (1.56)$$

Here $f(\tau, \mathbf{x}, \mathbf{p})$ is the particle number density in phase space, with $\mathbf{p} \equiv am\mathbf{v}$ the conjugate momentum to the comoving spatial coordinate \mathbf{x} of the particle. Moreover, $\mathbf{v} = \partial_{\tau} \mathbf{x}$ denotes its peculiar velocity. Notice that the conjugate momentum can be related to the spatial part of the four momentum as $p^i = \delta^{ij} P_j$. We denote the partial derivative with respect to x^i as ∇_i and raise its index with δ^{ij} . This notation is commonly used in the literature. We limit ourselves to sub-Hubble scales, such that we can use the Newtonian limit of the geodesic equation $\partial_{\tau} \mathbf{p} = -am\nabla\Phi$ and the Poisson's equation $\nabla^2\Phi = a^2\delta\rho/M_p$ (see [174, 175]).

The position space equations can be obtained by taking successive moments of the Vlasov equation and replacing

$$\frac{1}{a^3} \int d^3\mathbf{p} \, mf(\tau, \mathbf{x}, \mathbf{p}) = \rho(\tau, \mathbf{x}), \quad (1.57a)$$

$$\frac{1}{a^3} \int d^3\mathbf{p} \, \frac{p^i}{a} f(\tau, \mathbf{x}, \mathbf{p}) = \rho v^i(\tau, \mathbf{x}) \equiv \pi^i(\tau, \mathbf{x}), \quad (1.57b)$$

$$\frac{1}{a^3} \int d^3\mathbf{p} \, \frac{p^i p^j}{a^2 m} f(\tau, \mathbf{x}, \mathbf{p}) = \rho v^i v^j(\tau, \mathbf{x}) + \rho \sigma^{ij}(\tau, \mathbf{x}), \quad (1.57c)$$

etc

Here we implicitly assumed some spatial smoothing of the phase space distribution function¹⁶ in order to define the mass density ρ and mean peculiar momentum density π^i . This induces velocity dispersion σ^{ij} because there are multiple individual particle velocities in a given spatial patch. Moreover, there could already be a microscopic velocity dispersion if particle trajectories cross [168]. It has been estimated numerically [166] and theoretically [167] that velocity dispersion induces a percent level correction to large scale matter power spectrum at $k \sim 0.1h\text{Mpc}^{-1}$. This means that we have to take it into account in our theoretical description.

Taking moments of the Vlasov equation generates an infinite set of evolution equations (the *Boltzmann hierarchy*) which couple the moments of the distribution function to each other. To arrive at the fluid description, we neglect the higher moments contained in the dots of Eq. 1.57, such that the full Boltzmann hierarchy is reduced to the Euler equations. This is possible as long as we pick a smoothing scale larger than the distance that dark matter particles have travelled $1/k \gg v_p/aH$ [159, 168, 176]. Here v_p is the typical size of the peculiar velocity of the dark matter particles, and $\frac{\mathcal{O}(1)}{aH}$ represents the particle horizon since reheating. Using linear theory (see § 1.4.3) one can estimate $\Delta_v^2 \sim \frac{a^2 H^2}{k^2} \Delta_\delta^2$, and therefore, if we only consider scales larger than the non-linear scale, this suggests that the fluid description is applicable, taking $v_p \lesssim \sqrt{\Delta_v^2(k_{NL})} \sim aH/k_{NL}$. The intuitive interpretation of [159] is that dark matter has an effective mean free path of the order of the non-linear scale, because of the finite horizon induced by gravity. Higher moments are suppressed, because they did not have time to develop.

¹⁶The microscopic phase space distribution for a collection of particles is given by the Klimontovich density

$$f(\tau, \mathbf{x}, \mathbf{p}) = \sum_n \delta^{(3)}(\mathbf{p} - \mathbf{p}_{(n)}) \delta^{(3)}(\mathbf{x} - \mathbf{x}_{(n)}) . \quad (1.58)$$

To get the coarse-grained stress tensor in some comoving ball ΔV around \mathbf{x} we are effectively smoothing f . For instance, with a spherical top-hat window function

$$\langle T_{\mu\nu} \rangle(\mathbf{x}) = \int_{\Delta V} \frac{d^3 \mathbf{x}'}{\Delta V} \int \frac{\Pi_j dP_j}{\sqrt{-g}} \frac{P_\mu P_\nu}{P^0} f(\mathbf{x}', \mathbf{p}, \tau) = \frac{1}{a^4 \Delta V} \sum_n \left. \frac{P_\mu^{(n)} P_\nu^{(n)}}{P_{(n)}^0} \right|_{\mathbf{x}^{(n)} \in \Delta V} . \quad (1.59)$$

Euler Equations

Truncating the Boltzmann hierarchy at the second moment allows us to describe the dynamics of dark matter on sufficiently large scales by the continuity equation and Euler equation

$$\partial_\tau \rho + 3aH\rho + \nabla_i(\rho v^i) = 0, \quad (1.60a)$$

$$\rho \partial_\tau v^i + aH\rho v^i + \rho v^j \nabla_j v^i + \rho \nabla^i \Phi = -\nabla_j \tau_\Lambda^{ij}. \quad (1.60b)$$

The index Λ on the stress tensor τ_Λ^{ij} emphasizes that we have smoothed the Vlasov equation on a scale $\Lambda \ll k_{NL}$. The resulting long wavelength fields ρ , v and Φ variables should also carry an index ℓ (for *long*), but we dropped it to avoid cluttering of notation. Importantly, smoothing the Vlasov equation creates an *effective stress tensor* in the Euler equation. For instance, the smoothing of the part that is responsible for the term $\sim \rho v v$ induces velocity dispersion as we saw above. We will see below that a similar story applies to the term $\sim \rho \Phi$. The other terms in the Euler equations are linear in ρ or π^i . Therefore, we end up with an effective stress tensor τ_Λ^{ij} that captures the velocity dispersion and the gravitationally induced stress, and also contains contributions from the higher moments that we have neglected.

Smoothing the Vlasov Equation

We smooth the Vlasov equation because it allows us to truncate the Boltzmann hierarchy. To understand the appearance of the effective stress tensor and its structure, we follow [159, 162] and see what happens if we smooth with a Gaussian window function

$$W_\Lambda(\mathbf{x}) = \left(\frac{\Lambda}{\sqrt{2\pi}} \right)^3 e^{-\frac{1}{2}\Lambda^2 x^2} \text{ or in Fourier space } W_\Lambda(\mathbf{k}) = e^{-\frac{1}{2}\frac{k^2}{\Lambda^2}}. \quad (1.61)$$

This provides a smooth cut-off $k \sim \Lambda$ in Fourier space. Smoothed quantities are defined as

$$O_\Lambda(\mathbf{x}, \tau) \equiv \int d^3\mathbf{x}' W_\Lambda(\mathbf{x} - \mathbf{x}') O(\mathbf{x}', \tau) \quad \text{or} \quad O_\Lambda(\mathbf{k}, \tau) = W_\Lambda(\mathbf{k}) O(\mathbf{k}, \tau). \quad (1.62)$$

The goal is to express the Euler equations in terms of long wavelength fields ρ_ℓ , v_ℓ and Φ_ℓ only, such that we can perturbatively solve the fluid equations.

Notice that Eq. 1.57 naturally defines $\rho_\ell = \rho_\Lambda$, $\rho_\ell v_\ell = \pi_\Lambda$ and $\rho_\ell \sigma_\ell^{ij} = (\rho \sigma^{ij})_\Lambda$. It is what we get if we convolve the moments of the phase space distribution with the window function. This means that the long wavelength

peculiar velocity is implicitly defined as $v_\ell \equiv \frac{\pi\Lambda}{\rho_\Lambda}$. Furthermore, we define $\Phi_\ell = \Phi_\Lambda$. When smoothing the Vlasov equation we encounter the terms $(\rho\partial_i\Phi)_\Lambda$ and $(\rho v^i v^j)_\Lambda$ which are *not* the same as a products of the corresponding long wavelength fields. For instance, smoothing a bilinear quantity fg yields [159]

$$(fg)_\Lambda = f_\ell g_\ell + (f_s g_s)_\Lambda + \frac{1}{\Lambda^2} \nabla f_\ell \cdot \nabla g_\ell + \dots, \quad (1.63)$$

where the dots represent higher derivative terms. The short wavelength fields are defined as $f_s \equiv f - f_\ell$. The same can be computed for trilinears. It has been shown that with this smoothing procedure the stress tensor receives contributions of the form [159, 168, 177]

$$\tau_\Lambda^{ij} \supset - \left(\rho v_s^i v_s^j + \frac{M_p^2}{a^2} \left(\delta^{ij} \nabla_k \Phi_s \nabla^k \Phi_s - 2 \nabla^i \Phi_s \nabla^j \Phi_s \right) + \rho \sigma_s^{ij} \right)_\Lambda + \dots \quad (1.64)$$

The dots represent the derivative terms suppressed by $(k/\Lambda)^2$. These derivative terms are very small, because they contain two or more powers of the long wavelength fields, where at least one of them is v_ℓ or Φ_ℓ .

Effective Theory of Large Scale Structure

From Eq. 1.64 we see that the effective stress-tensor captures the backreaction of the short scale physics on the dynamics of dark matter on the largest scales. By construction, we cannot evaluate Eq. 1.64 within perturbation theory. Therefore, the next step is to parameterize our ignorance of the small scales by expanding the stress tensor in terms of the long wavelength fields in the most generic way allowed by symmetries. This is where the so-called Effective Theory of Large Scale Structure (EFT of LSS) [159, 162] comes in.

What kind of contributions do we expect from the microscopic theory? Let's zoom in on a random small comoving box of particles of size $\sim 1/\Lambda^3$ which evolves in a background of a long wavelength gravitational potential Φ_ℓ . The worldlines of the particles inside the box are affected by long wavelength tidal fields and this changes the distribution of velocities. They also experience a common acceleration by the gradient of the gravitational potential, but this is locally unobservable. Averaging over many of these boxes, given a realization of the long wavelength perturbations, this results in a shear tensor correlated with the tidal field [159]

$$\langle (\rho v_s^i v_s^j)_\Lambda | \Phi_\ell \rangle = \langle (\rho v_s^2)_\Lambda \rangle_0 \left(c_1 \delta_{ij} + c_2 \frac{\nabla_i \nabla_j \Phi_\ell}{a^2 H^2} + \dots \right). \quad (1.65)$$

The first term contributes to the isotropic pressure, which is generated even in the absence of long wavelength perturbations. The second term is an example

of how the short scale physics feels the background of long modes in which they evolve, and how this is memorized in their backreaction.

More generally, one can write down all operators on the right hand side of the Euler equation that are allowed by the equivalence principle [159, 162, 178]. Local observables can only be affected by tidal forces $\nabla_i \nabla_j \Phi$ (or equivalently the shear $\nabla_i v_j$ [179]) and their spatial derivatives. To linear order in perturbations, up to higher derivative terms, this results in the Navier-Stokes equations with an average stress-tensor

$$\langle (\tau_{ij})_\Lambda | \delta_\ell, \dots \rangle = p(\tau) \delta^{ij} + \rho \left[c_s^2(\Lambda, \tau) \delta_\ell \delta^{ij} - c_{bv}^2(\Lambda, \tau) \frac{\nabla_k v_\ell^k}{aH} \delta^{ij} - \frac{3}{2} c_{sv}^2(\Lambda, \tau) \frac{\nabla^{(i} v_\ell^{j)}) - \frac{1}{3} \delta^{ij} \nabla_k v_\ell^k}{aH} \right] \quad (1.66)$$

Here the contraction $\nabla^2 \Phi_\ell$ has been replaced by δ_ℓ , using the Poisson equation. It has been shown that the induced pressure $p(\tau) \sim 10^{-5}$ is tiny and has a negligible effect on the dynamics of the background [159], so we will ignore it further. For some observables a higher precision is required, and one needs to go beyond the linear expansion [180–182]. Moreover, when averaging over an ensemble of realizations of small domains, like we did in Eq. 1.65, we assumed that short scales only get correlated with the long modes through non-linear gravitational evolution. This is true as long as the initial short scale fluctuations are independent of the long wavelength perturbations. However, in the presence of primordial non-Gaussianities the short modes are correlated with the long wavelength perturbations through their dependence on the realization of the primordial potential [183]. The short scales memorize these initial conditions and this affects their backreaction on the evolution of the long modes.

The resulting EFT coefficients, such as $c_s^2(\Lambda, \tau)$, explicitly depend on the unphysical smoothing scale Λ . This is important, because if we compute observable quantities, such as the power spectrum, we will at the same time encounter loop integrals that are bounded by the scale Λ . For a well-defined perturbation theory, these dependencies on Λ should cancel. We will see that this is the case for the simplest contribution to the power spectrum in § 1.4.3. Internal consistency of the EFT of LSS has been demonstrated explicitly for the power spectrum up to two loop [178, 180, 184], for the one loop bispectrum [181–183] and eventually to all orders in perturbation theory in [185].

Scalings with k/k_{NL}

Let's find out the relative importance of terms in the Euler equation by comparing their scaling with k/k_{NL} . For simplicity we consider scales $k_H < k < k_{\text{eq}}$ such that $\Phi \sim 10^{-5}$. On smaller scales we saw in § 1.4.1 that its amplitude is even more suppressed. First of all, at linear scales we have (using the results from § 1.4.1 and § 1.4.3

$$\Delta_\delta^2 \sim \left(\frac{k}{aH}\right)^4 \Delta_\Phi^2 \quad \text{and} \quad \Delta_v^2 \sim \left(\frac{k}{aH}\right)^2 \Delta_\Phi^2 \quad \text{with} \quad , \quad (1.67)$$

indicating that velocities and the gravitational potential remain small even when Δ_δ^2 gets close to unity. This allows us to compare the various terms in the Euler equation, using that $\delta \sim (k/k_{\text{NL}})^2$ and $\frac{\nabla\Phi}{aH} \sim v \sim \sqrt{\delta\Phi}$ [159, 162]

- The non-linear terms on the left-hand side of the Euler equation Eq. 1.60b scale like δ compared to the friction term

$$\frac{v\nabla v}{aHv} \sim \frac{\delta\nabla\Phi}{aHv} \sim \delta \quad (1.68)$$

- The linear terms in the stress tensor

$$c_s^2 \frac{\nabla\delta}{aHv} \sim c_v^2 \frac{\nabla\nabla v}{a^2 H^2 v} \sim c^2 \frac{\delta}{\Phi}, \quad (1.69)$$

where the coefficients are a measure of kinetic and gravitational stress induced by the short scales. For instance, in the example of Eq. 1.65, we would get $c^2 \sim \langle v_s^2 \rangle$. At the non-linear scale we expect the velocity to approach $v^2 \sim \Phi$. The same applies to virialized scales, where the kinetic and gravitational energy are of the same order. Therefore, a rough estimate is $c^2 \sim \Phi \sim 10^{-5}$. This suggests that the leading terms in the stress tensor scale like δ as well.

- Higher order corrections to the stress tensor come with additional powers of derivatives $(k/k_{\text{NL}})^2$ and long wavelength fields, so they are additionally suppressed.

1.4.3 Perturbation Theory

With all the ingredients at hand we can do perturbation theory in the long wavelength perturbations ρ_ℓ , \mathbf{v}_ℓ and Φ_ℓ . For that purpose it is convenient to define the long wavelength density contrast, velocity divergence and vorticity as

$$\delta_\ell \equiv \frac{\rho_\ell - \bar{\rho}}{\bar{\rho}}, \quad \theta_\ell \equiv \nabla \cdot \mathbf{v}_\ell, \quad \text{and} \quad \mathbf{w}_\ell \equiv \nabla \times \mathbf{v}_\ell, \quad (1.70)$$

respectively. For notational convenience we drop the index ℓ from now on. Using Eq. 1.66, the leading order source term of the Euler equation Eq. 1.60b is given by

$$\frac{1}{\rho} \nabla_j \tau_{\Lambda}^{ij} \approx c_s^2 \nabla^i \delta - \frac{4c_{bv}^2 + c_{sv}^2}{4aH} \nabla^i \theta - \frac{3c_{sv}^2}{4aH} \nabla^2 v^i + J^i, \quad (1.71)$$

with $J^i = \frac{1}{\rho} \nabla_j \Delta \tau^{ij}$ a stochastic term to account for the fact that the stress tensor is not precisely equal to its ensemble average over the short scales, i.e. $(\tau^{ij})_{\Lambda} = \langle (\tau^{ij})_{\Lambda} | \delta_{\ell}, \dots \rangle + \Delta \tau^{ij}$. After subtracting the background evolution of $\bar{\rho}$ using Eq. 1.2b we obtain equations for the density contrast, velocity divergence and vorticity. First of all, the equation of motion for vorticity reads [177]

$$\left(\partial_{\tau} + aH - \frac{3c_{sv}^2}{4aH} \nabla^2 \right) \mathbf{w} = \nabla \times (\mathbf{v} \times \mathbf{w} - \mathbf{J}). \quad (1.72)$$

In the absence of the terms on the right hand side, vorticity decays [149] (assuming $c_{sv}^2 > 0$). If it is not present initially it may be generated by the stochastic term, or by the higher order terms in the effective stress tensor [177]. However, numerical studies [166] find a negligible amount of vorticity on the larger scales, which gives less than a 0.01% correction to the matter power spectrum at $k \sim 0.1 h \text{Mpc}^{-1}$, so we neglect it.

The Euler equations Eq. 1.60 simplify considerably and in Fourier space they read

$$\partial_{\tau} \delta + \theta = S_{\alpha}, \quad (1.73a)$$

$$(\partial_{\tau} + aH) \theta + \frac{3}{2} \Omega_m a^2 H^2 \delta = S_{\beta} + c_s^2 k^2 \delta - c_v^2 \frac{k^2 \theta}{aH} - J, \quad (1.73b)$$

where the two viscosity coefficients are combined into one $c_v^2 = c_{sv}^2 + c_{bv}^2$ and the stochastic term is contracted to the scalar $J = \frac{1}{\rho} \nabla_i \nabla_j \Delta \tau^{ij}$. The SPT source terms are given by [149]

$$S_{\alpha}(\mathbf{k}, \tau) = - \int_{\mathbf{p}} \frac{\mathbf{p} \cdot \mathbf{k}}{p^2} \theta(\mathbf{p}, \tau) \delta(\mathbf{k} - \mathbf{p}, \tau), \quad (1.74a)$$

$$S_{\beta}(\mathbf{k}, \tau) = - \int_{\mathbf{p}} \frac{k^2 \mathbf{p} \cdot (\mathbf{k} - \mathbf{p})}{p^2 (\mathbf{k} - \mathbf{p})^2} \theta(\mathbf{p}, \tau) \theta(\mathbf{k} - \mathbf{p}, \tau). \quad (1.74b)$$

The equations can be solved perturbatively using Green's functions, and the solution can be written in terms of a power series of the initial density contrast $\delta_1(\mathbf{k}) = \delta(\mathbf{k}, \tau_{\text{in}})$

$$\delta(\mathbf{k}, \tau) = \sum_n \delta^{(n)}(\mathbf{k}, \tau), \quad \theta(\mathbf{k}, \tau) = \sum_n \theta^{(n)}(\mathbf{k}, \tau). \quad (1.75)$$

The n -th order solution receives SPT, viscosity and noise contributions (and mixing terms, but they do not appear at the order we consider now)

$$\delta^{(n)}(\mathbf{k}, \tau) = \delta_{\text{SPT}}^{(n)}(\mathbf{k}, \tau) + \delta_c^{(n)}(\mathbf{k}, \tau) + \delta_J^{(n)}(\mathbf{k}, \tau) \quad (1.76)$$

The linear solutions are obtained by setting the right-hand side of Eq. 1.73 to zero, i.e. $\delta^{(1)}(\mathbf{k}, \tau) = \delta_{\text{SPT}}^{(1)}(\mathbf{k}, \tau)$. The contributions from the stress-tensor are treated as a next-to-leading-order correction, since they are derivatively suppressed.

SPT Solution

The SPT solutions are obtained by putting the effective stress tensor to zero. To a remarkable good approximation they are of the separable form [149]

$$\delta_{\text{SPT}}^{(n)}(\mathbf{k}, \tau) \approx D_1^n(\tau)\delta_n(\mathbf{k}), \quad (1.77a)$$

$$\theta_{\text{SPT}}^{(n)}(\mathbf{k}, \tau) \approx -aHf(\tau)D_1^n(\tau)\theta_n(\mathbf{k}). \quad (1.77b)$$

Here $D_1(\tau)$ is the linear growth factor and $f(\tau)$ its logarithmic derivative with respect to the scale factor. In a universe filled with matter and dark energy only ($\Omega_m + \Omega_\Lambda = 1$) the growing mode solutions are given by [186–188]

$$D_1(a) = \frac{5\Omega_m^0}{2}H_0^2H(a) \int_{a_{\text{in}}}^a \frac{d\tilde{a}}{\tilde{a}^3H^3(\tilde{a})} \quad \text{and} \quad f(a) \equiv \frac{d \ln D_1(a)}{d \ln a} \approx \Omega_m^{5/9}(a). \quad (1.78)$$

Furthermore, the initial conditions δ_n and θ_n are convolutions of multiple δ_1 's with the SPT kernel functions F_n and G_n (please see [149] for their explicit form)

$$\delta_n(\mathbf{k}) = \int_{\mathbf{p}_1} \cdots \int_{\mathbf{p}_n} (2\pi)^3 \delta_D^{(3)}(\mathbf{k} - \mathbf{p}_{1\dots n}) F_n(\mathbf{p}_1, \cdots, \mathbf{p}_n) \delta_1(\mathbf{p}_1) \cdots \delta_1(\mathbf{p}_n), \quad (1.79a)$$

$$\theta_n(\mathbf{k}) = \int_{\mathbf{p}_1} \cdots \int_{\mathbf{p}_n} (2\pi)^3 \delta_D^{(3)}(\mathbf{k} - \mathbf{p}_{1\dots n}) G_n(\mathbf{p}_1, \cdots, \mathbf{p}_n) \delta_1(\mathbf{p}_1) \cdots \delta_1(\mathbf{p}_n). \quad (1.79b)$$

Remember that we are integrating over the long wavelength fields only, therefore each $\delta_1(p)$ implicitly carries a factor $W_\Lambda(p)$.

Notice that we can express the linear solution for the velocity field in terms of the gravitational potential. First, we saw in Eq. 1.72 that the vorticity decays on linear scales. This means that the linearized Euler equation must give a solution of the form $v^i \sim \nabla^i \Phi$. More precisely, using the SPT results from above we find

$$v_{(1)}^i = \frac{2f(\tau)}{3aH\Omega_m} \nabla^i \Phi. \quad (1.80)$$

Therefore, at the linear order the tidal tensor and shear tensor are proportional to each other.

EFT Solution

To leading order the EFT contributions coming from the stress-tensor are given by

$$\delta_c^{(1)}(\mathbf{k}, \tau) = -\xi(\tau)k^2\delta_{\text{SPT}}^{(1)}(\mathbf{k}, \tau), \quad (1.81a)$$

$$\delta_J^{(1)}(\mathbf{k}, \tau) = N(\mathbf{k}, \tau). \quad (1.81b)$$

Here $\xi(\tau)$ is a time integral over the Green's function corresponding to the Euler equations [Eq. 1.73](#) convolved with a combination of the viscosity parameters $c_s^2(\tau)$ and $c_v^2(\tau)$. The viscosity contribution correlates with the long wavelength density contrast, whereas the noise term is uncorrelated.

Power Spectrum

Using the ingredients above we can compute the contributions to the matter power spectrum

$$\Delta_\delta^2(k, \tau) = \frac{k^3}{2\pi^2} \langle \delta(\mathbf{k}, \tau) \delta(-\mathbf{k}, \tau) \rangle' = \Delta_{11}^2 + \Delta_{13}^2 + \Delta_{22}^2 + \Delta_{1c}^2 + \Delta_{JJ}^2 + \dots, \quad (1.82)$$

where $\Delta_{mn}^2(k, \tau) \equiv \frac{k^3}{2\pi^2} \left(\langle \delta_{\text{SPT}}^{(m)}(\mathbf{k}, \tau) \delta_{\text{SPT}}^{(n)}(-\mathbf{k}, \tau) \rangle' + \text{perm.} \right)$, and the other two contributions come from contractions of $\delta_{\text{SPT}}^{(1)}$ with δ_c , and δ_J with itself, respectively. The dots represent higher order terms that we have neglected. Furthermore, in this example we have assumed Gaussian initial conditions, so after Wick contracting initial density contrasts (the δ_1 's), terms such as Δ_{12}^2 will vanish.

Notice that all the terms on the right-hand side of [Eq. 1.82](#) individually depend on the unphysical smoothing scale Λ , because they are constructed from the long wavelength density fields. The Λ dependence should cancel if we add all terms, otherwise the perturbative description makes no sense. Fortunately, this can be ensured. To see how this works, let's consider the 'renormalization' of the sum of Δ_{13}^2 and Δ_{1c}^2 . The renormalization of $\Delta_{22}^2 + \Delta_{JJ}^2$ proceeds in a similar way. The two individual contributions are given by

$$\Delta_{13}^2(k, \Lambda) = 6\Delta_{11}^2(k, \Lambda) \int_{\mathbf{q}} F_3(\mathbf{k}, \mathbf{q}, -\mathbf{q}) P_{11}(q, \Lambda), \quad (1.83a)$$

$$\Delta_{1c}^2(k, \Lambda) = -\xi(\tau, \Lambda)k^2\Delta_{11}^2(k, \Lambda). \quad (1.83b)$$

For notational convenience we suppressed the τ argument of the power spectra. It is understood that $\Delta_{11}^2(k, \Lambda) = D_1^2(\tau)W_\Lambda^2(k)\Delta_{\text{in}}^2(k)$. The initial conditions are set well in the epoch of matter domination, such that the assumption $\Omega_m + \Omega_\Lambda = 1$ applies. Let's switch from Gaussian smoothing to imposing

a hard cut-off, where the window function is replaced by a heaviside step function that becomes zero at the scale Λ . This means we can split $\Delta_{13}^2(k, \Lambda)$ and $\Delta_{1c}^2(k, \Lambda)$ in a cut-off dependent and a cut-off independent part

$$\Delta_{13}^2(k, \Lambda) = 6\Delta_{11}^2(k) \left(\int_{\mathbf{q},0}^{\infty} F_3(\mathbf{k}, \mathbf{q}, -\mathbf{q}) P_{11}(q) - \int_{\mathbf{q},\Lambda}^{\infty} F_3(\mathbf{k}, \mathbf{q}, -\mathbf{q}) P_{11}(q) \right), \quad (1.84a)$$

$$\Delta_{1c}^2(k, \Lambda) = -(\xi_{\text{ph}}(\tau) + \xi_{\infty}(\tau, \Lambda)) k^2 \Delta_{11}^2(k). \quad (1.84b)$$

For large wavenumbers $q \gg k$ the SPT kernel F_3 scales like $F_3(\mathbf{k}, \mathbf{q}, -\mathbf{q}) \sim \frac{k^2}{q^2}$ [189]. This ensures that we can indeed cancel the cut-off dependence with the EFT coefficient

$$\xi_{\infty}(\tau, \Lambda) = -6k^{-2} \int_{\mathbf{q},\Lambda}^{\infty} F_3(\mathbf{k}, \mathbf{q}, -\mathbf{q}) P_{11}(q), \quad (1.85)$$

as long as we pick $\Lambda^2 \gg k^2$. The final result is therefore independent of the cut-off scale and we are left with a physical coefficient $\xi_{\text{ph}}(\tau)$, that has to be fitted to observational data or numerics

$$\Delta_{13}^2(k, \Lambda) + \Delta_{1c}^2(k, \Lambda) = \Delta_{11}^2(k) \left(\xi_{\text{ph}}(\tau) k^2 + 6 \int_{\mathbf{q},0}^{\infty} F_3(\mathbf{k}, \mathbf{q}, -\mathbf{q}) P_{11}(q) \right) \quad (1.86)$$

Of course we do not necessarily need to integrate up to infinity. We could also integrate up to the non-linear scale k_{NL} , for instance, and this automatically redefines what we mean by $\xi_{\text{ph}}(\tau)$. However, from a computational perspective it does not matter what integration limit we choose, considering that the integral converges in our universe.

Finally, we should mention that the time dependence of the physical EFT coefficients is not known. This is not a problem necessarily, it just means that they have to be fitted at every redshift separately. On the other hand, this might be a bit too unrestrictive. In the limit of an Einstein-de-Sitter universe ($\Omega_m = 1$) with power law initial conditions $\Delta_{\delta, \text{in}}^2 \sim k^{n+3}$, the time dependence of the EFT coefficients is fixed by symmetry [184]. For instance, the time dependence of ξ_{ph} is given by $\xi_{\text{ph}} \sim D_1(\tau)^{\frac{1-n}{n+3}}$. Therefore, one may also consider to make an ansatz for its time dependence $\xi_{\text{ph}}(\tau) \sim D_1(\tau)^p$ for some value of p .

1.4.4 Bispectrum

In [Chapter 6](#) we focus on the bispectrum as observable to constrain primordial non-Gaussianities. One can apply the same machinery as in [§ 1.4.3](#) to find all contributions to the bispectrum up to a given order. The gravitational induced bispectrum with Gaussian initial conditions has been computed up to one loop in [\[190\]](#) (SPT) and [\[181, 182\]](#) (EFT). The bispectrum coming from primordial non-Gaussianities has been computed to one-loop in [\[191\]](#) (SPT) and [\[183\]](#) (EFT). Schematically, the perturbative theoretical prediction for the bispectrum is given by

$$B^{\text{th}} = B_{\text{SPT}}^{\text{G}} + B_{\text{EFT}}^{\text{G}} + f_{\text{NL}} (B_{\text{SPT}}^{\text{NG}} + B_{\text{EFT}}^{\text{NG}}). \quad (1.87)$$

With the introduction of primordial non-Gaussianities we are doing a double expansion in f_{NL} and k/k_{NL} . Therefore, the relative importance of the various contributions to the bispectrum is harder to assess. Moreover, the bispectrum is a three dimensional function. The one-loop correction to the primordial signal might be more relevant in the squeezed configuration than, say, in the equilateral configuration.

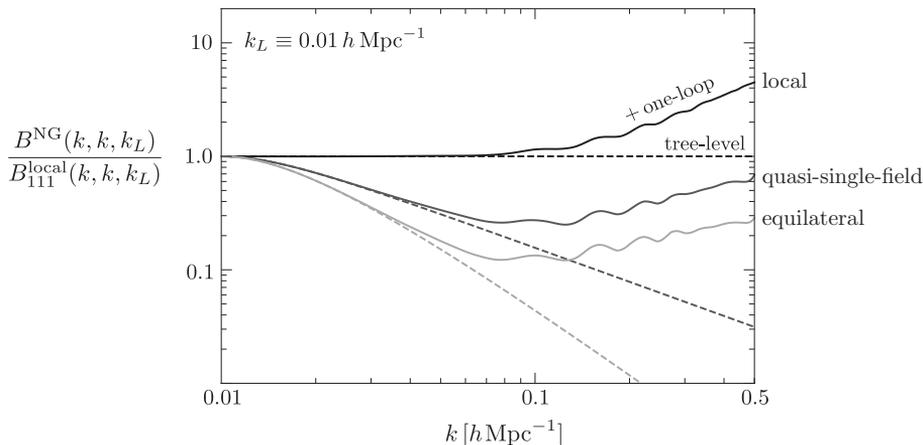


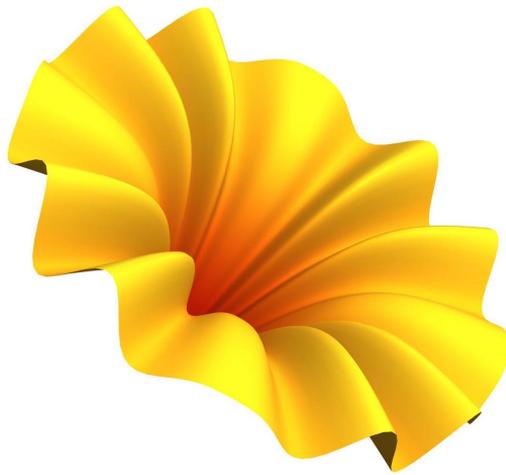
Figure 1.5: Spectroscopy: comparison of the SPT one-loop corrected non-Gaussian bispectra (solid lines) for local, quasi-single-field and equilateral type of primordial non-Gaussianities. The dashed lines show the linearly evolved primordial bispectra. The gravitational distortions tend to decrease the difference between the various shapes. Moreover, the loop corrections kick in on relatively large scales. This Figure is taken from [\[183\]](#).

The gravitational distortion (one-loop SPT) of the primordial non-Gaussian signal is shown in [Figure 1.5](#). We plot three different types of primordial non-Gaussianities, the local, equilateral and quasi-single-field templates defined in

Eq. 1.40, Eq. 1.41 and Eq. 1.42 (with $\nu = 1/2$), respectively. Figure 1.5 shows that the loop corrections become important already on relatively large scales. This was for us the main motivation to compute the EFT corrections to the one-loop bispectrum [183]. However, the gravitational distortions tend to decrease the difference between the three shapes. Therefore, we perform a simple statistical analysis in Chapter 6 to understand how much the EFT of LSS improves the modeling of the matter bispectrum, with the purpose of constraining primordial non-Gaussianities.

Deel I

Two-Field Inflation



Orbital Inflation

The purpose of this chapter is two-fold. First of all, it serves as a short introduction to multi-field inflation. Second, we introduce Orbital Inflation as one of the simplest way to extend the single field scenario.

We start with a brief recap of the linear perturbation analysis of multi-field inflation in § 2.2. If we consider more than one field, the comoving curvature perturbation interacts with *isocurvature perturbations* (quanta of the other fields). It turns out that the linear dynamics of multi-field inflation can be described in terms of a few kinematical and geometrical parameters. In case of two-field inflation, these are the field *radius of curvature*, the effective mass of isocurvature perturbations (*entropy mass*) and the Hubble slow-roll parameters. This motivates us to introduce Orbital Inflation: a family of two-field models of inflation where all these parameters are (approximately) constant. We discuss their phenomenology in § 2.3. Finally, by means of the Hamilton-Jacobi formalism we locally reconstruct potentials in § 2.4 that admit Orbital Inflation. This allows us to numerically test our predictions.

2.1 Introduction

What can we learn about inflation from the density perturbations that we observe in the sky? In case inflation is driven by a single light degree of freedom, the observational data can be described in terms of the Hubble slow roll parameters, the sound speed of the curvature perturbation, and its self-couplings. This suggests we should classify inflationary models by the behavior of these parameters, rather than by their potential. Essentially, this is what the effective field theory of single field inflation [162] does. If another light degree of freedom is active during inflation, we additionally may hope to infer its mass, its coupling to the inflaton and its self-interactions [192].

We eventually hope to come closer to the understanding of what drives inflation. Therefore, it is important to give the parameters that describe the data an interpretation in terms of the essential features of the UV embedding

of inflation. For instance, the data coefficients may inform us about the presence of extra dimensions or interactions with other fields with a certain spin, see e.g. [193–195]. For that reason, it is important to understand how various classes of inflationary theories affect the low energy dynamics of the curvature perturbation, possibly coupled to other light degrees of freedom.

In the context of multi-field inflation, the observational parameters are understood to be of kinematical and geometrical origin. In particular, in two-field inflation the dynamics of the inflaton coupled to the isocurvature perturbations are, to linear order, described in terms of the *radius of curvature* of the inflationary trajectory, the mass of the isocurvature perturbation (*entropy mass*) and the Hubble slow-roll parameters. This motivates us to introduce *Orbital Inflation*: a family of models in which the inflationary trajectory has a constant radius of curvature. Moreover, we take the entropy mass to be approximately constant. From the observational point of view, this is the simplest set-up to consider beyond single field.

Orbital Inflation is an attempt to realize ‘spontaneous symmetry probing’ [196] in the context of inflation. Moreover, it provides a realisation of quasi-single field inflation [100, 101]. We complement these studies in the following ways:

- We focus on the phenomenology in the regime of small entropy mass and relatively small radius of curvature. In particular, we allow the coupling strength to become much larger than typically considered in quasi-single field inflation (while staying in the perturbative regime.) Two results we would like to emphasize are:
 - The entropy mass dictates how the predictions for n_s change. For a decreasing radius of curvature, the value of r gets reduced. However, the shift in n_s depends on the value of the entropy mass. The predictions for (n_s, r) shift downwards and fan-out. This may help to distinguish the value of the entropy mass.
 - The larger the coupling to the isocurvature perturbations, the more they become suppressed relative to the curvature perturbations [138]. We discuss how the ratio of the power spectra at the end of inflation depends on both the entropy mass and the radius of curvature.
- We *locally* reconstruct potentials which provide exactly the kinematical properties of orbital inflation by means of a generalization of the Hamilton-Jacobi formalism [74, 75, 77, 78]. This reconstruction method

allows us to test our analytical predictions numerically. Moreover, this reconstruction method provides a playground for any quasi-single field model of inflation [100, 101] or spiral inflation [197], for which, to our knowledge, no exact models are known. Higher order couplings can be tuned as well. Interestingly, a vanishing entropy mass is reflected by a shift symmetry in the Hubble parameter rather than the resulting potential.

In this chapter we briefly review the kinematical analysis of multi-field inflation in § 2.2. This allows us to describe the phenomenology of Orbital Inflation in § 2.3. Finally, we show how to construct exact models of Orbital Inflation in § 2.4.

Throughout this chapter we work in Planck units $\hbar = c = 1$ and the reduced Planck mass is given by $M_p = (8\pi G)^{-1/2}$.

2.2 Kinematical analysis of multi-field inflation

In § 1.2 we have seen that the observable predictions of the simplest models of inflation¹ are, besides an overall energy scale, fully characterized by the kinematical Hubble slow-roll parameters

$$\epsilon \equiv -\frac{\dot{H}}{H^2}, \quad \epsilon_{n+1} \equiv \frac{\dot{\epsilon}_n}{H\epsilon_n} \quad \text{for } n \geq 2. \quad (2.1)$$

Here it is understood that $\epsilon = \epsilon_1$ and moreover, we will denote $\eta = \epsilon_2$ and $\xi = \epsilon_3$. To second order in perturbation theory we can characterize the data by means of power spectra and bispectra of the scalar and tensor perturbations. Canonical single field inflation predicts

$$r = 16\epsilon, \quad n_s = 1 - 2\epsilon - 2\eta, \quad n_t = -2\epsilon, \quad f_{NL}^{\text{sq, obs}} = 0. \quad (2.2)$$

Here r is the tensor-to-scalar ratio, n_s and n_t are the spectral indices of the scalar and tensor power spectrum respectively and $f_{NL}^{\text{sq, obs}}$ is the observable [109] amplitude of the bispectrum in the squeezed limit. For conciseness we don't show other predictions, such as the amplitude of scalar perturbations and the amplitude of the bispectrum in other configurations. Importantly, if we

¹With the simplest models of inflation we mean canonical single-field slow-roll inflationary models minimally coupled to gravity, and with Bunch Davies initial conditions.

were to detect a violation of the single field consistency relation [81, 105, 109] $f_{NL}^{\text{sq, obs}} = 0$, this would signal the presence of additional degrees of freedom.

To understand how the inflationary predictions Eq. 2.2 are affected if more degrees of freedom are active during inflation, we turn to the kinematical analysis of multi-field inflation, following the formalism developed in [198–201] (see also [202–204] and [205] for a comparative study). We identify the relevant kinematical parameters that characterise the dynamics of perturbations. For simplicity we restrict ourselves to linear perturbation theory. We start by assuming some general structure, namely minimal coupling to Einstein gravity, a field space, and as scalar field potential. In other words, we study models of multi-field inflation of the general form

$$S = \frac{1}{2} \int d^4x \sqrt{-g} \left[M_p^2 R - G_{ab} \partial_\mu \phi^a \partial^\mu \phi^b - 2V(\phi^a) \right]. \quad (2.3)$$

Here G_{ab} is the field metric characterizing the kinetic terms. Moreover, R is the Ricci scalar of spacetime and $V(\phi^a)$ the potential energy density of the scalar fields. Our aim is to eventually arrive at the perturbation equations in terms of kinematical quantities only. Furthermore, one also expects the appearance of geometrical parameters that characterise the field space.

2.2.1 Background dynamics

The background dynamics of the scalar fields follows from assuming a homogeneous, isotropic and flat Friedmann Lemaitre Robertson Walker spacetime $ds^2 = -dt^2 + a^2(t)dx^2$. The field equations and Friedmann equations are given by [198]

$$D_t^2 \phi^a + 3H D_t \phi^a + V^a = 0, \quad (2.4a)$$

$$3H^2 M_p^2 = \frac{1}{2} G_{ab} \dot{\phi}^a \dot{\phi}^b + V(\phi^a), \quad (2.4b)$$

$$\dot{H} M_p^2 = \frac{1}{2} G_{ab} \dot{\phi}^a \dot{\phi}^b, \quad (2.4c)$$

respectively, where $D_t \equiv \dot{\phi}^a \nabla_a$, with ∇_a the covariant field derivative with respect to the field metric. Moreover, the latin field indices are raised and lowered with the field metric. The last equation is not independent from the first two, but nevertheless useful later, when we construct exact models of Orbital Inflation in § 3.3.

Kinematical basis in field space

The inflationary background trajectory determines a natural basis of unit vectors in field space [199, 201, 206], which are defined iteratively²:

$$T^a = \frac{\dot{\phi}^a}{\dot{\phi}}, \quad (2.5a)$$

$$D_t T^a = -\Omega N^a, \quad (2.5b)$$

$$D_t N_j^a = \Omega_j N_{j-1}^a - \Omega_{j+1} N_{j+1}^a \quad \text{for } j \geq 1 \quad \text{with } N_i^a (N_i)_a = 1 \quad \forall i. \quad (2.5c)$$

Please see Figure 2.1 for an illustration. The first vector T^a is the tangent pointing along the inflationary trajectory. The other ones are the normal vectors, which are all normalized to unity. This uniquely determines the values of the turn rates Ω_i up to a sign. Furthermore, it is understood that $N_0^a = T^a$, $N^a = N_1^a$ and $\Omega = \Omega_1$. Finally, $\dot{\phi} \equiv \sqrt{G_{ab} \dot{\phi}^a \dot{\phi}^b}$ is the *proper field velocity*.

In § 2.3 and § 2.4 we will work with the field *radius of curvature* of the inflationary trajectory

$$\kappa \equiv \dot{\phi} (N_a D_t T^a)^{-1}. \quad (2.6)$$

Notice that this is related to the turn rate as $\kappa = -\frac{\sqrt{2\epsilon} M_p H}{\Omega}$.

Geodesics and turns

Using the tangent vector instead of $\dot{\phi}^a$, the field equations of motion can be written as

$$D_t T^a + \frac{V^a - V_T T^a}{\dot{\phi}} = 0. \quad (2.7)$$

This shows explicitly that the field trajectory is a geodesic only if $V^a = V_T T^a$, because in that case the tangent vector T^a is parallel transported, i.e. $D_t T^a \sim T^b \nabla_b T^a = 0$. Moreover, the deviation of a geodesic is parameterized by the turn rate Ω defined above. If we consider only two fields, we may choose N^a to have a fixed orientation with respect to the tangent vector, namely $N^a = \epsilon^{ab} T_b$. Then Ω will flip sign each time the inflationary trajectory changes from turning clockwise to anti-clockwise or vice versa. In the multi-field scenario it is perhaps more convenient to assign a definite sign to all Ω_i .

²This procedure becomes ill-defined as soon as a turn rate Ω_i becomes zero. In that case we can choose the remaining normal vectors N_n^a , for $n \geq i$, as any orthogonal normal set of vectors (and $\Omega_n = 0$).

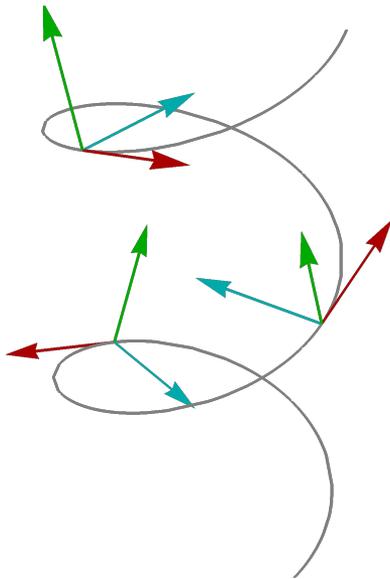


Figure 2.1: Illustration of an inflationary trajectory in a three-dimensional field space and the corresponding tangent (red), normal (cyan) and binormal (green) vector. In this example the inflationary trajectory is along a helix, and both turn rates Ω_1 and Ω_2 are non-zero.

Tangent and normal projections of the field equations

Projecting the equations of motion along the kinematical basis vectors we find

$$\ddot{\varphi} + 3H\dot{\varphi} + V_T = 0, \quad (2.8a)$$

$$V_N = \Omega\dot{\varphi}, \quad (2.8b)$$

$$V_{N_i} = 0 \quad \text{for } i > 1. \quad (2.8c)$$

This is an alternative way of writing the field equations [Eq. 2.4a](#). The first equation is the same as in single field inflation, and it suggests we should identify φ with the inflaton field. The remaining projections express a condition on the gradient of the potential to sustain centrifugal motion.

Kinematical expression for the gradient of the potential

The background equations allow us to write the gradient of the potential in terms of the slow roll parameters and the first turn rate. The Hubble slow roll parameters are defined in Eq. 2.1. Using Eq. 2.4, the first three slow-roll parameters can be written as

$$\epsilon \equiv -\frac{\dot{H}}{H^2} = \frac{\dot{\varphi}^2}{2H^2 M_p^2}, \quad (2.9a)$$

$$\eta \equiv \frac{\dot{\epsilon}}{\epsilon H} = \frac{2\ddot{\varphi}}{\dot{\varphi}H} + 2\epsilon, \quad (2.9b)$$

$$\xi \equiv \frac{\dot{\eta}}{\eta H}. \quad (2.9c)$$

The projections of the background equations Eq. 2.8 therefore imply the following kinematical expression for the potential gradient

$$V_a = -\dot{\varphi}H(3 - \epsilon + \frac{1}{2}\eta)T_a + \Omega\dot{\varphi}N_a. \quad (2.10)$$

2.2.2 Linear perturbations

Perturbing around the homogeneous background, the linearized equations of motion of field perturbations are given by [201, 207–210]

$$\frac{D^2}{dN^2}Q^a + (3 - \epsilon)\frac{D}{dN}Q^a + \frac{k^2}{a^2 H^2}Q^a + C^a_b Q^b = 0, \quad (2.11a)$$

$$\text{with } C^a_b \equiv \frac{\nabla_b V^a}{H^2} - 2\epsilon M_p^2 R^a_{\text{cdb}} T^c T^d + 2\epsilon(3 - \epsilon)T^a T_b + \frac{2\epsilon}{\varphi' H^2} (T^a V_b + T_b V^a). \quad (2.11b)$$

The covariant derivatives are with respect to e-folds $\frac{D}{dN} \equiv H D_t$. Moreover, Q^a are the gauge invariant field fluctuations.

These equations can be rewritten in terms of the *comoving curvature perturbation* $\mathcal{R} \equiv \frac{HT_a Q^a}{\dot{\varphi}}$ and *isocurvature perturbations* $\mathcal{S}_i \equiv \frac{(N_i)_a Q^a}{M_p}$ [202] by projecting Eq. 2.11 along the kinematical basis vectors Eq. 2.5. Since for the rest of this chapter we specialize to the two-field scenario, we only quote the corresponding two field equations

$$\left(\frac{\partial}{\partial N} + 3 - \epsilon - \eta\right) \left(\mathcal{R}' + \frac{2\Omega}{\sqrt{2\epsilon}H}\mathcal{S}\right) + \frac{k^2}{a^2 H^2}\mathcal{R} = 0, \quad (2.12a)$$

$$\left(\frac{\partial}{\partial N} + 3 - \epsilon\right) \mathcal{S}' - \frac{2\Omega\varphi'}{H} \left(\mathcal{R}' + \frac{2\Omega}{\sqrt{2\epsilon}H}\mathcal{S}\right) + \left(\frac{\mu^2}{H^2} + \frac{k^2}{a^2 H^2}\right) \mathcal{S} = 0. \quad (2.12b)$$

Here a prime denotes a derivative with respect to e-folds and $S \equiv S_1$. The perturbations are grouped in such a way that we can read off the *entropy mass* (the effective mass of the first isocurvature mode) most easily. The entropy mass is given by

$$\mu^2 \equiv V_{NN} + 2\epsilon H^2 M_p^2 R_{NTNT} + 3\Omega^2, \quad (2.13)$$

Here we use the notation $V_{NN} \equiv N^a N^b \nabla_a \nabla_b V$ and $R_{NTNT} \equiv N_a T^b N^c T^d R^a{}_{cdb}$. We explain in § 2.2.3 by means of a dispersion relation analysis why this is the correct interpretation of the mass of isocurvature perturbations.

2.2.3 Dispersion relation analysis

To understand why we interpret μ^2 (defined in Eq. 2.13) as the effective mass of isocurvature perturbations, we study the dispersion relations of the system of coupled perturbations Eq. 2.12 following [211–213]. We specialize to two fields and write the equations in terms of cosmic time t instead of e-folds

$$\partial_t \left(\dot{\mathcal{R}} + \frac{2\Omega}{\sqrt{2\epsilon}} \mathcal{S} \right) + (3 + \eta)H \left(\dot{\mathcal{R}} + \frac{2\Omega}{\sqrt{2\epsilon}} \mathcal{S} \right) + \frac{k^2}{a^2} \mathcal{R} = 0, \quad (2.14a)$$

$$\ddot{\mathcal{S}} + 3H\dot{\mathcal{S}} + \mu^2 \mathcal{S} + \frac{k^2}{a^2} \mathcal{S} = 2\sqrt{2\epsilon}\Omega \left(\dot{\mathcal{R}} + \frac{2\Omega}{\sqrt{2\epsilon}} \mathcal{S} \right). \quad (2.14b)$$

Let's write the solution as a sum of four normal modes

$$\begin{pmatrix} \mathcal{R} \\ \mathcal{S} \end{pmatrix} = \sum_{\omega} \begin{pmatrix} \mathcal{R}_{\omega} \\ \mathcal{S}_{\omega} \end{pmatrix} e^{-i \int dt \omega}. \quad (2.15)$$

Assuming the adiabatic condition $\dot{\omega}/\omega^2 \ll 1$, and neglecting both slow roll corrections and the rate of change of the turn rate $\dot{\Omega}$, we find the following linear system

$$\begin{pmatrix} -\omega^2 - 3H i \omega + \frac{k^2}{a^2} & -i\omega \frac{2\Omega}{\sqrt{2\epsilon}} + 3H \frac{2\Omega}{\sqrt{2\epsilon}} \\ 2i\omega \sqrt{2\epsilon}\Omega & -\omega^2 - 3H i \omega + \frac{k^2}{a^2} + \mu^2 - 4\Omega^2 \end{pmatrix} \begin{pmatrix} \mathcal{R}_{\omega} \\ \mathcal{S}_{\omega} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (2.16)$$

We can most easily solve for the normal modes of the system of coupled damped harmonic oscillators by introducing the variable $\omega_0^2 \equiv \omega^2 + 3H i \omega$, the eigenmodes of the undamped system. The solutions for ω_0^2 are given by

$$\omega_0^2 = \frac{k^2}{a^2} + \frac{1}{2}\mu^2 \pm \frac{1}{2}\sqrt{\mu^4 + 16\Omega^2 \frac{k^2}{a^2}}. \quad (2.17)$$

Therefore the four solutions are given by

$$\omega = \frac{-3H i \pm \sqrt{-9H^2 + 4\omega_0^2}}{2}, \quad (2.18)$$

two normal modes for each solution of ω_0^2 . Let's distinguish three cases:

- $\mu^2 \gg \frac{9}{4}H^2$. In this case we have heavy isocurvature perturbations which we can integrate out. Defining $c_s^2 \equiv 1 - \frac{4\Omega^2}{\mu^2}$, we have a few different regimes that the coupled oscillators will go through as the physical wavenumber $\frac{k}{a}$ redshifts during inflation.

- If Ω is non-zero we can go to very small scales $4(1 - c_s^2)\frac{k^2}{a^2} \gg \mu^2$ such that

$$\omega_0^2 \approx \frac{k^2}{a^2} \pm \frac{2\Omega k}{a} \quad (2.19)$$

The linear term is hardly ever important, because the inequality implies $\frac{k^2}{a^2} \gg \frac{\Omega^2}{(1-c_s^2)^2}$.

- Then as soon as $4(1 - c_s^2)\frac{k^2}{a^2} \ll \mu^2$ we find

$$\omega_0^2 \approx \left(1 \pm \frac{4\Omega^2}{\mu^2}\right) \frac{k^2}{a^2} + \frac{1}{2}\mu^2(1 \pm 1). \quad (2.20)$$

We can identify a dispersion relation ω_0^- corresponding to a massless excitation with reduced sound speed $(\omega_0^-)^2 = \frac{c_s^2 k^2}{a^2}$, and ω_0^+ , corresponding to an excitation of mass μ . The massive excitation yields the following two normal modes

$$\omega_\mu \approx -\frac{3Hi}{2} \pm \omega_0^+, \quad (2.21)$$

i.e. two decaying and rapidly oscillating solutions of mass μ . Moreover, these solutions have $\mathcal{R}_\omega = 0$, if we neglect the Hubble friction. Therefore, this motivates us to interpret μ^2 as the effective mass of the isocurvature mass perturbations. The behavior of the massless excitation splits in two regimes:

- * On subhorizon scales $\frac{c_s^2 k^2}{a^2} \gg \frac{9}{4}H^2$ we get

$$\omega \approx -\frac{3Hi}{2} \pm \frac{c_s^2 k^2}{a^2}. \quad (2.22)$$

These are two underdamped harmonic oscillators

- * On superhorizon scales $\frac{c_s^2 k^2}{a^2} \ll \frac{9}{4}H^2$ we find

$$\omega \approx -\frac{3Hi}{2}(1 \pm 1) \pm \frac{c_s^2 k^2}{a^2} \frac{i}{3H}, \quad (2.23)$$

one overdamped harmonic oscillator and one solution that converges to a constant. These are exactly the two solutions we expect for \mathcal{R} . The interaction with the isocurvature perturbations is encoded in an effective reduced speed of sound c_s [201, 212, 214, 215].

- $\mu^2 \ll \frac{9}{4}H^2$. Assuming a non-zero value of $\Omega^2 < H^2$ we find that the solution for ω_0^2 splits in two regimes as before:

– If we are at sufficiently small scales $16\Omega^2 \frac{k^2}{a^2} \gg \mu^4$ we get

$$\omega_0^2 \approx \frac{k^2}{a^2} + \frac{1}{2}\mu^2 \pm \frac{2\Omega k}{a} \quad (2.24)$$

The linear term is not important if we also ensure $\frac{k^2}{a^2} \gg 4\Omega^2$. Since $\mu^2 \ll H^2$ we can be either subhorizon or superhorizon:

- * Starting at subhorizon scales $\frac{k^2}{a^2} \gg \frac{9}{4}H^2 \gg \mu^2$ we find the solutions

$$\omega = -\frac{3Hi}{2} \pm \omega_0, \quad (2.25)$$

corresponding to underdamped coupled harmonic oscillators of negligible mass.

- * On superhorizon scales $\frac{k^2}{a^2} \ll \frac{9}{4}H^2$ on the other hand we have

$$\omega = -\frac{3Hi}{2}(1 \pm 1) \pm \frac{i\omega_0^2}{3H}. \quad (2.26)$$

We find two overdamped oscillators and two slowly decaying/growing solutions with exponential factor $-\frac{H}{3} \left(\frac{k^2}{a^2 H^2} + \frac{1}{2} \frac{\mu^2}{H^2} \pm \frac{2\Omega k}{a H^2} \right)$. It depends on the values of μ and Ω which term is most important.

– If $\mu \neq 0$, we can consider the regime $16\Omega^2 \frac{k^2}{a^2} \ll \mu^4$. This leads to

$$\omega_0^2 \approx \left(1 \pm \frac{4\Omega^2}{\mu^2} \right) \frac{k^2}{a^2} + \frac{1}{2}\mu^2 (1 \pm 1). \quad (2.27)$$

On super horizon scales $\frac{k^2}{a^2} \ll \frac{9}{4}H^2$ we find the following eigenmodes

$$\omega = -\frac{3Hi}{2}(1 \pm 1) \pm \frac{i\omega_0^2}{3H}. \quad (2.28)$$

In other words we find two overdamped oscillators, one slowly decaying solution with exponential factor $-\frac{H}{3} \frac{\mu^2}{H^2}$, and one solution that converges to a constant.

- $\mu^2 \sim H^2$. This is the typical quasi-single field regime [100]. Again, assuming that the turn rate is nonzero but $\Omega^2 < H^2$, we can distinguish two regimes.

- In case we have $\mu^4 \gg 16\Omega^2 \frac{k^2}{a^2}$, we find the same solution for ω_0 as in Eq. 2.27. On super-Hubble scales the four eigenmodes Eq. 2.18 contain a constant and decaying solution (when the \pm is a $-$ in Eq. 2.27). The other solutions corresponding to the isocurvature mode are both decaying

$$\omega \approx -\frac{3Hi}{2} \pm \Delta\omega \quad (2.29)$$

where $\Delta\omega$ is imaginary if $\mu^2 \lesssim 9/4H^2$ and real if $\mu^2 \gtrsim 9/4H^2$. On sub-Hubble scales we find the same form as Eq. 2.29, but now $\Delta\omega = \omega_0$. This solution corresponds to four underdamped harmonic oscillators of mass zero and mass μ .

- For values $\mu^4 \ll 16\Omega^2 \frac{k^2}{a^2}$ we find the same ω_0^2 as in Eq. 2.24. This condition forces us to be on sub-Hubble scales. The dispersion relations are again given by Eq. 2.29 with $\omega \sim \frac{k^2}{a^2}$, corresponding to four undamped harmonic oscillators.

We conclude that on sub-Hubble scales the isocurvature perturbations behave like an underdamped harmonic oscillator of mass μ . On super-Hubble scales the Hubble friction takes over and the isocurvature perturbations freeze out ($\mu = 0$), slowly decay ($\mu^2/H^2 \ll 9/4$) or rapidly decay ($\mu^2/H^2 \gtrsim 1$).

This concludes our brief review of the kinematical description of linear perturbations in multi-field inflation. From now on we specialize to the two field scenario.

2.3 Phenomenology of Orbital Inflation

The results from § 2.2 equip us with all the necessary tools to study Orbital Inflation. We define Orbital Inflation as two-field inflation with a slowly varying entropy mass μ/H and a constant radius of curvature κ of the inflationary trajectory in field space (see Eq. 2.6). In this section we study how the predictions for the spectral tilt n_s and the tensor-to-scalar ratio r depend on μ and κ . For that purpose we restrict ourselves to a small entropy mass $0 \leq \mu^2/H^2 \leq 0.2$. In the regime where $\kappa/M_p \lesssim 10^2$ we find non-trivial results, and as far as we know the phenomenology of this regime has not been studied before, except for the limiting case $\mu^2/H^2 = 0$, see [138] and Chapter 3. For the limit of large entropy mass $\mu^2/H^2 \gg 1$ we refer the reader to Chapter 5. Let us emphasize, though, that in the next section, where we construct explicit models of Orbital Inflation, the entropy mass can take any value.

Using the results from § 2.2 we first derive an analytical approximation for the power spectrum of curvature perturbations based on the evolution of the coupled perturbations on super-Hubble scales. We complement this with a numerical computation [216–219] when we present the results in Figure 2.2. To address the validity of the perturbative numerical computation and our simple analytical estimate we need the quadratic action of perturbations. Moreover, we also need it to have the right normalization of the fields when we quantize the theory. For two-field inflation, in e-folds, it is given by [138]

$$S^{(2)} = \frac{1}{2} \int dN d^3x a^3 \bar{M}_p^2 H \left[2\epsilon \left(\mathcal{R}' - \frac{2M_p}{\kappa} \mathcal{S} \right)^2 + (\mathcal{S}')^2 - \frac{\mu^2}{H^2} \mathcal{S}^2 + \dots \right] \quad (2.30)$$

The ellipses denote the gradient terms $-(\partial_i \mathcal{S})^2/H^2 - 2\epsilon(\partial_i \mathcal{R})^2/H^2$. The linearized system of coupled perturbations Eq. 2.12 for two fields reads

$$(\partial_N + 3 - \epsilon - \eta) \left(\mathcal{R}' - \frac{2M_p}{\kappa} \mathcal{S} \right) + \frac{k^2}{a^2 H^2} \mathcal{R} = 0, \quad (2.31a)$$

$$(\partial_N + 3 - \epsilon) \mathcal{S}' + \left(\frac{\mu^2}{H^2} + \frac{k^2}{a^2 H^2} \right) \mathcal{S} = -\frac{4\epsilon M_p}{\kappa} \left(\mathcal{R}' - \frac{2M_p}{\kappa} \mathcal{S} \right). \quad (2.31b)$$

Notice that we have rewritten the turn rate Ω in terms of the radius of curvature κ , defined in Eq. 2.6. On super-Hubble scales $k^2 \ll a^2 H^2$, the equations simplify considerably. The equation of motion for \mathcal{R} has the solution

$$\mathcal{R}' - \frac{2M_p}{\kappa} \mathcal{S} = 0 + \text{decaying part}. \quad (2.32)$$

Meanwhile, neglecting the decaying part on the right hand side of Eq. 2.31b, the equation for \mathcal{S} reduces to

$$\mathcal{S}'' + (3 - \epsilon) \mathcal{S}' + \frac{\mu^2}{H^2} \mathcal{S} = 0. \quad (2.33)$$

For approximately constant $0 < \mu^2/H^2 \ll 9/4$ and $\epsilon \ll 1$, this equation describes an overdamped oscillator with solution

$$\mathcal{S}(N) \approx \mathcal{S}_0 e^{-\frac{N-N_0}{3} \frac{\mu^2}{H^2}} + \text{decaying part}. \quad (2.34)$$

The isocurvature perturbations in turn source the curvature perturbations, through Eq. 2.32. Integrating this equation gives the superhorizon solution for \mathcal{R} :

$$\mathcal{R}(N) \approx \mathcal{R}_0 + \mathcal{S}_0 \frac{6M_p H^2}{\mu^2 \kappa} \left(1 - \exp \left(-\frac{N}{3} \frac{\mu^2}{H^2} \right) \right). \quad (2.35)$$

Here we made use of the fact that κ is constant. In the limit that $\mu^2/H^2 = 0$, the isocurvature perturbation freezes out at horizon crossing, conform Eq. 2.34.

In this case the second term in Eq. 2.35 becomes proportional to ΔN , which is in agreement with the expansion of this term for small μ^2/H^2 . In the quantum analysis of two-field inflation [198] there are two uncorrelated contributions to $\hat{\mathcal{R}}$. The first contribution is sourced by initial curvature perturbations where $\mathcal{S}_0 = 0$. This corresponds to the constant mode \mathcal{R}_0 that freezes out on super-Hubble scales. The second contribution is sourced by initial isocurvature perturbations, and corresponds to the second solution proportional to \mathcal{S}_0 . Using the typical amplitude of quantum perturbations at horizon crossing $\mathcal{S}_0 \sim \sqrt{2\epsilon}\mathcal{R}_0 \sim \frac{H}{2\pi}$, the power spectrum of curvature perturbations is given by³

$$P_{\mathcal{R}} \approx \frac{H^2}{8\pi^2\epsilon M_p^2} \left(1 + 2\epsilon \left(\frac{6M_p H^2}{\mu^2 \kappa} \right)^2 \left(1 - \exp \left(-\frac{\Delta N}{3} \frac{\mu^2}{H^2} \right) \right)^2 \right). \quad (2.37)$$

Here all variables are understood to be evaluated at horizon crossing, and ΔN denotes the number of e-folds counted from when the observable modes cross the horizon until the end of inflation. In the limit of small entropy mass μ^2/H^2 , we can expand the exponential in Eq. 2.37 to find that this term scales as $\Delta N^2/\kappa^2$. To get an improved analytical result, one can perform a full in-in computation similar to what is done in [138]. We checked our simple analytical estimate numerically by means of the exact models described in § 2.4.2, using the Python code developed by [218, 219], and found that it works well for $\kappa^2/M_p^2 \geq 1$. We plot the results for n_s and r in Figure 2.2. The analytical results are obtained by using

$$n_s = \frac{\partial \ln P_{\mathcal{R}}}{\partial N}, \quad r = \frac{2H^2}{\pi^2 M_p^2} \frac{1}{P_{\mathcal{R}}}, \quad \epsilon = \frac{p}{2\Delta N + p}. \quad (2.38)$$

The numerical results are computed using the following potential and kinetic term

$$V(\theta, \rho) = 3M_p^4 \left(\theta^2 + \frac{2M_p^2}{3f(\rho)} \right) \left(1 + \frac{\lambda}{12} \frac{(\rho - \rho_0)^2}{M_p^2} \right)^2, \quad (2.39a)$$

$$2K = f(\rho)(\partial\theta)^2 + (\partial\rho)^2 \quad \text{with} \quad f(\rho) = e^{2\rho/R_0}. \quad (2.39b)$$

³As a cross-check we compare the obtained power spectrum with that of quasi-single field inflation [100]. Matching with the numerical function $\mathcal{C}(\nu)$ defined in Eq. (3.8) of [100], we find

$$\mathcal{C}(\nu) = \frac{9}{2} \frac{1}{\left(\frac{9}{4} - \nu^2\right)^2} \left(1 - \exp \left(-\frac{\Delta N}{3} \left(\frac{9}{4} - \nu^2 \right) \right) \right)^2. \quad (2.36)$$

This seems to agree reasonably well if we take $\Delta N \sim 50 - 60$, even for μ^2/H^2 close to $9/4$. When $\kappa^2/M_p^2 \lesssim 10^2$ We don't find the same prediction for n_s as in Eq. (3.11) of [100], because the second term in Eq. 2.37 becomes important, and we have to take into account the ΔN -dependence of $\mathcal{C}(\nu)$.

The field metric is hyperbolic with curvature $\mathbb{R} = -\frac{2}{R_0^2}$. As explained in § 2.4.2 this model admits Orbital Inflation with $\kappa^2 = R_0^2$ and $\mu^2/H^2 = \lambda(1 - \frac{1}{3}\epsilon)$.

In Figure 2.2 we vary κ^2/M_p^2 between 1 and 10^5 . The lower bound comes from the validity of the perturbative approach we implicitly assumed. The numerical code [218] is performing a tree level in-in computation, and higher order tree level (and loop) corrections should be small compared to the leading result. Our simple analytical result captures the super-Hubble evolution of \mathcal{R} , and therefore provides an estimate of the leading order tree level computation. Using Eq. 2.30 we can estimate that $\alpha \sim \sqrt{8\epsilon} \frac{M_p}{\kappa}$ is the perturbation parameter that measures the relative size of the higher order tree level corrections compared to the leading tree level computation⁴. Therefore, we need to ensure that $\alpha \ll 1$, which is why we take $\kappa/M_p \geq 1$. At the same time we should be careful that quantum perturbations remain much smaller than the radius of curvature $\delta\rho \ll \kappa$. In particular, we should be careful in the limit that the isocurvature perturbations are very light $\mu^2/H^2 \ll 1$. Fortunately, this is the case. If we consider small values of κ , such that the second term in Eq. 2.37 dominates, we get $P_{\mathcal{R}} \sim \frac{H^2}{\kappa^2}$. Since the amplitude of the power spectrum is fixed by observations $A_{\mathcal{R}} \sim 10^{-9}$, this implies that the typical size of quantum fluctuations gets suppressed if we decrease κ . We find $\delta\rho^2 \sim H^2 \sim \kappa^2 A_{\mathcal{R}} \ll \kappa^2$, so we are always fine.

From Figure 2.2 we see that the observable predictions are already significantly modified for $\kappa^2/M_p^2 \lesssim 10^3$ (for $\mu^2/H^2 = 0$) or $\kappa^2/M_p^2 \lesssim 10^2$ (for $\mu^2/H^2 \approx 0.2$). Interestingly, the entropy mass μ^2/H^2 dictates how the inflationary predictions fan out in the (n_s, r) plane. In particular, the various entropy masses predict a different change in n_s . It would be very interesting to see if this effect may allow us to *distinguish* between the various entropy masses. This requires a complementary analysis of the bispectrum. For these models we expect to find local primordial non-Gaussianities, because the super-Hubble evolution of \mathcal{R} is the dominant contribution to its final amplitude. Therefore, in future work we plan to assess the amplitude and the scale-dependence of the bispectrum in the squeezed configuration. In particular, we would like to understand its dependence on the kinematical and geometrical parameters of multi-field inflation. This should give us more insight what to expect from the simplest modifications to the single field inflationary scenario.

⁴In the in-in computation the sourcing of \mathcal{R} by \mathcal{S} is captured by the interaction term $S_{\text{int}}^{(2)} = \int d\tau d^3x a^3 4\epsilon \frac{M_p^3 H}{\kappa} (\partial_\tau \mathcal{R}) \mathcal{S}$ (using Eq. 2.30 written in terms of conformal time $d\tau = dN/aH$.) Rewriting this in canonical variables $u \equiv \sqrt{2\epsilon} a M_p \mathcal{R}$ and $v \equiv a M_p \mathcal{S}$ we get $S_{\text{int}}^{(2)} \sim \int d\ln\tau d^3x \alpha (\partial_\tau u) v$, with $\alpha = \sqrt{8\epsilon} \frac{M_p}{\kappa}$.

Finally, we should also estimate the relative amplitude of isocurvature perturbations compared to the amplitude of curvature perturbations. How they are related to the late-time non-adiabatic perturbations depends on the mechanism of reheating, but we may still want to ensure that they are suppressed at the end of inflation. From [Eq. 2.34](#) we see that the isocurvature perturbations decay on super-Hubble scales if $\mu^2/H^2 \gtrsim 3/\Delta N$, so we can assume that isocurvature perturbations with entropy masses of $\mu^2/H^2 \gtrsim 0.06$ have decayed by the end of inflation. However, for $\mu^2/H^2 \lesssim 0.06$, we should be more careful. The ratio between curvature and isocurvature perturbations is given by

$$\beta_{\text{iso}} \equiv \frac{P_S}{2\epsilon P_{\mathcal{R}}} \approx \frac{1}{1 + 8\epsilon \frac{M_p^2 \Delta N^2}{\kappa^2}}. \quad (2.40)$$

To arrive at this result, we expanded the exponential in [Eq. 2.34](#). Therefore, our results are reliable in the regime

$$1 \gg 8\epsilon \frac{M_p^2}{\kappa^2} \gg \frac{1}{\Delta N^2} \quad (\text{or } \mu^2/H^2 \gtrsim 0.1) \quad (2.41)$$

Fortunately, this is contained in the regime where we find non-trivial results.

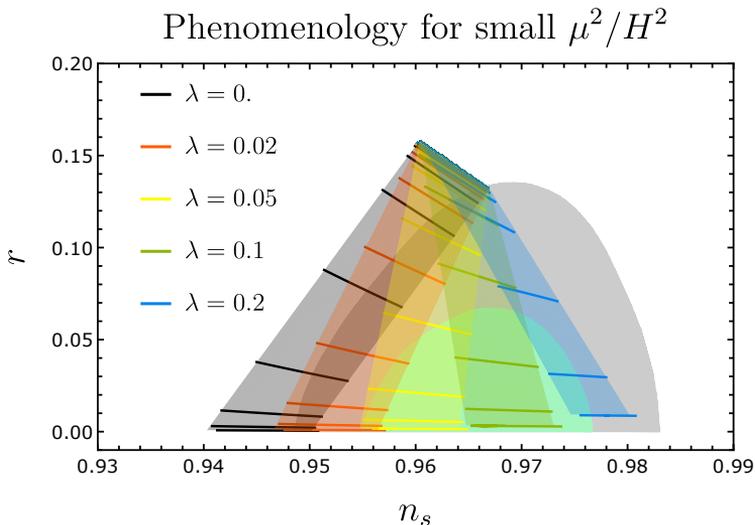


Figure 2.2: This figure shows the predictions of (n_s, r) for the model given in Eq. 2.39 using the numerical code [218, 219]. The entropy mass takes five different values, as indicated in the legend, with $\mu^2/H^2 = \lambda(1 - \frac{1}{3}\epsilon) \approx \lambda$. The solid lines correspond to $R_0^2/M_p^2 \in \{1, 4, 4^2, \dots, 4^8\}$ from bottom to top, and we let $\Delta N \in [50, 60]$. On top of that we plot our analytical results (coloured shaded regions) using Eq. 2.38, where we vary κ^2/M_p^2 between 1 and 10^5 . Furthermore, on the background we plotted the 1σ and 2σ confidence contours from *Planck* [115].

2.4 Exact models of Orbital Inflation

In this section we derive exact models of Orbital Inflation. For this we first generalize the Hamilton-Jacobi formalism [74, 75, 77, 78] to two-field inflation in § 2.4.1, which we use in § 2.4.2 to perform an explicit construction.

2.4.1 Hamilton-Jacobi for two-field inflation

We generalize the Hamilton-Jacobi formalism to two-field inflation, such that we can locally reconstruct the potential similar to the single field case described in § 1.2.3. The first step is to replace the time coordinate with the proper field distance φ along the inflationary trajectory

$$\varphi \equiv \int dt \sqrt{G_{ab} \frac{d\phi^a}{dt} \frac{d\phi^b}{dt}}. \quad (2.42)$$

Given the Hubble slow roll parameters, we know that we can reconstruct the potential *on the inflationary trajectory*, following § 1.2.3. However, in the two-field case we would like to reconstruct the potential in the neighborhood of the inflationary trajectory as well. In fact, from Eq. 2.8 we know that the gradient

of the potential in the orthogonal direction should be such that it counterbalances the centrifugal force. This provides an additional constraint in the two-field scenario. Generically, H is a function of the normal field coordinates σ_i as well. Therefore we might as well parameterize the Hubble parameter in terms of any set of field coordinates $H = H(\phi^a)$. This is fine as long as all coordinates are strictly increasing or decreasing on the inflationary trajectory.

Similarly to the single field case we have a solution for every initial value $H(\phi_0^a)$, however this cannot cover all possible initial data. In the case of two-field inflation we therefore need an additional function $F(\phi^a)$ to fully specify the system. Indeed, using the second Friedmann equation [Eq. 2.4c](#) we find

$$\dot{H} = \dot{\phi}^a H_a = -\frac{\dot{\phi}^a \dot{\phi}^b G_{ab}}{2M_p^2} \quad \longrightarrow \quad H_a = -\frac{G_{ab} \dot{\phi}^b}{2M_p^2} + F(\phi^b) N_a . \quad (2.43)$$

Note that we have the freedom to add to H_a any contribution proportional to the normal vector to the trajectory, because it is projected out when contracting with $\dot{\phi}^a$ to get \dot{H} . Using [Eq. 2.43](#), we rewrite the first Friedmann equation as the multi-field Hamilton-Jacobi equation

$$3H^2 M_p^2 = V + 2M_p^4 (H^a H_a - F^2) . \quad (2.44)$$

Notice that we need both H and F to fully determine the system. This equation contains the same information as the tangent projection of the field equations. To see this we write [Eq. 2.44](#) as

$$3H^2 M_p^2 = V + 2M_p^4 H_\varphi^2 \quad \text{with} \quad H_\varphi = H_T = -\frac{\dot{\varphi}}{2M_p^2} \quad (2.45)$$

and take a derivative with respect to φ to find [Eq. 2.8a](#). In addition we have to obey the normal projection of the field equations [Eq. 2.8b](#), which provides a constraint on F . We compute $D_t T^a$ using $T^a = -\frac{2M_p^2}{\dot{\varphi}} (H^a - F N^a)$. Moreover V_N can be computed using [Eq. 2.44](#). Together they lead to the constraint equation

$$3HF - 2M_p^2 H_T F_T + 2F (F_N - H_{NN}) = 0 . \quad (2.46)$$

In particular, $F = 0$ gives a solution compatible with the constraint equation. However, there are also solutions that have a non-trivial F .

At this point we would like to stress that we intend to apply the Hamilton-Jacobi formalism to reconstruct the potential in the neighborhood of a given inflationary trajectory only. Therefore, we only need to solve for [Eq. 2.44](#) and [Eq. 2.46](#) *on the trajectory*. These equations restrict the normal projection of

the gradient of the potential, but other than that they leave the potential free. For this local reconstruction it might be convenient to pick the natural kinematical field coordinates. This implies that each normal vector is parallel transported along its normal direction, i.e. $N_i^a \nabla_a N_i^b = 0$. In the two-field case we therefore have $H_{NN} \equiv N^a N^b \nabla_a \nabla_b H = N^a \nabla_a F \equiv F_N$ and the constraint equation further simplifies to

$$3HF - 2M_p^2 H_\varphi F_\varphi = 0. \quad (2.47)$$

This restricts the form of F only at $\sigma = \sigma_0$, i.e. on the trajectory. We will next see that this allows us to construct inflationary models with any isocurvature mass.

2.4.2 Orbital inflation

We are now ready to construct inflationary potentials with constant κ and (almost) constant μ^2/H^2 , using Eq. 2.44 and Eq. 2.47. A constant radius of curvature can be achieved by considering an inflationary trajectory that proceeds along an isometry direction of the field metric that is *not* a geodesic. This is the key characteristic of the exact models of Orbital Inflation. And in this sense it is an attempt of realizing spontaneously symmetry probing [196] in inflation. The existence of an isometry implies that we are free to choose our field coordinates (θ, ρ) , such that the field metric G_{ab} does not depend on θ . Moreover, we have also the freedom to put $G_{\theta\rho}$ to zero⁵. Furthermore, we denote $f(\rho) = G_{\theta\theta}$.

We would like to reconstruct the potential that admits Orbital Inflation, that is, solutions of the form

$$\dot{\rho} = 0, \quad \dot{\theta} < 0. \quad (2.48)$$

The sign of $\dot{\theta}$ is our choice of convention, also we take $\theta > 0$ on the inflationary trajectory. This means we can replace $\varphi \rightarrow \sqrt{f}\theta$ and $H_\varphi \rightarrow \frac{1}{\sqrt{f}}H_\theta$. Moreover, the relevant kinematical and geometrical inflationary background quantities simplify to

$$T^a = \frac{1}{\sqrt{f}}(-1, 0) \quad \text{and} \quad N^a = (0, 1), \quad (2.49a)$$

$$\dot{\theta} = -2M_p^2 \frac{H_\theta}{f}, \quad \epsilon = \frac{2M_p^2 H_\theta^2}{fH^2}, \quad \kappa = \frac{2f}{\partial_\rho f}, \quad \mathbb{R} = \frac{2}{\kappa^2} - \frac{f_{\rho\rho}}{f}. \quad (2.49b)$$

⁵If $G_{\theta\rho} \neq 0$, define $\tilde{\theta} = \theta + \int d\rho \frac{G_{\theta\rho}(\rho)}{G_{\theta\theta}(\rho)}$ such that $\tilde{G}_{\tilde{\theta}\rho} = 0$.

Here $f = f(\rho)$ and $H = H(\theta, \rho)$. On the desired inflationary trajectory the radial coordinate takes a constant value ρ_0 . Therefore, we expand $H(\rho, \theta)$ around $\rho = \rho_0$:

$$H(\rho, \theta) = M_p \left(W(\theta) + X(\theta) \frac{\rho - \rho_0}{M_p} + Y(\theta) \frac{(\rho - \rho_0)^2}{M_p^2} + \dots \right). \quad (2.50)$$

Moreover, we take $F(\rho, \theta) = \partial_\rho H(\rho, \theta)$. We insert this into the constraint equation Eq. 2.46 to identify the restrictions on $X(\theta)$. First of all, on the desired inflationary trajectory we have $F = H_\rho$ and $H_{\rho\rho} = F_\rho$, which allows us to use Eq. 2.47. Moreover, we express tangent derivatives in terms of derivatives with respect to θ . The constraint equation simplifies to

$$3W(\theta)X(\theta) - \frac{2W_\theta(\theta)X_\theta(\theta)M_p^2}{f(\rho_0)} = 0. \quad (2.51)$$

An obvious solution is given by $X(\theta) = 0$, which we will assume for simplicity. Next, we compute the effective mass of isocurvature perturbations. For the orbital inflationary trajectory we have $\mu^2 = V_{\rho\rho} + \epsilon M_p^2 H^2 \left(\mathbb{R} + \frac{6}{\kappa^2} \right)$, which can be written as

$$\mu^2 = 12M_p^2 Y(\theta)W(\theta) - \frac{8M_p^4}{f(\rho_0)} Y_\theta(\theta)W_\theta(\theta). \quad (2.52)$$

Therefore, we can construct a potential with any isocurvature mass we like. For instance, specializing to $Y(\theta) = \frac{1}{12}\lambda W(\theta)$ yields

$$\frac{\mu^2}{H^2} \Big|_{\rho=\rho_0} = \lambda \left(1 - \frac{2M_p^2}{3f(\rho_0)} \frac{W_\theta^2(\theta)}{W^2(\theta)} \right). \quad (2.53)$$

Notice that the second term in Eq. 2.53 is proportional to the slow roll parameter ϵ on the trajectory, using Eq. 2.49. Therefore, well within the slow-roll regime we approximately get $\frac{\mu^2}{H^2} \approx \lambda$.

It should be clear that if one continues in this fashion, combinations of the higher order derivatives of the potential can be tuned as well. For instance, this technique may be used to construct explicit quasi-single-field models of inflation with large $V_{\rho\rho\rho}$ [100]. Moreover, there are many more ways to arrive at a constant entropy mass. For now we put the higher order corrections to H and F to zero, and use Eq. 2.44 to find

$$V(\theta, \rho) = 3M_p^4 \left(W^2(\theta) + \frac{2M_p^2 W_\theta^2(\theta)}{3f(\rho)} \right) \left(1 + \frac{\lambda}{12} \frac{(\rho - \rho_0)^2}{M_p^2} \right)^2, \quad (2.54)$$

which constitutes a realization of Orbital Inflation. Finally, using the results from § 1.2.3, we may choose $W(\theta) = \theta^p$ such that ϵ is given by

$$\epsilon = \frac{p}{2\Delta N + p}, \quad (2.55)$$

where ΔN denotes the number of e-folds before the end of inflation where the observable modes cross the horizon. We used this potential with $p = 1$ in the numerical/analytical computations presented in Figure 2.2.

Finally, we would like to point out an interesting limit. Notice that we can locally achieve a zero entropy mass μ by choosing $Y(\theta) = 0$. If we demand Orbital Inflation at *any* radius ρ_0 with $\mu = 0$, this forces H to be a function of θ only. Therefore the masslessness of radial perturbations corresponds to a radial shift symmetry in H rather than in the resulting potential. We study this particular limit in Chapter 3.

2.5 Summary

In two-field inflation there are only a few kinematical and geometrical parameters that determine the evolution of linear perturbations. These are the *radius of curvature* of the inflationary trajectory in field space, the *entropy mass* (the mass of the isocurvature perturbations) and the Hubble slow-roll parameters. This motivates us to introduce Orbital Inflation, in which the radius of curvature is constant, and the remaining parameters are slowly varying. This is one of the simplest two-field extensions to single field inflation.

In § 2.3 we discussed the phenomenology of Orbital Inflation. In particular we focus on the regime of small entropy mass $\mu^2/H^2 \leq 0.2$ and we found that the predictions are substantially modified already for $\kappa^2/M_p^2 \lesssim 10^2$. Furthermore, in § 2.4 we showed how to explicitly construct exact models of Orbital Inflation. The key characteristic of these models is that inflation proceeds along an isometry direction of the field metric. We used a generalization of the Hamilton-Jacobi formalism to find the potential *locally* around the inflationary trajectory. Finally, we saw that an interesting limiting case is to have a vanishing entropy mass. This is reflected by a shift symmetry in the Hubble parameter rather than in the resulting potential. We investigate this particular case in Chapter 3.

For convenience of the reader, we collect the relevant formulas and elements that describe the two-field models we have reconstructed in § 2.4. The

procedure is more general, and can be used to reconstruct other multi-field models of inflation as well

- We specialized to a two-field model with two scalar fields θ and ρ described by the following action

$$S = \frac{1}{2} \int d^4x \sqrt{-g} [M_p^2 R - f(\rho) \partial_\mu \theta \partial^\mu \theta + \partial_\mu \rho \partial^\mu \rho - 2V(\theta, \rho)] . \quad (2.56)$$

with R the Ricci scalar of spacetime. The scalar kinetic term has an isometry in the θ direction.

- In order to achieve a constant radius of curvature, we force inflation to proceed exactly in the θ direction. This puts a restriction on the form of $H(\rho, \theta)$, namely that it has to be independent of θ on the trajectory $\rho = \rho_0$.
- Next, we compute the entropy mass in terms of $H(\rho, \theta)$ expanded around $\rho = \rho_0$. We took into account that the entropy mass, defined in Eq. 2.13, receives geometrical and kinematical corrections as well. This allowed us to reconstruct the following potential, which admits an approximately constant entropy mass $\mu^2/H^2 \approx \lambda$ up to slow-roll corrections

$$V = 3M_p^4 \left(W^2(\theta) + \frac{2M_p^2 W_\theta^2(\theta)}{3f(\rho)} \right) \left(1 + \frac{\lambda}{12} \frac{(\rho - \rho_0)^2}{M_p^2} + \dots \right)^2 , \quad (2.57)$$

The ellipses denote higher order terms in the expansion around $\rho = \rho_0$ we left unspecified. They determine higher order derivatives of the potential, such as $V_{\rho\rho\rho}$.

Orbital Inflation with ultra-light fields

In this chapter we present a class of inflationary models with two light fields that have predictions similar to those of single field inflation. Inflation proceeds along an ‘angular’ isometry direction in field space at arbitrary ‘radius’ and is a special case of Orbital Inflation discussed in [Chapter 2](#). More precisely, we study the Orbital Inflation in the limit of vanishing entropy mass¹. We dub this ‘ultra-light Orbital Inflation’, because it realizes the shift symmetry described in [\[138\]](#). If the field radius of curvature of the inflationary trajectory is sufficiently small, the amplitude of isocurvature perturbations and primordial non-Gaussianities are highly suppressed. Ultra-light Orbital Inflation mimics single field inflation, because only one degree of freedom is responsible for the observed perturbations.

We study a toy model of ultra-light Orbital Inflation in [§ 3.2](#). This allows us to intuitively understand its interesting properties. In the successive sections we make our intuitive arguments more precise. In [§ 3.3](#) we derive the family of two-field models which allow for ultra-light Orbital Inflation and give the corresponding exact solutions. We prove neutral stability of ultra-light Orbital Inflation in [§ 3.4](#). Then, in [§ 3.5](#) we recap the definition of mass and the consequences of having massless isocurvature perturbations. Finally, we study the phenomenology of ultra-light Orbital Inflation in [§ 3.6](#).

The results in this chapter are based on joint work with Ana Achúcarro, Edmund Copeland, Oksana Iargyina, Gonzalo Palma and Dong-Gang Wang.

¹ The entropy mass is the effective mass of isocurvature perturbations. The definition of mass is non-trivial in a time-dependent inflationary background. By computing the normal modes of the coupled system of perturbations, we find a dispersion relation of the isocurvature perturbations corresponding to modes of mass μ as defined in [Eq. 3.18](#). Therefore it is μ that we identify as the entropy mass. For more details, see [§ 2.2.3](#).

3.1 Introduction

The Planck data [115] reveal that inflationary perturbations are Gaussian and adiabatic to a high level of accuracy. A possible explanation for the observed simplicity is that the perturbations are generated by a single degree of freedom with small self-interactions. Do the observations therefore imply that, besides the inflaton, no other light fields are active during inflation? The answer is no. As pointed out in [138], in the limit that the other fields are massless, but coupled to the inflaton, the predictions mimic those of single field inflation.

Inflation with massless isocurvature modes behaves like single field inflation, because only one degree of freedom is relevant for the observed perturbations. The single field behavior is of dynamical origin. The key feature is that the isocurvature perturbations freeze out on superhorizon scales and constantly feed the curvature perturbations. Therefore, the isocurvature perturbations generate the temperature fluctuations we observe in the sky.

In this work we provide a realization of a family of two-field inflationary models in which the isocurvature perturbations become exactly massless (see footnote 1). We dub this ultra-light Orbital Inflation. We use the two-field generalization of the Hamilton-Jacobi formalism [74, 75, 77, 78] presented in Chapter 2 to derive the form of the potential. The key characteristic of *ultra-light* Orbital Inflation is that inflation proceeds along an isometry direction of the field metric at *arbitrary radius*. The resulting scalar field potential and kinetic term are given by

$$V = 3H^2 M_p^2 - 2M_p^4 \frac{H_\theta^2}{f(\rho)}, \quad 2K = f(\rho) \partial\theta^2 + \partial\rho^2, \quad (3.1)$$

where the fields are denoted by θ and ρ . Moreover, the Hubble parameter H is a function of θ only and $f(\rho) > 0$.

In this chapter we point out several interesting properties of ultra-light Orbital Inflation:

- Each attractor is an *exact* solution to the highly non-linear system of field equations and Friedmann equation. This is ensured by using the Hamilton-Jacobi formalism.
- This system is neutrally stable. A small perturbation orthogonal to a given attractor solution will bring us to one of the neighboring attractors.
- Because isocurvature (= normal) perturbations move us freely between attractors, this implies that they are exactly massless. Thanks to the

Hubble friction their velocity decays and therefore their amplitude freezes out.

- The quadratic action of perturbations has enhanced symmetry. On top of the usual shift symmetry of curvature perturbations $\mathcal{R} \rightarrow \mathcal{R} + \text{const}$, the masslessness of the isocurvature perturbation \mathcal{S} implies a combined shift symmetry [138]

$$\mathcal{S} \rightarrow \mathcal{S} + c, \quad \dot{\mathcal{R}} \rightarrow \dot{\mathcal{R}} - \lambda c. \quad (3.2)$$

We show that we can understand this as a result of the background dynamics. The symmetry transformation is related to a map of one background attractor to another, labeled by the continuous parameter c .

- For large enough λ this implies that curvature perturbations are dynamically enhanced and the predictions of the power spectra coincide with those of single field inflation [138]. Moreover, the final curvature perturbations are completely determined by the initial isocurvature perturbations. Therefore, these multi-field inflation scenarios mimic the predictions of single-field inflation.

We work in Planck units $\hbar = c = 1$ and the reduced Planck mass is given by $M_p = (8\pi G)^{-1/2}$.

3.2 A toy model with neutrally stable orbits

To illustrate the idea, we consider a simple toy model in flat field space in polar field coordinates $\phi^a = (\theta, \rho)$. The kinetic term and potential are given by

$$K = \frac{1}{2} (\rho^2 (\partial\theta)^2 + (\partial\rho)^2), \quad V = \frac{1}{2} m^2 M_p^2 \left(\theta^2 - \frac{2M_p^2}{3\rho^2} \right) \quad (3.3)$$

This potential allows for orbital solutions with a constant angular velocity, see [Figure 3.1](#). It explicitly breaks the shift symmetry of θ to overcome the Hubble friction. Moreover the ρ -dependent part of the potential provides a centripetal force that stabilizes the radial direction. Although the potential in this toy model is unbounded from below, inflation only takes place in the physically consistent regime where the potential energy is positive. Moreover, our analysis is restricted to radii $\rho > M_p$. Therefore, we only care about the local form of the potential close to the inflationary trajectory, which we assume is captured well by this toy potential. The full true potential should be well-behaved at smaller radii.

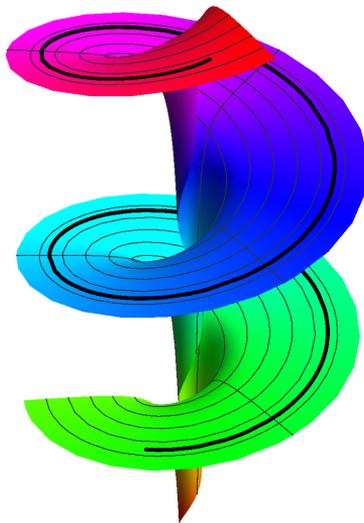


Figure 3.1: The potential given in Eq. 3.3 together with a typical inflationary trajectory. The black line corresponds to the full numerical solution.

3.2.1 Exact solution

We start with an analysis of the homogeneous background dynamics of $\theta(t)$ and $\rho(t)$. The field equations of motion and the Friedmann equation read

$$\rho^2 \ddot{\theta} + 3H\rho^2 \dot{\theta} + 2\rho \dot{\rho} \dot{\theta} + m^2 M_p^2 \theta = 0, \quad (3.4a)$$

$$\ddot{\rho} + 3H\dot{\rho} - \rho \dot{\theta}^2 + \frac{2m^2 M_p^4}{3\rho^3} = 0, \quad (3.4b)$$

$$6H^2 M_p^2 = \rho^2 \dot{\theta}^2 + \dot{\rho}^2 + m^2 M_p^2 \left(\theta^2 - \frac{2M_p^2}{3\rho^2} \right). \quad (3.4c)$$

Typically, it is very hard to find exact solutions to such a system of equations, because they are highly non-linear. In particular the friction terms in the

field equations involve the square root of the right-hand side of the Friedmann equation Eq. 3.4c. In this case, however, the system allows for *exact* stable solutions of the form

$$\rho = \rho_0, \quad \dot{\theta} = \pm \sqrt{\frac{2}{3}} \frac{mM_p^2}{\rho_0^2}, \quad H^2 = \frac{m^2\theta^2}{6}, \quad \epsilon = \frac{2M_p^2}{\rho_0^2\theta^2}, \quad (3.5)$$

for any ρ_0 . Here $\epsilon \equiv -\frac{\dot{H}}{H^2}$ is the slow roll parameter that measures the deviation from the de Sitter solution. Notice that the negative contribution to the squared Hubble parameter H^2 from the radial part of the potential is precisely cancelled by the angular kinetic energy. We will show neutral stability of these exact solutions in § 3.4. Therefore, in what follows we can assume the inflationary trajectory to be one of them.

3.2.2 Symmetry transformation and massless modes

The inflationary trajectory proceeds along an isometry of the field metric, namely the angular direction. This is clearly not a geodesic in flat field space. The radius of curvature of the trajectory is constant and given by $\kappa = \rho_0$. From Eq. 3.5 we can deduce $(\theta^2)' = \frac{2\theta\dot{\theta}}{H} = -\frac{4M_p^2}{\rho_0^2}$, where the prime denotes a derivative with respect to e-folds $(\cdot)' = \frac{d}{dN}(\cdot)$. Given some reference attractor, this implies that we can label each attractor by a continuous parameter c with the corresponding map

$$\rho_c = \rho_0 + cM_p, \quad (\theta_c^2)' = \frac{(\theta_0^2)'}{(1 + cM_p/\kappa)^2}, \quad (3.6)$$

This transformation identifies all attractors. Without loss of generality we take θ_0' to be negative. Let's figure out what this mapping tells us about the behavior of quantum fluctuations.

In the flat gauge, the isocurvature perturbations \mathcal{S} are associated with $\delta\rho/M_p$ and the curvature perturbations \mathcal{R} with $\frac{\rho}{\sqrt{2\epsilon}M_p}\delta\theta$. Using Eq. 3.5 we can rewrite this as $\mathcal{R} = \frac{\rho_0^2}{4M_p^2}\delta(\theta^2)$, which implies $\delta(\theta^2)' = \frac{4M_p^2}{\rho_0^2}\mathcal{R}' = -(\theta_0^2)'\mathcal{R}$. We aim to find the action of the transformation on the perturbations. For that purpose we split $\rho = \rho_0 + \mathcal{S}M_p$ and $(\theta^2)' = (\theta_0^2)'(1 - \mathcal{R}')$. We next determine how a small c changes \mathcal{S} and \mathcal{R}' . In the long wavelength limit every transformed set of perturbations $(\mathcal{S}_c, \mathcal{R}'_c)$ should provide a new solution to the equations of motion. This is because homogeneous perturbations map background solutions onto each other. Therefore, going back to cosmic time,

we expect to find the following *symmetry* for linearized perturbations

$$\mathcal{S} \rightarrow \mathcal{S} + c, \quad \dot{\mathcal{R}} \rightarrow \dot{\mathcal{R}} + \frac{2HM_p}{\kappa}c. \quad (3.7)$$

We will confirm this in § 3.2.3. Furthermore, we expect the isocurvature perturbations to be massless, as a consequence of the shift symmetry of \mathcal{S} .

3.2.3 Power spectrum and single field behavior

To get an intuitive notion of the behavior of perturbations, we employ the δN formalism [74, 220–223], see also Figure 3.2. First, we integrate the attractor equation for θ in Eq. 3.5 by changing time t to e-folds $dN = Hdt$:

$$N = \frac{\rho^2}{4M_p^2}\theta^2 - \frac{1}{2}, \quad (3.8)$$

with N the number of e-folds until the end of inflation (defined by $\epsilon_{\text{end}} = 1$). In the δN formalism, one computes the time evolution of a fluctuation for a given wavenumber k_* from horizon crossing $k_* \sim a_*H_*$ until the end of inflation, by treating it as a homogeneous fluctuation. Therefore, we can simply perturb equation Eq. 3.8 and the curvature perturbation at the end of inflation is given by

$$\mathcal{R}(k_*) = \delta N_* \approx \frac{1}{\sqrt{2\epsilon}M_p}(\rho\delta\theta)_* + \frac{2N_* + 1}{\kappa}\delta\rho_*. \quad (3.9)$$

Remember that the radius of curvature in our toy model is given by $\kappa = \rho_0$. The initial perturbations $(\rho\delta\theta)_*$ and $\delta\rho_*$ are random variables that arise from quantum fluctuations with typical amplitude $\frac{H_*}{2\pi}$ at horizon crossing. Neglecting their initial cross-correlation we find the following spectrum of curvature perturbations

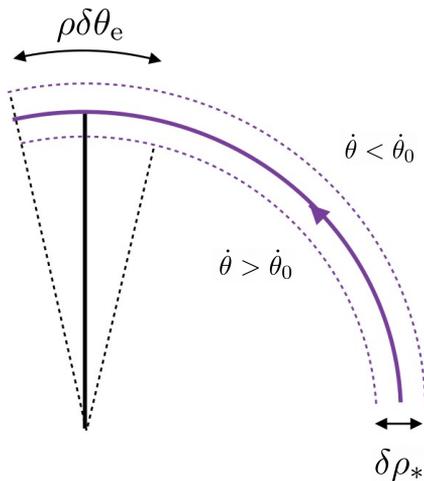
$$P_{\mathcal{R}}(k_*) \approx \frac{H_*^2}{4\pi^2 M_p^2} \left(\frac{1}{2\epsilon_*} + \frac{4N_*^2 M_p^2}{\kappa^2} \right). \quad (3.10)$$

We confirm this result in § 3.5 and compare with the *Planck* data in § 3.6.

In the limit of ‘small’ radius of curvature, i.e.

$$1 \leq \frac{\kappa^2}{M_p^2} \ll 8\epsilon_* N_*^2 \approx 4N_*, \quad (3.11)$$

the second term in Eq. 3.10 dominates and the final power spectrum is determined by the isocurvature perturbations. The lower bound comes from assuming the validity of our simple computation, see § 2.3. Moreover, since



Figuur 3.2: Visualization of the growth of curvature perturbations along the inflationary trajectory. The inflationary trajectory corresponds to the solid purple line and proceeds in the angular direction. In the δN formalism, we consider the effect of an initial constant perturbation $\delta\rho_*$ at the time of horizon crossing on the duration of inflation. The perturbed trajectory proceeds in the angular direction as well, but has smaller or larger radius of curvature (lower and upper purple dashed segment, respectively). The corresponding orbital velocity $\rho\dot{\theta}$ is larger for the lower curve, therefore inflation ends earlier compared to the background trajectory. On the other hand inflation ends later on the upper curve. This results in nonzero fluctuations in the curvature perturbations $\mathcal{R} = \delta N \sim \rho\delta\theta_e$. The amplitude of curvature fluctuations keeps growing after horizon crossing until the end of inflation. Moreover, the smaller the radius of curvature of the trajectory, the larger the final amplitude of curvature perturbations.

the isocurvature perturbations have a constant amplitude, they will be dynamically suppressed with respect to the curvature perturbations

$$\frac{P_S}{2\epsilon P_{\mathcal{R}}} = \frac{1}{1 + \frac{4M_p^2}{\kappa^2} N_*} \ll 1. \quad (3.12)$$

Hence, at the linear level we recover single field predictions. This happens because the final spectrum is generated by a single degree of freedom, namely the isocurvature perturbation. Therefore, this two-field model of inflation behaves like single-field inflation.

Finally, we estimate the amplitude of the bispectrum using the δN forma-

lism. In the limit of small radius of curvature Eq. 3.11, we find

$$f_{NL}^{\delta N} \approx \frac{5}{6} \frac{N_{\rho\rho}}{N_\rho^2} = \frac{5}{6} \epsilon_* . \quad (3.13)$$

It is slow roll suppressed. Therefore, the phenomenology is very similar to that of single field inflation. To understand whether these models obey the consistency relation $f_{NL} = \frac{5}{12}(1 - n_s)$ in the squeezed limit, we also need to compute the intrinsic bispectrum. We present the full in-in computation of the bispectrum in [224]. This will tell us if these models are distinguishable from single field inflation.

3.3 Ultra-light Orbital inflation

In this section we derive a general family of two-field models that admit ultra-light Orbital Inflation. The defining feature of ultra-light Orbital Inflation is that inflation proceeds along an isometry direction of the field metric at *arbitrary radius*. Therefore, we can follow the same derivation as Orbital Inflation presented in § 2.4.2 based on the Hamilton-Jacobi formalism [74, 75, 77, 78]. However, the difference is that we have to assume that the Hamilton-Jacobi equation Eq. 2.44 and constraint equation Eq. 2.46 are satisfied *globally*, or at least for a finite range of values of ρ .

The two equations we should solve simultaneously are the Hamilton-Jacobi equation Eq. 2.44 and constraint equation Eq. 2.46. Using Eq. 2.49 they simplify to

$$3H^2 M_p^2 = V + 2M_p^4 \frac{H_\theta^2}{f} , \quad (3.14a)$$

$$3HH_\rho - 2M_p^2 \frac{H_\theta H_{\rho\theta}}{f(\rho)} , \quad (3.14b)$$

for *any* value of ρ . The constraint equation implies that H has to be a function of θ only. Assuming the opposite, namely that $H_\rho \neq 0$ for *some* value of ρ , we can rewrite the constraint as

$$(\ln H)_\theta (\ln H_\rho)_\theta = \frac{3f}{2M_p^2} . \quad (3.15)$$

Since the right hand side is independent of θ , we can only allow for a solution of the form $(\ln H)_\theta \sim g(\theta)$ and $(\ln H_\rho)_\theta \sim \frac{1}{g(\theta)}$, which is impossible to solve. We conclude that H does not depend on ρ .

Therefore, we conclude that a two-field inflationary model of the form

$$V = 3H^2 M_p^2 - 2M_p^4 \frac{H_\theta^2}{f(\rho)}, \quad G_{ab} = \begin{pmatrix} f(\rho) & 0 \\ 0 & 1 \end{pmatrix} \quad (3.16)$$

with $H = H(\theta)$, admits ultra-light Orbital Inflation. The background dynamics of ultra-light Orbital Inflation is governed by [Eq. 2.49](#):

$$T^a = \frac{1}{\sqrt{f}}(-1, 0) \quad \text{and} \quad N^a = (0, 1) \quad (\text{because } \dot{\rho} = 0), \quad (3.17a)$$

$$\dot{\theta} = -2M_p^2 \frac{H_\theta}{f}, \quad \epsilon = \frac{2M_p^2 H_\theta^2}{f H^2}, \quad \kappa = \frac{2f}{\partial_\rho f}, \quad \mathbb{R} = \frac{2}{\kappa^2} - \frac{f_{\rho\rho}}{f}. \quad (3.17b)$$

3.3.1 Symmetry argument and massless isocurvature perturbations

As alluded to in the introduction, in ultra-light Orbital Inflation we expect the isocurvature perturbations to be massless (see [footnote 1](#)). Their effective mass is indeed zero

$$\mu^2 = V_{NN} + \epsilon H^2 M_p^2 \left(\mathbb{R} + \frac{6}{\kappa^2} \right) = 0. \quad (3.18)$$

The first term is the standard Hessian term plus geometrical corrections $V_{NN} \equiv N^a N^b (V_{ab} - \Gamma_{ab}^c V_c)$. The effective mass μ receives centrifugal and geometrical corrections, because the inflationary background solution is time-dependent. For ultra-light Orbital Inflation we have $V_{NN} = V_{\rho\rho}$. Using the properties of the inflationary background solution [Eq. 3.17](#) we find that the three terms cancel exactly.

The background dynamics hints about the existence of a shift symmetry for perturbations. Like we argued for the toy model in [§ 3.2.2](#), the map [Eq. 3.6](#) that relates all background trajectories is generalized to

$$\rho_c = \rho_0 + cM_p, \quad (G(\theta_c))' = (G(\theta_0))' \frac{f(\rho_0)}{f(\rho_0 + cM_p)}. \quad (3.19)$$

Here $G(\theta)$ is the primitive of $G_\theta = \frac{H}{H_\theta}$. Moreover, in the flat gauge we can write $\mathcal{R} = \sqrt{\frac{f(\rho)}{2cM_p^2}} \delta\theta = \frac{f(\rho)}{2M_p^2} \delta G$. This implies $\delta G' = -G'_0 \mathcal{R}$. We Taylor expand the transformation of the angular velocity for small $\frac{cM_p f_\rho}{f} = \frac{2cM_p}{\kappa}$ to linear order. This yields $G'_c = G'_0 \left(1 - \frac{2cM_p}{\kappa} \right)$. Therefore, the same arguments as in

§ 3.2.2 apply and we are led to expect the combined shift symmetry of linear perturbations

$$\mathcal{S} \rightarrow \mathcal{S} + c, \quad \dot{\mathcal{R}} \rightarrow \dot{\mathcal{R}} + \frac{2HM_p}{\kappa}c. \quad (3.20)$$

We confirm our intuition in § 3.5.

3.4 Stability

Our results rely on the fact that the inflationary trajectory is one of the exact solutions, which we said was an attractor. We now demonstrate neutral stability of the exact solutions.

What do we mean exactly with neutral stability in our dynamical system? We have seen that there is a continuous set of orbital solutions parametrized by ρ_0 . Moreover, we said that normal perturbations move us freely between these ‘attractors’, so the system is clearly not stable in the usual sense. The property we need to prove is that small perturbations shift us to another inflationary solution $\dot{\rho} = 0$.

Each attractor solution Eq. 3.17 corresponds to a point in the $(\dot{\rho}, \dot{\theta})$ plane. Unfortunately, these points are all different and lie on a curve. Moreover, we prefer to do the stability analysis in efolds rather than time, since we expect the Hubble friction to play a crucial role. Therefore, we introduce the variables

$$x(\theta, \rho, \theta', \rho') \equiv \frac{fH}{M_p^2 H_\theta} \theta' - 2 \frac{f}{f_\rho} \frac{\rho'}{M_p^2} + 2, \quad (3.21a)$$

$$y(\theta, \rho, \theta', \rho') \equiv \frac{fH}{M_p^2 H_\theta} \theta' + 2, \quad (3.21b)$$

$$z(\theta, \rho) \equiv \frac{fH^2}{M_p^2 H_\theta^2} - 2/3 \quad (3.21c)$$

Here a prime denotes a derivative with respect to efolds $(..)' = \frac{d}{dN}(..)$. Remember that $H = H(\theta)$ and $f = f(\rho)$. Our definition of stability now amounts to the presence of a fixed point at $(x, y) = (0, 0)$.

We choose these specific variables, because it turns out that the stability of this system is non-trivial to prove analytically. If we simply perturb the field equations we will find zero eigenvalues associated with the perturbations that move us between attractors. Moreover, it is not obvious how to find variables such that the linearized system of perturbations becomes diagonal.

The definition of x and y above are based on an observation for the models where the Hubble parameter is linear in θ . For $H \sim \theta$ the potential in Eq. 3.16 satisfies the following scaling relation

$$\theta V_\theta - 2 \frac{f}{f_\rho} V_\rho = 2V. \quad (3.22)$$

This ensures that the equations for x and y diagonalize at the linear level. This we can use to prove linear stability for the models $H \sim \theta$, which we show in Eq. 3.4.1. In fact it turns out to apply to any power law $H \sim \theta^n$. Moreover, we argue that neutral stability also applies to more general models.

3.4.1 Linear stability analysis

The first step is to rewrite the field equations and second Friedmann equation in terms of the x , y , ρ and z variables. The equations of motion read

$$\begin{aligned} x' + (3 - \epsilon)x + \left(2 \left(\frac{f}{f_\rho} \right)_\rho - g(\theta) \right) \left(\frac{\rho'}{M_p} \right)^2 + \frac{2(z + 2/3)}{z} g(\theta) (\epsilon - \epsilon_0) &= 0, \\ y' + (3 - \epsilon)y + \frac{2}{z} \left(-\frac{1}{3} \left(\frac{\rho'}{M_p} \right)^2 - \frac{1}{2} y^2 + 2y \right) - g(\theta) \left(\frac{\rho'}{M_p} \right)^2 + \frac{2(z + 2/3)}{z} g(\theta) (\epsilon - \epsilon_0) &= 0, \\ z' = 2(y - 2)(1 - g(\theta)) + \left(\frac{M_p f_\rho}{f} \right)^2 \frac{y - x}{2} \left(z + \frac{2}{3} \right), \\ \frac{\rho'}{M_p} = \frac{M_p f_\rho}{f} \frac{y - x}{2}, \\ \epsilon = \frac{1}{2} \frac{(y - 2)^2}{z + 2/3} + \frac{M_p^2 f_\rho^2}{f^2} \frac{(x - y)^2}{8}, \end{aligned}$$

where $\epsilon_0 = \frac{2}{z + 2/3}$. All the terms in brackets are combined to be manifestly zero on the attractor. We introduced the model specific function $g(\theta) \equiv \frac{H H_{\theta\theta}}{H_\theta^2}$. Note that $g(\theta)$ is in general a function of z and ρ , but it reduces to a constant in the case when we have a power law $H(\theta) \sim \theta^n$, and it is zero for $n = 1$.

In terms of the four variables, ultra-light Orbital Inflation is given by $(x, y, z', \rho') = (0, 0, -4(1 - g(\theta)), 0)$, and we would like to prove that this is the attractor solution. It will be sufficient to show that $(y, \rho') = (0, 0)$ is a fixed point. Note that the friction term is very large during inflation. We can already see that, without the friction, the system would be unstable. Let's figure out if the friction is large enough to render the system stable.

To study the stability of the point $(y, \rho') = (0, 0)$, we linearly perturb the

equations around the desired attractor with $\epsilon = \frac{2}{z+2/3}$. We get

$$\delta x' + \left(3 - \frac{2}{z+2/3}\right) \delta x - \frac{4g(\theta)}{z} \delta y = 0, \quad (3.24a)$$

$$\delta y' + \left(3 - \frac{2}{z+2/3} + \frac{4(1-g(\theta))}{z}\right) \delta y = 0, \quad (3.24b)$$

$$\delta z' = 2(1-g(\theta))\delta y + \left(\frac{M_p f_\rho}{f}\right)^2 \frac{\delta y - \delta x}{2} \left(z + \frac{2}{3}\right), \quad (3.24c)$$

$$\frac{\delta \rho'}{M_p} = \frac{M_p f_\rho}{f} \frac{\delta y - \delta x}{2}. \quad (3.24d)$$

Surprisingly, the linearized system of perturbations is very simple for any $g(\theta)$. In particular, for constant $g(\theta)$ we can explicitly prove stability. We show this in a moment in [Eq. 3.4.1](#). For a more general function we have to express $g(\theta)$ in terms of z and ρ and integrate the equations numerically. However, inspection by eye suggests that generically we may expect the system to be stable. If $(1-g(\theta))$ takes values of order 1 and does not vary too rapidly, then z will take large values during inflation and behave smoothly as well. In that case we see from [Eq. 3.24a](#) and [Eq. 3.24b](#) that $\delta x'$ and $\delta y'$ are dominated by the friction terms $-3\delta x$ and $-3\delta y$ respectively. Therefore, we expect both of them to decay like e^{-3N} . Finally [Eq. 3.24d](#) then implies that we quickly converge to the fixed point.

Power law inflation $H \sim \theta^n$

In the case of power law inflation with $1-g(\theta) = \frac{1}{n}$ we can integrate the δy equation [Eq. 3.24b](#), using $z = z_0 - \frac{4}{n}N$. This we can then use to solve for δx as well. We find the following solution

$$\delta y = \delta y_0 \frac{z}{z_0} \left(\frac{2+3z_0}{2+3z}\right)^{n/2} e^{-3N}, \quad (3.25a)$$

$$\delta x = \delta x_0 \left(\frac{2+3z_0}{2+3z}\right)^{n/2} e^{-3N} + \delta y_0 \frac{4(n-1)N}{n} \left(\frac{2+3z_0}{2+3z}\right)^{n/2} e^{-3N}. \quad (3.25b)$$

Plugging these solutions back into [Eq. 3.24d](#) we conclude that $(y, \rho') = (0, 0)$ is a fixed point. This proves stability for power law inflation.

3.5 Inflation with massless isocurvature perturbations

In this section we recap the relevant results from the linear perturbations analysis in two-field inflation. In particular, we focus on the limit that the isocurvature perturbations are massless [138] (see footnote 1).

3.5.1 Quadratic action of perturbations

The full action of our two-field models of inflation has the general form

$$S = \frac{1}{2} \int d^4x \sqrt{-g} \left[M_p^2 R - G_{ab} \partial\phi^a \partial\phi^b - 2V(\phi^a) \right]. \quad (3.26)$$

with the field metric and potential from Eq. 3.16. We study the dynamics of linearized perturbations around the exact homogeneous inflationary background solution.

In the flat gauge, the scalar metric perturbation is set to zero. The co-moving curvature perturbation \mathcal{R} is then given by the projection of the field perturbation along the inflationary trajectory: $\mathcal{R} = \frac{1}{\sqrt{2\epsilon}M_p} T_a \delta\phi^a$. The isocurvature perturbation \mathcal{S} corresponds to the remaining orthogonal field perturbation to the inflationary trajectory: $\mathcal{S} = \frac{N_a \delta\phi^a}{M_p}$. The quadratic action for perturbations takes the following form

$$S^{(2)} = \frac{1}{2} \int d^4x a^3 M_p^2 \left[2\epsilon \left(\dot{\mathcal{R}} - \frac{2HM_p}{\kappa} \mathcal{S} \right)^2 + \dot{\mathcal{S}}^2 - \mu^2 \mathcal{S}^2 + .. \right] \quad (3.27)$$

The ellipses denote the gradient terms $-(\partial_i \mathcal{S})^2 - 2\epsilon(\partial_i \mathcal{R})^2$. The perturbations are combined in such a way that the entropy mass μ is manifest (see footnote 1).

3.5.2 Massless isocurvature modes and symmetry

In the limit of $\mu^2 = 0$, the quadratic action Eq. 3.27 enjoys a combined shift symmetry [138]

$$\mathcal{S} \rightarrow \mathcal{S} + c, \quad \text{and} \quad \dot{\mathcal{R}} \rightarrow \dot{\mathcal{R}} + \frac{2HM_p}{\kappa} c. \quad (3.28)$$

The combined shift symmetry is exactly as we argued from the background dynamics in § 3.3.1 !

3.5.3 Power spectrum and effective single field behavior

The power spectra of perturbations in the massless limit can be well estimated from the coupled evolution of perturbations on superhorizon scales. The full derivation relies on an in-in computation to account for the coupled evolution of perturbations on subhorizon scales [138]. This gives the same results up to subleading corrections.

When $\mu = 0$, the linearized system of coupled perturbations simplify in the superhorizon limit. The isocurvature perturbation quickly converges to a constant on superhorizon scales where it sources the curvature perturbation. The final dimensionless power spectrum of curvature perturbations in the massless limit is given by (see § 2.3)

$$P_{\mathcal{R}} = \frac{H^2}{8\pi^2\epsilon_*} \left(1 + 2\epsilon_* \left(\int dN \frac{2M_p}{\kappa} \right)^2 \right). \quad (3.29)$$

Note that the power spectrum is completely determined by the isocurvature perturbations if the radius of curvature κ is sufficiently small:

$$2\epsilon_* \left(\int dN \frac{2M_p}{\kappa} \right)^2 \gg 1 \quad (3.30)$$

Using this condition, the spectrum of isocurvature perturbations will be dynamically suppressed as well

$$\frac{P_S}{2\epsilon P_{\mathcal{R}}} \approx \left(2\epsilon_* \int dN \frac{2M_p}{\kappa} \right)^{-2} \ll 1. \quad (3.31)$$

We confirm all intuitive results for the toy model in § 3.2.3 .

Summarizing, two-field inflation with massless isocurvature perturbations, together with a trajectory of sufficiently small radius of curvature mimics the predictions of single field inflation at the level of the power spectra.

3.6 Phenomenology

We now turn to the predictions for the spectral tilt n_s and the tensor-to-scalar ratio r . In this section we use the result from § 3.5 for the power spectrum in the limit that the isocurvature perturbations are massless and the radius of curvature of the trajectory is constant:

$$P_{\mathcal{R}} = \frac{H^2}{8\pi^2\epsilon_*} \left(1 + 2\epsilon_* \left(\frac{2M_p}{\kappa} \right)^2 N_*^2 \right). \quad (3.32)$$

During ultra-light Orbital Inflation we can use this expression, since $\kappa = \frac{2f}{\delta\rho f}$ is constant on the trajectory.

The spectral tilt and tensor-to-scalar ratio are (to good approximation) given by

$$n_s = \frac{\partial \ln P_{\mathcal{R}}}{\partial N}, \quad (3.33a)$$

$$r = \frac{16\epsilon_*}{\left(1 + \frac{8\epsilon_* M_p^2}{\kappa^2} N_*^2\right)}, \quad (3.33b)$$

where we have to be careful to use $\frac{\partial N_*}{\partial N} = -1$, since N_* counts the number of e-folds backwards. The predictions depend on the function $H(\theta)$. Like in single field slow roll inflation, this function determines how ϵ and η scale with N_* .

For concreteness, we consider power law inflation $H \sim \theta^p$. Using the exact solution of ultra-light Orbital Inflation Eq. 3.17, we integrate the equation of motion for θ and plug it back in the expression for ϵ to find

$$\epsilon_* = \frac{p}{2N_* + p}, \quad \eta_* \equiv \frac{\epsilon'_*}{\epsilon_*} = \frac{2}{2N_* + p}. \quad (3.34)$$

Note that these are exactly the same slow-roll parameters as the single field models $V \sim \varphi^{2p} - \frac{2}{3}$ (see § 1.2.3), which also have the exact homogeneous attractor solution $\dot{\phi} = -2pM_p^2\phi^{p-1}$. Using Eq. 3.33b, the predictions for n_s and r are therefore well approximated by

$$n_s \approx 1 - \frac{p+1}{N_*} - \frac{4p}{\frac{\kappa^2}{M_p^2} + 4pN_*}, \quad (3.35a)$$

$$r \approx \frac{8p\kappa^2}{N_*\kappa^2 + 4pM_p^2N_*^2}. \quad (3.35b)$$

We plot these analytical predictions against the Planck 1σ and 2σ contours [115] in Figure 3.3. The radius of curvature κ^2/M_p^2 varies between 1 and 10^5 . Moreover, we take N_* between 50 and 60. Our toy model from § 3.2 corresponds to the purple contour and resembles the predictions of chaotic inflation when $\kappa \rightarrow \infty$. Furthermore, we show the predictions of linear inflation (red contour) and $\phi^{2/3}$ inflation (orange contour).

The first thing to notice is that our results for n_s and r only depend on the value of κ and are therefore insensitive to the details of the field metric. We might expect that the higher order correlation functions retain this information. We estimate the amplitude of the bispectrum in the δN formalism like

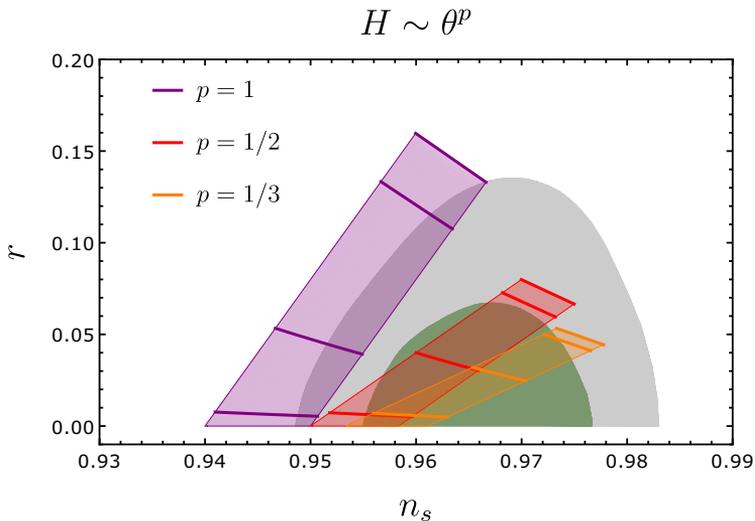


Figure 3.3: The analytical predictions Eq. 3.35 for (n_s, r) compared to the *Planck* 1σ and 2σ contours [115]. The colors of the contours correspond to different values of p , which determines the scaling of the Hubble parameter $H \sim \theta^p$. The slow roll parameters are exactly as those of a single field model with potential $V \sim \theta^{2p} - \frac{2}{3}$. We show the predictions for wavenumbers which cross the horizon 50 – 60 e-folds before the end of inflation. The predictions for (n_s, r) depend on the value of the radius of curvature κ of the inflationary trajectory. We vary κ^2/M_p^2 between 1 and 10^5 , and indicate the values $(10, 10^2, 10^3, 10^5)$ with thick lines (from bottom to top).

we did in § 3.2.3, which gives

$$f_{NL}^{\delta N} \approx \frac{5}{6} \frac{\epsilon_*}{p} \left(1 - \frac{\kappa^2 \mathbb{R}}{2} \right). \quad (3.36)$$

Here we assumed the limit of ‘small’ radius of curvature $\kappa^2 \ll \frac{2p^2}{\epsilon} M_p^2$. We see that the bispectrum has the potential to distinguish between different field spaces through its dependence on the Ricci scalar \mathbb{R} . We need a full in-in computation of the bispectrum [224] to verify this result.

Coming back to the predictions for n_s and r , we find they are pushed downwards and to the left in the (n_s, r) plane as the radius of curvature decreases. Therefore, in the case of power law inflation only for small p the predictions remain within the *Planck* contours. However, we saw in Chapter 2 that the predictions are sensitive to the value of μ^2/H^2 . In particular, if a small mass is generated, we arrive at a different conclusion.

Universality of multi-field α -attractors

In this chapter we investigate two-field cosmological α -attractors, which are characterized by a hyperbolic field metric. The important property of the single field realization of α -attractors is that, in the limit of small $\alpha < \mathcal{O}(10)$, their predictions converge to $n_s - 1 \simeq -\frac{2}{N}$ and $r \simeq \frac{12\alpha}{N^2}$, irrespective of the potential. In the two-field case, we find that the inflationary predictions show universal behavior too, insensitive to significant modifications of the potential.

In the simplest supergravity embedding of α -attractors, the potential depends on the complex scalar $Z = \rho e^{i\theta}$, living on a disk with $\rho < 1$. Moreover, the fields span a hyperbolic field space with Ricci curvature $\mathbb{R} = -\frac{2}{3\alpha}$. In the single field scenario, in which the angular field is stabilized, the universality of the predictions can be ultimately traced back to the radial stretching introduced by the hyperbolic geometry as we approach the boundary $\rho \sim 1$.

If both ρ and θ are light during inflation, the angular velocity $\dot{\theta}$ is exponentially suppressed, due to the hyperbolic geometry, and inflation proceeds (almost) in the radial direction. The angular field will *not* roll down to its minimum, but instead it is "*rolling on the ridge*". This is illustrated in Figures 4.3 and 4.4. Nevertheless, the trajectory is curved and the inflationary dynamics is truly multi-field. The multi-field effects conspire in such a way that the predictions remain unchanged with respect to the single field scenario.

This chapter is organized as follows. In Section 4.2 we present a new supergravity embedding of the $\alpha = 1/3$ two-field model. We study its inflationary dynamics, and elaborate on the "rolling on the ridge" behaviour in Section 4.3. Next, we work out the universal predictions for primordial perturbations in Section 4.4, and leave the details of the full multi-field analysis for Appendix 4.B. We extend this result to general values of α and work out the constraints on the potential to ensure the universality of the predictions in Section 4.5 and Appendix 4.A. Section 4.6 is for summary and conclusions.

This chapter is based on [139]:

Universality of multi-field α -attractors, A. Achúcarro, R. Kallosh, A. Linde,

D-G. Wang and Y. Welling, JCAP **1804** (2018) 07, 028, (arXiv:1711.09478 [hep-th]).

4.1 Introduction

UV embeddings of inflation typically contain multiple scalar fields beside the inflaton. If the additional fields are stabilized, we can integrate them out to find effectively single field inflation. On the other hand, if the additional fields remain light during inflation, we should take into account the full multi-field dynamics. Planck [115, 118] puts tight constraints on these inflationary models, therefore we should understand which model-building ingredients are important to ensure compatibility with the data. In particular, both the geometry of field space and the curvature of the inflationary trajectory play a very important role in determining the observables. In this paper we focus on the special role played by hyperbolic geometry.

A notable example are the α -attractor models, a relatively simple class of inflationary models that have a single scalar field driving inflation. In the simplest supergravity embedding of these models, the potential depends on the complex scalar $Z = \rho e^{i\theta}$, where Z belongs to the Poincaré disk with $|Z| = \rho < 1$ and the kinetic terms read¹

$$3\alpha \frac{\partial_\mu \bar{Z} \partial^\mu Z}{(1 - Z\bar{Z})^2} + \dots \quad (4.1)$$

In many versions of these models, the field θ is heavy and stabilized at $\theta = 0$, so that the inflationary trajectory corresponds to the evolution of the single field ρ . An important property of these models is that their cosmological predictions are stable with respect to considerable deformations of the choice of the potential of the field ρ : $n_s \approx 1 - \frac{2}{N}$, $r \approx \frac{12\alpha}{N^2}$ [225–233]. These predictions are consistent with the latest observational data for $\alpha < O(10)$.

In the single-field realizations, the universality of these predictions can be ultimately traced back to the radial stretching introduced by the geometry (4.1) as we approach the boundary $\rho \sim 1$. On the other hand it is clear that, in the two-field embedding in terms of Z , the stretching also affects the “angular” θ -direction and this begs the question whether perhaps there is a regime where the predictions for the inflationary observables are also fairly insensitive to the details of the angular dependence of the potential. In this paper we answer this question in the affirmative for sufficiently small $\alpha \lesssim O(1)$.

¹Alternatively, $3\alpha \frac{\partial T \partial \bar{T}}{(T + \bar{T})^2}$, where $T = \frac{1+Z}{1-Z}$.

A particularly interesting case is $\alpha = 1/3$, where a class of supergravity embeddings are known to possess an additional symmetry, which makes both ρ and θ light [230]. This means we cannot integrate out the angular field and we have to take into account the full multi-field dynamics. We will show that, in contrast with the naive expectation, the cosmological predictions of the simplest class of such models are very stable not only with respect to modifications of the potential of the field ρ , but also with respect to strong modifications of the potential of the field θ . Importantly, we have to account for the full multi-field dynamics [198–203, 234, 235] in order to obtain the right results². The predictions coincide with the predictions of the single-field α -attractors for $\alpha = 1/3$: $n_s \approx 1 - \frac{2}{N}$, $r \approx \frac{4}{N^2}$. It was emphasized in [230] that for $3\alpha = 1$, the geometric kinetic term

$$\frac{dZd\bar{Z}}{(1 - Z\bar{Z})^2} \quad (4.2)$$

has a fundamental origin from maximal $\mathcal{N} = 4$ superconformal symmetry and from maximal $\mathcal{N} = 8$ supergravity. Also the single unit size disk, $3\alpha = 1$, leads to the lowest B-mode target which can be associated with the maximal supersymmetry models, M-theory, string theory and N=8 supergravity, see [231, 232] and [233].

More generally, we will also show that, for sufficiently small values of $\alpha < O(1)$, the class of potentials exhibiting universal behaviour becomes very broad, and in particular it includes potentials with $\frac{1}{\rho}V_\theta \sim V_\rho \sim V$.

Our results lend support to the tantalizing idea, recently explored in some detail in [138] and building on earlier works in [215, 237–240], that multi-field inflation on a hyperbolic manifold may be compatible with current observational constraints *without the need to stabilize all other fields besides the inflaton*. Since axion-dilaton moduli systems with the geometry (4.1) are ubiquitous in string compactifications, this observation could have important implications for inflationary model building.

Although at first sight the universality found here resembles a similar result obtained in the theory of multi-field conformal attractors [241] for $\alpha = 1$, the reason for our new result is entirely different. In the model studied in [241], the light field θ evolved faster than the inflaton field, so it rapidly rolled down to the minimum of the potential with respect to the field θ , and the subsequent evolution became the single-field evolution driven by the inflaton field. The observable e-folds are in the latter, single-field regime. On the other hand, in the class of models to be discussed in our paper, the angular velocity $\dot{\theta}$ is exponentially suppressed, due to the hyperbolic geometry, and inflation

²See [236] for a recent review and references there.

proceeds (almost) in the radial direction. The angular field will *not* roll down to its minimum, but instead it is "*rolling on the ridge*". This is illustrated in Figures 4.3 and 4.4. Nevertheless, the trajectory is curved and the inflationary dynamics is truly multi-field.

Multi-field models of slow-roll inflation based on axion-dilaton systems have been studied for some time [242, 243]. However, it is only fairly recently that the very important role played by the hyperbolic geometry for multi-field inflation is being recognized (see, e.g. [138, 230, 238, 244–248]). Unlike in previous works, here we choose to be agnostic about the potential, and derive the conditions that will guarantee universality of the inflationary predictions for the two-field system.

4.2 α -attractors and their supergravity implementations

There are several different formulations of α -attractors in supergravity. One of the first formulations [227] was based on the theory of a chiral superfield Z with the K potential corresponding to the Poincaré disk of size 3α ,

$$K = -3\alpha \ln(1 - Z\bar{Z} - S\bar{S}), \quad (4.3)$$

and superpotential

$$W = S f(Z)(1 - Z^2)^{\frac{3\alpha-1}{2}}, \quad (4.4)$$

where $f(Z)$ is a real holomorphic function. It is possible to make the field S vanish during inflation, either by stabilizing it, or by making it nilpotent [249]. Either way, the kinetic term for Z is

$$3\alpha \frac{dZ d\bar{Z}}{(1 - Z\bar{Z})^2}. \quad (4.5)$$

The field Z can be represented as $e^{i\theta} \tanh \frac{\varphi}{\sqrt{6\alpha}}$, where φ is a canonically normalized inflaton field. In the simplest models of this class, the mass of the field θ in the vicinity of $\theta = 0$ during inflation is given by

$$m_\theta^2 = 2V \left(1 - \frac{1}{3\alpha}\right), \quad (4.6)$$

up to small corrections proportional to slow roll parameters. In particular, for the simplest models with $\alpha > 1/3$ one finds $m_\theta^2 > 0$, which means that the field θ is stabilized at $\theta = 0$. Meanwhile for $\alpha > 2/5$ one has $m_\theta^2 = V/3 \geq H^2$ where H is the Hubble constant. This means that the field θ for $\alpha \geq 2/5$ is

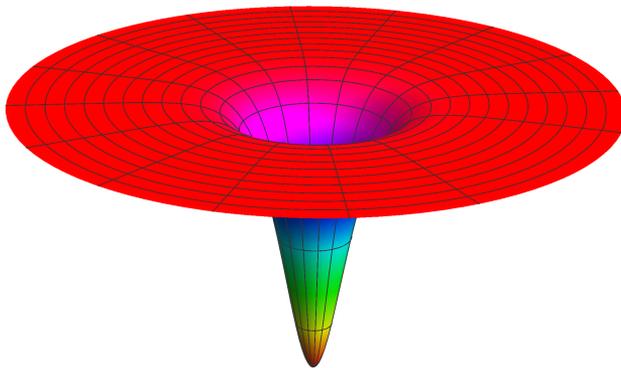


Figure 4.1: The θ -independent $3\alpha = 1$ T-model potential $V(\varphi) = m^2 \tanh^2 \frac{\varphi}{\sqrt{2}}$.

strongly stabilized, and the only dynamical field during inflation is the inflaton field φ with the potential

$$V = \left| f\left(\tanh \frac{\varphi}{\sqrt{6\alpha}}\right) \right|^2. \tag{4.7}$$

Meanwhile for $3\alpha \approx 1$ one finds that during inflation $|m_\theta^2| \ll H^2$. As an example, the potential V for $f(Z) = mZ$ does not depend on θ at all:

$$V = m^2 \tanh^2 \frac{\varphi}{\sqrt{6\alpha}}, \tag{4.8}$$

see Figure 4.1.

Later on, it was found [250] that one can strongly stabilize the field θ for all α and reduce investigation of the cosmological evolution to the study of the single inflaton field in the models with a somewhat different K potential,

$$K = -3\alpha \ln \frac{1 - Z\bar{Z}}{|1 - Z^2|} + S\bar{S}, \tag{4.9}$$

and superpotential

$$W = S f(Z), \tag{4.10}$$

which yields the same inflaton potential (4.7) for $\theta = 0$.

This considerably simplifies investigation of inflationary models. An advantage of this K potential is its manifest shift symmetry: it vanishes along the direction $Z = \bar{Z}$, corresponding to $\theta = 0$.

The next step was the construction of the anti-D3 brane induced geometric inflationary models with arbitrary α with a stabilized field θ [233] (see also

[251]). The K function is

$$\mathcal{G} = \ln |W_0|^2 - 3\alpha \ln \frac{1 - Z\bar{Z}}{|1 - Z^2|} + S + \bar{S} + G_{S\bar{S}}(Z, \bar{Z})S\bar{S} , \quad (4.11)$$

where the field S is nilpotent, with the metric

$$G_{S\bar{S}}(Z, \bar{Z}) = \frac{|W_0|^2}{\mathbf{V}(Z, \bar{Z}) + 3|W_0|^2} . \quad (4.12)$$

The bosonic part of the supergravity action is

$$g^{-1}\mathcal{L} = 3\alpha \frac{dZ d\bar{Z}}{(1 - Z\bar{Z})^2} - \mathbf{V}(Z, \bar{Z}) . \quad (4.13)$$

Note that the Z -part of the K potential has the inflaton shift symmetry at $Z = \bar{Z}$, as was shown in [250]. The potential is

$$\mathbf{V}(Z, \bar{Z}) = V(Z, \bar{Z}) + |F_S|^2 - 3|W_0|^2 = V(Z, \bar{Z}) + \Lambda . \quad (4.14)$$

Here, as in all models in [233], $V(Z, \bar{Z})$ is a function of Z and \bar{Z} which is regular at the boundary $Z\bar{Z} = 1$ and which vanishes at the minimum at $Z = 0$, so that

$$\mathbf{V}(Z, \bar{Z}) \Big|_{Z=0} = |F_S|^2 - 3|W_0|^2 \equiv \Lambda . \quad (4.15)$$

The scale of supersymmetry breaking due to the nilpotent field S is

$$e^{\mathcal{G}} \mathcal{G}_S \mathcal{G}^{S\bar{S}} \mathcal{G}_{\bar{S}} \Big|_{Z=0} = |F_S|^2 , \quad (4.16)$$

and the gravitino mass is $m_{3/2}^2 \Big|_{Z=0} = |W_0|^2$. The angular field in these models is heavy, by construction, inflation takes place at $Z = \bar{Z}$.

This formulation is valid for any α . However, subsequent investigations have revived interest in the specific models with $3\alpha = 1$ corresponding to the unit size disk [230], and in the possibility to describe models originating from merger of several unit size disks, which may lead to α -attractors with $3\alpha = 1, 2, 3, \dots, 7$ [231–233]. It has been argued that these models provide some of the better motivated targets for the future B-mode searches. Therefore it would be interesting to revisit all versions of these models, including the original versions with light, non-stabilized fields θ [230], since such models may exhibit a greater degree of symmetry, as shown in Figure 4.1. It would be particularly interesting to find the corresponding generalization of the anti-D3 brane induced geometric inflationary models described above, applicable specifically to models with $3\alpha = 1$.

We find this new formulation by returning to the original K frame with the axion shift symmetry, $K = -\ln(1 - Z\bar{Z})$, instead of $K = -\ln\frac{1-Z\bar{Z}}{|1-Z\bar{Z}|}$. In this way the mass of the θ -field will become light and we will have a two-field evolution on the disk of unit size $3\alpha = 1$. The K function which provides the action

$$g^{-1}\mathcal{L} = \frac{dZd\bar{Z}}{(1 - Z\bar{Z})^2} - V(Z, \bar{Z}) - \Lambda \quad (4.17)$$

will be taken in the following form:

$$\mathcal{G} = \ln|W_0|^2 - \ln(1 - Z\bar{Z}) + S + \bar{S} + G_{S\bar{S}}(Z, \bar{Z})S\bar{S}. \quad (4.18)$$

Here the metric of the nilpotent superfield is

$$G_{S\bar{S}}(Z, \bar{Z}) = \frac{|W_0|^2}{(1 - Z\bar{Z})\left(|F_S|^2 + V(Z, \bar{Z})\right) + 2|W_0|^2 Z\bar{Z}}. \quad (4.19)$$

It is different from the simpler version of $G_{S\bar{S}}$ in Equation (4.12), but the K potential $-\ln(1 - Z\bar{Z})$ as a function of Z, \bar{Z} is simpler here. Moreover, the Z -part of the K potential has an axion shift symmetry, it is θ -independent.

One can show that the expression for the scalar potential in this theory is given by

$$\mathbf{V}(Z, \bar{Z}) = V(Z, \bar{Z}) + |F_S|^2 - 3|W_0|^2 = V(Z, \bar{Z}) + \Lambda. \quad (4.20)$$

This result is very similar to Equation (4.14). However, (4.14) correctly represents the inflaton potential only along the inflaton direction $Z = \bar{Z}$. The potential for general values $Z \neq \bar{Z}$ must be calculated by the standard supergravity methods. This complication usually is not important for us since during inflation one can stabilize the fields along the inflaton direction $Z = \bar{Z}$. Meanwhile in our new approach, equation (4.20) gives the full expression for $\mathbf{V}(Z, \bar{Z})$, which is valid for any Z and \bar{Z} on the disk. This is a very special feature of the new formulation, which is valid for $3\alpha = 1$.

During inflation, one can safely ignore the tiny cosmological constant $\Lambda \sim 10^{-120}$, so the potential (4.20) is given by an arbitrary real function $V(Z, \bar{Z})$. In the simplest cases, where V is a function of $Z\bar{Z}$, it does not depend on the angular variable θ , just as the potential in the theory (4.3) (4.4) for $3\alpha = 1$ shown in Figure 4.1. For more general potentials, V may depend on θ , and the potentials can be quite steep with respect to ρ and θ .

The key feature of this class of models, as well as of the models (4.3) (4.4) for $3\alpha = 1$, is that they describe hyperbolic moduli space corresponding to the K potential $K = -\ln(1 - Z\bar{Z})$, with the metric of the type encountered in the description of an open universe, see Equation (4.24) below. As we will see, the slow roll regime is possible for these two classes of theories even for very steep potentials, because of the hyperbolic geometry of the moduli space.

4.3 Dynamics of multi-field α -attractors

Now we come to study inflation with the above theoretical construction. Our starting point is

$$g^{-1}\mathcal{L} = \frac{dZd\bar{Z}}{(1-Z\bar{Z})^2} - V(Z, \bar{Z}) . \quad (4.21)$$

The complex variable on the disk can be expressed as

$$Z = \rho e^{i\theta} , \quad (4.22)$$

where ρ is the radial field and θ is the angular field. In general, the potential $V(\rho, \theta)$ in these variables can be quite complicated and steep. For simplicity, in the following we assume the potential vanishes at the origin $Z = 0$ and is monotonic along the radial direction of the unit disk³, *i.e.* $V_\rho \geq 0$. One natural possibility is $V_\rho \sim V_\theta/\rho \sim V$, which at first glance cannot yield sufficient inflation. However, the hyperbolic geometry of the moduli space makes slow roll inflation possible even if the potential is quite steep.

To see this, and to connect this to a more familiar canonical field φ in $3\alpha = 1$ attractor models where the tanh argument is $\varphi/\sqrt{6\alpha}$, we can use the following relation

$$\rho = \tanh \frac{\varphi}{\sqrt{2}} . \quad (4.23)$$

Therefore, our cosmological models with geometric kinetic terms are based on the following Lagrangian of the axion-dilaton system

$$g^{-1}\mathcal{L} = \frac{1}{2}(\partial\varphi)^2 + \frac{1}{4}\sinh^2(\sqrt{2}\varphi)(\partial\theta)^2 - V(\varphi, \theta) , \quad (4.24)$$

where some choice of the potentials $V(\varphi, \theta)$ will be made depending on both moduli fields. In terms of this new field φ , the corresponding potential near the boundary $\rho = 1$ is exponentially stretched to form a plateau, where φ field becomes light and slow-roll inflation naturally occurs. If we further assume the potential is a function of the radial field only, then we recover the T-model as shown in Figure 4.1. Generally speaking, the potential may also depend on θ , and have ridges and valleys along the radial direction. One simple example is shown in Figure 4.2. Although the θ field can appear heavy in the unit disk coordinates, after stretching in the radial direction, the effective mass in the angular direction is also exponentially suppressed for $\varphi \gg 1$.

For a cosmological spacetime, the background dynamics is described by equations of motion of two scalar fields

$$\ddot{\varphi} + 3H\dot{\varphi} + V_\varphi - \frac{1}{2\sqrt{2}}\sinh\left(2\sqrt{2}\varphi\right)\dot{\theta}^2 = 0 , \quad (4.25)$$

³We leave other interesting cases with non-monotonic potential, such as the Mexican hat potential, for future work [252].

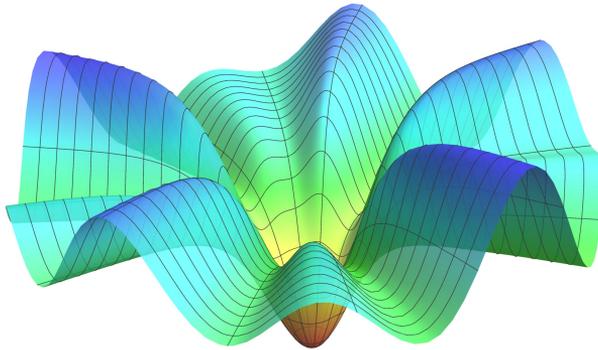


Figure 4.2: A stretched potential with angular dependence

$$\ddot{\theta} + 3H\dot{\theta} + \frac{V_{\theta}}{\frac{1}{2}\sinh^2(\sqrt{2}\varphi)} + \frac{2\dot{\theta}\dot{\varphi}}{\frac{1}{\sqrt{2}}\tanh(\sqrt{2}\varphi)} = 0, \quad (4.26)$$

and the Friedmann equation

$$3H^2 = \frac{1}{2}(\dot{\varphi}^2 + \frac{1}{2}\sinh^2\sqrt{2}\varphi\dot{\theta}^2) + V(\varphi, \theta), \quad (4.27)$$

where $H \equiv \dot{a}/a$ is the Hubble parameter. In such a two-field system with potential as shown in Figure 4.2, one may expect that the inflaton will first roll down from the ridge to the valley, and then slowly rolls down to the minimum along the valley. In the following we will demonstrate, due to the magic of hyperbolic geometry, the dynamics of moduli fields is totally different from this naive picture.

4.3.1 Rolling on the ridge

In single-field α -attractor models, inflation takes place near the edge of the Poincaré disk with $\rho \rightarrow 1$ (or equivalently $\varphi \gg 1$). Here we also focus on the large- φ regime where the potential in the radial direction is stretched to be very flat. As a consequence, the radial derivative of the potential is exponentially suppressed

$$V_{\varphi} \simeq 2\sqrt{2}V_{\rho}e^{-\sqrt{2}\varphi}. \quad (4.28)$$

After a quick relaxation, the fields can reach the slow-roll regime with the Hubble slow-roll parameters

$$\epsilon \equiv -\frac{\dot{H}}{H^2} = \frac{\dot{\varphi}^2 + \frac{1}{2}\sinh^2(\sqrt{2}\varphi)\dot{\theta}^2}{2H^2} \ll 1, \quad \eta \equiv \frac{\dot{\epsilon}}{H\epsilon} \ll 1. \quad (4.29)$$

Thus the kinetic energy of fields is much smaller than the potential, and the $\dot{\theta}\dot{\varphi}$ term in (4.25) is subdominant. Moreover, we assume that the field accelerations $\ddot{\varphi}$ and $\ddot{\theta}$ can be neglected with respect to the potential gradient. The equation of motion for θ is then simplified to

$$\frac{\dot{\theta}}{H} \simeq -8 \frac{V_\theta}{V} e^{-2\sqrt{2}\varphi}. \quad (4.30)$$

This gives us the velocity in the angular direction, which is highly suppressed in the large- φ regime. Substituting the above result in the equation of motion for φ (4.25), we can see that the centrifugal term proportional to $\dot{\theta}^2$ is also suppressed by $e^{-2\sqrt{2}\varphi}$. Thus for $\varphi \gg 1$ this term can be neglected compared to V_φ . Therefore the equation of motion for φ is approximately

$$3H\dot{\varphi} + V_\varphi \simeq 0, \quad (4.31)$$

which is the same as the single field case with slow-roll conditions. Similarly we get the field velocity in the radial direction $\dot{\varphi} \sim e^{-\sqrt{2}\varphi}$, which is much larger than the angular velocity $\dot{\theta}$. This is the main reason for the difference between the slow-roll regime in the present set of models, and in the multi-field conformal attractors studied in [241]. In the conformal attractors, the field θ was rapidly rolling down, whereas here instead of rolling down to the valley first, the scalar fields are *rolling on the ridge* with almost constant θ .

To see this counter-intuitive behaviour clearly, we can look at the flow $(\dot{\varphi}, \dot{\theta})$ in the polar coordinate system. The numerical result of the flow of the fields is shown in Figure 4.3 for the potential from Figure 4.2. As we see, although the potential *looks* chaotic in the angular direction, the fields always roll to the minimum along the ridge, no matter where they start.

However, it is crucial to emphasize that, although $\dot{\theta}$ is highly suppressed and θ is nearly constant, the angular motion is still quite important. In the curved field manifold, since the angular distance is also stretched for large φ , the proper velocity in the angular direction is given by $\frac{1}{\sqrt{2}} \sinh(\sqrt{2}\varphi)\dot{\theta}$. We are encouraged to define a new parameter γ as the ratio between the *physical* angular and radial velocity

$$\gamma \equiv \frac{\sinh(\sqrt{2}\varphi)\dot{\theta}}{\sqrt{2}\dot{\varphi}} \simeq \frac{V_\theta}{V_\varphi}, \quad (4.32)$$

where in the last step we have used large- φ and slow-roll approximations. Since θ hardly evolves and $\rho \simeq 1$ for $\varphi \gg 1$, γ is nearly constant during most period of inflation. This parameter captures the deviation from the single field scenario. For instance, let us look at the potential slow-roll parameter in the radial direction

$$\epsilon_\varphi \equiv \frac{1}{2} \left(\frac{V_\varphi}{V} \right)^2 \simeq \frac{\dot{\varphi}^2}{2H^2}, \quad (4.33)$$

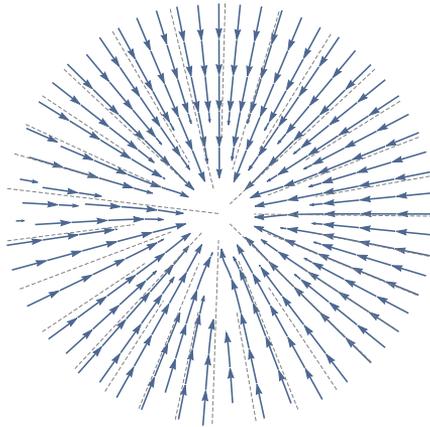


Figure 4.3: The stream of φ and θ fields on the potential with random angular dependence shown in Figure 4.2. The dashed gray lines show the radial directions, while the blue arrows correspond to the field flow, starting at $\varphi_i = 10$.

which is the same with the single field one. Then in our model the full Hubble slow-roll parameter (4.29) is approximately given by

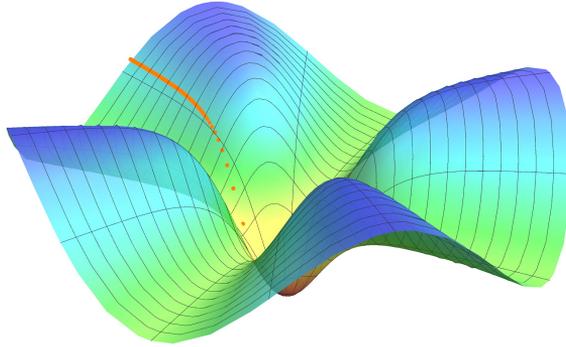
$$\epsilon = (1 + \gamma^2)\epsilon_\varphi . \quad (4.34)$$

Thus a nonzero γ demonstrates the contribution of the angular motion in the evolution of the two-field system. Furthermore, depending on the form of the potential, γ can be $\mathcal{O}(1)$ as we shall show in a toy model later. In such cases, the physical angular motion is comparable to the radial one, and the multi-field effects is particularly important⁴.

In summary, for multi-field α -attractors, there are two subtleties caused by the hyperbolic field space. First of all, the two-field evolution *looks* like the single field case without turning behaviour in the field space. On the other hand, the straight trajectory is an illusion, and the *multi-field* effect can still be significant. In Section 4.4, we will show how these surprising behaviours lead us to the universal predictions for primordial perturbations.

Concluding this subsection, we wish to further explain why “*rolling on the ridge*” is a quite general behaviour in multi-field α -attractors. Besides the aforementioned approximations, importantly we also neglect the centrifugal term in (4.25). Strictly speaking, this requires $V_\varphi \gg \frac{1}{2\sqrt{2}} \sinh(2\sqrt{2}\varphi) \dot{\theta}^2$, which in the large- φ regime is equivalent to the following condition of the

⁴To see the importance of multi-field behaviour, another way is to look at the nonzero turning parameter, which we will discuss in Appendix 4.B.



Figuur 4.4: Rolling on the ridge: the form of the potential is given by the toy model (4.37) with $A = 0.2$, $n = 4$ and initial angle $\theta_i = \pi/8$; the orange dots show a typical background trajectory, while the interval between the neighboring dots corresponds to one e-folding time.

potential

$$\frac{V_\rho}{V} \gg \frac{4}{3} \left(\frac{V_\theta}{V} \right)^2 e^{-\sqrt{2}\varphi} . \quad (4.35)$$

Now we can see, near the boundary of the disk, unless the angular dependence of the potential is exponentially stronger than the radial one, the above condition always holds true and the system evolves as we describe above. Finally let us stress that we have to ensure all our approximations are valid. We collect all conditions on the potential in Appendix 4.A. A natural choice of the potential with $V_\rho \sim V_\theta/\rho \sim V$ certainly satisfies these conditions.

4.3.2 A toy model

Before moving to the calculation of perturbations, let us work out a toy model to further confirm the above analysis. Consider the following potential on the unit disk

$$V(Z, \bar{Z}) = V_0 [Z\bar{Z} + A(Z^n + \bar{Z}^n)] . \quad (4.36)$$

To ensure that it is monotonic in the radial direction of the unit disk we need $A \leq \frac{1}{n}$. Then the condition (4.35) is certainly satisfied. In terms of φ and θ , the potential is given by

$$V(\varphi, \theta) = V_0 \tanh^2 \left(\frac{\varphi}{\sqrt{2}} \right) \left[1 + 2A \cos(n\theta) \tanh^{n-2} \left(\frac{\varphi}{\sqrt{2}} \right) \right] . \quad (4.37)$$

For demonstration, in the following we take $n = 4$, $A = 0.2$ and the initial angle $\theta_i = \pi/8$. We solve the background evolution of this two field system

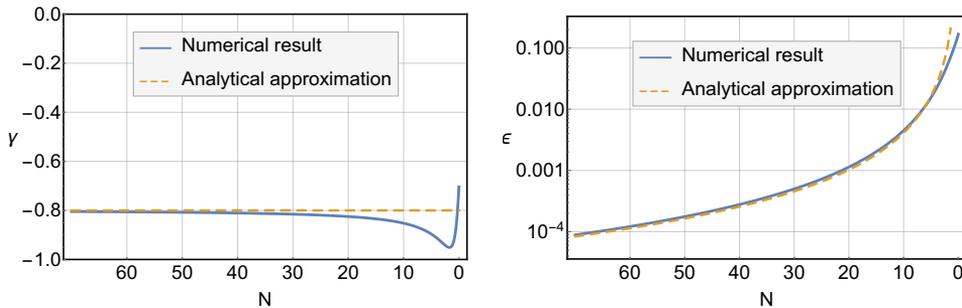


Figure 4.5: The evolution of γ and ϵ in the toy model (4.37) with $A = 0.2$, $n = 4$ and initial angle $\theta_i = \pi/8$.

numerically. Figure 4.4 shows the field trajectory on the toy model potential. We can see that the inflaton is rolling on the ridge with nearly constant θ .

Using the full numerical solution, we can check the validity of the large- φ and slow-roll approximations by looking at the evolution of background parameters. For example, within our analytical treatment, the γ parameter is given by (4.32) as

$$\gamma \simeq -\frac{nA \sin(n\theta)}{1 + nA \cos(n\theta)}. \quad (4.38)$$

It is nearly constant, since $\theta \simeq \theta_i$ during inflation. And the above choice of parameter values gives us $\gamma \simeq -0.8$, which agrees well with the numerical result as shown in Figure 4.5. Next, let us look at the slow-roll parameter ϵ . Solving (4.31) gives us its behaviour in terms of e-folding number as

$$\epsilon \simeq \frac{1 + \gamma^2}{4N^2}, \quad (4.39)$$

where (4.34) is used. As shown in Figure 4.5, this provides a good approximation until the last several e-foldings of inflation, where $\varphi \gg 1$ is not valid any more. It is interesting to notice that, during inflation the scalar field mainly rolls in the large- φ region, outside of which inflation would end very quickly. Therefore, the approximation $\varphi \gg 1$ does give a good description for the background dynamics. In the following section and in Appendix 4.B, we will come back to this toy model, and use it as an example to demonstrate other aspects of multi-field α -attractors.

4.4 Universal predictions of α -attractors

One of the most important properties of single field α -attractor inflation is the universal prediction for observations. For $\alpha \lesssim \mathcal{O}(1)$ and a broad class of potentials, as long as $V(\rho)$ is non-singular and rising near the boundary of the Poincaré disk, the resulting scalar tilt and tensor-to-scalar ratio converge to

$$n_s = 1 - \frac{2}{N} \quad \text{and} \quad r = \frac{12\alpha}{N^2}, \quad (4.40)$$

where $N \sim 50 - 60$ is the number of e-folds for modes we observe in the CMB.

One interesting question is whether the universal predictions are still valid in the multi-field regime. In multi-field scenarios the curvature perturbation is sourced by the isocurvature modes on superhorizon scales, thus their evolution is typically non-trivial and yields totally different results for n_s and r . As we show above, the angular dependence in the α -attractor potentials indeed leads to multi-field evolution. For the toy model we studied, the behaviour of perturbations can be computed using the numerical code `mTransport` [253]. We focus on one single k mode for curvature and isocurvature perturbations, and show their evolution in Figure 4.6. As expected, the curvature perturbation is enhanced during inflation, while the isocurvature modes decay. Therefore, naively one expects there would be corrections to the single field α -attractor predictions due to the multi-field effects.

In the following we will show that, surprisingly, the universal predictions are still valid in the multi-field regime. We use the δN formalism to derive the inflationary predictions for the multi-field α -attractor models studied in this paper. A full analysis of the perturbations is left for Appendix 4.B, where the evolution of the coupled system of curvature and isocurvature modes is solved via the first principle calculation .

The δN formalism [74,220–223] is an intuitive and simple approach to solve for the curvature perturbation in multi-field models. At the end of inflation, regardless of the various field trajectories, the amplitude of curvature perturbations is only determined by the perturbation of the e-folding number N , which is caused by the initial field fluctuations. Therefore, without studying details of the coupled system of curvature and isocurvature modes, as long as we know how the number of e-foldings N depends on the initial value of the two fields, the curvature perturbation can be calculated.

Let us therefore consider how the initial φ and θ determine N . In this paper, we define the e-folding number as the one counted backwards from the end of inflation, thus $dN = -Hdt$. In terms of N , the slow-roll equation (4.31) becomes

$$\frac{d\varphi}{dN} \simeq 2\sqrt{2}e^{-\sqrt{2}\varphi} \frac{V_\rho}{V}. \quad (4.41)$$

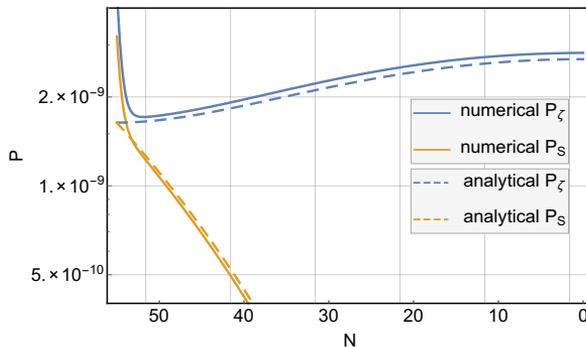


Figure 4.6: The evolution of curvature power spectrum P_ζ and isocurvature power spectrum P_S for perturbation modes which exit the horizon at $N = 55$. We use the toy model (4.37) with $A = 0.2$, $n = 4$ and initial angle $\theta_i = \pi/8$. The analytical solutions here are based on calculations in Appendix 4.B.

Since in the large φ regime $\rho \rightarrow 1$ and V_ρ/V is nearly constant for a given trajectory, the equation above yields the e-foldings from the end of inflation as

$$N = \frac{1}{B} e^{\sqrt{2}\varphi} + C(\theta) , \quad (4.42)$$

where $B \equiv 4V_\rho/V$ and $C(\theta)$ is an $\mathcal{O}(1)$ integration constant which can be fixed by setting $N = 0$ at the end of inflation. Thus, both two fields affect the duration of inflation as expected in multi-field models. By this expression, we can use the δN formalism to find curvature perturbation at the end of inflation

$$\zeta = \delta N = \frac{\partial N}{\partial \varphi} \delta \varphi + \frac{\partial N}{\partial \theta} \delta \theta = \frac{\sqrt{2} e^{\sqrt{2}\varphi}}{B} \delta \varphi + \left(C_\theta - \frac{B_\theta}{B^2} e^{\sqrt{2}\varphi} \right) \delta \theta . \quad (4.43)$$

As we see here, $\frac{\partial N}{\partial \varphi}$ and $\frac{\partial N}{\partial \theta}$ can be comparable to each other. However, one should keep in mind that θ field is non-canonical, thus to estimate the field fluctuation amplitudes at horizon-exit, one should consider the canonically normalized ones: $\delta \varphi$ and $\frac{1}{\sqrt{2}} \sinh(\sqrt{2}\varphi) \delta \theta$. Approximately in the large- φ region we have the following relation

$$\delta \varphi \simeq \frac{e^{\sqrt{2}\varphi}}{2\sqrt{2}} \delta \theta \simeq \frac{H}{2\pi} . \quad (4.44)$$

From here, we find that the field fluctuation $\delta \theta$ is exponentially suppressed, compared to the one from $\delta \varphi$. So we only need to take into account the first term in equation (4.43). In addition, equation (4.33) yields $\epsilon_\varphi = B^2 e^{-2\sqrt{2}\varphi}/4$, which further simplifies the δN formula to $\zeta \simeq \delta \varphi / \sqrt{2\epsilon_\varphi}$. In the end, the

power spectrum of curvature perturbation can be expressed as

$$P_\zeta \equiv \frac{k^3}{2\pi^2} |\zeta_k|^2 \simeq \frac{H^2}{8\pi^2 \epsilon_\varphi}. \quad (4.45)$$

It is interesting to note that this result does not depend on any parameters related to the multi-field effects (such as γ). Using (4.33) and (4.42), we also get $\epsilon_\varphi \simeq 1/(4N^2)$, which has the same behaviour with the single-field potential slow-roll parameter ϵ_V . Thus the power spectrum above is coincident with the single-field one. Then the predictions of scalar tilt and tensor-to-scalar ratio follow directly

$$n_s - 1 \simeq -\frac{2}{N} \quad \text{and} \quad r \simeq \frac{4}{N^2}. \quad (4.46)$$

These results are further confirmed by solving the full evolution of perturbations as shown in Appendix 4.B.

The δN calculation above also demonstrates the counter-intuitive properties of multi-field α -attractors. As we show in Section 4.3, the stretching effects of hyperbolic geometry not only flattens the potential in the radial direction, but also suppresses the angular velocity $\dot{\theta}$. At the level of perturbations, the similar effect occurs to the field fluctuations in the angular direction. While the canonically normalized angular field fluctuation has the same amplitude with $\delta\varphi$, the original field perturbation $\delta\theta$ is exponentially suppressed. Therefore, only the radial field fluctuation $\delta\varphi$ contributes to the final result.

Furthermore, the above results do not depend on the initial values of θ , which correspond to different field trajectories as shown in Figure 4.3. Certainly their respective e-folding number N and ϵ_φ can be different from each other. However, the N -dependence of ϵ_φ is the same for all the "rolling on the ridge" trajectories. Thus regardless of various initial values of θ , the multi-field α -attractors yield the same universal predictions for n_s and r .

Typically, another prediction in multi-field inflation is large local non-Gaussianity, which is disfavoured by the latest data [118]. Therefore it is also worthwhile to estimate the size of the bispectrum in our model. Here we expand the δN formula to the second order in field fluctuations

$$\zeta = \delta N = \frac{\partial N}{\partial \varphi} \delta\varphi + \frac{\partial N}{\partial \theta} \delta\theta + \frac{1}{2} \frac{\partial^2 N}{\partial \varphi^2} \delta\varphi^2 + \frac{1}{2} \frac{\partial^2 N}{\partial \theta^2} \delta\theta^2 + \frac{\partial^2 N}{\partial \theta \partial \varphi} \delta\theta \delta\varphi. \quad (4.47)$$

In principle, there are two contributions here: one captured by the δN expansion, and another one caused by field interactions of $\delta\varphi$ and $\delta\theta$. However, for the same reason shown in (4.44), the $\delta\theta$ -related terms in the expansion (4.47) are highly suppressed. Moreover, since there is no large coupling between field fluctuations, we expect that the second contribution to the bispectrum will

also be negligible. As a result, the local non-Gaussianity is approximately given by $\delta\varphi$ terms in (4.47)

$$f_{\text{NL}} \simeq \frac{5}{6} \frac{\partial^2 N}{\partial \varphi^2} \bigg/ \left(\frac{\partial N}{\partial \varphi} \right)^2 \simeq \frac{5}{6N}, \quad (4.48)$$

which is coincident with the single field consistency relation $f_{\text{NL}} = -\frac{5}{12}(n_s - 1)$ [81, 105]. Again, we find the multi-field α -attractor prediction returns to the single field one, which further demonstrates the scope of universality.

4.5 Universality conditions for more general α

Our investigation was stimulated by the realization that α -attractors have particularly interesting interpretation in supergravity models with $\alpha = 1/3$. A significant deviation from $\alpha = 1/3$ typically either makes the field θ tachyonic, or strongly stabilizes it at $\theta = 0$, which results in a single-field inflation driven by the field φ , see e.g. (4.6). One may wonder, however, what happens if we consider a more general class of two-field α -attractors, which may or may not have supergravity embedding, and concentrate on their general features related to the underlying hyperbolic geometry.

For general α , the canonically normalized field in the radial direction is defined by $\rho = \tanh(\varphi/\sqrt{6\alpha})$, which leads to the following kinetic term

$$\frac{1}{2} (\partial\varphi)^2 + \frac{3\alpha}{4} \sinh^2 \left(\sqrt{\frac{2}{3\alpha}} \varphi \right) (\partial\theta)^2. \quad (4.49)$$

The equations of motion (4.25) and (4.26) also change accordingly, see (4.57) and (4.58). Similarly as in Section 4.4, in the slow-roll and large- φ approximations these equations reduce to

$$\frac{\dot{\theta}}{H} \simeq -\frac{8}{3\alpha} \frac{V_\theta}{V} e^{-2\sqrt{\frac{2}{3\alpha}}\varphi}, \quad 3H\dot{\varphi} \simeq -\frac{2\sqrt{2}}{\sqrt{3\alpha}} V_\rho e^{-\sqrt{\frac{2}{3\alpha}}\varphi}. \quad (4.50)$$

As we see, the angular motion is also exponentially suppressed, compared to the radial one. So the rolling on the ridge behaviour is not unique for $\alpha = 1/3$, but quite general for any $\alpha \lesssim \mathcal{O}(1)$. Similarly to (4.35), we get the following condition to ensure its validity

$$\frac{V_\rho}{V} \gg \frac{4}{9\alpha} \left(\frac{V_\theta}{V} \right)^2 e^{-\sqrt{\frac{2}{3\alpha}}\varphi}, \quad (4.51)$$

which can be satisfied easily by many choices of potential, generalizing the bound (4.35) to other values of α . Therefore, the results follow directly just like we find in Section 4.3. For example, the ratio of proper velocities

$$\gamma = \frac{\sqrt{\frac{3\alpha}{2}} \sinh\left(\sqrt{\frac{2}{3\alpha}}\varphi\right) \dot{\theta}}{\dot{\varphi}} \quad (4.52)$$

is still nearly constant, while ϵ_φ evolves as

$$\epsilon_\varphi \simeq \frac{3\alpha}{4N^2}. \quad (4.53)$$

Repeating the same δN calculation for perturbations, we get $N \simeq 3\alpha e^{\sqrt{2/3\alpha}\varphi}/B$ and

$$\zeta = \delta N \simeq \frac{1}{\sqrt{2\epsilon_\varphi}} \delta\varphi, \quad (4.54)$$

which lead to the universal predictions (4.40) for generic α . Therefore in a broader class of α -attractors without supersymmetry, adding angular dependence to the potential will not modify the universal predictions either. Importantly, in order to validate the various assumptions we make to obtain the universal predictions, we need the potential to satisfy certain conditions. The most non-trivial condition is already given in (4.51). The additional constraints on the potential come from *assuming* the slow-roll, ‘slow-turn’ and large φ approximation. We give more detail about these approximations and collect the constraints on the potential in Appendix 4.A. Some of the conditions should also be satisfied for single field α -attractors. The smaller α becomes, the more pronounced the stretching of the hyperbolic field metric gets and it will be more likely to be within the large φ regime $\varphi \gtrsim \sqrt{\frac{3\alpha}{2}}$ and the slow-roll regime at the same time. Finally, there are some additional constraints on the potential because of the multi-field nature of our class of models. In particular, if we want to have suppressed field accelerations, we need to satisfy the slow-roll and the slow-turn conditions given in (4.59d) - (4.59f). A natural choice of the potential with $\frac{V_\theta}{\rho V} \sim \frac{V_\rho}{V} \sim \frac{V_{\theta\theta}}{\rho^2 V} \sim \frac{V_{\theta\rho}}{\rho V} \sim \frac{V_{\rho\rho}}{V} \sim 1$, evaluated at the boundary $\rho \sim 1$ amply satisfies all conditions for $\alpha \lesssim O(1)$.

4.6 Summary and Conclusions

In this paper we have studied the inflationary dynamics and predictions of a class of α -attractor models where both the radial and the angular component of the complex scalar field $Z = \rho e^{i\theta}$ are light during inflation. We concentrated on the special case $\alpha = 1/3$, where the model has a supergravity embedding with a high degree of symmetry from $\mathcal{N} = 4$ superconformal or $\mathcal{N} = 8$ supergravity. However, our results may have more general validity under the conditions specified in Appendix A. Under the weak assumptions that the potential is monotonic in the radial coordinate, and the angular gradient is not exponentially larger than the radial gradient (4.35), we find exactly the same predictions as in the theory of the single field α -attractors:

$$n_s - 1 \simeq -\frac{2}{N} \quad \text{and} \quad r \simeq \frac{12\alpha}{N^2} . \quad (4.55)$$

Universality of these predictions may make it difficult to distinguish between different versions of α -attractors by measuring n_s . However, from our perspective this universality is not a problem but an advantage of α -attractors, resembling universality of several other general predictions of inflationary cosmology, such as the flatness, homogeneity and isotropy of the universe, and the flatness, adiabaticity and gaussianity of inflationary perturbations in single field inflationary models.

The hyperbolic field metric plays a key role in finding these universal results. Let us summarize how we arrived at our new result and stress how the hyperbolic geometry dictates the analysis.

- First, the hyperbolic geometry effectively stretches and flattens the potential in the radial direction to a shape independent of the original radial potential. Independent - as long as the potential obeys the condition (4.35). The amplitude of this shape, however, varies along the angular direction.
- Next, the angular velocity $\dot{\theta}$ is exponentially suppressed, due to the hyperbolic geometry, and inflation proceeds (almost) in the radial direction. The inflaton is “*rolling on the ridge*” in the (φ, θ) plane. This is illustrated in Figures 4.3 and 4.4.
- The straight radial trajectory is an illusion, since the *physical* velocity in the axion θ direction is typically of the same order as the radial velocity. The angle between the inflationary trajectory and the radial direction is nonzero and practically constant in this regime. Moreover, although the field is following the gradient flow, the trajectory is curved in the hyperbolic geometry. Therefore, the perturbations are coupled and the multi-field effects have to be taken into account.

- Then, we use the δN formalism to compute the power spectrum of curvature perturbations (confirmed by a fully multi-field analysis in Appendix 4.B). The typical initial θ perturbations are very small and have a negligible effect on the number of e-folds. However, the initial value of θ of a given trajectory determines how much a perturbation in the radial direction affects the number of e-folds, since the trajectory is curved. At the same time the initial value of θ determines the renormalization of the slow-roll parameter ϵ . *These two effects cancel exactly*, leaving us with the same predictions as the single field α -attractors. Also the non-Gaussianity calculation recovers the single field result $f_{\text{NL}} \simeq -\frac{5}{12}(n_s - 1)$.
- Finally, in Section 4.5 and Appendix 4.A, we relax the condition $\alpha = 1/3$ and simply assume the hyperbolic geometry (4.1) with smooth potentials. We identify the conditions on the potential in order to exhibit the universal behaviour discussed in our paper, see (4.59). For $\alpha \lesssim O(1)$ these conditions are amply satisfied by a broad class of potentials $V(\rho, \theta)$, including natural ones without a hierarchy of scales: $\frac{V_\theta}{\rho V} \sim \frac{V_\rho}{V} \sim \frac{V_{\theta\theta}}{\rho^2 V} \sim \frac{V_{\theta\rho}}{\rho V} \sim \frac{V_{\rho\rho}}{V} \sim 1$, evaluated at the boundary $\rho \sim 1$.

In conclusion, the main result of our investigation is the stability of predictions of the cosmological α -attractors with respect to significant modifications of the potential in terms of the original geometric variables Z . Whereas the stability with respect to the dependence of the potential on the radial component of the field Z is well known [227], the stability with respect to the angular component of the field Z is a novel result which we did not anticipate when we began this investigation.

Our results could have important implications for constructing UV completions of inflation. We have confirmed again that multi-field models of inflation can be perfectly compatible with the current data, in particular when the additional fields are very light. This lends support to the idea that it is not always necessary to stabilize all moduli fields in order to have a successful model of inflation.

Note added: After submission, Ref. [254] appeared, which provides a supergravity embedding for the more general $\alpha < 1$ models discussed in Section 4.5.

4.A Constraints on the potential

In this Appendix we collect the conditions the potential has to obey in order to validate our approximations for any value of α . Let us first recap some relevant definitions and equation for general α . First of all, our three radial variables are given by φ and

$$R(\varphi) \equiv \sqrt{\frac{3\alpha}{2}} \sinh\left(\sqrt{\frac{2}{3\alpha}}\varphi\right), \quad \rho \equiv \tanh\left(\frac{\varphi}{\sqrt{6\alpha}}\right). \quad (4.56)$$

We introduced the radial variable $R(\varphi)$ because it appears naturally in the *physical* angular velocity $R(\varphi)\dot{\theta}$. The kinetic term can now be written in three equivalent ways

$$\frac{1}{2} \frac{(\partial\rho)^2 + \rho^2(\partial\theta)^2}{2(1-\rho^2)^2} = \frac{1}{2} (\partial\varphi)^2 + \frac{3\alpha}{4} \sinh^2\left(\sqrt{\frac{2}{3\alpha}}\varphi\right) (\partial\theta)^2 = \frac{1}{2} (\partial\varphi)^2 + \frac{1}{2} R(\varphi)^2 (\partial\theta)^2.$$

The equations of motion are generalized to

$$\ddot{\varphi} + 3H\dot{\varphi} + V_\varphi - \frac{1}{2} \sqrt{\frac{3\alpha}{2}} \sinh\left(2\sqrt{\frac{2}{3\alpha}}\varphi\right) \dot{\theta}^2 = 0, \quad (4.57)$$

$$\ddot{\theta} + 3H\dot{\theta} + \frac{V_\theta}{\frac{3\alpha}{2} \sinh^2\left(\sqrt{\frac{2}{3\alpha}}\varphi\right)} + \frac{2\dot{\theta}\dot{\varphi}}{\sqrt{\frac{3\alpha}{2}} \tanh\left(\sqrt{\frac{2}{3\alpha}}\varphi\right)} = 0. \quad (4.58)$$

Now we are ready to collect all constraints on the potential. In our derivation we *assume* that we can neglect $\ddot{\varphi}$ and that we can take the large- φ approximation. Moreover, it is important that we can neglect the centrifugal term proportional to $\dot{\theta}^2$ in Equation (4.57). We use the gradient flow to estimate the size of $\dot{\theta}$, and this leads to the first constraint (4.59a). For consistency, we have to ensure the validity of: ncy, we have to ensure the validity of:

- The slow-roll approximation, which gives rise to the next four constraints (4.59b) - (4.59e). This approximation ensures that the field velocities are small and that we can neglect their acceleration pointing along the corresponding field direction as well.
- The slow-turn approximation, which allows us to neglect the field accelerations pointing along the *other* field direction. If we can assume gradient flow for θ this leads to the condition (4.59f).
- The large- φ approximation, which requires us not to go to the extreme limit of a very shallow radial potential. We want to inflate sufficiently far from the origin in order to obtain enough e-folds of inflation, such

that we can use the large- φ approximation. In our analysis we work for simplicity with potentials $\left(\frac{V_\rho}{V}\right)^2 \gtrsim \frac{\alpha}{4}$ so this is automatically satisfied.

$$\frac{V_\rho}{V} \gg \frac{4}{9\alpha} \left(\frac{V_\theta}{V}\right)^2 e^{-\sqrt{2/3}\alpha\varphi}, \quad (4.59a)$$

$$\epsilon_\varphi \equiv \frac{1}{2} \left(\frac{V_\varphi}{V}\right)^2 = \frac{4}{3\alpha} \left(\frac{V_\rho}{V}\right)^2 e^{-2\sqrt{2/3}\alpha\varphi} \ll 1, \quad (4.59b)$$

$$\epsilon_\theta \equiv \frac{1}{2} \left(\frac{V_\theta}{RV}\right)^2 = \frac{4}{3\alpha} \left(\frac{V_\theta}{V}\right)^2 e^{-2\sqrt{2/3}\alpha\varphi} \ll 1, \quad (4.59c)$$

$$\eta_\varphi \equiv \frac{1}{3} \frac{V_{\varphi\varphi}}{V} = \frac{8}{9\alpha} \frac{V_{\rho\rho}}{V} e^{-2\sqrt{2/3}\alpha\varphi} \ll 1 \quad (4.59d)$$

$$\eta_\theta \equiv \frac{1}{3} \frac{V_{\theta\theta}}{R^2V} = \frac{8}{9\alpha} \frac{V_{\theta\theta}}{V} e^{-2\sqrt{2/3}\alpha\varphi} \ll 1, \quad (4.59e)$$

$$\omega_\phi \equiv \frac{V_{\theta\varphi}}{3RV} \frac{V_\theta}{RV_\varphi} = \frac{V_{\theta\rho}}{V} \frac{V_\theta}{V_\rho} \frac{8}{9\alpha} e^{-2\sqrt{2/3}\alpha\varphi} \ll 1. \quad (4.59f)$$

Please note that all constraints have to be evaluated at $\rho \sim 1$, i.e. at $\varphi \gg 6\alpha$. Our conditions are satisfied for simplest potentials, because in the large- φ regime all slow-roll and slow-turn parameters are exponentially suppressed. For instance, natural potentials which satisfy $\frac{V_\theta}{\rho V} \sim \frac{V_\rho}{V} \sim \frac{V_{\theta\theta}}{\rho^2 V} \sim \frac{V_{\theta\rho}}{\rho V} \sim \frac{V_{\rho\rho}}{V} \sim 1$ at the boundary $\rho \sim 1$, amply obey the conditions.

4.B Full analysis of perturbations

In this Appendix, we give a detailed study of turning trajectories in multi-field α -attractors and work out the full evolution of curvature and isocurvature perturbations.

4.B.1 Covariant formalism and large- φ approximations

For a general multi-field system spanned by coordinate ϕ^a with field metric G_{ab} , the equations of motion in the FRW background can be simply written as

$$D_t \dot{\phi}^a + 3H \dot{\phi}^a + V^a = 0, \quad 3H^2 = \frac{1}{2} \dot{\Phi}^2 + V \quad (4.60)$$

where D_t is the covariant derivative respect to cosmic time and $\dot{\Phi}^2 \equiv G_{ab} \dot{\phi}^a \dot{\phi}^b$. To describe the multi-field effects, it is convenient to define the tangent and orthogonal unit vectors along the trajectory as

$$T^a \equiv \frac{\dot{\phi}^a}{\dot{\Phi}}, \quad N_a \equiv \sqrt{\det G} \epsilon_{ab} T^b, \quad (4.61)$$

where ϵ_{ab} is the Levi-Civita symbol with $\epsilon_{12} = 1$. The rate of turning for the background trajectory is defined as

$$\Omega \equiv -N_a D_t T^a = \frac{V_N}{\dot{\Phi}}, \quad (4.62)$$

where for the second equality we have used the background equations of motion and $V_N = N^a \nabla_a V$ is the gradient of the potential along the normal direction of the trajectory. This quantity, which vanishes in single field models, is particularly important for the multi-field behaviour and evolution of perturbations. A dimensionless turning parameter is introduced as

$$\lambda \equiv -\frac{2\Omega}{H}. \quad (4.63)$$

Another important parameter is the field mass along the orthogonal direction defined as

$$V_{NN} \equiv N^a N^b \nabla_a \nabla_b V. \quad (4.64)$$

Now let us come back to our model with coordinates $\phi^a = (\varphi, \theta)$ and hyperbolic field metric

$$G_{ab} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \sinh^2(\sqrt{2}\varphi) \end{pmatrix}. \quad (4.65)$$

The Ricci scalar of this manifold is a negative constant $\mathbb{R} = -2$. By the definitions above, after some algebra, λ and V_{NN} here can be written into the following form

$$\lambda = \frac{1}{\epsilon H^3} \frac{\sinh(\sqrt{2}\varphi)}{\sqrt{2}} \left[\ddot{\varphi} \dot{\theta} - \ddot{\theta} \dot{\varphi} - \frac{2\dot{\theta} \dot{\varphi}^2}{\frac{1}{\sqrt{2}} \tanh(\sqrt{2}\varphi)} - \frac{1}{2\sqrt{2}} \sinh(2\sqrt{2}\varphi) \dot{\theta}^3 \right], \quad (4.66)$$

$$V_{NN} = \frac{1}{\dot{\Phi}^2} \left(\frac{V_{\theta\theta} \dot{\varphi}^2 + \frac{\sqrt{2}}{4} \sinh(2\sqrt{2}\varphi) V_{\varphi} \dot{\varphi}^2}{\frac{1}{2} \sinh^2(\sqrt{2}\varphi)} + 2\dot{\theta} \dot{\varphi} \left[\frac{\sqrt{2} V_{\theta}}{\tanh(\sqrt{2}\varphi)} - V_{\theta\varphi} \right] + \frac{1}{2} \sinh^2(\sqrt{2}\varphi) V_{\varphi\varphi} \dot{\theta}^2 \right). \quad (4.67)$$

These expressions look very complicated, but in the large- φ regime they can be efficiently simplified. First of all, since γ in (4.32) is nearly constant, we can use this parameter to replace $\dot{\theta}$ by $\dot{\varphi}$ in these expressions, for example $\dot{\Phi}^2 = (1 + \gamma^2) \dot{\varphi}^2$. Then we can use the relations of background quantities presented in Section 4.4 to further simplify the result. Finally the turning parameter λ can be expressed as

$$\lambda = \frac{-1}{\epsilon H^3} (1 + \gamma^2) \frac{\sqrt{2} \gamma \dot{\varphi}^3}{\tanh(\sqrt{2}\varphi)} \simeq \frac{2\sqrt{2} \gamma}{(1 + \gamma^2)^{1/2}} \cdot \sqrt{2\epsilon}, \quad (4.68)$$

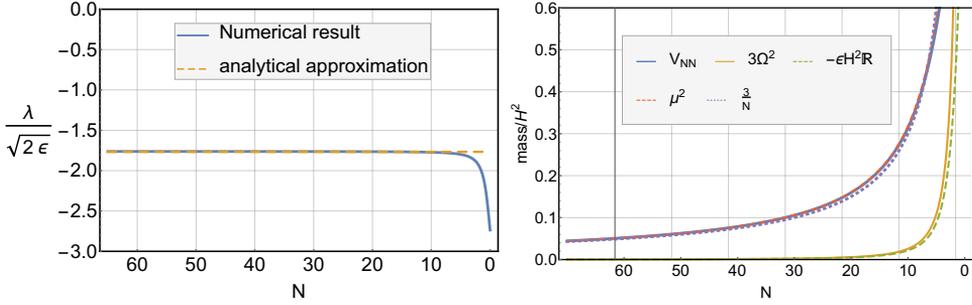


Figure 4.7: The evolution of the dimensionless turning parameter $\frac{\lambda}{\sqrt{2}\epsilon}$ and entropy masses. Here we use the toy model potential (4.37) with $n = 4$, $A = 0.2$ and $\theta_i = 8/\pi$.

where the large φ approximation is used in the last step. Therefore, at $\varphi \gg 1$, $\lambda/\sqrt{2}\epsilon$ is nearly constant. Similarly, we can work out the approximated expression for V_{NN} . Here we use the toy model potential for demonstration, which yields

$$V_{NN} \approx V_0 B e^{-\sqrt{2}\varphi}. \quad (4.69)$$

Therefore, V_{NN} is nearly zero at the beginning of inflation, but then grows up as φ rolls to the center. These analytical approximations are checked by using numerical solution of the toy model. In Figure 4.7 we show the numerical results versus the analytical ones for $n = 4$, $A = 0.2$ and $\theta_i = 8/\pi$. Indeed we see that $\frac{\lambda}{\sqrt{2}\epsilon}$ remains constant until the very end of inflation, where the large- φ approximation breaks down.

4.B.2 Primordial Perturbations

With the analytical approximations developed above, now we can move to study the behaviour of perturbations. In particular, we would like to derive the analytical expression for the power spectrum of curvature perturbations. At the linear level, the curvature perturbation ζ and the isocurvature modes σ are defined as

$$\delta\phi^a = \sqrt{2\epsilon}\zeta T^a + \sigma N^a \quad (4.70)$$

And their full equations of motion in terms of e-foldings are given by

$$\frac{d}{dN} \left(\frac{d\zeta}{dN} - \frac{\lambda}{\sqrt{2}\epsilon} \sigma \right) + (3 - \epsilon + \eta) \left(\frac{d\zeta}{dN} - \frac{\lambda}{\sqrt{2}\epsilon} \sigma \right) + \frac{k^2}{a^2 H^2} \zeta = 0, \quad (4.71)$$

$$\frac{d^2\sigma}{dN^2} + (3 - \epsilon) \frac{d\sigma}{dN} + \sqrt{2\epsilon}\lambda \left(\frac{d\zeta}{dN} - \frac{\lambda}{\sqrt{2}\epsilon} \sigma \right) + \frac{k^2}{a^2 H^2} \sigma + \frac{\mu^2}{H^2} \sigma = 0, \quad (4.72)$$

where μ^2 is the entropy mass of the isocurvature perturbations given by

$$\mu^2 \equiv V_{NN} + \epsilon H^2 \mathbb{R} + 3\Omega^2. \quad (4.73)$$

Thus besides V_{NN} , the turning effects and the curvature of the field manifold also contribute to the entropy mass. But in multi-field α -attractors here, as shown in Figure 4.7, μ^2 is mainly controlled by the V_{NN} term. Then by (4.69) and the solution of φ in (4.42), we get

$$\frac{\mu^2}{H^2} \approx \frac{V_{NN}}{H^2} \approx 3Be^{-\sqrt{2}\varphi} \approx \frac{3}{N}, \quad (4.74)$$

which provides a good analytical approximation as shown in Figure 4.7.

The exact solutions of the full equations (4.71) and (4.72) can be obtained only through numerical method, as we have shown in Figure 4.6. But notice that the leading effect here comes from the coupled evolution of curvature and isocurvature modes after horizon-exit. Thus for the analytical approximations, we can focus on super-horizon scales. There the isocurvature equation of motion reduces to

$$3 \frac{d\sigma}{dN} + \frac{\mu^2}{H^2} \sigma \simeq 0. \quad (4.75)$$

If we focus on the mode that exits horizon at N_* with amplitude σ_* , then we get the following solution for its evolution

$$\sigma(N) \approx \sigma_* \frac{N}{N_*}. \quad (4.76)$$

Remember that e-folding number is counted backwards from the end of inflation. Thus this solution shows the decay of the isocurvature perturbation outside of the horizon. The evolution of the normalized isocurvature power spectrum $P_{\mathcal{S}}$ is shown in Figure 4.6, where $\mathcal{S} = \sigma/\sqrt{2\epsilon}$. As we see, the analytical approximation above successfully captures the super-horizon decay, compared with the numerical result.

Next, we look at the curvature perturbation, which after horizon-exit is sourced by the the isocurvature modes in the following way

$$\frac{d\zeta}{dN} = \frac{\lambda}{\sqrt{2\epsilon}} \sigma. \quad (4.77)$$

Also for the mode exits horizon at N_* with amplitude ζ_* , we get the solution

$$\zeta(N) = \zeta_* + \int_{N_*}^N dN' \frac{\lambda}{\sqrt{2\epsilon}} \sigma(N'). \quad (4.78)$$

As we noticed in (4.68), $\lambda/\sqrt{2\epsilon}$ is nearly constant, thus it can be seen as unchanged after horizon-exit $\lambda/\sqrt{2\epsilon} = \lambda_*/\sqrt{2\epsilon_*}$. Meanwhile, notice that since σ is almost massless in the large- φ regime, one has the relation $\sqrt{2\epsilon_*}\zeta_* \simeq \sigma_* \simeq H/(2\pi)$. Then the evolution of ζ is given by

$$\zeta(N) = \zeta_* + \frac{\lambda_*}{2} \frac{N^2 - N_*^2}{N_*} \zeta_*. \quad (4.79)$$

These two contributions are uncorrelated with each other, since they come from the different parts of the quantized fluctuations. Thus finally we can write down the power spectrum at the end of inflation ($N = 0$)

$$P_\zeta = P_\zeta^{(1)} + P_\zeta^{(2)} = \frac{H^2}{4\pi^2} \frac{1}{2\epsilon_*} \left(1 + \frac{\lambda_*^2 N_*^2}{4} \right) = \frac{H^2}{8\pi^2 \epsilon_*} (1 + \gamma^2) = \frac{H^2}{8\pi^2 \epsilon_{\varphi_*}}, \quad (4.80)$$

where we used the relation (4.34), the expression of λ (4.68), and $\epsilon_\varphi \simeq 1/(4N^2)$. Therefore, we recover the same result as we got in δN calculation.

It is interesting to note that, although the turning effects play an important role in the intermediate calculation, they vanish in the final answer. There are two effects on the curvature perturbation in multi-field α -attractors: first, the growth on super-horizon scales gives an enhancement factor $(1 + \gamma^2)$; secondly, due to the motion in the θ direction, the slow-roll parameter ϵ is also enhanced by the same amount. Thus as a consequence, these two changes cancel each other, and the final power spectrum of curvature perturbation here becomes the same as the single field result.

Heavy fields and a reduced tensor-to-scalar ratio

We illustrate the impact of heavy fields on the inflationary observables by a simple two-field embedding of a few large-field models of inflation. In our setup the inflaton corresponds to the phase of a complex field with mildly broken $U(1)$ symmetry. This type of embedding affects the background evolution and modifies the effective sound speed of the curvature perturbation. The overall effect is that the tensor-to-scalar ratio is reduced, which improves the viability of the inflationary models under consideration.

The results in this chapter are based on [140]:
On the viability of $m^2\phi^2$ and natural inflation, A. Achúcarro, V. Atal and Y. Welling, JCAP **1507** (2015) 07, 008, (arXiv:1503.07486 [astro-ph.CO]).

5.1 Introduction

Precise observations of the cosmic microwave background and large scale structure allow today for a very accurate determination of the cosmological parameters [30]. This requires the theoretical predictions to be precise and robust, so that there are no uncertainties when interpreting the data, or, more realistically speaking, that any uncertainty is well understood. In the particular case of inflation [51, 57], the fact that this period might be driven by a single light scalar field, effectively uncoupled to any additional degree of freedom (at least during the time when observable perturbations are generated) is a very appealing scenario in terms of predictability¹. Indeed, any given potential has unique and precise predictions, with the only ambiguity coming from uncertainties on which e-foldings correspond to the observable scales. The fact that many models of inflation have been ruled out by measurements of the spectral tilt and bounds on the amount of tensor perturbation, reaffirms this claim. In particular some of the simplest large field models like $m^2\phi^2$ [62] and natural inflation [255] are in tension with the data [115]. In this chapter we show that a particular class of two-field embeddings, where the additional field is super heavy, can bring these models back into consistency with the data, by changing

¹At least if we ignore the issue of eternal inflation and the multiverse paradigm.

the value of the slow-roll parameters as well as by generating a reduced speed of sound $c_s < 1$ for the fluctuations. We emphasize that this deformation is coming entirely from super massive degrees of freedom (masses much heavier than the Hubble parameter), and is an example of how important heavy fields may be in determining the low energy effective description.

In general, a heavy field may influence the low energy dynamics by either affecting the background, the perturbations or a combination of both. We would like to stress that this does not require any high energy excitations or particle production, and our discussion in this chapter focuses on the regime in which they do not occur. First of all, changes in the background may come when evaluating the action at the vacuum expectation value (v.e.v.) of the heavy field. Provided the kinetic energy is dominated by the light field, it is possible to write a Lagrangian in which we recover single-field inflation described by some effective potential (see e.g. [256, 257]). Changing the background will result in different values for the slow-roll parameters ϵ and η , and thus we call this the *background* model. Secondly, for certain (derivative) couplings between the light and heavy field, perturbations of the heavy field contribute to the low frequency mode and therefore to the low energy - single field - effective field theory (EFT) for the perturbations². The effect of this coupling has been widely studied, both in the cases when it is small [258, 259], and the cases in which it is large [200, 201, 206, 213–215, 235, 260–265] (see [205] for a comparative study of some of these works). In the latter case, integrating out the heavy field results in a reduced speed of sound for the adiabatic fluctuations. This is a purely quantum effect that arises when considering the full two-field evolution for the perturbations.

As we will explicitly show through examples, a proper description of the system demands taking both effects into account. In order to do so, we embed inflation in a simple two-field realization such that inflation takes place on a turning trajectory whose radius of curvature is changing very slowly. In our embedding, the inflaton is the phase of a complex field where the $U(1)$ symmetry is mildly broken and the v.e.v. of the - massive - radial field is approximately the radius of curvature of the inflationary trajectory (see [266, 267] and [268] for realizations in field theory and supergravity respectively).

As already anticipated, we will show that within this framework quadratic inflation can be consistent with the data, and that subplanckian values of the effective (instantaneous) decay constant are no longer disfavored in natural inflation. This will require that the v.e.v. of the radial field takes subplanckian

²The isocurvature modes are heavy and decay fast, they do not source the curvature perturbations after horizon crossing. This situation is different from so-called multifield inflation, where there are multiple light fields [199, 202], and quasi-single field inflation, where the mass of the heavy field is order of the Hubble parameter [101].

values. Different deformations of standard natural and chaotic inflation were discussed in e.g. [269–272], and [272–278] respectively. Unlike previous studies that included additional heavy fields to improve agreement with the data due to a flattening of the potential, in all cases described here the effective potential *steepens*. However the speed of sound effects dominate and move the predictions *downward* in the (n_s, r) plane, towards the best fit region. In addition, in some cases the speed of sound slowly decreases along the trajectory, causing a shift towards higher values of n_s .

The outline of the Chapter is as follows. In section 5.2 we introduce a simple two-field embedding of inflation with an additional heavy field. We explain how both the background and the curvature perturbations are affected by the presence of the heavy field. Furthermore we provide analytical expressions for the observables. In sections 5.3, 5.4 and 5.5 we study the predictions of the observables of quadratic, linear and natural inflation embedded in this two-field scenario. We show explicitly how the predictions in the (n_s, r) plane move towards the 1σ allowed region of Planck. Finally in section 5.6 we discuss our findings and conclude.

5.2 General setup

5.2.1 Two-field embedding

We embed inflation in a simple two field realization, given by the following Lagrangian:

$$\mathcal{L} = \frac{1}{2} \partial_\nu \rho \partial^\nu \rho + \frac{1}{2} \rho^2 \partial_\nu \theta \partial^\nu \theta - \frac{m_\rho^2}{2} (\rho - \rho_0)^2 - V(\theta). \quad (5.1)$$

Here we assume the inflaton to be the phase of a complex field, and the $U(1)$ symmetry has been mildly broken by a potential $V(\theta)$. This model was already studied in [215] (with a different kinetic term), and in [213], for the case of linear inflation (that we discuss below). This is not the most general Lagrangian consistent with the symmetries invoked. In fact we may also have different choices for the potential or the field space metric, but we choose this form as the simplest starting case. We also notice that since the model becomes singular at $\rho = 0$, additional degrees of freedom should appear at some higher energy scale. Additionally, these models will have monodromy; the potential is not completely invariant after the phase θ has made a 2π cycle [279–282].

The standard logic (that we will show is inaccurate) is that, if the radial field is sufficiently heavy, the field ρ will rapidly reach its minimum at ρ_0 , so

that one may truncate the model and consider the single field Lagrangian

$$\mathcal{L} = \frac{1}{2}\rho_0^2\partial_\nu\theta\partial^\nu\theta - V(\theta), \quad (5.2)$$

which by a field redefinition becomes:

$$\mathcal{L} = \frac{1}{2}\partial_\nu\phi\partial^\nu\phi - V(\phi/\rho_0). \quad (5.3)$$

We will show that truncating the model in this manner yields inaccurate predictions. The reason is twofold. First, because of the kinetic coupling the radial field will have a minimum at $\bar{\rho} \neq \rho_0$. Plugging this solution back in the Lagrangian will result in an EFT in which we recover single-field inflation described by an effective potential [256, 257]. In general, this will result in different predictions for both ϵ and η (the slow-roll parameters) at the observable scales. This single field description is possible provided the kinetic energy is dominated by the angular field, or more specifically that $\dot{\rho}^2 + \rho^2\dot{\theta}^2 = ((d\rho/d\theta)^2 + \rho^2)\dot{\theta}^2 \sim \rho^2\dot{\theta}^2$. From this condition we will demand that $d\rho/d\theta \ll \rho$. Secondly, light and heavy field perturbations will be coupled through the angular velocity, $\dot{\theta}/H$, which, if large, will give rise to a low energy EFT with a reduced speed of sound c_s for the adiabatic fluctuations [212–214, 264].

The prediction for the system can then be computed with the usual relations³ $r = 16\epsilon c_s$ and $n_s = 1 - 2\epsilon - \eta - s$ with $\epsilon \equiv -\dot{H}/H^2$, $\eta \equiv \dot{\epsilon}/H\epsilon$ and $s \equiv \dot{c}_s/Hc_s$. From these expressions it is clear that a reduced speed of sound, $c_s < 1$, will contribute to moving the predictions of the model towards smaller values of r (taking into account that in general ϵ will also change while we go from $c_s = 1$ to $c_s \ll 1$). Additionally, n_s will also change making theories flow in the (n_s, r) plane⁴. In the following, we show how to exactly calculate these quantities.

³We actually use the more precise predictions for r as in [283], where the difference in freeze-out time between the scalar and tensor perturbation is considered [104]. This effect becomes relevant when $c_s \ll 1$. We also compute the power spectrum, from which we derive n_s , at second order in slow-roll, as in [283].

⁴In [284], the effect of c_s and s in the (n_s, r) plane for a phenomenological ansatz $c_s(N)$ was discussed.

5.2.2 Analytical predictions

The possibility to make analytical predictions for n_s and r depends on the ability to calculate the radius $\bar{\rho}(\theta)$ at which the radial field stabilizes, as a function of the parameters of the two-field embedding. While it is not generically possible to solve the full two-field model, there is a regime in which such analytical predictions are possible. This is the regime in which the time derivatives of ρ can be neglected in the equations of motion (e.o.m.). In order to show this explicitly, consider the e.o.m. for the system

$$\ddot{\rho} + 3H\dot{\rho} - \rho\dot{\theta}^2 + V_\rho = 0 \quad (5.4)$$

$$\rho^2\ddot{\theta} + 2\rho\dot{\rho}\dot{\theta} + 3H\rho^2\dot{\theta} + V_\theta = 0 \quad (5.5)$$

together with the Friedmann equation (from here on we set the reduced Planck mass $m_{pl} \equiv (8\pi)^{-1/2}M_{pl} = 1$)

$$3H^2 = \frac{1}{2} \left(\dot{\rho}^2 + \rho^2\dot{\theta}^2 \right) + V. \quad (5.6)$$

where $V_\theta = \partial V/\partial\theta$ and $V_\rho = \partial V/\partial\rho$.

First, we assume that we can neglect the derivatives of ρ in the previous equations (which is a good approximation in all of the cases studied here). We assume then $\ddot{\rho}, 3H\dot{\rho} \ll \rho\dot{\theta}^2, V_\rho$ and $2\rho\dot{\rho}\dot{\theta} \ll 3H\rho^2\dot{\theta}, V_\theta, \rho^2\ddot{\theta}$. Let us note that the previous inequalities demand that, at the same time, $\frac{1}{\rho} \frac{d\rho}{d\theta} \ll \frac{\dot{\theta}}{3H}$ and $\frac{1}{\bar{\rho}} \frac{d\bar{\rho}}{d\theta} \ll \frac{3H}{2\dot{\theta}}$. This directly implies that

$$\frac{1}{\bar{\rho}} \frac{d\bar{\rho}}{d\theta} \ll 1 \quad (5.7)$$

which is the condition for writing a single field model for the background (i.e. that the kinetic energy is dominated by the angular velocity). This condition does not mean that the field ρ has to be exactly constant, but rather that its time evolution is slow.

Furthermore, as θ plays - mainly - the role of the inflaton, we also drop $\ddot{\theta}$. This demands $\dot{\theta} < 3H\dot{\theta}$. The simplified system then reads:

$$\rho\dot{\theta}^2 = V_\rho \quad (5.8)$$

$$3H\rho^2\dot{\theta} + V_\theta = 0 \quad (5.9)$$

and

$$3H^2 = V. \quad (5.10)$$

Importantly, let us note that there is no bound on $\dot{\theta}/H$ (as long as $\dot{\theta} < m_\rho < M_{pl}$). This quantity plays an important role in determining both the

coupling of the perturbations between the light and heavy field, and - as we will see below - the slow-roll parameters.

In principle with these equations we obtain $\bar{\rho}(\theta)$, plug the solution in the potential and find a canonical variable so that we have a single-field effective potential. However, in situations in which the solution $\bar{\rho}(\theta)$ is a complicated function of θ , it may be too difficult to follow this procedure, the main reason being that we need to find a canonical variable ϕ such that $\bar{\rho}\dot{\theta} = \dot{\phi}$. Nonetheless, the system can still be solved semi-analytically in a single field approach. If the kinetic energy is dominated by θ , then ϵ is given by

$$\epsilon = \frac{1}{2\bar{\rho}^2} \left(\frac{V_\theta}{V} \right)^2. \quad (5.11)$$

With this we can calculate $\eta = \dot{\epsilon}/\epsilon H$ giving

$$\eta = \frac{2}{\bar{\rho}^2} \left(\frac{V_{\theta\theta}}{V} \right) + 4\epsilon - 2\frac{\dot{\bar{\rho}}}{\bar{\rho}H}. \quad (5.12)$$

Importantly, the last term cannot be neglected. Indeed

$$\delta \equiv \frac{\dot{\bar{\rho}}}{\bar{\rho}H} \quad (5.13)$$

$$= \frac{1}{\bar{\rho}} \frac{d\bar{\rho}}{d\theta} \frac{\dot{\theta}}{H} = \frac{1}{\bar{\rho}} \frac{d\bar{\rho}}{d\theta} \frac{\sqrt{2\epsilon}}{\bar{\rho}}. \quad (5.14)$$

While the reduced e.o.m. demands $\delta \ll 1$, δ may be $\mathcal{O}(\epsilon, \eta)$. Then, we can calculate all of the relevant quantities for the background with the following relations⁵:

$$\epsilon = \frac{1}{2\bar{\rho}^2} \left(\frac{V_\theta}{V} \right)^2, \quad \eta = \frac{2}{\bar{\rho}^2} \left(\frac{V_{\theta\theta}}{V} \right) + 4\epsilon - 2\delta, \quad \delta = \frac{\dot{\bar{\rho}}}{\bar{\rho}H}, \quad N = \int \bar{\rho}^2 \frac{V}{V_\theta} d\theta, \quad (5.15)$$

where $\bar{\rho}$ is the solution to $\rho^3 V_\rho = V_\theta^2/3V$ and N is the number of e-folds before the end of inflation. From here it is clear that the time dependence of $\bar{\rho}$ has to be explicitly taken into account in order to make accurate predictions. This mean that while we can neglect the derivatives of ρ in the e.o.m. its derivatives do play an important role in determining the observables of the model.

As for the perturbations, in a regime in which the angular acceleration $\ddot{\theta}$ is small in comparison with the effective mass of the heavy field [211] (given here

⁵These expressions were derived assuming a separable potential. For non separable potentials they remain approximately valid provided $V_{\theta\rho}(d\bar{\rho}/d\theta) \ll V_{\theta\theta}$

by $m_{\text{eff}}^2 = m_\rho^2 - \dot{\theta}^2$ and demanding $m_{\text{eff}} \gg H$), the low energy EFT develops a speed of sound of the fluctuations which is given by⁶

$$c_s^{-2} = 1 + 4 \frac{\dot{\theta}^2}{m_{\text{eff}}^2} \quad \text{where} \quad m_{\text{eff}}^2 = m_\rho^2 - \dot{\theta}^2. \quad (5.16)$$

We refer to [213] for a more detailed discussion. Moreover, it is easy to show that s can be written solely in terms of ϵ , η , δ and c_s :

$$s = \left(\epsilon - \frac{\eta}{2} + \delta \right) (1 - c_s^2) \left(\frac{3}{4} + \frac{1}{4c_s^2} \right). \quad (5.17)$$

With all these elements it is possible to compute all of the observables of the model, i.e. r and n_s , without having to solve any dynamical equation, as in the standard slow-roll computation. As can be seen from eq. (5.16), in order to have a substantial reduction in the speed of sound we will need large angular velocities, of the order of the effective heavy mass. This is consistent with slow-roll whenever the radius of curvature is small enough, and that is the reason why we will demand the condition $\rho_0 < 1$ to be satisfied.

Before closing this section, two comments are in order. First, it is important to ensure that the theory stays weakly coupled up to the scale where new physics cannot be further integrated out. In models with a reduced speed of sound, this places a theoretical lower bound on the speed of sound [102, 260]. Every case presented here is consistent with this bound, provided a scale dependent speed of sound - like the one we have - is taken into account.

Secondly, a reduction in the speed of sound unavoidably implies a cubic interaction for the adiabatic perturbation [99], producing potentially observable non-gaussianity. In particular, for the case of a nearly constant speed of sound we have [98] [213]:

$$f_{NL}^{(eq)} = \frac{125}{108} \frac{\epsilon}{c_s^2} + \frac{5}{81} \frac{c_s^2}{2} \left(1 - \frac{1}{c_s^2} \right)^2 + \frac{35}{108} \left(1 - \frac{1}{c_s^2} \right). \quad (5.18)$$

This means that in order to have a measurable non gaussianity $|f_{NL}^{eq}| > 5$, we need $c_s < 0.2$ (which is still consistent with weak coupling [102, 260]). While it is interesting to search for such values of c_s , we will notice that much milder reductions in the speed of sound can already leave big imprints in the power spectrum, and we are thus going to focus mainly on mild reductions of c_s .

⁶The speed of sound presented here is the $k \rightarrow 0$ limit of the speed of sound obtained by integrating out the heavy mode. The k -dependence is not extremely important to compute the observables of the theory - at least for moderate reductions in c_s - but it becomes important in order to assess the overall consistency and predictivity of the theory [102, 260]. Whenever needed, we use the full k -dependent c_s .

5.3 Quadratic inflation

Our first example to show how a heavy field may influence the low energy dynamics is a two-field embedding of the quadratic inflation model. The Lagrangian for the single field model [62] is given by:

$$\mathcal{L} = \frac{1}{2} \partial_\nu \phi \partial^\nu \phi - \frac{1}{2} m_\phi^2 \phi^2. \quad (5.19)$$

We embed this model in the two field scenario (5.1), and consider the following Lagrangian:

$$\mathcal{L} = \frac{1}{2} \partial_\nu \rho \partial^\nu \rho + \frac{1}{2} \rho^2 \partial_\nu \theta \partial^\nu \theta - \frac{m_\rho^2}{2} (\rho - \rho_0)^2 - \Lambda^4 \theta^2. \quad (5.20)$$

Assuming $\rho = \rho_0$, and defining $\phi = \rho_0 \theta$, we recover the single field Lagrangian with the mass m_ϕ given by $m_\phi^2 = 2\Lambda^4/\rho_0^2$. Thus, at the level of this truncation, both Lagrangians (5.19) and (5.20) are equivalent. Going beyond this simplification demands solving the full e.o.m. Fortunately we can rely on the reduced e.o.m. to find approximate solutions. Solving equations (5.8), (5.9) and (5.10), the minimum in the radial direction, $\bar{\rho}$, is given by the root of the following equation:

$$\bar{\rho}^3 (\bar{\rho} - \rho_0) - \frac{4}{3} \frac{\Lambda^4}{m_\rho^2} = 0, \quad (5.21)$$

while the angular velocity is given by

$$\dot{\theta} = -\frac{2}{\sqrt{3}} \frac{\Lambda^2}{\bar{\rho}^2}. \quad (5.22)$$

Here we have used $V = V(\theta)$, which, as we will remark below, is a very good approximation. With these solutions at hand we can then predict how the observables move in the (n_s, r) plane. In doing so we will split the effects on the background and perturbations.

Background model: All the relevant quantities for calculating the background can be found in equation (5.15). First of all, because $\bar{\rho} \neq \rho_0$, the potential V will have a contribution of the form $V_0 = \frac{m_\rho^2}{2} (\bar{\rho} - \rho_0)^2$. It is easy to show that this contribution is negligible in comparison with $V(\theta)$ in the computation of the slow-roll parameters at $N = 50 - 60$ ⁷. Thus, we can use $V \sim V(\theta)$. Under this simplification, and because $\bar{\rho} \sim cte$, the background model yields the same predictions as in the standard quadratic inflation, i.e. $\epsilon = 1/2N$ and $\eta = 1/N$.

⁷This approximation breaks down towards the end of inflation. The numerical results confirm that it does not affect the predictions at the observables scales.

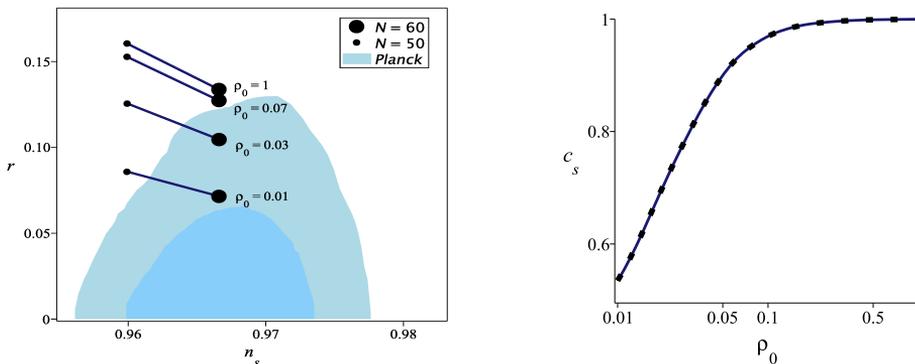


Figure 5.1: *Left:* The (n_s, r) plane for $m^2\phi^2$ inflation when embedded in the model given by (5.20), with the mass of the heavy field given by $m_{\text{eff}}^2 = 100H^2$. The predictions are calculated using the EFT, which is an excellent description for the full two-field system. The blue regions are the 1- σ and 2- σ allowed regions from Planck [115]. *Right:* Speed of sound of the adiabatic fluctuation, given by equation (5.16) for $N = 50 - 60$. The dotted line is computed using the numerical solution of the full two-field system (5.20), and the solid line is computed with the semi-analytical approximation of equation (5.15).

EFT for the perturbations: While the background does not change as we change ρ_0 , we find that perturbations develop a constant speed of sound which is noticeably different from 1 for values of $\rho_0 < 0.1$, as can be seen in the right panel of figure 5.1.

Putting all these elements together we compute the prediction for (n_s, r) . Since ϵ and η are unchanged, and $s \sim 0$, only the tensor to scalar ratio is going to be modified, and its modification will only be due to the change in c_s . We test these predictions with a numerical solution of the two-field system (partly done using the code from [253]), choosing ρ_0 ranging from 0.01 to 1 and m_ρ such that $(m_{\text{eff}}/H)^2 = 100$ in the observable scales. We fix Λ such that we have the right amplitude for the perturbations. Let us note that we have fixed the effective mass of the heavy field (which is always smaller than the bare mass m_ρ) such that it is much greater than the Hubble parameter. Our results are summarized in figure 5.1, displayed together with the experimental bounds from Planck⁸ [115]. First of all, there is very good agreement between the predictions of the analytical single field EFT and the full two-field system, and more importantly, there are sizeable effects in terms of where the predictions lie in the (n_s, r) plane.

⁸From the PLA-PR2-2015 official chains including TT,TE and low- l polarization data at <http://pla.esac.int/pla/>.

In particular, we see that when decreasing ρ_0 the quadratic potential becomes more consistent with Planck confidence regions. This is not a surprising result, since we know that reduced speeds of sound lead to smaller values for the tensor to scalar ratio, but the fact that this can be achieved with the simple quadratic potential “UV completed” with an additional very massive field is worth noting. Let us note that larger reductions in the speed of sound can easily be attainable provided we consider smaller values for ρ_0 . This will further increase the consistency with the data as well as generating potentially observable non gaussianities (e.g for, $\rho_0 = 10^{-3}$ we find $|f_{NL}^{eq}| \sim 5$).

5.4 Linear inflation

We repeat the analysis, this time for the linear inflaton potential [279]. We consider the following Lagrangian:

$$\mathcal{L} = \frac{1}{2} \partial_\nu \rho \partial^\nu \rho + \frac{1}{2} \rho^2 \partial_\nu \theta \partial^\nu \theta - \frac{m_\rho^2}{2} (\rho - \rho_0)^2 - \alpha \theta \quad (5.23)$$

This embedding was already studied in [213], where very small values of ρ_0 ($\rho_0 \sim 10^{-4}$) were considered in order to find large reductions in the speed of sound, as a working example of how decoupling works in this setup. We complement those results with the predictions for the values of ρ_0 considered here, in order to show how the theory flows from the vanilla linear potential to the new predictions. The v.e.v. of the radial field can be found by solving the reduced e.o.m.. Again, we can set $V = V(\theta)$, so that we have the following algebraic equation for the radial field $\bar{\rho}$:

$$\frac{\alpha}{3\bar{\rho}^4 \theta} = m_\rho r^2 \left(1 - \frac{\rho_0}{\bar{\rho}} \right) \quad (5.24)$$

Here, the solution for $\bar{\rho}$ will explicitly depend on both θ and ρ_0 . However, the dynamics in ρ is such that its time derivatives are still negligible in the e.o.m. Because $\bar{\rho}(\theta)$ is a complicated function of θ , it is not easy to find an effective potential. Fortunately, we can solve the system by considering the slow-roll parameters as given in (5.15). The solution of this system is such that ϵ becomes *bigger* as we decrease ρ_0 . In principle this is bad news since we would not like to move away from the 1σ contour of Planck data (see figure 5.2).

Fortunately, as ρ_0 decreases, c_s decreases, which dominates over the increase in ϵ . This means that the tensor to scalar ratio r decreases as ρ_0 decreases. We show these effects in fig. 5.2 where we plot the full EFT i.e.

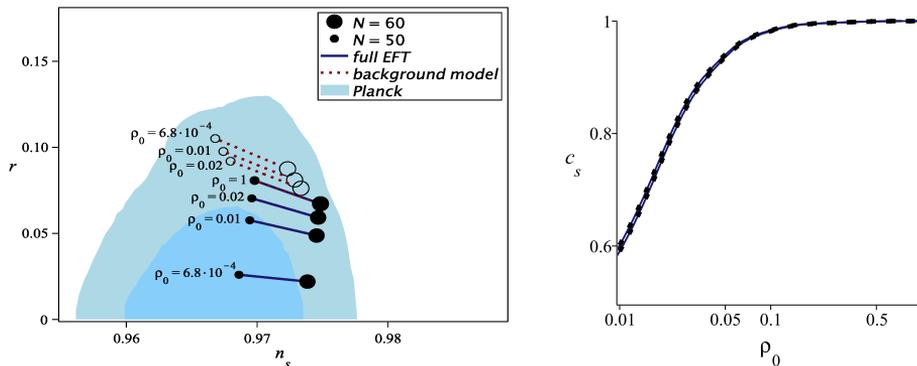


Figure 5.2: *Left:* The (n_s, r) plane for inflation with linear potential when embedded in the two-field model given by (5.23), with the mass of the heavy field given by $m_{\text{eff}}^2 = 100H^2$. The points are calculated using the EFT, which we have checked is an excellent approximation to the numerical two-field prediction. For comparison, we also plot the predictions if we had only considered the effects on the background (red dotted line). *Right:* Speed of sound of the adiabatic fluctuation equation (5.16) for $N = 50 - 60$. The dotted line is computed using the full two-field numerical solution of (5.23), and the solid line is the semi-analytical single-field approximation equation (5.15).

considering the combined effect of background and perturbations, and the predictions considering only the effects on the background. We also add the prediction for the parameters considered in [213] ($\rho_0 = 6.8 \times 10^{-4}$, $m_{\text{eff}}^2 = 250H^2$).

As for the quadratic potential, larger reductions in the speed of sound can easily be attainable provided we consider smaller values for ρ_0 . This will also increase the consistency with the data as well as generating potentially observable non gaussianities (e.g for, $\rho_0 = 10^{-3}$ we find $|f_{NL}^{eq}| \sim 5$).

5.5 Natural inflation

Finally, we consider the case for natural inflation [255]. The total Lagrangian is given by

$$\mathcal{L} = \frac{1}{2} \partial_\nu \rho \partial^\nu \rho + \frac{1}{2} \rho^2 \partial_\nu \theta \partial^\nu \theta - \frac{m_\rho^2}{2} (\rho - \rho_0)^2 - \Lambda^4 (1 + \cos(m\theta)). \quad (5.25)$$

This two-field completion is consistent with the original motivation of the inflaton being a Nambu-Goldstone boson, as the additional field ρ respects the $U(1)$ symmetry. If we assumed that the field ρ acquires its v.e.v. at the

minimum of the potential, and defining the canonical field as $\phi = \rho_0\theta$ we would get the standard potential, given by:

$$\mathcal{L} = \frac{1}{2}\partial_\nu\phi\partial^\nu\phi - \Lambda^4 \left(1 + \cos\left(\frac{\phi}{f_0}\right) \right). \quad (5.26)$$

where $f_0 = \rho_0/m$ would be the so-called decay constant. Again, we will show that in general the dynamics of the radial field cannot be neglected - even at energies below the spontaneous symmetry breaking scale - so that (5.26) is not a good description of (5.25).

While the natural - or axionic - potential is well motivated from the point of view of generic extensions of the Standard Model, the fact that decay constants larger than M_{pl} are needed in order to achieve successful inflation is a problem from the point of view of the UV completion [286]. Attempts to construct effectively super planckian decay constants by considering several axions coupled together were first considered in [269, 270]. Here instead, we take a different route. We will show that the presence of a heavy degree of freedom can improve the situation, in the sense that the overlap in the (n_s, r) plane between the predictions and the experimentally allowed region is greater. Since we cannot write an effective potential of the form (5.26), it is not completely fair to talk about a decay constant in our model. However, if we consider an *instantaneous* decay constant i.e. $f = \bar{\rho}/m$ (which is changing adiabatically as $\dot{f}/Hf = \dot{\bar{\rho}}/H\bar{\rho} \ll 1$), we will show that indeed subplanckian values of f could be consistent with the data.

As in the case of the linear potential, it is a difficult task to find an effective potential, but we can compute the single field predictions by using the expressions for ϵ , η and ΔN as in eqs. (5.15) and c_s and s in eqs. (5.16) and (5.17). We compare these predictions with the two-field model, in this case by considering two values for m , $m = \{0.01, 0.002\}$, and find an excellent agreement. The results are shown in figures 5.3 and 5.4.

Interestingly, as in the case of the linear and quadratic potential, the predictions move towards the best fit region. While the potential steepens due to the presence of the heavy field (so one might think the two-field model is disfavored with respect to the single field model), the reduced speed of sound and its variations causes the predictions to move towards the allowed experimental region. We also notice that $0.8M_{pl} < f < 1M_{pl}$ is consistent with Planck data for the case $m = 0.002$. In the region where sub-planckian decay constants overlap with Planck contours, $|f_{NL}^{eq}| \sim 1$. Larger values of $|f_{NL}^{eq}|$ can be achieved by considering smaller values of m . Whether the Lagrangian (5.25) arises in some UV completion remains however an open question.

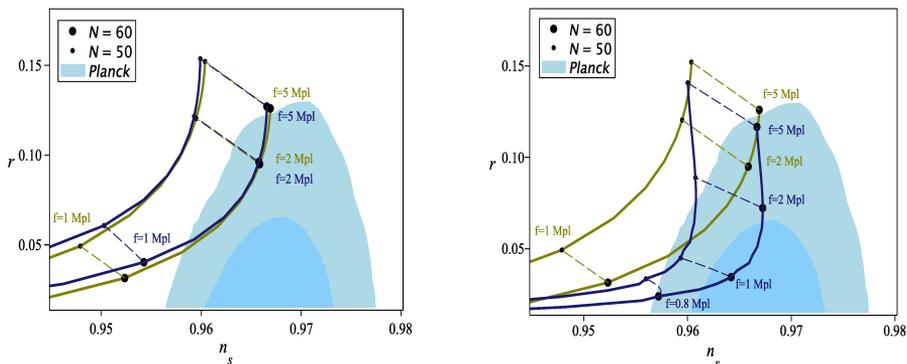


Figure 5.3: The (n_s, r) plane for inflation with the natural potential, with the mass of the heavy field given by $m_{\text{eff}}^2 \sim 100H^2$. The green line is the standard natural inflation scenario given by (5.26). The blue line is the two-field model in (5.25), for different values of the instantaneous decay constant $f = \bar{\rho}/m$. The agreement between the numerical and semi-analytical predictions is excellent. Although the potential steepens, the predictions moves towards the best fit region because of c_s . *Left)* $m = 0.01$ *Right)* $m = 0.002$

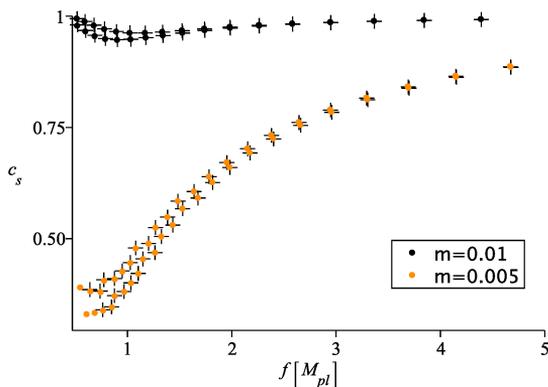


Figure 5.4: Speed of sound of the adiabatic fluctuations in the natural inflation model as a function of the instantaneous decay constant $f = \bar{\rho}/m$, for two values of m , $m = 0.01$ and $m = 0.002$. Each data point is double, representing the values for c_s at $N=50$ (lower) and $N=60$ (upper). The dots are computed using the full two-field model while the crosses are obtained using the semi-analytical single field approximation.

5.6 Conclusion

Natural inflation and quadratic inflation are theoretically appealing scenarios but their single field realizations are in tension with the data. In general, one expects the single field description to be an EFT in which the effects of massive

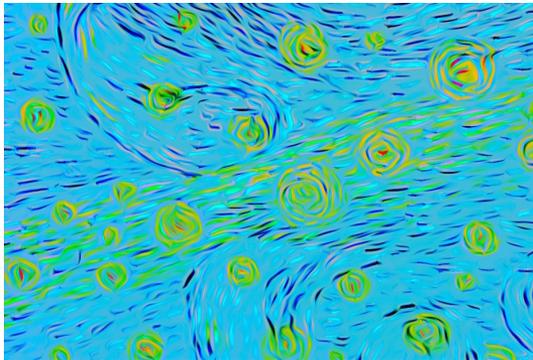
fields have been integrated out. Crucially, while in some cases truncating the heavy field is a good approximation to construct the EFT, there is a regime in which the inflationary predictions can be very different from those of the single field truncated theory. This can happen even if the additional fields have masses much greater than the Hubble parameter. Considering that the heavy field tracks its instantaneous, adiabatic, ground state along the inflationary trajectory (which, on a turning trajectory, is displaced from the minimum of the potential) leads to modifications of the background evolution, as well as reducing the speed of sound of the light mode fluctuations. As we have shown, both effects are crucial in obtaining the correct predictions for where a model lies in the (n_s, r) plane, as well as the expected level of non gaussianities.

In this chapter we have illustrated this idea in a very simple two-field embedding of various large field inflation models, in which the inflaton is approximately the phase of a complex field where a $U(1)$ symmetry is mildly broken. We find that this embedding can, in a weakly coupled regime, make models that are in tension with the data, viable. In particular, the quadratic potential and the natural potential with subplanckian values of the decay constant are no longer disfavored. Although the effective potential *steepens* in these examples, the effect of the speed of sound on the perturbations dominates: while ϵ increases, the tensor to scalar ratio goes *down* due to the reduced speed of sound, c_s . Furthermore, adiabatic changes in c_s along the trajectory can also modify of the spectral index. We have presented an analytic approximation which enables us to easily calculate all the relevant observables

Finally, we should add that this phenomenology is not restricted to the particular embedding studied here i.e. a flat field-space metric and a separable potential. The essential characteristic is that inflation happens on a curved trajectory (with respect to the field-space metric) with a large, sustained mass hierarchy. The same approximate symmetry that protects the inflaton mass and slow-roll parameters also keeps the radius of curvature and the mass of the heavy orthogonal direction approximately constant (for a recent discussion on adiabaticity and the slow-roll conditions see [213] and references therein). In this regime the radial field can be integrated out, and the isocurvature perturbations decay quickly outside the horizon. The resulting effective theory for the perturbations has a reduced speed of sound that changes slowly along the trajectory. As shown here, the effects of this reduction can be very important in obtaining the correct predictions for inflationary observables.

Deel II

Large Scale Structure



Lifting primordial non-Gaussianity above the noise

In this chapter we present a Fisher analysis to understand whether the ‘Effective Theory of Large Scale Structure’ can help us to improve the constraints on primordial non-Gaussianities in near future large scale structure surveys, such as Euclid. As a first step we focus exclusively on the matter bispectrum. Already in this simplified set-up we find that it is unlikely to reach the theoretical benchmarks quoted in Eq. 1.2.4. On the other hand, the EFT reduces the size of the error bars by a factor of 3 compared to standard perturbation theory in this set-up.

This chapter is organized as follows. In section 6.2, we review the results for the matter bispectrum in the Effective Theory of Large Scale Structure accounting for primordial non-Gaussianities. In section 6.3, we discuss the details of the Fisher forecast with particular emphasis on a consistent treatment of theoretical uncertainties. Section 6.4 is devoted to a discussion of the results of the Fisher forecast on primordial non-Gaussianities constraints from large scale structure surveys. We conclude in section 6.5. Several appendices contain more technical results. In Appendix 6.A, we summarize all relevant formulae to compute the bispectrum in the Effective Theory of Large Scale Structure. In Appendix 6.B, we present a detailed discussion of how to consistently account for theoretical errors. Appendix 6.C discusses the issue of binning the data for the Fisher forecast and finally, for the convenience of the reader, we collected all symbols used in this paper and their meaning in Appendix 6.D.

This chapter is based on [141]:

Lifting Primordial Non-Gaussianity Above the Noise, Y. Welling, D. van der Woude, and E. Pajer, JCAP **1608** (2016) 08, 044, (arXiv:1605.06426 [astro-ph.CO]).

6.1 Introduction and summary

Primordial deviations from Gaussianity are key to understand inflation and will soon be tested via a number of ambitious Large Scale Structure (LSS) surveys. It is therefore imperative to understand how late-time LSS observations can be related to the parameters that characterize *primordial non-Gaussianity* (PNG). This relation is complicated and non-linear. The degree to which we will be able to collect further primordial information from LSS survey will eventually be determined by our ability to model this non-linear relation. In this work, we focus on specific source of non-linearities, namely perturbative *matter non-linearities*. These are generated by the sub-horizon gravitational evolution of small initial matter inhomogeneities into larger ones, until the perturbations compete locally with the homogeneous background expansion. For concreteness, we study local, equilateral and quasi-single field non-Gaussianity, since these are well-motivated theoretically and represent signals that are complementary from the point of view of observations. Additional sources of non-linearity are also important, such as for example bias and redshift space distortion. In case of equilateral and quasi single field PNG, these are expected to further worsen our ability to constraint primordial parameters. In this sense, our results can be interpreted as *lower bounds* on the precision of future constraints. For local PNG, it is possible that non-linearities encapsulated in the biasing of tracers, if very well understood, might eventually help us improve on the bounds we find here (see [129] and e.g. [130, 287] for a recent estimate). We will discuss this possibility in subsection 6.4.1.

In our analysis, we will use the Effective Field Theory of Large Scale Structures (EFT of LSS) [159], which builds on Standard Perturbation Theory (SPT) [149], and provides a consistent perturbative approach to describe the evolution of matter distribution. We focus exclusively on the matter bispectrum, since it is a primary probe of PNG that is affected by matter non-linearities. Recent work on PNG and the bispectrum includes [130, 288–293]. Within the EFT approach, the bispectrum generated by the late-time gravitational evolution from otherwise Gaussian initial conditions has been studied in [181, 182]). This contribution plays the role of background noise in PNG searches. The signal, namely the primordial bispectrum, is also affected by gravitational non-linearities. This has been recently studied in [183]. Here, we use these two results and present a Fisher forecast for constraints on PNG. A key element of our forecast is the inclusion of theoretical error, employing and further developing the recent proposal of [294].

For the convenience of the reader, we collect here our major findings with references to where they are discussed in the rest paper.

- When using the EFT of LSS, the perturbative approach to model matter

non-linearities will not prevent upcoming LSS surveys to *improve upon the current bounds* from CMB anisotropies [118] (see Table 6.1).

- Our limited perturbative understanding of matter non-linearities limits the achievable bounds on equilateral non-Gaussianity from planned galaxy surveys to $\sigma(f_{NL}^{eq}) \gtrsim 10$ (see Table 6.1), *far from the theoretically interesting benchmark* $\sigma(f_{NL}^{eq}) \sim 1$ (see e.g. [103, 104] and references therein). We estimate that this remains true even if one included the full two-loop corrections (see Table 6.2). Local non-Gaussianity is more promising, and we find e.g. for Euclid $\sigma(f_{NL}^{loc}) \gtrsim 1$.
- The consistent treatment of short-scale effects within the EFT approach allows one to *improve PNG constraints by a factor of about 3* (see Table 6.3). This relies on two facts. First, the EFT parameters provide a better description of the late-time gravitational non-linearities (the “background” discussed in [181]). Second, for the specification of most upcoming experiments, the EFT parameters are only weakly correlated with PNG, and so marginalizing over them hardly degrades the constraints (see subsection 6.4.2).
- Both the SPT loops and the EFT corrections to primordial non-Gaussianity (the “signal” discussed in [183]) are small and their inclusion does *not* improve the PNG constraints appreciably (see first and second line of Table 6.3).
- We discuss several aspects of the method proposed in [294] to *model the theoretical error* inherent to the perturbative approach. We show that a wrong shape for the theoretical error can lead to a biased estimate for f_{NL} . This happens when it partly underestimates the error. Conversely, a conservatively large theoretical error leads to correct unbiased results. We thoroughly analyze the dependence of the Fisher forecast on the correlation length used in [294], and explain our results with a toy model.

Conventions Redshift z and conformal spatial coordinates \mathbf{x} are used as measures of time and position. We use the following convention for the Fourier transform

$$F(\mathbf{x}) = \int_{\mathbf{k}} \tilde{F}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}, \quad \text{where we use the shorthand} \quad \int_{\mathbf{k}} \equiv \int \frac{d^3\mathbf{k}}{(2\pi)^3}. \quad (6.1)$$

In particular, this implies that we have the following relation between any N -point equal-time correlation function and its spectrum

$$\langle \delta(\mathbf{k}_1) \dots \delta(\mathbf{k}_N) \rangle = (2\pi)^3 \delta_D(\mathbf{k}_1 + \dots + \mathbf{k}_N) S(\mathbf{k}_1, \dots, \mathbf{k}_N), \quad (6.2)$$

where we suppressed the time dependence.

For the numerical analysis, we compute the linear power spectrum with CAMB [173], where we assume a standard cosmological model with $\Omega_\Lambda^0 = 0.728$, $\Omega_m^0 = 0.272$, $h = 0.704$, $n_s = 0.967$ and $A_c = 2.46 \times 10^{-9}$.

6.2 Analytical predictions for the bispectrum

In this section, we review the analytical predictions for the late-time matter bispectrum within the Effective Field Theory of Large Scale Structures (EFT of LSS), accounting for non-Gaussian initial conditions. In subsection 6.2.1 we collect the contributions to the bispectrum up to first order in primordial non-Gaussianity and up to ‘one-loop’ order in perturbation theory. In subsection 6.2.2, we specify the types of primordial non-Gaussianity (PNG) we study in this paper. In subsection 6.2.3, we discuss the theoretical errors, which are intrinsic to the perturbative approach.

6.2.1 The bispectrum in the EFT of LSS

Despite being almost collisionless, cold dark matter on large scales behaves approximately as a fluid. This relies on the fact that the primordial universe is locally in (thermodynamical) equilibrium *and* that, during the age of the universe, dark matter particles move only over a distance that is small compared with the scales of interest. This displacement plays the same role as the mean free path in the more familiar interacting fluids. As long as we consider scales much larger than this displacement, an effective fluid description can be applied [159]. Here we follow the Effective Field Theory approach to Large Scale Structures (EFT of LSS). The dark matter bispectrum induced by gravity was discussed in [181, 182]. Non-Gaussian initial conditions were subsequently accounted for in [183]. The EFT of LSS allows to perturbatively compute non-linear correlators of the matter density contrast $\delta(\mathbf{x}, z)$ [162] and velocity $v(\mathbf{x}, z)$ [178], taking into account the effect of short-scale non-perturbative physics on the large-scale dynamics. In practice, one can use the results of Standard Perturbation Theory (SPT) [149], and correct them with additional effective terms, which will be denoted with the subscript ‘EFT’. We differentiate between contributions to the bispectrum coming from primordial non-Gaussianities (superscript “NG”) and those coming from the late-time gravitational evolution (superscript “G”). Schematically, the perturbative theoretical prediction for the bispectrum is

$$B^{\text{th}} = B_{\text{SPT}}^{\text{G}} + B_{\text{EFT}}^{\text{G}} + f_{\text{NL}} (B_{\text{SPT}}^{\text{NG}} + B_{\text{EFT}}^{\text{NG}}). \quad (6.3)$$

As we will see in section 6.3, for a Fisher forecast we do not need to specify¹ $B_{\text{SPT}}^{\text{G}}$. The leading order counterterms for Gaussian initial conditions have been computed in [181, 182] and read

$$B_{\text{EFT}}^{\text{G}} = \xi B_{\xi}^{\text{G}} + \sum_{i=1}^3 \epsilon_i B_{\epsilon_i}^{\text{G}}. \quad (6.4)$$

For non-Gaussian initial conditions, short modes and long modes are already correlated at the initial time. This leads to additional contributions to the matter bispectrum. To leading order, these are given by [183]

$$B_{\text{EFT}}^{\text{NG}} = \xi B_{\xi}^{\text{NG}} + \gamma B_{\gamma}^{\text{NG}} + \sum_{i=1}^2 \gamma_i B_{\gamma_i}^{\text{NG}}. \quad (6.5)$$

For convenience, we have adopted the notation of [183] and collected in appendix 6.A all the explicit expressions for the terms appearing in this subsection.

6.2.2 Primordial non-Gaussianity

To evaluate the non-Gaussian contributions to (6.3), we need to specify the primordial bispectrum. In this paper, we study the constraints on three types of primordial non-Gaussianity: local [95], equilateral [97] and quasi-single field [100]. In terms of the primordial potential ϕ , the primordial bispectra are given by the following shapes

$$B_{\phi}^{\text{loc}}(k_1, k_2, k_3) = 2f_{\text{NL}}^{\text{loc}} (P_{\phi}(k_1)P_{\phi}(k_2) + \text{perm}), \quad (6.6a)$$

$$B_{\phi}^{\text{eq}}(k_1, k_2, k_3) = 162f_{\text{NL}}^{\text{eq}} A_{\phi}^2 \frac{1}{k_1 k_2 k_3 K^3}, \quad (6.6b)$$

$$B_{\phi}^{\text{qsf}}(k_1, k_2, k_3) = 18\sqrt{3}f_{\text{NL}}^{\text{qsf}} A_{\phi}^2 \frac{1}{k_1 k_2 k_3 K^3} \frac{N_{\nu}(8\kappa)}{\sqrt{\kappa}N_{\nu}(8/27)}. \quad (6.6c)$$

Here we define $K = k_1 + k_2 + k_3$, and $\kappa = k_1 k_2 k_3 / K^3$. Moreover, N_{ν} is the Neumann function of order ν and we choose $\nu = \frac{1}{2}$. The normalization of the primordial power spectrum² is given by $A_{\phi} = 1.72 \cdot 10^{-8}$. To linearly evolve these to the late time matter bispectrum³ B_{111} , we use the transfer function $M(k, z)$, defined by

$$\delta_1(k, z) = M(k, z)\phi(k), \quad \text{with} \quad M^2(k, z) \equiv \frac{k^3 P_{11}(k, z)}{A_{\phi} \left(\frac{k}{k_{\star}}\right)^{n_s-1}}. \quad (6.7)$$

¹On the other hand, we do need to specify the SPT contributions to the power spectrum to compute the cosmic variance. Assuming it is dominated by the linearly evolved matter power spectrum, we do not have to specify additional ‘EFT’ parameters.

²Note that we define $A_{\phi} = 2\pi^2 \frac{9}{25} A_{\zeta}$.

³See appendix 6.A for relevant notation.

Here $k_* = 0.0028 \text{ hMpc}^{-1}$ and $n_s = 0.967$. This means we have

$$B_{111}(k_1, k_2, k_3, z) = M(k_1, z)M(k_2, z)M(k_3, z)B_\phi(k_1, k_2, k_3). \quad (6.8)$$

We collect all relevant higher order non-Gaussian contributions to the bispectrum in appendix 6.A.

6.2.3 Theoretical error

By definition, any results from perturbation theory are approximate - there is always an intrinsic theoretical error, typically estimated within perturbation theory itself. The true bispectrum is therefore given by

$$B^{true} = B^{th} + B^{er}, \quad (6.9)$$

where B^{th} is the perturbative theoretical prediction given in (6.3), and B^{er} represents the theoretical error. The strength of a well defined perturbation theory is of course that the error can be estimated within the perturbation theory itself.

In our case, there are in principle two perturbative schemes employed. First, we assume perturbative primordial non-Gaussianity. This means we assume the primordial potential can be schematically expanded as

$$\varphi_p = \varphi_p^G + f_{\text{NL}}\varphi_p^G \star \varphi_p^G + \dots \quad (6.10)$$

Here φ_p^G is a Gaussian field and \star denotes a convolution in Fourier space. This means we are effectively expanding in $f_{\text{NL}}\varphi_p \sim f_{\text{NL}}\sqrt{A_\phi}$, which is indeed very small according to current bounds. Hence we will not worry about corrections to this approximation for the rest of the paper.

Second, the EFT of LSS relies on the smallness of density perturbations on large scales, consistently taking into account our ignorance of short scale physics. Effectively, this comes down to an expansion in k/k_{NL} [159]. As argued in [181], the most relevant correction to B^{th} is the two-loop bispectrum. Since we have not computed the full two-loop bispectrum, we are forced to make an educated guess about its size and shape. One way to do this was proposed in [294], and relies on the scaling universe results of [184]. Here we use instead a different estimate. Unless indicated otherwise, we estimate the two-loop bispectrum by adding up the absolute values of the two two-loop diagrams we can compute, namely the so-called reducible diagrams, which we indicate by B_{332} . An explicit expression for B_{332} is again given in appendix 6.A. We compare our estimate to the scaling estimate of [294] in appendix 6.B.5.

The importance of keeping track of the theoretical error for forecasts has recently been stressed in [294], and we build on their approach. Qualitatively,

one expects not to be able to learn much about f_{NL} from bispectrum configurations for which B^{er} is larger than the non-Gaussian signal. To get an idea of the configurations for which this is the case, we plot the one-loop expressions for the three types of non-Gaussianities we consider (with $f_{\text{NL}} = 1$) and B_{332} as a function of scale for both squeezed and equilateral configurations in Figure 6.1. For reference we also plotted the one-loop Gaussian contribution to the bispectrum. As expected, for local PNG we can push to smaller scales in the squeezed configuration than for equilateral PNG. Note also that the naive k_{max} , beyond which we do not expect to gain any more signal, is *configuration dependent*. A detailed discussion on how to incorporate this theoretical error in a Fisher analysis is given in section 6.2.3, which proceeds along the lines of [294]. In appendix 6.B, we present further investigation of the validity of this method of treating the theoretical error.

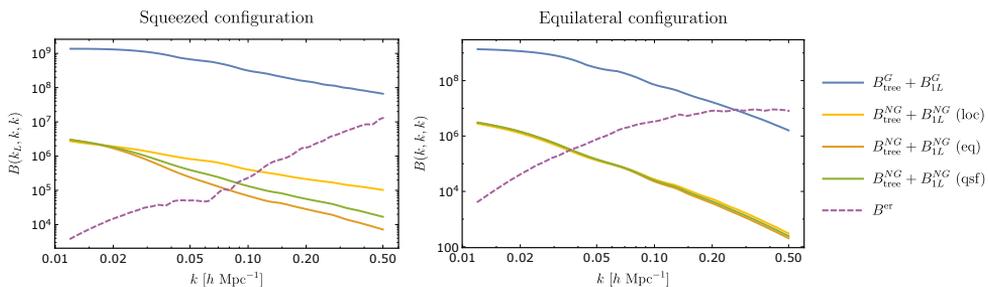


Figure 6.1: The SPT contributions to the bispectrum in the squeezed (left) and equilateral (right) configurations. The blue solid line denotes the Gaussian tree-level and one-loop contributions. The yellow, orange and green lines denote the one loop non-Gaussian contribution for $f_{\text{NL}} = 1$ for local, equilateral and quasi-single-field PNG, respectively. The dashed purple line corresponds to our order-of-magnitude estimate for the Gaussian two-loop correction B_{332} . In the squeezed configurations (left), we chose $k_L = 0.012 h\text{Mpc}^{-1}$.

6.3 Fisher analysis

In this section, we outline our method to forecast constraints on primordial non-Gaussianity. We have in mind a Gedankenexperiment that provides us with the matter distribution in space and time up to some maximal redshift. In this highly idealized scenario, we determine to what extent our inability to analytically describe the non-linear gravitational collapse of matter limits the information we can extract on primordial perturbations. We proceed along the lines of [288, 289, 295]. The outcome of the analysis for various surveys is presented in the next section.

6.3.1 Assumptions and approximations

For the convenience of the reader, we summarize the assumptions and approximations we make in the Fisher analysis.

- We assume we are given an idealized *survey of the late time dark matter density field*, instead of that of some biased tracer. This allows us to answer the question of whether further progress is needed in the modeling of the dark matter distribution to strengthen current bounds on PNG using upcoming LSS surveys.
- The idealized dark matter survey is characterized by a redshift range and the fraction of the sky covered. We divide the survey in redshift bins, to which we assign a fixed time that is equal to the mean redshift of the bin. Hence, we only need to know z_{bin} to predict the power spectrum and bispectrum. Observational redshift errors are neglected.
- We assume that each redshift bin can be approximated by a cube. Then we just need the volume of the bin $V(z_{\text{bin}})$ to account for cosmic variance.
- We compute correlation functions only within each bin. This does not seem to be a big drawback in the case of equilateral PNG. Instead, for local PNG, this might cause an unnecessary loss of information. We will discuss this issue elsewhere.
- We include shotnoise in the analysis to correctly remove weight from the higher redshift bins. For this, we use the specifications of specific upcoming surveys. We discuss this in section 6.4.1.
- We assume that the bispectra for different configurations are uncorrelated with each other. This means that we approximate the *bispectrum covariance matrix as diagonal*. In [295] it has been checked that this approximation works fine for the scales $k \leq 0.3h\text{Mpc}^{-1}$ at redshift zero. We assume it holds up to $k \leq 0.4h\text{Mpc}^{-1}$, since for local PNG we still gain signal up to this scale, as we see in Figure 6.9. Moreover, we assume that only the linear power spectrum determines the covariance matrix (see subsection 6.3.3 for more details). Finally, we neglect covariance due to observational effects, such as survey geometry and mask.

Importantly, we parameterize the theoretical error by treating the higher loop corrections to the bispectrum as a source of noise, which we integrate out. This contributes to the effective covariance matrix. Our estimate of the typical size of the two-loop corrections is given by B_{332} , defined in Appendix 6.A.

The time-dependence of the counterterms has been chosen to match the one

loop diagrams they are supposed to renormalize [296] (see also [297] for a related discussion). All the time dependence is absorbed in the contributions to the bispectrum, so that all the theoretical parameters become time-independent (see appendix 6.A). This means that we are measuring the same theoretical parameters in each redshift bin.

We need to discretize the bispectrum in order to compute the Fisher matrix. We will use logarithmic bins instead of linear bins, since we do not fully trust linear binning. We refer the reader to the appendix 6.C for more details. Finally, we do not marginalize over the standard cosmological parameters, but fix their values.

6.3.2 Fisher matrix

In a Fisher forecast, one computes the expected curvature around the maximum of the likelihood. The likelihood is given by

$$\mathcal{L}(\text{data}|\Theta, \text{priors}) = \frac{1}{\sqrt{(2\pi)^N \det(C_B)}} P_{\text{prior}}(\Theta) \quad (6.11)$$

$$\times \exp \left(-\frac{1}{2} \sum_{\hat{k}p} \Delta B(\hat{k}, \Theta) C_B^{-1}(\hat{k}, q, \Theta) \Delta B(p, \Theta) \right),$$

where Θ denote the set of theoretical parameters and N is the number of datapoints. We suppressed the time dependence. Here we use the Dutch calligraphic lower case symbols \hat{k} as a shortcut for a triplet of wavenumbers on which the bispectrum depends, i.e. $\hat{k} = (k_1, k_2, k_3)$. Furthermore, ΔB corresponds to the difference between estimator and theoretical prediction $\Delta B(\hat{k}, \Theta) = \hat{B}(\hat{k}) - B(\hat{k}, \Theta)$ and C_B is the covariance matrix of the bispectrum $C_B = \langle \Delta B \Delta B \rangle$.

Neglecting the theoretical error for the moment, the theoretical prediction for the bispectrum is given in equation (6.3). The 8 parameters we include in the Fisher analysis are therefore $\{f_{\text{NL}}, \xi, \epsilon_1, \epsilon_2, \epsilon_3, \tilde{\gamma}, \tilde{\gamma}_1, \tilde{\gamma}_2\}$ ⁴. The parameter ξ also appears in the power spectrum and has been measured to be⁵ $0.98 h^{-2} \text{Mpc}^2$ [296]. Therefore, we can put a sharp prior on this parameter. The other parameters are unknown, but we expect them to be of the same

⁴Here we denote $\tilde{\gamma} = f_{\text{NL}} \cdot \gamma$ and similarly for γ_i , so that the bispectrum is linear in all parameters. This is convenient for the Fisher analysis, as this makes the result independent of the fiducial values of the parameters. On the other hand, we effectively assume that the one-loop non-Gaussian counterterms have amplitudes independent of f_{NL} . Later in this paper we will find that these counterterms are negligibly small, therefore, this will not affect our results.

⁵Previous measurement gave $(1.62 \pm 0.03) h^{-2} \text{Mpc}^2$ [298] and $(1.5 \pm 0.03) h^{-2} \text{Mpc}^2$ [181], but neglected two-loop corrections.

order of magnitude (see [181, 183] for naive numerical estimates). Therefore, we can use a fiducial value of zero and a Gaussian prior with variance of 10.

For simplicity, we assume that all priors are Gaussian with covariance matrix C_Θ . Then the Fisher matrix is given by (see e.g. [299])

$$F_{ij} \equiv - \left\langle \frac{\partial^2 \log(\mathcal{L})}{\partial \Theta_i \partial \Theta_j} \right\rangle = \frac{1}{2} \text{Tr} [C_B^{-1} C_{B,i} C_B^{-1} C_{B,j}] + B_i^T C_B^{-1} B_j + (C_\Theta^{-1})_{ij}. \quad (6.12)$$

As we will see in a moment, our approximation for the covariance matrix does not depend on the theoretical parameters. Writing out the time dependence explicitly, the Fisher matrix simplifies to

$$F_{ij}(z) = \sum_{k,k'} B_i^T(k, z) C_B^{-1}(k, k', z) B_j(k', z) + (C_\Theta^{-1})_{ij}(z) \quad (6.13)$$

for each redshift bin. Since the bispectrum is linear in all parameters - taking into account that ξ has been measured - the Fisher matrix is independent of the fiducial value of the theoretical parameters⁶, which is very convenient for the analysis. To combine the results from the different redshift bins, we use that the parameters are the same in each bin, since we have fixed their time dependence. This time dependence is chosen to match the time dependence of the divergences they are supposed to cancel, motivated by [296]. The explicit expression can be found in Appendix 6.A. Therefore, we can compute the constraints on f_{NL} by summing the Fisher matrices and then marginalizing over the EFT parameters i.e.,

$$\sigma(f_{\text{NL}}) = \sqrt{\left(\sum_z F_{ij}(z) \right)_{11}^{-1}}, \quad (6.14)$$

where we assumed that the entry of the Fisher matrix belonging to f_{NL} is the first. Note that we have not included cross correlations between bins, which means we might be throwing away valuable information.

⁶To be more precise, the Fisher matrix is independent of the fiducial values of the parameters *to good approximation*. We choose a fiducial value for ξ of zero and in Section 6.4 we either specify a prior with $\sigma = 1$ for ξ or no prior at all. Therefore, the Fisher matrix has some dependence on the choice of ξ , but it will come exclusively from the non-Gaussian counterterm proportional to ξ . Again, since the non-Gaussian counterterms turn out to be extremely small, we expect this not to affect the results. Moreover, we have checked this explicitly by changing its fiducial value to 1.

6.3.3 Covariance matrix

To evaluate the Fisher matrix, we need to know the covariance matrix. Let us shortly review the derivation of the bispectrum covariance matrix. The estimator of the bispectrum is given by [190]

$$\hat{B}(k_1, k_2, k_3, z) = \frac{1}{V(z)V_{123}} \int_{\mathbf{q}_1} \int_{\mathbf{q}_2} \int_{\mathbf{q}_3} \delta(\mathbf{q}_1, z) \delta(\mathbf{q}_2, z) \delta(\mathbf{q}_3, z) \delta_D^3(\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3), \quad (6.15)$$

with $V(z)$ the volume of the bin. The integration is over logarithmic bins centered around the given wavenumbers, i.e. $\ln(|\mathbf{q}_i|) \in [\ln(k_i) - \frac{1}{2}\Delta \ln(k), \ln(k_i) + \frac{1}{2}\Delta \ln(k)]$. Moreover, V_{123} corresponds to the following k -space volume squared

$$V_{123} = \int_{\mathbf{q}_1} \int_{\mathbf{q}_2} \int_{\mathbf{q}_3} \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \approx \frac{8\pi^2}{(2\pi)^9} k_1^2 k_2^2 k_3^2 \sinh^3(\Delta \ln k). \quad (6.16)$$

This approximation becomes exact when we consider ‘internal’ bins, but it fails on the ‘edge’ bins. In the numerical analysis we compute the exact value of V_{123} for each bin, see appendix 6.C for more details. This allows us to compute the bispectrum covariance matrix, namely

$$C_B(k, k', z) = \left\langle \left(\hat{B}(k, z) - B(k, z) \right) \left(\hat{B}(k', z) - B(k', z) \right) \right\rangle \quad (6.17)$$

$$\approx \frac{s_{123}}{(2\pi)^3 V(z) V_{123}} P(k_1, z) P(k_2, z) P(k_3, z) \delta_{k, k'}. \quad (6.18)$$

There is a factor s_{123} which counts the number of non-vanishing contractions when computing $\langle \hat{B}(k) \hat{B}(k) \rangle$, which depends on the type of triangle that the triplet k forms. As each contraction comes with a delta function, this counting factor equals 6, 2 or 1 for equilateral, isosceles and scalene triangles respectively. If we include shotnoise in the covariance matrix, we replace $P(k_i, z)$ with $P(k_i, z) + \frac{1}{\bar{n}}$ in equation (6.18), where \bar{n} is the effective number density for the density contrast. This will be explained more when we include shotnoise in Section 6.4.1.

In this expression for the covariance matrix, we completely neglected higher order corrections beyond the power spectrum, making it approximately diagonal. In [295] it has been checked that this approximation works fine for the scales we are considering. The off-diagonal terms become important exactly when the higher order corrections to the power spectrum become important, since they are of the same order. Therefore, in order to be consistent, we only take into account the linear contribution P_{11} to the power spectrum $P(k_i, z)$. In particular, this means that the covariance matrix is independent of the theoretical parameters.

6.3.4 Theoretical error as nuisance parameters

To account for the theoretical error inherent to the perturbative expansion, we parameterize the bispectrum as

$$B(\boldsymbol{k}) = B^{\text{th}}(\boldsymbol{k}) + n(\boldsymbol{k})B^{\text{er}}(\boldsymbol{k}), \quad (6.19)$$

$$B^{\text{th}}(\boldsymbol{k}) = B_{\text{SPT}}^{\text{G}}(\boldsymbol{k}) + B_{\text{EFT}}^{\text{G}}(\boldsymbol{k}) + f_{\text{NL}} [B_{\text{SPT}}^{\text{NG}}(\boldsymbol{k}) + B_{\text{EFT}}^{\text{NG}}(\boldsymbol{k})], \quad (6.20)$$

where B^{th} represents the theoretical prediction up to some order in perturbation theory as before, and B^{er} is the estimate of the theoretical error. Following [294], we introduce a series of nuisance parameters $n(\boldsymbol{k})$, one per bin in k -space. The reason we implement the theoretical error this way, instead of proposing some k_{max} is that, as discussed in subsection 6.2.3, k_{max} depends on where the theoretical error and the signal become comparable. This complicates the analysis in two ways. First, k_{max} is configuration dependent, and second, it depends on the fiducial value of f_{NL} , which makes finding the error on f_{NL} a recursive problem. In the approach we take, the set of theoretical parameters thus becomes $\Theta = \{n(\boldsymbol{k})\}_{\boldsymbol{k}} \cup \{f_{\text{NL}}, \xi, \epsilon_1, \epsilon_2, \epsilon_3, \tilde{\gamma}, \tilde{\gamma}_1, \tilde{\gamma}_2\}$. Since the bispectrum remains linear in all parameters, expression (6.13) for the respective block of the Fisher matrix still applies.

We assume that the true corrections to the bispectrum are of similar size as $B^{\text{er}}(\boldsymbol{k})$. Therefore, we put a Gaussian prior on the parameters $n(\boldsymbol{k})$, with mean zero and variance one. Moreover, we expect the correction to have a smooth shape, which varies not too rapidly within the contours defined by $B^{\text{er}}(\boldsymbol{k})$. Therefore, the coefficients should have non-negligible cross correlations. Since we would like to have an increasing correlation for nearby points, we include cross-correlations with a typical correlation length as follows

$$N_{\alpha\beta} = \exp\left(-\frac{\sum_i \ln(|k_{\alpha}^i/k_{\beta}^i|)}{l}\right). \quad (6.21)$$

We replaced the label \boldsymbol{k}_{α} of the nuisance parameter of a given bin with the index α , so that we can reserve latin indices for the other theoretical parameters. Moreover, in order not to confuse this covariance matrix with the covariance matrix of the bispectrum C_B , we denote it as $N_{\alpha\beta}$. Note that we choose $\sigma_{\alpha} = 1$ for all α 's. Here, l denotes the logarithmic correlation length. We could have also chosen a quadratic correlation length, similar to modeling it as a random field [300]. The reason we opted for this form is that here the inverse matrix is very sparse, which is convenient for numerical purposes. Since our final results are quite insensitive to the correlation length (see 6.B.3), we do not believe this choice affects the results very much.

Since we introduced a set of new nuisance parameters, we should write down the full Fisher matrix $F_{\mu\nu}$ and invert it

$$F_{\mu\nu}^{-1} = \begin{pmatrix} F_{\alpha\beta} & F_{\alpha j} \\ F_{i\beta} & F_{ij} \end{pmatrix}^{-1} = \begin{pmatrix} \bullet & \bullet \\ \bullet & (F_{ij} - F_{i\gamma} F_{\gamma\delta}^{-1} F_{\delta j})^{-1} \end{pmatrix}, \quad (6.22)$$

where we use latin indices for the parameters $\{f_{\text{NL}}, \xi, \epsilon_1, \epsilon_2, \epsilon_3, \tilde{\gamma}, \tilde{\gamma}_1, \tilde{\gamma}_2\}$ and greek indices from the early alphabet for the theoretical error parameters $\{n(k_\alpha)\}_\alpha$. We did not write out explicitly the other entries, since we are only interested in the effective Fisher matrix, after marginalizing over the nuisance parameters coming from the theoretical uncertainty. To compute the effective Fisher matrix, we need to know $F_{\alpha\beta}$ and $F_{\alpha i}$. Since the derivative of the bispectrum with respect to the nuisance parameters Θ_α is only non-zero for the corresponding bin, these contributions to the Fisher matrix are particularly simple. We have

$$F_{\alpha\beta} = B^{\text{er}}(k_\alpha) C_B^{-1}(k_\alpha, k_\beta) B^{\text{er}}(k_\beta) + N_{\alpha\beta}^{-1} \equiv D_{\alpha\beta} + N_{\alpha\beta}^{-1}, \quad (6.23)$$

and similarly

$$F_{i\beta} = \sum_k B_i(k) C_B^{-1}(k, k_\beta) B^{\text{er}}(k_\beta) \quad (6.24a)$$

$$\text{and } F_{\alpha j} = \sum_k B^{\text{er}}(k_\alpha) C_B^{-1}(k_\alpha, k) B_j(k). \quad (6.24b)$$

This allows us to compute F_{ij}^{eff} by using (6.22). After some algebraic manipulations, we can rewrite it in the simple form

$$F_{ij}^{\text{eff}} = \sum_{k,p} B_i(k) (N^{\text{eff}}(k, p) + C_B(k, p))^{-1} B_j(p) + (C_\Theta^{-1})_{ij}, \quad (6.25)$$

with $N_{\alpha\beta}^{\text{eff}} = B^{\text{er}}(k_\alpha) N_{\alpha\beta} B^{\text{er}}(k_\beta)$. Again, the time dependence has been suppressed. In Appendix 6.B we present two alternative derivations of (6.25) and provide further detail.

In the next subsection we show the effectiveness of the current treatment of the theoretical error. However, we believe the interpretation of this method, and its relation to the actual situation, is a subtle matter. In particular, in Appendix 6.B.3 we argue by means of a simple toy model that this way of dealing with the theoretical error is certainly not the right way in the extremes of zero and maximal correlation among the parameters. Namely, on the one hand the theoretical error acts as shot noise per bin for zero correlation length, whereas for maximal correlation it acts as a single coefficient multiplying a fixed shape, effectively reducing the uncertainty about its shape to one number.

Neither of these cases correspond to the way we believe the theoretical error should act. At the same time, Appendix 6.B.3 shows that the effect of the correlation length on the results is very mild. This suggests that the main reason our method works so well is that our ansatz for the error is a much steeper function of k than the signal, so that the size of the error is much more important than its shape. Thus, even though our treatment of the theoretical error seems to work for the current case, we recommend a conservative use of the method. In this spirit, we use the correlation length that gives the most pessimistic results for the analysis, which we found to be $l \sim 0.5$.

6.3.5 Testing the effect of the theoretical error

To test the method of integrating out the theoretical error, we study its effect on the constraints on f_{NL} in a χ^2 -analysis. To that end, we compare two types of analyses, one which includes the theoretical error as outlined above, and one which does not. We generate a fake dataset with no primordial non-Gaussianity to test the theory. Our datapoints are given by

$$D(\hat{k}) = B_{112}(\hat{k}) + E_b(\hat{k}) + \text{cosmic noise}, \quad (6.26)$$

where we add some random noise, with variance equal to the cosmic variance, to each point. We consider a survey at redshift $z = 0$ with volume $V = 10 (h^{-1}\text{Gpc})^3$, and restrict (k_1, k_2, k_3) to be the central values of the binned range $[0.001, 1] h\text{Mpc}^{-1}$, where we take 27 logarithmic sized bins. The additional contribution to the bispectrum is given by

$$E_b(k_1, k_2, k_3) = 3B_{112}(k_1, k_2, k_3) \left(\frac{k_1 + k_2 + k_3}{3k_{NL}} \right)^{(3+n)l}, \quad (6.27)$$

with $n = -1.4$, $k_{NL} = 0.45$ and $l = 2$. This is exactly the ansatz for the two-loop contribution to the bispectrum used in [294] and it is based on scaling universes [184]. In appendix 6.B.5, we compare the ansatz for the higher loop corrections E_b with our ansatz B_{332} . As theoretical model for the bispectrum, we use

$$B^{\text{th}}(\hat{k}) = f_{NL} \cdot B_{111}(\hat{k}) + B_{112}(\hat{k}). \quad (6.28)$$

We now consider two analyses. In the first analysis, we neglect theoretical errors and take only cosmic variance into account. In the second analysis, we use our ansatz for the higher order corrections, namely B_{332} , and we account for both theoretical error and cosmic variance. In order to find the best fit value for f_{NL} , we minimize χ_B^2 , which is given by

$$\chi_B^2 = (D(\hat{k}) - B^{\text{th}}(\hat{k})) (C_B(\hat{k}, p) + N^{\text{eff}}(\hat{k}, p))^{-1} (D(p) - B^{\text{th}}(p)) + \text{const.}, \quad (6.29)$$

see (??) in Appendix 6.B. In the first case, we set N^{eff} to zero. Minimizing χ_B^2 yields

$$\text{Est}(f_{NL}) = \frac{B_{111}(k) (C_B(k, p) + N^{\text{eff}}(k, p))^{-1} (D(p) - B_{112}(p))}{B_{111}(k) (C_B(k, p) + N^{\text{eff}}(k, p))^{-1} B_{111}(p)}, \quad (6.30)$$

and taking another derivative with respect to f_{NL} allows us to compute the standard deviation

$$\sigma(f_{NL}) = \left(B_{111}(k) (C_B(k, p) + N^{\text{eff}}(k, p))^{-1} B_{111}(p) \right)^{-1/2}. \quad (6.31)$$

With the best-fit value of f_{NL} , we can evaluate $(\chi_B^2)_{\text{red}}$ and the p -value, which are given by

$$(\chi_B^2)_{\text{red}} = \frac{\chi_B^2}{N} \quad \text{and} \quad p\text{-value} = 1 - \text{CDF}_{\chi^2}(N, \chi_B^2), \quad (6.32)$$

with CDF_{χ^2} the cumulative distribution function of the χ^2 -distribution, and $N = N_{\text{bins}} - N_{\text{dofs}} - 1$ the number of datapoints minus one minus the number of fitting parameters. The p -value takes values between 0 and 1. It gives the probability of finding a higher value for χ_B^2 if it was drawn from a χ^2 -distribution. Therefore, it should take values around to 0.5. If the p -value is very close to zero, then the proposed theory vector is not a good description of the data. If the p -value is close to one, then either one is overfitting the data, or the estimate for the noise is too pessimistic.

In Figure 6.2, we plot the estimate for f_{NL} with errorbars, $(\chi_B^2)_{\text{red}}$ and the p -value, as given in equations (6.30), (6.31) and (6.32) respectively, for the two analyses. In the left panel, we show both the results for the analysis in which the higher order corrections are neglected, and the analysis in which we use B_{332} as an ansatz. In the right panel, we use $10 \times B_{332}$ as error estimate to make sure that our ansatz is always bigger than the true value of the higher order corrections. One can check that E_b has a different shape than B_{332} . For instance, in the equilateral configuration, E_b is smaller than B_{332} on small scales (more optimistic). On the other hand, on large scales in the equilateral configuration, and in the squeezed limit, it tends to be larger than B_{332} (more pessimistic). Upon multiplying the latter by a factor 10, we find a robust, conservative estimate (see 6.B.5).

In the left panel of Figure 6.2, we see that if we neglect the theoretical error (blue lines and contours), we get the wrong value for the best fit value for f_{NL} , because higher order corrections are mistakingly interpreted as signal. Fortunately, the p -value singles out where the theoretical description fails. Taking this into account, we get a reliable estimate for f_{NL} , albeit with larger

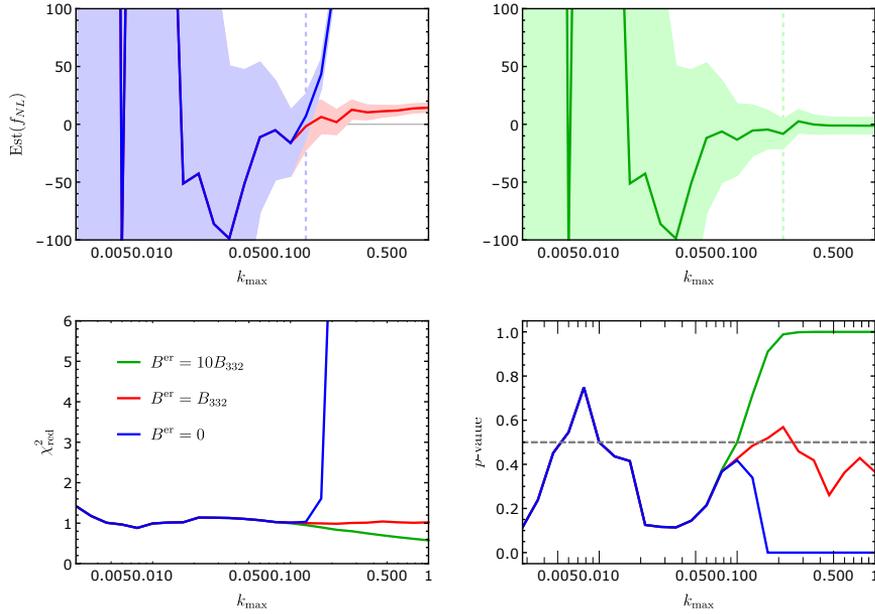


Figure 6.2: The figure shows the results from the χ -squared analysis for the data and theory given in equations (6.26) and (6.28). In the left panel, we show the results for both the analysis in which the higher order corrections are neglected (blue) and the one in which we use B_{332} as ansatz (red). In the right panel, we use $10B_{332}$ as ansatz (green) instead. In the upper panels, we show the best fit value for f_{NL} (solid line) as function of k_{max} and the lighter coloured regions correspond to the 2σ errorbars. The dashed vertical line corresponds to the largest value for k_{max} where the p -value is still between 0.01 and 0.99. The second and third row show $(\chi^2_B)_{red}$ and the p -value as function of k_{max} .

errorbars, since we have to stop already at a relatively small k_{max} . From the analysis that accounts for the theoretical error (red lines and contours), it seems we can continue the analysis to a higher k_{max} . However, the result we get for f_{NL} is biased, i.e. it is more than 5σ away from the actual value. The problem is that, in certain configurations of the bispectrum, our ansatz takes smaller values than the actual value in the data. This tends to decrease the p -value. At the same time, the p -value increases in the configurations where the theoretical error is overestimated. The interplay of these two effects can lead to a p -value, which is neither too small or too large, and this gives rise to a biased estimate. Hence, if one wants to use the p -value as diagnostic for k_{max} and avoid biased results, it is important to have a fairly good understanding of

the form of the theoretical noise. Alternatively, one can work with an ansatz which is consistently *underestimating* the theoretical error. In this case, the p -value should go to zero rapidly, as soon as the theoretical error kicks in. As a double check, we did the analysis using $0.1E_b$ as ansatz instead, which indeed gives unbiased results similarly to the case where we neglect the theoretical error altogether. In general, when performing a datafit, if one has insufficient information about the higher order corrections, it is therefore safer not to integrate out the theoretical error at all. Summarizing, *assuming the wrong shape for the theoretical error might lead to a false detection of primordial non-Gaussianity*.

In the right panel (green lines and contours), we show the same results, where now the ansatz is always more pessimistic than the actual theoretical error ($10 \times B_{332} > E_b$). In this case, the estimate for f_{NL} is equal to the real value within 2σ . The p -value now is very *large* and it would naively tells us to stop at some smaller k_{max} . However, since we are obviously not overfitting the data, this reflects the fact that our ansatz for the theoretical error is too pessimistic. Therefore, we can safely evaluate the estimate for f_{NL} and the corresponding errorbar at the highest value of k_{max} , where the errorbar is frozen to a finite value. As expected, we find better errorbars than in the case we neglect the theoretical error. This shows that integrating out the theoretical error helps constraining f_{NL} , as long as one is careful to take a conservative enough ansatz. In the next section, we take B_{332} as ansatz for the theoretical error⁷.

6.4 Results

In this section, we present the main results of our analysis. First, we give $\sigma(f_{NL})$ for various surveys, comparing our results to [294] and [287]. Next, we study the correlations among the EFT parameters for relevant surveys. Furthermore, we address the question of how much better the constraints would be, if we were able to compute the two-loop bispectrum. Finally, we show that the EFT of LSS clearly outperforms SPT in the constraints on f_{NL} , where we assume the EFT contributions to the bispectrum are part of the theoretical error in SPT.

⁷We checked that our results for $\sigma(f_{NL})$ change with less than a factor of 2, when we use $10B_{332}$ instead.

6.4.1 Constraints as function of z_{\max}

In this subsection, we compute the constraints on f_{NL} as function of maximum redshift for surveys similar to the ones studied in [294] and [287]. This allows us to study the effects of shotnoise and to compare our results with theirs. Furthermore, we show the effect of marginalizing over the EFT parameters for these surveys.

A large redshift survey (comparison with Baldauf et al. 2016)

First, to compare with [294], we focus on local and equilateral PNG. In the following, we list the specifications of the survey and the particular assumptions we make (in addition to general assumptions for the Fisher analysis discussed in subsection 6.3.1).

- We assume a survey with maximum redshift of $z = 5$. We divide the survey in redshift bins of equal volume, where the first bin runs from $z = 0$ to $z = 1$. We assume the survey covers 20000 deg^2 , which means that, with our choice of cosmological parameters, the volume of each bin is given by $V = 26.5(h^{-1}\text{Gpc})^3$. We approximate each bin to be a cube, so that $k_{\min} \approx 0.002 h\text{Mpc}^{-1}$. The maximum redshifts of all the redshift bins are given by⁸

$$\{1.00, 1.39, 1.71, 2.02, 2.31, 2.60, 2.89, 3.19, 3.49, 3.80, 4.11, 4.44, 4.78, 5.13\}.$$

- We restrict (k_1, k_2, k_3) to the values in the binned range $[0.002, 1] h\text{Mpc}^{-1}$, where we use 15 logarithmic bins per decade⁹.
- At high redshift, the late-time non-Gaussian background is particularly small and so the PNG signal becomes comparatively more pronounced. On the other hand, at high redshift, there are also much fewer tracers and this degrades our ability to measure the distribution of matter. To be able to capture this fact, we introduce a shotnoise that mimics what happens for example in galaxy surveys. For the purpose of comparison, we adopt the same convention for shotnoise as in [294], namely

$$\bar{n} = b_1^2 n_0 (1+z)^\alpha, \quad (6.33)$$

⁸Here, and in the next section, we make a particular choice of redshift bins. Larger redshift bins imply that we can include more configurations of the bispectrum in the analysis, in particular more configurations in the squeezed limit. At the same time, we fix the redshift of the bin to be the mean redshift, therefore, larger redshift bins imply a smaller maximum redshift. Therefore, the choice of binning might affect the final result.

⁹This is a little less than 45 bins over the full range, corresponding to $O(3000)$ triangles.

where we correct for the galaxy bias $b_1 = 2$, since the shotnoise in [294] applies to galaxies, while here it has been translated to the dark matter density field. We choose $n_0 = 10^{-3}h^3\text{Mpc}^{-3}$ and $\alpha = -1$.

- When we marginalize, we assume for each EFT coefficient a Gaussian prior with $\sigma = 10$, except for the EFT parameter ξ for which we take $\sigma = 1$.
- For B^{er} we consider both B_{332} and the ansatz E_b given in [294] (see (6.27)).

The results for $\sigma(f_{NL})$ are shown in Figure 6.3. We show the effects of the ansatz for B^{er} , shotnoise, and marginalization over the EFT parameters.

We can compare our results with those found in [294] by looking at the unmarginalized results, using their ansatz E_b for B^{er} . Thus, we should compare our dashed green lines with their dotted red lines in the last columns of their Figure 6 and 7. We indeed find a reasonably good agreement, given the fact that our analyses are not identical (different sky coverage, redshift bins and cosmological parameters, and moreover the translation of their shotnoise to ours is not perfect as we only took into account b_1). This check confirms that our code runs as expected.

Let us now study the effect of the EFT parameters. The solid lines in Figure 6.3 correspond to marginalization over the EFT parameters, where we assume a Gaussian prior with $\sigma = 10$ for each EFT coefficient, except for the EFT parameter ξ , for which we take $\sigma = 1$. We see that *the results for local PNG are almost unaffected by marginalizing over the EFT parameters*. The constraints on equilateral PNG weaken by a factor of about two, however.

Using B_{332} as an ansatz for B^{er} , we find slightly more optimistic results for local PNG, and slightly worse results for equilateral PNG, as compared with [294]. This can be understood from the comparison between B_{332} and E_b , shown in appendix 6.B.5. Local PNG peaks in the squeezed limit, and B_{332} is more optimistic than E_b in this configuration. On the other hand, equilateral PNG peaks in the equilateral configuration, and in this configuration E_b is more optimistic.

If we neglect shotnoise, we find that the differences at low redshifts are even bigger for the two ansätze. The difference is largest for local PNG, since B_{332} is an order of magnitude bigger than E_b in the squeezed configuration, whereas the difference in the equilateral configuration is only of the order of a few. However, at higher z_{max} the differences disappear. This might seem a bit strange at first. However, we should stress that what we find here is not the true result in case of zero shotnoise. It turns out that with no shotnoise, we can always gain information in the ultra squeezed limit, and at higher

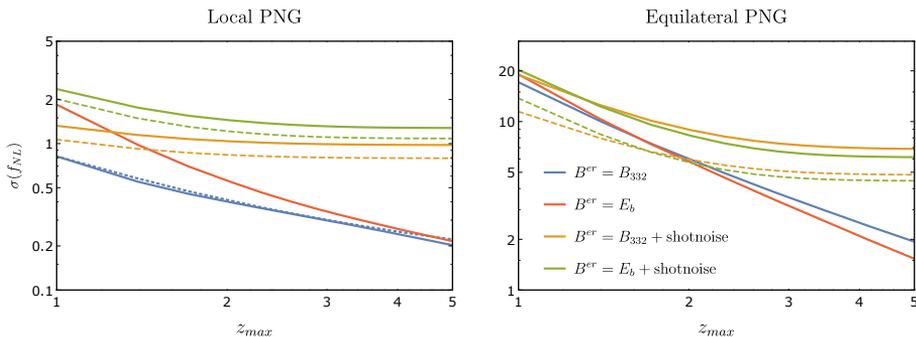


Figure 6.3: In these plots we show $\sigma(f_{NL})$ as function of maximum redshift z_{\max} for local PNG (left) and equilateral PNG (right). We include shotnoise with the same specifications as [294] (orange and green lines). For the green lines we used E_b as ansatz for B_{2L}^G whereas for the other lines we used B_{332} . The blue solid line shows the result without shotnoise. The solid lines show the marginalized results assuming Gaussian priors for the EFT parameters. The dashed lines correspond to the unmarginalized result with the same color. Finally, the dotted blue line in the left panel corresponds to the curve $\sigma(z_1)/\sqrt{N}$ with N the number of redshift bins.

redshift we can actually go to higher k than our choice $k_{\max} = 1h\text{Mpc}^{-1}$. Since $\sigma(f_{NL})$ does not freeze before we reach k_{\max} , this explains why the red and blue curves can get close for high z_{\max} . In appendix 6.B.5, we show these statements explicitly.

For comparison, note that if we were to always gain signal up to the same k_{\max} in each redshift bin, we would find roughly the same $\sigma(f_{NL})$ for each redshift bin¹⁰. Therefore, combining all the redshift bins, we should find a $1/\sqrt{N_{\text{bins}}}$ behavior if we neglect shotnoise. The dotdashed blue line in the figure corresponds to $\sigma(z_1)/\sqrt{N_{\text{bins}}}$, which indeed resembles the blue solid line quite well. This provides another indication that we can go up to higher k_{\max} . Interestingly, we find that we can also go to much smaller scales for equilateral PNG than suggested by scaling estimates for k_{\max} (see for instance [294]). The squeezed limit allows us to extract more information, also for equilateral PNG.

If we include shotnoise, it correctly cuts off the signal before we reach k_{\max} , so these results are reliable. However, one should keep in mind that, for more optimistic galaxy number densities, we might still extract more information from the ultra-squeezed limit. For the number densities we consider at the

¹⁰To good approximation: the entry of the Fisher matrix, corresponding to f_{NL} , will scale as $F \sim D(a)^6/D(a)^6 \sim 1$, if we neglect the loop corrections to B_{NG} . Then, forgetting about the marginalization over the EFT parameters, we find the same $\sigma(f_{NL})$ in each redshift bin, since we took the bins so that they have the same volume.

higher redshifts, shotnoise is the dominant source of noise. This is why we do include shotnoise in our analysis.

Current and upcoming surveys (comparison with Tellarini et al. 2016)

Next, we compare with [287], using the specifications of the surveys Euclid [150], BOSS [301], eBOSS [152] and DESI [151]. We have to consider the emission line galaxies (ELG), the luminous red galaxies (LRG) and quasars (QSO) separately, as they are measured at different redshifts, and have different number densities and bias coefficients. The specifications and assumptions are as follows.

- For the precise number densities and bias coefficients as function of redshift we refer to appendix D of [287]. Moreover, one can find here the fraction of sky covered by each survey.
- As before, we divide each survey in equal sized redshift bins. The boundaries of all the redshift bins are given by

eBOSS (ELG) : {0.6, 0.8, 0.95, 1.09, 1.21}	$V_{\text{bin}} = 2.8(h^{-1}\text{Gpc})^3,$
DESI (ELG) : {0.1, 0.6, 0.79, 0.94, 1.07, 1.19, 1.3, 1.4, 1.5, 1.59, 1.69, 1.78}	$V_{\text{bin}} = 5.4(h^{-1}\text{Gpc})^3,$
Euclid (ELG) : {0.6, 1., 1.28, 1.53, 1.75, 1.97}	$V_{\text{bin}} = 14.0(h^{-1}\text{Gpc})^3,$
eBOSS (LRG) : {0.6, 0.75, 0.87, 0.98}	$V_{\text{bin}} = 2.0(h^{-1}\text{Gpc})^3,$
DESI (LRG) : {0.1, 0.6, 0.79, 0.94, 1.07}	$V_{\text{bin}} = 5.4(h^{-1}\text{Gpc})^3,$
BOSS (LRG) : {0., 0.4, 0.52, 0.61, 0.68, 0.75, 0.8}	$V_{\text{bin}} = 1.3(h^{-1}\text{Gpc})^3,$
eBOSS (QSO) : {0.6, 1., 1.28, 1.53, 1.75, 1.97, 2.17}	$V_{\text{bin}} = 6.6(h^{-1}\text{Gpc})^3,$
DESI (QSO) : {0.1, 0.8, 1.08, 1.31, 1.51, 1.7, 1.89}	$V_{\text{bin}} = 11.0(h^{-1}\text{Gpc})^3.$

- We use k_{min} determined by the volume of each bin. Moreover, we choose the same binning of the k -range as in Section 6.4.1.
- The ansatz for shotnoise is now $\bar{n}(z) = b_1^2(z)n(z)$, where we correct for galaxy bias, similar as before.
- When we marginalize, we take the same prior as before. We assume for each EFT coefficient a Gaussian prior with $\sigma = 10$, except for the EFT parameter ξ for which we take $\sigma = 1$.
- As ansatz for the higher order corrections we use $B^{\text{er}} = B_{332}$.

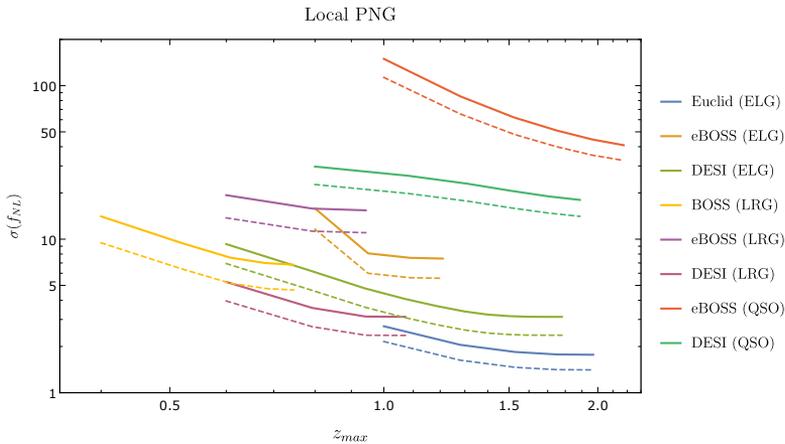


Figure 6.4: We show $\sigma(f_{NL})$ as function of z_{\max} for local PNG. We use the specification from Euclid (blue), BOSS (yellow), eBOSS (orange, purple and red) and DESI (green/yellow, pink and green). We show both the marginalized (solid lines) and unmarginalized results (dashed lines).

The results for local PNG are shown in Figure 6.4. We plot $\sigma(f_{NL})$ as function of z_{\max} for the four surveys. We show both the marginalized and unmarginalized results. After combining the different galaxy catalogs of a survey, we get $\sigma(f_{NL})$ for each survey, as summarized in Table 6.1a for local PNG and in Tables 6.1b and 6.1c for equilateral and quasi-single field PNG respectively.

We compare our results with [287] by looking at our unmarginalized results in Table 6.1a and their results in the last column in Table 1 of their paper. We find much weaker constraints, varying from 4 to 8 times smaller. This can be explained by the fact that we account for the theoretical error, which freezes the errorbars. Therefore, including the theoretical error gives rise to more conservative constraints. Moreover, scale-dependent bias could actually help us constrain local PNG also in the bispectrum. In fact, redoing the analysis for Euclid up to $k_{\max}(z) = 0.1D(z)$, as used in [287], with the same specifications, except that we ignore the theoretical error, gives $\sigma(f_{NL})$ equal to 0.57, 0.71 and 1.3 (unmarginalized, including and neglecting priors, respectively). This is roughly a factor three improvement from the results in Table 6.1a. The fact that this still does not challenge the results from [287] seems to indicate that scale-dependent bias helps to improve the constraints on primordial non-Gaussianity.

From the combined catalogs, we ultimately find $\sigma(f_{NL}^{\text{loc}}) = 1.8$, $\sigma(f_{NL}^{\text{eq}}) = 11.4$ and $\sigma(f_{NL}^{\text{qsf}}) = 8.9$, with priors on the EFT parameters, assuming the surveys are not independent. These results do not change dramatically if we do not put priors on the EFT parameters. If the surveys are independent, the

$\sigma(f_{NL}^{\text{loc}})$	unmarg.	with prior	no prior
BOSS	4.67	6.81	17.3
eBOSS	4.91	6.6	14.15
Euclid	1.41	1.77	3.66
DESI	1.66	2.18	4.68

(a) Local PNG

$\sigma(f_{NL}^{\text{eq}})$	unmarg.	with prior	no prior
BOSS	16.89	29.86	37.99
eBOSS	17.25	26.88	33.4
Euclid	7.46	11.37	13.66
DESI	7.18	11.4	13.48

(b) Equilateral PNG

$\sigma(f_{NL}^{\text{qsfl}})$	unmarg.	with prior	no prior
BOSS	12.57	23.65	27.26
eBOSS	13.1	21.43	23.49
Euclid	5.52	8.92	9.74
DESI	5.37	8.98	9.66

(c) Quasi-single-field PNG

Table 6.1: The final $\sigma(f_{NL})$ for local, equilateral and quasi-single field PNG for each survey, combining all expected galaxy catalogues. For the marginalized $\sigma(f_{NL})$, we put a Gaussian prior on each EFT coefficient with $\sigma = 10$, except for the EFT parameter ξ , for which we take $\sigma = 1$. In the last row, we also show the marginalized results, without prior on the EFT coefficients.

constraints improve approximately with a factor $1/\sqrt{2}$ upon combining Euclid and DESI.

6.4.2 Correlation coefficients

To gain intuition on how much the EFT parameters affect the constraints on f_{NL} for local, equilateral and quasi-single field PNG, we compute the correlation coefficients between parameters θ_i and θ_j . These are defined as

$$r_{ij} = \frac{F_{ij}^{-1}}{\sqrt{F_{ii}^{-1}F_{jj}^{-1}}}.$$

The correlation coefficient takes a value between 1 (perfectly correlated) and -1 (perfectly anti-correlated). In particular, the parameters are perfectly correlated with themselves. In Figure 6.5, we plot the absolute value of the correlation coefficients for each pair of parameters. We make the following assumptions

- We use the same binning in k -space as in Section 6.4.1.
- We use the redshift binning and shotnoise from Euclid, as given in Section 6.4.1.
- As ansatz for the theoretical error we use $B^{\text{er}} = B_{332}$.
- We do not marginalize over the EFT parameters. The marginalized results are quoted in the text below.

We find that the groups of parameters $\{\xi, \epsilon_1, \epsilon_2, \epsilon_3\}$ and $\{\gamma, \gamma_1, \gamma_2\}$ have strong correlation among themselves. The correlation between f_{NL} and the other parameters is, however, small.

For local PNG, we find f_{NL} is mainly correlated with γ, γ_1 and ϵ_3 , with correlation coefficients 0.44, -0.43 and 0.29 respectively. The other correlation coefficients are in absolute value smaller than 0.2. If we include a Gaussian prior on the EFT parameters, with the same variances as before, the correlation coefficients become 0.12, 0.08 and 0.14.

In case of equilateral PNG, we find f_{NL} has appreciable correlation with $\xi, \epsilon_1, \epsilon_3$ and γ_2 , with correlation coefficients 0.43, $-0.39, -0.47$ and -0.27 respectively. Including priors on the EFT parameters, we find they become 0.14, $-0.05, -0.30$ and approximately zero. This could motivate further study on the Gaussian EFT coefficients, see for instance [302]. It is surprising that f_{NL}^{eq} is not extremely degenerate with ξ , since the latter comes with an additional k^2 scaling, similar to equilateral non-Gaussianity. It turns out, however, that

the full shapes are sufficiently distinct. This will make it easier to constrain equilateral PNG from the bispectrum than naively thought. Then, for quasi-single-field PNG, we find that f_{NL} is mostly correlated with ϵ_3 and γ , with correlation coefficients -0.26 and 0.31 . Including the priors, they reduce to -0.18 and approximately zero.

Summarizing, although the ignorance about EFT parameters does affect the final result, for reasonable priors this is only a small effect, especially for local PNG. This indeed agrees with what is seen in Figures 6.3 and 6.4, and Table 6.1a.

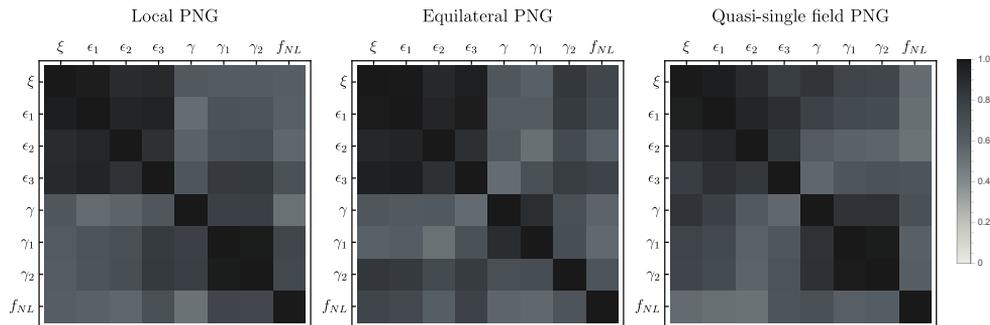


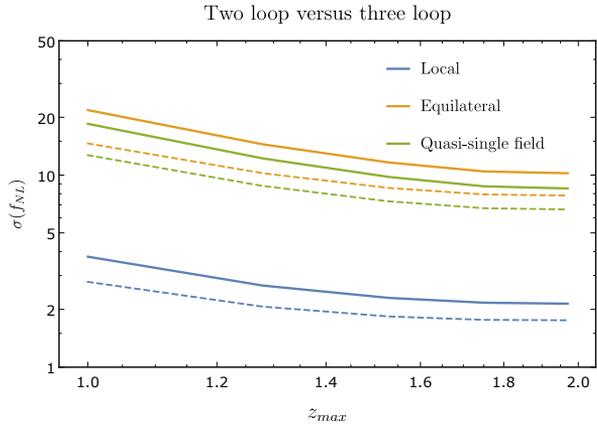
Figure 6.5: In these plots we show the correlation coefficients r_{ij} for each pair of theoretical parameters. We include shotnoise with the same specifications as Euclid and included all information up to redshift $z_{\max} = 2$. A value of 1 (black) corresponds to perfectly correlated or anti-correlated. A value of 0 (white) corresponds to no correlation.

6.4.3 Higher loop corrections

We can ask ourselves whether it is useful to compute the bispectrum up to two loops in gravitational non-linearities. Note that our analysis does not depend on the actual value of the two-loop bispectrum, as there is no theoretical parameter in front. This means we can simply assume that we have computed all diagrams, neglect the counterterms, and assume the theoretical error is given by the SPT three loop bispectrum B_{3L}^G . Again, we do not know what it is, so we have to make an ansatz for it. Here we use the ansatz for the higher loop corrections from [294], given in (6.27), since it is easy to compute¹¹. For

¹¹An alternative - more in line with our two loop ansatz - would be to compute the reducible diagram of B_{433} as order of magnitude estimate of B_{3L}^G . However, we point out that using only one diagram is dangerous. For instance, in [183] we considered only one of the two reducible two loop diagrams in our qualitative analysis, namely B_{332}^I . In the squeezed limit, the two loop contribution turns out to be much larger if we include B_{332}^I .

Figure 6.6: In this plot, we show $\sigma(f_{NL})$ as function of z_{\max} for local (blue), equilateral (orange) and quasi-single field (green) PNG. We use both B_{2L} (solid lines) and B_{3L} (dashed lines) as order of magnitude estimation for the theoretical error. We use the specifications of Euclid in the analysis.



a rough estimate this should suffice. We estimate B_{3L}^G using scaling universes [184]. We choose¹² E_b with $n = -1.5$, $k_{NL} = 0.5 h\text{Mpc}^{-1}$ and $l = 3$ (see (6.27)). Using the specifications from Euclid, we perform the Fisher analysis with both B_{332} and B_{3L} as order of magnitude estimates for the noise. We collect the result in Figure 6.6. This shows that the constraints would improve if one computed the two-loop corrections to the bispectrum. The precise values are given in Table 6.2, where we consider all surveys again. The tightest constraints on local, equilateral and quasi-single field PNG improve with a factor 1.2, 1.3 and 1.3 respectively with this particular choice for B_{3L}^G . If it turns out we can get constraints on PNG close to the theoretical benchmarks, it would then be worth computing the two-loop corrections. It might be time consuming, but otherwise much cheaper than doubling the survey volume.

¹²We choose a larger k_{NL} than for the two loop estimate, since this scale determines when the three loop correction equals the lower order corrections. Since we are doing a perturbative expansion, we assume this happens at a smaller scale than when the two loop correction becomes equal to its lower order corrections. Moreover, each loop will scale as k/k_{NL} to the power $3 + n$, where n will be of order of the scaling of the power spectrum at the non-linear scale k_{NL} . The power spectrum is steeper on smaller scales, therefore we take a more negative value for n . Each loop has this scaling, so we have to take $l = 3$, in case of three loops.

$\sigma(f_{NL}^{\text{loc}})$	2 loop	3 loop	$\sigma(f_{NL}^{\text{eq}})$	2 loop	3 loop
BOSS	8.73	6.05	BOSS	27.8	19.14
eBOSS	7.12	6.07	eBOSS	22.99	18.44
Euclid	2.14	1.75	Euclid	10.22	7.83
DESI	2.62	2.09	DESI	10.2	7.81

(a) Local PNG

$\sigma(f_{NL}^{\text{qsf}})$	2 loop	3 loop
BOSS	23.66	16.65
eBOSS	19.15	15.59
Euclid	8.52	6.62
DESI	8.46	6.6

(b) Equilateral PNG

(c) Quasi-single-field PNG

Table 6.2: The final $\sigma(f_{NL})$ for equilateral and quasi-single-field PNG (left and right), for each survey, combining all expected galaxy catalogues. For the marginalized $\sigma(f_{NL})$, we put a Gaussian prior on each EFT coefficient with $\sigma = 10$, except for the EFT parameter ξ , for which we take $\sigma = 1$. In the last row, we also show the marginalized results without prior on the EFT coefficients.

6.4.4 EFT of LSS versus SPT

In the EFT of LSS we are forced to include free parameters over which we have to marginalize. Above, we saw that this marginalization weakens the constraints on f_{NL} , be it only mildly. One might therefore wonder how much the improvement actually is over a more conservative approach, in which one uses only SPT results for B^{th} and moves all other gravitational contributions to the theoretical error. In this section we confirm that the EFT approach always performs sizably better. We consider a couple of options, for different choices of B^{th} and B^{er} , and compute the constraints. We use the specifications from Euclid as given in the previous section.

Table 6.3 shows our results, which include the usual Gaussian priors for the EFT parameters whenever they are included in B^{th} . The second and third columns give the theoretical description of the bispectrum B^{th} and what we consider to be the unknown B^{er} . For the latter, we sum the absolute values of all contributions indicated in the table. The EFT contributions, except for ξ , are multiplied by a factor 10, consistent with the priors we chose when we included them in B^{th} . The first row of Table 6.3 corresponds to the results we find in the ‘with prior’ columns of the ‘Euclid’ rows of Tables 6.1a, 6.1b and 6.1c.

We find that including the non-Gaussian counterterms does not improve the bounds on f_{NL} . This could have been anticipated from the qualitative

Approach	$B^{\text{th}} = B_0^G + \dots$	$B^{\text{er}} = B_{332} + \dots$	$\sigma(f_{NL}^{\text{loc}})$	$\sigma(f_{NL}^{\text{eq}})$	$\sigma(f_{NL}^{\text{qsf}})$
EFT (G+NG)	$+B_0^{NG} + B_{EFT}^G + B_{EFT}^{NG}$		1.77	11.37	8.92
EFT G+SPT NG	$+B_0^{NG} + B_{EFT}^G$	$B_{31}^G + B_{EFT}^{NG}$	1.78	11.37	8.92
SPT (G+NG)	$+B_0^{NG}$	$+B_{EFT}^{NG} + B_{EFT}^G$	6.11	27.61	21.76
SPT (G+NG tree)	$+B_{\text{tree}}^{NG}$	$+B_{EFT}^{NG} + B_{EFT}^G + B_{1L}^{NG}$	7.17	30.58	24.23

Table 6.3: We show the constraints on primordial non-Gaussianity of the local, equilateral and quasi-single-field type (last three columns). In the first row, we use the EFToLSS for both the Gaussian and non-Gaussian part of the bispectrum (‘EFT (G+NG)’). In the second row, we only use the EFT for the Gaussian part of the bispectrum, and include the non-Gaussian counterterms in the higher order corrections (‘EFT G + SPT NG’). Then, in the third row, we use only the SPT for describing the bispectrum (‘SPT (G+NG)’). In the last row, we only include the tree level non-Gaussian contribution to the bispectrum (‘SPT (G+NG tree)’). The second and third column denote all the contributions to the theoretical description of the bispectrum B^{th} , and the higher order corrections B^{er} respectively.

results in Figure 5 - 7 of [183]. We see that the counterterms are negligible in many configurations. Apparently, they are negligible in most configurations. However, using the EFT for the Gaussian part of the bispectrum performs significantly improves the results compared with just the SPT predictions. We find that the *EFT of LSS improves the constraints on PNG approximately by a factor 3*. Finally, neglecting the one-loop non-Gaussian contribution to the bispectrum makes only about a 10% difference. This is consistent with the observation that the non-Gaussian counterterms are not very important, as the non-Gaussian one loop correction itself is not very relevant.

6.5 Discussion and Outlook

In this work, we have presented how the EFT of LSS helps us improve the constraints on primordial non-Gaussianities (PNG), using the matter bispectrum as observable. We have accounted for intrinsic theoretical uncertainties in the perturbative description, and studied in details their modeling in a Fisher forecast.

Our main results are given in Table 6.3. The forecasted values for $\sigma(f_{NL})$ for the local, equilateral and quasi-single field types of PNG are presented. Moreover, we show that the EFT approach improves the constraints on PNG by almost a factor 3 with respect to the results from SPT.

Limitations Let us first discuss the limitations of these results. We would like to compare the constraints we find with theoretically interesting benchmarks and constraints coming from the CMB. However, we should be careful in making a direct comparison, as there are other sources of non-linearities and noise that we have not accounted for in our analysis. First, we have modeled the matter bispectrum. To relate it to the observed galaxy bispectrum, we have to include galaxy bias and redshift space distortions. These introduce new uncertainties, leading to worse constraints. However, considering the results found in [130, 287], scale-dependent bias might actually *improve* the constraints on (only) local PNG, since it enhances the non-Gaussian signal in the bispectrum. Second, except for shotnoise, we neglected all observational sources of noise. Survey geometry and survey mask may increase the errorbars as well. For instance the authors of [288] found that the errorbars increased by a factor of 4-5. Errors in determining the redshift of galaxies are another source of observational noise. Third, we made some simplifications in the Fisher analysis itself, such as neglecting the covariance between different points of the bispectrum. Combining this with the covariance induced by the survey geometry could further increase the errorbars by a factor of 8 [303].

Improvements On the other hand, there are also ways the constraints could be improved. First, we have used the specifications of Euclid to get a reasonable estimate for the limitations due to shotnoise. This determined our final forecasted result for $\sigma(f_{NL})$. It might well be possible, in a more futuristic survey, to optimize the number densities of galaxies and redshift range to be more suitable for constraining PNG (see for instance [130]). Moreover, we should perform a joint analysis of all large scale structure surveys. We have assumed for simplicity that we can do as well as the single best survey, which turned out to be Euclid for the four surveys we considered. In principle, we can do better if the surveys are not all precisely overlapping. Similarly, we should combine the results from different observables. For instance, we should perform a joint analysis of the power spectrum and bispectrum. This could improve the results by a factor of 2 for local PNG. Combining the results found in [287], and using the multitracer technique proposed in [304] instead, which could improve upon the constraints from the power spectrum by a factor of about 7. In addition, the trispectrum might turn out to be an important probe for non-Gaussianity, since linear theory works for a larger range of scales compared to the bispectrum [305]. The one loop corrections to the trispectrum in the EFT of LSS have recently been computed in [306]. Last, we divided the full redshift range in smaller redshift bins, and only considered correlations within each redshift bin. If we also include correlations among galaxies separated by a larger distance along the line of sight, we might extract more information

from a given survey. Finally, our focus here was on near future galaxy surveys, but of course our results will be relevant in the future also for 21 cm survey (see e.g. [307]).

One of our main results is that the EFT approach helps constraining PNG. The improvement comes completely from the EFT corrections to the late time gravitational non-linear evolution of matter. Both the SPT loops and EFT corrections to the primordial non-Gaussian signal, discussed in [183], do not help much improving the constraints.

Comparing our results with the theoretically interesting benchmark $\sigma(f_{NL}^{\text{eq}}) \sim 1$, we see that it does not look promising for equilateral PNG. Even with zero shotnoise, as we can see in Figure 6.3, we barely touch the theoretical targets. Our lack of understanding of matter non-linearities is already an important obstacle to reach $\sigma(f_{NL}^{\text{eq}}) \sim 1$. The same applies to quasi-single field PNG. Additional sources of non-linearities such as bias and redshift space distortions will make things worse. On the other hand, for local primordial non-Gaussianity, things look more promising. Matter non-linearities can be modeled well enough to get close to $\sigma(f_{NL}^{\text{loc}}) \sim 1$ from large scale structure experiments.

We can ask whether N -body simulations can help reaching better constraints on primordial non-Gaussianity. As pointed out in [294], using end to end simulations, without any perturbative input, will most likely be insufficient to reach $\sigma(f_{NL}^{\text{eq}}) \sim 1$. The reason is that simulations do not solve the exact problem but make a series of approximations, such as for example the particle mesh and tree approaches to solve Poisson equation, finite size effects and approximate initial conditions. Currently, simulations reach approximately 1% precision [308, 309]. Heuristically, looking at our Figure 6.1, we see that the PNG signal we are trying to extract is much smaller than that, so large improvements in the precision of simulations are needed. Alternatively, one can use N -body simulations to determine the EFT parameters¹³. We can then look directly at the unmarginalized columns in Table 6.1c. We see that, even in the very optimistic case that all relevant EFT parameters at one loop are fixed, $\sigma(f_{NL}^{\text{eq}})$ still remains around 7.

¹³In fact, in our analysis, we assumed that a one EFT parameter, ξ , was fixed by fitting the power spectrum to simulations.

Theoretical error Another goal of our paper is to clarify some aspects of the modeling of theoretical uncertainties in forecasting observational constraints, and, eventually, in analyzing data (see Section 6.3.5 and Appendix 6.B). We introduced the concept of correlation length in Section 6.3.4, along the lines of [294]. In Appendix 6.B, we argued that the choice of correlation length in integrating out the theoretical error is subtle and no “right” choice can be established a priori. However, in our particular analysis, we hardly find any dependence on the correlation length (see Figure 6.7). In future studies, with different observables and different perturbative approaches, we believe that an analysis on the choice of correlation length should be always performed.

In Figure 6.2 we have seen that assuming the wrong shape for the theoretical error can lead to biased results in a χ^2 -analysis. Therefore, if we want to fit to data, we need good estimates for the higher order corrections. For instance by using estimates from N-body simulations, or alternatively, by computing additional two-loop diagrams.

Outlook Our work can be extended and improved in different ways.

- Instead of dividing the survey volume in redshift bins and only consider correlations within these bins, it would be interesting to see how much we gain including all possible cross-correlations across redshift bins.
- It would be interesting to perform a similar Fisher analysis with an updated study of covariance effects due to geometry, masking and non-Gaussian gravitational evolution.
- We should join all forces. It would be interesting to do a joint analysis of multiple observations, such as the CMB, LSS surveys and the 21 cm observations. Moreover, all the different LSS surveys should be combined to have maximum constraining power. Furthermore, the results from the power spectrum, bispectrum and trispectrum should be combined too. Finally, on the theory side, one should also try to use results from N-body simulations as soon as our perturbative description starts to break down, i.e. when the theoretical error becomes dominant.

6.A Explicit Results for the Bispectrum

The purpose of this appendix is to provide, and to some extent clarify, explicit expressions for the bispectrum in the presence of primordial non-Gaussianity. This is essentially a summary of [183], so we refer the reader to that work for a thorough discussion and explanation of these results. We adopt the same notation, which we summarize in Appendix 6.D.

6.A.1 Perturbation Theory in the EFT of LSS

In the EFToLSS, the equations of motion for the density contrast δ and the velocity divergence $\theta = \partial_i v^i$ on large scales are

$$\delta_\tau \delta + \theta = \mathcal{S}_\alpha, \quad (6.34a)$$

$$(\delta_\tau + \mathcal{H})\theta + \frac{3}{2}\Omega_m \mathcal{H}^2 \delta = \mathcal{S}_\beta + \tau_\theta. \quad (6.34b)$$

Here the source terms $\mathcal{S}_{\alpha,\beta}$ are the standard nonlinear terms in the Euler equations, which, in Fourier space, are given by the following convolutions,

$$\mathcal{S}_\alpha(\mathbf{k}, \tau) \equiv - \int_{\mathbf{p}} \alpha(\mathbf{p}, \mathbf{k} - \mathbf{p}) \theta(\mathbf{p}, \tau) \delta(\mathbf{k} - \mathbf{p}, \tau), \quad (6.35a)$$

$$\mathcal{S}_\beta(\mathbf{k}, \tau) \equiv - \int_{\mathbf{p}} \beta(\mathbf{p}, \mathbf{k} - \mathbf{p}) \theta(\mathbf{p}, \tau) \theta(\mathbf{k} - \mathbf{p}, \tau), \quad (6.35b)$$

with

$$\alpha(\mathbf{k}_1, \mathbf{k}_2) \equiv \frac{\mathbf{k}_1 \cdot (\mathbf{k}_1 + \mathbf{k}_2)}{\mathbf{k}_1^2} \quad \text{and} \quad \beta(\mathbf{k}_1, \mathbf{k}_2) \equiv \frac{(\mathbf{k}_1 + \mathbf{k}_2)^2}{2} \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{\mathbf{k}_1^2 \mathbf{k}_2^2}. \quad (6.36)$$

Clearly, we neglected large scale vorticity and large scale velocity dispersion in (6.34). However, the backreaction from unknown short scale physics is taken into account through the effective stress tensor τ_θ . A complete description and motivation of this term in the presence of primordial non-Gaussianity was the main purpose of [183]. Here we just quote the leading contributions for the types of non-Gaussianities we consider, to first order in f_{NL} ,

$$\begin{aligned} \tau_\theta = & -d^2 \Delta \delta - e_1 \Delta(\delta^2) - e_2 \Delta(s^2) - e_3 \partial_i (s^{ij} \partial_j \delta) \\ & - f_{\text{NL}} [g(\Delta \Psi - \partial_i (\delta \partial^i \Psi)) + g_1 \Delta(\Psi \delta) + g_2 \partial_i \partial_j (\Psi s^{ij})], \end{aligned} \quad (6.37)$$

where Δ denotes the Laplacian, and the coefficients in this expression are functions of time only. The equations are formally solved using a Green's function method

$$\delta(\mathbf{k}, a) = D_1(a) \delta_1(\mathbf{k}) + \int_{a_{\text{in}}}^a G_\delta(a, a') [\mathcal{S}_\beta + \tau_\theta - \mathcal{H} \partial_a (a \mathcal{S}_\alpha)]. \quad (6.38)$$

Here $\delta_1(\mathbf{k})$ is the growing mode initial condition, and $D_1(a)$ is the linear growth factor

$$D_1(a) = \frac{5}{2} \mathcal{H}_0^2 \Omega_m^0 \frac{\mathcal{H}}{a} \int_{a_{in}}^a \frac{da'}{\mathcal{H}^3(a')}, \quad (6.39)$$

which in Einstein-de Sitter reduces to $D_1(a) = a/a_{in}$, where a is the scale factor. This equation can be solved perturbatively in terms of the linear solution, yielding

$$\delta(\mathbf{k}, a) = \sum_{n=1}^{\infty} \delta_{(n)}^{\text{SPT}}(\mathbf{k}, a) + \delta_{(n)}^c(\mathbf{k}, a), \quad (6.40)$$

where $\delta_{(n)}^{\text{SPT}}$ are the standard perturbation theory (SPT) terms (see [149]), sourced solely by $\mathcal{S}_{\alpha,\beta}$:

$$\delta_{(n)}^{\text{SPT}}(\mathbf{k}, a) \approx D_1^n(a) \int_{\mathbf{k}_1} \dots \int_{\mathbf{k}_n} (2\pi)^3 \delta_D(\mathbf{k} - \mathbf{k}_{1\dots n}) F_n(\mathbf{k}_1, \dots, \mathbf{k}_n) \delta_1(\mathbf{k}_1) \dots \delta_1(\mathbf{k}_n),$$

and $\delta_{(n)}^c$ is the ‘counterterm’ contribution, i.e. the terms proportional to one of the free parameters in (6.37), which we write as

$$\begin{aligned} \delta_{(n)}^c(\mathbf{k}, a) = & \int_{\mathbf{k}_1} \dots \int_{\mathbf{k}_n} (2\pi)^3 \delta_D(\mathbf{k} - \mathbf{k}_{1\dots n}) F_n^c(\mathbf{k}_1, \dots, \mathbf{k}_n|a) \delta_{(1)}(\mathbf{k}_1, a) \dots \delta_{(1)}(\mathbf{k}_n, a) \\ & + f_{\text{NL}} \int_{\mathbf{k}_1} \dots \int_{\mathbf{k}_n} (2\pi)^3 \delta_D(\mathbf{k} - \mathbf{k}_{1\dots n}) H_n^c(\mathbf{k}_1, \dots, \mathbf{k}_n|a) \psi(\mathbf{k}_1) \dots \delta_{(1)}(\mathbf{k}_n, a). \end{aligned} \quad (6.41)$$

Note that we have not specified the time dependence of the free parameters yet, which is why the counterterm kernels are still time dependent. It turns out that at first order in the perturbations, this is not really an issue, as the time dependence is just given by an integral of the Green’s function over some unknown function of time, which yields some other unknown function. However, in order to get the momentum dependence right at second order, we have to make an assumption about the first order terms. A convenient ansatz is

$$d^2(a) = [\mathcal{H}(a)f(a)]^2 [D_1(a)]^{m_d+1} \bar{d}^2, \quad (6.42)$$

$$g(a) = [\mathcal{H}(a)f(a)]^2 [D_1(a)]^{m_g+1} \bar{g}, \quad (6.43)$$

which in Einstein-de Sitter reduces to $(d^2, g) \propto a^{m_d}$. Then the expressions for the counterterm kernels at second order are as follows. We split up the kernels in the following way

$$F_2^c(\mathbf{k}_1, \mathbf{k}_2|a) = F_2^{\tau}(\mathbf{k}_1, \mathbf{k}_2|a) + F_2^{\alpha\beta}(\mathbf{k}_1, \mathbf{k}_2|a) + F_2^{\delta}(\mathbf{k}_1, \mathbf{k}_2|a), \quad (6.44)$$

where the terms coming from the new counterterms at second order are

$$F_2^T(\mathbf{k}_1, \mathbf{k}_2|a) = - \sum_{i=1}^3 \epsilon_i(a) E_i(\mathbf{k}_1, \mathbf{k}_2), \text{ with} \quad (6.45)$$

$$E_1(\mathbf{k}_1, \mathbf{k}_2) \equiv \mathbf{k}_{12}^2, \quad (6.46)$$

$$E_2(\mathbf{k}_1, \mathbf{k}_2) \equiv \mathbf{k}_{12}^2 \left[\frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{k_1^2 k_2^2} - \frac{1}{3} \right], \quad (6.47)$$

$$E_3(\mathbf{k}_1, \mathbf{k}_2) \equiv \left[-\frac{1}{6} \mathbf{k}_{12}^2 + \frac{1}{2} \mathbf{k}_1 \cdot \mathbf{k}_2 \left[\frac{\mathbf{k}_{12} \cdot \mathbf{k}_2}{k_2^2} + \frac{\mathbf{k}_{12} \cdot \mathbf{k}_1}{k_1^2} \right] \right], \text{ and} \quad (6.48)$$

$$\epsilon_i(a) \equiv - \frac{1}{[D_1(a)]^2} \int_{a_{in}}^a da' G_\delta(a, a') [D_1(a')]^2 e_i(a'). \quad (6.49)$$

Furthermore, the terms coming from plugging the lowest order counterterm back into the equations of motion are

$$F_2^{\alpha\beta}(\mathbf{k}_1, \mathbf{k}_2|a) = -\xi(a) E_{\alpha\beta}(\mathbf{k}_1, \mathbf{k}_2), \text{ with} \quad (6.50)$$

$$E_{\alpha\beta}(\mathbf{k}_1, \mathbf{k}_2) \equiv \frac{1}{2m_d + 9} \left[2\beta(\mathbf{k}_1, \mathbf{k}_2)(k_1^2 + k_2^2) \right. \quad (6.51)$$

$$\left. + \frac{2m_d + 7}{2(m_d + 2)} \left(\alpha(\mathbf{k}_1, \mathbf{k}_2)(k_2^2 + (m_d + 2)k_1^2) + \{1 \leftrightarrow 2\} \right) \right], \quad (6.52)$$

$$\xi(a) = \frac{2}{(m_d + 1)(2m_d + 7)} [D_1(a)]^{m_d+1} \bar{d}^2. \quad (6.53)$$

Finally, the term from the second order contribution to the density is

$$F_2^\delta(\mathbf{k}_1, \mathbf{k}_2|a) = -\xi(a) E_\delta(\mathbf{k}_1, \mathbf{k}_2), \text{ with} \quad (6.54)$$

$$E_\delta(\mathbf{k}_1, \mathbf{k}_2) = \frac{(m_d + 1)(2m_d + 7)}{(m_d + 2)(2m_d + 9)} \mathbf{k}_{12}^2 F_2(\mathbf{k}_1, \mathbf{k}_2). \quad (6.55)$$

Similarly, for the non-Gaussian kernels we have

$$H_2^c(\mathbf{k}_1, \mathbf{k}_2|a) = H_2^T(\mathbf{k}_1, \mathbf{k}_2|a) + H_2^{\alpha\beta}(\mathbf{k}_1, \mathbf{k}_2|a) + H_2^\Psi(\mathbf{k}_1, \mathbf{k}_2|a), \quad (6.56)$$

where second order counterterm contribution is

$$H_2^T(\mathbf{k}_1, \mathbf{k}_2|a) = - \sum_{i=1}^2 \gamma_i(a) G_i(\mathbf{k}_1, \mathbf{k}_2), \text{ with} \quad (6.57)$$

$$G_1(\mathbf{k}_1, \mathbf{k}_2) = \mathbf{k}_{12}^2, \quad (6.58)$$

$$G_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{(\mathbf{k}_{12} \cdot \mathbf{k}_2)^2}{k_2^2} - \frac{1}{3} \mathbf{k}_{12}^2, \text{ and} \quad (6.59)$$

$$\gamma_i(a) \equiv - \frac{1}{D_1(a)} \int_{a_{in}}^a da' G_\delta(a, a') D_1(a') g_i(a'). \quad (6.60)$$

Subsequently, the convoluted first order counterterms give

$$H_2^{\alpha\beta}(\mathbf{k}_1, \mathbf{k}_2|a) = -\gamma(a)G_{\alpha\beta}(\mathbf{k}_1, \mathbf{k}_2), \quad \text{with} \quad (6.61)$$

$$G_{\alpha\beta}(\mathbf{k}_1, \mathbf{k}_2) \equiv \frac{4}{2m_g + 7} \beta(\mathbf{k}_1, \mathbf{k}_2) k_1^2 \quad (6.62)$$

$$+ \frac{2m_g + 5}{(m_g + 1)(2m_g + 7)} [(m_g + 1)\alpha(\mathbf{k}_1, \mathbf{k}_2) + \alpha(\mathbf{k}_2, \mathbf{k}_1)] k_1^2, \quad (6.63)$$

$$\gamma(a) = \frac{2}{m_g(2m_g + 5)} [D_1(a)]^{m_g+1} \bar{g}. \quad (6.64)$$

Lastly, the second order contribution to the lowest order counterterm is given by

$$H_2^\Psi(\mathbf{k}_1, \mathbf{k}_2|a) = -\gamma(a)G_\Psi(\mathbf{k}_1, \mathbf{k}_2), \quad \text{with} \quad (6.65)$$

$$G_\Psi(\mathbf{k}_1, \mathbf{k}_2) = \frac{m_g(2m_g + 5)}{(m_g + 1)(2m_g + 7)} \left[\mathbf{k}_{12}^2 \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_2^2} - \mathbf{k}_{12} \cdot \mathbf{k}_1 \right]. \quad (6.66)$$

6.A.2 One Loop Bispectrum

Having obtained perturbative solutions for the evolved density contrast in terms of the initial field, we can compute correlation functions using the statistical properties of the initial distribution. Along the lines of the discussion above, we decompose the bispectrum as

$$B_{\text{tot}} = B_{\text{SPT}}^{\text{G}} + B_{\text{EFT}}^{\text{G}} + f_{\text{NL}} (B_{\text{SPT}}^{\text{NG}} + B_{\text{EFT}}^{\text{G}}). \quad (6.67)$$

The expressions for the Gaussian part of the bispectrum at one loop were given in [181] and [182], and read

$$B_{\text{SPT}}^{\text{G}} = B_{112} + \left[B_{114} + B_{123}^{(\text{I})} + B_{123}^{(\text{II})} + B_{222} \right], \quad (6.68)$$

$$B_{\text{EFT}}^{\text{G}} = \xi B_\xi^{\text{G}} + \sum_{i=1}^3 \epsilon_i B_{\epsilon_i}, \quad (6.69)$$

where

$$B_{112} = 2F_2(\mathbf{k}_1, \mathbf{k}_2)P_{11}(k_1)P_{11}(k_2) + 2 \text{ cycl. perms} \quad (6.70a)$$

$$B_{222} = 8 \int_{\mathbf{p}} F_2(-\mathbf{p}, \mathbf{p} + \mathbf{k}_1)F_2(\mathbf{p} + \mathbf{k}_1, -\mathbf{p} + \mathbf{k}_2)F_2(\mathbf{k}_2 + -\mathbf{p}, \mathbf{p}) \\ P_{11}(p)P_{11}(|\mathbf{p} + \mathbf{k}_1|)P_{11}(|\mathbf{p} - \mathbf{k}_2|), \quad (6.70b)$$

$$B_{321}^{(I)} = 6P_{11}(k_3) \int_{\mathbf{p}} F_3(-\mathbf{p}, \mathbf{p} - \mathbf{k}_2, -\mathbf{k}_3)F_2(\mathbf{p}, \mathbf{k}_2 - \mathbf{p})P_{11}(p)P_{11}(|\mathbf{p} - \mathbf{k}_2|) + 5 \text{ perms}, \quad (6.70c)$$

$$B_{321}^{(II)} = 6F_2(\mathbf{k}_2, \mathbf{k}_3)P_{11}(k_2)P_{11}(k_3) \int_{\mathbf{p}} F_3(-\mathbf{k}_3, \mathbf{p}, -\mathbf{p})P_{11}(p) + 5 \text{ perms}, \quad (6.70d)$$

$$B_{411} = 12P_{11}(k_2)P_{11}(k_3) \int_{\mathbf{p}} F_4(\mathbf{p}, -\mathbf{p}, -\mathbf{k}_2, -\mathbf{k}_3)P_{11}(p) + 2 \text{ cycl. perms}. \quad (6.70e)$$

Here $P_{11}(k)$ is the linear power spectrum, whose time dependence is implied. Furthermore

$$B_{\xi}^G \equiv -2[E_{\alpha\beta}(\mathbf{k}_1, \mathbf{k}_2) + E_{\delta}(\mathbf{k}_1, \mathbf{k}_2)]P_{11}(k_1)P_{11}(k_2) + 2 \text{ perms}, \quad (6.71a)$$

$$B_{e_i} \equiv -2E_i(\mathbf{k}_1, \mathbf{k}_2)P_{11}(k_1)P_{11}(k_2) + 2 \text{ perms}. \quad (6.71b)$$

The non-Gaussian contribution at one loop is

$$B_{\text{SPT}}^{\text{NG}} = B_{111} + [B_{113}^{(I)} + B_{113}^{(II)} + B_{122}^{(I)} + B_{122}^{(II)}], \quad (6.72)$$

$$B_{\text{EFT}}^{\text{NG}} = \xi B_{\xi}^{\text{NG}} + \gamma B_{\gamma} + \sum_{i=1}^2 \gamma_i B_{\gamma_i}, \quad (6.73)$$

where

$$B_{113}^{(I)} = 3P_{11}(k_2) \int_{\mathbf{p}} F_3(\mathbf{k}_1 + \mathbf{p}, -\mathbf{p}, \mathbf{k}_2)B_{111}(k_1, p, |\mathbf{k}_1 + \mathbf{p}|) + 5 \text{ perms}, \quad (6.74a)$$

$$B_{113}^{(II)} = 3B_{111}(k_1, k_2, k_3) \int_{\mathbf{p}} F_3(\mathbf{k}_1, \mathbf{p}, -\mathbf{p})P_{11}(p) + 2 \text{ perms}, \quad (6.74b)$$

$$B_{122}^{(I)} = 4 \int_{\mathbf{p}} F_2(\mathbf{k}_3 + \mathbf{p}, -\mathbf{p})F_2(\mathbf{p}, \mathbf{k}_2 - \mathbf{p})B_{111}(k_1, |\mathbf{k}_3 + \mathbf{p}|, |\mathbf{k}_2 - \mathbf{p}|)P_{11}(p) + 2 \text{ perms}, \quad (6.74c)$$

$$B_{122}^{(II)} = 2F_2(\mathbf{k}_1, \mathbf{k}_2)P_{11}(k_2) \int_{\mathbf{p}} F_2(\mathbf{p}, \mathbf{k}_1 - \mathbf{p})B_{111}(k_1, p, |\mathbf{k}_1 - \mathbf{p}|) + 5 \text{ perms}, \quad (6.74d)$$

and

$$B_{\xi}^{\text{NG}} \equiv -(k_1^2 + k_2^2 + k_3^2)B_{111}(k_1, k_2, k_3), \quad (6.75a)$$

$$B_{\gamma} \equiv -[G_{\alpha\beta}(\mathbf{k}_1, \mathbf{k}_2) + G_{\Psi}(\mathbf{k}_1, \mathbf{k}_2)]P_{11}(k_1)P_{1\psi}(k_2) + 5 \text{ perms}, \quad (6.75b)$$

$$B_{\gamma_i} \equiv -G_i(\mathbf{k}_1, \mathbf{k}_2)P_{11}(k_1)P_{1\psi}(k_2) + 5 \text{ perms}. \quad (6.75c)$$

Again, the time dependence of the correlation functions is implied. Finally, we have to specify the type of non-Gaussianity. In this work we consider the primordial bispectra given in 6.2.2. The corresponding correlation between the linear density field and the first order non-Gaussian counterterm ψ (see (6.41)) is given by

$$P_{1\psi}(k) \equiv \frac{(k/\mu)^\Delta}{M(k)} P_{11}(k), \quad (6.76)$$

where μ is some arbitrary momentum scale, introduced for dimensional reasons, which cancels when multiplied with the EFT parameter in the full contribution to the bispectrum. We therefore set it to unity in the numerical evaluation. In our case the k -dependence is respectively given by $\Delta = \{0, 1, 2\}$ for local, quasi-single field and equilateral type non-Gaussianities. The transfer function was defined in (6.7). For this work we choose the time dependence of the lowest order counterterms (6.42) to match the divergence it is supposed to cancel, which corresponds to the choice $m_d = m_g = 1$. This is was argued for in [296]. Moreover, our main results do not depend much on this assumption.

6.A.3 Ansatz Two Loop Bispectrum

As an ansatz for the two loop bispectrum we compute the two reducible two loop diagrams, given by [181],

$$B_{332}^I = 2F_2(\mathbf{k}_1, \mathbf{k}_2) \frac{P_{13}(k_1)}{2} \frac{P_{13}(k_2)}{2} + 2 \text{ cycl. perms} \quad (6.77a)$$

$$B_{332}^{II} = 6 \frac{P_{13}(k_3)}{2} \int_{\mathbf{p}} F_3(-\mathbf{p}, \mathbf{p} - \mathbf{k}_2, -\mathbf{k}_3) F_2(\mathbf{p}, \mathbf{k}_2 - \mathbf{p}) P_{11}(p) P_{11}(|\mathbf{p} - \mathbf{k}_2|) + 5 \text{ perms}, \quad (6.77b)$$

with

$$P_{13}(k) = 6P_{11}(k) \int_{\mathbf{p}} F_3(\mathbf{k}, \mathbf{p}, -\mathbf{p}) P_{11}(p). \quad (6.78)$$

As an estimate for the theoretical error we use

$$B_{332} = |B_{332}^I| + |B_{332}^{II}|. \quad (6.79)$$

6.B Theoretical noise

This appendix contains the details of the implementation of the theoretical error and further investigates some issues related to it. First, we give the intermediate steps to derive (6.25) and provide an alternative derivation of the effective Fisher matrix in the presence of theoretical error. Then, we study the effect of the correlation length, both by means of a toy model and by running the analysis for several correlations lengths. Finally, we discuss in more detail the effect of the two possible ansätze for B^{er} .

6.B.1 Derivation of (6.25)

Let us first show how to go from equations (6.22), (6.23) and (6.24b) to the effective Fisher matrix given in (6.25). We will need to use the Woodbury matrix identity several times, which relates the inverse of sums of matrices to their individual inverses

$$(A + B)^{-1} = A^{-1} - A^{-1} (A^{-1} + B^{-1})^{-1} A^{-1}. \quad (6.80)$$

By using this identity, we can rewrite (6.23) as

$$F_{\alpha\beta}^{-1} = (N_{\alpha\beta}^{-1} + D_{\alpha\beta})^{-1} = D_{\alpha\beta}^{-1} - D_{\alpha\gamma}^{-1} (N + D^{-1})_{\gamma\delta}^{-1} D_{\delta\beta}^{-1}. \quad (6.81)$$

This allows us to compute

$$\begin{aligned} F_{i\gamma} F_{\gamma\delta}^{-1} F_{\delta j} &= \sum_{k,p} B_i(k) C^{-1}(k, k_\gamma) B^{\text{er}}(k_\gamma) (D^{-1} - D^{-1} (N + D^{-1})^{-1} D^{-1})_{\gamma\delta} \\ &\quad \times B^{\text{er}}(k_\delta) C^{-1}(k_\delta, p) B_j(p) \\ &= \sum_{k,p} B_i(k) \delta_{k, k_\alpha} \left(C^{-1}(k_\alpha, k_\beta) - \frac{1}{B^{\text{er}}(k_\alpha)} (N + D^{-1})_{\alpha\beta}^{-1} \frac{1}{B^{\text{er}}(k_\beta)} \right) \\ &\quad \times \delta_{k_\beta, p} B_j(p) \\ &= \sum_{k_\alpha, k_\beta} \frac{B_i(k_\alpha)}{B^{\text{er}}(k_\alpha)} (D - (N + D^{-1})^{-1})_{\alpha\beta} \frac{B_i(k_\beta)}{B^{\text{er}}(k_\beta)} \\ &= \sum_{k_\alpha, k_\beta} \frac{B_i(k_\alpha)}{B_\alpha} (D(N^{-1} + D)^{-1} D)_{\alpha\beta} \frac{B_i(k_\beta)}{B_\beta}. \end{aligned}$$

Here summation over the Greek indices is understood. Upon applying the Woodbury identity again, the effective Fisher matrix then becomes

$$\begin{aligned}
F_{ij}^{\text{eff}} &= \sum_{k_\alpha, k_\beta} \frac{B_i(k_\alpha)}{B_\alpha} (D - D(N^{-1} + D)^{-1}D)_{\alpha\beta} \frac{B_j(k_\beta)}{B_\beta} + (C_\Theta^{-1})_{ij} \\
&= \sum_{k_\alpha, k_\beta} \frac{B_i(k_\alpha)}{B_\alpha} (N + D^{-1})_{\alpha\beta}^{-1} \frac{B_j(k_\beta)}{B_\beta} + (C_\Theta^{-1})_{ij} \\
&= \sum_{k_\alpha, k_\beta} B_i(k_\alpha) (N^{\text{eff}} + C_B)_{\alpha\beta}^{-1} B_j(k_\beta) + (C_\Theta^{-1})_{ij},
\end{aligned} \tag{6.82}$$

with $N_{\alpha\beta}^{\text{eff}} = B^{\text{er}}(k_\alpha)N_{\alpha\beta}B^{\text{er}}(k_\beta)$.

6.B.2 Alternative derivation of the effective Fisher matrix

Next, we present a slightly different derivation of the effective Fisher matrix, by marginalizing at the level of the likelihood function. Let us first expand $\chi^2 = -2\log(\mathcal{L})$ in the nuisance parameters Θ_α . We would like to expand about some value $\bar{\Theta}_\alpha$ to get

$$\begin{aligned}
\chi^2(\Theta_i, \Theta_\alpha) &= \chi^2(\Theta_i, \bar{\Theta}_\alpha) + (\Theta_\alpha - \bar{\Theta}_\alpha)\chi_\alpha^2(\Theta_i, \bar{\Theta}_\alpha) \\
&\quad + \frac{1}{2}(\Theta_\alpha - \bar{\Theta}_\alpha)(\Theta_\beta - \bar{\Theta}_\beta)\chi_{\alpha\beta}^2(\Theta_i, \bar{\Theta}_\alpha),
\end{aligned} \tag{6.83}$$

where summation over repeating indices is understood, and the index α on χ^2 denotes a derivative with respect to the corresponding nuisance parameter. It is an equality, since the variables are Gaussian distributed. We can rewrite this expression in more compact notation as

$$\chi^2 = \chi_0^2 + \delta\Theta^\alpha X_\alpha + \frac{1}{2}\delta\Theta^\alpha\delta\Theta^\beta Y_{\alpha\beta}, \tag{6.84}$$

where χ_0^2 is the chi-squared we would get if we ignored the presence of the nuisance parameters Θ_α . By completing the square and adding some prior information on the nuisance parameters (i.e. a covariance matrix), we can integrate them out to get an effective chi-squared. In other words, we would like to evaluate the following integral

$$\int d^N\Theta_\alpha \exp\left(-\frac{1}{2}\left(\chi_0^2 + \delta\Theta^\alpha X_\alpha + \frac{1}{2}\delta\Theta^\alpha\delta\Theta^\beta Y_{\alpha\beta}\right)\right) \tag{6.85}$$

$$\times \exp\left(-\frac{1}{2}\delta\Theta^\alpha(N^{-1})_{\alpha\beta}\delta\Theta^\beta\right). \tag{6.86}$$

where $N_{\alpha\beta}$ is the covariance matrix of the theoretical error parameters. The integration results into

$$\sqrt{(2\pi)^N \cdot \det((\frac{1}{2}Y + N^{-1})^{-1})} \exp\left(-\frac{1}{2}\chi_0^2 + \frac{1}{4}X_\gamma (Y + 2N^{-1})_{\gamma\delta}^{-1} X_\delta\right), \quad (6.87)$$

and therefore,

$$\chi_{\text{eff}}^2 = \chi_0^2 - \frac{1}{2}X_\gamma (Y + 2N^{-1})_{\gamma\delta}^{-1} X_\delta + \ln(\det((\frac{1}{2}Y + N^{-1}))). \quad (6.88)$$

Please note that all these terms do in general depend on Θ_i . Since the joint probability distribution of all parameters is a multivariate Gaussian, we know that Y is independent of Θ_i and X only depends linearly on Θ_i . In that case, we get

$$(\chi_{\text{eff}}^2)_{ij} = (\chi_0^2)_{ij} - \frac{1}{2}X_{i\gamma} (Y + 2N^{-1})_{\gamma\delta}^{-1} X_{\delta j}. \quad (6.89)$$

The full Fisher matrix is given by

$$F_{\mu\nu} = \begin{pmatrix} \frac{1}{2}Y_{\alpha\beta} + N_{\alpha\beta}^{-1} & \frac{1}{2}X_{\alpha j} \\ \frac{1}{2}X_{i\beta} & F_{ij} \end{pmatrix} \quad (6.90)$$

where we have to evaluate the matrices at the maximum likelihood value of the parameters. This means that the effective chi-squared is given by

$$(\chi_{\text{eff}}^2)_{ij} = (\chi_0^2)_{ij} - 2F_{i\gamma}F_{\gamma\delta}^{-1}F_{\delta j}, \quad (6.91)$$

or, in other words, the effective Fisher matrix for the theoretical parameters is given by

$$F_{ij}^{\text{eff}} = F_{ij} - F_{i\gamma}F_{\gamma\delta}^{-1}F_{\delta j}. \quad (6.92)$$

This is what we found before in the main text.

Alternatively, starting from (6.89) we can write down immediately the expression for the effective likelihood

$$\begin{aligned} \mathcal{L}^{\text{eff}} &= \frac{1}{\sqrt{\det(\frac{1}{2}Y + N^{-1})}} \exp\left[-\frac{1}{2}\left(\chi_0^2 - X_\gamma(\frac{1}{2}Y + N^{-1})_{\gamma\delta}^{-1}X_\delta\right)\right] \\ &= \frac{1}{\sqrt{\det(D + N^{-1})}} \exp\left[-\frac{1}{2}\sum_{\hat{k}, p} \Delta B(\hat{k}) (C_B^{-1}(\hat{k}, p) - C_B^{-1}(\hat{k}, \hat{k}_\gamma) \right. \\ &\quad \left. B_{2L}(k_\gamma)(D + N^{-1})_{\gamma\delta}^{-1} B^{\text{er}}(k_\delta)C_B^{-1}(\hat{k}_\delta, p)) \Delta B(p)\right] \quad (6.93) \\ &= \frac{1}{\sqrt{\det(D + N^{-1})}} \exp\left[-\frac{1}{2}\sum_{\hat{k}_\alpha, \hat{k}_\beta} \Delta B(\hat{k}_\alpha) (C_B + N^{\text{eff}})_{\alpha\beta}^{-1} \Delta B(\hat{k}_\beta)\right]. \end{aligned}$$

Here the difference between the data and theory vector ΔB is evaluated at the fiducial values for the nuisance parameters $\bar{\Theta}_\alpha$ (i.e. at zero in our case). Taking now two derivatives with respect to the remaining theoretical parameters Θ_i , we find the effective Fisher matrix

$$F_{ij}^{\text{eff}} = \sum_{k_\alpha, k_\beta} B_i(k_\alpha) (C_B + N^{\text{eff}})_{\alpha\beta}^{-1} B_j(k_\beta) \quad (6.94)$$

with $N_{\alpha\beta}^{\text{eff}} = B^{\text{er}}(k_\alpha) N_{\alpha\beta} B^{\text{er}}(k_\beta)$.

6.B.3 Theoretical error - a toy model

In our approach, the theoretical error on the value of the bispectrum is modeled in the following way. For every bin, we introduce a nuisance parameter that is drawn from a Gaussian distribution with average zero and variance set by the estimated size of the theoretical error for that bin. Importantly, we allow for non-vanishing correlations among these nuisance parameters, i.e. we allow for a non-diagonal covariance matrix for them. The purpose of this appendix is to show that both the limit of zero and maximal correlation among the parameters have a clear interpretation, neither of which resembles the way we think the theoretical error should act. To be more precise, we prove, by means of a simple toy model that still captures the essence of the real analysis, the intuitive statements that:

- for zero correlation length, the theoretical error just acts as shot noise per bin;
- for maximal correlation length, the theoretical error acts as some free coefficient multiplying a fixed shape function, which by definition we think is the wrong function.

Toy model

We consider measuring some observable d a total of N times and collecting the data d_i . Our model is $d_i = x + e_i$, with x a Gaussian random variable with variance σ_x^2 , whose average, \bar{x} , we would like to determine as well as possible. The e_i are additional Gaussian variables that represent the systematic error or theoretical uncertainty in every measurement. Their averages and variances are \bar{e}_i and $\sigma_{e_i}^2$, respectively. One can think of this scenario as determining the average weight of a group of people, knowing that their weights are Gaussian distributed with variance σ_x^2 , where we use a different weighing scale with a systematic error \bar{e}_i and some uncertainty in the measurement characterized by

$\sigma_{m_i}^2$ every time we weigh someone. Since the e_i are uncorrelated with x , this leads to the likelihood

$$\log L = - \sum_{i=1}^N \frac{(d_i - \bar{x} - \bar{e}_i)^2}{2\sigma_d^2}, \quad (6.95)$$

where $\sigma_d^2 = \sigma_x^2 + \sigma_m^2$, assuming for convenience that $\sigma_{m_i} = \sigma_m$ (the arguments below do not depend on this assumption). Without any prior information on the systematic errors, they are completely degenerate with \bar{x} , so we do not expect to be able to learn anything about \bar{x} in this case. This can be verified using a Fisher analysis. We have

$$F_{ab} = \begin{pmatrix} \frac{N}{\sigma_d^2} & \frac{1}{\sigma_d^2} \vec{1}^T \\ \frac{1}{\sigma_d^2} \vec{1} & \frac{1}{\sigma_d^2} \mathbf{1}_{N \times N} \end{pmatrix}, \quad (6.96)$$

where $a, b = \bar{x}, \bar{e}_i$. Since we are ignorant about the systematic errors, we compute the marginalized error on \bar{x} ,

$$\sigma_{\bar{x}, marg}^2 = (F^{-1})_{\bar{x}\bar{x}}, \quad (6.97)$$

which can be computed using the block matrix inversion formula (see also (6.22)):

$$\text{let } F = \begin{pmatrix} A & B^T \\ B & D \end{pmatrix}, \quad (6.98)$$

then

$$\sigma_{\bar{x}, marg}^2 = (A - B^T D^{-1} B)^{-1} = \left(\frac{N}{\sigma_d^2} - \frac{1}{\sigma_d^2} \vec{1}^T \mathbf{1}_{N \times N} \vec{1} \right)^{-1} = \frac{1}{0}, \quad (6.99)$$

as expected. In a realistic situation we do have some prior information about the systematic errors. Here, and in the paper, we assume they are also Gaussian random variables with some variance $\sigma_{\bar{e}_i}^2$. Moreover, we allow for non-trivial correlations among the \bar{e}_i , which for the scales could mean they were produced by the same machine for instance. This means we obtain the updated likelihood

$$\log L = - \sum_{i=1}^N \frac{(d_i - \bar{x} - \bar{e}_i)^2}{2\sigma_d^2} - \bar{e}_i (C^{-1})_{ij} \bar{e}_j, \quad (6.100)$$

where

$$C_{ij} = \langle \bar{e}_i \bar{e}_j \rangle. \quad (6.101)$$

In the following, we investigate the effect of zero and maximal correlation length on $\sigma_{\bar{x}, marg}^2$.

Zero correlation

Let us first assume zero correlation among the systematic errors, leading to a diagonal covariance matrix,

$$C_{ij} = \sigma_i^2 \delta_{ij}. \quad (6.102)$$

In terms of the weighing scales this could mean all scales really come from different companies with uncorrelated systematic errors. We now show that in this case the ignorance about the systematic errors acts as shot noise per bin; it simply updates the variance of the measurements $\sigma_d^2 \rightarrow \sigma_{d_i}^2$. For any covariance matrix, the Fisher matrix is

$$F = \begin{pmatrix} \frac{N}{\sigma_d^2} & \frac{1}{\sigma_d^2} \vec{1}^T \\ \frac{1}{\sigma_d^2} \vec{1} & \frac{1}{\sigma_d^2} \delta_{ij} + (C^{-1})_{ij} \end{pmatrix}. \quad (6.103)$$

Then, using the block matrix inversion formula, zero correlation leads to an error

$$\sigma_{\bar{x}, marg}^2 = \left(\frac{N}{\sigma_d^2} - \frac{1}{\sigma_d^4} \vec{1}^T \frac{1}{\frac{1}{\sigma_d^2} + \frac{1}{\sigma_i^2}} \delta_{ij} \vec{1} \right)^{-1} \quad (6.104)$$

$$= \left[\sum_{i=1}^N \left(\frac{1}{\sigma_d^2} - \frac{1}{\sigma_d^4} \frac{1}{\frac{1}{\sigma_d^2} + \frac{1}{\sigma_i^2}} \right) \right]^{-1} \quad (6.105)$$

$$= \left[\sum_{i=1}^N \left(\frac{1}{\sigma_d^2 + \sigma_i^2} \right) \right]^{-1} \equiv \left[\sum_{i=1}^N \frac{1}{\sigma_{d_i}^2} \right]^{-1}, \quad (6.106)$$

which is the same error one gets from assuming the likelihood function

$$\log L = - \sum_{i=1}^N \frac{(d_i - \bar{x})^2}{2\sigma_{d_i}^2}. \quad (6.107)$$

This shows that indeed for zero correlation the systematic errors acts as shot noise per bin. In particular, this means the error on \bar{x} can be made arbitrarily small by increasing the number of measurements (if the $\sigma_{\bar{e}_i}$ do not grow too fast for additional measurements). The intuitive reason is of course that in this model we expect the systematic errors to average out to zero in the long run. This is clearly not what is expected of the theoretical error in the measurement of the bispectrum.

Maximal correlation

Next we assume maximal correlation, which by definition means

$$C_{ij} = \langle \bar{e}_i \bar{e}_j \rangle = \sqrt{\langle \bar{e}_i^2 \rangle} \sqrt{\langle \bar{e}_j^2 \rangle} = \sigma_i \sigma_j. \quad (6.108)$$

Since this matrix has rank one (all columns are multiples of the same vector), it is not invertible in more than one dimension. One way to deal with this is to introduce a regulator, such as a small matrix $\epsilon \delta_{ij}$, to break the degeneracy. Using our block matrix inversion formula, this is however not necessary. In the notation of ((6.98)), we wish to compute D^{-1} . Let us write $D = S + C^{-1}$, where $S = \frac{1}{\sigma_d^2} \delta_{ij}$. The Woodbury identity then gives

$$(S + C^{-1})^{-1} = S^{-1} - S^{-1}(S^{-1} + C)^{-1}S^{-1}. \quad (6.109)$$

Hence, we have to compute the inverse of $S^{-1} + C$, where $S^{-1} = \sigma_d^2 \delta_{ij}$, and $C = \sigma_i \sigma_j$. Conveniently, since C is of the form $\vec{\sigma}(\vec{\sigma})^T$, we can use the Sherman-Morrison formula to compute the inverse

$$(S^{-1} + C)^{-1} = S - \frac{S(\sigma_i \sigma_j)S}{1 + \sigma_i \sigma_j \sigma_j}. \quad (6.110)$$

Plugging this into the previous formula, we find

$$D^{-1} = (S + C^{-1})^{-1} = S^{-1} - S^{-1} + \frac{\sigma_i \sigma_j}{1 + \frac{(\sum \sigma_i^2)}{\sigma_d^2}} = \frac{\sigma_i \sigma_j}{1 + \frac{(\sum \sigma_i^2)}{\sigma_d^2}}. \quad (6.111)$$

This finally leads to the error on \bar{x} :

$$\sigma_{\bar{x}, \text{marg}}^2 = \left[\frac{N}{\sigma_d^2} - \frac{(\sum \sigma_i)^2}{\sigma_d^2 + (\sum \sigma_i^2)} \right]^{-1}. \quad (6.112)$$

In order to interpret this result, let us rewrite this expression as follows

$$\sigma_{\bar{x}, \text{marg}}^2 = \left[\frac{N}{\sigma_d^2} - \frac{(\sum \frac{\sigma_i}{\sigma})^2}{\sigma_d^4} \frac{1}{\frac{1}{\sigma^2} + \frac{(\sum (\frac{\sigma_i}{\sigma})^2)}{\sigma_d^2}} \right]^{-1}, \quad (6.113)$$

where we have introduced the dimensionful parameter σ to keep the dimensions clean. Now observe that we get exactly the same error on \bar{x} from the following likelihood function

$$\log L = - \sum_{i=1}^N \frac{(d_i - \bar{x} - \frac{\sigma_i}{\sigma} \bar{e})^2}{2\sigma_d^2} - \frac{\bar{e}^2}{2\sigma^2}. \quad (6.114)$$

whose Fisher matrix is

$$F = \begin{pmatrix} \frac{N}{\sigma_d^2} & \frac{1}{\sigma_d^2} \sum \frac{\sigma_i}{\sigma} \\ \frac{1}{\sigma_d^2} \sum \frac{\sigma_i}{\sigma} & \frac{1}{\sigma^2} + \frac{\sum (\frac{\sigma_i}{\sigma})^2}{\sigma_d^2} \end{pmatrix}, \quad (6.115)$$

This means that the maximal correlation case is equivalent to having a single, unknown parameter multiplying a known ‘shape’ function σ_i/σ . In terms of the weighing scales this would mean that we know in advance exactly the ratios between the systematic errors of the scales. In terms of the bispectrum this would mean that we claim to know the theoretical error is exactly some number times the two loop estimate we put in, which it is clearly not. Finally note that if we choose all σ_i to be equal, which for convenience we take to be σ , we find

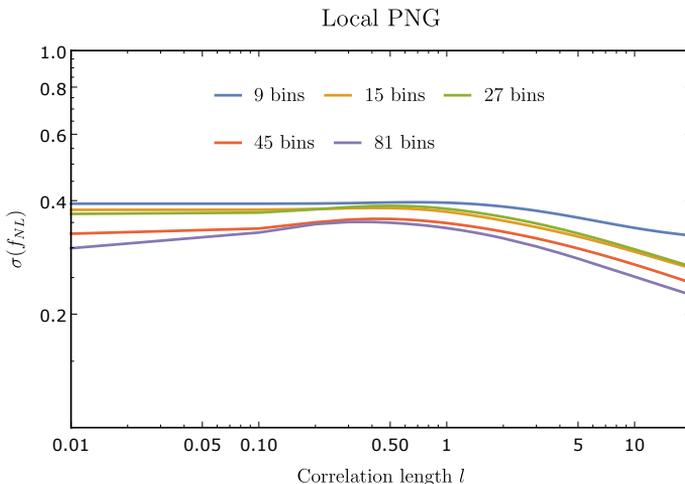
$$\sigma_{\bar{x}}^2 = \left[\frac{N}{\sigma_d^2} - \frac{N^2}{\sigma_d^4} \frac{1}{\frac{1}{\sigma^2} + \frac{N}{\sigma_d^2}} \right]^{-1} = \frac{\sigma_d^2}{N} + \sigma^2, \quad (6.116)$$

meaning the error on \bar{x} can never get below the uncertainty in the degenerate parameter \bar{e} . In terms of the weighing problem this makes perfect sense, as this case is equivalent to simply using one and the same scale for every measurement. In this case we never expect to beat the unknown systematic error in the scale. In terms of the bispectrum this shows the importance of the relation between the shapes of the non-Gaussian signal and the theoretical error. In fact, in the maximal correlation limit we treat the theoretical error exactly the same as the EFT terms.

Conclusions

From the above example it is clear that in neither limit the implementation of the theoretical error is completely satisfactory. Moreover, if the shapes are not too similar, the estimates from both limits are probably too optimistic. For this reason we recommend a conservative use of the method. In particular, we choose to use the correlation length that gives the weakest constraints on f_{NL} , as we show in the next subsection.

Figure 6.7: A test computation of $\sigma(f_{NL})$ at redshift $z = 0$ including theoretical error as function of correlation length. We choose $k_{\min} = 0.001 \text{ hMpc}^{-1}$ and $k_{\max} = 1 \text{ hMpc}^{-1}$ where we divide the k -range in 9 (blue), 15 (orange), 27 (green), 45 (red) and 81 (purple) bins.



6.B.4 Choice of correlation length

In order to find the most conservative correlation length to work with, we ran a test computation of $\sigma(f_{NL})$. We did this at redshift zero with $k_{\min} = 0.001 \text{ hMpc}^{-1}$ and $k_{\max} = 1 \text{ hMpc}^{-1}$, where we divide the k -range in 9 (blue), 15 (orange), 27 (green), 45 (red) and 81 (purple) bins. We find that the weakest constraints are obtained for $l \approx 0.5$, see Figure 6.7. This is therefore the value we take for the analysis in the paper.

Remarkably, the error is actually very insensitive to the correlation length, despite the very different nature of the effect of small and large correlation length. We believe the reason for this to be the fact that our ansatz for the theoretical error is a much steeper function of k than the non-Gaussian signal. The transition from the k 's for which the error is negligible to the region where it is completely dominant is therefore very small, and the shape of the error is therefore not very important in this case.

Another observation is that the error keeps increasing as we increase the correlation length beyond 10 decades, whereas the k 's we consider only run over a couple of decades. This makes the nuisance parameters almost maximally correlated for all these large correlation lengths. At the moment, we have no good explanation for the fact that the error seems to keep improving, other than it being a numerical fluke, perhaps related to the inversion of the correlation matrix.

6.B.5 Ansätze for higher loop corrections

We compare the ansätze for the higher loop corrections in Figure 6.8. It is a zoom-in of Figure 6.1, where we now show in addition the ansatz used in [294], for the two and three loop contribution to the bispectrum. Please note that in [183], we only showed one of the two reducible two loop diagrams contributing to B_{332} . Therefore, the plots look different now, in particular in the squeezed configuration of the bispectrum. We see that in the squeezed configuration, the ansatz B_{332} is an order of magnitude smaller than E_b . This explains why we have to multiply B_{332} by a factor 10 in section 6.3.5 to get reliable results. Furthermore, we note that at redshift zero, B_{332} allows one to go to higher k_{\max} in the squeezed configuration, whereas E_b allows one to go further in the equilateral configuration (for f_{NL} bigger than 10). This explains why using B_{332} as an ansatz gives more optimistic results for local PNG, whereas E_b gives more optimistic results for equilateral PNG (see section 6.4.1). Keep in mind also that the time dependence of the theoretical error terms is different from the signal, making the signal stronger at higher redshifts.

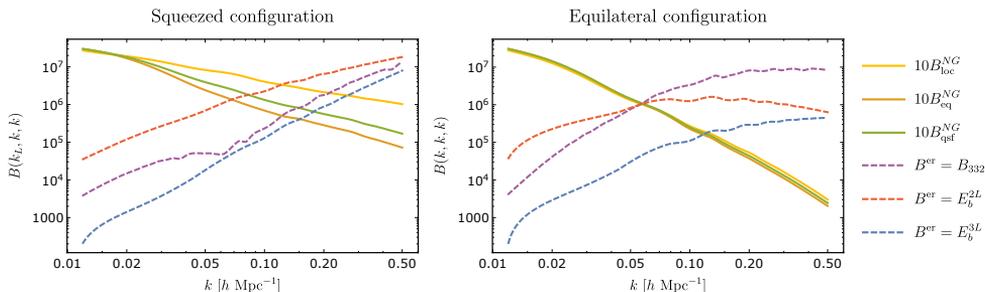


Figure 6.8: Comparison of the ansätze for the higher loop corrections. We plot E_b from equation (6.27) for two and three loops (yellow and green dashed lines) versus B_{332} (green dashed line) defined in equation (6.77). For the two loop ansatz using E_b we take $n = -1.4$, $k_{NL} = 0.45h\text{Mpc}^{-1}$ and $l = 2$ and for the three loop ansatz we use $n = -1.5$, $k_{NL} = 0.50h\text{Mpc}^{-1}$ and $l = 3$. We compare these ansätze with the non-Gaussian contribution to the bispectrum up to one loop with $f_{NL} = 10$ for local, equilateral and quasi-single-field PNG (red, blue and purple solid lines). In the left panel we compare the different contributions in the configuration $B(k_L, k, k)$ where we varied k and fixed $k_L = 0.012h\text{Mpc}^{-1}$. The smaller k the more squeezed the configuration is. In the right panel we show the equilateral configuration $B(k, k, k)$.

Next, we consider $\sigma(f_{NL})$ as function of k_{\max} at various redshifts in Figure 6.9. We do not include shotnoise, but we integrate out the theoretical error. The result for local PNG is shown in the left panel. We see that at redshift zero, the signal freezes out at some $k_{\max} < 1h\text{Mpc}^{-1}$. Furthermore, in agreement with what we expect from Figure 6.8, we see that using B_{332} as ansatz for the

two loop corrections gives more optimistic results. More specifically, $\sigma(f_{NL})$ is a factor 5 smaller. At higher redshifts, we find that $\sigma(f_{NL})$ does not freeze when we reach $k_{\max} = 1h\text{Mpc}^{-1}$. We think this is due to the fact that we keep gaining information as we go to more and more squeezed configurations. This is also important for equilateral PNG, shown in the right panel, even though, $\sigma(f_{NL})$ does freeze out in this case. Interestingly, compared to scaling estimates for k_{\max} for equilateral PNG (see for instance [294]) we find that we can go to much smaller scales than naively thought. The squeezed limit allows us to extract more information, also for equilateral PNG. The fact that $k_{\max} = 1h\text{Mpc}^{-1}$ is not large enough to ensure that $\sigma(f_{NL})$ is frozen when we ignore shotnoise explains the results we find in section 6.4.1.

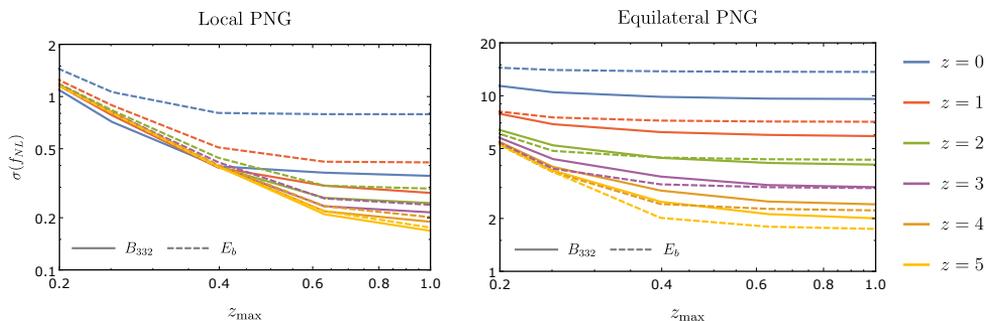


Figure 6.9: We show $\sigma(f_{NL})$ as function of k_{\max} using B_{332} and E_b as ansatz for the two loop corrections (solid and dashed lines). In the left panel we show the results for local PNG and in the right panel for equilateral PNG. The redshift takes values between $z = 0$ and $z = 5$. We use $k_{\min} = 0.001h\text{Mpc}^{-1}$ and $V = (2\pi/k_{\min})^3$ at each redshift.

6.C Choice of binning and volume of the bins

In this appendix, we motivate the decision of section 6.3.3 to use logarithmic binning and exactly computed values of V_{123} .

6.C.1 Exact computation of V_{123}

We will now explain how to compute V_{123} exactly and systematically by dividing all the bins in ‘interior’ and ‘edge’ bins. Moreover, the selection of bins is now determined by whether it contains at least some valid triangles instead of the usual selection rule that the central point should be a triangle.

Recall that V_{123} is defined as

$$V_{123} = \int_{\mathbf{q}_1} \int_{\mathbf{q}_2} \int_{\mathbf{q}_3} \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3). \quad (6.117)$$

We choose logarithmic binning, i.e. we have

$$q_i \equiv |\mathbf{q}|_i \in \left[k_i e^{-\frac{1}{2}\Delta \ln k}, k_i e^{\frac{1}{2}\Delta \ln k} \right]. \quad (6.118)$$

The integrand above only depends on the relative orientations of the vectors and their lengths. Fixing \mathbf{q}_1 along the \hat{z} -direction and \mathbf{q}_2 to be in the (x, z) -plane, their relative orientation is given by $\theta_{12} = \theta_2$. Now the lengths of these vectors, together with c_{12} , the cosine of θ_{12} , completely determine \mathbf{q}_3 . The length of \mathbf{q}_3 is then restricted to be in $\left[k_3 e^{-\frac{1}{2}\Delta \ln k}, k_3 e^{\frac{1}{2}\Delta \ln k} \right]$, which means

$$c_{12} \in [-1, 1] \cap \left[\frac{\left(k_3 e^{-\frac{1}{2}\Delta \ln k} \right)^2 - q_1^2 - q_2^2}{2q_1 q_2}, \frac{\left(k_3 e^{\frac{1}{2}\Delta \ln k} \right)^2 - q_1^2 - q_2^2}{2q_1 q_2} \right], \quad (6.119)$$

where q_1 and q_2 also take values within their bin. Then, if $[-1, 1]$ contains the range on the right for all values of q_1 and q_2 , we are dealing with an ‘interior bin’, and we get

$$\int dc_{12} dq_1 dq_2 q_1^2 q_2^2 = k_1^2 k_2^2 k_3^2 \sinh^3(\Delta \ln k). \quad (6.120)$$

Finally, accounting for the fact that we fixed θ_1 , $\phi_{1,2}$ and the factors of $(2\pi)^3$ we find

$$V_{123} = \frac{1}{(2\pi)^9} 8\pi^2 k_1^2 k_2^2 k_3^2 \sinh^3(\Delta \ln k). \quad (6.121)$$

This approximation breaks down when the two ranges are partly overlapping, in which case we have an ‘edge bin’.

let us evaluate V_{123} more precisely. We have seen that the integral simplifies to

$$V_{123} = \frac{8\pi^2}{(2\pi)^9} \int_{q_1} \int_{q_2} \int_{c_{12}} q_1^2 q_2^2 \quad (6.122)$$

where the c_{12} is restricted to be in the range given above. This integral can therefore be rewritten as

$$V_{123} = \frac{8\pi^2}{(2\pi)^9} \int_{q_1} \int_{q_2} \frac{1}{2} q_1 q_2 \max \left[0, \left(\min \left[(q_1 + q_2)^2, k_3^2 e^{\Delta \ln k} \right] - \max \left[(q_1 - q_2)^2, k_3^2 e^{-\Delta \ln k} \right] \right) \right].$$

In other words we integrate over the overlap

$$\left[(q_1 - q_2)^2, (q_1 + q_2)^2 \right] \cap \left[k_3^2 e^{-\Delta \ln k}, k_3^2 e^{\Delta \ln k} \right]. \quad (6.123)$$

There are multiple possibilities:

- The overlap is always zero.

This happens whenever $|q_1 + q_2|_{\max} \leq k_3 e^{-\frac{1}{2}\Delta \ln k}$ or $|q_1 - q_2|_{\min} \geq k_3 e^{\frac{1}{2}\Delta \ln k}$. This means we should exclude the cases $k_3 \geq (k_1 + k_2) e^{\Delta \ln k}$ and $k_3 < k_2$. The latter is already excluded since we have $k_1 \leq k_2 \leq k_3$. The first leads to a constraint to select the bins, namely

$$k_3 < (k_1 + k_2) e^{\Delta \ln k}. \quad (6.124)$$

- The first range always contains the second range.

This happens when $|q_1 - q_2|_{\max} \leq k_3 e^{-\frac{1}{2}\Delta \ln k}$ and $|q_1 + q_2|_{\min} \geq k_3 e^{\frac{1}{2}\Delta \ln k}$. So we need both

$$\begin{cases} k_3 \geq k_2 e^{\Delta \ln k} - k_1 \\ k_3 \leq (k_1 + k_2) e^{-\Delta \ln k} \end{cases} \quad (6.125)$$

In this case the volume takes the simple form

$$V_{123} = \frac{8\pi^2}{(2\pi)^9} k_1^2 k_2^2 k_3^2 \sinh^3(\Delta \ln k). \quad (6.126)$$

- Any other type of overlap.

For the other cases we have to compute the actual volume of the bin. We will numerically perform the integral given above. This is when one of the two options below is satisfied

$$\begin{cases} k_3 < k_2 e^{\Delta \ln k} - k_1 \\ k_3 > (k_1 + k_2) e^{-\Delta \ln k} \end{cases} \quad (6.127)$$

Not for all these edge bins the central point has to be a triangle, since there are some cases for which $k_3 > k_1 + k_2$, considering the second inequality. Thus, we can either decide to define another point in the bin

to represent the central triangle or we can merge these bins with one of their neighbors. In the first case a valid central triangle in the bin (k_1, k_2, k_3) is given by

$$\left(k_1 e^{\frac{k_3}{2(k_2+k_1)}}, k_2 e^{\frac{k_3}{2(k_2+k_1)}}, k_3 e^{-\frac{k_3}{2(k_2+k_1)}} \right) \quad (6.128)$$

The other option is to merge (i.e. we add the volumes) the bin with one of its neighbors, which has the advantage that we never have to change the representing triangle of a given bin. For practical reasons, we choose this option. We implement this by merging each bin \mathbf{k} for which $k_3 > k_1 + k_2$ with the bin \mathbf{p} which has $p_1 = k_1$, $p_2 = k_2$ and p_3 the biggest value below or equal to $k_1 + k_2$.

6.C.2 Logarithmic versus linear binning

Let us now compare logarithmic binning with linear binning. We will show two examples of a computation of a Fisher matrix and show that the linear binning might cause problems.

We assume the following form for the Fisher matrix

$$F = \sum_{k_1, k_2, k_3} f(k_1, k_2, k_3) \frac{V_{123}}{s_{123}}, \quad (6.129)$$

for some function $f(k_1, k_2, k_3)$. We will consider a ‘local’-type function f^{loc} and an ‘equilateral’-type function f^{eq} . The local function corresponds to assuming the late time power spectrum scales as $P(k) \sim k^{-3}$, where F represents the (f_{NL}, f_{NL}) -component of the Fisher matrix for local PNG. Forgetting about the right normalization, this gives

$$f^{\text{loc}}(k_1, k_2, k_3) = \left(\frac{k_1^{3/2}}{k_2^{3/2} k_3^{3/2}} + \frac{k_2^{3/2}}{k_1^{3/2} k_3^{3/2}} + \frac{k_3^{3/2}}{k_2^{3/2} k_1^{3/2}} \right)^2. \quad (6.130)$$

Similarly, we can define a function that corresponds to equilateral PNG

$$f^{\text{eq}}(k_1, k_2, k_3) = \frac{k_1 k_2 k_3}{(k_1 + k_2 + k_3)^6}. \quad (6.131)$$

We compute F over a range $k \in [0.003, 0.5] h\text{Mpc}^{-1}$ for both logarithmic and linear bins and for both the approximate and exact computation of V_{123} . In Figure 6.10 we plot F as function of number of bins (number of triangles) considered.

First of all, in Figure 6.10a, restricting ourselves to about 10.000 bins, we see

that if we use the approximate value for V_{123} (the usual assumption), the linear bins seem to converge quickly to the asymptotic value. The logarithmic bins seem to converge much slower. However, as we keep increasing the number of triangles, suddenly the graph of the linear bins jumps to the graph of the logarithmic bins. This shows that if we would have trusted the linear binning for a smaller number of bins we would have gotten the wrong result. This is quite unexpected and alarming as it seems we cannot always trust linear binning! If we now change to the exact V_{123} we see that both linear and logarithmic binning converge much faster and both to the same value. In fact, it turns out that they reach one percent agreement for about 15.000 triangles. Let us try to understand why this happens. As we are summing over a function which peaks in the squeezed limit we do in fact get most signal from k -triplets which satisfy $k_3 \sim k_2 \gg k_1$. In particular the edge bins will contribute an important part to the final result. We know that precisely for these bins the approximate value for V_{123} does not work, which is probably why the results improve dramatically when using the exact V_{123} . Now one can still wonder why the linear binning performs so badly in this case. A reason might be that we are sampling the values for k_1 much better in case of logarithmic binning. However one could argue exactly the opposite, namely that linear binning samples the values of k_2 and k_3 much better. We have not found a convincing argument why linear binning fails, this remains an open question. As we are also studying the Fisher matrix for local PNG in our paper, we decided to stick to logarithmic binning. Even with the exact value of V_{123} the result converges quite slowly for the local function. In order to be within a couple of percent of the actual outcome of the Fisher analysis we need quite some triangles. For the analysis therefore we divide the k -axis over three logarithmic decades in 45 bins. By this we mean that if for instance each k_i from the triplet can take values in the range $[0.001, 1] h\text{Mpc}^{-1}$, it can take one of the 45 logarithmically separated values.

We did the same analysis for the equilateral function. In Figure 6.10b we see again a jump of the graph corresponding to linear binning. This time we do not expect to gain most signal from the edge bins. However, when we use the exact value for V_{123} everything seems to be fine again. The jump takes place at a comparable value of N_{bins} . The graph of the logarithmic binning remains a bit wiggly, but we find one percent agreement between linear and logarithmic binning already for 1000 triangles. For equilateral PNG we therefore divide the k -axis over three logarithmic decades in either 27 or 45 bins.

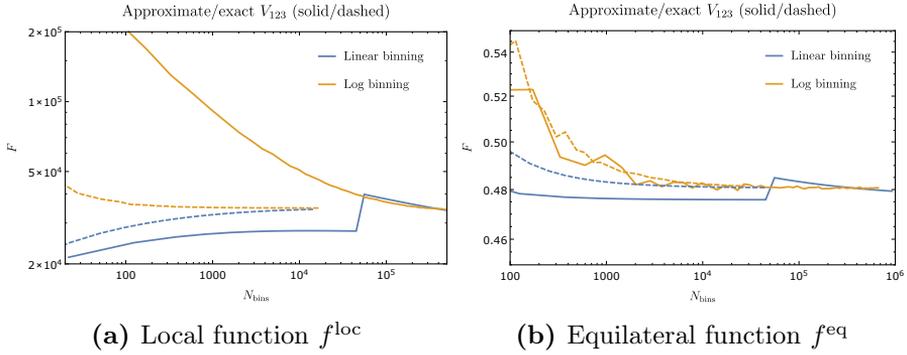


Figure 6.10: Computation of F as function of the number of triangles N_{bins} for (a) f^{loc} and (b) f^{eq} as given in equations (6.130) and (6.131). We show the results for both the approximate (solid) and exact (dashed) expression for V_{123} . Moreover we denote the results from linear binning with a blue line and logarithmic binning with a orange line.

6.D Table of parameters

Symbol	Relation	Meaning
a		scale factor
τ	$a \, d\tau = dt$	conformal time
\mathcal{H}	$\equiv d \ln(a)/d\tau$	conformal Hubble parameter
\mathcal{H}_0		present value of \mathcal{H}
\mathbf{x}		comoving coordinate
\mathbf{k}		momentum
Ω_m		matter density in units of the critical density
Ω_Λ		dark energy density
h		dimensionless Hubble constant
ρ		dark matter density
δ	$\equiv \delta\rho/\rho$	dark matter density contrast
θ	$\equiv \partial_i v^i$	velocity divergence
$\delta_{(n)}$		density contrast in SPT at order n
F_n		kernel function in $\delta_{(n)}$
P_{mn}	$\equiv \langle \delta_{(m)} \delta_{(n)} \rangle'$	power spectrum in SPT
B_{lmn}	$\equiv \langle \delta_{(l)} \delta_{(m)} \delta_{(n)} \rangle'$	bispectrum in SPT
ϕ		Newtonian potential
Φ	$\Delta \Phi = \delta$	rescaled Newtonian potential

Symbol	Relation	Meaning
φ	$\phi = T(k)\varphi$	primordial potential
φ_g		Gaussian primordial potential
$T(k)$		transfer function
$M(k)$		transfer function in the Poisson equation
D_1		linear growth factor
f	$\equiv d \ln D_1 / \ln a$	growth rate
P_φ		primordial power spectrum
B_φ		primordial bispectrum
n_s		scalar spectral index
ψ		correlation in the initial conditions
Ψ	$\Psi(\mathbf{x}) \equiv \psi(\mathbf{q}(\mathbf{x}))$	Eulerian definition of ψ
Δ		scaling dimension in K_{NL}
f_{NL}		amplitude of the primordial bispectrum
d^2	$\equiv c_s^2 + f(c_{vis}^2 + \hat{c}_{vis}^2)$	parameter in τ_v
$e_i, g,$ g_i		parameters in τ_v
ξ		parameter in $\delta_{(1)}^c$
γ		parameter in $\delta_{(1)}^c$
ϵ_i		parameter in $\delta_{(2)}^c$
γ_i		parameter in $\delta_{(2)}^c$
$\mathcal{S}_{\alpha,\beta}$		SPT quadratic source terms
τ_θ		EFT source in Euler equation
$\delta_{(n)}^c$		viscosity counterterm at order n
F_n^c		kernel function in $\delta_{(n)}^c$
H_n^c		kernel function in $\delta_{(n)}^c$
G_δ		Green's function for δ
\mathcal{D}_δ		evolution operator in the fluid equation
$P_{1\psi}$	$\equiv \langle \delta_{(1)} \psi \rangle'$	correlation of $\delta_{(1)}$ and ψ
$B_{\text{SPT}}^{\text{G}}$		Gaussian SPT contributions to B_δ
$B_{\text{SPT}}^{\text{NG}}$		non-Gaussian SPT contributions to B_δ
$B_{\text{EFT}}^{\text{G}}$		sum of Gaussian EFT counterterms
$B_{\text{EFT}}^{\text{NG}}$		sum of non-Gaussian EFT counterterms

Bibliografie

- [1] J.-P. Luminet, “Lemaitre’s Big Bang,” *PoS*, vol. FFP14, p. 214, 2016, 1503.08304.
- [2] J. Einasto, “Dark Matter,” *ArXiv e-prints*, Jan. 2009, 0901.0632.
- [3] A. D. Linde, “Particle physics and inflationary cosmology,” *Contemp. Concepts Phys.*, vol. 5, pp. 1–362, 1990, hep-th/0503203.
- [4] V. Mukhanov, *Physical Foundations of Cosmology*. Oxford: Cambridge University Press, 2005.
- [5] D. Baumann, “Inflation,” in *Physics of the large and the small, TASI 09, proceedings of the Theoretical Advanced Study Institute in Elementary Particle Physics, Boulder, Colorado, USA, 1-26 June 2009*, pp. 523–686, 2011, 0907.5424.
- [6] S. Weinberg, *Cosmology*. 2008.
- [7] A. Einstein, “The Foundation of the General Theory of Relativity,” *Annalen Phys.*, vol. 49, no. 7, pp. 769–822, 1916. [65(1916)].
- [8] A. Einstein, “Cosmological Considerations in the General Theory of Relativity,” *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, vol. 1917, pp. 142–152, 1917.
- [9] A. S. Eddington, “On the Instability of Einstein’s Spherical World,” *Mon. Not. Roy. Astron. Soc.*, vol. 90, pp. 668–678, 1930.
- [10] W. de Sitter, “Einstein’s theory of gravitation and its astronomical consequences,” *Mon. Not. Roy. Astron. Soc.*, vol. 78, pp. 3–28, 1917.
- [11] V. M. Slipher, “Nebulae,” *Proceedings of the American Philosophical Society*, vol. 56, pp. 403–409, 1917.

- [12] A. Friedmann, "On the curvature of space," *Z. Phys.*, vol. 10, p. 377–386, 1922. [Gen. Rel. Grav.31,1991 (1999)].
- [13] A. Friedmann, "On the Possibility of a world with constant negative curvature of space," *Z. Phys.*, vol. 21, pp. 326–332, 1924. [Gen. Rel. Grav.31,2001(1999)].
- [14] G. Lemaitre, "A Homogeneous Universe of Constant Mass and Growing Radius Accounting for the Radial Velocity of Extragalactic Nebulae," *Annales Soc. Sci. Bruxelles A*, vol. 47, pp. 49–59, 1927. [Gen. Rel. Grav.45,no.8,1635(2013)].
- [15] H. P. Robertson, "Kinematics and World-Structure," *Astrophys. J.*, vol. 82, pp. 284–301, 1935.
- [16] A. G. Walker, "On Milne's theory of world-structure," *Proceedings of the London Mathematical Society*, vol. 42, p. 90–127, 1937.
- [17] E. P. Hubble, "Extragalactic nebulae," *Astrophys. J.*, vol. 64, pp. 321–369, 1926.
- [18] A. S. Eddington, *The Mathematical Theory of Relativity*. 1924.
- [19] G. Lemaitre, "Note on de sitter's universe," *Journal of mathematics and physics*, vol. 4, no. 1-4, pp. 188–192, 1925.
- [20] E. Hubble, "A relation between distance and radial velocity among extragalactic nebulae," *Proc. Nat. Acad. Sci.*, vol. 15, pp. 168–173, 1929.
- [21] G. Lemaitre, "The expanding universe," *Gen. Rel. Grav.*, vol. 29, pp. 641–680, 1997. [Annales Soc. Sci. Bruxelles A53,51(1933)].
- [22] R. A. Alpher, H. Bethe, and G. Gamow, "The origin of chemical elements," *Phys. Rev.*, vol. 73, pp. 803–804, 1948.
- [23] R. H. Cyburt, B. D. Fields, and K. A. Olive, "Primordial nucleosynthesis in light of WMAP," *Phys.Lett.*, vol. B567, pp. 227–234, 2003, astro-ph/0302431.
- [24] A. Coc, J.-P. Uzan, and E. Vangioni, "Standard Big-Bang Nucleosynthesis after Planck," 2013, 1307.6955.
- [25] A. Coc, P. Petitjean, J.-P. Uzan, E. Vangioni, P. Descouvemont, C. Iliadis, and R. Longland, "New reaction rates for improved primordial D/H calculation and the cosmic evolution of deuterium," *Phys. Rev.*, vol. D92, no. 12, p. 123526, 2015, 1511.03843.

-
- [26] R. A. Alpher and H. Herman, “Evolution of the Universe,” *Nature*, vol. 162, pp. 774–775, 1948.
- [27] A. A. Penzias and R. W. Wilson, “A Measurement of excess antenna temperature at 4080-Mc/s,” *Astrophys. J.*, vol. 142, pp. 419–421, 1965.
- [28] J. C. Mather *et al.*, “Measurement of the Cosmic Microwave Background spectrum by the COBE FIRAS instrument,” *Astrophys. J.*, vol. 420, pp. 439–444, 1994.
- [29] G. F. Smoot, C. Bennett, A. Kogut, E. Wright, J. Aymon, *et al.*, “Structure in the COBE differential microwave radiometer first year maps,” *Astrophys. J.*, vol. 396, pp. L1–L5, 1992.
- [30] R. Adam *et al.*, “Planck 2015 results. I. Overview of products and scientific results,” *Astron. Astrophys.*, vol. 594, p. A1, 2016, 1502.01582.
- [31] H. W. Babcock, “The rotation of the Andromeda Nebula,” *Lick Observatory Bulletin*, vol. 19, pp. 41–51, 1939.
- [32] J. H. Oort, “Some Problems Concerning the Structure and Dynamics of the Galactic System and the Elliptical Nebulae NGC 3115 and 4494,” *ApJ*, vol. 91, p. 273, Apr. 1940.
- [33] P. J. E. Peebles, *Physical cosmology*. 1971.
- [34] Y. Parijskij, “Search for Primordial Perturbations of the Universe: Observations with Ratan-600 Radio Telescope,” *The Large Scale Structure of the Universe. International Astronomical Union*, vol. 79, 1978.
- [35] J. Kapteyn, “First Attempt at a Theory of the Arrangement and Motion of the Sidereal System,” *ApJ*, vol. 55, p. 302, May 1922.
- [36] J. H. Jeans, “The Motions of Stars in a Kapteyn Universe,” *MNRAS*, vol. 82, pp. 122–132, Jan. 1922.
- [37] F. Zwicky, “Die Rotverschiebung von extragalaktischen Nebeln,” *Helvetica Physica Acta*, vol. 6, pp. 110–127, 1933.
- [38] V. C. Rubin, W. K. Ford, Jr., and N. Thonnard, “Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 / $R = 4\text{kpc}/$ to UGC 2885 / $R = 122\text{ kpc}/$,” *ApJ*, vol. 238, pp. 471–487, June 1980.
- [39] G. R. Blumenthal, H. Pagels, and J. R. Primack, “Galaxy formation by dissipationless particles heavier than neutrinos,” *Nature*, vol. 299, p. 37, Sept. 1982.

- [40] J. R. Bond, A. S. Szalay, and M. S. Turner, "Formation of Galaxies in a Gravitino Dominated Universe," *Phys. Rev. Lett.*, vol. 48, p. 1636, 1982.
- [41] H. Pagels and J. R. Primack, "Supersymmetry, cosmology, and new physics at teraelectronvolt energies," *Physical Review Letters*, vol. 48, pp. 223–226, Jan. 1982.
- [42] P. J. E. Peebles, "Primeval adiabatic perturbations - Effect of massive neutrinos," *ApJ*, vol. 258, pp. 415–424, July 1982.
- [43] J. R. Bond and A. S. Szalay, "The collisionless damping of density fluctuations in an expanding universe," *ApJ*, vol. 274, pp. 443–468, Nov. 1983.
- [44] J. E. Gunn and B. M. Tinsley, "An accelerating Universe," *Nature*, vol. 257, pp. 454–457, Oct. 1975.
- [45] M. S. Turner, G. Steigman, and L. M. Krauss, "Flatness of the universe - Reconciling theoretical prejudices with observational data," *Physical Review Letters*, vol. 52, pp. 2090–2093, June 1984.
- [46] L. A. Kofman and A. A. Starobinskii, "Effect of the Cosmological Constant on Largescale Anisotropies in the Microwave Background," *Soviet Astronomy Letters*, vol. 11, pp. 271–274, Sept. 1985.
- [47] A. G. Riess *et al.*, "Observational evidence from supernovae for an accelerating universe and a cosmological constant," *Astron. J.*, vol. 116, pp. 1009–1038, 1998, astro-ph/9805201.
- [48] S. Perlmutter *et al.*, "Measurements of Omega and Lambda from 42 high redshift supernovae," *Astrophys. J.*, vol. 517, pp. 565–586, 1999, astro-ph/9812133.
- [49] D. N. Spergel *et al.*, "First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Determination of cosmological parameters," *Astrophys. J. Suppl.*, vol. 148, pp. 175–194, 2003, astro-ph/0302209.
- [50] A. R. Liddle, "An Introduction to cosmological inflation," in *Proceedings, Summer School in High-energy physics and cosmology: Trieste, Italy, June 29-July 17, 1998*, pp. 260–295, 1999, astro-ph/9901124.
- [51] A. H. Guth, "The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems," *Phys. Rev.*, vol. D23, pp. 347–356, 1981.

-
- [52] S. W. Hawking, I. G. Moss, and J. M. Stewart, “Bubble collisions in the very early universe,” *Phys. Rev. D*, vol. 26, pp. 2681–2693, Nov. 1982.
- [53] A. H. Guth and E. J. Weinberg, “Could the Universe Have Recovered from a Slow First Order Phase Transition?,” *Nucl. Phys.*, vol. B212, pp. 321–364, 1983.
- [54] A. D. Linde, “A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems,” *Phys. Lett.*, vol. 108B, pp. 389–393, 1982.
- [55] A. Albrecht and P. J. Steinhardt, “Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking,” *Phys. Rev. Lett.*, vol. 48, pp. 1220–1223, 1982.
- [56] V. F. Mukhanov and G. V. Chibisov, “Quantum Fluctuation and Non-singular Universe. (In Russian),” *JETP Lett.*, vol. 33, pp. 532–535, 1981.
- [57] A. A. Starobinsky, “Spectrum of relict gravitational radiation and the early state of the universe,” *JETP Lett.*, vol. 30, pp. 682–685, 1979.
- [58] S. Hawking, “The Development of Irregularities in a Single Bubble Inflationary Universe,” *Phys.Lett.*, vol. B115, p. 295, 1982.
- [59] A. A. Starobinsky, “Dynamics of Phase Transition in the New Inflationary Universe Scenario and Generation of Perturbations,” *Phys.Lett.*, vol. B117, pp. 175–178, 1982.
- [60] A. H. Guth and S. Pi, “Fluctuations in the New Inflationary Universe,” *Phys.Rev.Lett.*, vol. 49, pp. 1110–1113, 1982.
- [61] J. M. Bardeen, P. J. Steinhardt, and M. S. Turner, “Spontaneous Creation of Almost Scale - Free Density Perturbations in an Inflationary Universe,” *Phys. Rev.*, vol. D28, p. 679, 1983.
- [62] A. D. Linde, “Chaotic Inflation,” *Phys. Lett.*, vol. 129B, pp. 177–181, 1983.
- [63] P. A. R. Ade *et al.*, “Planck 2015 results. XIII. Cosmological parameters,” *Astron. Astrophys.*, vol. 594, p. A13, 2016, 1502.01589.
- [64] A. G. Riess *et al.*, “A 2.4% Determination of the Local Value of the Hubble Constant,” *Astrophys. J.*, vol. 826, no. 1, p. 56, 2016, 1604.01424.
- [65] T. M. C. Abbott *et al.*, “Dark Energy Survey Year 1 Results: Cosmological Constraints from Galaxy Clustering and Weak Lensing,” 2017, 1708.01530.

- [66] H. Hildebrandt *et al.*, “KiDS-450: Cosmological parameter constraints from tomographic weak gravitational lensing,” *Mon. Not. Roy. Astron. Soc.*, vol. 465, p. 1454, 2017, 1606.05338.
- [67] K. N. Abazajian *et al.*, “CMB-S4 Science Book, First Edition,” 2016, 1610.02743.
- [68] LSST Science Collaboration, P. A. Abell, J. Allison, S. F. Anderson, J. R. Andrew, J. R. P. Angel, L. Armus, D. Arnett, S. J. Asztalos, T. S. Axelrod, and *et al.*, “LSST Science Book, Version 2.0,” *ArXiv e-prints*, Dec. 2009, 0912.0201.
- [69] R. Laureijs, J. Amiaux, S. Arduini, J. . Auguères, J. Brinchmann, R. Cole, M. Cropper, C. Dabin, L. Duvet, A. Ealet, and *et al.*, “Euclid Definition Study Report,” *ArXiv e-prints*, Oct. 2011, 1110.3193.
- [70] D. Baumann, “Cosmology,” <http://www.damtp.cam.ac.uk/user/db275/Cosmology/Le>
- [71] W. E. East, M. Kleban, A. Linde, and L. Senatore, “Beginning inflation in an inhomogeneous universe,” *JCAP*, vol. 1609, no. 09, p. 010, 2016, 1511.05143.
- [72] K. Clough, E. A. Lim, B. S. DiNunno, W. Fischler, R. Flauger, and S. Paban, “Robustness of Inflation to Inhomogeneous Initial Conditions,” *JCAP*, vol. 1709, no. 09, p. 025, 2017, 1608.04408.
- [73] R. Brandenberger, “Initial conditions for inflation — A short review,” *Int. J. Mod. Phys.*, vol. D26, no. 01, p. 1740002, 2016, 1601.01918.
- [74] D. S. Salopek and J. R. Bond, “Nonlinear evolution of long wavelength metric fluctuations in inflationary models,” *Phys. Rev.*, vol. D42, pp. 3936–3962, 1990.
- [75] E. J. Copeland, E. W. Kolb, A. R. Liddle, and J. E. Lidsey, “Reconstructing the inflation potential, in principle and in practice,” *Phys. Rev.*, vol. D48, pp. 2529–2547, 1993, hep-ph/9303288.
- [76] A. R. Liddle, P. Parsons, and J. D. Barrow, “Formalizing the slow roll approximation in inflation,” *Phys. Rev.*, vol. D50, pp. 7222–7232, 1994, astro-ph/9408015.
- [77] A. G. Muslimov, “On the Scalar Field Dynamics in a Spatially Flat Friedman Universe,” *Class. Quant. Grav.*, vol. 7, pp. 231–237, 1990.
- [78] J. E. Lidsey, “The Scalar field as dynamical variable in inflation,” *Phys. Lett.*, vol. B273, pp. 42–46, 1991.

-
- [79] J. D. Barrow, “New types of inflationary universe,” *Phys. Rev.*, vol. D48, pp. 1585–1590, 1993.
- [80] J. D. Barrow, “Exact inflationary universes with potential minima,” *Phys. Rev.*, vol. D49, pp. 3055–3058, 1994.
- [81] J. M. Maldacena, “Non-Gaussian features of primordial fluctuations in single field inflationary models,” *JHEP*, vol. 05, p. 013, 2003, astro-ph/0210603.
- [82] W. H. Kinney, “Horizon crossing and inflation with large η ,” *Phys. Rev.*, vol. D72, p. 023515, 2005, gr-qc/0503017.
- [83] R. L. Arnowitt, S. Deser, and C. W. Misner, “The Dynamics of general relativity,” *Gen. Rel. Grav.*, vol. 40, pp. 1997–2027, 2008, gr-qc/0405109.
- [84] M. Sasaki, “Large Scale Quantum Fluctuations in the Inflationary Universe,” *Prog. Theor. Phys.*, vol. 76, p. 1036, 1986.
- [85] V. F. Mukhanov, “Quantum Theory of Gauge Invariant Cosmological Perturbations,” *Sov. Phys. JETP*, vol. 67, pp. 1297–1302, 1988. [Zh. Eksp. Teor. Fiz.94N7,1(1988)].
- [86] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger, “Theory of cosmological perturbations. Part 1. Classical perturbations. Part 2. Quantum theory of perturbations. Part 3. Extensions,” *Phys. Rept.*, vol. 215, pp. 203–333, 1992.
- [87] N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space*. Cambridge Monographs on Mathematical Physics, Cambridge, UK: Cambridge Univ. Press, 1984.
- [88] E. D. Stewart and D. H. Lyth, “A More accurate analytic calculation of the spectrum of cosmological perturbations produced during inflation,” *Phys. Lett.*, vol. B302, pp. 171–175, 1993, gr-qc/9302019.
- [89] D. H. Lyth, “What would we learn by detecting a gravitational wave signal in the cosmic microwave background anisotropy?,” *Phys. Rev. Lett.*, vol. 78, pp. 1861–1863, 1997, hep-ph/9606387.
- [90] C. Armendariz-Picon, T. Damour, and V. F. Mukhanov, “k - inflation,” *Phys. Lett.*, vol. B458, pp. 209–218, 1999, hep-th/9904075.
- [91] V. Acquaviva, N. Bartolo, S. Matarrese, and A. Riotto, “Second order cosmological perturbations from inflation,” *Nucl. Phys.*, vol. B667, pp. 119–148, 2003, astro-ph/0209156.

-
- [92] N. Bartolo, E. Komatsu, S. Matarrese, and A. Riotto, “Non-Gaussianity from inflation: Theory and observations,” *Phys. Rept.*, vol. 402, pp. 103–266, 2004, astro-ph/0406398.
- [93] X. Chen, “Primordial Non-Gaussianities from Inflation Models,” *Adv. Astron.*, vol. 2010, p. 638979, 2010, 1002.1416.
- [94] E. Komatsu and D. N. Spergel, “Acoustic signatures in the primary microwave background bispectrum,” *Phys. Rev.*, vol. D63, p. 063002, 2001, astro-ph/0005036.
- [95] N. Bartolo, S. Matarrese, and A. Riotto, “Nongaussianity from inflation,” *Phys. Rev.*, vol. D65, p. 103505, 2002, hep-ph/0112261.
- [96] D. Wands, “Local non-Gaussianity from inflation,” *Class. Quant. Grav.*, vol. 27, p. 124002, 2010, 1004.0818.
- [97] P. Creminelli, “On non-Gaussianities in single-field inflation,” *JCAP*, vol. 0310, p. 003, 2003, astro-ph/0306122.
- [98] X. Chen, M.-x. Huang, S. Kachru, and G. Shiu, “Observational signatures and non-Gaussianities of general single field inflation,” *JCAP*, vol. 0701, p. 002, 2007, hep-th/0605045.
- [99] C. Cheung, P. Creminelli, A. L. Fitzpatrick, J. Kaplan, and L. Senatore, “The Effective Field Theory of Inflation,” *JHEP*, vol. 03, p. 014, 2008, 0709.0293.
- [100] X. Chen and Y. Wang, “Quasi-Single Field Inflation and Non-Gaussianities,” *JCAP*, vol. 1004, p. 027, 2010, 0911.3380.
- [101] X. Chen and Y. Wang, “Large non-Gaussianities with Intermediate Shapes from Quasi-Single Field Inflation,” *Phys. Rev.*, vol. D81, p. 063511, 2010, 0909.0496.
- [102] R. Gwyn, G. A. Palma, M. Sakellariadou, and S. Sypsas, “Effective field theory of weakly coupled inflationary models,” *JCAP*, vol. 1304, p. 004, 2013, 1210.3020.
- [103] M. Alvarez *et al.*, “Testing Inflation with Large Scale Structure: Connecting Hopes with Reality,” 2014, 1412.4671.
- [104] D. Baumann, D. Green, and R. A. Porto, “B-modes and the Nature of Inflation,” *JCAP*, vol. 1501, no. 01, p. 016, 2015, 1407.2621.

-
- [105] P. Creminelli and M. Zaldarriaga, “Single field consistency relation for the 3-point function,” *JCAP*, vol. 0410, p. 006, 2004, astro-ph/0407059.
- [106] C. Cheung, A. L. Fitzpatrick, J. Kaplan, and L. Senatore, “On the consistency relation of the 3-point function in single field inflation,” *JCAP*, vol. 0802, p. 021, 2008, 0709.0295.
- [107] Y. Tada and V. Vennin, “Squeezed bispectrum in the δN formalism: local observer effect in field space,” *JCAP*, vol. 1702, no. 02, p. 021, 2017, 1609.08876.
- [108] T. Tanaka and Y. Urakawa, “Dominance of gauge artifact in the consistency relation for the primordial bispectrum,” *JCAP*, vol. 1105, p. 014, 2011, 1103.1251.
- [109] E. Pajer, F. Schmidt, and M. Zaldarriaga, “The Observed Squeezed Limit of Cosmological Three-Point Functions,” *Phys. Rev.*, vol. D88, no. 8, p. 083502, 2013, 1305.0824.
- [110] G. Cabass, E. Pajer, and F. Schmidt, “How Gaussian can our Universe be?,” *JCAP*, vol. 1701, no. 01, p. 003, 2017, 1612.00033.
- [111] S. Weinberg, “Adiabatic modes in cosmology,” *Phys. Rev.*, vol. D67, p. 123504, 2003, astro-ph/0302326.
- [112] S. Dodelson, “Coherent phase argument for inflation,” *AIP Conf. Proc.*, vol. 689, pp. 184–196, 2003, hep-ph/0309057. [,184(2003)].
- [113] W. Hu and S. Dodelson, “Cosmic microwave background anisotropies,” *Ann.Rev.Astron.Astrophys.*, vol. 40, pp. 171–216, 2002, astro-ph/0110414.
- [114] S. Dodelson, *Modern Cosmology*. Amsterdam: Academic Press, 2003.
- [115] P. A. R. Ade *et al.*, “Planck 2015 results. XX. Constraints on inflation,” *Astron. Astrophys.*, vol. 594, p. A20, 2016, 1502.02114.
- [116] A. R. Liddle and S. M. Leach, “How long before the end of inflation were observable perturbations produced?,” *Phys. Rev.*, vol. D68, p. 103503, 2003, astro-ph/0305263.
- [117] S. Weinberg, “Must cosmological perturbations remain non-adiabatic after multi-field inflation?,” *Phys. Rev.*, vol. D70, p. 083522, 2004, astro-ph/0405397.

- [118] P. A. R. Ade *et al.*, “Planck 2015 results. XVII. Constraints on primordial non-Gaussianity,” *Astron. Astrophys.*, vol. 594, p. A17, 2016, 1502.01592.
- [119] W. Hu and M. J. White, “A CMB polarization primer,” *New Astron.*, vol. 2, p. 323, 1997, astro-ph/9706147.
- [120] U. Seljak, “Measuring polarization in cosmic microwave background,” *Astrophys. J.*, vol. 482, p. 6, 1997, astro-ph/9608131.
- [121] M. Zaldarriaga and U. Seljak, “An all sky analysis of polarization in the microwave background,” *Phys.Rev.*, vol. D55, pp. 1830–1840, 1997, astro-ph/9609170.
- [122] M. Kamionkowski, A. Kosowsky, and A. Stebbins, “Statistics of cosmic microwave background polarization,” *Phys.Rev.*, vol. D55, pp. 7368–7388, 1997, astro-ph/9611125.
- [123] D. Baumann *et al.*, “CMBPol Mission Concept Study: Probing Inflation with CMB Polarization,” *AIP Conf. Proc.*, vol. 1141, pp. 10–120, 2009, 0811.3919.
- [124] E. Pajer and M. Zaldarriaga, “A New Window on Primordial non-Gaussianity,” *Phys. Rev. Lett.*, vol. 109, p. 021302, 2012, 1201.5375.
- [125] R. Emami, E. Dimastrogiovanni, J. Chluba, and M. Kamionkowski, “Probing the scale dependence of non-Gaussianity with spectral distortions of the cosmic microwave background,” *Phys. Rev.*, vol. D91, no. 12, p. 123531, 2015, 1504.00675.
- [126] W. Hu, D. Scott, and J. Silk, “Power spectrum constraints from spectral distortions in the cosmic microwave background,” *Astrophys. J.*, vol. 430, pp. L5–L8, 1994, astro-ph/9402045.
- [127] J. Chluba, R. Khatri, and R. A. Sunyaev, “CMB at 2x2 order: The dissipation of primordial acoustic waves and the observable part of the associated energy release,” *Mon. Not. Roy. Astron. Soc.*, vol. 425, pp. 1129–1169, 2012, 1202.0057.
- [128] G. Cabass, E. Pajer, and D. van der Woude, “Spectral distortion anisotropies from single-field inflation,” 2018, 1805.08775.
- [129] N. Dalal, O. Dore, D. Huterer, and A. Shirokov, “The imprints of primordial non-gaussianities on large-scale structure: scale dependent bias and abundance of virialized objects,” *Phys. Rev.*, vol. D77, p. 123514, 2008, 0710.4560.

-
- [130] O. Doré *et al.*, “Cosmology with the SPHEREX All-Sky Spectral Survey,” 2014, 1412.4872.
- [131] P. E. Dewdney, P. J. Hall, R. T. Schilizzi, and T. J. L. W. Lazio, “The Square Kilometre Array,” *IEEE Proceedings*, vol. 97, pp. 1482–1496, Aug. 2009.
- [132] Y.-C. Li and Y.-Z. Ma, “Constraints on Primordial non-Gaussianity from Future HI Intensity Mapping Experiments,” *Phys. Rev.*, vol. D96, no. 6, p. 063525, 2017, 1701.00221.
- [133] B. P. Abbott *et al.*, “Observation of Gravitational Waves from a Binary Black Hole Merger,” *Phys. Rev. Lett.*, vol. 116, no. 6, p. 061102, 2016, 1602.03837.
- [134] P. D. Lasky *et al.*, “Gravitational-wave cosmology across 29 decades in frequency,” *Phys. Rev.*, vol. X6, no. 1, p. 011035, 2016, 1511.05994.
- [135] M. C. Guzzetti, N. Bartolo, M. Liguori, and S. Matarrese, “Gravitational waves from inflation,” *Riv. Nuovo Cim.*, vol. 39, no. 9, pp. 399–495, 2016, 1605.01615.
- [136] B. Carr, F. Kuhnel, and M. Sandstad, “Primordial Black Holes as Dark Matter,” *Phys. Rev.*, vol. D94, no. 8, p. 083504, 2016, 1607.06077.
- [137] D. Baumann and L. McAllister, *Inflation and String Theory*. Cambridge Monographs on Mathematical Physics, Cambridge University Press, 2015, 1404.2601.
- [138] A. Achúcarro, V. Atal, C. Germani, and G. A. Palma, “Cumulative effects in inflation with ultra-light entropy modes,” *JCAP*, vol. 1702, no. 02, p. 013, 2017, 1607.08609.
- [139] A. Achúcarro, R. Kallosh, A. Linde, D.-G. Wang, and Y. Welling, “Universality of multi-field α -attractors,” *JCAP*, vol. 1804, no. 04, p. 028, 2018, 1711.09478.
- [140] A. Achúcarro, V. Atal, and Y. Welling, “On the viability of $m^2\phi^2$ and natural inflation,” *JCAP*, vol. 1507, p. 008, 2015, 1503.07486.
- [141] Y. Welling, D. van der Woude, and E. Pajer, “Lifting Primordial Non-Gaussianity Above the Noise,” *JCAP*, vol. 1608, no. 08, p. 044, 2016, 1605.06426.
- [142] Y. B. Zel’dovich, “Gravitational instability: An approximate theory for large density perturbations,” *A & A*, vol. 5, pp. 84–89, Mar. 1970.

- [143] P. J. E. Peebles and J. T. Yu, “Primeval Adiabatic Perturbation in an Expanding Universe,” *ApJ*, vol. 162, p. 815, Dec. 1970.
- [144] G. R. Blumenthal, S. M. Faber, J. R. Primack, and M. J. Rees, “Formation of galaxies and large-scale structure with cold dark matter,” *Nature*, vol. 311, pp. 517–525, Oct. 1984.
- [145] N. Kaiser, “Clustering in real space and in redshift space,” *Mon. Not. Roy. Astron. Soc.*, vol. 227, pp. 1–27, 1987.
- [146] T. Matsubara, “The Correlation function in redshift space: General formula with wide angle effects and cosmological distortions,” *Astrophys. J.*, vol. 535, p. 1, 2000, astro-ph/9908056.
- [147] V. Desjacques, D. Jeong, and F. Schmidt, “Large-Scale Galaxy Bias,” *Phys. Rept.*, vol. 733, pp. 1–193, 2018, 1611.09787.
- [148] D. J. Eisenstein *et al.*, “Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies,” *Astrophys. J.*, vol. 633, pp. 560–574, 2005, astro-ph/0501171.
- [149] F. Bernardeau, S. Colombi, E. Gaztanaga, and R. Scoccimarro, “Large scale structure of the universe and cosmological perturbation theory,” *Phys. Rept.*, vol. 367, pp. 1–248, 2002, astro-ph/0112551.
- [150] R. Laureijs *et al.*, “Euclid Definition Study Report,” 2011, 1110.3193.
- [151] M. Levi *et al.*, “The DESI Experiment, a whitepaper for Snowmass 2013,” 2013, 1308.0847.
- [152] K. S. Dawson *et al.*, “The SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Overview and Early Data,” *Astron. J.*, vol. 151, p. 44, 2016, 1508.04473.
- [153] M. Crocce and R. Scoccimarro, “Renormalized cosmological perturbation theory,” *Phys. Rev.*, vol. D73, p. 063519, 2006, astro-ph/0509418.
- [154] P. McDonald, “Dark matter clustering: a simple renormalization group approach,” *Phys. Rev.*, vol. D75, p. 043514, 2007, astro-ph/0606028.
- [155] S. Matarrese and M. Pietroni, “Resumming Cosmic Perturbations,” *JCAP*, vol. 0706, p. 026, 2007, astro-ph/0703563.
- [156] T. Matsubara, “Nonlinear perturbation theory with halo bias and redshift-space distortions via the Lagrangian picture,” *Phys. Rev.*, vol. D78, p. 083519, 2008, 0807.1733. [Erratum: *Phys. Rev.*D78,109901(2008)].

-
- [157] A. Taruya, F. Bernardeau, T. Nishimichi, and S. Codis, “RegPT: Direct and fast calculation of regularized cosmological power spectrum at two-loop order,” *Phys. Rev.*, vol. D86, p. 103528, 2012, 1208.1191.
- [158] J. Carlson, M. White, and N. Padmanabhan, “A critical look at cosmological perturbation theory techniques,” *Phys. Rev.*, vol. D80, p. 043531, 2009, 0905.0479.
- [159] D. Baumann, A. Nicolis, L. Senatore, and M. Zaldarriaga, “Cosmological Non-Linearities as an Effective Fluid,” *JCAP*, vol. 1207, p. 051, 2012, 1004.2488.
- [160] X. Wang, M. Neyrinck, I. Szapudi, A. Szalay, X. Chen, J. Lesgourgues, A. Riotto, and M. Sloth, “Perturbation Theory of the Cosmological Log-Density Field,” *Astrophys. J.*, vol. 735, p. 32, 2011, 1103.2166.
- [161] F. Bernardeau, N. van de Rijt, and F. Vernizzi, “Resummed propagators in multicomponent cosmic fluids with the eikonal approximation,” *PrD*, vol. 85, p. 063509, Mar. 2012, 1109.3400.
- [162] J. J. M. Carrasco, M. P. Hertzberg, and L. Senatore, “The Effective Field Theory of Cosmological Large Scale Structures,” *JHEP*, vol. 09, p. 082, 2012, 1206.2926.
- [163] M. Crocce, R. Scoccimarro, and F. Bernardeau, “MPTbreeze: A fast renormalized perturbative scheme,” *Mon. Not. Roy. Astron. Soc.*, vol. 427, p. 2537, 2012, 1207.1465.
- [164] D. Blas, S. Floerchinger, M. Garny, N. Tetradis, and U. A. Wiedemann, “Large scale structure from viscous dark matter,” *JCAP*, vol. 1511, p. 049, 2015, 1507.06665.
- [165] D. Blas, M. Garny, M. M. Ivanov, and S. Sibiryakov, “Time-Sliced Perturbation Theory for Large Scale Structure I: General Formalism,” *JCAP*, vol. 1607, no. 07, p. 052, 2016, 1512.05807.
- [166] S. Pueblas and R. Scoccimarro, “Generation of Vorticity and Velocity Dispersion by Orbit Crossing,” *Phys. Rev.*, vol. D80, p. 043504, 2009, 0809.4606.
- [167] P. Valageas, “Impact of shell crossing and scope of perturbative approaches, in real and redshift space,” *Astron. Astrophys.*, vol. 526, p. A67, Feb. 2011, 1009.0106.

- [168] M. Pietroni, G. Mangano, N. Saviano, and M. Viel, “Coarse-Grained Cosmological Perturbation Theory,” *JCAP*, vol. 1201, p. 019, 2012, 1108.5203.
- [169] J. M. Bardeen, “Gauge Invariant Cosmological Perturbations,” *Phys. Rev.*, vol. D22, pp. 1882–1905, 1980.
- [170] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, “The statistics of peaks of Gaussian random fields,” *ApJ*, vol. 304, pp. 15–61, May 1986.
- [171] D. J. Eisenstein and W. Hu, “Baryonic features in the matter transfer function,” *Astrophys. J.*, vol. 496, p. 605, 1998, astro-ph/9709112.
- [172] U. Seljak and M. Zaldarriaga, “A line of sight approach to cosmic microwave background anisotropies,” *Astrophys. J.*, vol. 469, pp. 437–444, 1996, astro-ph/9603033.
- [173] A. Lewis and S. Bridle, “Cosmological parameters from CMB and other data: A Monte Carlo approach,” *Phys. Rev.*, vol. D66, p. 103511, 2002, astro-ph/0205436.
- [174] P. J. E. Peebles, *The large-scale structure of the universe*. 1980.
- [175] E. Bertschinger, “Cosmological dynamics: Course 1,” in *Proceedings, Les Houches Summer School on Cosmology and Large Scale Structure (Session 60): Les Houches, France, August 1-28, 1993*, pp. 273–348, 1993, astro-ph/9503125.
- [176] C.-P. Ma and E. Bertschinger, “Cosmological perturbation theory in the synchronous and conformal Newtonian gauges,” *Astrophys. J.*, vol. 455, pp. 7–25, 1995, astro-ph/9506072.
- [177] M. P. Hertzberg, “Effective field theory of dark matter and structure formation: Semianalytical results,” *Phys. Rev.*, vol. D89, no. 4, p. 043521, 2014, 1208.0839.
- [178] L. Mercolli and E. Pajer, “On the velocity in the Effective Field Theory of Large Scale Structures,” *JCAP*, vol. 1403, p. 006, 2014, 1307.3220.
- [179] M. Mirbabayi, F. Schmidt, and M. Zaldarriaga, “Biased Tracers and Time Evolution,” *JCAP*, vol. 1507, no. 07, p. 030, 2015, 1412.5169.
- [180] J. J. M. Carrasco, S. Foreman, D. Green, and L. Senatore, “The 2-loop matter power spectrum and the IR-safe integrand,” *JCAP*, vol. 1407, p. 056, 2014, 1304.4946.

-
- [181] T. Baldauf, L. Mercolli, M. Mirbabayi, and E. Pajer, “The Bispectrum in the Effective Field Theory of Large Scale Structure,” *JCAP*, vol. 1505, no. 05, p. 007, 2015, 1406.4135.
- [182] R. E. Angulo, S. Foreman, M. Schmittfull, and L. Senatore, “The One-Loop Matter Bispectrum in the Effective Field Theory of Large Scale Structures,” *JCAP*, vol. 1510, no. 10, p. 039, 2015, 1406.4143.
- [183] V. Assassi, D. Baumann, E. Pajer, Y. Welling, and D. van der Woude, “Effective theory of large-scale structure with primordial non-Gaussianity,” *JCAP*, vol. 1511, p. 024, 2015, 1505.06668.
- [184] E. Pajer and M. Zaldarriaga, “On the Renormalization of the Effective Field Theory of Large Scale Structures,” *JCAP*, vol. 1308, p. 037, 2013, 1301.7182.
- [185] A. A. Abolhasani, M. Mirbabayi, and E. Pajer, “Systematic Renormalization of the Effective Theory of Large Scale Structure,” *JCAP*, vol. 1605, no. 05, p. 063, 2016, 1509.07886.
- [186] D. J. Heath, “The growth of density perturbations in zero pressure Friedmann-Lemaitre universes,” *MNRAS*, vol. 179, pp. 351–358, May 1977.
- [187] O. Lahav, P. B. Lilje, J. R. Primack, and M. J. Rees, “Dynamical effects of the cosmological constant,” *MNRAS*, vol. 251, pp. 128–136, July 1991.
- [188] S. M. Carroll, W. H. Press, and E. L. Turner, “The cosmological constant,” *Annual review of astronomy and astrophysics*, vol. 30, pp. 499–542, 1992.
- [189] M. H. Goroff, B. Grinstein, S.-J. Rey, and M. B. Wise, “Coupling of modes of cosmological mass density fluctuations,” *ApJ*, vol. 311, pp. 6–14, Dec. 1986.
- [190] R. Scoccimarro, S. Colombi, J. N. Fry, J. A. Frieman, E. Hivon, and A. Melott, “Nonlinear evolution of the bispectrum of cosmological perturbations,” *Astrophys. J.*, vol. 496, p. 586, 1998, astro-ph/9704075.
- [191] E. Sefusatti, “1-loop Perturbative Corrections to the Matter and Galaxy Bispectrum with non-Gaussian Initial Conditions,” *Phys. Rev.*, vol. D80, p. 123002, 2009, 0905.0717.
- [192] L. Senatore and M. Zaldarriaga, “The Effective Field Theory of Multi-field Inflation,” *JHEP*, vol. 04, p. 024, 2012, 1009.2093.

-
- [193] M. Alishahiha, E. Silverstein, and D. Tong, “DBI in the sky,” *Phys. Rev.*, vol. D70, p. 123505, 2004, hep-th/0404084.
- [194] N. Arkani-Hamed and J. Maldacena, “Cosmological Collider Physics,” 2015, 1503.08043.
- [195] H. Lee, D. Baumann, and G. L. Pimentel, “Non-Gaussianity as a Particle Detector,” *JHEP*, vol. 12, p. 040, 2016, 1607.03735.
- [196] A. Nicolis and F. Piazza, “Spontaneous Symmetry Probing,” *JHEP*, vol. 06, p. 025, 2012, 1112.5174.
- [197] I.-S. Yang, “The Strong Multifield Slowroll Condition and Spiral Inflation,” *Phys. Rev.*, vol. D85, p. 123532, 2012, 1202.3388.
- [198] S. Groot Nibbelink and B. J. W. van Tent, “Density perturbations arising from multiple field slow roll inflation,” 2000, hep-ph/0011325.
- [199] S. Groot Nibbelink and B. J. W. van Tent, “Scalar perturbations during multiple field slow-roll inflation,” *Class. Quant. Grav.*, vol. 19, pp. 613–640, 2002, hep-ph/0107272.
- [200] A. Achúcarro, J.-O. Gong, S. Hardeman, G. A. Palma, and S. P. Patil, “Mass hierarchies and non-decoupling in multi-scalar field dynamics,” *Phys. Rev.*, vol. D84, p. 043502, 2011, 1005.3848.
- [201] A. Achúcarro, J.-O. Gong, S. Hardeman, G. A. Palma, and S. P. Patil, “Features of heavy physics in the CMB power spectrum,” *JCAP*, vol. 1101, p. 030, 2011, 1010.3693.
- [202] C. Gordon, D. Wands, B. A. Bassett, and R. Maartens, “Adiabatic and entropy perturbations from inflation,” *Phys. Rev.*, vol. D63, p. 023506, 2001, astro-ph/0009131.
- [203] Z. Lalak, D. Langlois, S. Pokorski, and K. Turzyski, “Curvature and isocurvature perturbations in two-field inflation,” *JCAP*, vol. 0707, p. 014, 2007, 0704.0212.
- [204] C. M. Peterson and M. Tegmark, “Testing multifield inflation: A geometric approach,” *Phys. Rev.*, vol. D87, no. 10, p. 103507, 2013, 1111.0927.
- [205] Y. Welling, “Multiple Field Inflation and Signatures of Heavy Physics in the CMB,” Master’s thesis, Utrecht U., 2015, 1502.04369.
- [206] S. Cespedes and G. A. Palma, “Cosmic inflation in a landscape of heavy-fields,” *JCAP*, vol. 1310, p. 051, 2013, 1303.4703.

-
- [207] D. Seery and J. E. Lidsey, “Primordial non-Gaussianities from multiple-field inflation,” *JCAP*, vol. 0509, p. 011, 2005, astro-ph/0506056.
- [208] D. Langlois and F. Vernizzi, “A geometrical approach to nonlinear perturbations in relativistic cosmology,” *Class. Quant. Grav.*, vol. 27, p. 124007, 2010, 1003.3270.
- [209] J.-O. Gong and T. Tanaka, “A covariant approach to general field space metric in multi-field inflation,” *JCAP*, vol. 1103, p. 015, 2011, 1101.4809. [Erratum: *JCAP*1202,E01(2012)].
- [210] J. Elliston, D. Seery, and R. Tavakol, “The inflationary bispectrum with curved field-space,” *JCAP*, vol. 1211, p. 060, 2012, 1208.6011.
- [211] S. Cespedes, V. Atal, and G. A. Palma, “On the importance of heavy fields during inflation,” *JCAP*, vol. 1205, p. 008, 2012, 1201.4848.
- [212] A. Achucarro, J.-O. Gong, S. Hardeman, G. A. Palma, and S. P. Patil, “Effective theories of single field inflation when heavy fields matter,” *JHEP*, vol. 05, p. 066, 2012, 1201.6342.
- [213] A. Achucarro, V. Atal, S. Cespedes, J.-O. Gong, G. A. Palma, and S. P. Patil, “Heavy fields, reduced speeds of sound and decoupling during inflation,” *Phys. Rev.*, vol. D86, p. 121301, 2012, 1205.0710.
- [214] A. J. Tolley and M. Wyman, “The Gelaton Scenario: Equilateral non-Gaussianity from multi-field dynamics,” *Phys. Rev.*, vol. D81, p. 043502, 2010, 0910.1853.
- [215] S. Cremonini, Z. Lalak, and K. Turzynski, “Strongly Coupled Perturbations in Two-Field Inflationary Models,” *JCAP*, vol. 1103, p. 016, 2011, 1010.3021.
- [216] M. Dias, J. Frazer, D. J. Mulryne, and D. Seery, “Numerical evaluation of the bispectrum in multiple field inflation—the transport approach with code,” *JCAP*, vol. 1612, no. 12, p. 033, 2016, 1609.00379.
- [217] D. Seery, “CppTransport: a platform to automate calculation of inflationary correlation functions,” 2016, 1609.00380.
- [218] D. J. Mulryne and J. W. Ronayne, “PyTransport: A Python package for the calculation of inflationary correlation functions,” 2016, 1609.00381.
- [219] J. W. Ronayne and D. J. Mulryne, “Numerically evaluating the bispectrum in curved field-space— with PyTransport 2.0,” *JCAP*, vol. 1801, no. 01, p. 023, 2018, 1708.07130.

- [220] A. A. Starobinsky, “Multicomponent de Sitter (Inflationary) Stages and the Generation of Perturbations,” *JETP Lett.*, vol. 42, pp. 152–155, 1985. [Pisma Zh. Eksp. Teor. Fiz.42,124(1985)].
- [221] M. Sasaki and E. D. Stewart, “A General analytic formula for the spectral index of the density perturbations produced during inflation,” *Prog. Theor. Phys.*, vol. 95, pp. 71–78, 1996, astro-ph/9507001.
- [222] M. Sasaki and T. Tanaka, “Superhorizon scale dynamics of multiscalar inflation,” *Prog. Theor. Phys.*, vol. 99, pp. 763–782, 1998, gr-qc/9801017.
- [223] H.-C. Lee, M. Sasaki, E. D. Stewart, T. Tanaka, and S. Yokoyama, “A New delta N formalism for multi-component inflation,” *JCAP*, vol. 0510, p. 004, 2005, astro-ph/0506262.
- [224] A. Achúcarro, V. Vardanyan, D.-G. Wang, and Y. Welling, “In Preparation,” 2018, 18xx.xxxxx.
- [225] R. Kallosh and A. Linde, “Universality Class in Conformal Inflation,” *JCAP*, vol. 1307, p. 002, 2013, 1306.5220.
- [226] S. Ferrara, R. Kallosh, A. Linde, and M. Porrati, “Minimal Supergravity Models of Inflation,” *Phys. Rev.*, vol. D88, no. 8, p. 085038, 2013, 1307.7696.
- [227] R. Kallosh, A. Linde, and D. Roest, “Superconformal Inflationary α -Attractors,” *JHEP*, vol. 11, p. 198, 2013, 1311.0472.
- [228] S. Cecotti and R. Kallosh, “Cosmological Attractor Models and Higher Curvature Supergravity,” *JHEP*, vol. 05, p. 114, 2014, 1403.2932.
- [229] M. Galante, R. Kallosh, A. Linde, and D. Roest, “Unity of Cosmological Inflation Attractors,” *Phys. Rev. Lett.*, vol. 114, no. 14, p. 141302, 2015, 1412.3797.
- [230] R. Kallosh and A. Linde, “Escher in the Sky,” *Comptes Rendus Physique*, vol. 16, pp. 914–927, 2015, 1503.06785.
- [231] S. Ferrara and R. Kallosh, “Seven-Disk Manifold, α -attractors and B-modes,” 2016, 1610.04163.
- [232] R. Kallosh, A. Linde, T. Wrase, and Y. Yamada, “Maximal Supersymmetry and B-Mode Targets,” *JHEP*, vol. 04, p. 144, 2017, 1704.04829.
- [233] R. Kallosh, A. Linde, D. Roest, and Y. Yamada, “ $\overline{D3}$ induced geometric inflation,” *JHEP*, vol. 07, p. 057, 2017, 1705.09247.

-
- [234] N. Bartolo, S. Matarrese, and A. Riotto, “Adiabatic and isocurvature perturbations from inflation: Power spectra and consistency relations,” *Phys. Rev.*, vol. D64, p. 123504, 2001, astro-ph/0107502.
- [235] C. M. Peterson and M. Tegmark, “Testing Two-Field Inflation,” *Phys. Rev.*, vol. D83, p. 023522, 2011, 1005.4056.
- [236] Y. Welling, “Multiple Field Inflation and Signatures of Heavy Physics in the CMB,” 2015, 1502.04369.
- [237] T. Kobayashi and S. Mukohyama, “Effects of Light Fields During Inflation,” *Phys. Rev.*, vol. D81, p. 103504, 2010, 1003.0076.
- [238] S. Renaux-Petel and K. Turzyski, “On reaching the adiabatic limit in multi-field inflation,” *JCAP*, vol. 1506, no. 06, p. 010, 2015, 1405.6195.
- [239] S. Cremonini, Z. Lalak, and K. Turzyski, “On Non-Canonical Kinetic Terms and the Tilt of the Power Spectrum,” *Phys. Rev.*, vol. D82, p. 047301, 2010, 1005.4347.
- [240] C. van de Bruck and M. Robinson, “Power Spectra beyond the Slow Roll Approximation in Theories with Non-Canonical Kinetic Terms,” *JCAP*, vol. 1408, p. 024, 2014, 1404.7806.
- [241] R. Kallosh and A. Linde, “Multi-field Conformal Cosmological Attractors,” *JCAP*, vol. 1312, p. 006, 2013, 1309.2015.
- [242] A. A. Starobinsky, S. Tsujikawa, and J. Yokoyama, “Cosmological perturbations from multifield inflation in generalized Einstein theories,” *Nucl. Phys.*, vol. B610, pp. 383–410, 2001, astro-ph/0107555.
- [243] F. Di Marco, F. Finelli, and R. Brandenberger, “Adiabatic and isocurvature perturbations for multifield generalized Einstein models,” *Phys. Rev.*, vol. D67, p. 063512, 2003, astro-ph/0211276.
- [244] S. Renaux-Petel and K. Turzyski, “Geometrical Destabilization of Inflation,” *Phys. Rev. Lett.*, vol. 117, no. 14, p. 141301, 2016, 1510.01281.
- [245] J. Ellis, M. A. G. García, D. V. Nanopoulos, and K. A. Olive, “Two-Field Analysis of No-Scale Supergravity Inflation,” *JCAP*, vol. 1501, p. 010, 2015, 1409.8197.
- [246] A. R. Brown, “Hyperinflation,” 2017, 1705.03023.
- [247] S. Mizuno and S. Mukohyama, “Primordial perturbations from inflation with a hyperbolic field-space,” 2017, 1707.05125.

- [248] A. Achúcarro, O. Iarygina, G. A. Palma, D.-G. Wang, and Y. Welling, “In Preparation,” 2018, 18xx.xxxxx.
- [249] S. Ferrara, R. Kallosh, and A. Linde, “Cosmology with Nilpotent Superfields,” *JHEP*, vol. 10, p. 143, 2014, 1408.4096.
- [250] J. J. M. Carrasco, R. Kallosh, A. Linde, and D. Roest, “Hyperbolic geometry of cosmological attractors,” *Phys. Rev.*, vol. D92, no. 4, p. 041301, 2015, 1504.05557.
- [251] E. McDonough and M. Scalisi, “Inflation from Nilpotent Kähler Corrections,” *JCAP*, vol. 1611, no. 11, p. 028, 2016, 1609.00364.
- [252] A. Linde, D.-G. Wang, Y. Welling, Y. Yamada, and A. Achúcarro, “Hypernatural inflation,” 2018, 1803.09911.
- [253] M. Dias, J. Frazer, and D. Seery, “Computing observables in curved multifield models of inflation? A guide (with code) to the transport method,” *JCAP*, vol. 1512, no. 12, p. 030, 2015, 1502.03125.
- [254] Y. Yamada, “U(1) symmetric α -attractors,” 2018, 1802.04848.
- [255] K. Freese, J. A. Frieman, and A. V. Olinto, “Natural inflation with pseudo - Nambu-Goldstone bosons,” *Phys. Rev. Lett.*, vol. 65, pp. 3233–3236, 1990.
- [256] S. G. Rubin, “Effect of massive fields on inflation,” *JETP Lett.*, vol. 74, pp. 247–250, 2001, hep-ph/0110132. [Pisma Zh. Eksp. Teor. Fiz.74,275(2001)].
- [257] X. Dong, B. Horn, E. Silverstein, and A. Westphal, “Simple exercises to flatten your potential,” *Phys. Rev.*, vol. D84, p. 026011, 2011, 1011.4521.
- [258] S. Pi and M. Sasaki, “Curvature Perturbation Spectrum in Two-field Inflation with a Turning Trajectory,” *JCAP*, vol. 1210, p. 051, 2012, 1205.0161.
- [259] X. Chen and Y. Wang, “Quasi-Single Field Inflation with Large Mass,” *JCAP*, vol. 1209, p. 021, 2012, 1205.0160.
- [260] D. Baumann and D. Green, “Equilateral Non-Gaussianity and New Physics on the Horizon,” *JCAP*, vol. 1109, p. 014, 2011, 1102.5343.
- [261] G. Shiu and J. Xu, “Effective Field Theory and Decoupling in Multi-field Inflation: An Illustrative Case Study,” *Phys. Rev.*, vol. D84, p. 103509, 2011, 1108.0981.

-
- [262] A. Avgoustidis, S. Cremonini, A.-C. Davis, R. H. Ribeiro, K. Turzynski, and S. Watson, “Decoupling Survives Inflation: A Critical Look at Effective Field Theory Violations During Inflation,” *JCAP*, vol. 1206, p. 025, 2012, 1203.0016.
- [263] C. P. Burgess, M. W. Horbatsch, and S. Patil, “Inflating in a Trough: Single-Field Effective Theory from Multiple-Field Curved Valleys,” *JHEP*, vol. 01, p. 133, 2013, 1209.5701.
- [264] X. Gao, D. Langlois, and S. Mizuno, “Influence of heavy modes on perturbations in multiple field inflation,” *JCAP*, vol. 1210, p. 040, 2012, 1205.5275.
- [265] D. Baumann, D. Green, H. Lee, and R. A. Porto, “Signs of Analyticity in Single-Field Inflation,” *Phys. Rev.*, vol. D93, no. 2, p. 023523, 2016, 1502.07304.
- [266] J. McDonald, “Sub-Planckian Two-Field Inflation Consistent with the Lyth Bound,” *JCAP*, vol. 1409, no. 09, p. 027, 2014, 1404.4620.
- [267] G. Barenboim and W.-I. Park, “Spiral Inflation,” *Phys. Lett.*, vol. B741, pp. 252–255, 2015, 1412.2724.
- [268] T. Li, Z. Li, and D. V. Nanopoulos, “Helical Phase Inflation,” *Phys. Rev.*, vol. D91, no. 6, p. 061303, 2015, 1409.3267.
- [269] J. E. Kim, H. P. Nilles, and M. Peloso, “Completing natural inflation,” *JCAP*, vol. 0501, p. 005, 2005, hep-ph/0409138.
- [270] S. Dimopoulos, S. Kachru, J. McGreevy, and J. G. Wacker, “N-flation,” *JCAP*, vol. 0808, p. 003, 2008, hep-th/0507205.
- [271] R. Kappl, H. P. Nilles, and M. W. Winkler, “Natural Inflation and Low Energy Supersymmetry,” *Phys. Lett.*, vol. B746, pp. 15–21, 2015, 1503.01777.
- [272] T. Li, Z. Li, and D. V. Nanopoulos, “Symmetry Breaking Indication for Supergravity Inflation in Light of the Planck 2015,” *JCAP*, vol. 1509, no. 09, p. 006, 2015, 1502.05005.
- [273] C. Pallis, “Non-Minimally Gravity-Coupled Inflationary Models,” *Phys. Lett.*, vol. B692, pp. 287–296, 2010, 1002.4765.
- [274] R. Kallosh, A. Linde, and D. Roest, “Universal Attractor for Inflation at Strong Coupling,” *Phys. Rev. Lett.*, vol. 112, no. 1, p. 011303, 2014, 1310.3950.

- [275] A. Ashoorioon, K. Dimopoulos, M. M. Sheikh-Jabbari, and G. Shiu, “Reconciliation of High Energy Scale Models of Inflation with Planck,” *JCAP*, vol. 1402, p. 025, 2014, 1306.4914.
- [276] K. Kannike, G. Hütsi, L. Pizza, A. Racioppi, M. Raidal, A. Salvio, and A. Strumia, “Dynamically Induced Planck Scale and Inflation,” *JHEP*, vol. 05, p. 065, 2015, 1502.01334.
- [277] L. Boubekur, E. Giusarma, O. Mena, and H. Ramírez, “Does Current Data Prefer a Non-minimally Coupled Inflaton?,” *Phys. Rev.*, vol. D91, p. 103004, 2015, 1502.05193.
- [278] W. Buchmuller, E. Dudas, L. Heurtier, A. Westphal, C. Wieck, and M. W. Winkler, “Challenges for Large-Field Inflation and Moduli Stabilization,” *JHEP*, vol. 04, p. 058, 2015, 1501.05812.
- [279] L. McAllister, E. Silverstein, and A. Westphal, “Gravity Waves and Linear Inflation from Axion Monodromy,” *Phys. Rev.*, vol. D82, p. 046003, 2010, 0808.0706.
- [280] N. Kaloper and L. Sorbo, “A Natural Framework for Chaotic Inflation,” *Phys. Rev. Lett.*, vol. 102, p. 121301, 2009, 0811.1989.
- [281] N. Kaloper, A. Lawrence, and L. Sorbo, “An Ignoble Approach to Large Field Inflation,” *JCAP*, vol. 1103, p. 023, 2011, 1101.0026.
- [282] K. Harigaya and M. Ibe, “Simple realization of inflaton potential on a Riemann surface,” *Phys. Lett.*, vol. B738, pp. 301–304, 2014, 1404.3511.
- [283] G. A. Palma and A. Soto, “B-modes and the sound speed of primordial fluctuations,” *Phys. Rev.*, vol. D91, p. 063525, 2015, 1412.0033.
- [284] I. Zavala, “Effects of the speed of sound at large- N ,” *Phys. Rev.*, vol. D91, no. 6, p. 063005, 2015, 1412.3732.
- [285] R. Gwyn, G. A. Palma, M. Sakellariadou, and S. Sypsas, “On degenerate models of cosmic inflation,” *JCAP*, vol. 1410, no. 10, p. 005, 2014, 1406.1947.
- [286] T. Banks, M. Dine, P. J. Fox, and E. Gorbatov, “On the possibility of large axion decay constants,” *JCAP*, vol. 0306, p. 001, 2003, hep-th/0303252.
- [287] M. Tellarini, A. J. Ross, G. Tasinato, and D. Wands, “Galaxy bispectrum, primordial non-Gaussianity and redshift space distortions,” *JCAP*, vol. 1606, no. 06, p. 014, 2016, 1603.06814.

-
- [288] R. Scoccimarro, E. Sefusatti, and M. Zaldarriaga, “Probing primordial non-Gaussianity with large - scale structure,” *Phys. Rev.*, vol. D69, p. 103513, 2004, astro-ph/0312286.
- [289] E. Sefusatti and E. Komatsu, “The bispectrum of galaxies from high-redshift galaxy surveys: Primordial non-Gaussianity and non-linear galaxy bias,” *Phys. Rev.*, vol. D76, p. 083004, 2007, 0705.0343.
- [290] T. Baldauf, U. Seljak, and L. Senatore, “Primordial non-Gaussianity in the Bispectrum of the Halo Density Field,” *JCAP*, vol. 1104, p. 006, 2011, 1011.1513.
- [291] D. Jeong and E. Komatsu, “Primordial non-Gaussianity, scale-dependent bias, and the bispectrum of galaxies,” *Astrophys. J.*, vol. 703, pp. 1230–1248, 2009, 0904.0497.
- [292] G. Tasinato, M. Tellarini, A. J. Ross, and D. Wands, “Primordial non-Gaussianity in the bispectra of large-scale structure,” *JCAP*, vol. 1403, p. 032, 2014, 1310.7482.
- [293] N. Roth and C. Porciani, “Can we really measure f_{NL} from the galaxy power spectrum?,” *Mon. Not. Roy. Astron. Soc.*, vol. 425, pp. L81–L85, 2012, 1205.3165.
- [294] T. Baldauf, M. Mirbabayi, M. Simonović, and M. Zaldarriaga, “LSS constraints with controlled theoretical uncertainties,” 2016, 1602.00674.
- [295] E. Sefusatti, M. Crocce, S. Pueblas, and R. Scoccimarro, “Cosmology and the Bispectrum,” *Phys. Rev.*, vol. D74, p. 023522, 2006, astro-ph/0604505.
- [296] T. Baldauf, L. Mercolli, and M. Zaldarriaga, “Effective field theory of large scale structure at two loops: The apparent scale dependence of the speed of sound,” *Phys. Rev.*, vol. D92, no. 12, p. 123007, 2015, 1507.02256.
- [297] S. Foreman, H. Perrier, and L. Senatore, “Precision Comparison of the Power Spectrum in the EFTofLSS with Simulations,” *JCAP*, vol. 1605, no. 05, p. 027, 2016, 1507.05326.
- [298] J. J. M. Carrasco, S. Foreman, D. Green, and L. Senatore, “The Effective Field Theory of Large Scale Structures at Two Loops,” *JCAP*, vol. 1407, p. 057, 2014, 1310.0464.
- [299] A. Heavens, “Statistical techniques in cosmology,” 2009, 0906.0664.

- [300] C. E. Powell, “Generating Realisations of Stationary Gaussian Random Fields by Circulant Embedding.” https://www.nag.co.uk/doc/techrep/pdf/tr1_14.pdf.
- [301] K. S. Dawson *et al.*, “The Baryon Oscillation Spectroscopic Survey of SDSS-III,” *Astron. J.*, vol. 145, p. 10, 2013, 1208.0022.
- [302] M. Garny, T. Konstandin, R. A. Porto, and L. Sagunski, “On the Soft Limit of the Large Scale Structure Power Spectrum: UV Dependence,” *JCAP*, vol. 1511, no. 11, p. 032, 2015, 1508.06306.
- [303] E. Sefusatti and R. Scoccimarro, “Galaxy bias and halo-occupation numbers from large-scale clustering,” *Phys. Rev.*, vol. D71, p. 063001, 2005, astro-ph/0412626.
- [304] U. Seljak, “Extracting primordial non-gaussianity without cosmic variance,” *Phys. Rev. Lett.*, vol. 102, p. 021302, 2009, 0807.1770.
- [305] L. Verde and A. F. Heavens, “On the trispectrum as a Gaussian test for cosmology,” *Astrophys. J.*, vol. 553, p. 14, 2001, astro-ph/0101143.
- [306] D. Bertolini, K. Schutz, M. P. Solon, and K. M. Zurek, “The Trispectrum in the Effective Field Theory of Large Scale Structure,” *JCAP*, vol. 1606, no. 06, p. 052, 2016, 1604.01770.
- [307] A. Cooray, “21-cm Background Anisotropies Can Discern Primordial Non-Gaussianity,” *Phys. Rev. Lett.*, vol. 97, p. 261301, 2006, astro-ph/0610257.
- [308] K. Heitmann, M. White, C. Wagner, S. Habib, and D. Higdon, “The Coyote Universe I: Precision Determination of the Nonlinear Matter Power Spectrum,” *Astrophys. J.*, vol. 715, pp. 104–121, 2010, 0812.1052.
- [309] A. Schneider, R. Teyssier, D. Potter, J. Stadel, J. Onions, D. S. Reed, R. E. Smith, V. Springel, F. R. Pearce, and R. Scoccimarro, “Matter power spectrum and the challenge of percent accuracy,” *JCAP*, vol. 1604, no. 04, p. 047, 2016, 1503.05920.

Publications

- *Hypersnatural inflation*, A. Linde, D.-G. Wang, Y. Welling, Y. Yamada, and A. Achúcarro, JCAP **1807** (2018) 07, 035, (arXiv:1803.09911 [hep-th]).
- *Universality of multi-field α -attractors*, A. Achúcarro, R. Kallosh, A. Linde, D.-G. Wang, and Y. Welling, JCAP **1804** (2018) 04, 028, (arXiv:1711.09478 [hep-th]).
- *Lifting Primordial Non-Gaussianity Above the Noise*, Y. Welling, D. van der Woude, and E. Pajer, JCAP **1608** (2016) 08, 044, (arXiv:1605.06426 [astro-ph.CO]).
- *Effective Theory of Large-Scale Structure with Primordial Non-Gaussianity*, V. Assassi, D. Baumann, E. Pajer, Y. Welling, and D. van der Woude, JCAP **1511** (2015) 11, 024, (arXiv:1505.06668 [astro-ph.CO]).
- *On the viability of $m^2\phi^2$ and natural inflation*, A. Achúcarro, V. Atal, and Y. Welling, JCAP **1507** (2015) 07, 008, (arXiv:1503.07486 [astro-ph.CO]).

Samenvatting

Ik heb enorme mazzel dat ik natuurkundige ben, want puzzelen is één van mijn favoriete bezigheden. Samen met mijn kosmollega's bewandelen we duizenden wegen om een nieuw inzicht te vinden en soms is het raak. Dit is één van de meest vervullende gevoelens die ik ken, nog beter dan de kick die je krijgt als je drie uur lang een berg opfietst.

Uiteindelijk doen we dit werk natuurlijk ook om bij te dragen aan de Grote Speurtocht naar kennis en begrip van de bouwstenen waaruit wij en alles om ons heen zijn opgebouwd. Wie weet kan één van onze inzichten ooit gebruikt worden voor een grote stap voorwaarts in ons begrip van de fundamentele natuurkunde. Dat zou fantastisch zijn. Maar goed, voordat het zover is, moeten we eerst nog een hoop mini-puzzels oplossen. Verderop zal ik bespreken aan welke puzzels ik heb gewerkt.

Voordat ik dat doe, wil ik een gedachte delen die voor mij nog steeds verbluffend is: we gebruiken het allergrootste, de kosmos of het heelal waarvan de omvang nauwelijks te bevatten is, om iets te weten te komen over het allerkleinste, de elementaire deeltjes. We zijn in staat waarnemingen op kosmische schaal te relateren aan de deeltjessoep van het heelal, toen het nog jonger dan een miljardste van een miljardste van een miljardste seconde oud was. Best indrukwekkend!

Waar we nu staan is het resultaat van eeuwenlange inspanningen van sterrenkundigen, kosmologen, en theoretische natuurkundigen. Een zeer recente ontdekking is het bestaan van zwaartekrachtsgolven. Die kunnen we waarnemen wanneer twee zware objecten (veel zwaarder dan de aarde) met elkaar versmelten en een enorme hoeveelheid energie uitstoten. Tijdens de laatste fase van de versmelting is de zwaartekracht zo sterk dat we dit experiment onmogelijk kunnen nabootsen in een aards laboratorium. De kosmos leert ons zo meer over het gedrag van materie en krachten in extreme omstandigheden. Wij kunnen vervolgens onze fundamentele deeltjestheorieën hieraan toetsen

om tot een betere beschrijving te komen. Eenzelfde idee geldt voor de theorie van inflatie, waarmee ik me voornamelijk heb beziggehouden tijdens mijn promotieonderzoek. Hier ga ik verder op in op het einde van deze samenvatting, nadat ik eerst een paar belangrijke ideeën zal schetsen die aan mijn onderzoeksgebied ten grondslag liggen.

Inflatie

Waar komen we vandaan? Leven we op een speciale plek? Ik gok dat iedereen weleens op de één of andere manier over deze filosofische kwesties heeft nagedacht. In de kosmologie worden dit soort vragen ook gesteld, maar om een eenduidig antwoord te krijgen, moeten we deze filosofische gedachten eerst in een toetsbaar vraagstuk omzetten. We kunnen bijvoorbeeld de tweede vraag formuleren als: is het *Kosmologische Principe* correct? Het Kosmologische Principe zegt dat het heelal statistisch gezien homogeen en isotroop is op grote schaal. Dit betekent niet alleen dat wij niet op een speciale plek leven, maar ook dat er überhaupt geen speciale plekken bestaan in het heelal. Sinds Copernicus weten we dat de aarde niet het centrum van ons zonnestelsel is, en dat de zon meebeweegt met de ‘dans der sterren’. In hoeverre kunnen we het Kosmologische Principe verder testen? Gemeten kan worden (en is) hoeveel sterrenstelsels zich om ons heen bevinden en hoe ver die van ons verwijderd zijn. Het blijkt dat statistisch gezien de verdeling van waargenomen sterrenstelsels op grote schaal uniform is. Hiermee voldoet ons waarneembare heelal aan het Kosmologische Principe. Wat dat betreft leven we dus niet op een speciale plek, en buitenaards leven evenmin.

Het Kosmologische Principe is één van de belangrijkste hoekstenen van de kosmologie, omdat we hiermee experimenten kunnen herhalen door het waarneembare heelal op te delen in kleinere stukken. We gaan er bijvoorbeeld vanuit dat de onderlinge locaties van sterrenstelsels een bepaalde kansverdeling volgen. In ons heelal zijn de locaties van sterrenstelsels vastgelegd, en daarom vormt het heelal slechts één realisatie van die kansverdeling. Volgens het Kosmologische Principe kunnen we het heelal echter opdelen in kleinere delen die allemaal een betrouwbare steekproef bieden van dezelfde onderliggende kansverdeling. Het heelal fungeert daarom als een laboratorium waarin we metingen op diverse locaties kunnen opvatten als een ensemble van realisaties van (sub)heelallen.

Een ander vraagstuk dat een belangrijke rol heeft gespeeld in de ontwikkeling van de kosmologie is of het heelal krimpt, even groot blijft, of uitdijt.

We zouden ons het heelal voor kunnen stellen als een ballon met daarop vastgeplakte stukjes papier: de sterrenstelsels. Aangenomen dat de sterrenstelsels zichzelf niet over de ballon verplaatsen, wordt de afstand tussen hen groter of kleiner als je de ballon verder opblaast of juist een beetje leeg laat lopen. Als ons melkwegstelsel één van die papiertjes is, dan zien we de andere sterrenstelsels dus allemaal van ons vandaan bewegen of allemaal op ons afkomen. Om de snelheid van de verwegelegen sterrenstelsels in onze kijkrichting te bepalen, maken astronomen gebruik van het verschijnsel *kosmische roodverschuiving*. Net zoals de verandering in de toonhoogte die je hoort als er een ambulance voorbij scheurt (Doppler effect), verschuift de waargenomen kleur van een sterrenstelsel naar het rood als het heelal uitzet, danwel naar het blauw als het heelal inkrimpt. Het gevolg van de rood(blauw)verschuiving is een af(toe)name in de helderheid van het sterrenstelsel.

Zonder de meting te doen, hebben we al een aanwijzing voor het juiste antwoord op de vraag of de grootte van het heelal verandert. Laten we aannemen dat het heelal ontzettend groot is en lang genoeg bestaat, en dat alle sterren uniform verdeeld zijn over het heelal. Als er dan geen roodverschuiving plaatsvindt, is er in elke kijkrichting wel een ster te vinden met min of meer dezelfde helderheid. Oftewel, we hebben geen verschil tussen dag en nacht (en we verbranden allemaal, maar dat terzijde.) Dit is de paradox van Olbers. Een voor de hand liggende oplossing is dat het heelal uitzet, en er dus kosmische roodverschuiving plaatsvindt, waardoor de helderheden van sterren afnemen met de afstand. Een andere mogelijke uitweg is dat het heelal nog maar kort bestaat en we nog niet al het licht van ver verwijderde sterren hebben kunnen ontvangen. Inmiddels hebben astronomen de kosmische roodverschuiving van miljoenen sterrenstelsels waargenomen, en daarmee is de uitdijning van het heelal een feit.

Stel dat we teruggaan in de tijd zodat het heelal juist inkrimpt. Er treedt dan blauwverschuiving op van alle straling. Oftewel, hoe jonger het heelal is, des te dichter en heter het is. Bij temperaturen hoger dan zo'n $3000K$ is het te heet voor de elektronen om zich te binden aan de atoomkernen. Licht (fotonen) gaat dan voortdurend interacties aan met de vrije elektronen. Het vroege universum was een ondoorzichtige soep van deeltjes tot op het moment dat neutraal waterstof gevormd kon worden en het licht overal in de ruimte vrij kon reizen. Uit deze gedachtengang voorspelden Alpher en Hermann dat we nu nog steeds een gloed zouden moeten zien van het moment dat het heelal doorzichtig werd. De temperatuur was toen $3000K$, ongeveer de helft van de oppervlaktetemperatuur van de zon. Inmiddels is het heelal zo'n duizendmaal

groter¹⁴ en kouder geworden en zouden we nu van alle kanten straling moeten ontvangen met een temperatuur van ongeveer $3K$. Deze straling is inderdaad in 1965 door Penzias en Wilson voor het eerst waargenomen en staat nu bekend als de *kosmische achtergrondstraling*.

De kosmische achtergrondstraling is in feite een babyfoto van het heelal. Het heelal was toen nog maar 380.000 jaar oud, slechts een fractie van zijn huidige leeftijd van 13,8 miljard jaar. Rond die tijd waren er nog geen moleculen, laat staan sterrenstelsels. Nog verder terug in de tijd, toen het heelal nog maar een seconde oud was en een temperatuur had van tien miljard Kelvin, waren er de perfecte omstandigheden om een aantal lichte, chemische isotopen (onder andere van waterstof en helium) te vormen, de *nucleosynthese*. De verhouding waarmee de verschillende isotopen volgens berekeningen van Gamow gevormd konden worden, klopt zeer goed met de huidige verdeling in de sterrenstelsels. De ontdekking van de kosmische achtergrondstraling en de juiste voorspelling van de nucleosynthese waren een grote triomf voor de zogenaamde *hete oerknaltheorie*.

Maar met de nauwkeurigere meting van de kosmische achtergrondstraling kwam er iets gekks aan het licht: de temperatuur van de straling is wel héél uniform verdeeld! Ten tijde van het doorzichtig worden was het heelal nog zo jong dat er zo'n duizend delen aan te wijzen waren die niet met elkaar in causaal contact stonden. Dat wil zeggen dat die delen vanwege onderlinge afstand en de eindigheid van de lichtsnelheid nooit via uitwisseling van hitte een thermisch evenwicht konden bereiken. Hoe kan het dan dat deze delen allemaal ongeveer even heet blijken te zijn geweest? Simpelweg aannemen dat de begincondities van alle afzonderlijke delen van het heelal nou eenmaal hetzelfde waren, is niet echt bevredigend. Als wetenschappers zoeken we liever naar een verklaring. Uiteindelijk stelde Guth voor dat het vroege heelal een periode van exponentieel snelle uitdijning heeft gekend, een periode van *inflatie*. Hiermee faalt de eerdere berekening van de grootte van de causale delen, ze zijn namelijk exponentieel groter. Het heelal dat we nu kunnen zien bevindt zich in één opgeblazen deel.

De vraag is nu wat voor een soort materie we nodig hebben om inflatie op gang te brengen. Kunnen we misschien materie gebruiken die we al kennen vanuit de deeltjesfysica? Of spelen nog onbekende deeltjes een rol? De energieën van deeltjes in het vroege heelal liggen ver voorbij het bereik van deeltjesversnellers zoals de LHC van CERN. Inflatie biedt een enorme kans om

¹⁴Preciezer gezegd is het heelal duizendmaal groter geworden in elke ruimtelijke richting, dus het volume is met een factor miljard gegroeid.

nieuwe deeltjes te ontdekken en hun gedrag te beschrijven en daarmee meer kennis over de fundamentele natuurkunde te vergaren!

De onbekende materie die inflatie veroorzaakt, noemen we de *inflaton*. De inflaton fungeert als een klok, bij elke slag van de wijzer verdubbelt het heelal in grootte. Dit gaat door tot het einde van inflatie. Vanwege kwantumeffecten is er een kleine onzekerheid in de klok waardoor sommige delen van het heelal iets later en andere delen juist iets eerder de periode van inflatie verlaten. Dit resulteert in minuscule dichtheidsvariaties van materie in het heelal. Door de zwaartekracht bundelen deze kleine materiehoopjes zich tot alsmaar grotere structuren, en vormen uiteindelijk de sterrenstelsels en het netwerk van donkere materie daaromheen. Inflatie creëert dus de zaadjes van alles wat we om ons heen zien! Omdat de vorming van deze zaadjes een kansproces is dat de wetten van de kwantummechanica volgt, kunnen we alleen statistische voorspellingen maken. Gegeven de eigenschappen van de inflaton, kunnen we bijvoorbeeld de verdeling van sterrenstelsels voorspellen. Inflatie voorspelt verder ook kleine temperatuurvariaties in de kosmische achtergrondstraling van éénduizendste procent. Deze zijn inmiddels met grote precisie vastgelegd en de voorspelde eigenschappen van de bijbehorende kansverdeling zijn ook daadwerkelijk waargenomen. De volgende hint waar iedereen op hoopt zijn kleine afwijkingen in de kansverdeling die ons iets kunnen vertellen over de gedetailleerde eigenschappen van de inflaton en zijn interactie met andere deeltjes.

Zo zie je dus dat het heelal op de allergrootste schaal (verdeling van sterrenstelsels en kosmische achtergrondstraling) informatie bevat over de periode van inflatie, en daarmee gekoppeld is aan de allerkleinste schaal (elementaire deeltjes).

Dit proefschrift

Tijdens mijn promotieonderzoek heb ik nagedacht over inflatiemodellen waarin de inflaton interacties aangaat met een ander deeltje zonder spin. Dit is gemotiveerd vanuit de snaartheorie, die de aanwezigheid van veel van dat soort deeltjes voorspelt. De uiteindelijke voorspellingen hangen onder andere af van de massa van het nieuwe deeltje en de koppelingsconstante tussen de inflaton en het nieuwe deeltje. Theoretisch gezien is een belangrijke vraag of je met typische ingrediënten uit de snaartheorie een werkend model voor inflatie krijgt, en of bovendien de voorspellingen overeenstemmen met de waarnemingen tot nu toe. Aan de andere kant is het voor een experimentalist interessanter om te bepalen welke observabelen ons mogelijk iets kunnen vertellen over de aan-

wezigheid, en zo ja, de massa en koppelingsconstante van nieuwe deeltjes. In dit proefschrift heb ik me met beide vraagstukken beziggehouden.

Exacte modellen

In Hoofdstuk 2 en 3 construeren we inflatiemodellen waarin we de koppelingsconstante en massa van het nieuwe deeltje precies kunnen kiezen zoals we willen. In inflatie is dit niet op een heel eenvoudige manier te realiseren, en naar ons weten is dit nooit eerder gedaan.

Allereerst levert dit nieuwe inzichten op over ‘toegestane structuren’ in het deeltjesmodel van inflatie. In het bijzonder blijkt de combinatie kleine massa en grote koppelingsconstante (efficiënte overdracht van energie) ook compatibel met de waarnemingen. Dit opent mogelijk nieuwe wegen om inflatie te beschrijven in snaartheorie. We laten zien welke vorm van de ‘potentiaal’ nodig is om de juiste combinatie van parameters te garanderen. Als snaartheoretici een model vinden dat hierop lijkt, dan weten we of de voorspellingen consistent zijn met de data. Ten tweede kunnen we met deze exacte modellen onze analytische voorspellingen toetsen, aangezien we het model nu ook in kunnen voeren in de computer. We vinden dat de massa bepaalt hoe de ‘spectrale richtingscoëfficiënt’ afhangt van de koppelingsconstante als de energieoverdracht voldoende efficiënt is. Deze observabele zou daarom kunnen helpen om de massa van het andere deeltje te bepalen.

Universaliteit door hyperbolische kromming

In hoofdstuk 4 bestuderen we het effect van een veelvoorkomend element uit de snaartheorie, namelijk een hyperbolisch gekromde veldenruimte. Het komt erop neer dat hoe vroeger we ons in de periode van inflatie bevinden, des te meer de beweging van de deeltjes geremd wordt.

We vinden dat deze remming zowel de koppelingsconstante als de massa van het andere deeltje beïnvloedt en dat de voorspellingen niet meer afhangen van de vorm van de potentiaal. Er treedt universeel gedrag op zolang de hyperbolische kromming sterk genoeg is. Een hyperbolisch gekromde veldenruimte zou daarom een belangrijke rol kunnen spelen in het realiseren van inflatie binnen de snaartheorie.

Zware deeltjes

In Hoofdstuk 5 illustreren we de impact van zware deeltjes als ze interacties aangaan met de inflaton met een simpele aanpassing van drie inflatiemodellen. Het belangrijkste waarneembare effect is dat de 'tensor-to-scalar-ratio' kleiner wordt.

Driehoeksmetingen

Ten slotte onderzoeken we in Hoofdstuk 6 of we bepaalde interacties van de inflaton met een ander deeltje kunnen meten in de volgende generatie experimenten die sterrenstelsels in kaart brengen. Deze experimenten noemen we ook wel surveys van het kosmische web. Het blijkt dat, vanaf een bepaalde grootte van de massa van het andere deeltje, sterrenstelsels een voorkeur hebben om patronen te vormen die opgebouwd zijn uit gelijkzijdige driehoeken. Bij kleine massa van het andere deeltje, zullen ze juist gelijkbenige driehoeken willen vormen waarbij de derde zijde vele malen korter is dan de gelijke zijden. Door hier een statistische analyse op los te laten kunnen we zien hoe goed dit te onderscheiden is van de driehoeken die de zwaartekracht zelf al neigt te vormen tijdens het formatieproces van alle structuren op kosmische schalen. Dit resulteert uiteindelijk in een verwachting van de grootte van de foutenmarges die we met de foutenmarges van de kosmische achtergrondstraling kunnen vergelijken.

We vinden dat het type interactie dat geassocieerd is met gelijkzijdige driehoeken nauwkeuriger bepaald kan worden met behulp van de kosmische achtergrondstraling dan met de verdeling van sterrenstelsels. De volgende generatie surveys van het kosmische web zal helaas de huidige foutenmarges niet kunnen verkleinen. De mate waarin we het andere type interactie kunnen bepalen is veelbelovend, maar zal uiteindelijk met een andere techniek makkelijker gemeten kunnen worden. In een voorgaand onderzoek hebben we een theoretische methode van het formatieproces van grote structuren verder ontwikkeld. In dit hoofdstuk laten we verder zien dat deze methode de foutenmarges met een factor drie vermindert.

Curriculum Vitae

I was born on March 3th, 1989 in Almelo, the Netherlands. In 2007 I graduated from high school at the "OSG Erasmus". In the same year I won the Dutch Physics Olympiad, which allowed me to represent the Netherlands at the International Physics Olympiad in Isfahan, Iran.

In 2010 I completed the TWIN physics and mathematics bachelor studies at Utrecht University. Besides that I also participated in a honors program on experimental physics and I wrote my thesis on the mechanical properties of vimentin under the supervision of Prof. Gijsje Koenderink at Amolf in Amsterdam. From 2011 to 2014 I studied theoretical physics at Utrecht University, and for my final research project I specialized in cosmology by writing a master thesis on multi-field inflation under the supervision of Prof. Ana Achúcarro and Dr. Tomislav Prokopec.

In 2014 I was awarded the doctoral de Sitter fellowship at Leiden University. During my PhD I was supervised by Prof. Ana Achúcarro and Dr. Enrico Pajer. The work I did with Prof. Ana Achúcarro and her group resulted in three scientific publications and three in preparation, forming the first part of this thesis. The collaboration with Dr. Enrico Pajer has led to two scientific publications, that constitute the second part of this thesis.

From Oktober 2018 I will be a postdoctoral researcher at DESY in Hamburg in the group of Dr. Alexander Westphal.

Acknowledgements

Ana, I am extremely lucky with you as my advisor, in particular because of the endless time you dedicated to our discussions. Without doubt these discussions led to the ideas that seeded the main part of this thesis. It was very exciting to work with you and to learn from your great physical insights and strong values. Also, you are the best coach I could have wished for to guide me through all the highs and lows of scientific puberty. I am grateful that you ‘put your finger on the wound’ (and poke it inside) when I was lost and frustrated. Thank you for your open-mindedness and letting me develop into the scientist I want to be.

Enrico, thanks for embedding me in your group during your time in Utrecht. The discussions and journal clubs we had were a steep learning curve for me and made me feel very excited. You always expressed your thoughts, opinions and vision very clearly. Moreover, you showed me time after time your very structured way of approaching any problem. I am grateful to both you and Ana for getting in touch with many great physicists around the world and for all the opportunities to speak at workshops and conferences.

I am indebted to Eric Eliel, Koenraad Schalm, Daniel Baumann, Gonzalo Palma and Alexander Westphal for carefully reading the manuscript of my thesis. Moreover, I would like to thank Martijn, Marian, Marleen and Valentijn for their valuable feedback on the Dutch summary.

I am grateful to my collaborators Andrei, Daniel, Drian, Ed, Evangelos, Gang, Gonzalo, Irene, Miguel, Oksana, Renata, Thomas, Valentin, Valeri, Vicente and Yusuke for the valuable insights you gave me and for sharing your enthusiasm with me. Further, I am happy to thank all my cosmoleagues in Leiden and Utrecht for creating a very pleasant working atmosphere. Thanks to Anita, Barry, Els, Evelijn, Erik, Fran, Marianne, Marjan, Monica and Wanda for taking care of everything so smoothly that I didn’t even notice.

Ik wil mijn vrienden en sportmaatjes bedanken voor het bieden van zoveel plezier en de nodige afleiding van mijn werk. Aafke, Tessa, Valentijn en Wilke bedankt voor de lol en avonturen die we samen beleefd hebben, en jullie luisterend oor en de diepe gesprekken die we gevoerd hebben. Jullie staan altijd voor me klaar.

Mama, bedankt voor je onbegrensde liefde en de steun voor het najagen van mijn dromen. Jouw wijsheid en gekkigheid hebben me gevormd tot wie ik ben. Jasper, mijn lieve broer, bedankt voor het delen van de mooiste momenten van het leven en je geruststellende onbezorgdheid. Reinier, grote dank voor het aanwakkeren en motiveren van mijn interesse in natuurkunde. Piet, Lieke, Marco, Margreet, Angelo, Marleen, Wietske, Henk, Leon, Marcel, Roos, Sanne, Tijmen, Simon, Richelle en Wendy, jullie zijn de meest fantastische familie die ik me kan wensen. Papa, mijn mooiste jeugdherinnering is dat we samen raadsels en wiskunde problemen oplossen. Jij bent altijd bij me. Dit proefschrift is voor jou.

Martijn, je hebt me laten opbloeien met jouw liefde en steun, en met de lol die we elke dag samen beleven. Dankje voor je incasseringvermogen en begrip op de onmogelijke tijden dat ik met mijn werk bezig was. Wat heb ik een geluk met jou, en dat je alles achterlaat om samen ons volgende avontuur in Hamburg aan te gaan.