Phenomenology of Particle Production during Inflation

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Ryo Namba

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Prof. Marco Peloso, Adviser

September 2013

Acknowledgments

First and foremost, I would like to express my ineffable gratitude to my adviser, Marco Peloso, for his continuous guidance, instructions, and passion for research, from which I have learned so much.

I am also extremely grateful to A. Emir Gümrükçüoğlu for having mentored me at my earliest stages of research in the field, and for his continuous friendship and generosity.

One of my first collaborators, Neil Barnaby, has given me countless insights into physics and cosmology, and I appreciate his thoughtful discussions on many different subjects.

In the course of my graduate study, I had the opportunity to collaborate with several people. I am thankful to Gary Shiu for invaluable discussions and support of my future career, Shinji Mukohyama for valuable correspondence and his hospitality during my visit to the Kavli IPMU in Japan, and Emanuela Dimastrogiovanni, Peng Zhou and Jordan Moxon for exciting collaborations.

Importantly, I acknowledge S. G. Crowder and V. Mandic at the University of Minnesota for performing the data analysis and forecast in the study of the polarized stochastic gravitaional wave background signals presented in Section IV 4 of this thesis.

I thank my close friends in the department, N. Joe Pastika, who has been my roommate for the last two years, Alexey Finkel, David Tran, Michael Schecter, Adam Schreckenberger, Tianran Chen, Kevin Christie, Chad Geppert, and Pamela Vo, and those outside, Ayako Carter, Katsuya Okada, Yoshikazu Suzuki-Lambrecht, and Yuka Miyazato, for their company, without which I would not have survived the last five years.

Finally, I would like to thank my family and relatives for their bottomless support during my entire life.

Dedication

To my mother, whose understanding and support have led me this far in academic pursuit, and whose way of living has always inspired me.

Abstract

This thesis is devoted to the study on particle production during the era of primordial inflation and its phenomenological impacts. The simplest models of inflation typically assume only one dynamical degree of freedom, the inflaton, that is responsible for all the inflationary dynamics and predictions. Yet, it is a natural expectation that the inflaton should be coupled to some other fields, for successful reheating of the universe after inflation.

We first consider the models in which the inflaton is coupled to a U(1) gauge field. For a pseudo-scalar inflaton, its natural coupling induces tachyonic growth of the gauge quanta, which then inverse-decay to the inflaton perturbations. This imprints non-Gaussianity in the cosmic microwave background (CMB) anisotropies. This non-Gaussianity has a nearly equilateral shape, and the fact that we have not observed it with Planck provides a bound on the axion decay constant ($f \gtrsim 10^{16}$ GeV), which is in the range naturally obtained in UV complete theories. The produced gauge quanta also source gravitational waves (GWs). Future GW interferometer experiments can improve over the CMB non-Gaussianity limits.

We then study a different model characterized by a scalar inflaton coupled to gauge fields via a dilation-like interaction. This coupling can result in a nearly scale-invariant spectrum for the gauge field. Also in this case, the produced gauge quanta source inflaton perturbations, but the resulting non-Gaussianity now has a shape peaked for squeezed triangles, and which exhibits a peculiar angular dependence, that, if detected, would be a smoking gun of the higher-spin fields involved.

In the above two models, the GW signals are always subdominant at the CMB scales, due to the non-Gaussianity bounds from the scalar perturbations (namely, from the perturbations generated by the inflaton quanta produced by the gauge fields). We study the radically different situation in which some field other than the inflaton produces the gauge quanta, and these quanta have no direct coupling (apart from the unavoidable gravitational interaction) to the inflaton. We study whether this production can result in a detectable GW signal at CMB scales, without conflicting with the bounds from non-Gaussianity of the scalar perturbations. We study two possibilities: (i) gauge quanta production due to a sudden variation of their mass, and (ii) gauge quanta production from a rolling pseudo scalar. In case (i), we find that GW signals are unlikely to be detectable, due to the suppressed quadrupole moment of non-relativistic quanta. In case (ii), we instead find that GWs from particle production can actually exceed the usual inflationary vacuum fluctuations. Observable B-mode polarization can be obtained for any choice of inflaton potential, and the amplitude of the signal is not necessarily correlated with the scale of inflation.

Contents

List of Tables

List of Figures

Part I

Introduction

Recent advances in the observations of the cosmic microwave background (CMB) and large-scale structure (LSS) have provided compelling evidence for inflationary cosmology and have stimulated extensive studies of the detailed physics of the primordial universe and of the subsequent stages of its evolution. After the Cosmic Background Explorer (COBE) established the thermal nature of the CMB and found anisotropies in the CMB in 1992 [282], the Wilkinson Microwave Anisotropy Probe (WMAP) (see [61, 167] for their final analysis) and the most recent Planck spacecraft (see e.g. [11–13]), together with ground-based and balloon-borne missions, have probed these anisotropies to unprecedented precision.

The inflationary paradigm is tested against observations through its strict predictions. It was originally proposed in order to account for the fine-tuned initial conditions of the standard Big Bang cosmology and for the non-observation of unwanted relics which are expected from symmetry breaking in the early universe [16, 159, 209]. This can be achieved if the universe experienced a period of accelerating expansion, during which a small patch of the universe in causal contact is inflated to a size much larger than the current observable universe. It was soon realized that inflation naturally seeds small inhomogeneities in the CMB and LSS, originating from quantum vacuum fluctuations. The observational missions have so far shown that the anisotropies seen in the CMB and LSS are to a large extent what would be expected on the basis of inflationary theory.

In the simplest models, inflation is driven by a single scalar field. This scalar field, dubbed *inflaton*, needs to have a potential that is sufficiently flat to sustain a long enough duration of inflation. While it slowly rolls down its potential with its energy density dominating the energy density of the universe, it drives a quasi-exponential expansion of the space. Its vacuum fluctuations, on the other hand, become classical after the wavelengths of their modes are stretched beyond the horizon scale, and seed the cosmological inhomogeneities due to their coupling with the spacetime geometry through gravity. These simplest models predict a nearly scale-invariant, almost Gaussian spectrum of curvature perturbations, which is in good agreement with the current observations.

In this thesis, we study the effects of interactions between the inflaton and another field (or, in some cases, between two fields other than the inflaton) during inflation. We focus our attention to interactions involving vector fields, and these result in a significant production of vector quanta. These produced quanta in turn feed back to the inflaton perturbations (either from the direct coupling to the inflaton that was the origin of their production, or – if this coupling is absent – gravitationally), thus sourcing a new contribution to the primordial density perturbations. The same gauge quanta also source gravitational waves (GWs). Depending on the amount of quanta produced (and, therefore, on the strength of the interactions), the effects from particle production can significantly alter the phenomenological predictions of inflationary models over the simplest scenarios.

The inflaton must be coupled to other fields to decay and reheat the universe after inflation (namely, to form the quanta that characterized the Hot Big Bang phase that follows inflation). This decay cannot take place during inflation (or inflation would terminate) and therefore the inflaton couplings to other fields are typically disregarded during inflation. On the contrary, in most of the cases studied in this thesis we study the effect of gauge quanta produced by the motion of the unperturbed inflaton field during inflation. This extracts some of the energy out of the unperturbed inflaton. However, even the extraction of a small fraction of this energy, so that the background evolution is substantially unaffected, can drastically change the amount and the nature of the primordial perturbations and of the GWs. This is because the energy of the unperturbed inflaton is about 10 orders of magnitude greater than the energy in its vacuum perturbations and in the vacuum GWs (these are the standard modes present in inflationary models, irrespective of the gauge field production that we study). This can strongly constrain the strength of the inflaton interactions during inflation, and give rise to new signatures. In particular, the perturbations sourced by the produced quanta are highly non-Gaussian. Non-Gaussianity is typically signaled by a non-vanishing three-point correlation functions of the CMB anisotropies. Different triplets of points in the sky form triangles of different shapes, and many different shapes (in principle, infinitely many) of non-Gaussianity exist, according to which triangular configurations exhibit greater correlations. Any given mechanism typically produces a specific shape of non-Gaussianity, and this can provide specific signatures for specific models. Therefore, particular attention will be paid in this thesis to obtain the precise non-Gaussian signal for the different mechanisms of particle production that we study.

At the same time, we also compute the amount of GWs that are also sourced by the produced quanta. The limits on GWs are rapidly improving at the moment, and the improvement is expected to continue in the near future, both from the Planck satellite, and balloon-borne and ground-based experiments [85, 100, 194, 247]. These experiments are extremely important, since, in the standard models the amount of GWs is directly related to the energy scale of inflation, which is still unknown. Some of the models that we study have a sourced GW signal that can potentially be observed by these experiments. Moreover, in some of the models the amount of produced GWs increases during inflation. The resulting signal can be several orders of magnitude greater than the unsourced one (namely, the standard signal simply due to the expansion of the universe), and can potentially be observed even by the terrestrial interferometers such as for instance Advanced LIGO [162], Advanced Virgo [3], GEO-HF [309], and KAGRA [201].

Probing the physics of the primordial universe is to probe physical processes at very high energies. Since the Standard Model of elementary particles was well established by the 1980s, there has been continuous work in both theoretical and experimental particle physics. However, the new theoretical ideas involve energy scales far beyond current and near-future experimental reach. Primordial cosmology offers the data to test the predictions in some of those energy scales. Therefore, cosmology provides an exciting ground not only for cosmologists but also for elementary particle physicists.

The outline of this thesis is as follows. In Part II, we briefly review the motivations for the inflationary paradigm, i.e. the conceptual difficulties in the standard Big Bang cosmology and how inflation solves them. We also review the predictions from the simplest inflationary models. We then summarize the most recent observational constraints from the Planck satellite. In Part III, we obtain the general formalism of particle production and its effects on the cosmological perturbations. This formalism is readily applicable to the models in the following Parts. In Part IV, we study a very natural class of inflationary models, called axion inflation. In this class of models, a pseudo-scalar inflaton couples to a gauge field, and the production of the gauge quanta leads to surprisingly rich phenomenology. In Part V, we instead consider a scalar inflaton which is coupled to a gauge field through a characteristic interaction typical of dilaton-like fields. Non-Gaussianity of this model takes a nearly local shape with a distinct angular dependence depicting the nature of higher-spin fields. Part VI studies the models in which a gauge field is coupled to another field that is not the inflaton and interacts with both the inflaton and graviton only gravitationally. This is the most optimistic scenario, without exotic ingredients, in favor of the GW signals. In Part VII, we provide final reflections of our findings in the thesis and conclude.

Throughout this thesis, we follow the following conventions of notations: $\hbar = c = k_B = 1$, $M_p \equiv 1/\sqrt{8\pi G_{\rm N}} \simeq 2.436 \times 10^{18}$ GeV is the reduced Planck mass, where $G_{\rm N}$ is the Newtonian gravitational constant, Greek indices $(\mu, \nu, \dots)$ run through the 4-dimensional spacetime coordinate labels while Latin indices $(i, j, k, \dots)$ only the three spatial coordinate labels, and the summation is assumed for repeated indices. We denote physical time by t and conformal time by τ, and derivative with respect to physical/conformal time by a dot/prime. The scale factor and the physical Hubble rate are denoted by a and H, respectively. Also, I

we take the $(-+++)$ convention of the spacetime metric everywhere in the thesis apart from the model of fermionic production (Subsection VI 1.4 and Appendix F) where we switch to the opposite signature.

Part II

Inflationary Cosmology

1 Motivations

Primordial inflation, first introduced by Guth in 1981 [159], has been extensively studied in the past few decades, providing a firm picture of the very early stages of the universe. Inflation takes place before nucleosynthesis, the generation of the light elements, so that the Hot Big Bang model subsequently gives an essentially correct description of the observed universe. In inflationary scenarios, the universe undergoes a rapid, accelerated spatial expansion, which quickly dilutes all the contents in it which are decoupled from the inflationary dynamics. This (nearly) exponential expansion makes it possible that the scales relevant to the current largest-scale observations were once in a causally connected region in the earlier stage of inflation. Such scenarios can be realized in a surprisingly easy setup and provides a compelling mechanism to resolve conceptual difficulties in the standard Hot Big Bang cosmology.

Moreover, the inhomogeneous perturbations of the metric and other relevant degrees of freedom during inflation can seed the observable fluctuations of the cosmic microwave background (CMB) and large-scale structure (LSS). The recent technological advances of the observations of CMB and LSS allow the inflationary theory to be a framework of testable predictions. Inflationary cosmology has in this regard turned out to be an important source of information for elementary particle theory, as the inflationary energy scale may be as high as $\sim 10^{16}$ GeV [12], which may therefore allow us to probe a regime inaccessible in particle accelerators. In this respect, inflation can serve as the only laboratory that could probe the physical processes in such ultra-high energies.

In this Section, we briefly review the Hot Big Bang cosmology without inflation and state the conceptual problems that inevitably arise in this standard picture, the very reason why inflation was first demanded. This introductory Part is based mainly on [206, 211, 239, 306].

1.1 Hot Big Bang Cosmology

The central premise of modern cosmology is that the visible universe on large scales is isotropic and homogeneous: it looks the same in all directions at any given point in space.[1]

[1] One has to take an appropriate reference frame to make this statement. The universe would not appear isotropic if an observer were moving close to the speed of light. In considering the CMB, the CMB dipole

This has been tested by a variety of observations, such as galaxy distributions (e.g. [312, 313]), but the most notable one is the nearly identical temperature of the CMB coming from different parts of the sky. For smaller scales, the universe is highly inhomogeneous, with materials clumped into stars, galaxies, and galaxy clusters. However, presumably these inhomogeneities have grown over time due to gravitational attraction and the distribution was more homogeneous at earlier times.

The assumption that the visible universe is isotropic and homogeneous on large scales leads us to choose a spacetime coordinate system where metric takes a simple form. This metric was derived separately by Friedmann [133], Robertson [259], and Walker [301], and is called the FRW metric.[2] Including the expansion of the universe, which was observationally found by the early 1930s by Hubble, it takes the form

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

where ds is the differential line element, $a(t)$ is the scale factor, and r and Ω are the radial coordinate and solid angle, respectively, of the spherical coordinate system. The constant K is related to the curvature of spatial geometry, and the scale factor can always be normalized, without loss of generality, such that $K = -1, 0, +1$, corresponding to an open, flat, or closed Friedmann universe, respectively

The expansion of the universe is governed by the Einstein field equations

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{1}{M_p^2}T_{\mu\nu} \tag{2}$$

where $R_{\mu\nu}$ is the Ricci tensor, constructed from the metric $g_{\mu\nu}$, R is the Ricci scalar $R \equiv R^{\mu}{}_{\mu}$, and $T_{\mu\nu}$ is the energy-momentum tensor of the contents of the universe. These equations with the FRW metric (1) determine the evolution of the scale factor through

$$H^2 = \frac{\rho}{3M_p^2} - \frac{K}{a^2} \tag{3}$$

$$\frac{\ddot{a}}{a} = -\frac{\rho + 3p}{6M_p^2} \tag{4}$$

where dots denote the derivative with respect to physical time t, $H \equiv \frac{\dot{a}}{a}$ is the Hubble

observations tell us that there is a frame in which its dipole anisotropy vanishes, which we call the CMB rest frame, and that we (the earth) move with respect to the CMB at a speed of order $10^{-3}c$ [15].

[2] Different conventions call this metric RW or FLRW metric (with Lemaître [203] included), but "FRW" is taken throughout this thesis.

parameter, $\rho \equiv -T^0{}_0$ is the energy density of matter in the universe, and $p \equiv \frac{1}{3} T^i{}_i$ is its pressure. These two equations combined derive the energy conservation law [3]

$$\dot{\rho} + 3\,\frac{\dot{a}}{a}\,(\rho + p) = 0 \tag{5}$$

and each decoupled matter content obeys this equation separately. We define the equation of state w as

$$w \equiv \frac{p}{\rho} \tag{6}$$

and if w is constant, (5) can be solved as

$$\rho \propto a^{-3(1+w)}\ . \tag{7}$$

If the energy content of the form (7) dominates the total energy density of the universe, we can solve (3) to find the time dependence of a,

$$a \propto \begin{cases} t^{\frac{2}{3(1+w)}} & w \neq -1 \\ \mathrm{e}^{Ht} & w = -1 \end{cases} \tag{8}$$

Some physically important energy contents are (i) ultra-relativistic gas (radiation) $w = \frac{1}{3}$, (ii) non-relativistic dust (matter) $w = 0$, and (iii) cosmological constant $w = -1$. Their energy densities behave as

$$\rho \propto \begin{cases} a^{-4}\ , & \text{Radiation} \\ a^{-3}\ , & \text{Matter} \\ \text{const.}\ , & \text{Cosmological constant} \end{cases} \tag{9}$$

These dependences are physically understandable; as the volume of space expands as $\propto a^3$, the density of matter is diluted by the same proportion, while the energy of radiation is further decreased by $\propto a^{-1}$ due to its wavelength stretched by the expansion (cosmological constant is simply as its name stands). A comparison of these different behaviors is schematically illustrated in Figure 1 (in log scale). If one of them dominates the total

[3] Alternatively, the energy conservation equation (5) can also be derived from the $\nu = 0$ part of the conservation equation of the energy-momentum tensor, $\nabla^\mu T_{\mu\nu} = 0$, where ∇^μ is the covariant derivative. This is no accidental; if we act ∇^μ on the both sides of (2), the left-hand side is identically zero, due to the Bianchi identity. This means that the Einstein field equations (2) already contain the energy and momentum conservation, and the latter does not introduce any additional information.

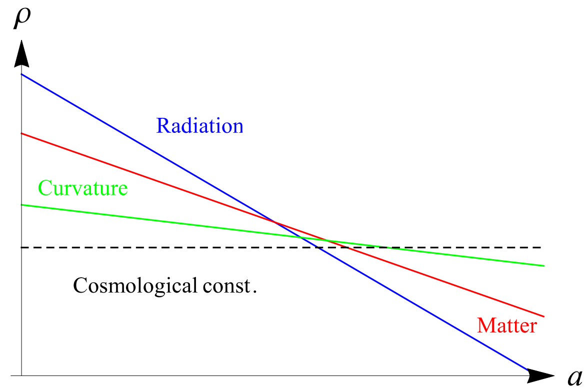

Figure 1: A schematic illustration of the evolution of energy densities.

density, the time dependences of the scale factor and of the Hubble parameter are

$$\text{Radiation domination} \; : \quad a \propto t^{1/2} \; , \quad H = \frac{1}{2t} \tag{10}$$

$$\text{Matter domination} \; : \quad a \propto t^{2/3} \; , \quad H = \frac{2}{3t} \tag{11}$$

$$\text{Cosm. const. domination} \; : \quad a \propto \mathrm{e}^{Ht} \; , \quad H = \text{const.} \tag{12}$$

These observations tell us that unless the cosmological constant already dominates at the beginning, regardless of the value of K, the scale factor vanishes at some "initial" time $t = 0$, and the energy density becomes infinite. This is the time of the initial cosmological singularity, historically called Big Bang.

If cosmological constant is absent, the universe continues to expand forever if it is open ($K = -1$) or flat ($K = 0$).[4] If it is closed ($K = +1$) with $w > -\frac{1}{3}$, on the other hand, then the two terms in (3) become equal at some time in the expansion; after this point, the scale factor a decreases and eventually vanishes, which leads to the future singularity, called Big Crunch. However, if cosmological constant, or anything else with $w < -\frac{1}{3}$, is present, the fate of the universe is different. Once it dominates the energy density, it keeps dominating, and the universe expands indefinitely.

Once the initial contents of the universe are known, its subsequent evolution can be determined through (3–5). It was, however, not clear whether the early universe had been filled with non-relativistic dust (cold) or ultra-relativistic gas (hot), until the 2.7 K CMB

[4] For $K = -1$, after the curvature term $\frac{K}{a^2}$ in the right-hand side of (3) dominates over the other term, the scale factor goes as $a \propto t$.

was found in 1965 by Penzias and Wilson [250], followed by another group [108]. This discovery quickly led to widespread acceptance of the hot universe model, as it predicts the CMB radiation emitted in the early universe [118, 134].

In this model, the universe was in its earliest stages of evolution filled with ultra-relativistic plasma of elementary particles, such as quarks, leptons, gauge bosons, Higgs bosons, and other beyond-Standard-Model particles. During this epoch, temperature is extremely high (typically $T \gtrsim 10^{16}$ GeV), and one may assume that some underlying symmetry is restored and that interactions do not distinguish particle species; such theories are called grand unified theories (GUTs). As the universe expands, the temperature drops as

$$T \propto a^{-1} \tag{13}$$

due to the entropy conservation for adiabatic expansion. As it cools down, a number of phase transitions are expected to take place during the course of its early history. The GUT phase transition spontaneously breaks the unifying symmetry (if it existed), down to $\mathrm{SU_c}(3) \times \mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$ at the temperatures $T \sim 10^{14} - 10^{16}$ GeV, where $\mathrm{SU_c}(3)$ corresponds to the strong interaction gauge group, $\mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$ to the electroweak. While $\mathrm{SU_c}(3)$ remains unbroken, $\mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$ goes through further spontaneous symmetry breaking at $T \sim \mathcal{O}(10^2)$ GeV, so that $\mathrm{SU_L}(2) \times \mathrm{U_Y}(1) \to \mathrm{U_Q}(1)$ through the Higgs mechanism.

At $T \sim 100 - 300$ MeV ($t \sim 10^{-5}$ sec), a transition associated with chiral symmetry breaking and color confinement takes place, which separates the strongly interacting particles into baryons and mesons. The details of this transition could possibly affect the following processes, but primordial (Big Bang) nucleosynthesis should occur at $T \sim 10 - 0.1$ MeV. Nucleosynthesis is the epoch of the generation of light elements and is a milestone of the current cosmology. At $T \sim 1$ eV ($t \sim 10^{11}$ sec), the universe enters from the radiation-dominated to the matter-dominated epoch. This is called matter-radiation equality and is also the start of structure formation. Finally, the ions and electrons combine to form neutral atoms, and the radiation (photons) decouple from the matter. This photon decoupling, also called the last scattering, takes place at $T \sim 0.1$ eV ($t \sim 10^{13}$, red shift $z \sim 1100$). The CMB detected by Penzias and Wilson is the remnant of such decoupled photons, which have been cooled by the expansion of the universe.

This is a very brief description of the early history of the universe within Hot Big Bang cosmology. While it gained widespread acceptance immediately after the discovery of the CMB and has since then provided an essentially correct picture of the cosmological evolution, it suffers several conceptual difficulties, along with the uncertainties in the initial conditions of the universe. In the following, we list these issues, giving way to inflationary cosmology.

1.2 Problems with the Standard Picture

The Hot Big Bang theory can confidently depict the evolution of the universe after nucleosynthesis. Yet, the earlier epochs are still a subject of speculation and contain uncertainties for the initial conditions. There are several issues associated with this theory which cannot be solved within its own framework. Below, we briefly discuss these issues, for most of which inflation will provide solutions (see e.g. [206, 211, 306]).

(i) *Flatness problem*:

The Friedmann equation (3) can be rewritten as

$$\Omega - 1 = \frac{K}{a^2 H^2} \tag{14}$$

where $\Omega \equiv \frac{\rho}{\rho_c}$ is the fractional density of the total matters, and $\rho_c \equiv 3M_p^2 H^2$ is called the critical density. If the universe is exactly flat, i.e. $K = 0$ and $\Omega = 1$, then it remains so for all time; otherwise, Ω evolves over time due to the difference in time dependence between Ω and the $\frac{K}{a^2H^2}$ term. Using (10, 11), we have $|\Omega - 1| \propto t$ for a radiation-dominated universe and $|\Omega - 1| \propto t^{2/3}$ for a matter-dominated universe. In either case (more generally $w < -1$ or $w > -\frac{1}{3}$), $|\Omega - 1|$ increases as time goes on. Current observations set a stringent bound on this value [11],

$$1 - \Omega_0 = -0.0005^{+0.0065}_{-0.0066} \tag{15}$$

at 95% CL, where the subscript 0 denotes the present time. The fact that the total fractional density Ω is close to 1 implies that at earlier times, it must have been extremely close to 1, e.g. $|\Omega - 1| \lesssim 10^{-16}$ at most at the time of nucleosynthesis and even smaller at earlier times, in order to obtain the currently observed universe. This is not a paradox in itself, i.e. there is no reason why the curvature should not have been so small or 0; yet, one may argue that such finely tuned initial conditions seem unlikely, which is what physicists typically would like to explain if possible.

(ii) *Horizon problem*:

From the first detection of the CMB, its high degree of isotropy raised a question. The particle horizon, the distance beyond which no past events can be observed, in a

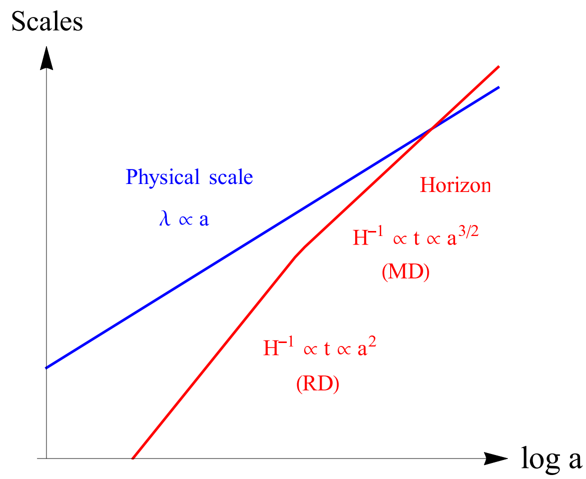

Figure 2: The evolution of physical scale λ and of the horizon $R_H \sim H^{-1}$.

radiation-dominated (RD) and matter-dominated (MD) universe is [5]

$$R_H(t) = a(t) \int_0^t \frac{dt'}{a(t')} = \begin{cases} 2t \propto a^2 \,, & \text{RD} \\ 3t \propto a^{3/2} \,, & \text{MD} \end{cases} \tag{16}$$

using (10, 11). On the other hand, physical scales scale as $\lambda \propto a$, or $\lambda \propto t^{1/2}$ in the RD and $\lambda \propto t^{2/3}$ in the MD. This scaling difference between λ and R_H is illustrated in Figure 2. This difference implies that the horizon size has increased at a higher rate than physical scales since the time of last scattering, and thus regions causally disconnected in the past have been entering inside the horizon. Therefore, the CMB radiation coming from the regions separated by more than the horizon scale at the time of last scattering (i.e. when it started free propagation), which now subtends an angle in the sky about $\mathcal{O}(1^\circ)$, could not have been previously smoothed out by any physical processes. The Hot Big Bang model can provide no prospect of explaining why the temperature of the CMB is nearly perfectly isotropic and homogeneous.

(iii) *Unwanted relics problem*:

In grand unified theories, local symmetry under a simple symmetry group is spontaneously broken at typically an energy $\sim 10^{14} - 10^{16}$ GeV to the Standard Model gauge symmetry under the group $\mathrm{SU_c}(3) \times \mathrm{SU_L}(2) \times \mathrm{SU_Y}(1)$. In all such models, the

[5] In the RD and MD universe, the event horizon has an infinite radius, that is, there is no limitation on the distance that light will be able to travel in the future.

scalar fields that break the symmetry can be left with non-zero magnetic charge and cannot be diluted through any continuous processes to the observed level (see [304] for discussion). This poses a problem for some cosmological models (see [254] for a review). Unless they can find and annihilate each other, such magnetic monopoles would be found by the amount of roughly one monopole per nucleon today, which is in disagreement with the observations. This potential difficulty was one of the leading motivations for inflationary cosmological models. Other topological defects, such as cosmic strings, domain walls, and textures, may also be problematic, but this potential harassment is more model-dependent (see e.g. [298, 315] for reviews).

In models beyond the Standard Model, another potentially dangerous relic is the gravitino, the spin-$\frac{3}{2}$ partner of the graviton in supergravity theory. In most versions of supergravity, the presence of gravitino upsets nucleosynthesis if the Hot Big Bang starts before $T \gtrsim 10^5$ GeV [186]. There are other potential classes of unwanted relics in models such as superstring theory [31, 106]. In any of these cases, problems arise due to the relics produced in the early universe that are in conflict with observations.

(iv) *Seed of inhomogeneity*:

The universe is observed to be highly isotropic and homogeneous on large scales. However, it is not completely so; the fluctuations in the isotropic and homogeneous background are of order 10^{-5} in the CMB temperature and the LSS density. Within the Hot Big Bang model, a simple interpretation of the origin of such fluctuations is that they correspond to inhomogeneities present at the surface of last scattering and further evolved over time. Their corresponding scale is, however, much larger than the horizon size at that time, and therefore they could not be generated causally and would have to be part of the initial conditions [171, 205]. However, the Hot Big Bang model does not provide a mechanism to implement such initial conditions. Although it remains possible that the large-scale inhomogeneities are generated well after last scattering, an alternative way beyond the standard Hot Big Bang model seems more natural. Inflation automatically implements a mechanism to produce initial inhomogeneities as quantum fluctuations, which are in a good agreement with the CMB and LSS observation.

The most serious of the above issues is the horizon problem, as there are possible solutions for the other ones which do not have to rely on inflation. Historically, the issues (i)–(iii) were the main motivations for inflation. As it has turned out, a further interesting aspect is that inflation can generate irregularities in the very early universe that persist to the present, addressing (iv) of the issues. These inhomogeneities are originated from quantum fluctuations and may lead to structure formation. In the following section, we briefly summarize the basic properties of inflation and how it resolves the above issues.

2 Inflationary Paradigm

Inflationary cosmology is not a replacement for the previously existing Hot Big Bang cosmology; rather, the former adds an inflating phase to the cosmological evolution, well before the time the latter successfully gives a relevant description of the current universe. The formal definition of inflation is an epoch during which the universe undergoes an accelerating expansion, or in terms of the scale factor, introduced in (1),

$$\text{Inflation} \quad \Longleftrightarrow \quad \ddot{a} > 0 \tag{17}$$

which is achieved under the condition of the energy contents

$$\text{Inflation} \quad \Longleftrightarrow \quad \rho + 3p < 0 \quad \text{i.e.} \quad w < -\frac{1}{3} \tag{18}$$

as can be seen from (4). As we always assume that the energy density ρ is positive, the pressure p has to be negative in order to satisfy this condition, which is independent of the curvature K of the universe.

If inflation occurs, all the issues within the Hot Big Bang model that are discussed in the previous section can be resolved. The condition (17) is precisely the one that drives the fractional density Ω toward unity. The time derivative of the right-hand side of (14) is $-\frac{2K}{a^3H^3}\,\ddot{a}$. As $a > 0$ and $H > 0$, the condition (17) implies that the right-hand side of (14) decreases over time, leading to $\Omega \to 1$, consistent with the current observation (15). The abundances of unwanted relics can be easily reduced to an acceptable level by the expansion during inflation, as long as they are produced before the inflationary epoch. The horizon problem can also be solved due to the decrease of the comoving Hubble length $(aH)^{-1}$ during inflation. In the case of an exponential expansion, for example, physical scales, which are proportional to a, extend exponentially, while the horizon H^{-1} stays constant. This situation is depicted in Figure 3. As a result, any observable scales, which used to be outside the horizon in the past, could have once been within a causally connected region inside the horizon, as long as inflation lasted sufficiently long. The overwhelming inhomogeneities and anisotropies that might have been present before inflation could be smoothed out before

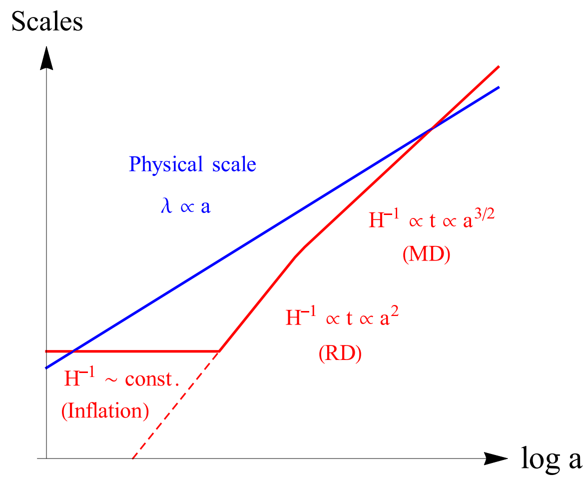

Figure 3: The evolution of physical scale λ and of the horizon $R_H \sim H^{-1}$, with initial inflationary phase.

the horizon exit during inflation. Yet another attractive consequence of inflation is that the quantum fluctuations present initially and stretched by the expansion can seed the small perturbations that are observed in the CMB and are necessary for structure formation.

Thanks to all these merits, the relatively simple idea of inflation has become a dominant paradigm for the cosmology of the very early universe. In the following Subsection, we briefly discuss the class of simplest physical realizations of inflation.

2.1 Slow-roll Inflation

Guth's original work [159] was that a scalar field was initially trapped in a local minimum of some potential, supporting an exponential expansion, and then leaked through the potential barrier and rolled down to a true minimum. It was soon realized [160, 163] that this idea suffers from the so-called graceful exit problem. The first-order phase transition from the initial "false vacuum" phase to the lower-energy "true vacuum" phase could not have occurred everywhere in space simultaneously, but instead in the small bubbles of true vacuum. The latent heat released in the phase transition would have spread only on the bubble walls, while their interiors are essentially empty. Thus the distribution of the energy that could grow into the present contents of the universe would be highly inhomogeneous and anisotropic. The background space of false vacuum continued to inflate at a rate too fast for the bubble walls ever to coalesce to smooth out those irregularities. Guth's version of inflation was then modified by Linde [209, 210] and by Albrecht and Steinhardt [16], known as "new inflation." In this new model, bubbles are again formed in the phase transition; however, low but finite temperature lowers the potential barrier significantly. The scalar field then slowly rolls down the potential, while the universe, including the bubbles, expands exponentially. Eventually the field is converted into ordinary particles, filling the bubble, a small part of which our universe is supposed to occupy.

It was soon understood that in these models, the physical observables are the consequences of the slow roll of the scalar field, rather than the bubble nucleation itself. Indeed, the important aspects of inflation depend only on the requirement that a scalar field φ, called the *inflaton*, has a large but flat potential $V(\varphi)$. Scalar fields are peculiar in that if their energy density is dominated by the potential term over the kinetic, their equation of state automatically becomes $w \simeq -1$, leading to a nearly exponential expansion and an almost constant Hubble parameter, until the field value starts changing fast.

There are a huge number of variations of inflationary models (see e.g. [59, 213, 219, 248] for reviews). Here we illustrate the simplest single-scalar-field models of inflation, on which the most of the models studied in the following parts of this thesis are based. This class of models is characterized by a simple action

$$S = \int d^4x \left[\frac{M_p^2}{2} R - \frac{1}{2}\, g^{\mu\nu}\, \partial_\mu\varphi\, \partial_\nu\varphi - V(\varphi) \right] \tag{19}$$

where $g_{\mu\nu}$ is the metric tensor of (1), R is the corresponding Ricci scalar, and $V(\varphi)$ is the inflaton potential. The energy density $\rho = -T^0{}_0$ and the pressure $p = \frac{1}{3} T^i{}_i$ of a spatially homogeneous inflaton field in a FRW spacetime take the form, respectively,[6]

$$\rho = \frac{1}{2}\dot{\varphi}^{(0)\,2} + V \tag{20}$$

$$p = \frac{1}{2}\dot{\varphi}^{(0)\,2} - V \tag{21}$$

which immediately shows that $w \simeq -1$ if $V \gg \frac{1}{2}\dot{\varphi}^{(0)\,2}$. Inflation lasts as long as $\rho + 3p < 0$, that is $\dot{\varphi}^{(0)\,2} < V$. With a suitably flat potential, even if this condition is not satisfied initially, $\varphi^{(0)}$ quickly runs into a so-called *attractor* to satisfy it. Once inflation starts, the curvature term in the Friedmann equation (3) quickly becomes negligible; we thus ignore

[6] Notational note: in this thesis, $\varphi(t, \vec{x})$ denotes the total value of the inflaton, while $\varphi^{(0)}(t)$ denotes its homogeneous background value. Later we consider its perturbation around $\varphi^{(0)}(t)$, denoting it as $\delta\varphi(t, \vec{x}) = \varphi(t, \vec{x}) - \varphi^{(0)}(t)$.

this term from the start. Eq. (3) with a single scalar inflaton then takes the form

$$H^2 = \frac{1}{3M_p^2}\left[\frac{1}{2}\dot{\varphi}^{(0)\,2} + V\right] \tag{22}$$

and the energy conservation law (5) becomes

$$\ddot{\varphi}^{(0)} + 3H\dot{\varphi}^{(0)} + V_\varphi = 0 \tag{23}$$

where the subscript φ denotes the derivative with respect to φ. The above two equations determine the dynamics of inflation, as the $\ddot{a}$ equation, (4), is their direct consequence. The slow-roll conditions are quantified by two parameters

$$\epsilon \equiv \frac{M_p^2}{2}\left(\frac{V_\varphi}{V}\right)^2 \ , \quad \eta \equiv M_p^2 \frac{V_{\varphi\varphi}}{V} \tag{24}$$

and are satisfied by requiring $\epsilon \ll 1$ and $|\eta| \ll 1$.[7] These are the two "flatness" conditions needed to insure the slow roll of both ϕ and $\dot{\phi}$. Under these conditions, (22) and (23) are approximated as

$$H^2 \simeq \frac{V}{3M_p^2} \tag{27}$$

$$3H\dot{\varphi}^{(0)} \simeq -V_\varphi \ . \tag{28}$$

At a glance, the conditions $\epsilon \ll 1, |\eta| \ll 1$ only restrict the form of the potential, and since the full equation of motion for the inflaton, (23), is a second-order differential equation, the value of $\dot{\varphi}^{(0)}$ can still be chosen freely. Therefore, it is in principle an additional assumption that the solution for a given potential satisfies (28), or $|\ddot{\varphi}^{(0)}| \ll 3H|\dot{\varphi}^{(0)}| \simeq |V_\varphi|$. However, it can be shown that (27) and (28) are the attractor of the full system. As long as the conditions $\epsilon \ll 1, |\eta| \ll 1$ are satisfied, even if the system starts off with the value of $\dot{\varphi}^{(0)}$ violating the slow roll, it quickly moves into the attractor to satisfy (27) and (28).

[7] An alternative set of slow-roll parameters in terms of the Hubble parameter is

$$\epsilon_H \equiv -\frac{\dot{H}}{H^2} \ , \quad \delta_H \equiv \frac{\ddot{H}}{2H\dot{H}} \ . \tag{25}$$

These parameters are related to the ones in terms of inflaton potential, ϵ, η, through

$$\epsilon_H = \frac{\dot{\varphi}^{(0)\,2}}{2M_p^2 H^2} \simeq \epsilon \ , \quad \delta_H = \frac{\ddot{\varphi}^{(0)}}{H\dot{\varphi}^{(0)}} \simeq \epsilon - \eta \ . \tag{26}$$

Therefore, the slow-roll conditions $\epsilon \ll 1, |\eta| \ll 1$ are equivalent to requiring $\epsilon_H \ll 1, |\delta_H| \ll 1$.

This makes the condition $|\ddot{\varphi}^{(0)}| \ll 3H|\dot{\varphi}^{(0)}|$ satisfied without a further explicit assumption. In fact, the attractor behavior is vital for the success of the slow-roll approximations, as it enables us to make generic predictions from inflation using (27) and (28).

Seen from (17), inflation continues as long as

$$\frac{\ddot{a}}{a} = \dot{H} + H^2 \simeq (1-\epsilon)\, H^2 > 0 \tag{29}$$

and thus if the slow-roll condition $\epsilon \ll 1$ is satisfied, inflation is guaranteed. One way for inflation to end is by violation of the slow roll as the field approaches a minimum of its potential. In such cases, it can be considered that inflation comes to an end when $\epsilon(\varphi) = 1$ is reached.

The amount of expansion during inflation is often quantified by the number of e-foldings N, the logarithm of the ratio of the scale factor at the end of inflation to its value at some time during inflation, i.e.

$$N(t) \equiv \ln \frac{a(t_{\mathrm{end}})}{a(t)} \tag{30}$$

where t_{end} denotes the time of the end of inflation. Since a is a strictly increasing function of time, $N(t)$ can also be considered a time variable, which measures the amount of inflation that still has to occur after time t. Under the slow-roll conditions, the expansion is generally not strictly exponential, but it can still be very large. Suppose that $\varphi^{(0)}$ moves down its potential during inflation toward the end value $\phi_{\mathrm{end}} \equiv \varphi^{(0)}(t_{\mathrm{end}})$, defined by $\epsilon(\phi_{\mathrm{end}}) = 1$, with $0 < V(\phi_{\mathrm{end}}) < V(t)$. Under the slow-roll approximations, we then have

$$N = \int_t^{t_{\mathrm{end}}} H\, dt = \int_{\varphi^{(0)}}^{\phi_{\mathrm{end}}} \frac{H}{\dot{\tilde{\varphi}}^{(0)}}\, d\tilde{\varphi}^{(0)} \simeq \frac{1}{M_p^2} \int_{\phi_{\mathrm{end}}}^{\varphi^{(0)}} \frac{V}{V_\varphi}\, d\tilde{\varphi}^{(0)} \, . \tag{31}$$

For example, for the potential as a monomial function of φ, i.e. $V(\varphi) \propto \varphi^p$, the number of e-foldings is

$$N \simeq \frac{\varphi^{(0)\,2} - \phi_{\mathrm{end}}^2}{2pM_p^2} = \frac{p}{4}\left[\frac{1}{\epsilon(t)} - 1\right] \simeq \frac{p}{4\epsilon(t)} \tag{32}$$

where the last approximation is valid well before the end of inflation. Thus, an arbitrarily large number of e-foldings can be achieved as long as $\epsilon \ll 1$. With the standard evolution, the number of e-foldings required to generate the scales that correspond to the present Hubble radius is $N = 50 - 60$, depending on the precise thermal history of the universe.

It is worth noting that in large-field models, of which the simplest examples are those with a monomial potential $V(\varphi) \propto \varphi^p$, a large N requires $\varphi^{(0)} > M_p$ (if $p \gtrsim (2N)^{-1}$ in the monomial potential case, see (32)). Such large values of $\varphi^{(0)}$ do not necessarily rule out the classical treatment of gravity on which we have been relying. The condition that the corrections from quantum gravity can be neglected is that the energy density should be much smaller than the Planck energy density, $V(\varphi) \ll M_p^4$. This condition can be satisfied even for $\varphi^{(0)} > M_p$, if $V(\varphi)$ is proportional to a sufficiently small coupling constant or mass. Such a constant appears neither in the slow-roll parameters (24) nor in the growth (31) of a and is irrelevant for the background dynamics of inflation. Its value can only be determined by the consideration of the perturbations. The current observations of the CMB and LSS indeed support a small constant that fulfills the condition $V(\varphi) \ll M_p^4$.

To enjoy the success of the Hot Big Bang model, the universe has to be filled with ultra-relativistic ordinary particles (radiation) some time between the end of inflation and nucleosynthesis. This process is called *reheating*. While reheating typically has little impact on the predictions from the inflationary scenario and thus is not a focus of this thesis, an understanding of it is of great physical importance. A brief description of what is supposed to happen in the process is typically the following. First, after inflation ends, the inflaton field starts oscillating coherently around a minimum of its potential at a rapid rate on the Hubble timescale. During this period, the inflaton effectively behaves like matter, with its energy density $\rho \propto a^{-3}$. This oscillating phase could last for considerable time, depending on the particle decay time. Secondly, the inflaton undergoes coherent decays, once the Hubble time reaches the decay time. One notable process for this phase is *preheating*, where an extremely efficient decay takes place due to broad parametric resonance between the inflaton and other bosonic particles [189, 190]. The last stage of reheating process is thermalization. The bosonic particles should decay, interact, and finally reach thermal equilibrium [19,55,252]. The details are strongly model-dependent, but this series of phases ultimately determines the temperature at which the universe becomes in equilibrium and enters the standard Hot Big Bang behavior.

For inflation to successfully account for the observed isotropy and homogeneity, any given observable scale must have been within the Hubble radius some time during inflation. In this respect, we are interested in the evolution of comoving wavenumber k, arising from a Fourier transform of perturbations, related to physical scale by $\lambda \equiv \frac{a}{k}$. The comoving Hubble length $(aH)^{-1}$ is by definition decreasing during inflation, whereas it is increasing at all other times. A fixed comoving scale k^{-1} may therefore be inside the horizon at some initial time and be well outside of it by the end of inflation. For any such scale, an important moment is the time when it crosses the horizon, i.e. $k = aH$. This can be related to the number of e-foldings N, defined in (30), which now depends on k, and to the scale now equaling the Hubble radius, $k = a_0 H_0$. One simple yet viable scenario of the evolution is that the universe goes through (i) the inflationary epoch (ending at $t_{\rm end}$), (ii) the stage of inflaton coherent oscillation (from $t_{\rm end}$ to $t_{\rm reh}$), (iii) reheating (at $t_{\rm reh}$), (iv) the radiation-dominated era (from $t_{\rm reh}$ to $t_{\rm eq}$), and (v) the matter-dominated era (from $t_{\rm eq}$ to t_0). Assuming instantaneous transitions between the regimes, we can write [206],

$$\frac{k}{a_0 H_0} = \frac{a_k H_k}{a_0 H_0} = \frac{a_k}{a_{\rm end}} \frac{a_{\rm end}}{a_{\rm reh}} \frac{a_{\rm reh}}{a_{\rm eq}} \frac{a_{\rm eq}}{a_0} \frac{H_k}{H_0} \tag{33}$$

where the subscript k denotes the time when $k = aH$. The fraction $\frac{a_k}{a_{\rm end}}$ gives the number of e-foldings $N(k)$ of inflation remaining when the scale k crosses the horizon. This leads to, assuming typical evolution,

$$N(k) = 62 - \ln \frac{k}{a_0 H_0} - \ln \frac{10^{16}\ \text{GeV}}{V_{\rm CMB}^{1/4}} + \ln \frac{V_{\rm CMB}^{1/4}}{V_{\rm end}^{1/4}} - \frac{1}{3} \ln \frac{V_{\rm end}^{1/4}}{\rho_{\rm reh}^{1/4}} \, . \tag{34}$$

In typical models of inflation, the uncertainties regarding the precise evolution is not expected to be large, and generally we can take $N(k) = 50 - 60$. The smallest scale accessible to LSS observations is about 1 Mpc, which would correspond to the Hubble length about 9 e-foldings after the largest scale does. Thus, all scales relevant to the LSS and CMB crossed the horizon within a limited number of e-foldings well before the end of inflation. As long as inflation lasted this long, the issues raised in Subsection II 1.2 can be solved simultaneously.

2.2 Perturbation from Inflation

The last Subsection was devoted to the description of the background dynamics of inflation and consequently to explaining the currently observed flatness, isotropy, and homogeneity of the universe. There are, however, small inhomogeneities ($\sim 10^{-5}$ level) in the CMB temperature and in the LSS density distribution. The inflationary cosmology, discussed in the previous Subsection, can in fact account for such inhomogeneities, without any additional ingredients. This is one of the major features that inflation has attracted considerable attention since the idea was introduced.

A robust treatment of perturbations from inflation requires several technical details. In order to avoid losing track of the context, we will leave them to the following Parts of the thesis, where we study different models in detail. Here we only introduce the basics and observable quantities important for phenomenology on generic grounds.

As seen in the previous Subsection, all the observable scales crossed the horizon during inflation. Well within the horizon, the perturbations of the inflaton, metric, and any other relevant fields behave as freely propagating quantum fields, once properly quantized. These quantum fluctuations are stretched by the expansion and go through the transition from quantum to classical perturbations [253, 287]. Among different types of perturbations, the ones of most phenomenological interest are the *curvature perturbation* (scalar mode) and the *gravitational wave* (tensor mode). These modes are conserved outside the horizon, which is the very property that allows us to connect the distant past (at the horizon exit during inflation) to the relatively recent past (at the horizon reentry) (see [306] and references therein).

The curvature perturbation, ζ, is related to one of the scalar modes of the spatial components of metric perturbations [33, 302].[8] The goal of the theory of cosmological perturbations is to compute their spectral functions of primordial origin. These functions convey various information of the physics in the very early universe, which is testable by the observations of the CMB and LSS. The two-point correlation function of ζ defines its power spectrum P_ζ by

$$\langle \zeta_{\vec{k}}\, \zeta_{\vec{k}'} \rangle \equiv \frac{2\pi^2}{k^3}\, P_\zeta(\vec{k})\, \delta^{(3)}(\vec{k}+\vec{k}') \tag{35}$$

where $\zeta_{\vec{k}}$ is the Fourier transform of $\zeta(\vec{x})$, $k \equiv |\vec{k}|$, and $\delta^{(3)}$ is the 3-dimensional Dirac delta function, which is the consequence of the translational invariance (momentum conservation) in the theory.[9] If the background dynamics respects rotational invariance, the power spectrum will depend only on the magnitude of $\vec{k}$, i.e. $P_\zeta(\vec{k}) = P_\zeta(k)$. The power spectrum is often parametrized as

$$P_\zeta = \mathcal{A}_s \left(\frac{k}{k_*}\right)^{n_s - 1} \tag{36}$$

where $\mathcal{A}_s$ is the amplitude of the spectrum, n_s is the scalar spectral index, and k_* is a pivot momentum of the observation, which typically is taken to be $k_* = 0.05\ \mathrm{Mpc}^{-1}$ or $0.002\ \mathrm{Mpc}^{-1}$. In the simplest single-scalar-field models discussed in the previous Subsection, the spectrum is found to be

$$P_\zeta \cong \left(\frac{H^2}{2\pi\dot{\varphi}^{(0)}}\right)^2 \left(\frac{k}{k_*}\right)^{2\eta-6\epsilon} \simeq \frac{H^2}{8\pi^2\epsilon M_p^2}\left(\frac{k}{k_*}\right)^{2\eta-6\epsilon} \tag{37}$$

[8] There is another standard conventional definition of curvature perturbation, $\mathcal{R}$, which is also conserved outside the horizon [32, 217]. The difference between the definitions of ζ and $\mathcal{R}$ is the choice of coordinates on which they are defined: ζ is defined on the uniform-density hypersurfaces, while $\mathcal{R}$ is on the co-moving hypersurfaces. Nonetheless, they are equivalent quantities outside the horizon, i.e. $k \ll aH$.

[9] What we suffice to assume is in fact the probability distribution function that is translationally invariant, not necessarily the translationally invariant initial conditions of the perturbations.

giving $\mathcal{A}_s \cong \frac{H^2}{8\pi^2\epsilon M_p^2} \simeq \frac{V}{24\pi^2\epsilon M_p^4}$ and $n_s - 1 \cong 2\eta - 6\epsilon \ll 1$.[10] A nearly scale-invariant power spectrum, $n_s \simeq 1$, is a generic prediction of the simplest models and is in a good agreement with the observations, $n_s = 0.9603 \pm 0.0073$ [12].

The observed fluctuations are consistent with Gaussian spectra. If they are completely Gaussian, then all the even-number correlation functions can be expressed in terms only of combinations of two-point correlation functions, and all the odd-number correlation functions will vanish. There have been a large number of ongoing activities to search for the signals deviated from such Gaussian spectra, which can potentially provide rich phenomenology about the physics in the early universe.

The leading indicator of non-Gaussianity is the three-point correlation function of curvature perturbations, which defines the bispectrum B_ζ as

$$\left\langle \zeta_{\vec{k}_1}\, \zeta_{\vec{k}_2}\, \zeta_{\vec{k}_3} \right\rangle \equiv B_\zeta\big(k_1, k_2, k_3\big)\, \delta^{(3)}\big(\vec{k}_1 + \vec{k}_2 + \vec{k}_3\big) \tag{38}$$

where $k_i \equiv |\vec{k}_i|$, and the delta function is due to translational invariance of a theory. Due to the constraint $\vec{k}_1 + \vec{k}_2 + \vec{k}_3 = 0$, the modes form a triangle. As a result, B_ζ depends not only on the amplitude, but also on the angle in this triangle and the overall scaling with respect to the momenta. The bispectrum thus provides much more informative fingerprints of the origin of perturbations than the power spectrum. A standard parametrization of the level of non-Gaussianity is with the use of *non-linearity parameter* $f_{\rm NL}$. Under local ansatz, the curvature perturbation in the coordinate space may be expanded as [193]

$$\zeta(x) = \zeta_g(x) + \frac{3}{5} f_{\rm NL} \left[\zeta_g^2(x) - \langle \zeta_g^2(x) \rangle\right] + \dots \tag{39}$$

where ζ_g is a Gaussian random field, and the $\dots$ indicates the higher-order terms in the expansion. In the case of single-scalar-field inflationary models, $f_{\rm NL}$ is suppressed by the slow-roll parameters and is too small to be detectable [8, 52, 135, 223, 263, 269]. In fact, perturbations of (almost) free fields originated from the ground state of an adiabatic vacuum are expected to be (nearly) Gaussian. To have a large non-Gaussianity, violation of a single field, canonical kinetic terms, slow roll, or initial adiabatic Bunch-Davies vacuum is typically necessary [191].[11] In this thesis, we explore yet another possibility of the effects

[10] The recent observation [12] finds $\mathcal{A}_s \cong 2.23 \times 10^{-9}$, which in the single-scalar models indicates $V^{1/4} \simeq 0.03\, \epsilon^{1/4} M_p \ll M_p$, which justifies the semi-classical treatment of gravity, neglecting the corrections from quantum gravity, briefly discussed in Subsection II 2.1.

[11] A small subset of a number of models that have been proposed to produce a detectable non-Gaussian signature is: using sound speed effects [88–90], higher derivatives [36, 39, 41], non-vacuum initial conditions [88–90, 170, 233, 234], sharp potential features [88, 89], post-inflationary effects [38, 40, 68]

from interactions between fields during inflation and show that they can feed detectable levels of non-Gaussianity. Within this context, $f_{\rm NL}$ serves as a measure of the strength of such interactions. Non-Gaussianity is a powerful discriminator between different models that would otherwise give degenerate predictions, which is the very reason why it has evoked a vast amount of explorations in the recent years.

The expansion (39) with local ansatz gives the bispectrum of the form

$$B_\zeta^{\rm local} \propto \frac{1}{k_1^3 k_2^3} + \frac{1}{k_2^3 k_3^3} + \frac{1}{k_3^3 k_1^3} \tag{40}$$

which is the case for the single-field models. This form of bispectrum is peaked at squeezed triangles, e.g. $k_1 \ll k_2 \simeq k_3$ (or permutations). However, models of different physics give different forms. Analyzing data for each form one by one is computationally challenging, so in the CMB data analysis, some "templates" of bispectrum are prepared and employed to give the observational bounds. Two other commonly used templates are "equilateral" and "orthogonal" configurations, whose templates are [192, 272]

$$B_\zeta^{\rm equil} \propto -\frac{1}{k_1^3 k_2^3} - \frac{1}{k_2^3 k_3^3} - \frac{1}{k_3^3 k_1^3} - \frac{2}{k_1^2 k_2^2 k_3^2} + \frac{1}{k_1 k_2^2 k_3^3} + (5 \text{ perms}) \tag{41}$$

$$B_\zeta^{\rm orth} \propto -\frac{3}{k_1^3 k_2^3} - \frac{3}{k_2^3 k_3^3} - \frac{3}{k_3^3 k_1^3} - \frac{8}{k_1^2 k_2^2 k_3^2} + \frac{3}{k_1 k_2^2 k_3^3} + (5 \text{ perms}) \tag{42}$$

where the permutations act only on the last term of each equation. The "equilateral" bispectrum is peaked at the equilateral shape of the triangle, i.e. $k_1 = k_2 = k_3$. The observations constrain the values of non-linearity parameter using these templates (40–42), denoting $f_{\rm NL}^{\rm local}$, $f_{\rm NL}^{\rm equil}$, and $f_{\rm NL}^{\rm orth}$, respectively.

Another observable quantity in the cosmological perturbations is the spectrum of gravitational waves (GWs). They are the tensor modes of the metric perturbations; they have two polarization states, as graviton is massless, propagating at the speed of light. We denote them as h_λ, where $\lambda = L, R$, or $\lambda = +, \times$, is the polarization index. The two-point correlation function again defines the power spectrum of GWs as

$$\left\langle h_\lambda(\vec{k})\, h_{\lambda'}(\vec{k}')\right\rangle = \frac{2\pi^2}{k^3}\, P_\lambda(k)\, \delta_{\lambda\lambda'}\, \delta^{(3)}(\vec{k} + \vec{k}') \tag{43}$$

where $\delta_{\lambda\lambda'}$ is the Kronecker delta, showing that the two polarization modes in GWs do not correlate, which is generic due to the statistical isotropy and the spin-2 nature of GWs. Here, both translational and rotational invariances are assumed. The total GW spectrum is then

$$P_{\rm GW} = P_L + P_R \tag{44}$$

and the strength of the tensor spectrum is quantified by the tensor-to-scalar ratio

$$r \equiv \frac{P_{\rm GW}}{P_\zeta} \ . \tag{45}$$

In the single-scalar-field inflationary models, we have $P^L = P^R$ and

$$P_{\rm GW} \cong \frac{2H^2}{\pi^2 M_p^2} \left(\frac{k}{k_*}\right)^{2\epsilon} \tag{46}$$

giving

$$r \cong 16\epsilon \ . \tag{47}$$

Notice from (37) that in the single-scalar models, the observed value of $\mathcal{A}_s$ determines the combination $\frac{V}{\epsilon}$, and that of n_s the combination $2\eta - 6\epsilon$; thus, we need the detection of r, in order to completely determine all the inflationary parameters of the simplest models.

These predictions of curvature perturbations and GWs have made the inflationary cosmology a testable theoretical framework. Exploring the models of inflation is not only an attempt to reveal the physics of the primordial universe, but also the study of the elementary particle theory of extremely high energies, which the current, or near future, accelerators cannot possibly achieve. In the next section, we briefly summarize the results of the recent observations relevant for the constraints on inflationary parameters.

3 Recent Observational Results

The CMB anisotropies are widely considered one of the most powerful probes of the early-universe cosmology. Recent advances of the CMB observations are largely based on a series of space-based missions: the Cosmic Background Explorer (COBE) [282], the Wilkinson Microwave Anisotropy Probe (WMAP) [60, 285], and Planck [11]. The COBE results first established the firm evidence of a nearly scale-invariant spectrum of primordial fluctuations in the CMB, on angular scales larger than 7°. These findings were consistent with the predictions from inflationary models and stimulated further investigation of the subject. The WMAP results (see [61, 167] for the final WMAP analysis results), together with the results from high-resolution ground-based CMB experiments (e.g. [257, 280, 289]), are in an excellent agreement with the predictions from the so-called ΛCDM model, which is now often regarded as a "standard" cosmological model.[12] The combination of precise measurements and accurate theoretical predictions allowed the WMAP results to set a tight constraint on cosmological parameters.

Planck is the third-generation CMB space mission, primarily aiming the measurements of the fluctuations of the temperature and polarization in the CMB radiation at μK sensitivity per resolution over the entire sky [5]. Planck covers a wide frequency range $(30 - 857$ GHz) to accurately discriminate Galactic emission from the primordial fluctuations. Planck's high sensitivity, wide frequency range and all-sky coverage have enabled to put the most stringent observational bounds on cosmological parameters, including the ones of inflationary cosmology.

The first Planck results were released in March 2013.[13] We below list the summaries of the most recent observational findings and parametric constraints that are relevant for inflation [10–13].

- The amplitude of the scalar power spectrum, defined in (36) is measured to be $\mathcal{A}_s = (2.23 \pm 0.16) \times 10^{-9}$ with Planck only, and $A_s = 2.23^{+0.051}_{-0.060} \times 10^{-9}$ with Planck + WMAP polarization, both at 1σ. The pivot scale is taken at $k_* = 0.05\,\mathrm{Mpc}^{-1}$.

- The scalar spectral index of curvature perturbations, defined in (36), is found to be $n_s = 0.9603 \pm 0.0073$ at 1σ using the data from Planck and WMAP polarization. The pivot scale is again $k_* = 0.05\,\mathrm{Mpc}^{-1}$. This value of n_s rules out exact scale invariance $n_s = 1$ at over 5σ.

- There is no significant running of n_s observed, with $\frac{dn_s}{d\ln k} = -0.0134 \pm 0.0090$ at the pivot scale $k_* = 0.05\,\mathrm{Mpc}^{-1}$.

- The fitting of the data to the ΛCDM model with r being as an additional parameter gives the constraint on the tensor-to-scalar ratio (45): $r < 0.11$ at 95% CL with the pivot scale $k_* = 0.002\,\mathrm{Mpc}^{-1}$.

- The Planck data shrinks the allowed parameter space of standard inflationary models. The inflaton potential $V(\varphi)$ is preferred to have the form with $V_{\varphi\varphi} < 0$. Moreover,

[12] The ΛCDM model is based on the assumptions of a spatially-flat, expanding universe. The dynamics of the universe is governed by General Relativity, and its energy content is dominated by cold dark matter (CDM) and a cosmological constant (Λ), from which its name comes. The seeds of structure formation are taken to be Gaussian-distributed adiabatic fluctuations with a nearly scale-invariant spectrum of primordial origin. The model is specified by only 6 key parameters, and yet it has shown to be a successful description of various cosmological data.

[13] This release does not include the CMB polarization data.

the models of an exponential potential, the simplest hybrid inflationary models, and the models of a monomial potential $V(\varphi) \propto \varphi^p$ with $p \geq 2$ are disfavored as they do not give a good fit to the Planck data.

- Planck, together with other preceding missions, has put the most stringent constraints on non-Gaussianity at an unprecedented level. The bounds on the amplitudes of non-linearity parameter (39) at local, equilateral, and orthogonal configurations are now, respectively, $f_{\rm NL}^{\rm local} = 2.7 \pm 5.8$, $f_{\rm NL}^{\rm equil} = -42 \pm 75$, and $f_{\rm NL}^{\rm orth} = -25 \pm 39$, all at 68% CL.

- If the adiabatic condition is violated,[14] isocurvature modes can also be excited, possibly correlated among themselves and with the curvature (adiabatic) modes. The Planck results put upper bound on the fractional primordial contributions of CDM isocurvature modes of the types expected in the curvaton and axion scenarios, which are 0.25% and 3.9%, respectively, both at 95% CL, at the scale $k = 0.05$ Mpc^{-1}. The Planck data are consistent with adiabatic initial conditions.

The simplest inflationary models have passed the test with the stringent Planck results, with a narrowed parameter space [12]. The full mission data including the Planck polarization data, which have not been included in the current Planck results, will be expected to improves precision and statistical significance. Yet, anomalies such as the low-multipole (corresponding to largest scales) signals in the temperature fluctuations that are significantly lower than the predicted values may indicate deviations from the simplest scenarios and extensions to more complex models. Although non-Gaussianity is still consistent with non-detection, the bounds on it are not yet strong enough to be decisive, especially for the equilateral and orthogonal configurations. Non-Gaussianity is one of the major indicators of deviations from the simplest models and can be a powerful tool to discriminate between different inflationary models. In the following parts of this thesis, we explore the possibilities of interaction effects that lead to particle production during inflation, which potentially generate large non-Gaussianity or detectable gravitational wave signals. If such signals should be detected, it would greatly enhance our understanding of the physics in the very early universe and, equivalently, the physics of extremely high energies.

[14] The adiabatic condition is to say that at very early stages the universe obeyed a common, spatially-uniform equation of state, and all components initially shared a common velocity field [12]. In other words, the quantity $\frac{\delta \rho_i}{\dot{\rho}_i}$ takes the same value for all the matter component individually of the universe (the subscript i denotes different components). If this condition is satisfied, there is no particle flow and hence no heat flow between different locations, which justifies its name.

Part III

General Formalism of Particle Production

This thesis is devoted to the study of the mechanism of particle production during inflation and its phenomenology. In particular, we consider production of vector fields due to their coupling to rolling scalar/pseudo scalar fields, either the inflaton or some non-inflaton field. These fields modify the dispersion relation of the vector field, which leads to tachyonic or non-adiabatic growth, depending on the type of the coupling. The abundantly produced vector particles in turn source the cosmological perturbations through the same coupling or gravitational interaction. These sourced perturbations are uncorrelated with the standard modes originated from vacuum fluctuations and are expected to be highly non-Gaussian. Primordial non-Gaussianity, if detected in the CMB and LSS observations, can potentially be a powerful tool to probe the features of particle interactions during inflation. In this Part, we formulate the general formalism of vector particle production and its sourcing effects on the cosmological perturbations.

We assume that quanta of a vector field are produced through some specific coupling during inflation (either at some discrete moment t_*, or throughout inflation). The vector field A_μ can be decomposed as, taking the Coulomb gauge,

$$
\begin{aligned}
&A_0 = 0\,, \quad A_i\big(\tau,\vec{x}\big) = \int \frac{d^3k}{\left(2\pi\right)^{3/2}} \mathrm{e}^{i\vec{k}\cdot\vec{x}}\,\tilde{A}_i\big(\tau,\,\vec{k}\big) \\
&\tilde{A}_i\big(\tau,\,\vec{k}\big) = \sum_{\lambda=\pm} \epsilon_i^{(\lambda)}\big(\hat{k}\big)\left[a_\lambda\big(\vec{k}\big)\,A_\lambda\big(\tau,k\big) + a_\lambda^\dagger\big(-\vec{k}\big)\,A_\lambda^*\big(\tau,k\big)\right]
\end{aligned} \tag{48}
$$

where we have included only the transverse vector polarization (in Section VI 1, we actually study also the longitudinal polarization, as it is present, and may be relevant, in that model). Here we denote $k \equiv |\vec{k}|$, and $\vec{\epsilon}_\lambda$ are circular polarization vectors satisfying $\vec{k}\cdot\vec{\epsilon}^{\,(\pm)}\big(\hat{k}\big) = 0$, $\vec{k}\times\vec{\epsilon}^{\,(\pm)}\big(\hat{k}\big) = \mp ik\vec{\epsilon}^{\,(\pm)}\big(\hat{k}\big)$, $\vec{\epsilon}^{\,(\pm)}\big(-\hat{k}\big) = \vec{\epsilon}^{\,(\pm)\,*}\big(\hat{k}\big)$, and normalized according to $\vec{\epsilon}^{(\lambda)\,*}\big(\hat{k}\big)\cdot\vec{\epsilon}^{(\lambda')}\big(\hat{k}\big) = \delta_{\lambda\lambda'}$. Also, a_λ and $a_\lambda^\dagger$ are the annihilation and creation operators, respectively, for the transverse vector modes, which satisfy the commutation relation $\left[a_\lambda\big(\vec{k}\big), a_{\lambda'}^\dagger\big(\vec{k}'\big)\right] = \delta_{\lambda\lambda'}\,\delta^{(3)}\big(\vec{k}-\vec{k}'\big)$, where $\delta_{\lambda\lambda'}$ is the Kronecker delta function.

In the following Sections, we obtain the formal expressions for the observable correlators of the scalar (ζ) and tensor perturbations (h_λ). These expressions are generic for any types of couplings we study, and can be readily applied to any given model. We will apply this formalism in the remainder of the thesis. The discussion is divided into three sections; in Section III 1 we present the structure of the equations for the scalar and tensor perturbations in presence of the source; in Section III 2 we present the formal solutions, and the structures of the correlators; in Section III 3 we show how these expressions enter in observable quantities.

1 Decompositions, and Sourced Equations for ζ, h_λ

We compute scalar cosmological perturbations in the spatially flat gauge $\delta g_{ij,\text{scalar}} = 0$. In this gauge, the curvature perturbation ζ is (up to negligible corrections) $\zeta = -\frac{H}{\dot{\varphi}^{(0)}}\,\delta\varphi$, where $\varphi^{(0)}$ and $\delta\varphi$ are the unperturbed and perturbed part of the inflaton, respectively. We decompose

$$\delta\varphi(\tau, \vec{x}) = \int \frac{d^3k}{(2\pi)^{3/2}}\, \mathrm{e}^{i\vec{k}\cdot\vec{x}}\, \frac{Q_\varphi(\tau, \vec{k})}{a(\tau)} \tag{49}$$

As we show in the following parts of this thesis, the effect of the vector fields on the inflaton modes can be encoded in the approximate equation

$$\left[\partial_\tau^2 + \left(k^2 - \frac{a''}{a}\right)\right] Q_\varphi(\tau,\, \vec{k}) \simeq J_\varphi(\tau,\, \vec{k}) \tag{50}$$

where $\partial_\tau \equiv \frac{\partial}{\partial\tau}$, and the source is formally of the type

$$J_\varphi(\tau,\, \vec{k}) \equiv \int \frac{d^3p}{(2\pi)^{3/2}}\, \hat{O}_{\varphi,ij}(\tau,\, \vec{k},\, \vec{p})\, \tilde{A}_i(\tau,\, \vec{p})\tilde{A}_j(\tau,\, \vec{k} - \vec{p}) \tag{51}$$

where $\hat{O}_{\varphi,ij}$ is a model-dependent operator.

In the tensor sector, the gravity waves are encoded in $g_{ij} = a^2\,(\delta_{ij} + h_{ij})$, where the modes h_{ij} are transverse and traceless. We introduce the canonical modes

$$h_{ij}(\tau, \vec{x}) = \frac{2}{M_p\, a(\tau)} \int \frac{d^3k}{(2\pi)^{3/2}}\, \mathrm{e}^{i\vec{k}\cdot\vec{x}} \sum_{\lambda=L,R} \Pi_{ij}^{(\lambda)}(\hat{k})\, Q_\lambda(\tau, \vec{k}) \tag{52}$$

where $\Pi_{ij}^{(R/L)}(\hat{k}) = \epsilon_i^{(\pm)}(\hat{k})\, \epsilon_j^{(\pm)}(\hat{k})$ are the helicity tensors. The canonical modes Q_λ obey the equation of motion

$$\left[\partial_\tau^2 + \left(k^2 - \frac{a''}{a}\right)\right] Q_\lambda(\tau,\, \vec{k}) = J_\lambda(\tau,\, \vec{k}) \tag{53}$$

The source is obtained by starting from the transverse and traceless spatial part of the energy momentum tensor and projecting along the λ polarization with $\Pi_{ij}^{(\lambda)}$. One finds

$$J_\lambda\big(\tau,\,\vec{k}\big) = \Pi_{ij}^{(\lambda)\,*}\big(\hat{k}\big)\int\frac{d^3x}{(2\pi)^{3/2}}\,\mathrm{e}^{-i\vec{k}\cdot\vec{x}}\,\frac{a}{M_p}\,T_{ij}\,(\tau,\,\vec{x}) \tag{54}$$

We note that, the multiplication by $\Pi_{ij}^{(\lambda)\,*}$ automatically projects on the transverse and traceless part of T_{ij}. If the energy-momentum tensor is quadratic in the gauge fields, as in the cases that we will consider, we recover an expression formally identical to (51), with a different (and model-dependent) operator $\hat{O}_{\lambda,ij}\big(\tau,\,\vec{k},\,\vec{p}\big)$.

In summary, we have formally identical equations for the scalar and tensor canonical modes

$$\begin{aligned}\left[\partial_\tau^2+\left(k^2-\frac{a''}{a}\right)\right]Q_X\big(\tau,\,\vec{k}\big) &\simeq J_X\big(\tau,\,\vec{k}\big)\\ J_X\big(\tau,\,\vec{k}\big) &\equiv \int\frac{d^3p}{(2\pi)^{3/2}}\,\hat{O}_{X,ij}\big(\tau,\,\vec{k},\,\vec{p}\big)\,\tilde{A}_i\big(\tau,\,\vec{p}\big)\tilde{A}_j\big(\tau,\,\vec{k}-\vec{p}\big)\end{aligned} \tag{55}$$

where $X=\{\varphi,\,\lambda=L/R\}$. The operators $\hat{O}_X$ depend on what X is, and also on the model. In all cases, they are invariant under the simultaneous $i\leftrightarrow j$ and $\vec{p}\to\vec{k}-\vec{p}$ operations.

2 Formal Solutions and Correlators

The equation (55) is formally solved by

$$Q_X\big(\tau,\,\vec{k}\big) = Q_X^{\rm vac}\big(\tau,\,\vec{k}\big)+Q_X^{\rm sourced}\big(\tau,\,\vec{k}\big) \tag{56}$$

where $Q_X^{\rm vac}$ is the standard vacuum solution, obtained in absence of the vector source, and $Q_X^{\rm sourced}$ is the particular solution of (55), corresponding to the sourced contribution.

The homogeneous vacuum term is expanded as

$$Q_X^{\rm vac}(\tau,\vec{k}) = b_X(\vec{k})\,\phi_X(\tau,k)+b_X^\dagger(-\vec{k})\,\phi_X^*(\tau,k)\;. \tag{57}$$

The creation/annihilation operators of the particle $X=\varphi,L/R$ obey

$$\left[b_X(\vec{k}),b_Y^\dagger(\vec{k}')\right] = \delta_{XY}\delta^{(3)}(\vec{k}-\vec{k}') \tag{58}$$

and commute with the creation/annihilation operators of the gauge field

$$\left[b_X(\vec{k}), a_\lambda(\vec{k}')\right] = \left[b_X(\vec{k}), a_\lambda^\dagger(\vec{k}')\right] = 0 \tag{59}$$

The properly normalized homogeneous solutions of (55) are given by the well-known result

$$\phi_X(\tau, k) = i\frac{\sqrt{\pi}}{2}\sqrt{-\tau}H_\nu^{(1)}(-k\tau)\ , \quad \nu \cong \frac{3}{2} + \mathcal{O}(\epsilon, \eta) \tag{60}$$

where $H_\nu^{(1)}$ is the Hankel function of the first kind, and we have chosen the (arbitrary) phase so that $\phi_X(\tau, k)$ is real in the limit $-k\tau \to 0$.

The vacuum modes (60) are employed to construct the retarded Green function associated with (55),

$$\begin{aligned} G_k(\tau, \tau') &= i\Theta(\tau - \tau')\left[\phi_X(\tau, k)\phi_X^*(\tau', k) - \phi_X^*(\tau, k)\phi_X(\tau', k)\right] \\ &\simeq \frac{1}{k^3\tau\tau'}\left[k\tau' \cos\left(k\tau'\right) - \sin\left(k\tau'\right)\right] \end{aligned} \tag{61}$$

where Θ is the step function. The Green function obeys $\left[\partial_\tau^2 + \left(k^2 - \frac{a''}{a}\right)\right] G_k(\tau, \tau') = \delta\left(\tau - \tau'\right)$, and its approximate form (second line) is valid in the super-horizon regime $-k\tau \ll 1$. Using the Green function (61) the particular solution of (55) takes the form

$$Q_X^{\text{sourced}}(\tau, \vec{k}) = \int_{-\infty}^{\tau} d\tau' G_k(\tau, \tau') J_X(\tau', \vec{k}) \tag{62}$$

where the source term was defined in (55). We note that this particular solution is statistically independent of the homogeneous solution (57). In fact, the particular solution can be expanded in terms of the annihilation/creation operators $a_\lambda(\vec{k}), a_\lambda^\dagger(\vec{k})$ associated with the gauge field, while the homogeneous solution can be expanded in terms of the annihilation/creation operators $b_X(\vec{k}), b_X^\dagger(\vec{k})$ associated with the vacuum fluctuations of particle X. As we pointed out, these two sets of operators commute with one another.

We are interested in the two- and three-point correlators of the sourced solutions. For the sourced two-point correlator, we find

$$\begin{aligned} \left\langle Q_X^{\text{sourced}}\left(\tau, \vec{k}_1\right) Q_Y^{\text{sourced}}\left(\tau, \vec{k}_2\right)\right\rangle &= \int^\tau d\tau_1\, G_{k_1}\left(\tau, \tau_1\right) \int^\tau d\tau_2\, G_{k_2}\left(\tau, \tau_2\right) \\ &\quad \times \left\langle J_X\left(\tau_1, \vec{k}_1\right) J_Y\left(\tau_2, \vec{k}_2\right)\right\rangle \end{aligned} \tag{63}$$

where

$$\left\langle J_X(\tau_1,\,\vec{k}_1)J_Y(\tau_2,\,\vec{k}_2)\right\rangle \;=\; 2\int\frac{d^3p_1d^3p_2}{(2\pi)^3}\,\hat{O}_{X,ij}(\tau_1,\,\vec{k}_1,\,\vec{p}_1)\hat{O}_{Y,lm}(\tau_2,\,\vec{k}_2,\,\vec{p}_2)$$
$$\times\left\langle \tilde{A}_i(\tau_1,\,\vec{p}_1)\tilde{A}_m(\tau_2,\,\vec{k}_2-\vec{p}_2)\right\rangle\left\langle \tilde{A}_j(\tau_1,\,\vec{k}_1-\vec{p}_1)\tilde{A}_l(\tau_2,\,\vec{p}_2)\right\rangle \tag{64}$$

In writing this expression, we Wick-decomposed the $\langle A^4\rangle$ correlator coming from $\langle J^2\rangle$, and then used the symmetry of the $\hat{O}_{Y,lm}$ operator. From (48) we then see that the sourced part of the two-point correlator of the gauge field is formally of the type

$$\left\langle \tilde{A}_i\left(\tau_1,\,\vec{q}_1\right)\tilde{A}_j\left(\tau_2,\,\vec{q}_2\right)\right\rangle = \sum_{\sigma=\pm}\mathcal{P}_{ij}^{(\sigma)}\left(\hat{q}_1\right)\mathcal{D}_{(0,0)}^{(\sigma)}\left[\tau_1,\,\tau_2;\,q_1\right]\delta^{(3)}\left(\vec{q}_1+\vec{q}_2\right) \tag{65}$$

where

$$\mathcal{P}_{ij}^{(\pm)}\left(\hat{p}\right)\equiv\epsilon_i^{(\pm)}\left(\vec{p}\right)\epsilon_j^{(\pm)*}\left(\vec{p}\right)=\frac{1}{2}\left(\delta_{ij}-\hat{p}_i\,\hat{p}_j\right)\mp\frac{i}{2}\epsilon_{ijk}\,\hat{p}_k \tag{66}$$

The index $(0,0)$ on $\mathcal{D}$ indicates that no time derivative is acting on the modes $\tilde{A}$ on the left-hand side of (65), and it is introduced for later convenience. Using the hermiticity of $A_\mu\left(x\right)$, one can show that $\mathcal{D}_{(0,0)}^{(\sigma)}\left[\tau_1,\,\tau_2;\,q\right]=\mathcal{D}_{(0,0)}^{(\sigma)*}\left[\tau_2,\,\tau_1;\,q\right]$. As we show in the following Parts, in the models we study, $\mathcal{D}_{(0,0)}^{(\sigma)}$ becomes real and symmetric under $\tau_1\leftrightarrow\tau_2$ for all times after the particle production. Inserting (65) into (64) we obtain

$$\left\langle J_X(\tau_1,\,\vec{k}_1)J_Y(\tau_2,\,\vec{k}_2)\right\rangle = 2\,\delta^{(3)}\big(\vec{k}_1+\vec{k}_2\big)\int\frac{d^3p}{(2\pi)^3}\sum_{\sigma,\sigma'}\mathcal{P}_{im}^{(\sigma)}(\hat{p})\mathcal{P}_{lj}^{(\sigma')}(\widehat{p-k_1})$$
$$\times\hat{O}_{X,ij}(\tau_1,\,\vec{k}_1,\,\vec{p})\,\hat{O}_{Y,lm}(\tau_2,\,-\vec{k}_1,\,\vec{p}-\vec{k}_1)\,\mathcal{D}_{(0,0)}^{(\sigma)}\left[\tau_1,\,\tau_2;\,p\right]\mathcal{D}_{(0,0)}^{(\sigma')}\left[\tau_1,\,\tau_2;\,|\vec{p}-\vec{k}_1|\right] \tag{67}$$

where we denote $\widehat{p-k_i}\equiv\frac{\vec{p}-\vec{k}_i}{|\vec{p}-\vec{k}_i|}$.

Finally, for the sourced three-point function, we only need to compute the connected three-point correlator of Q_φ. Proceeding as for the two-point function we obtain

$$\left\langle\prod_{i=1}^{3}Q_\varphi^{\rm sourced}(\tau,\,\vec{k}_i)\right\rangle=\int^\tau d\tau_1 G_{k_1}\left(\tau,\,\tau_1\right)\int^\tau d\tau_2 G_{k_2}\left(\tau,\,\tau_2\right)\int^\tau d\tau_3 G_{k_3}\left(\tau,\,\tau_3\right)\left\langle\prod_{i=1}^{3}J_\varphi(\tau_i,\,\vec{k}_i)\right\rangle \tag{68}$$

with

$$\left\langle \prod_{i=1}^{3} J_\varphi(\tau_i,\, \vec{k}_i) \right\rangle = 8\, \delta^{(3)}(\vec{k}_1 + \vec{k}_2 + \vec{k}_3) \int \frac{d^3p}{(2\pi)^{9/2}} \sum_{\sigma,\sigma',\sigma''} \mathcal{P}^{(\sigma)}_{im}(\hat{p}) \mathcal{P}^{(\sigma')}_{nj}(\widehat{p - k_1}) \mathcal{P}^{(\sigma'')}_{lo}(\widehat{p + k_2})$$
$$\times \hat{O}_{\varphi,ij}(\tau_1,\, \vec{k}_1,\, \vec{p})\, \hat{O}_{\varphi,lm}(\tau_2,\, \vec{k}_2,\, \vec{p} + \vec{k}_2)\, \hat{O}_{\varphi,no}(\tau_3,\, \vec{k}_3,\, \vec{p} - \vec{k}_1)$$
$$\times \mathcal{D}^{(\sigma)}_{(0,0)}[\tau_1,\, \tau_2;\, p]\, \mathcal{D}^{(\sigma')}_{(0,0)}[\tau_1,\, \tau_3;\, |\vec{p} - \vec{k}_1|]\, \mathcal{D}^{(\sigma'')}_{(0,0)}[\tau_2,\, \tau_3;\, |\vec{p} + \vec{k}_2|]\, . \tag{69}$$

where the factor 8 again comes from the Wick contraction and the symmetry of the $\hat{O}_{\varphi,ij}$ operator.

Once we specify a model, we can find the model-dependent operators $\hat{O}_{X,ij}$ and then compute the scalar and tensor correlation functions by (63 and (68) together with (67) and (69). Using these correlators, we can now compute the quantities relevant to observations.

3 Phenomenology

From the two-point scalar correlator we obtain the curvature power spectrum P_ζ,

$$P_\zeta\, \delta^{(3)}(\vec{k}_1 + \vec{k}_2) = \frac{k_1^3}{2\pi^2} \left\langle \zeta_{\vec{k}_1}\, \zeta_{\vec{k}_2} \right\rangle\, . \tag{70}$$

where $\zeta_{\vec{k}}$ is the Fourier transform of the curvature perturbations $\zeta = -\frac{H}{\dot{\varphi}^{(0)}}\delta\varphi$. The two terms in the solutions (56) are uncorrelated, and therefore their power spectra add up. Recalling the standard slow-roll result (37) for the vacuum mode, and using the two expressions (63) and (67), we obtain

$$P_\zeta = P_\zeta^{\rm vac} + P_\zeta^{\rm sourced}$$
$$P_\zeta^{\rm vac} = \left(\frac{H^2}{2\pi \dot{\varphi}^{(0)}} \right)^2$$
$$P_\zeta^{\rm sourced} = \frac{k_1^3}{\pi^2} \frac{H^2}{a^2 \dot{\varphi}^{(0)\,2}} \int^\tau d\tau_1 G_{k_1}(\tau,\, \tau_1) \int^\tau d\tau_2 G_{k_1}(\tau,\, \tau_2) \int \frac{d^3p}{(2\pi)^3} \sum_{\sigma,\sigma'} \mathcal{P}^{(\sigma)}_{im}(\hat{p}) \mathcal{P}^{(\sigma')}_{lj}(\widehat{p - k_1})$$
$$\times \hat{O}_{\varphi,ij}(\tau_1,\, \vec{k}_1,\, \vec{p})\, \hat{O}_{\varphi,lm}(\tau_2,\, -\vec{k}_1,\, \vec{p} - \vec{k}_1)\, \mathcal{D}^{(\sigma)}_{(0,0)}[\tau_1,\, \tau_2;\, p]\, \mathcal{D}^{(\sigma')}_{(0,0)}[\tau_1,\, \tau_2;\, |\vec{p} - \vec{k}_1|] \tag{71}$$

For the tensor mode, starting from the decomposition (52), the GW power spectrum P_λ is

$$P_\lambda\, \delta_{\lambda\lambda'}\, \delta^{(3)}(\vec{k}_1 + \vec{k}_2) = \frac{k_1^3}{2\pi^2} \frac{4}{a^2 M_p^2} \left\langle Q_\lambda(\vec{k}_1)\, Q_{\lambda'}(\vec{k}_2) \right\rangle \tag{72}$$

where $\delta_{\lambda\lambda'}$ implies that modes of different polarizations are uncorrelated, which is a generic feature due to the statistical isotropy and the spin-2 nature of GWs. The vacuum and the sourced modes in the power spectrum add up as in the scalar case, giving

$$
\begin{aligned}
& P_\lambda = P_\lambda^{\rm vac} + P_\lambda^{\rm sourced} \\
& P_\lambda^{\rm vac} = \frac{H^2}{\pi^2 M_p^2} \\
& P_\lambda^{\rm sourced} = \frac{k_1^3}{\pi^2} \frac{4}{a^2 M_p^2} \int^\tau d\tau_1 G_{k_1}(\tau,\, \tau_1) \int^\tau d\tau_2 G_{k_1}(\tau,\, \tau_2) \int \frac{d^3p}{(2\pi)^3} \sum_{\sigma,\sigma'} \mathcal{P}_{im}^{(\sigma)}(\hat{p}) \mathcal{P}_{lj}^{(\sigma')}(\widehat{p-k_1}) \\
& \quad \times \hat{O}_{\lambda,ij}(\tau_1,\, \vec{k}_1,\, \vec{p})\, \hat{O}_{\lambda,lm}(\tau_2,\, -\vec{k}_1,\, \vec{p}-\vec{k}_1)\, \mathcal{D}_{(0,0)}^{(\sigma)}[\tau_1,\, \tau_2;\, p]\, \mathcal{D}_{(0,0)}^{(\sigma')}[\tau_1,\, \tau_2;\, |\vec{p}-\vec{k}_1|]
\end{aligned}
\tag{73}
$$

The tensor-to-scalar ratio r is defined as

$$
r = \frac{\sum_\lambda P_\lambda}{P_\zeta}
\tag{74}
$$

If the sourced term is absent, one recovers the standard vacuum slow-roll result $r_{\rm v} = \frac{8\,\dot{\varphi}^{(0)2}}{H^2 M_p^2} = 16\epsilon$, (47). However, we see that the sourced contribution can modify this result.

Finally, we are interested in the bispectum B_ζ of the scalar modes, defined as in (38),

$$
B_\zeta(\vec{k}_i)\, \delta^{(3)}(\vec{k}_1 + \vec{k}_2 + \vec{k}_3) = \left\langle \zeta_{\vec{k}_1}\, \zeta_{\vec{k}_2}\, \zeta_{\vec{k}_3} \right\rangle .
\tag{75}
$$

We disregard the contribution from the vacuum modes, as it is unobservable for standard slow roll inflation. Using (68) and (69) we obtain

$$
\begin{aligned}
B_\zeta\left(\vec{k}_i\right) &= 8 \int^\tau d\tau_1\, G_{k_1}(\tau,\, \tau_1) \int^\tau d\tau_2\, G_{k_2}(\tau,\, \tau_2) \int^\tau d\tau_3\, G_{k_3}(\tau,\, \tau_3) \\
& \times \int \frac{d^3p}{(2\pi)^{9/2}} \sum_{\sigma,\sigma',\sigma''} \mathcal{P}_{im}^{(\sigma)}(\hat{p})\, \mathcal{P}_{nj}^{(\sigma')}(\widehat{p-k_1})\, \mathcal{P}_{lo}^{(\sigma'')}(\widehat{p+k_2}) \\
& \times \hat{O}_{\varphi,ij}(\tau_1,\, \vec{k}_1,\, \vec{p})\, \hat{O}_{\varphi,lm}(\tau_2,\, \vec{k}_2,\, \vec{p}+\vec{k}_2)\, \hat{O}_{\varphi,no}(\tau_3,\, \vec{k}_3,\, \vec{p}-\vec{k}_1) \\
& \times \mathcal{D}_{(0,0)}^{(\sigma)}[\tau_1,\, \tau_2;\, p]\, \mathcal{D}_{(0,0)}^{(\sigma')}[\tau_1,\, \tau_3;\, |\vec{p}-\vec{k}_1|]\, \mathcal{D}_{(0,0)}^{(\sigma'')}[\tau_2,\, \tau_3;\, |\vec{p}+\vec{k}_2|] .
\end{aligned}
\tag{76}
$$

Then the non-linearity parameter, defined in (39), can be expressed in terms of B_ζ and P_ζ as

$$
f_{\rm NL} = \frac{10}{3\,(2\pi)^{5/2}} \frac{\prod_i k_i^3}{\sum_i k_i^3} \frac{B_\zeta}{P_\zeta^2} .
\tag{77}
$$

In the following Parts, we apply this generic formalism to specific models and study their observational signatures. The effects from particle production during inflation, encoded mostly in r and f_{NL}, can potentially leave very distinct signals and thus provide us rich phenomenology of the physics in the primordial universe.

Part IV
Axion Inflation

Primordial inflation has become the dominant paradigm for the early universe, as it provides a simple mechanism to resolve the conceptual difficulties of the standard Big Bang cosmology and has enjoyed great phenomenological success in accounting for the properties of the observed CMB anisotropies. In spite of this success, however, a compelling particle physics realization is still lacking. A serious obstacle is the requirement of an extremely flat potential of inflaton, $V(\varphi)$, and its stability against quantum corrections. Successful inflation requires $\epsilon, |\eta| \ll 1$, where the slow roll parameters are defined in (24). These parameters are notoriously sensitive to Ultra-Violet (UV) physics. For example, even generic Planck-suppressed corrections to $V(\varphi)$ may contribute $\Delta\eta = \mathcal{O}(1)$, thus spoiling inflation. This UV sensitivity represents a technical fine tuning problem which must be addressed in any particle physics model of inflation.

It is conceivable that dangerous corrections to ϵ, η may be absent as a result of fine-tuning [57], or that the requirement of a flat potential can be evaded by invoking somewhat exotic effects such as dissipation [62, 150], small sound speed [18] or higher derivative corrections [36, 39, 41]. However, perhaps the simplest and most cogent way to realize $\epsilon, |\eta| \ll 1$ in a natural way is by assuming that the inflaton φ is a Pseudo-Nambu-Goldstone-Boson (PNGB) [9, 22, 25, 116, 125, 129, 132, 178, 187, 231]. In this case the inflaton enjoys a shift symmetry $\varphi \to \varphi + \text{const}$, which is broken either explicitly or by quantum effects. In the limit of exact symmetry we must have $\epsilon = \eta = 0$, thus dangerous contributions to the slow roll parameters are controlled by the smallness of symmetry breaking. Moreover, PNGBs like the axion are ubiquitous in particle physics: they arise whenever an approximate global symmetry is spontaneously broken and are plentiful in string theory compactifications.

The idea of invoking a PNGB to obtain a natural realization of inflation is more than 20 years old [132]. Besides advancing this general idea, Ref. [132] also studied the simplest situation in which there is a single axion with potential $V(\varphi) \propto 1 - \cos\left(\frac{\varphi}{f}\right)$, which arises from nonperturbative effects and breaks the continuous shift symmetry down to a discrete subgroup $\varphi \to \varphi + (2\pi)f$. However, this model agrees with observations only if the axion decay constant is $f > M_p$ [265]. It has been debated in the literature weather a trans-Planckian breaking may be compatible with gravity, in the case where the symmetry is global [141–143,169,175,180]. The shift symmetry may emerge from a gauge symmetry, as typically in string theory. However, also in this case, a trans-Planckian f is considered problematic, since all known controlled string theory constructions are characterized by $f < M_p$ [30,291]. Fortunately, a number of models offer the mechanisms to have a scale $f < M_p$ compatible with inflation, while several of them also lead to an interesting phenomenology (see [248] for a recent review). For instance, one can consider more than one axion [63, 116, 187], require nontrivial compactifications in string theory [174, 176, 231, 281], couple the axion to a 4-form [178], modify the axion kinetic term [140, 222, 245], or slow down the axion through particle production [22, 237, 299]. Currently, there exist a number of natural, controlled realizations of axion inflation with an axion scale f a few orders of magnitude smaller than M_p but that nevertheless behave effectively as large field inflation models.

In any axion inflation model, the inflaton is expected to couple to some gauge field A_μ via interactions of the type

$$\mathcal{L}_{\text{int}} = -\frac{\alpha}{4f}\,\varphi\,F_{\mu\nu}\tilde{F}^{\mu\nu} \tag{78}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the field strength tensor and $\tilde{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\alpha\beta}F_{\alpha\beta}$ is its dual. The strength of the interaction is controlled by the decay constant, f, and by the dimensionless parameter α. While α is in principle a model dependent quantity, from the perspective of effective field theory we generally expect it to be order unity. Thus, for controlled effective field theory realizations of axion inflation it is very natural to have $\alpha/f \gtrsim \mathcal{O}\left(10^2\right) M_p^{-1}$, in which case the interaction (78) is much stronger than gravitationally suppressed. In this Part of the thesis, we consider in detail the cosmological implications of the interaction (78). Our analysis is quite general: the general logic of effective field theory requires the inclusion of an interaction (78) whenever φ is pseudo-scalar. We here consider the Abelian case for simplicity, but the same consideration applies to non-Abelian theory, as long as the gauge coupling is not too strong (so that the "gluons" are light).

We here adopt a conservative approach, where the interaction (78) can have a profound impact on the phenomenology of the model, even in the conventional slow roll regime.[15] The underlying physics is as follows. The motion of the inflaton amplifies the fluctuations of the gauge field, δA, which in turn produce inflaton fluctuations via the *same* interaction, through the process of *inverse decay*: $\delta A + \delta A \to \delta\varphi$. When $f/\alpha \lesssim 10^{-2}M_p$, which is very natural in models that admit a UV completion, this new source of perturbations actually dominates over the usual fluctuations from the vacuum. In this regime, all previous studies of axion inflation are invalid. Our analysis is phenomenological: we use the CMB data to place observational limits on the coupling α/f without making any specific assumptions

[15] One can consider some non-standard possibilities. In [22], for example, it was shown that energy dissipation into gauge fields, via the interaction (78), can slow down the motion of the inflaton on a very steep potential.

about the microphysical origins of the model.

Non-Gaussian statistics, such as the bispectrum (38), provide a powerful tool to discriminate between a plethora of inflationary models in the literature and may provide a valuable window into the detailed physics of the very early universe. However, a single decoupled scalar field in slow roll is well known to produce an undetectably small signal [8, 223, 269]. The reason for this is intuitively easy to understand: non-Gaussianity is a measure of the strength of interactions, while in the vanilla scenario the requirement of a flat potential typically also constrains interactions to be weak. To evade this no-go result and obtain an observably large signal, previous studies have introduced non-standard ingredients (refer to Subsection II 2.2). In the models of axion inflation, on the other hand, the no-go results of [8, 223, 269] are circumvented in a very novel way:[16] the interaction that gives rise to large non-Gaussianity, eq. (78), does not play any role in the background dynamics and is thus unconstrained by the requirement of slow roll. In general, interactions between the inflaton and other fields should be small; otherwise, they would give large corrections to the inflaton self interactions $V(\varphi)$ through loops. One thus expect that these interactions should be small, sourcing only small non-Gaussianity. However, if $V(\varphi)$ is protected by a symmetry, such as the shift symmetry for our current consideration, then the interactions with other fields are not restricted to be small, as long as they are perturbative, while the quantum corrections do not spoil the flatness of $V(\varphi)$. Axion inflation is therefore a natural framework for potentially large, observable non-Gaussianity, without appeal to any exotic phenomena or fine tuning.

In terms of the broader picture, this result suggests that the conventional lore concerning the difficulty of obtaining large non-Gaussianity may have been excessively conservative. One generically expects inflaton to couple to some fields which do not play any role in driving inflation, since such interactions are unavoidable from an effective field theory perspective and are (at least to some extent) necessary in order to successfully reheat the universe after inflation. Refs. [35, 42, 43, 48] provide explicit examples demonstrating how the consistent inclusion of such interactions can radically modify the phenomenology of inflation, via particle production effects. The present study represents a challenge to the conventional lore that non-Gaussianity is a "smoking gun" signature of non-standard inflationary dynamics by illustrating explicitly that perhaps the simplest and best-motivated particle physics models of inflation are *already* constrained by existing observational limits on non-Gaussianity.

Another potentially unique observable quantity is gravitational waves (GWs) and their feature. The produced gauge quanta source the GWs in the way similar to the inverse decay

[16] The original work on this was done in [48], followed by a more extensive work [45].

process to curvature perturbations. This process $\delta A + \delta A \to h$ occurs through gravitational interaction, which is weaker than the axion-gauge interaction. Consequently, at the scales relevant to the CMB observations, once the bound on non-Gaussianity is imposed, the standard modes of GWs from vacuum fluctuations dominate over this sourced contribution, and therefore the GW signal at the CMB scales would be only standard. However, the particle production becomes significantly more effective towards the end of inflation. The modes of the cosmological perturbations that are sourced by these produced quanta at later stages of inflation correspond to the scales much smaller than the CMB, more specifically to the scales relevant to terrestrial GW detectors. At small scales, the scalar perturbations may potentially be strongly constrained by primordial black holes, but these constraints carry large theoretical uncertainties [22, 47, 208]. The GW signals contain no such uncertainties and can provide solid constraints on the model.

The *parity-violating* axion-gauge coupling induces the production of only one of the polarization states of the gauge fields, which in turn encodes the *polarized* GWs, under spin conservation. Such a polarized GW signal can be searched by the networks of multiple terrestrial GW detectors. Since the stochastic GW background signals probed by the GW detectors are expected to be unpolarized, detecting polarization asymmetry in the signals is potentially an excellent way of distinguishing the cosmological component from the possibly dominant astrophysical one. The axion inflation models provide a natural mechanism to produce detectable polarized GWs of cosmological origion.

This Part of the thesis is mostly based on the work done in [45]. In Section IV 1, we provide a summary of the mechanism. In Section IV 2, we compute the correlation functions of the curvature fluctuations and tensor (gravitational wave, GW) perturbations. Section IV 3 studies the resulting phenomenology. In Section IV 4, which is based on the work [101], we consider the possibility to detect the *polarized* GWs at the terrestrial GW detectors and discuss the constraint on the axion-gauge coupling strength from the forthcoming detectors. In Section IV 5, we review most of the existing models of axion inflation, and briefly discuss the implications of our findings for these models. Finally, in Section IV 6, we summarize the findings of axion inflation.

1 The Mechanism

In this section we briefly summarize the mechanism of the production of gauge field fluctuations and the subsequent inverse decay effects in axion inflation.

We consider a simple theory of a PNGB inflaton interacting with a U(1) gauge field[17] via the interaction (78). The action is

$$\mathcal{S} = \int d^4x\sqrt{-g}\left[\frac{M_p^2}{2}R - \frac{1}{2}(\partial\varphi)^2 - V(\varphi) - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{\alpha}{4f}\varphi\,\tilde{F}^{\mu\nu}\,F_{\mu\nu}\right] \tag{79}$$

where R is the Ricci scalar, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ the field strength, and $\tilde{F}^{\mu\nu} \equiv \frac{1}{2}\frac{\eta^{\mu\nu\alpha\beta}}{\sqrt{-g}}\,F_{\alpha\beta}$ its dual, with $\eta^{0123} = 1$. We separate the inflaton into a homogeneous (background) part plus its fluctuations

$$\varphi = \varphi^{(0)}(t) + \delta\varphi(t,\,\vec{x}) \tag{80}$$

We leave the potential $V(\varphi)$ arbitrary, except to assume that it is sufficiently flat to support the required amount of inflation ($N_e \gtrsim 60$). We assume a spatially flat Friedmann-Robertson-Walker (FRW) space-time with metric

$$\begin{aligned} ds^2 \equiv g_{\mu\nu}dx^\mu dx^\nu &= -dt^2 + a^2(t)\,d\vec{x}\cdot d\vec{x} && (81)\\ &= a^2(\tau)\left[-d\tau^2 + d\vec{x}\cdot d\vec{x}\right] && (82)\end{aligned}$$

where on the second line we have introduced conformal time τ, related to cosmic time as $ad\tau = dt$. We denote the derivatives as $\partial_t \equiv \dot{f}$ and $\partial_\tau \equiv f'$, where f is any function of time. The Hubble rate $H \equiv \dot{a}/a$ has conformal time analogue $\mathcal{H} \equiv a'/a$.

We are first interested in the gauge quanta which are produced by the homogeneous rolling inflaton $\varphi^{(0)}(\tau)$. To this end, we can ignore the inflaton and metric perturbations in the equations of motion of the gauge field. Extremizing the action with respect to A_0, and choosing the Coulomb gauge $A_0 = 0$, then gives $\left(\vec{\nabla}\cdot\vec{A}\right)' = 0$, from which we set $\vec{\nabla}\cdot\vec{A} = 0$. The equations of motion for $\vec{A}$ then read

$$\vec{A}'' - \nabla^2\vec{A} - \frac{\alpha}{f}\,\varphi^{(0)\,\prime}\,\vec{\nabla}\times\vec{A} = 0 \tag{83}$$

where $\nabla^2 \equiv \partial_i\partial_i$. As we discuss in Subsection IV 1.1, this equation describes the production of the quanta of the gauge fields that results from the motion of the inflaton.

The produced gauge quanta have two key effects: they backreact on the homogeneous background dynamics (see Subsection IV 1.2) and also source inflaton perturbations (see Subsection IV 1.3). Both effects are governed by the equation of motion of the inflaton,

[17] The generalization to non-Abelian gauge groups is straightforward.

and the (00) Einstein equation, which read, respectively

$$\varphi'' + 2\mathcal{H}\varphi' - \nabla^2\varphi + a^2\,\frac{dV}{d\varphi} = a^2\frac{\alpha}{f}\vec{E}\cdot\vec{B}$$
$$\mathcal{H}^2 = \frac{1}{3M_p^2}\left[\frac{1}{2}\varphi'^2 + \frac{1}{2}\left(\vec{\nabla}\varphi\right)^2 + a^2\,V + \frac{a^2}{2}\left(\vec{E}^2 + \vec{B}^2\right)\right] \tag{84}$$

In these equations we have retained the spatial dependence of φ (due to the inflaton perturbations), and we have introduced the physical "electric" and "magnetic" fields[18]

$$\vec{B} = \frac{1}{a^2}\vec{\nabla}\times\vec{A}\,,\quad \vec{E} = -\frac{1}{a^2}\vec{A}'\,. \tag{85}$$

1.1 Production of Gauge Field Fluctuations

During inflation, the motion of the inflaton leads to an instability for the fluctuations of the gauge field. To see this effect, we start from the equation of motion for A_μ in the background of the homogeneous inflaton $\varphi^{(0)}(t)$, eq. (83) above. We decompose the field $\vec{A}(\tau,\vec{x})$ as in (48),

$$\vec{A}(\tau,\vec{x}) = \sum_{\lambda=\pm}\int\frac{d^3k}{(2\pi)^{3/2}}\left[\vec{\epsilon}^{\,(\lambda)}(\hat{k})\,a_\lambda(\vec{k})\,A_\lambda(\tau,\vec{k})\,\mathrm{e}^{i\vec{k}\cdot\vec{x}} + \mathrm{h.c.}\right] \tag{86}$$

where "h.c." denotes the Hermitian conjugate of the preceding term. The annihilation/creation operators and the circular polarization vectors $\vec{\epsilon}^{\,(\pm)}$ obey the properties introduced below (48).

Inserting the decomposition (86) into eq. (83) results in the equation of motion

$$\left[\frac{\partial^2}{\partial\tau^2} + k^2 \pm \frac{2k\xi}{\tau}\right]A_\pm(\tau,k) = 0,\qquad \xi \equiv \frac{\alpha\dot{\varphi}^{(0)}}{2fH} \tag{87}$$

for the mode functions $A_\pm$. During inflation the parameter ξ may be treated as constant, as its time variation is subleading in a slow roll expansion.

From equation (87) we see that one of the polarizations of $\vec{A}_\lambda$ experiences a tachyonic instability for $k/(aH) \lesssim 2\xi$. Without loss of generality, we assume that $\dot{\varphi}^{(0)} > 0$ during inflation, so that the mode exhibiting the instability is A_+. In Appendix B we review the

[18] We do not assume that A_μ necessarily corresponds to the Standard Model electro-magnetic gauge potential.

solutions of (87) and show that the growth of fluctuations is well described by [22]

$$A_+(\tau,k) \cong \frac{1}{\sqrt{2k}}\left(\frac{k}{2\xi aH}\right)^{1/4} e^{\pi\xi - 2\sqrt{2\xi k/(aH)}} \tag{88}$$

in the interval $(8\xi)^{-1} \lesssim k/(aH) \lesssim 2\xi$ [48] of phase space that accounts for most of the power in the produced gauge fluctuations. The phase space of growing modes is non-vanishing for $\xi \gtrsim \mathcal{O}(1)$, which we assume throughout. Notice the exponential enhancement $e^{\pi\xi}$ in the solution (88), which arises due to tachyonic instability, and reflects significant nonperturbative gauge particle production in the regime $\xi \gtrsim 1$. On the other hand, the production of gauge field fluctuations is uninterestingly small for $\xi < 1$. Note also that the other polarization state, $A_-(\tau,k)$, is not produced and can therefore be ignored.

We have thus seen that the motion of the homogeneous inflaton $\varphi^{(0)}(t)$ leads to production of gauge field quanta δA_μ. There are two key physical effects associated with the interactions of these produced quanta with the inflaton. The first effect is the backreaction of the produced quanta on the homogeneous dynamics of $\varphi^{(0)}(t)$, $a(t)$. In the next subsection we study the conditions under which backreaction effects are negligible. The second key physical effect is the production of inflaton fluctuations via *inverse decay*; this is the subject of Subsection IV 1.3.

1.2 Backreaction Effects

Backreaction effects can be accounted for using the mean of the field equations (84):

$$\ddot{\varphi}^{(0)} + 3H\dot{\varphi}^{(0)} + V_\varphi = \frac{\alpha}{f}\langle \vec{E}\cdot\vec{B}\rangle \tag{89}$$

$$3H^2 = \frac{1}{M_p^2}\left[\frac{1}{2}\dot{\varphi}^{(0)\,2} + V + \frac{1}{2}\langle \vec{E}^2 + \vec{B}^2\rangle\right] \tag{90}$$

where we have switched to physical time. The expectation values appearing in (89,90) encode the backreaction of the produced gauge quanta on the homogeneous dynamics of $\varphi^{(0)}(t)$, $a(t)$. From (85) and (86), we have

$$\begin{aligned} \langle \vec{E}\cdot\vec{B}\rangle &= -\frac{1}{4\pi^2 a^4}\int dk\, k^3 \frac{d}{d\tau}|A_+|^2 \\ \frac{1}{2}\langle \vec{E}^2 + \vec{B}^2\rangle &= \frac{1}{4\pi^2 a^4}\int dk\, k^2\left[|A_+'|^2 + k^2|A_+|^2\right] \end{aligned} \tag{91}$$

Since we are studying the backreaction of the produced quanta, we should disregard the modes that do not experience the growth discussed in the previous subsection, namely all modes A_- and the large momentum ($k > 2\xi aH$) modes of A_+. As can be seen from (87), modes of A_+ with $k/(aH) \gg 2\xi$ remain in their vacuum state and do not experience any tachyonic instability. Such modes contribute to the vacuum energy of the U(1) field (in eqn. (90); an analogous consideration applies to $\langle \vec{E} \cdot \vec{B} \rangle$ in eqn. (89)) that we assume is canceled by a bare vacuum energy, as is customary in QFT. (In short, we have nothing to add to the cosmological constant problem.) This prescription provides a UV cutoff $k/\,(aH) < 2\xi$ in the integrals (91); see the paragraph after (87). From a direct inspection of the solutions of (87), we verified that the integrals (91) converge in the infrared $k \to 0$ region, and that they receive their support almost entirely from the region $(8\xi)^{-1} \lesssim k/(aH) \lesssim 2\xi$ in which (88) is a very good approximation of the exact solution. Therefore, (91) can be evaluated by using the expression (88) for A_+, and by integrating only over this region of momenta. In fact, we verified that the momentum interval can be extended from 0 to ∞, since the expression (88) rapidly decreases outside the $(8\xi)^{-1} \lesssim k/(aH) \lesssim 2\xi$ interval, and the contribution of the "outer" regions to (91) can be neglected. Proceeding in this way allows for an analytic result:

$$\langle \vec{E} \cdot \vec{B} \rangle \simeq -2.4 \cdot 10^{-4}\, \frac{H^4}{\xi^4}\, \mathrm{e}^{2\pi\xi} \quad , \quad \langle \frac{\vec{E}^2 + \vec{B}^2}{2} \rangle \simeq 1.4 \cdot 10^{-4}\, \frac{H^4}{\xi^3}\, \mathrm{e}^{2\pi\xi} \tag{92}$$

This procedure follows Ref. [22], where these expressions were first derived. We expect that any sensible renormalization prescription will yield results in agreement with (92).

From (89,90) we can distinguish two distinct kinds of backreaction effects. First, the gauge field fluctuations are produced at the expense of the kinetic energy of $\varphi^{(0)}(t)$. This contributes a new source of dissipation into the homogeneous Klein-Gordon equation (89). In order to trust the usual slow-roll inflationary solution, we require that $|3H\dot{\varphi}^{(0)}| \cong |-V_\varphi| \gg |\frac{\alpha}{f}\langle \vec{E} \cdot \vec{B} \rangle|$. Using (92), this condition reads

$$\frac{H^2}{2\pi|\dot{\varphi}^{(0)}|} \ll 13\, \xi^{3/2} \mathrm{e}^{-\pi\xi} \tag{93}$$

The second kind of backreaction effect arises because the energy density in produced gauge field fluctuations contributes to the Friedmann equation (90). To ensure that the expansion of the universe is dominated by the potential energy of the inflaton we require

$3M_p^2H^2 \cong V \gg \frac{1}{2}\langle \vec{E}^2 + \vec{B}^2\rangle$. Once eq. (92) is taken into account, this condition reads

$$\frac{H}{M_p} \ll 146\,\xi^{3/2}\mathrm{e}^{-\pi\xi} \tag{94}$$

Taken together, the constraints (93,94) ensure that the unstable growth of gauge field fluctuations does not modify the usual homogeneous inflationary dynamics.

1.3 Inverse Decay Effects and Inflaton Perturbations

Even when (93,94) are satisfied, the coupling $\varphi F\tilde{F}$ may still have a profound impact on the cosmological fluctuations in the model (79). The perturbations of the inflaton are described by the equation [19]

$$\left[\frac{\partial^2}{\partial\tau^2} + 2\mathcal{H}\frac{\partial}{\partial\tau} - \nabla^2 + a^2V_{\varphi\varphi}\right]\delta\varphi(\tau,\vec{x}) = a^2\frac{\alpha}{f}\,\left(\vec{E}\cdot\vec{B} - \langle\vec{E}\cdot\vec{B}\rangle\right)\ . \tag{95}$$

The solution of (95) splits into two parts: the solution of the homogeneous equation and the particular solution which is due to the source. Schematically, we have (equivalent to (56))

$$\delta\varphi(\tau,\vec{x}) = \underbrace{\delta\varphi_{\mathrm{vac}}(\tau,\vec{x})}_{\text{homogeneous}} + \underbrace{\delta\varphi_{\mathrm{inv.decay}}(\tau,\vec{x})}_{\text{particular}} \tag{96}$$

The homogeneous solution corresponds, physically, to the usual vacuum fluctuations from inflation. The particular solution, on the other hand, can be interpreted as arising due to inverse decay processes $\delta A + \delta A \to \delta\varphi$. This new source of inflaton fluctuations contributes directly to the observable curvature perturbation on uniform density hypersurfaces, owing to the relation $\zeta = -\frac{H}{\dot{\varphi}^{(0)}}\delta\varphi$. The inverse decay contribution to the cosmological fluctuations may actually *dominate* over the usual vacuum fluctuations in the regime $f \lesssim 10^{-2}M_p$. This radically modifies the phenomenology of axion inflation. In particular, the inverse decay contribution to the primordial cosmological fluctuations is highly non-Gaussian; this is evident already from (95) since the particular solution $\delta\varphi_{\mathrm{inv.decay}}$ is bilinear in the Gaussian field δA_μ.

We proceed with the computation in the way described in Part III. Decomposing $\delta\varphi(\tau,\vec{x})$ as in (49), eq. (95) then results in (we note that the last term of (95) has no effect on the

[19] See Appendix A.1 for the full equation, with the consistent inclusion of the metric perturbations. In the same Appendix, it is shown that the other source terms, arising due to integrating out the non-dynamical modes of the metric perturbations, are completely negligible compared to the source term in (95).

mode functions with momentum different from zero) [20]

$$\left[\partial_\tau^2 + k^2 - \frac{a''}{a} + a^2 m^2\right] Q_\varphi(\tau, \vec{k}) = J_\varphi(\tau, \vec{k}) \tag{97}$$

$$J_\varphi(\tau, \vec{k}) \equiv a^3(\tau)\frac{\alpha}{f}\int \frac{d^3k}{(2\pi)^{3/2}} e^{-i\vec{k}\cdot\vec{x}}\vec{E}\cdot\vec{B} \tag{98}$$

where the a^2m^2 term in (97) is slow-roll suppressed and thus can be neglected in our computation. We separate the mode functions of the two terms in (96) as

$$Q_\varphi(\tau, \vec{k}) = Q_\varphi^{\rm vac}(\tau, \vec{k}) + Q_\varphi^{\rm inv.decay}(\tau, \vec{k}) \tag{99}$$

and we discuss each contribution separately. Following the procedure in Section III 2, we expand the homogeneous term as [21]

$$Q_\varphi^{\rm vac}(\tau, \vec{k}) = b_\varphi(\vec{k})\,\varphi_k(\tau) + b_\varphi^\dagger(-\vec{k})\,\varphi_k^*(\tau) \tag{100}$$

with

$$\varphi_k(\tau) = i\frac{\sqrt{\pi}}{2}\sqrt{-\tau}H_\nu^{(1)}(-k\tau), \qquad \nu \cong \frac{3}{2} + \mathcal{O}(\epsilon, \eta) \tag{101}$$

where $\varphi_k(\tau)$ is chosen to be real in the limit $-k\tau \to 0$.

The particular solution, on the other hand, takes the form

$$Q_\varphi^{\rm inv.decay}(\tau, \vec{k}) = \int_{-\infty}^{\tau} d\tau' G_k(\tau, \tau') J_\varphi(\tau'\,\vec{k}) \tag{102}$$

$$G_k(\tau, \tau') = i\Theta(\tau - \tau')\left[\varphi_k(\tau)\,\varphi_k^*(\tau') - \varphi_k^*(\tau)\,\varphi_k(\tau')\right] \tag{103}$$

where the source term was defined in eq. (98). We remind that this particular solution is statistically independent of the homogeneous solution (100).

We are now in a position to compute the correlation functions for the perturbations $\delta\varphi$. We present this computation in Section IV 2. In Section IV 3 we discuss the resulting phenomenology. Finally, in Appendix A.1 we show that these result are valid also when metric perturbations are also consistently taken into account.

[20] In Appendix A.1 we present the complete computation, including also scalar metric perturbations. We show explicitly that equation (97) still holds, with only the addition of subdominant (Planck-suppressed) terms in the source J_φ.

[21] Corresponding to the notation in III, we note $\varphi_k(\tau) \equiv \phi_\varphi(\tau, k)$.

2 Correlation Functions

In this section we compute the main phenomenological signatures of the model (79). Specifically: in Subsection IV 2.1 we compute the two point correlation function of the density perturbation $\zeta = -\frac{H}{\dot{\varphi}^{(0)}}\,\delta\varphi$; in Subsection IV 2.2 we present the corresponding power spectrum; in Subsection IV 2.3 we compute the three point correlation function of ζ; in Subsection IV 2.4 we present the corresponding bispectrum; in Subsection IV 2.5 we finally summarize the computation of the power spectrum of the gravity waves (GW) modes. In this section we neglect scalar metric perturbations for simplicity, however, in Appendix A.1 we show explicitly that their consistent inclusion does not modify our results.

2.1 Two-Point Correlation Function

We start from the relation $\zeta\,(\tau,\,\vec{x}) = -\frac{H}{\dot{\varphi}^{(0)}}\,\delta\varphi\,(\tau,\,\vec{x})$ between the inflaton perturbations, and the curvature perturbation on uniform density hypersurfaces. We decompose the latter as

$$\zeta(\tau, \vec{x}) = \int \frac{d^3k}{(2\pi)^{3/2}}\,\zeta_{\vec{k}}(\tau)\,\mathrm{e}^{i\vec{k}\cdot\vec{x}} \tag{104}$$

so that $\zeta_{\vec{k}} = -\frac{H}{\dot{\varphi}^{(0)}}\frac{Q_\varphi}{a}$. As we discussed in Subsection IV 1.3, the inflaton perturbations comprise of two terms, one being the vacuum fluctuations, and one the fluctuations sourced by the gauge quanta; since these two terms are statistically independent, we have

$$\langle\zeta_{\vec{k}}\,\zeta_{\vec{k}'}\rangle = \frac{H^2}{a^2\dot{\varphi}^{(0)\,2}}\left[\langle Q_\varphi^{\rm vac}(\vec{k})\,Q_\varphi^{\rm vac}(\vec{k}')\rangle + \langle Q_\varphi^{\rm inv.decay}(\vec{k})\,Q_\varphi^{\rm inv.decay}(\vec{k}')\rangle\right] \tag{105}$$

The contribution from the vacuum modes is standard. Using eqs. (100) and (101), one obtains the well-known result

$$\begin{aligned}\langle\zeta_{\vec{k}}^{\rm vac}\zeta_{\vec{k}'}^{\rm vac}\rangle &= \frac{H^4}{2\dot{\varphi}^{(0)\,2}}\left(\frac{k}{aH}\right)^{n_s-1}\frac{1}{k^3}\,\delta^{(3)}\left(\vec{k}+\vec{k}'\right)\\ &= \frac{2\pi^2}{k^3}\,\mathcal{P}\left(\frac{k}{aH}\right)^{n_s-1}\delta^{(3)}\left(\vec{k}+\vec{k}'\right)\end{aligned} \tag{106}$$

in the late time / large scales limit, $-k\tau \ll 1$. In the second line of (106) we have introduced

$$n_s = 1 + 3 - 2\nu = 1 + \mathcal{O}(\epsilon,\eta) \tag{107}$$

$$\mathcal{P}^{1/2} \equiv \frac{H^2}{2\pi|\dot{\varphi}^{(0)}|} \tag{108}$$

We now compute the 2-point correlator of the particular solution (102):

$$\langle\zeta^{\rm inv.decay}_{\vec{k}}\,(\tau)\;\zeta^{\rm inv.decay}_{\vec{k}'}\,(\tau)\rangle = \frac{H^2}{\dot{\varphi}^{(0)\,2}}\int d\tau'\,d\tau''\frac{G_k\,(\tau,\,\tau')}{a(\tau)}\;\frac{G_{k'}\,(\tau,\,\tau'')}{a(\tau)}\;\langle J_\varphi(\tau',\vec{k})\,J_\varphi(\tau'',\vec{k}')\rangle \tag{109}$$

(In the following, we temporarily omit the superscript 'inv.decay' for notational convenience.) In evaluating this expression, we can make some approximations for the mode functions $\varphi_k\,(\tau)$ and $\varphi_k\,(\tau')$ appearing in the Green function (103). Since we are interested in the power spectrum of modes well outside the horizon we can use the small argument limit $-k\tau \ll 1$ for the modes (101) entering in the Green function (103),

$$\varphi_k\,(\tau) \simeq \frac{a\,(\tau)\,H}{\sqrt{2}}\,\frac{1}{k^{3/2}}\,\left(-k\tau\right)^{\frac{n_s-1}{2}}\,, \qquad \text{for} - k\tau \ll 1 \tag{110}$$

Notice that we disregard the small slow roll corrections to the amplitude, but we retain them in the momentum dependence (since this controls the departure of the spectrum from scale invariance). This gives

$$\langle\zeta_{\vec{k}}\,(\tau)\,\zeta_{\vec{k}'}\,(\tau)\rangle \cong \frac{2H^4}{\dot{\varphi}^{(0)\,2}}\,\frac{(-k\tau)^{n_s-1}}{k^3}\int_{-\infty}^{\tau} d\tau'\,d\tau''\,{\rm Im}\left[\varphi_k\left(\tau'\right)\right]\,{\rm Im}\left[\varphi_k\left(\tau''\right)\right]\,\langle J_\varphi(\tau',\vec{k})\,J_\varphi(\tau'',\vec{k}')\rangle \tag{111}$$

In this relation, we have already used the fact that the correlator is nonvanishing only for $|\vec{k}| = |\vec{k}'|$. We stress that, in (111) we do *not* employ the approximation (110) for the modes $\varphi_k(\tau')$ which appear under the integral, since τ' must be integrated over.

To evaluate (111), we require the source correlator $\langle J_\varphi(\tau',\vec{k})\,J_\varphi(\tau'',\vec{k}')\rangle$. Explicit evaluation gives

$$\begin{aligned}\langle J_\varphi(\tau',\vec{k})\,J_\varphi(\tau'',\vec{k}')\rangle = \frac{\alpha^2\delta^{(3)}(\vec{k}+\vec{k}')}{8f^2a\,(\tau')\,a\,(\tau'')}\int\frac{d^3q}{(2\pi)^3}\,\left[1+\frac{|\vec{q}|^2-\vec{q}\cdot\vec{k}}{|\vec{q}|\,|\vec{k}-\vec{q}|}\right]^2 \\ \times\mathcal{A}\left[\tau',\,|\vec{q}|,\,|\vec{q}-\vec{k}|\right]\,\mathcal{A}^*\left[\tau'',\,|\vec{q}|,\,|\vec{q}-\vec{k}|\right]\end{aligned} \tag{112}$$

where

$$\mathcal{A}\left[\tau',\,|\vec{q}|,\,|\vec{q}-\vec{k}|\right] \equiv |\vec{q}|A_+'\left(\tau',|\vec{q}-\vec{k}|\right)\,A_+\left(\tau',|\vec{q}|\right) + |\vec{q}-\vec{k}|A_+'\left(\tau',|\vec{q}|\right)\,A_+\left(\tau',\,|\vec{q}-\vec{k}|\right) \tag{113}$$

and where the relations

$$\left|\vec{\epsilon}^{\,(+)}\left(\vec{q}\right)\cdot\vec{\epsilon}^{\,(+)}\left(\vec{k}-\vec{q}\right)\right|^2=\frac{1}{4}\left[1+\frac{|\vec{q}|^2-\vec{q}\cdot\vec{k}}{|\vec{q}|\,|\vec{k}-\vec{q}|}\right]^2\;,\quad \vec{\epsilon}^{\,(\lambda)}(-\vec{k})=\vec{\epsilon}^{\,(\lambda)*}(\vec{k}) \tag{114}$$

have been used.

Putting all together, we have

$$\begin{aligned}\langle\zeta_{\vec{k}}\left(\tau\right)\zeta_{\vec{k}'}\left(\tau\right)\rangle\;&\cong\;\frac{\alpha^2\,H^6}{4\,f^2\dot{\varphi}^{(0)\,2}}\,\frac{\left(-k\tau\right)^{n_s-1}}{k^3}\,\delta^{(3)}\left(\vec{k}+\vec{k}'\right)\int\frac{d^3q}{(2\pi)^3}\left[1+\frac{|\vec{q}|^2-\vec{q}\cdot\vec{k}}{|\vec{q}|\,|\vec{k}-\vec{q}|}\right]^2\\&\qquad\times\left|\int_{-\infty}^{\tau}d\tau'\left(-\tau'\right)\,\mathrm{Im}\left[\varphi_k\left(\tau'\right)\right]\mathcal{A}\left[\tau'\,|\vec{q}|,\,|\vec{q}-\vec{k}|\right]\right|^2\end{aligned} \tag{115}$$

Following similar arguments to those made in Subsection IV 1.2, we can use the approximation (88) for the gauge field mode functions (we discuss this step more in details at the end of this subsection). This leads to

$$\begin{aligned}\langle\zeta_{\vec{k}}\left(\tau\right)\zeta_{\vec{k}'}\left(\tau\right)\rangle\;&\cong\;\frac{\alpha^2H^6\,\mathrm{e}^{4\pi\xi}}{2^8\,\pi^3\,f^2\dot{\varphi}^{(0)\,2}}\,\frac{\left(-k\tau\right)^{n_s-1}}{k^3}\,\delta^{(3)}\left(\vec{k}+\vec{k}'\right)\\&\times\int d^3q_*\left[1+\frac{|\vec{q}_*|^2-\vec{q}_*\cdot\hat{k}}{|\vec{q}_*|\,|\hat{k}-\vec{q}_*|}\right]^2|\vec{q}_*|^{1/2}\,|\vec{q}_*-\hat{k}|^{1/2}\left[|\vec{q}_*|^{1/2}+|\vec{q}_*-\hat{k}|^{1/2}\right]^2\\&\times\mathcal{I}^2\left[2\sqrt{2\xi}\left(\sqrt{|\vec{q}_*|}+\sqrt{|\vec{q}_*-\hat{k}|}\right)\right]\end{aligned} \tag{116}$$

where the integration variable $\vec{q}_*\equiv\vec{q}/|\vec{k}|$ is dimensionless, and where

$$\mathcal{I}\left[z\right]\equiv\sqrt{\frac{\pi}{2}}\int_{-k\tau}^{\infty}dx\,x^{3/2}\,\mathrm{Re}\left[H_\nu^{(1)}\left(x\right)\right]\,\mathrm{e}^{-z\sqrt{x}} \tag{117}$$

(notice that $x\equiv-k\tau$). As we are interested only in super horizon modes, $-k\tau\ll1$, we can set to zero the lower extreme of integration of $\mathcal{I}$. It is then manifest that we can set $\nu=3/2$ in the argument of the Hankel function, since the slow roll corrections appearing there only modify (in a negligible amount) the amplitude of the correlator, but not its scale dependence. This leads to

$$\mathcal{I}\left(z\right)\simeq\int_0^\infty dx\left(\sin x-x\cos x\right)\mathrm{e}^{-z\sqrt{x}} \tag{118}$$

For future convenience, we rewrite the correlator (116) as

$$\langle \zeta_{\vec{k}}^{\rm inv.decay}(\tau)\, \zeta_{\vec{k}'}^{\rm inv.decay}(\tau)\rangle \equiv \frac{2\pi^2}{k^3}(-k\tau)^{n_s-1}\mathcal{P}^2 f_2(\xi) e^{4\pi\xi}\delta^{(3)}(\vec{k}+\vec{k}') \tag{119}$$

where $\mathcal{P}$ was defined in (108) and the dimensionless function $f_2(\xi)$ is

$$f_2(\xi) \equiv \frac{\xi^2}{8\,\pi}\int d^3q_* \left[1+\frac{|\vec{q}_*|^2-\vec{q}_*\cdot\hat{k}}{|\vec{q}_*|\,|\hat{k}-\vec{q}_*|}\right]^2 |\vec{q}_*|^{1/2}\,|\vec{q}_*-\hat{k}|^{1/2}\left[|\vec{q}_*|^{1/2}+|\vec{q}_*-\hat{k}|^{1/2}\right]^2$$
$$\times\,\mathcal{I}^2\left[2\sqrt{2\xi}\left(\sqrt{|\vec{q}_*|}+\sqrt{|\vec{q}_*-\hat{k}|}\right)\right] \tag{120}$$

In general, the function $f_2(\xi)$ needs to be evaluated numerically. However, a simplification is achieved when the argument of $\mathcal{I}$ is much greater than one.

$$\mathcal{I}\left[z\right] \cong \int_0^\infty dx\frac{x^3}{3}{\rm e}^{-z\sqrt{x}} = \frac{3360}{z^8} \quad \text{for } z \gg 1 \tag{121}$$

As $\sqrt{|\vec{q}_*|}+\sqrt{|\vec{q}_*-\hat{k}|} \geq 1$, this approximation is certainly appropriate at large ξ. One is then left with a two dimensional integral that numerically evaluates to

$$f_2(\xi) \cong \frac{7.5\cdot 10^{-5}}{\xi^6} \quad , \quad \xi \gg 1 \tag{122}$$

The degree of accuracy of this approximation can be seen in Figure 4. It is also useful to have a fit for f_2 in the range $2 \leq \xi \leq 3$, as this is the most relevant one for phenomenology (as we discussed in Section IV 3). The best monomial fit to f_2 in this range is

$$f_2(\xi) \cong \frac{3\cdot 10^{-5}}{\xi^{5.4}} \quad , \quad 2 \leq \xi \leq 3 \tag{123}$$

We conclude this subsection with the justification of the use of (88) into the integral (115). We know that the expressions (88) are accurate approximations to the modes $A_+\,(\tau,\,k)$ in the range $1/8\xi \leq -\tau\, k \leq 2\xi$. As discussed in Subsection IV 1.2, for any given mode k, this interval corresponds to the times for which the amplification of the gauge field is maximal. The key point is to ensure that there is a common region in the integration space of (115) for which both the functions A_+ that enter in the expression of $\mathcal{A}$ can be

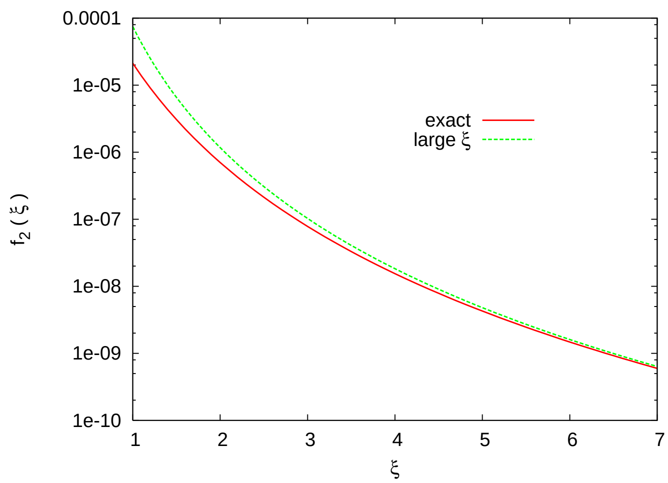

Figure 4: The function $f_2(\xi)$.

approximated by (88). This requires that

$$\frac{1}{8\xi} \le -|\vec{q}|\,\tau' \le 2\xi \quad \text{and} \quad \frac{1}{8\xi} \le -|\vec{q}-\vec{k}|\,\tau' \le 2\xi \tag{124}$$

or, equivalently,

$$\frac{1}{8\xi} \le |\vec{q}_*|\,x \le 2\xi \quad \text{and} \quad \frac{1}{8\xi} \le |\vec{q}_* - \hat{k}|x \le 2\xi \tag{125}$$

($x = -k\tau'$ is the integration variable in (117), and $\vec{q}_* = \vec{q}/|\vec{k}|$). For the super-horizon modes that are relevant for phenomenology, we know that x extends from 0 to ∞. Therefore, for any value of $\vec{q}_*$, there are always values of the rescaled time x for which either $A_+(\tau',\, |\vec{q}_*|)$ or $A_+(\tau',\, |\vec{q}_* - \hat{k}|)$ is maximal, and approximated by (88). Only when $|\vec{q}_*| \simeq |\vec{q}_* - \hat{k}|$, both conditions are valid for the same values of x, and the approximations (88) can be used in the whole integrand of (115). Therefore, our result is correct only if the integrand of (115) is strongly peaked at $|\vec{q}_*| = \mathcal{O}\,(1)$. We have verified with direct inspection that this is indeed the case (We have verified this claim also using the representation (B.2) of the gauge field modes which is valid arbitrarily deep in the IR.)

There is a clear physical reason why the integrand is strongly peaked in the region of phase space where the conditions (124) are satisfied. Notice that, for the values $z = \mathcal{O}\,(1)$ which are relevant for the present computation, the expression (118) has most of its support at $x = \mathcal{O}(1)$. This means that the "imprint" of the fluctuations ζ_k from the gauge modes occurs when the wavelength of the fluctuation is of the order of the horizon scale (x of order one), and it is caused by the inverse decay of gauge field modes, whose wavelength is also of the order of the horizon scale ($|\vec{q}_*|$ of order one). This is a natural outcome of causality/locality. Identical considerations apply in the computation of the three point function that we perform in Subsection IV 2.3.

2.2 The Primordial Power Spectrum

The two point correlation function in momentum space is related to the power spectrum by the standard expression

$$\langle \zeta_{\vec{k}} \, \zeta_{\vec{k}'} \rangle \equiv P_\zeta(k) \, \frac{2\pi^2}{k^3} \delta^{(3)}(\vec{k} + \vec{k}') \tag{126}$$

As we have seen in the previous subsection, $\zeta_{\vec{k}} = \zeta_{\vec{k}}^{\rm vac} + \zeta_{\vec{k}}^{\rm inv.decay}$, and the two terms are uncorrelated. From (106) and (119) we get the power spectrum at late times

$$P_\zeta(k) = \mathcal{P} \left(\frac{k}{k_0} \right)^{n_s - 1} \left[1 + \mathcal{P} \, f_2 \left(\xi \right) \, \mathrm{e}^{4\pi\xi} \right] \tag{127}$$

where $\mathcal{P}^{1/2} = \frac{H^2}{2\pi |\dot{\varphi}^{(0)}|}$.

2.3 Three-Point Correlation Function

In this subsection we calculate the three point function of $\zeta_{\vec{k}}^{\rm inv.decay}$,

$$\begin{aligned} \langle \zeta_{\vec{k}_1} \left(\tau \right) \, \zeta_{\vec{k}_2} \left(\tau \right) \, \zeta_{\vec{k}_3} \left(\tau \right) \rangle \quad = \quad & - \frac{H^3}{\dot{\varphi}^{(0)\,3}} \int d\tau_1 \, d\tau_2 \, d\tau_3 \, \frac{G_{k_1} \left(\tau , \, \tau_1 \right)}{a(\tau)} \, \frac{G_{k_2} \left(\tau , \, \tau_2 \right)}{a(\tau)} \, \frac{G_{k_3} \left(\tau , \, \tau_3 \right)}{a(\tau)} \\ & \times \langle J_\varphi(\tau_1, \vec{k}_1) \, J_\varphi(\tau_2, \vec{k}_2) \, J_\varphi(\tau_3, \vec{k}_3) \rangle \end{aligned} \tag{128}$$

Proceeding as in Subsection IV 2.1, we arrive to

$$\begin{aligned}
\langle \zeta_{\vec{k}_1}\, \zeta_{\vec{k}_2}\, \zeta_{\vec{k}_3} \rangle \;\; &= \;\; -\frac{\alpha^3 H^9}{f^3\, \dot{\varphi}^{(0)\,3}\, k_1^3\, k_2^3\, k_3^3}\, \delta^{(3)}\left(\vec{k}_1 + \vec{k}_2 + \vec{k}_3\right) \int \frac{d^3 q}{(2\pi)^{9/2}} \\
&\times \left[\vec{\epsilon}\,(\vec{q}) \cdot \vec{\epsilon}\left(\vec{k}_1 - \vec{q}\right)\right] \left[\vec{\epsilon}\left(\vec{q} - \vec{k}_1\right) \cdot \vec{\epsilon}\left(-\vec{q} - \vec{k}_3\right)\right] \left[\vec{\epsilon}\left(\vec{q} + \vec{k}_3\right) \cdot \vec{\epsilon}\,(-\vec{q})\right] \\
&\times \prod_i \int_{-\infty}^{0} d\tau_i \left[k_i \tau_i\, \cos\left(k_i \tau_i\right) - \sin\left(k_i \tau_i\right)\right] \\
&\quad \times \mathcal{A}\left[\tau_1,\, |\vec{q}|,\, |\vec{k}_1 - \vec{q}|\right] \mathcal{A}\left[\tau_2,\, |\vec{k}_1 - \vec{q}|,\, |\vec{k}_3 + \vec{q}|\right] \mathcal{A}\left[\tau_3,\, |\vec{k}_3 + \vec{q}|,\, |\vec{q}|\right]
\end{aligned} \tag{129}$$

where $\mathcal{A}$ was defined in (113). In this expression we have used the fact that the mode functions A_+ are real (which is true both in the approximations (B.2) and (88)), and we have disregarded any (mild) scale dependence.[22]

The correlator depends on the size and the shape of the triangle formed by the vectors $\vec{k}_i$. We denote

$$|\vec{k}_1| = k \;\;,\;\; |\vec{k}_2| = x_2\, k \;\;,\;\; |\vec{k}_3| = x_3\, k \tag{130}$$

and, for future convenience, we parametrize the correlator as

$$\begin{aligned}
\langle \zeta_{\vec{k}_1}\, \zeta_{\vec{k}_2}\, \zeta_{\vec{k}_3} \rangle \;\; &\equiv \;\; \frac{3}{80\,(2\pi)^{7/2}}\, \frac{\alpha^3 H^9}{f^3\, \dot{\varphi}^{(0)\,3}}\, \frac{\mathrm{e}^{6\pi\xi}}{\xi^3}\, \frac{\delta^{(3)}\left(\vec{k}_1 + \vec{k}_2 + \vec{k}_3\right)}{k^6}\, \frac{1 + x_2^3 + x_3^3}{x_2^3\, x_3^3}\, f_3\left(\xi,\, x_2,\, x_3\right) \\
&= \;\; \frac{3}{10}\,(2\pi)^{5/2}\, \mathcal{P}^3\, \mathrm{e}^{6\pi\xi}\, \frac{\delta^{(3)}\left(\vec{k}_1 + \vec{k}_2 + \vec{k}_3\right)}{k^6}\, \frac{1 + x_2^3 + x_3^3}{x_2^3\, x_3^3}\, f_3\left(\xi,\, x_2,\, x_3\right)
\end{aligned} \tag{131}$$

We proceed from eq. (129) as we did in Subsection IV 2.1. Using the last expression,

[22] Specifically, we have set $\nu = 3/2$ in the mode functions entering in the Green function (103)).

eq. (113), and eqs. (88), we arrive to

$$
\begin{aligned}
f_3\left(\xi;\, x_2,\, x_3\right) \;=\; & \frac{5}{3\pi}\,\frac{\xi^3}{x_2\,x_3\left[1+x_2^3+x_3^3\right]}\int d^3q_*\,|\vec{q}_*|^{1/2}\,|\hat{k}_1-\vec{q}_*|^{1/2}\,|\vec{q}_*+x_3\,\hat{k}_3|^{1/2} \\
& \times\left[\vec{\epsilon}\left(\vec{q}_*\right)\cdot\vec{\epsilon}\left(\hat{k}_1-\vec{q}_*\right)\right]\left[|\vec{q}_*|^{1/2}+|\hat{k}_1-\vec{q}_*|^{1/2}\right]\mathcal{I}\left[2\sqrt{2\xi}\left(|\vec{q}_*|^{1/2}+|\hat{k}_1-\vec{q}_*|^{1/2}\right)\right] \\
& \times\left[\vec{\epsilon}\left(\vec{q}_*-\hat{k}_1\right)\cdot\vec{\epsilon}\left(-\vec{q}_*-x_3\,\hat{k}_3\right)\right]\left[|\hat{k}_1-\vec{q}_*|^{1/2}+|\vec{q}_*+x_3\,\hat{k}_3|^{1/2}\right] \\
& \qquad\qquad\times\mathcal{I}\left[2\sqrt{\frac{2\xi}{x_2}}\left(|\hat{k}_1-\vec{q}_*|^{1/2}+|\vec{q}_*+x_3\,\hat{k}_3|^{1/2}\right)\right] \\
& \times\left[\vec{\epsilon}\left(\vec{q}_*+x_3\,\hat{k}_3\right)\cdot\vec{\epsilon}\left(-\vec{q}_*\right)\right]\left[|\vec{q}_*+x_3\,\hat{k}_3|^{1/2}+|\vec{q}_*|^{1/2}\right] \\
& \qquad\qquad\times\mathcal{I}\left[2\sqrt{\frac{2\xi}{x_3}}\left(|\vec{q}_*+x_3\,\hat{k}_3|^{1/2}+|\vec{q}_*|^{1/2}\right)\right]
\end{aligned}
\tag{132}
$$

To continue, we can set

$$
\begin{aligned}
\hat{k}_1 &= (1,\,0,\,0) \\
x_3\hat{k}_3 &= -\frac{1}{2}\left(1-x_2^2+x_3^2,\;\sqrt{-\left(1-x_2+x_3\right)\left(1+x_2-x_3\right)\left(1-x_2-x_3\right)\left(1+x_2+x_3\right)},\;0\right)
\end{aligned}
\tag{133}
$$

and we note that, to a generic vector $\vec{k}=k\left(\sin\theta\,\cos\phi,\,\sin\theta\sin\phi,\,\cos\theta\right)$, corresponds the polarization operator

$$
\vec{\epsilon}^{\,(+)}(\hat{k})=\frac{1}{\sqrt{2}}\left(\cos\theta\,\cos\phi-i\sin\phi,\,\cos\theta\sin\phi+i\,\cos\phi,\,-\sin\theta\right)
\tag{134}
$$

(it is immediate to verify that this expression satisfies all the properties listed after eq. (48)).

In the reminder of this subsection we evaluate f_3 for the equilateral configuration $x_2 = x_3 = 1$. In the large ξ limit, we can use the analytic approximation (121) for $\mathcal{I}$, and perform the momentum integral numerically. We obtain

$$
f_3\left(\xi;\,1,\,1\right)\cong\frac{2.8\cdot10^{-7}}{\xi^9}\quad,\quad\xi\gg1
\tag{135}
$$

The degree of accuracy of this approximation can be seen in Figure 5. It is also useful to have a fit for f_3 in the range $2\leq\xi\leq3$, as this is the most relevant one for phenomenology

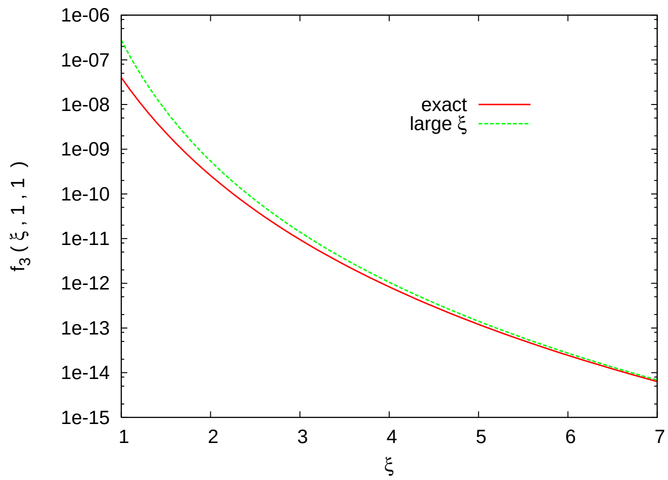

Figure 5: The function $f_3(\xi, 1, 1)$.

(as we discuss in Section IV 3). The best monomial fit to f_3 in this range is

$$f_3(\xi; 1, 1) \cong \frac{7.4 \cdot 10^{-8}}{\xi^{8.1}} \quad , \quad 2 \leq \xi \leq 3 \tag{136}$$

2.4 The Bispectrum and Non-linearity Parameter

A popular parametrization of non-Gaussianity is the non-linearity parameter $f_{\rm NL}$, introduced by assuming that the curvature perturbation may be expanded as in (39),

$$\zeta\left(\vec{x}\right) = \zeta_g\left(\vec{x}\right) + \frac{3}{5}\, f_{\rm NL}\, \left[\zeta_g^2\left(\vec{x}\right) - \langle \zeta_g^2\left(\vec{x}\right)\rangle\right] \tag{137}$$

where $\zeta_g(x)$ is a Gaussian random field (see [215] for a careful discussion of sign conventions). Both ζ and ζ_g may be decomposed as in (104) so that the relation between the q-modes of the Fourier decomposition is

$$\zeta_{\vec{k}} = \zeta_{g,\vec{k}} + \frac{3}{5}\, f_{\rm NL} \int \frac{d^3p}{\left(2\pi\right)^{3/2}}\, \zeta_{g,\vec{p}}\, \zeta_{g,\vec{k}-\vec{p}} \tag{138}$$

By definition, the three point correlator of ζ_g vanishes. However, due to the quadratic term in (235), the three point correlator of ζ is nonvanishing, and can be expressed through a sum of two point correlators of ζ_g. One finds[23]

$$\langle \zeta_{\vec{k}_1} \, \zeta_{\vec{k}_2} \, \zeta_{\vec{k}_3} \rangle = \frac{3}{10} \, (2\pi)^{5/2} \, f_{\rm NL} P_\zeta \, (k)^2 \, \delta^{(3)} \left(\vec{k}_1 + \vec{k}_2 + \vec{k}_3 \right) \frac{\sum_i k_i^3}{\Pi_i k_i^3} \tag{139}$$

where the power spectrum was defined in (126). To obtain this expression, recall that the ladder operators are normalized according to (58), and that the power spectrum $P\,(k)$ is related to the two point function as in (126). One should also identify the two point function of ζ with that of ζ_g (as the difference in subleading in a perturbative expansion), and disregard the the mild scale dependence of the power spectrum.

By comparison to (237), one may define an "effective" (momentum dependent) non-linearity parameter, even when the intrinsic non-Gaussianity is not of the local form (235). For axion inflation, using the parametrization (131) of the three point correlator, we can write

$$f_{\rm NL}^{\rm eff}(\xi; x_2, x_3) = \frac{f_3 \left(\xi; \, x_2, \, x_3\right) \, \mathcal{P}^3 \, {\rm e}^{6\pi\xi}}{P_\zeta \, (k)^2} \tag{140}$$

where we recall that $\mathcal{P}^{1/2} = \frac{H^2}{2\pi |\dot{\varphi}^{(0)}|}$.

2.5 Power Spectrum of the Tensor Modes

The produced gauge quanta also source tensor metric perturbations (gravity waves). The total GW power spectrum was first given in [48]. Ref. [284] then pointed out that the chiral nature of the GW produced by the gauge modes can be probed through the resulting nonvanishing $\langle BE \rangle$ and $\langle BT \rangle$ correlators of the CMB, as studied in [147, 262]. For the present model, a positive detection of parity violation would only be possible in a cosmic variance limited experiment, and for a limited portion of the parameter space [284]. In particular, one needs to be in a regime in which the GW production from the gauge modes dominates over that from the vacuum. In the minimal version of the model studied here, this region is ruled out by the non-Gaussianity limit that we discuss in the next section [48]. Ref. [284] circumvented this problem by considering the presence of many ($\mathcal{N} \gtrsim 10^3$) gauge fields, or by the use of a curvaton.

As the computation of the GW power spectrum has been presented in details in [284], we only provide a quick summary here. The tensor modes enter in the spatial components of the metric as $g_{ij} = a^2 \left(\delta_{ij} + h_{ij} \right)$, with $h_{ii} = \partial_i h_{ij} = 0$. From the Einstein equations, one

[23] Note that the factors of $(2\pi)^{n/2}$ differ from [90]. This stems from a different convention for the normalization of the Fourier transform (104).

then finds

$$\frac{1}{2a^2}\left(\partial_\tau^2+\frac{2a'}{a}\,\partial_\tau-\partial_{\mathbf{x}}^2\right)h_{ij}=\frac{1}{M_p^2}\left(-E_iE_j-B_iB_j\right)^{TT} \tag{141}$$

where TT denotes the transverse and traceless projection of the spatial components of the energy-momentum tensor of the gauge field. The computation of the GW production is performed analogously to that of the density perturbations that we have presented in details in the previous subsections. As for density perturbations, the GW modes produced by the gauge quanta are uncorrelated with those from the vacuum, and the two contributions add up incoherently in the power spectrum. The two GW helicities are obtained from the projectors $\Pi_{ij,R/L}(\hat{k})=\epsilon_i^{(\pm)}(\hat{k})\,\epsilon_j^{(\pm)}(\hat{k})$. One finds the two power spectra

$$P_{L/R}\cong\frac{H^2}{\pi^2\,M_p^2}\left(\frac{k}{k_0}\right)^{n_T}\left[1+\frac{2H^2}{M_p^2}\,f_{h,L/R}\left(\xi\right)\,\mathrm{e}^{4\pi\xi}\right] \tag{142}$$

where

$$\begin{aligned} f_{h,L/R} &= \frac{1}{\xi}\int\frac{d^3q_*}{(2\pi)^3}\sqrt{q_*\,|\hat{k}-\vec{q}_*|}\,\frac{(1\pm\cos\theta)^2\left(1-q_*\cos\theta\pm\sqrt{1-2q_*\cos\theta+q_*^2}\right)^2}{16\left(1-2q_*\cos\theta+q_*^2\right)} \\ &\times\left\{\int_0^\infty dx\sqrt{x}\left[\sin x-x\,\cos x\right]\left[\frac{2\xi}{x}+\sqrt{q_*\,|\hat{k}-\vec{q}_*|}\right]\mathrm{e}^{-2\sqrt{2\xi x}\left[\sqrt{\vec{q}_*}+\sqrt{|\hat{k}-\vec{q}_*|}\right]}\right\}^2 \end{aligned} \tag{143}$$

and where q_* and θ are, respectively, the magnitude of the (dimensionless) integration momentum $\vec{q}_*$, and the angle between this vector and the momentum $\vec{k}$ of the mode.

In eq. (142), $n_T=-2\epsilon$. We note that the contribution to the spectrum of the modes from the vacuum, and of the modes sourced from the gauge field have the same scale dependence. The reason for this is that the scale dependence of the second term originates from the homogeneous solutions of (141) – which are the vacuum solutions - employed in the Green function. For the same reason, also the two terms in the scalar power spectrum have identical scale dependence, see Subsection IV 2.1.

At large ξ, the integral over x can be performed as in the previous subsections, and one finds [284]

$$f_{h,L}\cong\frac{4.3\cdot10^{-7}}{\xi^6}\quad,\quad f_{h,R}\cong\frac{9.2\cdot10^{-10}}{\xi^6}\quad,\quad\xi\gg1 \tag{144}$$

The degree of accuracy of this approximation can be seen in Figure 6. It is also useful to have a fit for $f_{h,L}$ in the range $2\leq\xi\leq3$, as this is the most relevant one for phenomenology

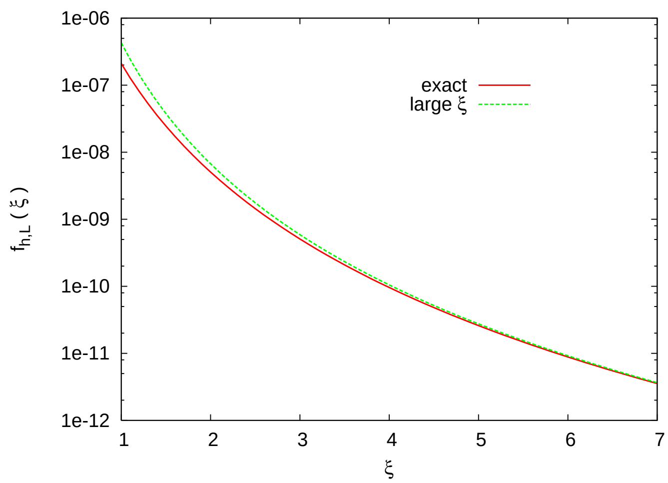

Figure 6: The function $f_{h,L}(\xi)$.

(as we discussed in Section IV 3). The best monomial fit to $f_{h,L}$ in this range is

$$f_{h,L}(\xi) \cong \frac{2.6 \cdot 10^{-7}}{\xi^{5.7}} \quad , \quad 2 \leq \xi \leq 3 \tag{145}$$

From (144) we see that, for the purpose of computing the tensor-to-scalar ratio, one can disregard the right helicity GW modes produced by the gauge fields. One then obtains the result

$$P_{\rm GW} = P_{h,L} + P_{h,R} \cong \frac{2H^2}{\pi^2 \, M_p^2} \left(\frac{k}{k_0} \right)^{n_T} \left[1 + \frac{H^2}{M_p^2} \, f_{h,L} \left(\xi \right) \, {\rm e}^{4\pi\xi} \right] \tag{146}$$

first reported in [48].

3 Phenomenology of Axion Inflation

From the observational perspective, the key quantities which characterize any model of inflation are the spectrum of scalar and tensor perturbations, P_ζ and P_{GW}, along with the bispectrum of scalar perturbations B_ζ that encodes the leading departures from Gaussian statistics. The explicit computations of these quantities was presented in Section IV 2. In this Section, we study the resulting observational signatures.

3.1 COBE Normalization and Spectral Tilt

In Subsections IV 2.1 and IV 2.2 we found that two uncorrelated terms contribute to the power spectrum in axion inflation. These are the usual fluctuations generated from the vacuum, along with the modes produced by the inverse decay of the gauge quanta excited by the motion of the inflaton. Taking into account both contributions results in a power spectrum of the form

$$P_\zeta(k) = \mathcal{P}\left(\frac{k}{k_0}\right)^{n_s-1}\left[1+\mathcal{P}\,f_2\left(\xi\right)\,\mathrm{e}^{4\pi\xi}\right] \tag{147}$$

where $n_s = 1+2\eta-6\epsilon$ is the spectral index, the pivot is $k_0 = 0.002\,\mathrm{Mpc}^{-1}$, and we have defined

$$\mathcal{P}^{1/2} \equiv \frac{H^2}{2\pi|\dot{\varphi}^{(0)}|} \tag{148}$$

It is worth noting that both terms in (147) have the *same* scale dependence. Thus, we recover the standard prediction for the scalar spectral tilt in single field inflation. The function $f_2(\xi)$ which appears in (147) is plotted in Figure 4. For $\xi \gg 1$, eq. (122) provides an asymptotic expression for f_2 while, on the other hand, eq. (123) provides a good fit in the $2 \leq \xi \leq 3$ interval (as we discuss below, this is the most interesting interval for phenomenology). The COBE normalization $P_\zeta(k) \cong 22\cdot 10^{-10}$ is satisfied along the curve

$$\mathcal{P}_{\mathrm{COBE}}\left(\xi\right) \cong \frac{\mathrm{e}^{-4\pi\xi}}{2f_2(\xi)}\left[-1+\sqrt{1+10^{-8}f_2(\xi)\mathrm{e}^{4\pi\xi}}\right] \tag{149}$$

in the $\xi - \mathcal{P}$ plane. We see that, at low ξ, the contribution to (147) sourced by the gauge field is subdominant and we recover the standard result $\mathcal{P}^{1/2}_{\mathrm{COBE}} \cong 5\cdot 10^{-5}$. The two contributions in (147) become equal at $\xi \cong 2.9$. As ξ is increased, inverse decay effects dominate the spectrum and the value of $\mathcal{P}^{1/2}_{\mathrm{COBE}}$ must be (exponentially) decreased to avoid over-producing density perturbations.

The curve (149) is shown (red solid line) in Figure 7. The black dashed line shown in the Figure separates the (lower) region in which the vacuum fluctuations dominate from the (upper) region in which the fluctuations from inverse decay dominate. In the region above the solid black line in Figure 7 the backreaction bound (93) is violated. In that region of parameter space, the production of gauge field fluctuations is so strong that dissipative effects (rather than the potential V) dominate the motion of $\dot{\varphi}^{(0)}$. From Figure 7 we see that this backreaction effect can be safely disregarded after the COBE normalization

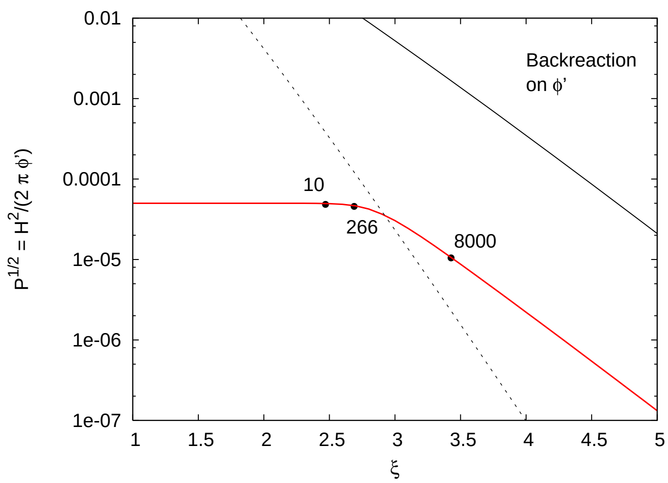

Figure 7: COBE normalization of the power spectrum in axion inflation.

is imposed. [24] Finally, Figure 7 shows some reference values of $f_{\rm NL}^{\rm equil}$ along the COBE normalized curve; we discuss this in Subsection IV 3.3.

3.2 Tensor-to-Scalar Ratio

Similarly to the result (147), the GW spectrum is also the sum of two uncorrelated contributions, one due the modes generated from the vacuum, and one due to the modes sources by the gauge field quanta. In Subsection IV 2.5 we derived the result

$$P_{\rm GW} \cong \frac{2H^2}{\pi^2\,M_p^2}\left(\frac{k}{k_*}\right)^{n_T}\left[1+\frac{H^2}{M_p^2}\,f_{h,L}\left(\xi\right)\,{\rm e}^{4\pi\xi}\right] \tag{150}$$

where $n_T = -2\epsilon$. The second term in the square braces corresponds to the gravitational waves sourced by gauge field quanta. The function $f_{h,L}$ is plotted in Figure 6. eq. (144) provides a large argument expansion while eq. (145) provides a fit in the $2 \leq \xi \leq 3$ interval (as we discuss below, this is the most interesting interval for phenomenology).

We define the tensor-to-scalar ratio in the usual way, by normalizing the amplitude of

[24] One should also ensure that the energy density of the produced quanta gives a negligible contribution to the expansion of the universe. The resulting condition, eq. (94), cannot be shown in the $\xi - \mathcal{P}$ plane, and therefore needs to be studied on a case by case basis. We have verified that this condition is satisfied for the models studied in this section.

the power in gravitational waves to that in scalar fluctuations

$$r \equiv \frac{P_{\rm GW}}{P_\zeta} \tag{151}$$

For $\xi \lesssim 1$, inverse decay effects are negligible and we recover the standard consistency relation for r, familiar from single field inflation. At $\xi \to \infty$, on the other hand, r tends to a different constant value, which is smaller than the asyptotic $\xi \to 0$ value. From (147), (150), and from the result $n_T = -2\epsilon$ it is easy to show that

$$r = \begin{cases} -8\, n_T & \text{if } \xi \lesssim 3; \\ 1.8\, n_T^2 & \text{as } \xi \to \infty. \end{cases} \tag{152}$$

Therefore, r is independent of ξ if either vacuum fluctuations or inverse decay effects dominate, while it interpolates between these two asymptotic values for intermediate ξ.

3.3 Non-Gaussianity

As discussed in Section IV 1, the cosmological fluctuations generated by inverse decay effects are highly non-Gaussian. There are many different ways to parametrize departures from Gaussianity. A standard work-horse is the local ansatz:

$$\zeta(x) = \zeta_g(x) + \frac{3}{5} f_{\rm NL}^{\rm local} \left[\zeta_g^2(x) - \langle \zeta_g^2(x) \rangle\right] \tag{153}$$

where ζ_g is a Gaussian random field and $f_{\rm NL}^{\rm local}$ quantifies the amount of non-Gaussianity. Although this simple parametrization has received considerable attention, it is certainly not the only well-motivated model for a non-Gaussian curvature perturbation. More generally, one should consider the bispectrum, $B_\zeta(k_i)$, which is the 3-point correlation function of ζ in Fourier space:

$$\langle \zeta_{\vec{k}_1} \zeta_{\vec{k}_2} \zeta_{\vec{k}_3} \rangle = B_\zeta(k_i)\, \delta^{(3)}\left(\vec{k}_1 + \vec{k}_2 + \vec{k}_3\right) \tag{154}$$

The bispectrum is a function of three momenta, $\vec{k}_i$, that form a triangle: $\vec{k}_1 + \vec{k}_2 + \vec{k}_3 = 0$. Hence a generic bispectrum may be characterized by specifying three interesting properties: the magnitude of the function, its dependence on the shape of the triangle, and its dependence of the size of the triangle. These properties are usually referred to as the *size*, *shape* and *running* of the non-Gaussianity, respectively. To characterize these properties, we find it convenient to introduce

$$|\vec{k}_1| = k \ , \ |\vec{k}_2| = x_2\, k \ , \ |\vec{k}_3| = x_3\, k \tag{155}$$

so that k encodes the overall size of the triangle while the dimensionless quantities x_2, x_3 encode its shape.

If we assume a local ansatz (153) then the bispectrum has a very particular dependence on momenta:

$$\begin{aligned} B_\zeta^{\rm local}(k_i) &= \frac{3}{10}\,(2\pi)^{5/2}\, f_{\rm NL}^{\rm local} P_\zeta(k)^2 \frac{\sum_i k_i^3}{\prod_i k_i^3} \\ &= \frac{3(2\pi)^{5/2}}{10}\frac{P_\zeta(k)^2}{k^6}\frac{1+x_2^3+x_3^3}{x_2^3 x_3^3} f_{\rm NL}^{\rm local} \end{aligned} \tag{156}$$

This function peaks in the squeezed limit where one of the wave-numbers is much smaller than the other two (for example $k_1 \ll k_2, k_3$).

3.4 The Size and Running of the Non-Gaussianity

The bispectrum from axion inflation contains two uncorrelated contributions corresponding, respectively, to the usual vacuum fluctuations and to the fluctuations generated by inverse decay processes. As is well known, the former contribution gives rise to undetectably small non-Gaussianity and may be ignored. The second contribution, however, is more interesting and this was computed in Subsection IV 2.4. We found that the bispectrum from axion inflation is very different from the local form (156). The bispectrum from axion inflation peaks on equilateral, rather than squeezed, triangles. Nevertheless, it is conventional to characterize the size of non-Gaussianity by matching to (156) on equilateral triangles $|\vec{k}_1| = |\vec{k}_2| = |\vec{k}_3|$. Proceeding in this way we find

$$f_{\rm NL}^{\rm equil} \equiv \frac{f_3\,(\xi;\,1,\,1)\;\mathcal{P}^3\,{\rm e}^{6\pi\xi}}{P_\zeta\,(k)^2} \tag{157}$$

The function $f_3\,(\xi;\,1,\,1)$ is plotted in Figure 5. Eq. (135) provides a large argument expansion of this function, while eq. (136) provides a fit in the $2 \leq \xi \leq 3$ interval. We stress that this result does not include the negligible contribution from $\langle \delta\phi_{\rm vac}^3 \rangle$ and is accurate as long as $|f_{\rm NL}^{\rm equil}| \gtrsim 1$.

From eqs. (149) and (157), from the observed power $P_\zeta \cong 22 \cdot 10^{-10}$ (we disregard the slow-roll suppressed scale dependence of $f_{\rm NL}$), and from the expressions of f_2 and f_3

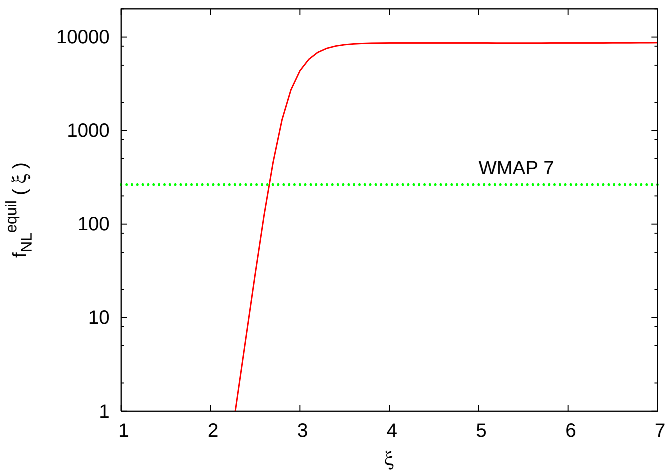

Figure 8: Non-linearity paramter $f_{\rm NL}^{\rm equil}(\xi)$ in axion inflation.

computed in the previous section, we can compute the value of $f_{\rm NL}^{\rm equil}(\xi)$ along the COBE normalized curve. We show this in Figure 8 (see also the reference values shown in Figure 7). We notice that $f_{\rm NL}^{\rm equil}(\xi)$ saturates to a constant value at large ξ (in the region where the vacuum contribution to P_ζ is negligible). From the large value asymptotic expressions (122) and (135), we find $f_{\rm NL}^{\rm equil}(\infty) \cong 8,600$. This value is already above the 95% CL upper $-214 < f_{\rm NL}^{\rm equil} < 266$ obtained from the WMAP 7 data [192]. This limit rules out $\xi \gtrsim 2.65$.[25]

Having quantified the size of non-Gaussianity in axion inflation, we now turn our attention to its running. From eq. (131) we can see that

$$B_\zeta(k_1, k_2, k_3) = k^{-6} B_\zeta(1, x_2, x_3) \tag{158}$$

disregarding the mild, slow roll suppressed scale dependence of the vacuum solutions. The overall k^{-6} behavior reflects the near scale invariance of the bispectrum from axion inflation. Slight departures from scale invariance are quantified by the index $n_{\rm NG} - 1 = \frac{d\ln|f_{\rm NL}|}{d\ln k}$ which is easily seen to be proportional to the slow roll parameters ϵ, η, and hence negligible whenever the observational bounds on n_s are satisfied. We conclude that the running of non-Gaussianity is uninterestingly small in axion inflation.

[25] The bound by the Planck satellite is $f_{\rm NL}^{\rm equil} = -42 \pm 75$ at 68% CL [13]. This essentially does not change the limit on f/α.

3.5 The Shape of the Non-Gaussianity

In order to discuss the shape of the bispectrum, it is natural to extract the strong k^{-6} scaling in (158) and define a "shape function" of the form

$$S(k_i) = N(k_1 k_2 k_3)^2 B_\zeta(k_i) \tag{159}$$

where the constant of proportionality, N, is arbitrary. This shape function coincides with the quantity that was plotted in many previous works, including [26] for example. For the case of interest, we have

$$S\left(\xi; x_2,\, x_3\right) \equiv \frac{1 + x_2^3 + x_3^3}{x_2\, x_3} \frac{f_3\left(\xi;\, x_2,\, x_3\right)}{3\, f_3\left(\xi;\, 1,\, 1\right)} \tag{160}$$

which is normalized so that $S(1,1) = 1$. Note that the bispectrum is defined only in the region $x_2 + x_3 \geq 1$, which follows from the triangle inequality. Moreover, the bispectrum is symmetric under interchange of any two momenta, and therefore we can restrict to the region $x_3 \leq x_2 \leq 1$ to avoid considering the same configuration more than once.

We plot the shape function $S(x_2, x_3)$ from axion inflation in the left panel of Figure 9. The bispectrum in this model depends on the parameter ξ. In practice, however, we find that only the size of the non-Gaussianity (quantified by $f_{\rm NL}^{\rm equil}$) depends strongly on ξ. The shape function $S(x_2, x_3)$, on the other hand, is very mildly dependent on ξ. In Figure 9 we work in the $\xi \to \infty$ limit, in which case the shape becomes independent of model parameters. (This can be seen by using the large argument expansion (121) of $\mathcal{I}$ in the expression (132) for f_3.) For $\xi \sim \mathcal{O}(1)$ this figure would be nearly indistinguishable.

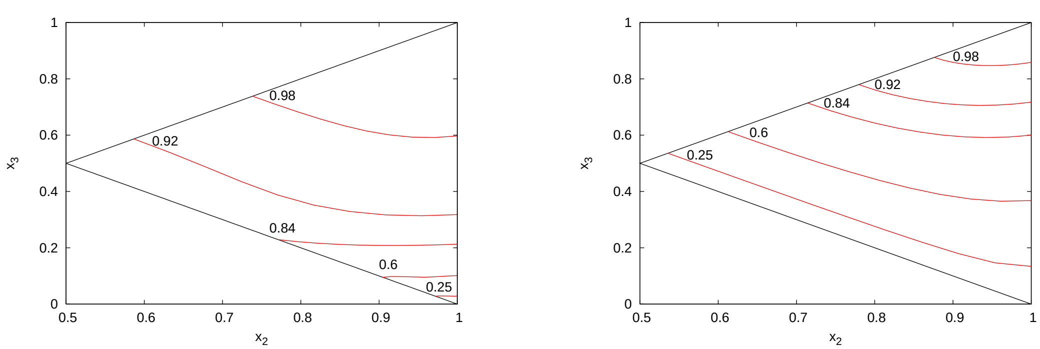

Figure 9: The shape function $S(x_2,\, x_3)$ in axion inflation.

From Figure 9 we see that the bispectrum from axion inflation peaks on equilateral triangles (corresponding to $x_2 = x_3 = 1$) and is thus qualitatively similar to the so-called equilateral template which is often employed to analyze CMB data [192, 272]

$$B_{\rm equil}(k_i) \propto -\frac{1}{k_1^3 k_2^3} - \frac{1}{k_1^3 k_3^3} - \frac{1}{k_2^3 k_3^3} - \frac{2}{k_1^2 k_2^2 k_3^2} + \frac{1}{k_1 k_2^2 k_3^3} + (5\ \text{ perms}) \tag{161}$$

(where the permutations only act on the last term). Eq. (161) is the template that is used to obtain the WMAP7 limit on $f_{\rm NL}^{\rm equil} < 266$ which we employed in Subsection IV 3.4. The shape function associated with the template (161) is plotted in the right panel of Figure 9 for comparison with the analogous result from axion inflation.

To quantitatively compare the bispectrum from axion inflation to the equilateral template (161) we follow [26] and define a scalar product between any two bi-spectra as

$$B_1 \cdot B_2 \equiv \sum_{\vec{k}_i} B_1(\vec{k}_1, \vec{k}_2, \vec{k}_3) B_2(\vec{k}_1, \vec{k}_2, \vec{k}_3) / \left(\sigma_{k_1}^2 \sigma_{k_2}^2 \sigma_{k_3}^2\right) \tag{162}$$

where $\sigma_{k_i}^2$ is the variance of a given mode and the summation runs over all possible triangles. As shown in [26], this product is the best estimator for the overlapping of any two distributions: if we assume that the real data have the bi-spectrum B_1, the template B_2 will produce a higher / lower value of non-Gaussianity according to how large / small the product (162) is [26]. To be quantitative, one defines the cosine of the "angle" between the two bi-spectra as [26]

$$\cos\left(B_1, B_2\right) \equiv \frac{B_1 \cdot B_2}{(B_1 \cdot B_1)^{1/2} (B_2 \cdot B_2)^{1/2}} \tag{163}$$

The cosine can be used to quickly estimate how well the limit given in the literature on some given template applies to a different shape [26].

We have computed the cosine of the "angle" between the bispectrum from axion inflation and the equilateral template for several different values of ξ. These results are reported in Table 1. There we see that the cosine depends only very weakly on ξ and saturates to a value $\cong 0.93$ in the limit $\xi \to \infty$. This confirms quantitatively our previous claim that the shape of non-Gaussianity is insensitive to ξ.

Table 1 shows that $\cos\left(B_{\rm inv.decay}, B_{\rm equil}\right)$ is very close to unity and hence we expect that the limit on $f_{\rm NL}^{\rm equil}$ can be applied also to axion inflation (at least to first approximation). This justifies our interpretation of the observational limit on non-Gaussianity in Subsection IV 3.4.

ξ	$\cos(B_{\text{inv.decay}}, B_{\text{equil}})$	$\cos(B_{\text{inv.decay}}, B_{\text{orth}})$
2	0.94	-0.093
3	0.94	-0.12
5	0.93	-0.13
∞	0.93	-0.15

Table 1: Cosine of the "overlapping angle" between the non-Gaussian shape in axion inflation and the equilateral and orthogonal templates.

Although the bispectrum from axion inflation is very similar to the equilateral template, the two shapes are not identical. It may be interesting to characterize the difference between these two shapes – indeed, this would become pressing in the event that Planck, or some other future mission, should detect a non-vanishing bispectrum on equilateral triangles. From Figure 9, we see that the bispectrum from inverse decay mostly differs from the equilateral template for $x_2 \simeq x_3 \simeq 1/2$, corresponding to "flattened" triangles, where one side is twice the length of the other two. For such triangles, the bispectrum from inverse decay is significantly greater than the equilateral template. Hence, this provides a natural limit in which the two shapes may be distinguished.

The fact that axion inflation gives a large non-Gaussianity on flattened triangles implies a small but nontrivial overlap with the so-called "orthogonal template" that was introduced in [272]:

$$B_{\text{orth}}(k_i) \propto -\frac{3}{k_1^3 k_2^3} - \frac{3}{k_1^3 k_3^3} - \frac{3}{k_2^3 k_3^3} - \frac{8}{k_1^2 k_2^2 k_3^2} + \frac{3}{k_1 k_2^2 k_3^3} + (5 \text{ perms}) \tag{164}$$

(where the permutations only act on the last term). The corresponding shape function evaluates to $+1$ in the equilateral limit, and to -2 along the $x_3 = 1 - x_2$ boundary (which includes the flatten triangle configurations). In Table 1 we have computed the cosine of the angle between the bispectrum from axion inflation and the template (164). We find $\cos(B_{\text{inv.decay}}, B_{\text{orth}}) \cong -0.15$ (at large ξ, see Table 1 for intermediate ξ), while $\cos(B_{\text{equil}}, B_{\text{orth}}) \cong 0.21$. Therefore, the overlap with the orthogonal template may provide a useful tool to discriminate observationally between the non-Gaussianity from axion inflation and that of the equilateral template.

3.6 Large Field Inflation

Once the COBE normalization is imposed, the key phenomenological predictions of any inflationary model are the spectral index, n_s, the tensor-to-scalar ratio r, and the non-linearity parameter $f_{\rm NL}$. As we discussed previously, the spectral index in axion inflation has the standard form $n_s = 1 + 2\eta - 6\epsilon$ and requires no further discussion. The remaining observables depend on the coupling α/f, the background axion velocity $\dot{\varphi}^{(0)}$ and the Hubble rate H. Out of these three quantities, we have defined the two combinations

$$\xi \equiv \frac{\alpha \dot{\varphi}^{(0)}}{2 f H} \quad , \quad \mathcal{P}^{1/2} \equiv \frac{H^2}{2\pi |\dot{\varphi}^{(0)}|} \tag{165}$$

Both the two and three point correlation functions of ζ can be written solely in terms of these two combinations. Therefore both the power spectrum and the bi-spectrum are a functions of ξ and $\mathcal{P}$ only. The COBE normalization fixes $\mathcal{P}^{1/2}$ in terms of ξ, see Figure 7. Therefore, the predicted non-linearity parameter $f_{\rm NL}$ is function of ξ only, see Figure 8.

The tensor to scalar ratio, on the other hand, is a function of a different combination of parameters, see eq. (150). For this reason, we cannot present it in a plot as a function of ξ only. This can only be achieved once the potential for the axion is specified, since this provides one additional (slow-roll) relation between the parameters of the model.

For a specific choice of inflationary potential $V(\varphi)$, it is of interest to determine how the combinations (165), along with the tensor-to-scalar ratio, depend on parameters of the underlying theory. As we discuss in [48] and also in Section IV 5, in the models of axion inflation of interest, the axion/inflaton dynamics effectively occurs as in a large field inflationary model with potential:

$$V = \frac{\lambda_p}{p}\, \varphi^p \tag{166}$$

where λ_p has mass dimension $4 - p$. Using slow roll approximation, we find

$$H = \sqrt{\frac{\lambda_p}{3\,p}\, \frac{(2\,p\,N)^{p/4}}{M_p^{1-p/2}}} \quad , \quad \dot{\varphi}^{(0)} = \sqrt{\frac{p}{2N}}\, H\, M_p \quad , \quad \xi = \frac{\alpha\, M_p}{2 f} \sqrt{\frac{p}{2N}} \tag{167}$$

where N is the number of e-folds between the moment the CMB scales left the horizon and the end of inflation.

The value of λ_p is fixed by COBE normalization and we assume $N = 60$ e-foldings of inflation. Then, for any given value of p, all observational predictions can be written in terms of f/α only. In axion monodromy [231], $p = 1$; in most of the other models one expands the potential close to the minimum, where it is quadratic, $p = 2$. We therefore show in Figure 10 the predicted values of $f_{\rm NL}$ and of r for these two cases. We notice that, once $f_{\rm NL}$ is required to be below the CMB bound, the standard value for r is recovered.

It is interesting to note that axion inflation models generically predict the *same* values

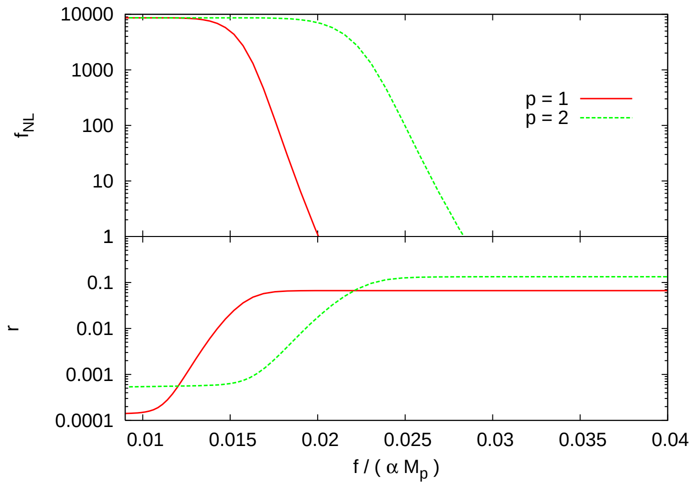

Figure 10: Predicted values for the equilateral $f_{\rm NL}$ parameter and tensor-to-scalar ratio r in axion inflation models.

of n_s and r as would be obtained in vanilla chaotic inflation. However, our scenario predicts also a large non-Gaussianity with a (nearly) equilateral shape. Axion inflation provides a rare example of a theory which predicts both a detectable tensor-to-scalar ratio *and* a large equilateral bispectrum. Note that, if such a non-Gaussian signal is eventually detected, then it will immediately fix the value of the coupling α/f. On the other hand, if it is not detected, then we will have a surprisingly stringent bound on the strongest axion-type couplings between the inflaton and *any* gauge field.

4 Polarized GWs from Axion Inflation at GW Detectors

Another possible phenomenological signature from axion inflation is *polarized* GWs. As seen in Subsection IV 2.5, the explicit calculation of $f_{h,L/R}$ shows that one of the polarization states is produced by a significantly larger amount than the other polarization in the sourced contribution – see (144). Physically, this is related to spin conservation in the sourcing process $A + A \to h$; since only one of the circular polarizations of the spin-1 gauge field (A_+) is enhanced by the tachyonic growth, the corresponding circular polarization of the spin-2 gravitational wave (h_L) is sourced much more effectively than the other polarization state.

This chiral feature is unique in the sourced contribution to the GW signal, that is, the contribution from the standard vacuum fluctuations, $h_\lambda^{\rm vac}$, is unpolarized ($h_L^{\rm vac} = h_R^{\rm vac}$). As discussed in Subsection IV 3.6, once the constraint on the scalar perturbations at the CMB scales is imposed, the tensor-to-scalar ratio r is standard at the same scales, $r = 16\epsilon$, and thus is unobservable in the minimal implementation of the model.[26] On the other hand, the phenomenological signatures of the sourcing mechanism depend on the strength of the (effective) inflaton-gauge coupling ξ, given in (87), and since $\xi \propto \frac{\dot{\varphi}^{(0)}}{H}$, its strength slowly grows during inflation. The growth of ξ results in a greater gauge quanta production at progressively smaller scales. The copiously produced gauge quanta then source the GWs (and also curvature perturbations). The GWs produced by these quanta can reach a large amplitude at small scales and can be dtected at terrestrial GW detectors [47, 99].

The more effective production of gauge quanta at smaller scales can also enhance the effects on the curvature perturbations through the inverse decay, extensively studied in Subsections IV 2.1–2.4. This could potentially lead to another effect, namely to the possible overproduction of primordial black holes [208]. However, this theoretical scenario carries a large uncertainty [22, 47, 208], while our present consideration of the GW signals at small scales does not.

In this Section, we consider the polarized GW signals at the scales relevant to terrestrial GW detectors and generate forecasted constraints on the strength of ξ for future detectors. The polarization asymmetry can be measured by tracking the amplitude of different polarization modes as a function of time at a given GW detector and by comparing it to the same measurement made by other detectors at different locations. We follow the formalism developed in [20], modified to address a polarized stochastic GW background (SGWB) signals as discussed in [275].[27] Estimating the sensitivity of the future GW detectors to the SGWB polarization, we study how the axion inflation model could be constrained. In particular, we consider the upcoming second-generation GW detectors — Advanced LIGO (aLIGO) [162] detectors at Hanford, WA (H1) and Livingston, LA (L1), Advanced Virgo [3] in Italy (V1), GEO-HF [309] in Germany, and KAGRA [201, 296] in Japan (K1) are ex-

[26] In Subsection IV 3.6, the large-field cases were considered. In the models of small-field inflation, the value of ϵ is even smaller, and so is the value of r. Thus, once the CMB constraint on non-Gaussianity is respected, the GW signal is completely unobservable in either case.

[27] A SGWB is expected to arise from the superposition of gravitational waves (GWs) from many uncorrelated and unresolved sources. Numerous cosmological SGWB models have been proposed, including inflationary models [58, 64, 155, 286], models based on cosmic (super)strings [105, 279], and models of alternative cosmologies [139]. Furthermore, various astrophysical models have been proposed based on integrating contributions from astrophysical objects across the universe, such as compact binary coalescences of binary neutron stars and/or black holes [251, 311], magnetars [102, 228], or rotating neutron stars [256].

pected to have $\sim 10\times$ better strain sensitivities than the first-generation detectors and to produce first science-quality data in 2015. We also consider an example configuration of a pair of third-generation GW detectors, with strain sensitivity similar to the proposed Einstein Telescope [266].

The formalism we apply here to axion inflation is more generic and can be applied to any GW signals in search of circularly polarized isotropic SGWB [101]. Polarization asymmetry in the SGWB could be generated in the early universe if parity is violated either explicitly or spontaneously (as in the axion inflation models).[28] In all of these models, parity violation breaks the symmetry between the two circular polarization modes. Since astrophysical sources are unlikely to induce such asymmetry in isotropic SGWB, detecting circularly polarized SGWB is a potentially excellent way of distinguishing between the cosmological and astrophysical SGWB contributions.

4.1 Background Evolution in the Strong Backreaction Regime

In the previous Sections, we always assumed that the conditions (93) and (94) were satisfied at the CMB scales and that therefore the backreaction from the produced gauge field had negligible effects on the background evolution (89, 90).[29] This assumption is certainly correct within the parameter range ($\xi \lesssim 2.66$) allowed by the non-Gaussianity constraints. However, $\xi \propto \frac{\dot{\varphi}^{(0)}}{H}$ is a slowly growing function of time, and this makes the production, which is exponentially sensitive to the value of ξ, more efficient at later stages of inflation. Toward the end of inflation, this efficient production can lead to non-negligible effects on the background dynamics, even for the values of ξ that are allowed by the CMB constraints.

The gauge field fluctuations are produced at the expense of the kinetic energy of the homogeneous inflaton $\varphi^{(0)}(t)$, introducing a new source of dissipation into the equation of motion of $\varphi^{(0)}$. Moreover, these produced fluctuations contribute to the energy density of the universe and thus modify the Friedmann equation. These effects are encoded in the mean field equations (89) and (90), with $\langle \vec{E} \cdot \vec{B} \rangle$ and $\frac{1}{2}\langle \vec{E}^2 + \vec{B}^2 \rangle$ given in (92).

The effects were fully treated in [47], in which eqs. (89) and (90) were numerically

[28] Explicit breaking mechanisms are typically due to quantum gravity effects, such as the imaginary part of the Immirzi parameter [221] and higher curvature terms in some power-counting renormalizable theories of gravity [292]. On the other hand, spontaneous breaking mechanisms are typically due to axial couplings of an axion (or an axion-like, pseudo-scalar field) to gravity [17, 216, 275] and/or a gauge field [284], which is of our current focus.

[29] Using the relation $|\dot{\varphi}^{(0)}| \simeq \sqrt{2\epsilon}\, H M_p$, we see that the condition (93) is more stringent than (94). If we take the standard result for the scalar spectrum, $P_\zeta \simeq \left(\frac{H^2}{2\pi\dot{\varphi}^{(0)}}\right)^2 \simeq 2.2 \cdot 10^{-9}$, then the condition reads $\xi \lesssim 4.7$.

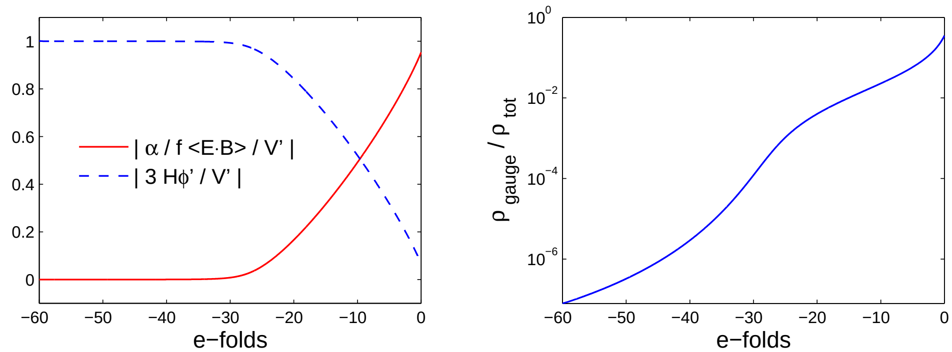

Figure 11: Friction terms in the equation of motion for $\varphi^{(0)}$ (left panel), and the fractional energy density of the produced quanta (right panel) [47].

evolved,[30] taking into account the backreaction term in the equation for $\varphi^{(0)}$. The energy density of the gauge quanta in the Friedmann equation (90) was still disregarded, which was the assumption that was justified by the numerical results. The right panel of Figure 11 shows that this is a good approximation all the way until the end of inflation [47]. Here, $\xi = 2.5$ is taken at $|\varphi^{(0)}| \simeq 11\, M_p$, which is the value of the inflaton about 60 e-foldings before the end of inflation, under the standard slow-roll assumption.

On the other hand, the production of gauge quanta introduces an additional friction on the inflaton motion that prolongs the duration of inflation. The left panel of Figure 11 compares the evolution of the two "friction" terms in the equation (89) for $\varphi^{(0)}$ as a function of the number of $e-folds$ to the end of inflation [47]. As seen in the plot, the backreaction of the produced quanta on the $\varphi^{(0)}$ dynamics becomes noticeable during the last 25 e-folds of inflation, while it is negligible at earlier times. The standard Hubble friction controls the early stages, but the friction due to the produced gauge quanta gradually increases and eventually dominates the inflaton evolution from about 10 e-folds before the end of inflation. This in fact results in the increase of the duration of inflation by about 10 e-folds [47].

The parameters are chosen such that, initially, we are in the standard slow-roll regime, and therefore the results for the power spectrum and for $f_{\rm NL}$ obtained in the previous Sections IV 2-3 still apply. The strong backreaction regime is reached only in the last $\mathcal{O}(10)$ e-folds of inflation.

Since the inverse-decay effects on the cosmological perturbations from the produced

[30] Monomial functions are considered for the inflaton potential, i.e. $V(\varphi) \propto \varphi^p$. Ref. [47] chose $p = 1$, typical of axion monodromy models in string theory construction [129, 231]. Later in this Section, we will give the forecast for the values of ξ from the GW signals at the future-generation GW detectors in the cases of $p = 1, 2$.

gauge quanta are encoded near the horizon exit of each mode, the modes that leave the horizon much later than the CMB modes (but still during inflation) are significantly affected by the increase in the coupling strength ξ. The scales relevant to the terrestrial GW detectors are orders of magnitude smaller than the CMB scales and leave the horizon during the time when the backreaction onto the background dynamics plays a significant role in determining the evolution of the homogeneous background, $\varphi^{(0)}(t)$ and $H(t)$. To consider the GW signals at the detector scales, it is therefore necessary to include the consistent treatment of this effect, which we do in the current Section of this thesis.

The tensor perturbation obeys the equation of motion given in (141), and its power spectrum is computed in (142), for left- and right-handed helicity states, where $f_{h,L/R}$ is computed from (143). In order to compare our results with the sensitivities of the GW detectors, it is convenient to introduce the fractional GW energy density per logarithmic wavenumber interval

$$\Omega_{\rm GW} \equiv \frac{1}{3H_0^2 M_p^2} \frac{\partial \rho_{\rm GW,0}}{\partial \log k} \tag{168}$$

where the subscripts 0 denote the quantities evaluated today and

$$\rho_{\rm GW,0} \equiv \frac{M_p^2}{8a^2} \langle h'_{ij} h'_{ij} + \partial_k h_{ij}\, \partial_k h_{ij} \rangle \; . \tag{169}$$

The evaluation of $\rho_{\rm GW,0}$ requires to evolve the primordial signal (142) after horizon re-entry until today. Using the tensor transfer function (see e.g. [70, 306]), it is then found [47],

$$\Omega_{\rm GW} = \frac{\Omega_{R,0}}{24} \left(P_L + P_R \right) \tag{170}$$

where $\Omega_{R,0} \equiv \frac{\rho_{R,0}}{3H_0^2 M_p^2} \simeq 8.6 \cdot 10^{-5}$ is the fractional density of radiation today.

In considering the GW signal at the terrestrial GW detector scales, it is useful to see the values of $\Omega_{\rm GW}$ as a function of frequency $f = \frac{k}{2\pi}$. The detailed history of reheating is necessary to relate the size of any given mode to the number of e-folds, N, at which this mode exits the horizon before the end of inflation; this relation is given in (34). As we have seen from the right panel of Figure 11, the gauge fields may carry a non-negligible fraction of the total energy density, already by the end of inflation. One may therefore expect that the equation of state of the universe will have an intermediate value between that of radiation and matter [252], leading to the number of e-folds between 59 and 62 for the CMB pivot scale [47]. Denoting by $N_{\rm CMB}$ the number of e-folds between the horizon

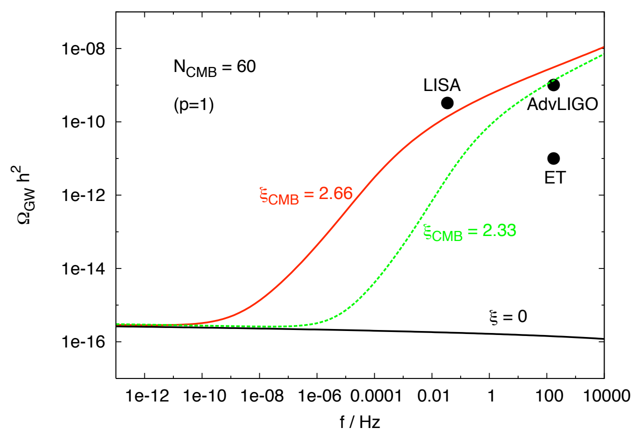

Figure 12: $\Omega_{\mathrm{GW}}h^2$ as a function of the frequency f, for 60 e-folds of observable inflation, with a linear inflaton potential [47].

exit of the CMB modes and the end of inflation, we have the relation

$$\begin{aligned} N - N_{\mathrm{CMB}} &= \ln \frac{k_{\mathrm{CMB}}}{0.002\,\mathrm{Mpc}^{-1}} - \ln \frac{k}{0.002\,\mathrm{Mpc}^{-1}} \\ &= \ln \frac{k_{\mathrm{CMB}}}{0.002\,\mathrm{Mpc}^{-1}} - 44.9 - \ln \frac{f}{10^2\,\mathrm{Hz}} \,. \end{aligned} \tag{171}$$

The resulting GW signal as a function of frequency f, with the inclusion of the backreaction effects, is presented in Figure 12 for the case of a linear inflaton potential, together with three reference values of the inflaton-gauge field coupling: $\xi_{\mathrm{CMB}} = 2.66$ corresponds roughly to the highest value allowed by the CMB limits on equilateral non-Gaussianity, $\xi_{\mathrm{CMB}} = 2.33$ roughly to the lowest possible value that can be detected at Advanced LIGO/VIRGO, and $\xi_{\mathrm{CMB}} = 0$ to the standard case without any inflaton-gauge coupling [47]. In the Figure, the expected sensitivities of LISA, Advanced LIGO/VIRGO and Einstein Telescope (at their most sensitive frequency) are also shown.

There are two competing effects during the phase of increasing values of ξ. One effect can be seen in Figure 12: from left to right, one can see the vacuum contribution, then the fast growth of the sourced contribution and finally the slow-down of the growth when the backreaction onto the motion of $\varphi^{(0)}$ becomes sizable. Secondly, for larger ξ_{CMB}, the whole signal is effectively shifted toward lower frequencies (to the left) due to the larger number of additional e-folds of strong backreaction to the end of inflation. Since the second effect is more prominent, although the two effects act in opposite directions, the strong backreaction phase effectively results in an increase of the signal. This increase is crucial and makes the natural window $N_{\rm CMB} = 50-62$ potentially visible at the (future-generation) GW detectors.

The sourced part of the GWs in axion inflation is circularly polarized, i.e. $P_L \gg P_R$, as found in Subsection IV 2.5. Thus the total GW signals will show the same feature whenever the sourced contribution dominates over the vacuum one (this is necessary for any hope for a detection of the primordial GWs from axion inflation models at the terrestrial GW detectors). In the next Subsection, we develop the formalism to search for a generic SGWB signal that is circularly polarized. Then in the following Subsection, we apply this formalism to provide the forecast sensitivities for ξ from the future-generation GW detectors.

4.2 Search Formalism

Following [275], we start from the plane-wave expansion of the metric at time t and position $\vec{x}$:

$$h_{ab}(t,\vec{x}) = \sum_A \int_{-\infty}^{\infty} df \int_{S^2} d\hat{\Omega} h_A(f,\hat{\Omega}) e^{-2\pi i f(t-\vec{x}\cdot\hat{\Omega})} e^A_{ab}(\hat{\Omega}), \tag{172}$$

where $e^A_{ab}(\hat{\Omega})$ is the polarization tensor associated with a wave traveling in the direction $\hat{\Omega}$, and f is frequency (we use natural units $c = \hbar = 1$). We consider the left- and right-handed correlators [275]:

$$\langle h_{R/L}(f,\hat{\Omega}) h^*_{R/L}(f',\hat{\Omega}')\rangle = \frac{\delta(f-f')\delta^2(\hat{\Omega}-\hat{\Omega}')}{4\pi}\left[I(f) \pm V(f)\right] \tag{173}$$

where $h_L = (h_+ + ih_\times)/\sqrt{2}$, $h_R = (h_+ - ih_\times)/\sqrt{2}$, and $+$ and $\times$ are the standard plus and cross polarizations.

Note this is the point of departure from the past searches for unpolarized isotropic SGWB, which assume $V = 0$. Further note that $\langle h_R h^*_L\rangle$ vanishes due to statistical isotropy. The normalized energy density is then given by [20, 275]:

$$\Omega_{\rm GW}(f) = \frac{f}{\rho_c}\frac{d\rho_{\rm GW}}{df} = \frac{\pi f^3}{G_N \rho_c} I(f) \tag{174}$$

where $d\rho_{\rm GW}$ is the energy density in the range $[f,\ f+df]$, G_N is Newton's constant, and ρ_c

the critical energy density of the universe.[31] We also compute the standard cross-correlation estimator [20]:

$$\begin{aligned}\langle\hat{Y}\rangle &= \int_{-\infty}^{+\infty} df \int_{-\infty}^{+\infty} df' \delta_T(f-f')\langle(s_1^*(f)s_2(f')\rangle\tilde{Q}(f') \\ &= \frac{3H_0^2 T}{10\pi^2}\int_0^{\infty} df \frac{\Omega'_{\rm GW}(f)\gamma_I(f)\tilde{Q}(f)}{f^3}, \end{aligned} \tag{175}$$

where

$$\begin{aligned}&\Omega'_{\rm GW}(f)\gamma_I(f) = \Omega_{\rm GW}(f)\left[\gamma_I(f) + \Pi(f)\gamma_V(f)\right] \\ &\gamma_I(f) = \frac{5}{8\pi}\int d\hat{\Omega}(F_1^+F_2^{+*} + F_1^\times F_2^{\times *})e^{2\pi i f\hat{\Omega}\cdot\Delta\vec{x}} \\ &\gamma_V(f) = -\frac{5}{8\pi}\int d\hat{\Omega} i(F_1^+F_2^{\times *} - F_1^\times F_2^{+*})e^{2\pi i f\hat{\Omega}\cdot\Delta\vec{x}}.\end{aligned} \tag{176}$$

Here, T is the measurement time, $\delta_T(f) \equiv \sin(\pi f T)/(\pi f)$, $\tilde{s}_1(f)$ and $\tilde{s}_2(f)$ are Fourier transforms of the strain time-series of two GW detectors, $\tilde{Q}(f)$ is a filter, and $F_n^A = e_{ab}^A d_n^{ab}$ is the contraction of the tensor mode of polarization A, e_{ab}^A, with the response of the detector n, d_n^{ab}.[32] The factor $\gamma_I(f)$ is the standard overlap reduction function arising from different locations and orientations of the two detectors, and $\gamma_V(f)$ is a new function, associated with the parity violating term and first computed in [275]. Figure 13 shows these functions for two real detector pairs. Finally, $\Pi(f) = V(f)/I(f)$ encodes the parity violation, with maximal values $\Pi = \pm 1$ corresponding to fully right- or left-handed polarizations. Setting $\Pi = 0$ reproduces the standard unpolarized SGWB search [20].

Assuming stationary Gaussian detector noise (uncorrelated between two detectors), the estimator for the variance associated with $\hat{Y}$ is [20]:

$$\sigma^2 = \frac{T}{4}\int_0^{\infty} df P_1(f)P_2(f)|\tilde{Q}(f)|^2\,, \tag{177}$$

where $P_n(f)$ are the one-sided noise power spectral densities of the two GW detectors. In practice, we divide the sensitive frequency band of the GW detectors into bins $\Delta f = 0.25$ Hz wide [6]. We then compute the estimator $\hat{Y}_i$ and the variance σ_i^2 for each frequency bin i assuming a frequency independent spectrum template $\Omega_{\rm GW}(f_i) = \Omega_0$ for each bin. Optimization of the signal-to-noise ratio then leads to the following optimal filter for a

[31] Note that the similar eq. (3) of [275] contains an additional factor of 4, which we believe is incorrect.

[32] Note that the minus sign in the equation for $\gamma_V(f)$ is missing in the computation of [275].

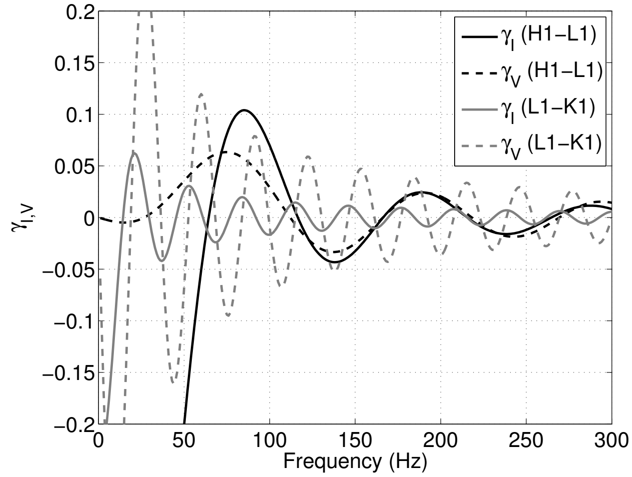

Figure 13: Overlap reduction functions for detector pairs H1-L1 and L1-K1.

frequency-independent GW spectrum in the frequency bin f [20]:

$$\tilde{Q}(f_i) = \mathcal{N} \, \frac{\gamma_I(f)}{f_i^3 P_1(f_i) P_2(f_i)} \, , \tag{178}$$

with normalization constant $\mathcal{N}$ chosen so that $\langle \hat{Y}_i \rangle = \Omega_0$.

To perform parameter estimation in the parity violating models, we adopt the Bayesian approach introduced in [227]. We define the following likelihood function:

$$L(\hat{Y}_i, \sigma_i | \vec{\theta}) \propto \exp\left[-\frac{1}{2} \sum_i \frac{(\hat{Y}_i - \Omega'_{\rm M}(f_i; \vec{\theta}))^2}{\sigma_i^2} \right] \tag{179}$$

where the sum runs over frequency bins f_i and $\Omega'_{\rm M}(f_i;\vec{\theta})$ is the SGWB energy density spectrum $\Omega_{\rm M}(f_i;\vec{\theta}_1)$ specified by some of the free parameters $\vec{\theta}_1$, multiplied by the parity-violating correction $(1 + \Pi(f_i;\vec{\theta}_2)\gamma_V(f_i)/\gamma_I(f_i))$ (see eq. 176), where $\vec{\theta}_2$ is another set of parameters specifying the parity violation term Π, and $\vec{\theta} = (\vec{\theta}_1, \vec{\theta}_2)$. Multiplying the likelihood with the prior distribution for $\vec{\theta}$ yields the Bayesian posterior distribution, which can then be used to extract confidence intervals for $\vec{\theta}$. In the subsequent results, we take all priors to be flat within the plotted range of $\vec{\theta}$ (and zero elsewhere).

Note that the choices of $\Omega_{\rm M}(f;\vec{\theta}_1)$ and $\Pi(f;\vec{\theta}_2)$ (or equivalently of $I(f;\vec{\theta})$ and $V(f;\vec{\theta})$) are assumptions underlying the analysis. It is clear that if the analysis is performed using a single detector pair and a single frequency bin, it will not be possible to estimate multiple parameters, and in particular to detect parity violation. If the analysis is done for a single detector pair, but for a series of frequency bins, then in principle it is possible to break degeneracies between different parameters of a given model — in particular, *for the given model* it may be possible to measure parity violation. However, with a single detector pair it is not possible to detect parity violation *independently of the model*. For example, with a single detector pair it is not possible to distinguish between a SGWB spectrum $\Omega_{\rm M}(f;\vec{\theta}_1)$ with a parity violating term $\Pi = 1$ from a parity-conserving model ($\Pi = 0$) with a different SGWB spectrum, given by $\Omega_{\rm M,2}(f;\vec{\theta}_1) = \Omega_{\rm M}(f;\vec{\theta}_1)(1 + G(f))$, where the correction factor coincidentally happens to be $G(f) = \gamma_V(f)/\gamma_I(f)$ for the overlap reduction functions $\gamma_V(f)$ and $\gamma_I(f)$ *for that specific detector pair*. While such a coincidence is already unlikely for any reasonable SGWB models (due to the complex frequency dependence of $\gamma_I(f)$ and $\gamma_V(f)$), it could be definitively ruled out by inclusion of additional detector pairs. In other words, while such a coincidence may happen for a given detector pair, it cannot happen for multiple detector pairs simultaneously, since different pairs are associated with different overlap reduction functions.

4.3 Parity Violation from Axion Inflation

We now apply the formalism of Section IV 4.2 to the parity violating GWs in axion inflation. The last term in the dispersion relation $\omega_\pm = k^2 \pm \frac{2k\xi}{\tau}$, see (87), becomes dominant as the Compton wavelength $\lambda \sim a/k$ becomes greater than the Hubble horizon H^{-1}, at which point one of the two helicities experiences a tachyonic growth. As in the previous Sections, we assume $\xi > 0$ so that the growing helicity is the + one. The gauge quanta produce chiral gravitational waves [45, 284] through the $A_+A_+ \to h$ process. These "sourced" modes superpose incoherently with the corresponding "vacuum" modes. If the sourced gravitational waves dominate over the vacuum ones, $\Pi \simeq -1$.

The phenomenological signatures of this mechanism depend on the strength of the inflaton-gauge field coupling (encoded by the parameter ξ given in eq. (87)), and on the inflaton potential. For definiteness, a monomial inflaton potential is typically assumed, $V \sim \varphi^p$. For each p, $\Omega_{\rm M}(f;\xi)$ in eq. (179) is a function of the single parameter ξ. Ref. [232] computed the limits on ξ from the non-observation of the sourced scalar perturbations at the CMB scales. This analysis was based on the existing WMAP [192] and ACT [122] data. For $p = 2$, they obtained the limit $\xi < 2.41$ (at 95% CL) on the value of ξ when the CMB modes left the horizon [232]. They also performed a more general search over different inflationary potentials by marginalizing over the slow roll parameters in the fit. This leads

Table 2: 2σ forecasted constraints on ξ for future gravitational wave interferometers

Detector Network	p	ξ (sensitivity)	ξ (exclude $\Pi = 0$)
2^{nd} Gen H1-L1-V1-K1	1	2.3	2.8
2^{nd} Gen H1-L1-V1-K1	2	2.2	2.6
3^{rd} Gen	1	1.8	2.0
3^{rd} Gen	2	1.9	2.0

to $\xi < 2.5$ (at 95% CL). These results are based on a flat prior in the interval $0 < \xi < 10$.

In this analysis, we present results for $p = 1$ (typical of axion monodromy [231]) and $p = 2$ (a generic Taylor expansion around the minimum of a potential). The growth of ξ during inflation, $\xi \propto \frac{\dot{\varphi}^{(0)}}{H}$, results in a greater gauge quanta production at progressively smaller scales. The GW produced by these quanta can reach a large amplitude at small scales, and can be detectable at terrestrial GW detectors [47, 99]. We apply the formalism developed in Section IV 4.2 to evaluate the current LIGO limits and the future detection prospect (both for the second-generation H1-L1-V1-K1 detector network and for the example third-generation detector pair) for such a signal. We assumed a flat prior distribution for ξ over $0 \leq \xi \leq 4$. Similarly to the CMB limit [232], the sourced signal grows exponentially at $\xi > 1$ and the posterior-distribution drops to zero extremely fast at $\xi > 3$. Therefore, the posterior distribution is insensitive to the upper bound on the prior distribution, as long as it is ~ 4 or greater.

In order to be directly comparable to the CMB bound, we also present our results in terms of the value assumed by ξ when the CMB modes left the horizon. We found that using the recent LIGO results set the limit $\xi \leq 3.5$ (at 95% CL) for $p = 1$ and $\xi \leq 3.4$ (at 95% CL) for $p = 2$. These limits are weaker than the current CMB limits of [232]. However, our predictions for future interferometer results (Table 2) show that second-generation detectors will begin to constrain the ξ parameter beyond the CMB bound and third-generation detectors may reach $\xi \sim 1.8$. Somewhat larger values of ξ are required in order to detect parity violation in the observed SGWB and rule out $\Pi = 0$ at the 2σ level for this model.

The results summarized in Table 2 show that future SGWB searches will firmly improve over the limit $\xi < 2.41$ that is based on the CMB measurements of the power spectrum of scalar perturbations. However, the growth of scalar perturbations could lead to another effect, i.e. the possible overproduction of primordial black holes [208]. The bound from the primordial black holes, $\xi \lesssim 1.5$ (where, again, the numerical value refers to the value assumed by ξ when the CMB modes left the horizon), is potentially stronger than those obtained here, but it carries an uncertainty associated with how many scalar perturbations are produced at such small scales [22, 47, 208] and with how many black holes they create [208]. No such uncertainties are present for the SGWB signal that we have studied here.

5 Models of Axion Inflation

Our results concerning non-Gaussianity and inverse decay effects are quite general and may be applied directly to a variety axion inflationary models. Moreover, we expect that our qualitative results may have also implications more broadly, for example in any multi-field scenario that involves a dynamical axion.[33] In this section we survey some interesting microscopic constructions, both from field theory and string theory, and comment on the possibility of large non-Gaussianity. The key input which our analysis requires from a microscopic computation are the inflaton potential $V(\varphi)$, the decay constant f, and the dimensionless parameter α. Together, these determine the quantity

$$\xi \cong \frac{\alpha M_p^2}{2f} \left| \frac{V_\varphi}{V} \right| \tag{180}$$

that measures the strength of inverse decay effects. As we have seen, inverse decay processes cannot be neglected when $\xi \gtrsim \mathcal{O}(1)$ which roughly translated into $f/(\alpha M_p) \lesssim 10^{-2}$ for the most interesting models; see Figure 10. Whenever this inequality is satisfied, our findings strongly affect the phenomenology of these models.

In most of the scenarios that we survey, both $V(\varphi)$ and f tend to be fairly well understood while, on the other hand, the coefficient α is rather more model dependent (although calculable in principle). A detailed computation of α in each interesting scenario is beyond the scope of this thesis and we leave such an analysis to future work. Although we generically expect $\alpha = \mathcal{O}(1)$, it should nevertheless be noted that this parameter contributes a source of theoretical uncertainty to what follows. We believe that our phenomenological results should provide a motivation for a detailed microscopic computation of α in the various scenarios discussed below.

[33] This seems to be a generic expectation for closed string inflation models. In the effective SUGRA description the Kahler moduli τ_i are typically paired with axions θ_i into complex fields $T_i = \tau_i + i\theta_i$. Absent tuning, one expects the curvature of the scalar potential to be comparable in the τ_i and θ_i directions. Hence, nontrivial modular dynamics is typically also associated with nontrivial dynamics in the axion sector. See, for example, reference [174] for a review.

5.1 Natural Inflation

The original natural inflation proposal [132] was based on the potential[34]

$$V(\varphi) \cong \Lambda^4 \left[1 - \cos\left(\frac{\varphi}{f}\right)\right] \tag{181}$$

For a spectral index $n_s \gtrsim 0.95$ this model requires $f \gtrsim \sqrt{8\pi}\, M_p$ [265], for which the inflaton dynamics shows little difference from that standard chaotic inflation with $V(\varphi) \cong \frac{1}{2}m^2\varphi^2$, at least for the present considerations. Such large values of f weaken the pseudo-scalar coupling $\varphi F\tilde{F}$ and inverse decay effects are negligible unless $\alpha \gtrsim$ few $\times\, 10^2$. While specific situations have been constructed that can result in a large α – for instance in the some extra-dimensional model [22] – such large values conflicts with our general expectation that $\alpha = \mathcal{O}(1)$. We conclude that large non-Gaussianity seems unlikely in the simplest models of natural inflation.

In spite of apparent simplicity, however, the original natural inflation model [132] seems incompatible with UV completion. If we interpret φ as a PNGB then $f > M_p$ suggests a global symmetry broken *above* the quantum gravity scale, where effective field theory is presumably not valid. Moreover, $f > M_p$ does not seem possible in a controlled limit of string theory [30] (which is the only known framework wherein such questions may be addressed). Hence, requiring the existence of a sensible UV completion automatically pushes us towards the regime $f \ll M_p$ where inverse decay effects are important. We will illustrate how this works in several explicit microscopic realizations below, however, it is clear that this trend applies more generally.

5.2 Double-Axion Inflation

Perhaps the simplest scenario to realize natural inflation with $f < M_p$ is the double-axion model proposed in [187]. This model is characterized by two axions, θ and ρ, whose potential

$$V(\theta, \rho) = \sum_{i=1}^{2} \Lambda_i^4 \left[1 - \cos\left(\frac{\theta}{f_i} + \frac{\rho}{g_i}\right)\right] \tag{182}$$

arises from pseudo-scalar couplings to two different gauge groups: $\frac{\theta}{f_i} F_i\tilde{F}_i$ and $\frac{\rho}{g_i} F_i\tilde{F}_i$. For $f_1 g_2 = g_1 f_2$ one linear combination of θ and ρ becomes a flat direction of the potential,

[34] See [9] for some proposed particle physics realizations. In [174] this model was realized in supergravity – the so-called "axion valley" model.

corresponding to an enhanced symmetry of the theory. Taking $f_1 g_2 \cong g_1 f_2$ the curvature in this direction becomes controllably flat and one may realize natural inflation with $f_i, g_i \ll M_p$. This scenario therefore provides a simple and compelling illustration of how the requirement of a UV completion leads to large non-Gaussianity in natural inflation.

5.3 N-flation

The N-flation model [116] is based on N axion fields φ_i, each with its own softly broken shift symmetry, resulting in a separable potential of the form

$$V(\varphi_i) \cong \sum_i \Lambda_i^4 \left[1 - \cos\left(\frac{\varphi_i}{f_i}\right)\right] \cong \sum_i \frac{1}{2} m_i^2 \varphi_i^2 \tag{183}$$

where, in the second equality, we have expanded in small field values $\varphi_i \ll f_i$. For $N \gg 1$ the assisted inflation mechanism [207] allows inflation to proceed even while all the decay constants are sub-Planckian, $f_i < M_p$. To a first approximation, the dynamics may be captured by a simple quadratic potential $V_{\text{eff}} \cong \frac{1}{2} m^2 \Phi^2$ for the collective field $\Phi^2 \equiv \sum_i \varphi_i^2$ [125]. Successful inflation requires that the collective field traverses a super-Planckian distance in field space, $\Delta\Phi \gtrsim M_p$. This is achievable with sub-Planckian φ_i provided the number of axions is sufficiently large. Roughly speaking we require [176]

$$N \sim 240 \left(\frac{M_p}{f_{\text{avg}}}\right)^2 \tag{184}$$

For $f_{\text{avg}} \sim 10^{-1} M_p$ we have $N \sim 10^3$ while $f_{\text{avg}} \sim 10^{-2} M_p$ would require $N \sim 10^6$.

N-flation makes sense as a purely field theoretical construction; however, much of the interest in this scenario arises because the exponentially large values of N that are required may be rather generic in string theory compactifications. In this context, axions arise on dimensional reduction, from integrating p-form gauge potentials over p-cycles (see [59, 174] for a review). For example, in type IIB string theory one has axions b_i and c_i which arise, respectively, from integrating the Neveu-Schwarz (NS) and Ramond-Ramond (RR) 2-forms B_{MN} and C_{MN} over compact 2-cycles. Generically, instanton effects break the shift symmetry down to a subgroup $b_i \to b_i + (2\pi) f_i$, leading to periodic contributions to the effective potential (similarly for c_i).

The general framework described above shows how N-flation may arise within string theory: the low energy theory contains one axion for each independent cycle that B_{MN} can wrap and generic Calabi-Yau compactifications may contain exponentially large numbers of such cycles. There exist know examples with N as large as $\sim 10^5$ [80], however, to our knowledge there is no general theorem that prohibits finding Calabi-Yau manifolds with even larger values.

There may be many ways to realize N-flation in an explicit, stabilized string theory compactification. The first efforts in this direction were undertaken in [125] and further analyzed in [176]. See also [154] for an alternative construction.

In order to quantify inverse decay effects, we are most interested in the effective coupling between the collective field Φ and a given gauge field. The detailed derivation of such interactions is rather complicated and dependent on model building details (see [21,71,151] for a more detailed discussion). We will not attempt to estimate the effective value of α/f, but rather note that a generic spectrum is expected to contain some axions with $f_i \ll M_p$, hence strong inverse decay effects (and their associated non-Gaussianities) are at least plausible in N-flation.

5.4 Axion Monodromy Inflation

The axion monodromy model [231] is a string theoretic construction based on a single axion field. The key ingredient in this construction is a suitably wrapped brane which leads to a non-periodic contribution to the inflaton potential, explicitly breaking the shift symmetry. This monodromy in the moduli space allows the axion to develop a kinematically unbounded field range and accommodate the super-Planckian excursions required for large-field inflation. In the effective field theory description of this scenario, the inflaton potential has the form

$$V(\varphi) = \mu^3 \varphi + \Lambda^4 \cos\left(\frac{\varphi}{f}\right) \tag{185}$$

where the linear term arises from wrapped branes while the (subdominant) periodic modulation arises from instanton effects.

In [129] the decay constant for axion monodromy inflation was studied in detail. It was shown that microphysics bounds the allowed values as

$$0.06 \, \frac{g_s^{1/4}}{\mathcal{V}^{1/2}} < \frac{f}{M_p} < 0.9 \, g_s \tag{186}$$

where $g_s < 1$ is the string coupling and $\mathcal{V} \gg 1$ is the compactification volume in string units. This bound illustrates that $f \ll M_p$ in this model. If $\alpha = \mathcal{O}(1)$ can be realized in a consistent string compactification, then large non-Gaussianity is easily accommodated by axion monodromy inflation.

The small periodic modulation of the potential (185) is not important for determining the number of e-foldings of inflation, however, this term may nevertheless have an important role of the phenomenology of the model. In [129], and also in [161], it was shown that axion monodromy models can give rise to large resonant-type non-Gaussianities [88, 89, 130, 202]. Depending on the parameters, either resonant effects or inverse decay effects may dominate the bispectrum in the regime $f \ll M_p$. It would be interesting to study the combined observational impact of these two effects in future work.

5.5 Axion/4-Form Mixing

In [178] a realization of chaotic inflation was proposed which shares many features of the axion monodromy model discussed above. Here the axion φ mixes with a 4-form field strength through a term like $\varphi\, \epsilon^{\mu\nu\alpha\beta} F_{\mu\nu\alpha\beta}$. The theory also includes charged membranes which source a background for the conjugate momenta of the gauge field and break the axion shift symmetry, giving rise to a non-periodic potential $V(\varphi) \cong \frac{1}{2} m^2 \varphi^2$ which is robust against a wide variety of corrections [177]. There is no obstruction to realizing inflation with $f \ll M_p$ in this model, and hence we generically expect that inverse decay processes may play an important role.

As in the case of axion monodromy, the power-law potential will generically be modulated by subdominant oscillatory features arising from instanton effects. In [177] non-Gaussianities were discussed. In general resonant effects may operate in concert with inverse decay processes, and it would be interesting to study their combined impact.

5.6 Dante's Inferno

In [63] a model was proposed which consists of two axions, r and θ, with decay constants $f_r < f_\theta \ll M_p$. It is assumed that a linear combination of these receives a periodic potential from nonperturbative effects and, moreover, that some explicit shift symmetry breaking effect generates a non-periodic contribution $W(r)$. The effective potential therefore takes the form

$$V(r,\theta) = W(r) + \Lambda^4 \left[1 - \cos\left(\frac{r}{f_r} - \frac{\theta}{f_\theta} \right) \right] \tag{187}$$

For a power-law $W(r) = \mu^{4-p}\, r^p$ the dynamics of this model are well approximated by a single field with effective potential $V_{\rm eff} = (f_r/f_\theta)^p \mu^{4-p} \varphi_{\rm eff}^p$ and our analysis of inverse decay processes is directly applicable. The possibility of large non-Gaussianity follows immediately from $f_r < f_\theta \ll M_p$.

In [63] it was shown how to embed the model (187) within string theory as a modest extension of axion monodromy [129, 231]. It may be possible to realize such a potential from axion/4-form mixing, along the lines of [177, 178].

5.7 Multi-Field Scenarios

So far we have focused our attention on models which can, at least to first approximation, be described in terms of the dynamics of a single field (that may represent a collective excitation or be related non-trivially to the axions of the original theory). Such scenarios are appealing, because our analysis of inverse decay processes applies more-or-less without modification. However, there are also a number of interesting inflation models which involve dynamical axions but are *not* well described in terms of a single dynamical field. (As discussed in footnote 33, this scenario seems especially natural in supergravity models.) In multi-field models with $f \ll M_p$ strong inverse decay effects are clearly possible, however, the detailed phenomenology is more complicated than what we have presented. Nevertheless, we expect that large non-Gaussianity should be possible. It would be interesting to explore multi-field inverse decay processes in future works.

There are a variety of interesting multi-field models involving axions, for example race-track inflation [65, 66] and the axionic D3/D7 model [73]. Possibly, interesting models from the perspective of obtaining strong inverse decay effects, are based on the large volume compactification of [27, 28, 98]. For example, roulette inflation [69] – which generalizes [96] to incorporate the dynamics of the axionic partner of the Kahler modulus – is characterized by significant motion in the axion direction during the early stages of inflation. See [37, 97] for a discussion of axion/moduli couplings and decays in large volume inflationary scenarios. See also [236] for a related model.

6 Summary of Axion Inflation

Axion inflation is a compelling class of inflationary models, since they are characterized by a (broken) shift symmetry that protects the required flatness of the inflaton potential against quantum corrections. Axions are coupled to gauge fields through $\frac{\alpha}{f}\varphi F\tilde{F}$. More generally, from the perspective of effective field theory, a coupling $\varphi F\tilde{F}$ must be included whenever φ is pseudo-scalar. The dimensionless coupling α is a model-dependent quantity; however, from an effective field theoretical point of view, one does not expect it to be $\ll 1$. In this Part of the thesis, we studied the phenomenological signatures induced by this coupling, keeping α/f as a free parameter. The crucial point of our study is that, in presence of this coupling, the motion of φ induces a tachyonic growth for one polarization of the gauge field. This amplification is most important for modes with physical wavelength comparable to the horizon $\lambda \sim 1/H$. These produced gauge quanta inverse decay to produce inflaton fluctuations of a comparable wavelength, consistent with the general expectation from causality/locality. These inflaton modes then leave the horizon, and the corresponding density fluctuations, $\zeta_{\rm inv.decay}$, become frozen and contribute to the observable cosmological perturbations. An analogous process leads to production of gravitational waves. These scalar and tensor perturbations add up incoherently with those generated by the expansion of the universe (the standard "vacuum" modes, $\zeta_{\rm vac}$ and $h_{\rm vac}$).

We find that the amplitude of the perturbations generated by the inverse decay is an exponentially growing function of α/f. These modes dominate over the vacuum ones for $\alpha/f \gtrsim 10^{-2}\, M_p^{-1}$ (the precise value depending on the inflaton potential). Due to the exponential sensitivity on the coupling, there is only a small range in α/f for which both $\zeta_{\rm inv.decay}$, and $\zeta_{\rm vac}$ are relevant. For smaller values of α/f this new effect is completely negligible, and the standard results are recovered. For larger values, drastically new predictions are obtained. In particular, the main characteristic of $\zeta_{\rm inv.decay}$ is that they are highly non-Gaussian: this is due to the fact that two gauge quanta participate in the inverse decay, and that the initial distribution of these quanta is itself Gaussian (loosely speaking, $\zeta_{\rm inv.decay}$ behaves as the square of a Gaussian field, which is obviously not Gaussian). As a consequence of this effect, the non-Gaussianity parameter $f_{\rm NL}$ also grows exponentially with α/f in the region of parameter space for which both $\zeta_{\rm vac}$ and $\zeta_{\rm inv.decay}$ are comparable. When $\zeta_{\rm vac}$ can be neglected, $f_{\rm NL}$ saturates to about $8,600$ in the equilateral configuration, well beyond the WMAP limit $f_{\rm NL}^{\rm equil} < 266$ (95% CL, which was used for our analysis) and the current Planck limit $f_{\rm NL}^{\rm equil} < 33$ (68% CL). We found that these limits allow only values of α/f for which $\zeta_{\rm inv.decay}$ can contribute $\lesssim 10\%$ to the power spectrum (see Figure 7). In this regime, also the spectrum of gravitational waves produced by the inverse decay at CMB scales is much smaller than that from vacuum.

We have seen that non-Gaussianity from axion inflation is greatest for the equilateral configuration. This is related to causality/locality of the underlying inverse decay production mechanism. To understand this, intuitively, recall that a mode $\zeta_{\vec{k}}$ is sourced by the inverse decay of gauge perturbations of comparable wavelength. Consequently, the three-point correlation $\langle \zeta_{\vec{k}_1} \zeta_{\vec{k}_3} \zeta_{\vec{k}_3} \rangle$ is suppressed between modes of very different scale. We found that the shape of non-Gaussianity produced by this mechanism is very well reproduced by the equilateral template that has been widely studied in the literature. Quantitatively, the "cosine" between the two shapes (a measure of how much the two shapes coincide [26]) is $\cong 0.93$. Therefore, it is sensible to use the limits on $f_{\rm NL}^{\rm equil}$ (obtained with the use of the equilateral template) to probe the current mechanism. On the other hand, the specific shape that we have computed is nevertheless distinct from the equilateral template and this fact might be useful to distinguish axion inflation from other models (for instance, the shape we found has a significant overlapping also with the orthogonal template). This issue can be more thoroughly explored when / if a nonzero value for $f_{\rm NL}^{\rm equil}$ will be found in the data. The non-Gaussian signature of axion inflation, together with the requirement of standard results for the spectral tilt and the tensor-to-scalar ratio tensor of large field inflation, will allow to falsify the mechanism (or, at least, the simplest version that we have studied here) in the near future. This is due to the fact that large field inflationary model provide much larger – and detectable – values of $n_s - 1$ and of r than many other classes of models.

We stress that the non-Gaussianity which we have studied is very different from the so called resonant non-Gaussianity [88, 89], that has been discussed previously [129, 130, 161, 202] in the context of axion monodromy inflation [231]. Axion monodromy is one of the particle physics realizations of axion inflation; in this model the periodic potential typical of axions is only a subdominant term in the inflaton potential, and it provides a non-Gaussian modulation of the inflaton perturbations. Depending on model parameters, either this effect or inverse decay processes may dominate the bispectrum.

It is remarkable that, in the mechanism we have studied, the large non-Gaussianity is obtained in a rather minimal way (simply by considering a coupling of the pseudo-scalar allowed by the symmetries of the model, and therefore expected in an effective field theory context). Non-Gaussianity is a measure of the strength of interactions of the inflaton, which are typically constrained to be small by the requirement of a flat potential. To obtain observable non-Gaussianity, previous studies have invoked non-standard field theories (involving small sounds speed [90] or higher derivatives [36, 39, 41]), or initial conditions [90, 170, 233, 234], potentials with sharp features [88, 89], dissipative effects [150], fine-tuned inflationary trajectories [75, 77, 91, 258, 297] or post-inflationary effects (such as preheating [38, 40, 68]). The mechanism we have studies circumvents the common lore result in a very novel way: the interaction that gives rise to large non-Gaussianity, eqn. (78), does not play any role in the background dynamics and is thus unconstrained by the requirement of slow roll. At the same time, the effect of this interaction (namely, the amplification of gauge quanta) persists during the entire inflationary phase, leading to a (nearly) scale invariant signature (as opposed to sporadic episodes of particle productions [35, 42, 43] that would lead to a highly localized non-Gaussianity in momentum space).

At the scales relevant to CMB observations, the constraints on the axion-gauge coupling from non-Gaussianity is more stringent than those from gravitational waves, as mentioned already. Namely, once the non-Gaussianity limit is imposed, the standard vacuum mode of GWs dominates over the sourced contribution, and thus the GW signal would only be standard. However, the effective coupling strength ξ increases during inflation. Although the rate of this growth is slow-roll suppressed, the amount of particle production is exponentially sensitive to ξ. This implies that the production becomes significantly more effective and that the feed-back effects on cosmological perturbations are stronger for the modes that cross the horizon at later stages of inflation, corresponding to the scales much smaller than the CMB scales [47,99]. The growth of scalar perturbations at small scales could potentially lead to overproduction of primordial black holes [208]; however, this effect inherently carries a large uncertainty. On the other hand, the feed-back on GWs does not contain such uncertainties, and as discussed in Section IV 4, the future-generation terrestrial GW detectors can put a strong constraint on the coupling strength ξ of the axion inflation models.

The axion-gauge coupling induces a parity-violating interaction, which is encoded in the polarization feature of GWs. Such polarized GWs can, if present, be detectable at the terrestrial GW detector networks. Since a polarized isotropic signal may be a unique way to discriminate a stochastic GW background (SGWB) of cosmological origin from one of astrophysical origin, a detection of such a signal would open up the new observational possibility to probe the physics in the very early universe. For a definite detection of parity violation, it is critical for the detector network to include multiple non-collocated detector pairs, as this is the only way to definitively separate parity violation effects from frequency dependence in the SGWB. Applying the search formalism for polarized GWs to the axion inflation signals, we have quantitatively shown that the current limits from the SGWB are weaker than the current CMB limits, but that future GW detectors can improve over them.

Finally, it is remarkable that current observational limit on non-Gaussianity *already* place a surprisingly stringent bound on pseudo-scalar interactions of the form (78), and the forthcoming terrestrial GW detectors can even improve it. As a comparison, the bound on the coupling of the much lighter QCD (or QCD-like) axion to photons is $\frac{\alpha_{\rm photons}}{f} \lesssim \mathcal{O}\left(10^{-11}\right)$ GeV^{-1}, from energy loss in stars [241]. On the other hand, the mechanism that we have studied provides the bound $\alpha/f \lesssim \mathcal{O}\left(10^{-16}\right)$ GeV^{-1} on the coupling of the pseudo-scalar inflaton with *any* gauge field. This provides a unique window for constraining – or perhaps probing – a large class of inflationary models. As we have reviewed in Section IV 5, there are a variety of interesting multi-field models involving axions, many of which can be realized in string theory. While most of these studies provide the value of axion decay constant f

and of the potential $V(\varphi)$, comparatively less attention has been devoted to compute the dimensionless coupling α.

Part V

Non-Gaussianity from Dilaton-like Coupling during Inflation

Focusing on inflaton interactions with gauge fields, for a singlet inflaton φ there are two very natural classes of gauge field interactions to consider, depending on the parity. In the previous Part, we studied the case where the inflaton is a pseudo scalar field, which has a natural coupling to gauge fields through $\mathcal{L}_{\rm int}^{\rm pseudo} \propto \varphi F_{\mu\nu}\tilde{F}^{\mu\nu}$, where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the field strength tensor associated with some U(1) gauge field A_μ and $\tilde{F}^{\mu\nu}$ is its dual. In this Part, on the other hand, we study the case of a scalar inflaton. In this case, one may expect couplings of the form

$$\mathcal{L}_{\rm int}^{\rm scalar} = -\frac{I^2(\varphi)}{4} F_{\mu\nu}F^{\mu\nu} \,, \tag{188}$$

where $I(\varphi)$ plays the role of a field dependent gauge coupling. The interaction (188) is typical of moduli or dilaton-like fields in string theory and supergravity frameworks. As noted in [128], coupling the inflaton to a gauge field may actually be the only way to reheat for some of these models.

In this Part of the thesis, we show that rich phenomenology is possible in simple models with a single *scalar* inflaton in slow roll, via the coupling (188). The underlying mechanism is quite novel. The time dependence of the inflaton condensate $\varphi^{(0)}(t) = \langle \varphi(t,\vec{x}) \rangle$ breaks the conformal invariance of the gauge field sector and leads to amplification of the quantum fluctuations of A_μ, similarly to the well-known mechanism that produces scale invariant curvature fluctuations during inflation. For simplicity, we focus our attention on the case where large-scale fluctuations of the gauge field are produced during inflation with a scale-invariant "magnetic" component. We notice that the *same* coupling that leads to production of gauge field fluctuations also implies that these produced fluctuations must, in turn, couple to the cosmological perturbations of the inflaton, $\delta\varphi(t,\vec{x}) = \varphi(t,\vec{x}) - \varphi^{(0)}(t)$. We find that the feedback of produced gauge fluctuations on $\delta\varphi$ contributes a new component to the observable curvature fluctuations that is highly non-Gaussian and is uncorrelated with the usual spectrum from quantum vacuum fluctuations. As we later show the details, for reasonable parameters we obtain nearly local-type non-Gaussianity with shape

$$\langle \zeta_{\vec{k}_1} \, \zeta_{\vec{k}_2} \, \zeta_{\vec{k}_3} \rangle \propto \frac{1 + \cos^2\left(\vec{k}_1, \, \vec{k}_2\right)}{k_1^3 \, k_2^3} + \text{permutations} \tag{189}$$

and at the level $f_{\rm NL} = \mathcal{O}(10 - 100)$. Such values are close to current observational limits. The shape (189) has a strong overlapping with the local template (40) in the squeezed limit, where both shapes are enhanced. However, in this limit, (189) has a quadrupolar dependence on the angle between the shorter side and either of the longer sides. Contrary to what typically happens for non-Gaussianity sourced by scalar fields [204], this angular dependence does not vanish in the squeezed limit, and it contributes to about $1/5$ of the amplitude of (189). This therefore appears as a signature of non-Gaussianity from higher spin fields, and it may be an important distinguishing feature when the model is confronted with observations.

As pointed out in [53], the results obtained in the present study should be interpreted as the theoretical expectation values, averaged over several realizations of the mechanism. The novel contribution to the curvature perturbation is the feedback of the gauge quanta produced through the interaction (188). In our setup, these quanta result in a nearly scale-invariant spectrum for the "magnetic" perturbations.[35] Once modes leave the horizon, they add up in a "random walk" manner to form a classical background [53]. Let us denote by $N_{\rm tot}$ the total number of e-folds of inflation and by N some given number of e-folds. The modes that were produced and left the horizon in the first $N_{\rm tot} - N$ e-folds of inflation then add up as a classical field and are experienced as a homogeneous background by the modes of smaller size, which leave the horizon in the last N e-folds of inflation. A homogeneous vector background points in one direction in a given frame that is determined by the random addition of the super-horizon mode, breaking isotropy. Ref. [53] shows that this vector encodes a strong anisotropy in both the power spectrum and bispectrum of curvature perturbations in a single realization of the model (188) and that the isotropic results obtained in the present study are indeed the values averaged over many realizations of those anisotropic signals. We discuss this issue in Section V 4.

Models of vector fields producing non-Gaussianity were also proposed in [50, 51, 109, 110, 112–115, 148, 183, 184, 224, 300]. In these models the vector field is amplified either from the time dependence of its kinetic term, as we consider here, or from a non-minimal coupling

[35] The interaction (188) enjoys an "electric" $\leftrightarrow$ "magnetic" duality under $I \leftrightarrow \frac{1}{2}$. All of our discussion for the scalar and tensor perturbations in this Part still holds if we simultaneously change $I \to \frac{1}{I}$ and "magnetic" $\leftrightarrow$ "electric."

(as an effective mass term) to the Ricci scalar R. This second case, however, introduces a longitudinal vector component, which turns out to be a ghost [164, 165]. (In some of these models, the energy in the vector field sources anisotropic inflation; we do not study this possibility here, but we refer to Ref. [283] for a review.) In these models, the vector acts as a curvaton, and it contributes to the spectrum and the bispectrum through its energy density. We compute instead the non-Gaussianity resulting from the *same* coupling that leads to production of gauge field fluctuations, and which remains imprinted in the inflaton perturbations even when the energy density in the gauge field is very negligible at reheating. The model of ref. [314] studied non-Gaussianity from a gauge field amplified analogously to what we do here, but imprinted through the waterfall field of hybrid inflaton [212].

It is interesting to contrast the mechanism considered here to the closely related physics of inverse decay that was studied in Part IV. For axial couplings in Part IV, the relevant production of inflaton perturbations arises near horizon crossing, and then the bispectrum is instead nearly equilateral. In the case at hand, on the other hand, the feed-back of the produced gauge field fluctuations on the inflaton fluctuations leads to a super-horizon growth of the curvature perturbation. Such growth is consistent, since the produced gauge field fluctuations provide a source of large scale isocurvature perturbations. Consequently, we find a bispectrum which is very close to the local shape.

Models of the type which we study have received considerable attention in the literature, in connection with primordial magnetogenesis [29, 79, 107, 138, 182, 229, 255, 268, 290, 307]. At face value, our choice of $I^2(\varphi)$ can produce large scale magnetic fields with a sufficient amplitude to account for observations at galaxy and cluster scales; see [144,149,181,261,308] for reviews. This would open the interesting possibility of correlating the magnetic field with the primordial perturbations [79], although the correlation would only involve the component of metric perturbations that are sourced by the vector field, and that is typically subdominant with respect to the vacuum part. Moreover, the magnetic field would induce non-Gaussianiy from its direct coupling to the CMB photons [72, 78, 84, 274, 278, 293]. This effect can be observed for a magnetic field at the $\sim$ nG $-$ 10 nG level, while the non-Gaussianity we have obtained arises from the direct coupling to the inflaton that generated the gauge field, and can be observable even if the current "magnetic" field is significantly smaller. A list of works that study the general signatures of a magnetic field on the CMB can be found in the review [308]. In particular, "magnetic" fields continue to source scalar and tensor perturbations until neutrino decoupling. Refs. [277, 294] shows that this effect constrains the "magnetic" field to be $\lesssim$ few nG at present. We show in Subsection V 1.3 that the "magnetic field" generated in the cases of our interest is of $\mathcal{O}\left(10^{-10}\right)$ nG or less.

One major problem with identifying the gauge field with the electromagnetic one is due to the fact that a scale invariant magnetic field can only be obtained if the effective gauge coupling, $g(t) = I\left[\varphi^{(0)}(t)\right]^{-1}$, decreases by *many* orders of magnitude during inflation (specifically, by a factor of $\mathrm{e}^{2N_{\rm tot}}$, where $N_{\rm tot}$ is the number of e-folds of inflation). If one starts from $g_{\rm in} = \mathcal{O}(1)$, then the gauge coupling at the end of inflation will be *much* too small to identify A_μ with the Standard Model (SM) photon. Normalizing instead the coupling constant to be the electromagnetic one at the end of inflation would instead entail an unacceptable breakdown of perturbation theory. This problem was stressed in [107], and it is a serious obstacle in identifying the gauge field of the mechanism with the photon. In the following, we discuss some unsuccessful attempts of solving this problem. We cannot of course rule out that a solution of the problem can be found, but we believe that it would require a substantial modification of the model.

This Part is mostly based on the work in [46]. We organize the following of this Part as follows. In Section V 1 we introduce our model and compute the production of gauge field fluctuations. We then discuss the challenging connection to magnetogenesis. In Section V 2 we present a detailed computation of the 2-point and 3-point correlation functions of the scalar perturbations. In Section V 3 we present the computation of the 2-point correlation function of gravitational wave fluctuations. In Section V 4 we give a succinct review of the phenomenological predictions from the model. Finally, in Section V 5, we summarize the results of the model.

1 Gauge Field Production

1.1 The Model

We consider a simple model with a dilaton-like coupling of the inflaton to a U(1) gauge field A_μ

$$S = \int d^4x\sqrt{-g}\left[\frac{M_p^2}{2}R - \frac{1}{2}\,\partial_\mu\varphi\,\partial^\mu\varphi - V\left(\varphi\right) - \frac{I^2\left(\varphi\right)}{4}F_{\mu\nu}F^{\mu\nu}\right] \tag{190}$$

In this action, $F_{\mu\nu}$ is the field strength of A_μ, and I and V are functions of the inflaton φ. We assume that the potential $V(\varphi)$ is sufficiently flat to support a long phase of quasi de Sitter expansion. As usual, we require that $\epsilon, |\eta| \ll 1$ where the slow roll parameters are (24)

$$\epsilon \equiv \frac{M_p^2}{2}\left(\frac{V_\varphi}{V}\right)^2 \qquad \eta \equiv M_p^2\frac{V_{\varphi\varphi}}{V} \tag{191}$$

To simplify the calculation, we choose $I(\varphi)$ and $V(\varphi)$ to be related to each other in such a way as to obtain the solution $I \propto a^n$, where a is the scale factor of the Universe, during inflation. The required relation can be obtained [303] by taking the ratio of the slow roll relations (27, 28)

$$H^2 \simeq \frac{1}{3M_p^2} V \ , \quad 3\, H\dot{\varphi} \simeq -V_\varphi \tag{192}$$

(assuming that the interaction with the gauge field provides negligible corrections to the dynamics, see below). In this way, one forms a differential equation for $da/d\varphi$, that is integrated into $a \propto \exp\left[-\int \frac{V d\varphi}{V_\varphi M_p^2}\right]$. The functional form of $I\left(\varphi\right)$ can be then set to the n-th power of the right hand side expression. For definiteness, we consider a monomial inflaton potential

$$V = \mu^{4-r} \varphi^r \ \ , \ \ I = I_{\rm end} \exp\left[-\frac{n\varphi^2}{2rM_p^2}\right] \tag{193}$$

although other choices are clearly possible. Here $I_{\rm end}$ is the value of the coupling function at the end of inflation, when the inflaton is in the vacuum $\varphi = 0$. We therefore assume that, after inflation, the function I sets to a constant. We could then normalize $I_{\rm end} = 1$.

We found it algebraically convenient to define "electric" and "magnetic" components of the gauge field as

$$E_i \equiv -\frac{\langle I \rangle}{a^2} A_i' \ \ , \ \ B_i \equiv \frac{\langle I \rangle}{a^2} \epsilon_{ijk} \partial_j A_k \tag{194}$$

where here and in the remainder of this Part the Coulomb gauge $A_0 = 0$ is assumed. We do not necessarily assume A_μ is the Standard Model photon (more on this later); however, we will sometimes use the language "electric field" and "magnetic field", by analogy with standard electromagnetism.

As the gauge field has no classical expectation value, its perturbations do not couple to that of the inflaton or of the geometry at linearized order. We can therefore solve for these perturbations by assuming a FRW background, and by treating I as a classical function. Therefore, we can simply set $I \propto a^n$ for the remainder of this Section. The time component of the vector equation of motion is solved by $\partial_i A_i = 0$, and we can decompose the vector potential as in (48)

$$\vec{A} = \sum_{\lambda=\pm} \int \frac{d^3k}{\left(2\pi\right)^{3/2}} \vec{\epsilon}^{\,(\lambda)} \left(\vec{k}\right) {\rm e}^{i\vec{k}\cdot\vec{x}} \left[a_\lambda \left(\vec{k}\right) \, A_\lambda \left(k\right) + a_\lambda^\dagger \left(-\vec{k}\right) \, A_\lambda^* \left(k\right)\right] \tag{195}$$

where $\vec{\epsilon}^{\,(\pm)}$ are the circular polarization operators, and a_λ and $a_\lambda^\dagger$ are the annihilation/creation operators of the gauge field (see the sentences after (48) for their properties and normalization).

The vector mode functions then satisfy

$$V_\lambda'' + \left(k^2 - \frac{I''}{I}\right) V_\lambda = 0 \;, \quad V_\lambda \equiv I\, A_\lambda \tag{196}$$

where prime denotes derivative with respect to conformal time τ. For a constant I, one recovers the typical Minkowski $\mathrm{e}^{-ik\tau}/\sqrt{2k}$ solution for the mode functions, due to the fact that the gauge field is conformally coupled to the FRW metric. For our purposes, it is sufficient to obtain the leading expression of the mode functions in a slow roll expansion. Namely, for the de Sitter limit $a = -\frac{1}{H\tau}$, the mode solution that, up to an arbitrary constant phase, is normalized to $\mathrm{e}^{-ik\tau}/\sqrt{2k}$ (the so called adiabatic vacuum) in the asymptotic early time / high momentum regime is

$$V_\lambda = i\frac{\sqrt{\pi}}{2}\,\sqrt{-\tau}\,H^{(1)}_{n+1/2}\left(-k\tau\right) \tag{197}$$

where $H^{(1)}_\nu$ is the Hankel function of the first kind. This solution has been discussed at length in [107] for all values of n, and it is unnecessary to review all their study here. We only discuss the $n = 2$ case, which results in a scale invariant "magnetic" field. Interesting non-Gaussian properties of the primordial perturbations may be possible also for other values of n. The arbitrary phase in (197) has been chosen so that the function V_λ is real and positive in the super horizon limit.

In this case, the "electric" and "magnetic" gauge field operators during inflation reduce to

$$\begin{aligned} \vec{E} &= \int \tfrac{d^3k}{(2\pi)^{3/2}}\,\mathrm{e}^{i\vec{k}\cdot\vec{x}}\,\vec{E}_{\vec{k}} \;\;,\;\; \vec{E}_{\vec{k}} = -\tfrac{H^2\tau}{\sqrt{2}}\sum_\lambda \tfrac{1}{k^{1/2}}\vec{\epsilon}^{\,(\lambda)}\left(\vec{k}\right)\left[a_\lambda\left(\vec{k}\right) + a_\lambda^\dagger\left(-\vec{k}\right)\right] \\ \vec{B} &= \int \tfrac{d^3k}{(2\pi)^{3/2}}\,\mathrm{e}^{i\vec{k}\cdot\vec{x}}\,\vec{B}_{\vec{k}} \;\;,\;\; \vec{B}_{\vec{k}} = \tfrac{3H^2}{\sqrt{2}}\sum_\lambda \lambda\tfrac{1}{k^{3/2}}\vec{\epsilon}^{\,(\lambda)}\left(\vec{k}\right)\left[a_\lambda\left(\vec{k}\right) + a_\lambda^\dagger\left(-\vec{k}\right)\right] \end{aligned} \tag{198}$$

in the super-horizon limit $-k\tau \ll 1$. As for the standard scalar case, the mode function does not oscillate in the super-horizon regime, which is a signal that the field has become classical [253]. The energy densities are given by

$$\rho_E = \frac{\langle\vec{E}^2\rangle}{2} \simeq \frac{H^4\,\tau^2}{4\pi^2}\int dk\,k\left(1 + k^2\tau^2\right) \;\;,\;\; \rho_B = \frac{\langle\vec{B}^2\rangle}{2} \simeq \frac{9\,H^4}{4\pi^2}\int\frac{dk}{k} \tag{199}$$

We must only compute the energy in the classical fields, and we therefore limit the integrals to momenta that have exited the horizon during inflation. Modes of smaller wavelength remain in their vacuum state and their contribution to the energy density must be renormalized away (this is part of the cosmological constant problem). The smallest momentum $k_{\rm min} \simeq \frac{1}{-\tau_{\rm in}}$ corresponds to modes that have exited the horizon at the start of inflation (we stress that we are computing theoretical expectations under a constrained and finite value for the total number of e-foldings of inflation). For any moment τ, the largest momentum $k_{\rm max} \simeq \frac{1}{-\tau}$ corresponds to modes that have just exited the horizon at that moment. Therefore, for any $\tau \gg \tau_{\rm in}$ during inflation,

$$\rho_E \simeq \frac{3\,H^4}{16\pi^2} \quad , \quad \rho_B \simeq \frac{9\,H^4}{4\pi^2} \ln \frac{a\,(\tau)}{a\,(\tau_{\rm in})} \tag{200}$$

These behaviors are very different from the usual behavior $\rho \propto a^{-4}$ for radiation. This shows that energy is being transferred from the inflaton to the gauge field through the I^2F^2 coupling. We also note that the energy in the "magnetic" component is greater than that in the "electric" one. The energy in the "magnetic" component is scale invariant, and the logarithmic increase in the final result is due to the increase of the phase space of the modes that have become classical. We also note, that, despite for most of the super-horizon modes the density $\frac{d\rho_E}{dk}$ is several orders of magnitude smaller than $\frac{d\rho_B}{dk}$, and decreases with time, the total value of ρ_E is "only" logarithmically suppressed with respect to that of ρ_B. This is due to the fact that the integral for ρ_E has most of its support in the UV region, where the "electric" and "magnetic" energy densities are not too different from each other.

In passing, notice that our choice to produce scale-invariant "magnetic" fields – as opposed to "electric" fields – is essentially arbitrary from the perspective of primordial non-Gaussianity. Indeed, there is an electric/magnetic duality that leaves the Maxwell equations invariant under the replacement $\vec{E} \to -\vec{B}$, $\vec{B} \to \vec{E}$ and $I \to 1/I$; see [145] for more discussion. In this case at hand, this means that we can interchange the "electric" and "magnetic" spectra simply by taking $n = -2$ rather than $n = 2$. The feed-back of these produced fluctuations on the scalar inflaton is essentially unchanged under such a replacement.

1.2 Backreaction Bounds

Throughout the discussion above, we have assumed that the produced gauge field fluctuations have a negligible effect on the homogeneous background dynamics during inflation. To ensure that this assumption is consistent, we must verify several backreaction constraints.

First, we require that the energy density in the "magnetic" field is much smaller than the potential energy driving inflation. Hence, we require

$$\frac{\rho_B}{V} \simeq \frac{3}{4\pi^2}\frac{H^2}{M_p^2}\ln\frac{a(\tau)}{a(\tau_{\rm in})} \ll 1 \tag{201}$$

A second constraint arises because the produced gauge field fluctuations modify the homogeneous Klein-Gordon equation for the inflaton condensate $\varphi^{(0)}(t)$. To ensure that the usual slow roll equations (192) are reliable, we require that the "driving force" in the inflaton equation of motion is dominated by the derivative of the inflaton potential. That is, we require:

$$|V_\varphi| \gg \left|\frac{I_\varphi}{I}\langle B^2\rangle\right| \tag{202}$$

Finally, we note that the backreaction of produced gauge fields can also lead to a correction for the effective mass of the inflaton. The easiest way to see this effect is to note that the action (190) contains a term of the form $(I^2)_{\varphi\varphi}(\delta\varphi)^2\langle F^2\rangle$ once we expand $\varphi = \varphi^{(0)} + \delta\varphi$ and replace F^2 with its vacuum average, to estimate the magnitude of backreaction effects. We require that this "new" correction to the inflaton mass is much smaller than the Hubble scale so that we do not spoil the scale invariance of the spectrum. This amounts to a constraint

$$\frac{1}{\epsilon M_p^2}\langle B^2\rangle \ll H^2 \tag{203}$$

The conditions (202) and (203) are more stringent than (201). In both cases, we obtain a constraint on the total number of e-foldings of inflation of the form

$$N_{\rm tot} \ll 10^{-1}\mathcal{P}^{-1} \tag{204}$$

where $N_{\rm tot} = \ln(a_{\rm end}/a_{\rm in})$ and we have defined

$$\mathcal{P} \equiv \left(\frac{H^2}{2\pi\dot{\varphi}^{(0)}}\right)^2 \simeq \frac{H^2}{8\pi^2 M_p^2\epsilon} \simeq 2.5\cdot 10^{-9} \tag{205}$$

which gives the amplitude of the power spectrum from the usual vacuum fluctuations [58]. (In this model there is also an additional contribution to the power spectrum coming from second order effects; however, we will always work in a regime where this is subdominant. More on this later.)

The condition (204) then indicates that backreaction is negligible provided that $N_{\rm tot} \ll$

$\mathrm{O}\left(10^{7}\right)$. Note that $N_{\rm tot}$ enters in this condition because the "magnetic" field energy grows (proportionally to the number of e-folds) during the *whole* duration of inflation. We stress that $N_{\rm tot}$ may be much greater than the number of e-folds $N_{\rm CMB} \simeq 50-60$ which separates the moment at which the largest CMB scales exited the horizon to the end of inflation.

1.3 Connection with Magnetogenesis

Magnetic fields have been observed at many scales. They are present in structures (e.g. galaxies, galaxy clusters and high redshift protogalactic structures) with strength $\sim 10^{-6} - 10^{-3}\,\mathrm{G}$, and in the low density intergalactic space with strength $\sim 10^{-14} - 10^{-17}\,\mathrm{G}$. See [144, 149, 181, 308] for reviews. The origin of these fields is not well understood. Although an astrophysical mechanism is not ruled out, the observed homogeneity and large coherence length ($\sim$ kpc $-$ Mpc) could suggest a primordial origin.

For a standard electromagnetic action, the photon is conformally coupled to a FRW geometry; loosely speaking, the scale factor drops from the action term $\sqrt{-g}F^2$, and the photon remains in its vacuum state. Mechanisms for generation of magnetic field during inflation break the conformal invariance by introducing some extra-term. For instance, in [295] couplings to the curvature invariants of the type RA^2 and $R_{\mu\nu}A^\mu A^\nu$ were considered. These couplings however break the U(1) invariance associated to the electromagnetic field, and one should worry about the longitudinal photon component that they introduce. It was shown in [164, 165] that, with the R^2A^2 coupling introduced in [295], the longitudinal photon is a ghost. It is safer to consider models that preserve the U(1) invariance. Axial couplings $\frac{1}{f}\varphi F\tilde{F}$ have been considered in [21, 76, 123]. In such models it is typically difficult to produce a sufficiently large field. Note that any attempt to generate primordial magnetic fields via an axial coupling must take into account the limits on f due to non-Gaussianity from inverse decay effects, discussed in Part IV in this thesis and in [45, 47, 48], which are much more stringent than backreaction bounds.

The model (190) has been studied in connection with primordial magnetogenesis [79, 107, 138, 229, 255, 268, 290]. It is indeed tempting to identify A_μ with the standard model photon. If we do so, and we assume that the electromagnetic energy density scales in the standard way $\rho \propto a^{-4}$ from the end of inflation on, we find

$$\left(\frac{d\rho_B}{d\ln k}\right)^{1/2}_{\rm today} \simeq 10^{-15}\,\mathrm{G}\left(\frac{H}{10^{15}\,\mathrm{GeV}}\right)^{2/3}\left(\frac{T_{\rm rh}}{10^{9}\,\mathrm{GeV}}\right)^{2/3} \tag{206}$$

where we have assumed matter domination due to the coherent inflaton oscillations until reheating takes place at the temperature $T_{\rm rh}$. We have also disregarded the current departure from matter domination (this gives a negligible correction to the estimate), and treated the value of H as constant during inflation. We note that a lower reheating temperature results in a smaller value of the magnetic field today [107]; the estimate obtained in [79] assumes that radiation domination starts immediately after inflaton. In this case the expression (206) evaluates to $10^{-10}\,{\rm G}\left(H/10^{15}\,{\rm GeV}\right)$.

The problem arising in associating the field A_μ with the electromagnetic photon has already been stressed in [107]. The model (190) must be supplemented by the action for the matter fields. The most minimal approach is to assume that $I^2\left(\varphi\right)$ only enters in the F^2 term, so that the relevant terms for the electromagnetic coupling of the (Standard Model) fermions are

$$\mathcal{L}_{\rm matter} = -\frac{I^2(\varphi)}{4}F^2 - \bar{\psi}\gamma^\mu(\partial_\mu + ie\,A_\mu)\psi \tag{207}$$

If this is the case, the "instantaneous" electric coupling constant is

$$e_{\rm physical} \equiv e\,I^{-1}\left[\varphi^{(0)}(t)\right] \tag{208}$$

During inflation we have $I \propto a^n$ and, consequently, for $n > 0$, the electric coupling constant decreases by a huge factor during inflation (we recall that the scale invariant B field is obtained for $n = 2$). Thus, if we take $e_{\rm in} \lesssim \mathcal{O}(1)$ at the start of inflation, then the gauge coupling at the end of inflation will be *extremely* tiny. Assuming no further evolution of $e_{\rm physical}\,(t)$ in the post-inflationary epoch, we clearly cannot identify A_μ with our photon. Alternatively, if we normalize I such that $e_{\rm physical}$ after inflation coincides with the present value, we necessarily imply that $e_{\rm physical}$ was extremely large all throughout inflation, apart from the very last stages. Even if there were no real charged particles during inflation, this would lead to strong quantum effects from the vacuum fluctuations of these fields, and to a quantum theory (at the very least) out of computational control. This poses serious questions on any result obtained from the model. We stress that this problem is not present if A_μ is a hidden sector gauge field, since in this case one may assume that its associated physical coupling constant is $\lesssim \mathcal{O}(1)$ at the start of inflation.

We briefly comment on a few (unsuccessful) attempts to solve this problem. Firstly, we note that moving the function $I^2\left(\varphi\right)$ outside the entire electromagnetic-sector Lagrangian does not affect this issue. Indeed, multiplying the second term of (207) by any factor $\tilde{I}$ affects both the fermionic kinetic term $\bar{\psi}\partial_\mu\psi$ and the vertex $\bar{\psi}A_\mu\psi$; however, the fermionic field enters quadratically in both expressions. After canonical normalization of the fermionic field, the factor $\tilde{I}$ drops out from the physical value of the electric coupling constant. It is also difficult to imagine how one may try to modify the structure of the covariant derivative without spoiling gauge invariance.

Secondly, one may try to arrange for a time evolution of I during reheating in such a way that $e_{\rm physical}$ is brought from a very tiny value at the end of inflation (so to avoid the strong coupling problem during inflation) to the present value before the onset of Big-Bang Nucleosynthesis (given that only a fractional discrepancy from the current value can be tolerated then [95, 103]). We stress that this requires a huge change of I, which can be difficult to accomplish without disrupting the result for the magnetic field achieved during inflation. The comoving energy densities

$$\begin{aligned} \bar{\rho}_B &= \frac{1}{2\pi^2}\int dk\, k^4 V^2 \equiv \int dk\, \bar{\rho}_{Bk} \\ \bar{\rho}_E &= \frac{1}{2\pi^2}\int dk\, k^2 \left(V' - \frac{I'}{I}\, V\right)^2 \equiv \int dk\, \bar{\rho}_{Ek} \end{aligned} \tag{209}$$

need to satisfy

$$\frac{d}{d\tau}\left(\bar{\rho}_{Ek} + \bar{\rho}_{Bk}\right) = -2\frac{I'}{I}\left(\bar{\rho}_{Ek} - \bar{\rho}_{Bk}\right) \tag{210}$$

If the electric component in this expression can be neglected, one finds $\bar{\rho}_{Bk} \propto I^2$; alternatively, if the magnetic component can be neglected, one finds $\bar{\rho}_{Ek} \propto I^{-2}$. In general, achieving such a large change in I during reheating does not appear feasible.

Thirdly, one may abandon the idea of identifying A_μ with the electromagnetic field, but still generate a large scale value of a hidden sector and weakly coupled A_μ, and then try to convert it to an electromagnetic field through some coupling. For instance, gauge invariance allows for

$$\Delta\mathcal{L} = \frac{\chi}{2} F^{\mu\nu}\, \mathcal{F}^{\rm em}_{\mu\nu} \quad \text{or} \quad \Delta\mathcal{L} = \frac{\chi}{2}\epsilon^{\mu\nu\alpha\beta} F_{\mu\nu}\, \mathcal{F}^{\rm em}_{\alpha\beta} \tag{211}$$

The first coupling was originally proposed in [168], and one can promote χ from a constant parameter to the expectation value of a scalar field; in the second case, χ is a pseudo-scalar function. One could imagine that χ experiences a quick transition from zero to a nonvanishing value χ_* at some given time τ_* after inflation, when I has set to one (we model the transition with a step function; clearly this approximation will break at very small scales). Solving the equations of motion in vacuum at leading order in χ_*, and requiring continuity of the vector potentials at τ_*, one finds that the turning on of χ results in a partial conversion of the "electric" or of the "magnetic" component of A_μ into our electric field: $\vec{\mathcal{E}}^{\rm em} \simeq \chi_*\, \vec{E}$ in the first case in (211), and $\vec{\mathcal{E}}^{\rm em} \simeq \chi_*\, \vec{B}$ in the second case. This solution is obtained in absence of any charged particle. However, as soon as a plasma is formed, it shuts off the electric field well before it can convert into a magnetic field. If already present at τ_*, the plasma would prevent any electric field generation at all.

In conclusion, none of these attempts appears to provide a solution to the strong coupling problem. We believe that a solution, if at all possible, will require a more radical modification of the model than those mentioned here.

2 Scalar Perturbations

To encode the effect of the gauge field on the cosmological perturbations we need to study the perturbations up to second order. Therefore, we decompose

$$\varphi = \varphi_0 + \delta_1\varphi + \delta_2\varphi \quad , \quad g_{\mu\nu} = g_{\mu\nu\,0} + \delta_1 g_{\mu\nu} + \delta_2 g_{\mu\nu} \quad , \quad A_\mu \tag{212}$$

The gauge field has no zero order part, and, as we discuss below, we do not need to evaluate it at second order.

Since the gauge field has no zero order part, the metric/inflaton perturbations do not mix with the gauge field modes at linear order; this is because the gauge field enters already quadratically in the action for the perturbations, through the expansion of the last term in (190). This is the same reason that in the previous Section allowed us to compute A_μ disregarding inflaton and metric perturbations. We note that the gauge field can still affect the first order metric/inflaton perturbations through its backreaction on the background evolution. This can be disregarded under the assumption that the two conditions (201) and (204) hold.

Therefore, at the linearized level the standard results of scalar field inflation hold. We work in the spatially flat gauge for the scalar perturbations, $\delta g_{ij,\mathrm{scalar}} = 0$. In this gauge, the curvature perturbation on uniform density hypersurfaces is $\zeta = -\frac{H}{\dot{\varphi}^{(0)}}\,\delta\varphi$. As we show in Appendix A, one finds

$$\left[\frac{\partial^2}{\partial\tau^2} + 2\mathcal{H}\frac{\partial}{\partial\tau} - \nabla^2 + \left(a^2 V_{\varphi\varphi} - 3\frac{\varphi^{(0)\prime\,2}}{M_p^2}\right)\right]\delta_1\varphi = 0\,, \tag{213}$$

where $\nabla^2 \equiv \partial_i\partial_i$. This is the standard equation for the Mukhanov-Sasaki [240,264] variable rewritten in terms of ζ; the last term in this equation has been simplified using slow roll approximation. Finally, $\mathcal{H} = \frac{a'}{a} = aH$. Here we have disregarded corrections to the effective mass due to backreaction effects; see Section V 1. In the end we will only perform computations at leading order in slow roll parameters, hence this neglect has no impact on our final results.

To compute the effect of the gauge fields on the cosmological perturbations we expand all the equations of the system at second order in the perturbations. We combine these equations in the same formal way that they are combined to obtain (213). In this way we obtain a "master equation" for $\delta_2\varphi$ that does not contain any $\delta_2 g_{\mu\nu}$ mode. As we show in Appendix A.2, this equation reads

$$\left[\frac{\partial^2}{\partial\tau^2}+2\mathcal{H}\frac{\partial}{\partial\tau}-\nabla^2+\left(a^2V_{\varphi\varphi}-3\frac{\varphi^{(0)\prime\,2}}{M_p^2}\right)\right]\delta_2\varphi=J_1\left[A^2\right]+J_2\left[\left(\delta_1\varphi\right)^2,\,\left(\delta_1 g\right)^2,\,\delta_1\varphi\times\delta_1 g\right] \tag{214}$$

where

$$J_1\left[A^2\right]\equiv a^2\,\frac{I_\varphi}{I}\left(\vec{E}^2-\vec{B}^2\right)-\frac{a^2\varphi^{(0)\prime}}{2M_p^2\mathcal{H}}\left[\frac{\vec{E}^2+\vec{B}^2}{2}+\frac{1}{a^4}\nabla^{-2}\partial_\tau\left(a^4\vec{\nabla}\cdot(\vec{E}\times\vec{B})\right)\right] \tag{215}$$

As for the linear theory equation, we have disregarded a backreaction-induced correction to the effective mass.

Note that our model (190) contains higher dimension interactions between the inflaton and gauge fields of the form $\sum_n c_n(\delta\varphi)^n F^2$ which arise from expanding the coupling function $I^2(\varphi)$ in powers of $\delta\varphi$. These couplings will enter into the calculation explicitly at higher order in perturbation theory. Using the in-in formalism, we have verified that such high dimension operators do not modify our leading order results for the spectrum and bispectrum.

The right hand side of (214) comprises of two sources for $\delta_2\varphi$; the first source contains terms at second order in the gauge perturbations, and has been completely given in (215). The second source contains terms that are the product of two first order inflaton perturbations, or of two first order metric perturbations (not only the scalar ones), or of one first order inflaton perturbations times one first order metric perturbation. We note that no "mix source" of the type $\delta_1\varphi\times A$ or of the type $\delta_1 g_{\mu\nu}\times A_\mu$ present in (214), because A_μ does not enter linearly in (190).

Expression (214)-(215) was first obtained in [268] by extremizing the cubic order action of the perturbations. This is equivalent to working directly with the equations expanded at second order, and we have verified that our result coincides with that of [268]. The source J_2 is the standard result obtained at second order in single scalar field inflation. The scalar part of this expression in the gauge we have adopted is explicitly given in [225].

Therefore, the inflaton perturbation is formally given by

$$\delta\varphi = \left(\delta_1\varphi + \delta_2\varphi|_{\text{sourced by } J_2}\right) + \delta_2\varphi|_{\text{sourced by } J_1} \tag{216}$$

The part in parenthesis is the standard result obtained in single scalar field inflation, with only negligible corrections coming from the backreaction of the gauge field on the background dynamics. The $\delta_2\varphi|_{\text{sourced by } J_2}$ term is clearly negligible in the primordial power spectrum, and also leads to unobservable non-Gaussianity [8, 223, 269, 270]. We therefore disregard it in our calculations. The last term in (216) encodes the effect of the gauge field on the inflaton perturbations. We note that this term is uncorrelated with the other two, since the quantum/statistical operators entering in J_1 are those of the gauge field. Therefore, we are interested in computing

$$\left\langle \delta\varphi^2 \right\rangle \simeq \left\langle \left(\delta\varphi_{\text{vacuum}}\right)^2 \right\rangle + \left\langle \left(\delta\varphi_{\text{sourced}}\right)^2 \right\rangle \;, \quad \left\langle \delta\varphi^3 \right\rangle \simeq \left\langle \left(\delta\varphi_{\text{sourced}}\right)^3 \right\rangle \tag{217}$$

where

$$\delta\varphi_{\text{vacuum}} \equiv \delta_1\varphi \;, \quad \delta\varphi_{\text{sourced}} \equiv \delta_2\varphi|_{\text{sourced by } J_1} \tag{218}$$

We combine (213) and (214) in a unique equation for $\delta\varphi = \delta\varphi_{\text{vacuum}} + \delta\varphi_{\text{sourced}}$, where the two quantities are defined in (218). We approximate this equation in slow roll approximation and we keep only the leading source term that arises from the direct I^2F^2 coupling. This gives

$$\left[\partial_\tau^2 + 2\frac{a'}{a}\partial_\tau - \nabla^2\right]\delta\varphi \simeq J \;, \quad J = a^2\,\frac{I_\varphi}{I}\left[E^2 - B^2\right] \simeq -\sqrt{\frac{2}{\epsilon}}\,\frac{a^2}{M_p}\,\left[E^2 - B^2\right] \tag{219}$$

We note that the source $E^2 - B^2$ interacts with the inflaton perturbation with a strength that is gravitationally suppressed but slow roll $1/\sqrt{\epsilon}$ enhanced (this is one of the enhancements that make non-Gaussianity visible in the model). The remaining terms in (215) have the same scale dependence, but an interaction strength that is both gravitationally and slow roll $\sqrt{\epsilon}$ suppressed. The same suppression characterizes all the terms in J_2. Therefore, the dominant source is that one arising from the direct inflaton-gauge field coupling $I^2(\varphi)\,F^2$. We explicitly see that simply computing the effect on $\delta\varphi$ from the gauge fields in an unperturbed metric reproduces the leading results for $\delta_2\varphi$.

In this calculation, we have estimated the curvature perturbation in flat gauge as $\zeta = -\frac{H}{\dot{\varphi}^{(0)}}\delta\varphi$. This equation actually receives corrections at second order; however, these corrections are (1) subdominant at the end of inflation, and (2) become even smaller (by several orders of magnitude) during reheating, when the energy in the gauge field decreases faster than the one of the dominating plasma (the equation of state of the dominating plasma is either the one of matter – for perturbative reheating – or intermediate between the one of matter and radiation [252] – in the nonperturbative case). We explicitly show this in Appendix C.

2.1 Two-point Correlation Function, and Correction to Power Spectrum

We are interested in the primordial curvature perturbation, given by

$$\zeta(t,\vec{x}) = -\frac{H}{\dot{\varphi}^{(0)}}\delta\varphi(t,\vec{x}) \tag{220}$$

The two point correlation function in momentum space is related to the power spectrum by the standard expression

$$\frac{H^2}{\dot{\varphi}^{(0)\,2}}\left\langle \frac{Q_\varphi(\vec{k})}{a}\,\frac{Q_\varphi(\vec{k}')}{a}\right\rangle = \langle \zeta_{\vec{k}}\,\zeta_{\vec{k}'}\rangle \equiv P_\zeta\,(k)\,\frac{2\pi^2}{k^3}\delta^{(3)}\left(\vec{k}+\vec{k}'\right) \tag{221}$$

where all quantities are evaluated at some time τ, and where we use the convention (49),

$$\delta\varphi(\tau,\vec{x}) = \int \frac{d^3k}{(2\pi)^{3/2}}\,\frac{Q_\varphi(\tau,\vec{k})}{a(\tau)}\,e^{i\vec{k}\cdot\vec{x}} \tag{222}$$

We define the Fourier transform of the source as

$$J_\varphi\left(\vec{k}\right) \equiv a\int \frac{d^3x}{(2\pi)^{3/2}}\,\mathrm{e}^{-i\vec{k}\cdot\vec{x}}J = -\sqrt{\frac{2}{\epsilon}}\,\frac{a^3}{M_p}\int\frac{d^3p}{(2\pi)^{3/2}}\left[\vec{E}_{\vec{p}}\cdot\vec{E}_{\vec{k}-\vec{p}} - \vec{B}_{\vec{p}}\cdot\vec{B}_{\vec{k}-\vec{p}}\right] \tag{223}$$

where (198) and (219) have been used. We find

$$\begin{aligned} J_\varphi\left(\tau,\vec{k}\right) \;\; &\simeq \;\; -\frac{H^4\,a^3\,(\tau)}{\sqrt{2\epsilon}\,M_p}\sum_{\lambda,\lambda'}\int\frac{d^3p}{(2\pi)^{3/2}}\,\frac{\vec{\epsilon}^{\,(\lambda)}\,(\vec{p})\cdot\vec{\epsilon}^{\,(\lambda')}\left(\vec{k}-\vec{p}\right)}{|\vec{p}|^{1/2}\,|\vec{k}-\vec{p}|^{1/2}}\left[\tau^2-\frac{9\lambda\lambda'}{|\vec{p}|\,|\vec{k}-\vec{p}|}\right] \\ &\times\left[a_\lambda\,(\vec{p})+a_\lambda^\dagger\,(-\vec{p})\right]\,\left[a_{\lambda'}\left(\vec{k}-\vec{p}\right)+a_{\lambda'}^\dagger\left(-\vec{k}+\vec{p}\right)\right] \end{aligned} \tag{224}$$

We stress that the first term in the square parenthesis in the first line of (224) is the E-contribution to the source, while the second term is the $B-$contribution.

Combining eqs. (219), (222), and (223), the equation for the Fourier modes of the inflaton perturbations is of the type (55)

$$Q''_\varphi + \left(k^2 - \frac{a''}{a}\right) Q_\varphi \simeq J_\varphi \tag{225}$$

where all quantities are evaluated at the time τ.

The solution for (225) is outlined in Section III 2. The homogeneous solution to this equation is the standard vacuum solution, which leads to the standard result (205) for the power spectrum (221). The particular solution is obtained with the Green function method

$$Q_\varphi^{\rm sourced}\left(\tau, \vec{k}\right) = \int_{\tau_{\rm in}}^{\tau} d\tau' G_k\left(\tau,\, \tau'\right) J_\varphi\left(\tau', \vec{k}\right) \tag{226}$$

$$\begin{aligned} G_k\left(\tau, \tau'\right) &\simeq \frac{1}{k^3\,\tau\,\tau'}\left[k\tau'\,\cos\left(k\tau'\right) - \sin\left(k\tau'\right)\right] \quad , \quad \vert k\,\tau\vert \ll 1 \\ &\simeq -\frac{\tau'^2}{3\,\tau} \quad , \quad \vert k\,\tau\vert\,,\,\left\vert k\,\tau'\right\vert \ll 1 \end{aligned} \tag{227}$$

The first approximation for G_k is valid since we are interested in superhorizon scales $\vert k\tau\vert \ll 1$ ($\vec{k}$ is the external momentum), and the second approximation is also valid because the sourcing effect occurs only outside the horizon.

Using the source (224), and the identity

$$\left\vert\vec{\epsilon}_\lambda\left(\vec{p}\right)\cdot\vec{\epsilon}_{\lambda'}\left(\vec{q}\right)\right\vert^2 = \frac{1}{4}\left(1 - \lambda\,\lambda'\,\hat{p}\cdot\hat{q}\right)^2 \tag{228}$$

(where hat denotes a unit vector) we obtain

$$\begin{aligned} \left\langle \frac{Q_\varphi^{\rm sourced}(\vec{k})}{a}\,\frac{Q_\varphi^{\rm sourced}(\vec{k}')}{a}\right\rangle &\simeq \frac{H^4\,\delta^{(3)}\left(\vec{k}+\vec{k}'\right)}{9\epsilon M_p^2\,k^3}\int\frac{d^3q}{(2\pi)^3}\int\frac{dy'}{y'}\int\frac{dy''}{y''} \\ \times\Bigg\{\left[1+\cos^2\left(\vec{q},\,\hat{k}-\vec{q}\right)\right] &\left[\frac{y'^2\,y''^2}{\vert\vec{q}\vert\,\left\vert\hat{k}-\vec{q}\right\vert} + \frac{81}{\vert\vec{q}\vert^3\,\left\vert\hat{k}-\vec{q}\right\vert^3}\right] + 18\cos\left(\vec{q},\,\hat{k}-\vec{q}\right)\,\frac{y'^2+y''^2}{\vert\vec{q}\vert^2\,\left\vert\hat{k}-\vec{q}\right\vert^2}\Bigg\} \end{aligned} \tag{229}$$

where we have introduced the dimensionless integration variables $\vec{q} = \vec{p}/k$, $y' = -k\tau'$, and $y'' = -k\tau''$. The momentum integral need to be restricted so that the gauge modes participating in the original convolution were inside the horizon at the start of inflation (otherwise they would not be produced by this mechanism). This means

$$|\vec{q}|\ ,\ |\hat{k}-\vec{q}| > \frac{1}{k\,|\tau_{\rm in}|} \tag{230}$$

The time integrations are instead restricted to times which are between $\tau_{\rm in}$ and τ, and for which the sourcing modes have exited the horizon. This means $\tau_{\rm in}$, $-\frac{1}{|\vec{p}|}$, $-\frac{1}{|\vec{k}-\vec{p}|} < \tau'$, $\tau'' < \tau$. We do not need to include $\tau_{\rm in}$ in this condition thanks to (230). Therefore, the time integrals in (229) are restricted to

$$k\,|\tau| < y',\, y'' < {\rm Min}\left[\frac{1}{|\vec{q}|},\, \frac{1}{|\hat{k}-\vec{q}|}\right] \tag{231}$$

In Section V 1 we saw that the energy density in the produced $B-$field is logarithmically enhanced with respect to that in the $E-$field. Eq. (229) shows that the an analogous logarithmic enhancement takes place in the source of $\delta\varphi$. In this case, all the three integrals in (229) present an enhancement. Disregarding the subdominant $E-$contribution, and performing the time integrals, the expression (229) gives

$$\left\langle \frac{Q_\varphi^{\rm sourced}(\vec{k})}{a}\,\frac{Q_\varphi^{\rm sourced}(\vec{k}')}{a}\right\rangle \simeq \frac{9H^4}{\epsilon M_p^2}\,\frac{\delta^{(3)}\left(\vec{k}+\vec{k}'\right)}{k^3} \times \int \frac{d^3q}{(2\pi)^3}\,\frac{1+\cos^2\left(\vec{q},\,\hat{k}-\vec{q}\right)}{|\vec{q}|^3\,\left|\hat{k}-\vec{q}\right|^3}\,\ln^2\frac{{\rm Min}\left[\frac{1}{|\vec{q}|},\,\frac{1}{|\hat{k}-\vec{q}|}\right]}{k|\tau|} \tag{232}$$

The momentum integral has most of its support at the two logarithmic poles; due to the symmetry between the two poles, we can simply evaluate the integral for $|\vec{q}| \simeq \frac{1}{k|\tau_{\rm in}|} \ll 1$ and multiply the result by two:

$$\left\langle \frac{Q_\varphi^{\rm sourced}(\vec{k})}{a}\,\frac{Q_\varphi^{\rm sourced}(\vec{k}')}{a}\right\rangle \simeq \frac{12}{\pi^2}\,\frac{H^4}{\epsilon\,M_p^2}\,\frac{\delta^{(3)}\left(\vec{k}+\vec{k}'\right)}{k^3}\,\ln^2\frac{1}{k|\tau|}\,\ln\left(k|\tau_{\rm in}|\right) \tag{233}$$

The first logarithmic enhancement is due to the growth of the two modes $\frac{Q_\varphi(\vec{k})}{a}$ and $\frac{Q_\varphi(\vec{k}')}{a}$ in the super-horizon regime. The growth is due to the presence of the entropy modes A_μ. At the end of inflation $\ln^2\frac{1}{k|\tau|} = N_{\rm CMB}^2$. The second enhancement $\ln\left(k|\tau_{\rm in}|\right) = N_{\rm tot} - N_{\rm CMB}$ is due to the number of gauge field modes that source the inflaton perturbation. More specifically, we have seen that each sourced inflaton mode is obtained as a convolution of two gauge field modes. The momenta of these two modes need to add up to the momentum of the inflaton mode. The second enhancement occurs in the IR limit of one of the two gauge modes. This enhancement is the counterpart of the enhancement taking place for ρ_B, which is also due to the number of large wavelength modes produced during inflation.

As we discussed, the contribution (233) adds up incoherently with the vacuum one in (221). Using the relation $\zeta = -\frac{H}{\dot{\varphi}^{(0)}}\delta\varphi$, the sum gives

$$P_\zeta|_{\text{end inflation}} \simeq \mathcal{P}\left[1 + 192\,\mathcal{P}\,N_{\text{CMB}}^2\,\left(N_{\text{tot}} - N_{\text{CMB}}\right)\right] \tag{234}$$

where we recall that $\mathcal{P}$ is the contribution from the vacuum term, for which the standard slow roll expression (205) holds. We also remind that $N_{\text{CMB}} \simeq 50 - 60$ is the number of e-folds before the end of inflation when the largest scale CMB modes left the horizon, while N_{tot} is the total number of e-folds of inflation. The enhancement from the momentum integral takes place for $N_{\text{tot}} \gg N_{\text{CMB}}$; if inflation only lasted about the observed number of e-folds, then we estimate that the final momentum integral produces an order one result, so that the result (234) remains valid as an order of magnitude estimate.

For $N_{\text{tot}} \gg N_{\text{CMB}} \sim 60$, the ratio between the sourced and the standard power spectrum is $\simeq 1.7 \cdot 10^{-3}\,N_{\text{tot}}$. The standard term dominates provided that inflation lasted less than about 600 e-folds. Here we assume that this is the case.

2.2 Three-point Correlation Function, and Observable Non-Gaussianity

We are interested in the three point correlation function of ζ as a measure of non-Gaussianity. A common parametrization of non-Gaussianity is the nonlinearity parameter f_{NL}, introduced by assuming that the curvature perturbation may be expanded as in (39)

$$\zeta\left(\vec{x}\right) = \zeta_g\left(\vec{x}\right) + \frac{3}{5}\,f_{\text{NL}}\,\left[\zeta_g^2\left(\vec{x}\right) - \langle\zeta_g^2\left(\vec{x}\right)\rangle\right] \tag{235}$$

where $\zeta_g(x)$ is a Gaussian random field. Both ζ and ζ_g may be decomposed as in (222) so that the relation between the q-modes of the Fourier decomposition is

$$\zeta_{\vec{k}} = \zeta_{g,\vec{k}} + \frac{3}{5}\,f_{\text{NL}}\int\frac{d^3p}{\left(2\pi\right)^{3/2}}\,\zeta_{g,\vec{k}}\,\zeta_{g,\vec{k}-\vec{p}} \tag{236}$$

The three point correlator of ζ_g vanishes, as this field is Gaussian. However, due to the quadratic term in (235), the three point correlator of ζ is nonvanishing, and can be expressed through a sum of two point correlators of ζ_g. One finds

$$\langle \zeta_{\vec{k}_1} \, \zeta_{\vec{k}_2} \, \zeta_{\vec{k}_3} \rangle = \frac{3}{10} \, (2\pi)^{5/2} \, f_{\rm NL} P_\zeta \, (k)^2 \, \delta^{(3)} \left(\vec{k}_1 + \vec{k}_2 + \vec{k}_3 \right) \frac{\sum_i k_i^3}{\Pi_i k_i^3} \tag{237}$$

where the power spectrum was defined in (221). To obtain this expression, one identifies the two point function of ζ with that of ζ_g (as the difference is subleading in a perturbative expansion), and disregards the mild scale dependence of the power spectrum.

By evaluating $\langle \zeta^3 \rangle$, and by inserting it in (237), one defines an "effective" (momentum dependent) nonlinearity parameter, even when the intrinsic non-Gaussianity is not of the local form (235). The dependence of $f_{\rm NL}$ on the relative size of the momenta is denoted as "shape" of the non-Gaussianity. Ref. [26] provides a method to evaluate whether the shape obtained in a given model is well reproduced by the local template (namely, $f_{\rm NL}$ constant in (237)) or by any other template employed in data analysis.

As non-Gaussianity from the vacuum term is negligible, we need to compute

$$\left\langle \zeta_{\vec{k}_1} \, \zeta_{\vec{k}_2} \, \zeta_{\vec{k}_3} \right\rangle \simeq - \frac{H^3}{\dot{\varphi}^{(0)\,3}} \left\langle \frac{Q_\varphi^{\rm sourced}(\vec{k}_1)}{a} \, \frac{Q_\varphi^{\rm sourced}(\vec{k}_2)}{a} \, \frac{Q_\varphi^{\rm sourced}(\vec{k}_3)}{a} \right\rangle \tag{238}$$

where each expression is evaluated at some given time τ.

We proceed as in the previous Subsection by inserting the source (224) into (226) and by evaluating the correlator. Keeping only the dominant "magnetic" source, we obtain

$$\begin{aligned} \left\langle \prod_{i=1}^{3} \frac{Q_\varphi^{\rm sourced}(\vec{k}_i)}{a} \right\rangle \simeq & \; \frac{1}{a^3} \, \frac{729}{8\pi^{9/2}} \, \frac{H^3}{\epsilon^{3/2} M_p^3} \prod_{i=1}^{3} \int d\tau_i \frac{G_{k_i}(\tau, \, \tau_i)}{(-\tau_i)^3} \sum_{\lambda_i \, \sigma_i} \int d^3 p_i \, \frac{\vec{\epsilon}^{(\lambda_i)}\,(\vec{p}_i) \cdot \vec{\epsilon}^{(\sigma_i)} \left(\vec{k}_i - \vec{p}_i \right)}{|\vec{p}_i|^{3/2} \, |\vec{k}_i - \vec{p}_i|^{3/2}} \\ & \times \delta_{\sigma_1 \lambda_2} \delta^{(3)} \left(\vec{k}_1 - \vec{p}_1 + \vec{p}_2 \right) \delta_{\sigma_2 \lambda_3} \delta^{(3)} \left(\vec{k}_2 - \vec{p}_2 + \vec{p}_3 \right) \delta_{\sigma_3 \lambda_1} \delta^{(3)} \left(\vec{k}_3 - \vec{p}_3 + \vec{p}_1 \right) \end{aligned} \tag{239}$$

(where the delta functions emerge from commutators between a and $a^\dagger$ gauge field operators in the standard way) where the integration regions are bounded in an analogous way as for the two point function:

$$|\vec{p}_i| , \; |\vec{k}_i - \vec{p}_i| > \frac{1}{|\tau_{\rm in}|} \; , \quad |\tau| < |\tau_i| < {\rm Min} \left[\frac{1}{|\vec{p}_i|} , \, \frac{1}{|\vec{k}_i - \vec{p}_i|} \right] \tag{240}$$

Performing the time integrals and employing the delta functions, we obtain

$$\left\langle \prod_{i=1}^{3} \frac{Q_\varphi^{\rm sourced}(\vec{k}_i)}{a} \right\rangle \simeq \frac{27}{8\pi^{9/2}} \frac{H^6}{\epsilon^{3/2} M_p^3} \delta^{(3)}\left(\vec{k}_1+\vec{k}_2+\vec{k}_3\right) \int d^3p$$
$$\times \ln {\rm Min}\left[\frac{1}{|\vec{p}|\,|\tau|},\, \frac{1}{|\vec{k}_1-\vec{p}|\,|\tau|}\right] \ln {\rm Min}\left[\frac{1}{|\vec{k}_1-\vec{p}|\,|\tau|},\, \frac{1}{|\vec{k}_3+\vec{p}|\,|\tau|}\right] \ln {\rm Min}\left[\frac{1}{|\vec{k}_3+\vec{p}|\,|\tau|},\, \frac{1}{|\vec{p}|\,|\tau|}\right]$$
$$\times \frac{\sum_{\lambda_1} \epsilon_k^{(\lambda_1)*}\left(\vec{p}\right) \epsilon_i^{(\lambda_1)}\left(\vec{p}\right) \sum_{\lambda_2} \epsilon_i^{(\lambda_2)*}\left(\vec{p}-\vec{k}_1\right) \epsilon_j^{(\lambda_2)}\left(\vec{p}-\vec{k}_1\right) \sum_{\lambda_3} \epsilon_j^{(\lambda_3)*}\left(\vec{p}+\vec{k}_3\right) \epsilon_k^{(\lambda_3)}\left(\vec{p}+\vec{k}_3\right)}{|\vec{p}|^3\,|\vec{p}-\vec{k}_1|^3\,|\vec{p}+\vec{k}_3|^3} \tag{241}$$

where, due to (240), the integration region is delimited by

$$|\vec{p}|,\ |\vec{p}-\vec{k}_1|,\ |\vec{p}+\vec{k}_3| > \frac{1}{|\tau_{\rm in}|} \tag{242}$$

The momentum integral in (241) has most of its support at the three logarithmic poles; each pole occurs when one of the three quantities in (242) reaches its minimal value $\frac{1}{|\tau_{\rm in}|}$. Formally,

$$\left\langle \prod_{i=1}^{3} \frac{Q_\varphi^{\rm sourced}(\vec{k}_i)}{a} \right\rangle \simeq \mathcal{C}_{|\vec{p}|\simeq\frac{1}{|\tau_{\rm in}|}}\left[\vec{k}_1,\, \vec{k}_2,\, \vec{k}_3\right] + \mathcal{C}_{|\vec{p}-\vec{k}_1|\simeq\frac{1}{|\tau_{\rm in}|}}\left[\vec{k}_1,\, \vec{k}_2,\, \vec{k}_3\right] + \mathcal{C}_{|\vec{p}+\vec{k}_3|\simeq\frac{1}{|\tau_{\rm in}|}}\left[\vec{k}_1,\, \vec{k}_2,\, \vec{k}_3\right] \tag{243}$$

where $\mathcal{C}$ refers to the contribution to the integral in (241) from the region close to the pole indicated by the suffix.

To evaluate the contribution from the second region, we redefine the integration variable as $\vec{p} \to \vec{p}+\vec{k}_1$; we then see that this contribution is formally equal to the contribution from the first region, provided the external momenta in the first region are changed as $\vec{k}_1 \to \vec{k}_2$, $\vec{k}_2 \to \vec{k}_3$, and $\vec{k}_3 \to \vec{k}_1$. Analogously, to evaluate the contribution from the third region, we redefine the integration variable as $\vec{p} \to \vec{p}-\vec{k}_3$; we then see that this contribution is formally equal to the contribution from the first region, provided the external momenta in the first region are changed as $\vec{k}_1 \to \vec{k}_3$, $\vec{k}_2 \to \vec{k}_1$, and $\vec{k}_3 \to \vec{k}_2$. We therefore have

$$\left\langle \prod_{i=1}^{3} \frac{Q_\varphi^{\rm sourced}(\vec{k}_i)}{a} \right\rangle \simeq \mathcal{C}_{|\vec{p}|\simeq\frac{1}{|\tau_{\rm in}|}}\left[\vec{k}_1,\, \vec{k}_2,\, \vec{k}_3\right] + \text{permutations} \tag{244}$$

To evaluate this contribution, we use the identity

$$\sum_\lambda \epsilon_i^{(\lambda)*}\left(\vec{p}\right)\,\epsilon_j^{(\lambda)}\left(\vec{p}\right) = \delta_{ij} - \hat{p}_i\,\hat{p}_j \tag{245}$$

and we obtain

$$\begin{aligned}\left\langle \prod_{i=1}^{3} \frac{Q_\varphi^{\rm sourced}(\vec{k}_i)}{a} \right\rangle &\simeq \frac{9}{\pi^{7/2}}\,\frac{H^6}{\epsilon^{3/2} M_p^3}\,\delta^{(3)}\left(\vec{k}_1+\vec{k}_2+\vec{k}_3\right) \\ &\times \Bigg\{ \ln\frac{1}{k_1\left|\tau\right|}\,\ln\frac{1}{k_3\left|\tau\right|}\,\ln{\rm Min}\left[\frac{1}{k_1\left|\tau\right|},\,\frac{1}{k_3\left|\tau\right|}\right]\left[1+\cos^2\left(\vec{k}_1,\,\vec{k}_3\right)\right]\,\frac{\ln\,{\rm Min}\left[k_1\left|\tau_{\rm in}\right|,\,k_3\left|\tau_{\rm in}\right|\right]}{k_1^3 k_3^3} \\ &\quad + {\rm permutations}\Bigg\}\end{aligned} \tag{246}$$

Assuming that the external momenta are not too hierarchical (see below), we disregard the difference among them in the argument of the logarithms. The expression for the correlator, evaluated at the end of inflation, then simplifies to

$$\begin{aligned}\left\langle \prod_{i=1}^{3} \frac{Q_\varphi^{\rm sourced}(\vec{k}_i)}{a} \right\rangle &\simeq 144\sqrt{\frac{2}{\pi}}\,\mathcal{P}^{3/2}\,H^3\,\delta^{(3)}\left(\vec{k}_1+\vec{k}_2+\vec{k}_3\right) N_{\rm CMB}^3\left(N_{\rm tot}-N_{\rm CMB}\right) \\ &\times \left[\frac{1+\cos^2\left(\vec{k}_1,\,\vec{k}_2\right)}{k_1^3\,k_2^3} + \frac{1+\cos^2\left(\vec{k}_1,\,\vec{k}_3\right)}{k_1^3\,k_3^3} + \frac{1+\cos^2\left(\vec{k}_2,\,\vec{k}_3\right)}{k_2^3\,k_3^3}\right]\end{aligned} \tag{247}$$

We can now evaluate the 3-point function of the curvature perturbation using the relation $\zeta = -\frac{H}{\dot{\varphi}^{(0)}}\delta\varphi$ and introduce a momentum-dependent nonlinearity parameter by comparison with (237). We find:

$$\begin{aligned} f_{\rm NL}(k_i) &\simeq f_{\rm NL}^{\rm equiv.\ local} \times \frac{3}{4}\,\frac{\frac{1+\cos^2\left(\vec{k}_1,\vec{k}_2\right)}{k_1^3\,k_2^3} + \frac{1+\cos^2\left(\vec{k}_1,\vec{k}_3\right)}{k_1^3\,k_3^3} + \frac{1+\cos^2\left(\vec{k}_2,\vec{k}_3\right)}{k_2^3\,k_3^3}}{\frac{1}{k_1^3\,k_2^3}+\frac{1}{k_1^3\,k_3^3}+\frac{1}{k_2^3\,k_3^3}} \\ f_{\rm NL}^{\rm equiv.\ local} &\simeq 1280\,\frac{\mathcal{P}^3}{P_\zeta(k)^2}\,N_{\rm CMB}^3\left(N_{\rm tot}-N_{\rm CMB}\right)\end{aligned} \tag{248}$$

If the power spectrum is dominated by the vacuum fluctuations (which we assume here) then we have $P_\zeta(k) \approx \mathcal{P}$ and $f_{\rm NL}^{\rm equiv.\ local} \simeq 1280\,\mathcal{P} N_{\rm CMB}^3\left(N_{\rm tot}-N_{\rm CMB}\right)$.

We conclude by noting that a more precise shape than (248) can be readily obtained from

(246) also for hierarchical momenta; for instance, in the limit $k_1 \ll k_2 \simeq k_3$, the third term in the numerator of $f_{\rm NL}(k_i)$ becomes irrelevant, and the $N_{\rm CMB}$ factors entering in $f_{\rm NL}^{\rm equiv.\ local}$ should read $N_{{\rm CMB},1}\, N_{{\rm CMB},2}^2\, (N_{\rm tot} - N_{{\rm CMB},1})$, where $N_{{\rm CMB},i}$ refers to the horizon-exit of the mode with momentum k_i.

3 Tensor Modes

Production of gauge field fluctuations during inflation and its effect on the curvature perturbation have been discussed in the previous Sections. The produced gauge quanta, however, affect not only the scalar but also tensor perturbations (gravity waves). Metric perturbations couple to each content of the universe and are inevitably sourced by the produced gauge field fluctuations. To see the effect, we consider the transverse and traceless components of the spatial metric perturbations: $g_{ij} = a^2\,(\delta_{ij} + h_{ij})$, with $h_{ii} = 0$ and $\partial_i h_{ij} = 0$. As the matter content is scalar and vector, h_{ij} is the only tensor perturbations in the model.

From the Einstein equations, one finds the same equations for h_{ij} in terms of the physical E_i and B_i as in the axion inflation model (See Subsection IV 2.5),

$$\frac{1}{2a^2}\left(\partial_\tau^2 + 2\frac{a'}{a}\partial_\tau - \nabla^2\right) h_{ij} = -\frac{1}{M_p^2}\,(E_iE_j + B_iB_j)^{TT} \tag{249}$$

where TT denotes the transverse and traceless projection of the spatial components of the energy-momentum tensor of the gauge field.[36] Since the gauge field has no expectation value, there is no coupling to the tensor perturbations at linearized level. Thus (249) is in fact up to second order, and the right-hand side should in principle contain the source terms from squares of the first-order inflaton and metric perturbations. As for $\delta_2\varphi$, such source terms are uncorrelated and subdominant to those from the gauge field.

Tensor modes (or GW) are transverse and traceless part of δg_{ij} and have two physical degrees of freedom, the left-handed (L) and right-handed (R). It is convenient to decompose tensor perturbations as

$$h_{ij}\,(\tau, \vec{x}) = \int \frac{d^3k}{(2\pi)^{3/2}}\, \mathrm{e}^{i\vec{k}\cdot\vec{x}} \sum_{\lambda=L,R} \Pi_{ij}^{(\lambda)}(\vec{k})\hat{h}_\lambda(\vec{k}) \tag{250}$$

where $\hat{h}_\lambda(\vec{k}) = h_\lambda\,(k)\, b_\lambda(\vec{k}) + h_\lambda^*\,(k)\, b_\lambda^\dagger(-\vec{k})$, and the helicity projectors are $\Pi_{ij}^{(R\mathcal{L})}(\vec{k}) =$

[36] This projection can be done by an operator $\mathcal{O}_{ij,lm}$, which can be expressed in the momentum space as $\mathcal{O}_{ij,lm}(\hat{k}) = P_{il}(\hat{k})P_{jm}(\hat{k}) - \frac{1}{2}P_{ij}(\hat{k})P_{lm}(\hat{k})$, where $P_{ij}(\hat{k}) = \delta_{ij} - \hat{k}_i\hat{k}_j$.

$e_i^{(\pm)}(\vec{k})\, e_j^{(\pm)}(\vec{k})$, which clearly have the properties $\Pi_{ii}^{(\lambda)}(\vec{k}) = \hat{k}_i\, \Pi_{ij}^{(\lambda)}(\vec{k}) = 0$. Note that $h_\lambda(k)$ depends only on the magnitude of $\vec{k}$. Since the mechanism of GW production is analogous to that of the curvature perturbations, presented in detail in the previous Sections, we merely show the result of our computation here. Define GW power spectrum $P_{L/R}$ in the usual way (72),

$$P_\lambda(k)\ \delta_{\lambda\lambda'}\, \delta^{(3)}\left(\vec{k}+\vec{k}'\right) \equiv \frac{k^3}{2\pi^2}\left\langle \hat{h}_\lambda(\vec{k})\, \hat{h}_{\lambda'}(\vec{k}')\right\rangle \tag{251}$$

where $\lambda = L/R$. As for the scalar perturbations, the GW modes produced by the gauge quanta are uncorrelated with those from vacuum fluctuations, and so the two contributions simply add up in the power spectrum. The two helicity states are produced in the same amount, and their sum gives

$$P_{\rm GW}(k) \cong \frac{2H^2}{\pi^2 M_p^2}\left[1 + \frac{6H^2}{\pi^2 M_p^2}\ln^2\frac{a(\tau)\,H}{k}\ \ln\frac{k}{a(\tau_{\rm in})\,H}\right] \tag{252}$$

where the first term in the square brackets is the contribution from the vacuum and the latter from the source. Evaluating this expression at the end of inflation gives

$$P_{\rm GW}\big|_{\text{end inflation}} = \frac{2H^2}{\pi^2 M_p^2}\left[1 + \frac{6H^2}{\pi^2 M_p^2} N_{\rm CMB}^2\left(N_{\rm tot} - N_{\rm CMB}\right)\right] \tag{253}$$

As expected, the standard vacuum part dominates in the regime we are interested in: when the vacuum contribution to P_ζ is dominant over that from the source, the same is also true for the GW power spectrum; this is due to the fact that the tensor modes are only produced gravitationally, while the dominant source of the scalar modes is the direct inflaton-gauge field coupling; (this coupling is mathematically enhanced with respect to the gravitational one by $1/\epsilon$). In this regime, the tensor-to-scalar ratio r reproduces the standard result:[37]

$$r \approx 16\epsilon \tag{254}$$

4 Phenomenology of the Dilaton-like Coupling Model

In this Section, we summarize the phenomenological predictions of the model (190). The actual computations were performed in Sections V 2 and V 3.

[37] We define the tensor-to-scalar ratio in the usual way, by normalizing the power in GW to that in curvature perturbations: $r \equiv P_{\rm GW}/P_\zeta = (P_L + P_R)/P_\zeta$.

We define the power spectrum of curvature fluctuations via (35),

$$\langle \zeta_{\vec{k}} \, \zeta_{\vec{k}'} \rangle = \frac{2\pi^2}{k^3} P_\zeta(k) \delta^{(3)} \left(\vec{k} + \vec{k}' \right) \tag{255}$$

The final result for P_ζ, evaluated on large scales and at the end of inflation, is

$$P_\zeta = \mathcal{P} \left[1 + 192 \, \mathcal{P} \, N_{\rm CMB}^2 \, \left(N_{\rm tot} - N_{\rm CMB} \right) \right] \tag{256}$$

where $\mathcal{P}^{1/2} \equiv \frac{H^2}{2\pi|\dot{\varphi}^{(0)}|}$ is the amplitude of the power spectrum of the vacuum modes, $N_{\rm CMB}$ denotes the number of e-folds between the moment at which the large scales CMB modes leave the horizon and the end of inflation, and $N_{\rm tot}$ is the total number of e-folds of inflation. The first term in (256) is the usual contribution from the vacuum while the second term is the "sourced" contribution described heuristically in the last Subsection. Here we work to leading order in slow roll parameters, so the spectrum (256) is exactly flat. In a more complete computation we would see small departures from scale invariance $\propto k^{n_s-1}$.

We can physically understand the structure of (256) as follows. The suppression $\mathcal{P} \ll 1$ in the second term arises simply because we are computing an effect which is higher order in perturbation theory.[38] The factors of $N_{\rm CMB}$, on the other hand, arise due to the logarithmic time evolution, $\zeta_{\rm sourced} \sim \ln a$, outside the horizon. Such growth is consistent since we have large scale entropy perturbations playing an important role in the dynamics. Finally, the factor $N_{\rm tot} - N_{\rm CMB}$ is related to the phase space of contributing gauge field fluctuations. It is related to the number of B-modes that source the inflaton perturbation, and it is the counterpart of the logarithmic enhancement in the background density (200). We explain this in details after eq. (233). We should stress that eq. (256) is valid only when $N_{\rm tot} \gg N_{\rm CMB}$. Otherwise the factor $N_{\rm tot} - N_{\rm CMB}$ is replaced by an order one factor.

We require that the sourced contribution to the power spectrum is subdominant:

$$192 \, \mathcal{P} \, N_{\rm CMB}^2 \, \left(N_{\rm tot} - N_{\rm CMB} \right) < 1 \tag{257}$$

This yields a constraint on the total number of e-foldings $N_{\rm tot} - N_{\rm CMB} < \mathcal{O}(10^{-6})\mathcal{P}^{-1}$ (taking $N_{\rm CMB} \sim 60$ for illustration) which is considerably more stringent than the backreaction bounds discussed in the last Section.

The bispectrum is given by the 3-point correlation function (38)

$$\langle \zeta_{\vec{k}_1} \zeta_{\vec{k}_2} \zeta_{\vec{k}_3} \rangle \equiv B_\zeta \left(k_i \right) \, \delta^{(3)} \left(\vec{k}_1 + \vec{k}_2 + \vec{k}_3 \right) \tag{258}$$

[38] See [49] for more discussion on the counting of such factors in models with particle production.

We have found that our bispectrum is very close to the local shape. Indeed, the "cosine" between our bispectrum and the local shape, as defined in [26] (this is a measure on how well a template reproduced a given bi-spectrum), is about 0.98. Nevertheless, here we will retain the full momentum dependence of the bispectrum, since it has a simple analytical shape.

As is conventional in the literature (see e.g. [193]), we defined a k-dependent nonlinearity parameter from computing the bispectrum and the power spectrum of ζ, and by comparing them with those obtained from (39)

$$\zeta\left(\vec{x}\right)=\zeta_g\left(\vec{x}\right)+\frac{3}{5}\,f_{\rm NL}\,\left[\zeta_g^2\left(\vec{x}\right)-\langle\zeta_g^2\left(\vec{x}\right)\rangle\right] \tag{259}$$

where ζ_g is Gaussian. An explicit computation gives the result

$$\begin{aligned}
f_{\rm NL}(k_i) &\simeq f_{\rm NL}^{\rm equiv.\ local}\times\frac{3}{4}\,\frac{\frac{1+\cos^2\left(\vec{k}_1,\vec{k}_2\right)}{k_1^3\,k_2^3}+\frac{1+\cos^2\left(\vec{k}_1,\vec{k}_3\right)}{k_1^3\,k_3^3}+\frac{1+\cos^2\left(\vec{k}_2,\vec{k}_3\right)}{k_2^3\,k_3^3}}{\frac{1}{k_1^3\,k_2^3}+\frac{1}{k_1^3\,k_3^3}+\frac{1}{k_2^3\,k_3^3}}\\
f_{\rm NL}^{\rm equiv.\ local} &\simeq 1280\,\mathcal{P}N_{\rm CMB}^3\left(N_{\rm tot}-N_{\rm CMB}\right)\\
&\simeq 0.7\left(\frac{N_{\rm CMB}}{60}\right)^3\left(N_{\rm tot}-N_{\rm CMB}\right)
\end{aligned} \tag{260}$$

where we have assumed that (257) is satisfied. We can a-posteriori see that this condition is indeed always satisfied whenever the result (260) is within the observational limits.

We have defined our "equivalent" local nonlinearity parameter as follows: we first note that both our bispectrum and the local template (40) are enhanced in the squeezed limit. In this limit, our bispectrum satisfies

$$\begin{aligned}
k_1^3\,k_3^3\,\langle\zeta_{\vec{k}_1}\,\zeta_{\vec{k}_2}\,\zeta_{\vec{k}_3}\rangle &\propto 1+\cos^2\left(\vec{k}_1,\,\vec{k}_3\right) \quad , \quad k_3\ll k_1\simeq k_2\\
&\propto \cos\epsilon\,Y_0^0+\sin\epsilon\,Y_2^0 \quad , \quad \epsilon\equiv\tan^{-1}\frac{1}{2\sqrt{5}}\simeq 0.22
\end{aligned} \tag{261}$$

where Y_l^0 are normalized spherical harmonics, characterized by the angle between $\vec{k}_1$ and $\vec{k}_3$. The average of $f_{\rm NL}$ over all possible values of this angle is then equal to the average of $f_{\rm NL}^{\rm equiv.\ local}$. Note that the current CMB limit on local shape non-Gaussianity is $f_{\rm NL}^{\rm local}=2.7\pm 5.8$ at 68% CL [13].

Using the local template to study this signature is clearly a good approximation, given that the local template is characterized by the monopole part only in (261), and that $\cos\epsilon\simeq 0.98$. However, we note more than 1/5 of the amplitude in (261) is contributed by the quadrupole part. This is in contrast to what typically happens in non-Gaussianity from scalar fields only, where the quadrupole and higher harmonic terms in the squeezed limit expansion are suppressed by powers of k_3/k_1 and give a negligible contribution in this limit [204]. The angular dependence is imprinted by a "directionality" generated by the largest wavelength mode k_3^{-1}, seen by the smaller modes when they leave the horizon. In the scalar case, the directionality is typically due to a gradient, and therefore it vanishes in the $k_3 \to 0$ limit. In our case, the directionality is due to the polarization of the vectors, and it therefore remains finite in the limit.

Let us also discuss the parametrical dependence of $f_{\rm NL}^{\rm equiv.\ local}$. The factor $N_{\rm CMB}^3$ is due to the super-horizon growth of the three modes used in computing the three point function. The factor $N_{\rm tot} - N_{\rm CMB}$ is due to the number of super-horizon modes that contribute to the correlator, analogously to what we have described in relation to (256). Also in this case, this factor is replaced by an order one factor if $N_{\rm tot}$ is close to $N_{\rm CMB}$. The source term is most significant on very large scales, and this large-scale entropy mode leads to the nearly local shape of the bispectrum.

The "magnetic" fields also produce gravity wave modes, which add incoherently with the vacuum ones. The power of gravitational waves produced during inflation is conventionally parametrized by the ratio r of their power divided by the scalar power. We find

$$r \equiv \frac{P_{\rm GW}}{P_\zeta} \simeq 16\,\epsilon\,\frac{1 + 48\,\epsilon\,\mathcal{P}\,N_{\rm CMB}^2\,(N_{\rm tot} - N_{\rm CMB})}{1 + 192\,\mathcal{P}\,N_{\rm CMB}^2\,(N_{\rm tot} - N_{\rm CMB})} \tag{262}$$

which gives the standard result $r \approx 16\epsilon$ when the vacuum modes dominate the power spectrum of curvature fluctuations.

The phenomenological results of the model (190) are summarized above in this Section: particle production of the gauge field through the inflaton-gauge interaction can modify the result of non-Gaussianity (260) over its standard result (the curvature power spectrum (256) and the tensor-to-scalar ratio (262) are standard in our consideration). Before ending the section, we note that the contributions to these results from the effects of particle production should however be interpreted as the theoretical expectation values, averaged over several realizations of the mechanism, as discussed in [53]. In [53], a careful and physically consistent study is performed on generic models of the dilaton-like coupling (188) during inflation. While their results are applicable to other relevant models [255, 303, 314], it also gives important insights to the results we obtained in the previous Sections.[39]

[39] Ref. [53] does not explicitly compute the tensor power spectrum, but it is the natural expectation that the the same be true in the tensor sector. Phenomenologically speaking, once the constraints on the scalar

With our choice of $I \propto a^2$, a nearly scale-invariant spectrum for the "magnetic" field is produced – see (199) – after the modes exit the horizon. Once produced, the modes become classical and behave as a homogeneous background in the point of view of the modes that cross the horizon at later stages during inflation. Viewed by the CMB modes, which exit the horizon $N_{\rm CMB}$ e-folds before the end of inflation, the larger-wavelength modes that left the horizon in the first $N_{\rm tot} - N_{\rm CMB}$ e-folds of inflation add up as such a background field.

Let us denote by $\vec{B}_{\rm classical}(\tau)$ the classical and homogeneous "magnetic" field, observed by a local observer at some time τ.[40] In our calculations, we assumed no vev of the gauge field which would obey the classical equation of motion, and therefore the classical fields consist only of the gauge fields that were produced through the I^2F^2 coupling. One can assume that no classical fluctuations are present at the initial time of inflation, $\tau_{\rm in}$, so that $\vec{B}_{\rm classical}(\tau_{\rm in}) = 0$. For $\tau > \tau_{\rm in}$, the modes that left the horizon become classical and homogeneous, seen by the modes of smaller scales. Any single realization of the mechanism of I^2F^2 leads to $\vec{B}_{\rm classical}$ obeying a Gaussian statistics with a vanishing mean and with variance

$$|\vec{B}_{\rm classical}|^2 \equiv \langle \vec{B}^2(\vec{x}) \rangle = 2\rho_B \simeq \frac{9H^4}{2\pi^2} N \tag{263}$$

where N is the number of e-folds between the times $\tau_{\rm in}$ and τ.

This produced homogeneous vector background by nature introduces a privileged direction in space, namely $\hat{B}_{\rm classical}$, which in turn inevitably encodes the anisotropy in the spectrum of the curvature perturbations through the sourcing mechanism. Consistently taking this classical field $\vec{B}_{\rm classical}$ into account, Ref. [53] finds the curvature power spectrum

$$P_\zeta \simeq \mathcal{P}\left[1 + \frac{24}{\epsilon}\,\frac{|\vec{B}_{\rm classical}|^2}{V(\varphi)}\,\mathcal{P}\,N_{\rm CMB}^2\,\sin^2\theta_{\hat{k},\hat{B}_{\rm classical}}\right] \tag{264}$$

where $\theta_{\hat{V}_1,\hat{V}_2}$ denotes the angle between the two (unit) vectors $\hat{V}_1$ and $\hat{V}_2$. Now averaging over all the possible directions of $\hat{B}_{\rm classical}$ and using (263), we find $|\vec{B}_{\rm classical}|^2 \sin^2\theta_{\hat{k},\hat{B}_{\rm classical}} \to \frac{3H^4}{\pi^2}N$. Here N is the classical modes viewed by the CMB modes, that is $N = N_{\rm tot} - N_{\rm CMB}$. Using the slow-roll relation $V \simeq 3M_p^2H^2$ (192) and the definition of $\mathcal{P}$ (205), we see that our result (256) is recovered.

perturbations are imposed, the tensor-to-scalar ratio r is standard and thus does not supply any further constraints.

[40] In [53], the authors take the choice $I \propto a^{-2}$ (our choice is $I \propto a^2$), and consequently the "electric" component is the dominant one, while with our choice the "magnetic" component dominates. However, due to the "magnetic" $\leftrightarrow$ "electric" duality under $I \leftrightarrow 1/I$, mentioned at the end of Subsection V 1.1, the phenomenological results from the produced vector fields are identical with either choice.

A similar argument is drawn for non-Gaussianity. The bispectrum B_ζ is enhanced in the squeezed limit precisely as the local form (40), as in our result (247). In this limit $k_1 \ll k_2 \simeq k_3$, Ref. [53] finds the non-linearity parameter

$$f_{\rm NL}^{\rm local} \simeq \frac{240}{\epsilon} \frac{\mathcal{P}^2}{P_\zeta^2} \frac{|\vec{B}_{\rm classical}|^2}{V(\varphi)} N_{\rm CMB}^3 \, \mathcal{C}_{\hat{k}_1,\hat{k}_2,\hat{B}_{\rm classical}} \,, \quad k_1 \ll k_2 \simeq k_3 \tag{265}$$

where

$$\mathcal{C}_{\hat{k}_1,\hat{k}_2,\hat{V}} \equiv 1 - \cos^2\theta_{\hat{k}_1,\hat{V}} - \cos^2\theta_{\hat{k}_2,\hat{V}} + \cos\theta_{\hat{k}_1,\hat{V}} \, \cos\theta_{\hat{k}_2,\hat{V}} \, \cos\theta_{\hat{k}_1,\hat{k}_2} \,. \tag{266}$$

Averaging over all the possible directions of $\hat{B}_{\rm classical}$, one finds $|\vec{B}_{\rm classical}|^2 \mathcal{C}_{\hat{k}_1,\hat{k}_2,\hat{B}_{\rm classical}} \to \frac{3H^4}{2\pi^2} \left(N_{\rm tot} - N_{\rm CMB}\right) \left(1 + \cos^2\theta_{\hat{k}_1,\hat{k}_2}\right)$, arriving at the identical expression to our result (260) in the same squeezed limit.

We have observed that our results (256) and (260) are the theoretical expectation values, averaged over ensemble of the realizations of the mechanism, as discussed in [53]. The modes of the vector field that left the horizon in the first $N_{\rm tot} - N_{\rm CMB}$ e-folds of inflation contribute to the curvature perturbations as a classical homogeneous background in the point of view of the modes at the CMB scales. Since in the history of our universe (assuming this mechanism describes it) the CMB modes exit the horizon after a single realization of the first $N_{\rm tot} - N_{\rm CMB}$ e-folds of inflation, they are affected by the value of $\vec{B}_{\rm classical}$ that happened to be taken in that single realization, instead of the isotropic average $\langle |\vec{B}_{\rm classical}|^2 \rangle$. However, for the data that are reduced to the level of an isotropic measurement, the isotropic predictions (256) and (260) of the signal are the quantities that should be used for the data analysis.

5 Summary

In this Part of the thesis, we have considered a simple model where the scalar inflaton is coupled to some U(1) gauge field in a way that would be typical for moduli or dilaton-like fields. We have seen that the time dependence of the inflaton condensate during inflation leads to a production of large-scale gauge field fluctuations, analogous to the usual mechanism that amplifies the quantum vacuum fluctuations of the inflaton or gravity wave perturbations. Focusing on the case where the spectrum of produced "magnetic" fields is scale invariant, we have shown that (nearly) local non-Gaussianity is very naturally generated at the level $f_{\rm NL} \gtrsim \mathcal{O}(10)$, which is being probed by the ongoing Planck mission [13]. Although we have neglected slow roll corrections to the running of the spectrum and bispectrum, we can still see a logarithmic running of the effective $f_{\rm NL}$ parameter with scale. This arises since the "sourced" contribution to the curvature perturbation experiences a super-horizon evolution during inflation, due to the presence of large-scale iso-curvature perturbations. Logarithmic running of this type may be of observational interest; see [200, 215, 276] for example. Moreover, since the non-Gaussian part of ζ is uncorrelated with the Gaussian part, we have a non-hierarchical scaling which can lead to interesting signatures in probes that are sensitive to the global structure of the PDF [49].

A novel feature of our result is the dependence of $f_{\rm NL}$ on the $N_{\rm tot} - N_{\rm CMB}$ which measures the number of "extra" e-foldings of inflation, beyond the minimal $N_{\rm CMB} \sim 60$ e-foldings between the end of inflation and horizon exit for CMB scales. Ordinarily, one would expect that such "extra" e-foldings are completely unobservable, since scalar modes which leave the horizon prior to $N_{\rm CMB}$ should just be absorbed into a renormalization of the homogeneous background. However, in the current model the total duration of the quasi de Sitter phase impacts the energy density of produced gauge field [107]. The modes that were produced through the interaction I^2F^2 and left the horizon become classical and generate the scale-invariant spectrum of "magnetic" field density. The modes relevant to the CMB scales, which exit the horizon roughly $N_{\rm CMB} \sim 60$ e-folds before the end of inflation, experience as a classical homogeneous background the modes that left the horizon at earlier stages of inflation [53]. The logarithmic enhancement $N_{\rm tot} - N_{\rm CMB}$ in the spectra (power spectrum in (234) and non-Gaussianity in (248)) is precisely the contribution from the modes that left the horizon during the first $N_{\rm tot} - N_{\rm CMB}$ e-folds of inflation.

The statistically isotropic results we obtained, (234, 248, 253), should be interpreted as the theoretical average over several realizations of the mechanism. A classical vector background breaks isotropy, and thus any single realization is expected to give anisotropic power spectrum and bispectrum that carry some non-trivial angular dependence [53]. One would obtain our results (234, 248, 253) only if one could superpose many realizations of such anisotropic signals. Considering our universe, the CMB modes "see" the homogeneous background modes that left the horizon after a single realization of the first $N_{\rm tot} - N_{\rm CMB}$ e-folds of inflation, and therefore they should be affected by the classical fields that follow the statistics with non-vanishing angular dependence, not by their isotropic average. Interestingly, the level of statistical anisotropy in the power spectrum is associated with a firm prediction for non-Gaussianity; the natural range of statistical anisotropy will likely correspond to the detectable level of local non-Gaussianity, $f_{\rm NL}^{\rm local} \sim \mathcal{O}(10)$ [53].

We have seen that the correlation functions of scalar and tensor cosmological perturbations exhibit a logarithmic time dependence which is related to the produced gauge fluctuations that have contributed to the classical, homogeneous background. As is clear from the above discussions, these logarithms should not be confused with the IR logs that have been discussed extensively in the literature in association with loop effects during inflation [74, 267, 273, 305] (see also [200] for a related discussion). In the case at hand the interpretation of this logarithmic time dependence is straightforward. The production of gauge field fluctuations in our model arises simply because the effective gauge coupling is time dependent. This time dependence leads to a growth in the (scale-invariant) energy density of gauge fluctuations, which is drained from the scalar condensate. The energy transfer provides a physical clock, producing the classical background of the vector field, and its logarithmic growth is a real physical effect that is not related to the de Sitter background or to the quantization of gravitational fluctuations.

Ref. [53] has already extended the study of the model (188) in a very generic inflationary setup. It has included the effects that arise when the vector field has a vacuum expectation value already at 0th order in perturbation theory and their effects on the angular dependence of the power spectrum and bispectrum. On the other hand, it might still be interesting to consider different choices of coupling function, for example $n \neq 2, -2$, and to explore in more detail the non-Gaussian phenomenology of our model, in particular as regards higher moments and scale dependence. Also, it would be worth to investigate whether this mechanism for the amplification of the gauge "magnetic" modes can be consistently modified so to avoid the strong coupling problem of [107], and therefore be used for magnetogenesis.

Part VI

Gravitational Waves and Non-Gaussianity from a Sector Gravitationally Coupled to the Inflaton

One of the particularly important probes of the physics of inflation is provided by primordial gravitational waves (GWs). A number of observational searches for GW perturbations are proposed or are currently underway; these may be probed through the B-mode polarization of the CMB [58, 67, 126] or, on smaller scales, through interferometers such as LIGO [1], VIRGO [3], DECIGO [185], Einstein Telescope [266], or LISA [2].

During inflation, GW fluctuations are inevitably generated by quantum fluctuations of the tensor part of the metric. These have an amplitude controlled by $\frac{H^2}{M_p^2}$, where H is the Hubble scale and $M_p \approx 2.4 \cdot 10^{18}\,\mathrm{GeV}$ is the reduced Planck mass. The GW signal from vacuum fluctuations is detectable only when the inflaton field range is trans-Planckian [218], which might be challenging to realize in a controlled effective field theory.

Additional sources of GWs, which are uncorrelated with the usual quantum vacuum fluctuations, may be present in the early universe (see, for instance, [64] for a recent review). Two broad categories of mechanisms are:

1. **Models involving phase transitions:** For example, in first order phase transitions, vacuum bubble collisions [24, 82, 83, 92, 172, 179, 196, 197, 235, 288, 310] and the subsequent turbulence [81,117,156,195,243] could source GW. The generation and decay of cosmic strings can also give rise to large GW [104, 121, 230, 246]. Along similar lines, the self-ordering of a scalar field after a second order phase transition has also been considered [127, 173, 198, 199].

2. **Models involving particle production:** Generically the inflaton should be expected to couple to some additional degrees of freedom, as would seem to be necessary for successful reheating [37, 71, 189, 190, 252]. In this case, there is a natural possibility that the time-dependence of the inflaton condensate during inflation leads to the production of some other degrees of freedom which may, in turn, provide an important new source of GW [47, 99, 271]. A variety of different models have been proposed.[41]

[41] For GWs from particle production at the *end* of inflation, see [119,120,124,136,137]; effects on the scalar fluctuations were discussed in [38, 40, 68, 86].

Given the importance of GWs as a probe of inflation, it is important to understand if such mechanisms could be competitive with the usual spectrum of GW from vacuum fluctuations. [42]

Any mechanism which is invoked to source GWs might also source scalar metric perturbations. Therefore, one must take care not to spoil the usual prediction of a nearly scale invariant spectrum of Gaussian scalar curvature perturbations. Sometimes this concern is evaded by restricting attention to effects that take place on the small scales relevant for interferometers where the scalar fluctuations are not strongly constrained. Here we will mostly be interested in models of particle production during inflation, where GWs are sourced on CMB scales. To test the feasibility of such scenarios, it is crucial to study also the spectra of the scalar fluctuations, to ensure that these are consistent with observations.

Models of particle production during inflation have received considerable attention in the literature; see for example [22,34,35,42,43,45,47,48,62,94,99,150,158,214,238,260,271,284]. One class of models involves instantaneous bursts of particle production, leading to localized features in the cosmological perturbations. Early studies focused on the production of fermion [94] or scalar [260] particles, neglecting the feed-back of the produced quanta on the perturbations of the inflaton. In Ref. [43] it was however shown that this feed-back effect actually dominates observables for the case of scalar particle production. Ref. [43] considered a simple model where scalar χ particles are produced through a coupling $g^2\varphi^2\chi^2$ to the inflaton, φ. Models of this same type were subsequently analyzed in the context of trapped inflation [150]. More recently, models of instantaneous vector particle production have been studied in connection with GW at interferometer scales [99]. Several scenarios of GW from scalar and string production were instead studied in [271].

Another possibility is that particle production occurs continuously during inflation. This is quite natural in the context of axion inflation [48]. Even in the simplest models of inflation driven by a single axion in slow roll on a smooth flat potential, pseudoscalar couplings $\varphi F\tilde{F}$ to gauge fields are ubiquitous. As shown in Part IV of this thesis, this interaction leads to a continuous tachyonic production of gauge field fluctuations during inflation. There are a host of interesting phenomenological signatures: observable equilateral non-Gaussianity [45, 48, 49], GW at interferometer scales [47, 99], and excess power at small scales [93, 232]. Additionally, backreaction effects in such models can assist inflation by dissipating the kinetic energy of the inflaton [22,23], analogously to warm inflation [62] and trapped inflation

[42] A GW signal may also be left at the largest scales as an imprint of a pre-inflationary era if inflation had only a minimal duration [157].

[150].[43] As in Part V, a model of gauge field production through the scalar coupling $I^2(\varphi)F^2$ was considered in [46, 79, 268], in connection with primordial non-Gaussianity and, perhaps, magnetogenesis (this last application is problematic [46, 107]).

So far we discussed some simple models of particle production during inflation where quanta are produced by a *direct coupling* between the inflaton and some additional degrees of freedom. In such a scenario the produced particles interact with the inflaton through couplings that are typically much stronger than gravitational. Hence, these produced particles will tend to source scalar curvature fluctuations much more efficiently than GW. It is not surprising, therefore, that the most stringent CMB constraints on such models often come from features and non-Gaussianity in the scalar fluctuations, rather than from GW.

In this Part, we investigate the possibility to source a significant GW signal on CMB scales without ruining the spectrum of curvature fluctuations. To minimize the impact on the inflaton fluctuations, we assume that particle production takes place in a "hidden" sector that is only gravitationally coupled to the inflaton.[44] We here compute the scalar perturbations induced by the gravitational coupling, and we compare their phenomenological impact with that of the produced GW. Since gravitational interactions are unavoidable, the case we investigate can be thought of as a kind of "best case" scenario for the production of GW while minimizing the effect on the scalar fluctuations. Concretely, we will focus on two models involving the production of spin-1 particles:

1. **Model I:** Particle production takes place in a hidden sector where a local U(1) invariance is spontaneously broken by the expectation value of some complex scalar ψ. Here gauge fields are produced at an isolated moment - namely, when the mass of the gauge field crosses zero during the evolution of ψ – and we have localized features in the scalar spectrum and bispectrum, in addition to a localized feature in the tensor spectrum. For a single isolated burst of production, we find that observational constraints on the scalar perturbations exclude any interesting effect in GW. However, in a concrete model there may be many bursts of particle production and their resonance could enhance the effect. We note that this model could lead to interesting phenomenological signatures in the scalar fluctuations; localized features in the

[43] The pseudo-scalar interaction is also used to [14] to dissipate the inflaton kinetic energy from its classical interaction to a non-Abelian vector field.

[44] Our scenario differs from the curvaton model [220], in which the inflaton provide the energy density to support inflation while a light curvaton field provides the seed of scalar perturbation. The GW spectrum in the curvaton model has been considered at tree-level [131, 242] and at 1-loop level [54], with possible detectability at future interferometer experiments [54]. See [50, 51, 109–115, 183, 184, 300] for examples of vector curvaton.

spectrum and bispectrum are discussed.

2. **Model II:** Particle production takes place in a hidden sector where a rolling pseudoscalar sources gauge field fluctuations continuously. Here we find that GW from particle production can be competitive with, or even larger than, the vacuum fluctuations, *without* violating observational bounds on non-Gaussianity. In this model the primordial tensor spectrum can be detectable for *any* choice of inflaton potential (that is, also for small field inflation). In this model, GW from particle production are chiral. We show that parity violation in the tensor sector can be almost maximal, which provides a distinctive observable signature of this scenario; see also [23, 284].

In summary: we find that the possibility to source interesting GW from particle production is rather model dependent. Even if particle production occurs in a hidden sector, coupled only gravitationally to the inflationary one, this is not a sufficient condition to ensure that observable GW can be sourced without ruining the scalar spectrum. On the other hand, we find that in some models GW from particle production can actually *exceed* the usual vacuum fluctuations. The reason for the smaller GW production in Model I with respect to Model II is that in Model I the gauge quanta are highly non-relativistic after their production, which suppresses their quadrupole moment. We verified that the same suppression takes place if, instead of gauge fields, the particles produced in this mechanism have spins 0 or $\frac{1}{2}$.[45] As we mentioned, a greater GW effect may be obtained for multiple instances of particle production, or for a more complicated evolution of the gauge field mass, or if the massive particles decay into massless ones short after they are produced [271].

Therefore, due to the possibility of GW from particle production, a measurement of primordial B-modes does not necessarily constitute a measurement of the scale of inflation. Nor does a detectable B-mode signal necessarily require super-Planckian excursions in field space. Fortunately, as we discuss below, the produced GW may be distinguished from the vacuum signal, either by their localization at some given wavelength (assuming that Model I can be modified so to enhance the GW signal), or by their violation of parity.

In terms of observational prospects of GWs on CMB scales, Model II is certainly more promising, as the gauge quanta are automatically relativistic and their quadrupole moment is not suppressed. While the present focus is to examine the signatures and constraints of GWs produced during inflation as a result of hidden sector particle production, the UV

[45] We find that the results agree in all three cases, with only order unity differences coming from counting the number of degrees of freedom and also from spin statistics. These order unity effects can be encoded in a very simple formula, which we present.

sensitivity of inflation motivates us to ask whether such models can be realized in string theory. Indeed, the 4D low energy spectrum of string theory contains a myriad of axion-like fields, e.g., those arise from the reduction of antisymmetric form fields on cycles of the internal space. These closed string axions couple to U(1) gauge fields on the world volume of D-branes via $\varphi F\tilde{F}$ couplings. The axion decay constant f one typically finds in string theory constructions is of the order of the GUT scale $M_{\rm GUT} \sim 10^{16}$ GeV (see, e.g., [30, 291]) which sits comfortably within the allowed window for consistency of our model (see eqs. (356) and (357)). It is not difficult to arrange the inflaton to have no direct coupling to the axion and the gauge field. We outline some ideas to realize Model II in string theory in the concluding section. We leave, however, a detailed study of these and other string theory embeddings for future work.

This Part, based on the study done in [44], is organized as follows. In Section VI 1 we consider Model I, discussed above, utilizing the formalism we developed in Part III to Model I, discussed above. In Section VI 2 we instead consider Model II, discussed above. In Section VI 3, we summarize the results from Models I and II.

1 Model I: Vector Produced by Non-adiabatic Change of its Mass

We consider a model where the sector in which particle production takes place has a local U(1) invariance spontaneously broken by the expectation value of a complex scalar Ψ:

$$S = \int d^4x\sqrt{-g}\left[\frac{M_p^2}{2}R - \underbrace{\frac{1}{2}(\partial\varphi)^2 - V(\varphi)}_{\text{inflaton sector}} - \underbrace{\left|\left(\partial_\mu - ieA_\mu\right)\Psi\right|^2 - U\left(|\Psi|\right) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}}_{\text{hidden sector}}\right]. \tag{267}$$

The field φ is the inflaton field, which is assumed to be only gravitationally coupled to the $\Psi - A_\mu$ sector. We work in the unitary gauge $\Psi = \frac{\psi}{\sqrt{2}}$, where ψ is real. We assume that the background value $\psi^{(0)}(t)$ crosses zero at the time t_* during inflation. Close to this moment, the gauge field mass can be approximated by

$$m = e\,\psi^{(0)} \simeq e\dot{\psi}_*^{(0)}\,(t - t_*) \equiv \dot{m}_*\,(t - t_*) \tag{268}$$

For definiteness, we take $\dot{m}_* > 0$. The (comoving) frequency of a gauge mode, $\omega = \sqrt{k^2 + a^2\,m^2}$ varies nonadiabatically ($\omega' > \omega^2$) when the mass vanishes. As we will show later, the gauge field modes that dominate the observational signatures have $k \sim \sqrt{\dot{m}_*}$. For such modes, the frequency changes nonadiabatically in the time interval $\sim t_* \pm 1/\sqrt{\dot{m}_*}$, provided that the expansion of the universe can be disregarded in this time interval. This is the case for $\frac{1}{\sqrt{\dot{m}_*}} \ll \frac{1}{H}$. In the remainder of this Subsection we concentrate on the production during this interval, and disregard the expansion of the universe (we use physical and conformal time interchangeably). We can eliminate any reference to the scale factor by normalizing $a(t_*) = 1$. In the following Subsections the expansion is taken into account.

The non-adiabatic change of the frequency causes non-perturbative production of the gauge modes. To compute this, we decompose the vector field as in (48),

$$A_i(\tau, \vec{x}) = \int \frac{d^3k}{(2\pi)^{3/2}} \, \mathrm{e}^{i\vec{k}\cdot\vec{x}} \, \tilde{A}_i\left(\tau, \vec{k}\right)$$
$$\tilde{A}_i\left(\tau, \vec{k}\right) = \sum_{\lambda=\pm} \epsilon_i^{(\lambda)}(\hat{k}) \left[a_\lambda\left(\vec{k}\right) A_\lambda(\tau, k) + a_\lambda^\dagger\left(-\vec{k}\right) A_\lambda^*(\tau, k)\right] . \tag{269}$$

We further decompose the mode functions in positive and negative frequency modes

$$A_\lambda(k) = \alpha_k(\tau) f_k(\tau) + \beta_k(\tau) f_k^*(\tau) \quad , \quad f_k \equiv \frac{\mathrm{e}^{-i\int^\tau d\tau' \, \omega(\tau')}}{\sqrt{2\omega(\tau)}}$$
$$A'_\lambda(k) = -i\omega\left[\alpha_k(\tau) f_k(\tau) - \beta_k(\tau) f_k^*(\tau)\right] \tag{270}$$

(the second line is the decomposition of the modes of the conjugate momentum to A_i). The decomposition (269) disregards the longitudinal vector mode; we compute this mode in Appendix E, and we discuss its effects below in this section. We have suppressed the index λ in this decomposition, since the Bogolyubov coefficients are the same for both helicities. For $t \ll t_*$, the mode is in the adiabatic vacuum $\alpha = 1$ (up to an arbitrary phase); the quantity $|\beta|^2$ is the occupation number of the gauge modes.

The gauge field modes satisfy $A'' + \omega^2 A = 0$. With the approximated expression (268), this equation can be analytically solved in terms of two parabolic cylinder functions. The linear combination satisfying the proper initial condition $\alpha = 1$ gives (up to an arbitrary phase) [99, 190]

$$\alpha_k(t \gg t_*) \simeq \sqrt{1 + \mathrm{e}^{-\frac{\pi k^2}{\dot{m}_*}}} \quad , \quad \beta_k(t \gg t_*) \simeq -\mathrm{e}^{-\frac{\pi k^2}{2\,\dot{m}_*}} \tag{271}$$

Namely, one finds an $\mathcal{O}(1)$ occupation number for modes up to $k \simeq \sqrt{\dot{m}_*}$, while the production is exponentially suppressed at higher momenta. An analogous result is obtained for scalar [190] and fermion [94] fields produced by a mass varying as in (268). This result is valid provided that

$$H^2 \ll \dot{m}_* \ll \sqrt{6} e M_p H \tag{272}$$

where (as we mentioned previously) the first condition ensures that the expansion of the universe can be disregarded during the interval of particle production, while the second condition imposes that the kinetic energy of ψ is negligible with respect to the inflaton energy density.

Having the mode functions, we can now compute the correlators (65), $\langle \tilde{A}_i\left(\tau_1, \vec{q}_1\right)\, \tilde{A}_j\left(\tau_2, \vec{q}_2\right)\rangle$. The correlators need to be regularized. The regularization can be performed by normal ordering with respect to the time dependent annihilation creation operators

$$\begin{aligned} \bar{a}_\lambda\left(\tau,\, \vec{k}\right) &\equiv \alpha_k\left(\tau\right) a_\lambda\left(\vec{k}\right) + \beta_k^*\left(\tau\right) a_\lambda^\dagger\left(-\vec{k}\right) \\ \bar{a}_\lambda^\dagger\left(\tau,\, -\vec{k}\right) &\equiv \beta_k\left(\tau\right) a_\lambda\left(\vec{k}\right) + \alpha_k^*\left(\tau\right) a_\lambda^\dagger\left(-\vec{k}\right) \end{aligned} \tag{273}$$

These are the "physical" operators of the system, as one can show that they diagonalize the Hamiltonian at all times. Using (269) and (270) it is immediate to show that, in terms of these operators,

$$\tilde{A}_i\left(\tau,\, \vec{k}\right) = \sum_\lambda \epsilon_i^{(\lambda)}\left(\hat{k}\right)\left[f_k\left(\tau\right)\bar{a}_\lambda\left(\tau,\, \vec{k}\right) + f_k^*\left(\tau\right)\bar{a}_\lambda^\dagger\left(\tau,\, -\vec{k}\right)\right] \tag{274}$$

We then compute $\langle :\, \tilde{A}_i(\tau,\, \vec{k})\tilde{A}_i(\tau',\, \vec{k}')\, :\rangle$ by normal ordering with respect to the $\bar{a}$, and $\bar{a}^\dagger$, but by recalling that the vacuum state is annihilated by the original time-independent a (we are working in the Heisenberg picture, in which the states are constant).[46] Proceeding in this way, and casting the result as in (65), we obtain

$$\begin{aligned} \mathcal{D}_{(0,0)}^{(\sigma)}\left[\tau,\, \tau';\, k\right] \;=\;& \alpha_k\left(\tau\right)\beta_k^*\left(\tau'\right) f_k\left(\tau\right) f_k\left(\tau'\right) + \beta_k\left(\tau\right)\alpha_k^*\left(\tau'\right) f_k^*\left(\tau\right) f_k^*\left(\tau'\right) \\ & + \beta_k\left(\tau\right)\beta_k^*\left(\tau'\right) f_k^*\left(\tau\right) f_k\left(\tau'\right) + \beta_k^*\left(\tau\right)\beta_k\left(\tau'\right) f_k\left(\tau\right) f_k^*\left(\tau'\right) \end{aligned} \tag{275}$$

Clearly, the result is σ-independent, since both helicities are produced in the same

[46] This procedure is conventionally adopted in Bogolyubov computations, although it is not always spelled out; indeed, the initial Hamiltonian, without normal ordering, can be cast in the form $\hat{H} = \int d^3k\, \omega_k\, \hat{N}$, where the counting operator is $\hat{N} = \frac{\bar{a}\bar{a}^\dagger + \bar{a}^\dagger\bar{a}}{2}$. One then defines normal ordering with respect to the barred operators, $:\hat{N} := \bar{a}^\dagger\, \bar{a}$, and evaluates $\langle :\, \hat{N}\, :\rangle$ using (273) and recalling that the vacuum is annihilated by the original time-independent a; only in this way one obtains $\langle :\, \hat{N}\, :\rangle = |\beta|^2$. We employ the exact same procedure for evaluating the correlators $\mathcal{D}$.

amount. After the particle production, this expression simplifies into

$$\begin{aligned} \mathcal{D}^{(\sigma)}_{(0,0)}\left[\tau,\,\tau';\,k\right] &= \alpha_k\,\beta_k^*\,f_k\left(\tau\right)f_k\left(\tau'\right)+\alpha_k^*\,\beta_k\,f_k^*\left(\tau\right)f_k^*\left(\tau'\right)\\ &\quad+\left|\beta_k\right|^2\left[f_k^*\left(\tau\right)f_k\left(\tau'\right)+f_k\left(\tau\right)f_k^*\left(\tau'\right)\right] \end{aligned} \tag{276}$$

where α and β assume the asymptotic values (271). As we anticipated, this expression is real and symmetric under $\tau \leftrightarrow \tau'$.

The same result (276) is obtained if one computes $\langle \tilde{A}_i(\tau,\,\vec{k})\,\tilde{A}_i(\tau',\,\vec{k}')\rangle$ without normal ordering, and then subtracts the term that one would have in absence of particle production ($\alpha = 1,\ \beta = 0$), as done in [43]. Therefore these two regularizations are equivalent after the particle production has taken place.

For further convenience, we extend the definition (65) to

$$\left\langle :\left(\frac{\partial}{\partial\tau_1}\right)^a\tilde{A}_i\left(\tau_1,\,\vec{q}_1\right)\left(\frac{\partial}{\partial\tau_2}\right)^b\tilde{A}_j\left(\tau_2,\,\vec{q}_2\right):\right\rangle \equiv \sum_\sigma \mathcal{P}^{(\sigma)}_{ij}\left(\hat{q}_1\right)\mathcal{D}^{(\sigma)}_{(a,b)}\left[\tau_1,\,\tau_2;\,q_1\right]\,\delta^{(3)}\left(\vec{q}_1+\vec{q}_2\right) \tag{277}$$

We compute these quantities by using (270). After α and β have assumed the asymptotic values (271), the correlators of our interest become

$$\begin{aligned} \mathcal{D}^{(\sigma)}_{(1,0)}\left[\tau,\,\tau';\,k\right] &= -i\omega_k\left(\tau\right)\Big\{\alpha_k\,\beta_k^*\,f_k\left(\tau\right)f_k\left(\tau'\right)-\alpha_k^*\,\beta_k\,f_k^*\left(\tau\right)f_k^*\left(\tau'\right)\\ &\quad+\left|\beta_k\right|^2\left[-f_k^*\left(\tau\right)f_k\left(\tau'\right)+f_k\left(\tau\right)f_k^*\left(\tau'\right)\right]\Big\}\\ \mathcal{D}^{(\sigma)}_{(0,1)}\left[\tau,\,\tau';\,k\right] &= -i\omega_k\left(\tau'\right)\Big\{\alpha_k\,\beta_k^*\,f_k\left(\tau\right)f_k\left(\tau'\right)-\alpha_k^*\,\beta_k\,f_k^*\left(\tau\right)f_k^*\left(\tau'\right)\\ &\quad+\left|\beta_k\right|^2\left[f_k^*\left(\tau\right)f_k\left(\tau'\right)-f_k\left(\tau\right)f_k^*\left(\tau'\right)\right]\Big\}\\ \mathcal{D}^{(\sigma)}_{(1,1)}\left[\tau,\,\tau';\,k\right] &= -\omega_k\left(\tau\right)\,\omega_k\left(\tau'\right)\Big\{\alpha_k\,\beta_k^*\,f_k\left(\tau\right)f_k\left(\tau'\right)+\alpha_k^*\,\beta_k\,f_k^*\left(\tau\right)f_k^*\left(\tau'\right)\\ &\quad-\left|\beta_k\right|^2\left[f_k^*\left(\tau\right)f_k\left(\tau'\right)+f_k\left(\tau\right)f_k^*\left(\tau'\right)\right]\Big\} \end{aligned} \tag{278}$$

We note that in the regime in which (276) and (278) are valid, $\partial_\tau\mathcal{D}^{(\sigma)}_{(0,0)}\left[\tau,\tau';k\right]=\mathcal{D}^{(\sigma)}_{(1,0)}\left[\tau,\tau';k\right]$, and analogously for the other terms. Namely, in this regime, time derivatives can be equivalently taken before or after evaluating the correlator. One can verify that (for the regularized correlators) this is not the case during the particle production, when α and β are still functions of time.

1.1 Scalar Perturbations Sourced by the Vector Modes

We are interested in the phenomenological consequences of the vector field production in the model (267). As mentioned in the Introduction, we work under the assumption that the $\Psi - A_\mu$ sector has a negligible energy density with respect to the inflationary sector, and that the two sectors are coupled only gravitationally. Under these assumptions, the main signature of the particle production is encoded in how the gauge quanta enter in the gravitational equations. In this subsection we study the effect on the scalar metric and inflaton perturbations; the effect on the tensor modes is computed in the next subsection.

We divide the computation in 3 parts; in the first one, we present the master equation for the density perturbation ζ; in the second and the third part we compute, respectively, the power spectrum and the bispectrum of the part of ζ sourced by the vector quanta.

1.1.1 Master Equation for ζ

We perform computations in the spatially flat gauge $\delta g_{ij,\text{scalar}} = 0$. We then decompose the fields into background + first order + second order perturbations; for the inflaton field we write

$$\varphi(t, \vec{x}) = \varphi^{(0)}(t) + \delta_1\varphi(t, \vec{x}) + \delta_2\varphi(t, \vec{x}) + \ldots \tag{279}$$

(dots denotes perturbations of higher order, that we ignore). Under our working assumption that Ψ gives a negligible contribution both to the background energy density, and to the cosmological perturbations (which is certainly the case, if ρ_Ψ is sufficiently small), the curvature perturbations ζ in this gauge is

$$\zeta = -\frac{H}{\dot{\varphi}^{(0)}}\,\left(\delta_1\varphi + \delta_2\varphi\right)\;. \tag{280}$$

The master equation for $\delta_i\varphi$, and thus for ζ, is derived in Appendix A, with the metric perturbations consistently included. The master equation at first order in perturbation theory reproduces the standard first-order expression

$$\left(\partial_\tau^2 - \nabla^2\right)\delta_1\varphi + 2\mathcal{H}\delta_1\varphi' \simeq 0 \tag{281}$$

where $\nabla^2 \equiv \partial_i\partial_i$ and $\mathcal{H} \equiv \frac{a'}{a}$. The approximation symbol arises because on the left-hand side of this expression we have disregarded a "mass term" (see (A.8) in Appendix A) for

$\delta_1\varphi$ which is proportional to the slow roll parameters (the same is true for eq. (283))

$$\epsilon \equiv \frac{M_p^2}{2}\left(\frac{V_\varphi}{V}\right)^2 \ll 1 \ , \ \ \eta \equiv M_p^2 \frac{V_{\varphi\varphi}}{V} \ll 1 \tag{282}$$

The reason for this is that we are not interested in slow roll corrections to our results.

At next order in perturbation theory we find in Appendix A.4

$$\left(\partial_\tau^2 - \nabla^2\right)\delta_2\varphi + 2\mathcal{H}\delta_2\varphi' \simeq -\frac{a^2\varphi^{(0)\prime}}{2M_p^2\mathcal{H}}\left\{\frac{E^2+B^2}{2} + \frac{1}{a^4}\nabla^{-2}\partial_\tau\left[a^4\vec{\nabla}\cdot\left(\vec{E}\times\vec{B}\right)\right] + \frac{m^2}{2a^2}A_iA_i\right\} + \ldots \tag{283}$$

Dots denote terms which are proportional to squares of the first order inflaton and tensor metric perturbations, namely they are of $\mathcal{O}\left(\delta_1\varphi^2\right)$ and of $\mathcal{O}\left(h_{1,ij}^2\right)$ (we recall that $\delta\Psi$ can be disregarded); these terms are explicitly given in [225]. The "electric" and "magnetic" fileds are defined in analogy to the electromagnetic expressions, namely

$$E_i = -\frac{1}{a^2}A_i' \ , \ \ B_i = \frac{1}{a^2}\epsilon_{ijk}\partial_j A_k \tag{284}$$

A more detailed derivation of (283) is shown in Appendix A.4, where the slow-roll suppressed terms are also included. The expression (283) should be compared with eq. (95) of Part IV. Besides neglecting the slow-roll suppressed inflaton "mass term" in (283), the major difference is that since the direct inflaton-gauge coupling is absent in the current model, the dominant source is replaced by the one from the Planck-suppressed gravitational interaction. [47]

1.1.2 Scalar Source and Power Spectrum

Identifying the decompositions (279) together with (49),

$$\delta\varphi\left(\tau,\vec{x}\right) = \int \frac{d^3k}{(2\pi)^{3/2}}\,\mathrm{e}^{i\vec{k}\cdot\vec{x}}\,\frac{Q_\varphi\left(\tau,\vec{k}\right)}{a(\tau)} \ , \tag{285}$$

[47] This provides the main difference between the scalar field cosmological perturbations obtained in this Part and in Part IV; the computations of Part IV were done under the working assumption that the direct interaction is stronger than the gravitational coupling, so that all the other terms were disregarded. In this Part VI, we instead compute the more complicated contribution of the gravitational interactions.

the two equations (281) and (283) can then be combined into a unique linear non-homogeneous differential equation for Q_φ, which is formally of the type (50),

$$\left[\partial_\tau^2 + \left(k^2 - \frac{a''}{a}\right)\right] Q_\varphi \simeq J_\varphi \tag{286}$$

in terms of the two sources

$$J_\varphi\left[A_\mu^2\right] + \tilde{J}_\varphi\left[\delta_1\varphi^2,\, h_{1,ij}^2\right] \tag{287}$$

Namely, the source J_φ is obtained from the terms explicitly written on right hand side of (283), while $\tilde{J}_\varphi$ is obtained from the terms denoted with dots. The two sources are uncorrelated, and therefore the two particular solutions sourced by them can be obtained independently. The solution sourced by $\tilde{J}_\varphi$ is the one emerging from cosmological second order perturbation theory in absence of gauge fields. For standard scalar field slow roll inflation, this term provides a negligible contribution to the power spectrum, and unobservable non-Gaussianity. Therefore, we disregard this term in our calculations. In Appendix A.4 we show that

$$J_\varphi \simeq \frac{\dot{\varphi}^{(0)}}{2M_p^2 H a} \int \frac{d^3p}{(2\pi)^{3/2}} \left[\hat{k}_i\hat{k}_j\left(M^2 - \partial_\tau^{(1)}\,\partial_\tau^{(2)}\right) - M^2\,\delta_{ij}\right] \tilde{A}_i\left(\tau,\,\vec{p}\right) \tilde{A}_j\left(\tau,\,\vec{k}-\vec{p}\right) \tag{288}$$

where $M \equiv a\,m = a\,e\,\psi^{(0)}$ and we are using the notation

$$\partial_\tau^{(1)}\,\partial_\tau^{(2)}\; f\; g \equiv \partial_\tau f\;\partial_\tau g \tag{289}$$

It is worth noting that the source does not diverge as $k \to 0$; this is not immediate from (283), since one term contains an inverse laplacian. See Appendix A.4 for details.

From (288) we see that, for this model, the source is formally of the type (51), with

$$\hat{O}_{\varphi,ij}\left(\tau,\,\vec{k},\,\vec{p}\right) = \frac{\dot{\varphi}^{(0)}}{2M_p^2 H a}\left[\hat{k}_i\hat{k}_j\left(M^2 - \partial_\tau^{(1)}\,\partial_\tau^{(2)}\right) - M^2\,\delta_{ij}\right] \tag{290}$$

We insert this operator in (71); we also use the fact that $\mathcal{D}_{(0,0)}^{(\sigma)}$ is actually helicity independent in this model, so that the sum over the helicity σ is limited to $\sum_\sigma \mathcal{P}_{im}^{(\sigma)}\left(\hat{p}\right) = \delta_{im} - \hat{p}_i\,\hat{p}_m$ (see (66)) - and analogously for the sum over σ'. Moreover, in the integrand we approximate $\vec{p}-\vec{k}\simeq\vec{p}$.[48] We obtain

$$\begin{aligned} P_\zeta^{\rm sourced}(k) \simeq & \frac{k^3}{4\pi^2 a^2 M_p^4}\int_{\tau_*}^{\tau} d\tau_1 \frac{G_k(\tau,\,\tau_1)}{a(\tau_1)}\int_{\tau_*}^{\tau} d\tau_2 \frac{G_k(\tau,\,\tau_2)}{a(\tau_2)} \\ & \times \int \frac{d^3p}{(2\pi)^3}\Bigg\{\left[1+\left(\hat{k}\cdot\hat{p}\right)^4\right] M^2(\tau_1)\, M^2(\tau_2)\,\delta_0^a\,\delta_0^b + \left[1-\left(\hat{k}\cdot\hat{p}\right)^2\right]^2 \delta_1^a\,\delta_1^b \\ & + \left(\hat{k}\cdot\hat{p}\right)^2\left[1-\left(\hat{k}\cdot\hat{p}\right)^2\right]\left[M^2(\tau_1)\,\delta_0^a\,\delta_1^b + M^2(\tau_2)\,\delta_1^a\,\delta_0^b\right]\Bigg\}\mathcal{D}_{(a,b)}\left[\tau_1,\,\tau_2;\,p\right]^2 \end{aligned} \tag{291}$$

where we have omitted the suffix (σ) from the correlators (since they are σ-independent in this model). By inserting the expressions (276) and (278) for the correlators, and by performing the angular integrals, we obtain

$$\begin{aligned} P_\zeta^{\rm sourced}(k) \simeq & \frac{k^3}{4\pi^2 a^2 M_p^4}\int_{\tau_*}^{\tau} d\tau_1 \frac{G_k(\tau,\,\tau_1)}{a(\tau_1)}\int_{\tau_*}^{\tau} d\tau_2 \frac{G_k(\tau,\,\tau_2)}{a(\tau_2)}\,\frac{M(\tau_1)\,M(\tau_2)}{2\pi^2} \\ & \times\int dp\,p^2\Bigg\{|\beta_p|^2\left[|\alpha_p|^2+|\beta_p|^2\right]+\frac{2|\beta_p|^2}{3}\,{\rm Re}\left[\alpha_p^*\beta_p\left({\rm e}^{i\gamma(\tau_1)}+{\rm e}^{i\gamma(\tau_2)}\right)\right] \\ & +\frac{11}{15}\,{\rm Re}\left[\alpha_p^{*2}\beta_p^2{\rm e}^{i[\gamma(\tau_1)+\gamma(\tau_2)]}+|\beta_p|^4{\rm e}^{i[\gamma(\tau_1)-\gamma(\tau_2)]}\right]\Bigg\} \end{aligned} \tag{292}$$

where $\gamma(\tau)\equiv 2\int_{\tau_*}^{\tau} d\tau'\,M(\tau')$, and where the Bogolyubov coefficients are evaluated after the particle production, and given in (271). To obtain this expression we have disregarded p with respect to M inside the mode functions f_p present inside the correlators. This is appropriate since, as we shall see, the integrand is peaked at $p\sim\sqrt{\dot{m}_*}\ll M$.

The curly parenthesis in (292) contains several terms proportional to the fast oscillating phase ${\rm e}^{i\gamma}$. All these terms give a negligible contribution to the final result. To see this, we

[48]The reason for this is that, as we shall see, the signal from particle production is maximal at $k\sim H$ (namely, for modes of the size of the horizon when particle production occurs; recall that the scale factor is normalized to one at the moment of particle production). We shall also see that the source integrand is peaked at $p\sim\sqrt{\dot{m}_*}$. We therefore have $k\ll p$ due to (272).

can separate the momentum and time integrals in this expression, and obtain

$$
\begin{aligned}
P_\zeta^{\rm sourced}(k) \;\simeq\; & \frac{k^3}{8\pi^4 a^2 M_p^4} \int dp p^2 \Big\{ |\beta_p|^2 \left[|\alpha_p|^2 + |\beta_p|^2 \right] \mathcal{T}_k^2 \\
& + \frac{4}{3} |\beta_p|^2 \mathcal{T}_k \, {\rm Re}\left(\alpha_p^* \beta_p \mathcal{E}_k \right) + \frac{11}{15} |\beta_p|^4 |\mathcal{E}_k|^2 + \frac{11}{15} {\rm Re}\left(\alpha_p^{*2} \beta_p^2 \mathcal{E}_k^2 \right) \Big\}
\end{aligned}
\tag{293}
$$

where

$$
\mathcal{T}_k \equiv \int_{\tau_*}^{\tau} d\tau' G_k\left(\tau, \tau'\right) m\left(\tau'\right) \;, \quad \mathcal{E}_k \equiv \int_{\tau_*}^{\tau} d\tau' G_k\left(\tau, \tau'\right) m\left(\tau'\right) {\rm e}^{2i \int_{\tau_*}^{\tau'} d\tau'' M(\tau'')} \tag{294}
$$

Using the expression (61) for the Green function and the expression

$$
m(\tau) = \dot{m}_* \left(t - t_*\right) = \frac{\dot{m}_*}{H} \ln\left(\frac{-\tau_*}{-\tau}\right) = -\frac{\dot{m}_*}{H} \ln\left(-H\tau\right) \tag{295}
$$

(in the last equality we have used the fact that $a(\tau_*) = 1$) we find, in the super-horizon $-k\tau \ll 1$ limit,

$$
\mathcal{T}_k \simeq \frac{a(\tau)\, \dot{m}_*}{27\, H^3}\, {}_2F_3\left(\frac{3}{2}, \frac{3}{2}; \frac{5}{2}, \frac{5}{2}, \frac{5}{2}; \frac{-k^2}{4H^2}\right) \tag{296}
$$

where the generalized hypergeometric function evaluates to ~ 1 for $k \ll H$ and to $\sim \frac{27\pi}{2} \frac{H^3}{k^3} \left(\ln\frac{k}{H} - 0.423\right)$ for $k \gg H$ (in practice, we introduce the integration variable $y' = -k\tau'$, and we integrate it from 0 to $-k\tau_*$, rather than from $-k\tau$ to $-k\tau_*$).

The time integral in the exponent of $\mathcal{E}$ can be done analytically, and, leads to

$$
\begin{aligned}
\mathcal{E}_k \;=\; & \frac{a(\tau)\, \dot{m}_*}{k^3} \int_{-H\tau}^{1} \frac{dy}{y} \left[\frac{k}{H} y \cos\left(\frac{k}{H} y\right) - \sin\left(\frac{k}{H} y\right) \right] \ln(y) \\
& \times \left\{ \cos\left[\frac{\dot{m}_*}{H^2} \ln^2 y\right] + i \sin\left[\frac{\dot{m}_*}{H^2} \ln^2 y\right] \right\}
\end{aligned}
\tag{297}
$$

Since $\dot{m}_* \gg H$, the term in curly parenthesis is rapidly oscillating unless $y \equiv \frac{\tau'}{\tau_*} \simeq 1$. This however means that the integrand is peaked at the moment in which the particle production is taking place. In our computation we used the asymptotic values (271) for α, β (valid once particle production has completed). Therefore we only provide an upper bound on $\mathcal{E}_k$. It is easy to verify that instead the integrand of $\mathcal{T}_k$ is dominated by times at which (271) hold.

Provided that $k \ll \sqrt{\dot{m}_*}$ (which is true in the range of our interest, since the occupation number of the gauge field is exponentially suppressed in the opposite regime), we can keep in the integral only the fast oscillating term, and the log term, and obtain the estimate

$$\begin{aligned} |\mathcal{E}_k| &< \left| \frac{a(\tau)\,\dot{m}_*}{k^3} \left[\frac{k}{H} \cos\left(\frac{k}{H}\right) - \sin\left(\frac{k}{H}\right)\right] \right. \\ &\qquad \left. \times \int_{-H\tau}^{1} dy \, \ln(y) \left\{ \cos\left[\frac{\dot{m}_*}{H^2} \ln^2 y\right] + i \sin\left[\frac{\dot{m}_*}{H^2} \ln^2 y\right]\right\} \right| \\ &\simeq \frac{a(\tau)\,H^2}{2k^3} \left| \frac{k}{H} \cos\frac{k}{H} - \sin\frac{k}{H} \right| \end{aligned} \tag{298}$$

for the upper bound. We indeed see that $\frac{|\mathcal{E}_k|}{\mathcal{T}_k} < \frac{H^2}{\dot{m}_*} \ll 1$.

This confirms that the oscillatory terms in (292), or (293), provide a negligible contribution to the final result. Performing the momentum integral we finally obtain

$$P_\zeta^{\rm sourced}(k) \simeq \frac{2+\sqrt{2}}{2^6\,3^6\,\pi^5} \frac{k^3\,\dot{m}_*^{7/2}}{H^6 M_p^4} \left[{}_2F_3\left(\frac{3}{2}, \frac{3}{2}; \frac{5}{2}, \frac{5}{2}, \frac{5}{2}; \frac{-k^2}{4H^2}\right)\right]^2 \,, \quad k \ll \sqrt{\dot{m}_*} \tag{299}$$

while the result is exponentially suppressed and uninteresting at larger momenta. The exponential suppression of the spectrum is due to the fact that the occupation numbers of the vector particles are exponentially suppressed at such large momenta, see eq. (271). We discuss this result in Subsection VI 1.3.

1.1.3 Scalar Bispectrum

We now evaluate the formal expression (76) for the bispectrum in this model. This expression contains various intermediate quantities that have been given above. Explicitly, we write the green functions in eq. (61), the projection operators in eq. (66), the operator $\hat{O}_\varphi$ in eq. (290), and the correlators in eqs. (276) and (278). The computation follows the same steps presented in the previous Subsection for P_ζ. Also in this case we use the approximation $k \ll p \ll M$, and we find that the terms with a fast oscillating phase in the final time integral can be disregarded. We obtain

$$\begin{aligned} B_\zeta(k_1, k_2, k_3) &\simeq \frac{1}{2M_p^6 a(\tau)^3} \mathcal{T}_{k_1} \mathcal{T}_{k_2} \mathcal{T}_{k_3} \int \frac{d^3p}{(2\pi)^{9/2}} |\beta|^4 \left(3|\alpha|^2 + |\beta|^2\right) \\ &\simeq \frac{27+8\sqrt{6}}{22{,}674{,}816\,\pi^{9/2}} \frac{\dot{m}_*^{9/2}}{H^9 M_p^6} \prod_{i=1}^{3} {}_2F_3\left(\frac{3}{2}, \frac{3}{2}; \frac{5}{2}, \frac{5}{2}, \frac{5}{2}; \frac{-k_i^2}{4H^2}\right) \end{aligned} \tag{300}$$

where we are interested in the modes $k_i \ll \sqrt{\dot{m}_*}$.

This corresponds to the effective equilateral nonlinear parameter (77)

$$f_{\rm NL}^{\rm equil.eff.}(k) \simeq 2.1 \cdot 10^7 \frac{k^6\, \dot{m}_*^{9/2}}{H^9 M_p^6} \left[{}_2F_3 \left(\frac{3}{2}, \frac{3}{2}; \frac{5}{2}, \frac{5}{2}, \frac{5}{2}; \frac{-k^2}{4H^2} \right) \right]^3 , \quad k \ll \sqrt{\dot{m}_*} \tag{301}$$

where we have used the numerical value $P_\zeta \simeq 2.5 \cdot 10^{-9}$ for the power spectrum. Analogous to the sourced part of the power spectrum (299), the bispectrum and the nonlinear parameter are exponentially suppressed and uninteresting at larger momenta.

1.2 Gravitational waves sourced by the vector modes

We now compute the amount of GWs sourced by the produced gauge fields. Inserting the energy momentum tensor of the gauge field

$$T_{\mu\nu,{\rm gauge}} = F_{\mu\alpha}\, F_\nu{}^\alpha + m^2\, A_\mu A_\nu + g_{\mu\nu} \left(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{m^2}{2} A_\mu A^\mu \right) \tag{302}$$

for this model into the general expression (54), we see that the source is formally of the type (55) with

$$\hat{O}_{\lambda,ij} = \frac{\Pi_{mn}^{(\lambda)*}\left(\hat{k}\right)}{a M_p} \left[\delta_{mi}\delta_{nj} \left(-\partial_\tau^{(1)} \partial_\tau^{(2)} + M^2 \right) + \epsilon_{mai}\epsilon_{nbj} p_a \left(k-p\right)_b \right] \tag{303}$$

where we recall our notation (289).

It is instructive to compare this expression with the analogous operator in the source of the scalar perturbations, given in (290). In writing (290) we disregarded terms proportional to spatial momenta $\vec{p}$ and $\vec{p} - \vec{k} \simeq \vec{p}$ with respect to time derivatives and the mass M. The reason for this is that $k \ll p \ll M$, as we explained after (292). As a consequence, time derivatives acting on a mode also give $\partial_\tau A \sim MA \gg pA$, and should be retained in the operator, when compared to spatial momenta. Disregarding the momentum terms is correct for the scalar source. However, we see that in (303) the time derivatives enter with an opposite sign to M^2. As a consequence, the dominant contributions from the first two terms in (303) cancel against each other, and therefore we need to keep the complete structure. We now show explicitly how the cancellation arises.

We insert the operator (303) in the formal expression (73) for the power in the tensor modes. The other intermediate quantities entering in (73) are the Green functions, given in

(61), the projection operators, given in (66), and the correlators, given in (276) and (278). We obtain [49]

$$\begin{aligned} P_\lambda^{\rm sourced} &= \frac{2k^3}{3\pi^4 a^2 M_p^4} \int_{\tau_*}^{\tau} d\tau_1 \frac{G_k(\tau,\,\tau_1)}{a(\tau_1)} \int_{\tau_*}^{\tau} d\tau_2 \frac{G_k(\tau,\,\tau_2)}{a(\tau_2)} \\ & \times \int dp\, p^2 \left\{ \frac{7}{5} \left[M(\tau_1)^2 \delta_{a0} - \delta_{a1} \right] \left[M(\tau_2)^2 \delta_{b0} - \delta_{b1} \right] + \frac{7}{5} p^4 \delta_{a0}\delta_{b0} \right. \\ & \left. + p^2\, \delta_{b0} \left[M(\tau_1)^2 \delta_{a0} - \delta_{a1} \right] + p^2\, \delta_{a0} \left[M(\tau_2)^2 \delta_{b0} - \delta_{b1} \right] \right\} \mathcal{D}_{(a,b)} [\tau_1,\, \tau_2;\, p]^2 \end{aligned} \quad (304)$$

where we have disregarded k as compared to p.

This expression is analogous to eq. (291) for P_ζ, with the difference that here we have already performed the trivial angular integrals of d^3p. Terms without p in the curly parenthesis are obtained from the square of the first term $\propto \left(-\partial_\tau^{(1)}\partial_\tau^{(2)} + M^2\right)$ in (303) (recall that two $\hat{O}_\lambda$ enter in P_λ) . We note that the structure $-\partial_\tau^{(1)}\partial_\tau^{(2)} + M^2$ is preserved in (304) since the suffix 0 (1) on $\mathcal{D}$ indicates that $\tilde{A}$ ($\tilde{A}'$) is present in the correlator. The term $\propto p^4$ in the curly parenthesis is obtained from square of the other term in (303). The terms $\propto p^2$ are the mixed terms. The different coefficients ($\frac{7}{5}$ vs 1) follow from the angular integrals.

The result (304) is a sum of squares of correlators. In each square, most terms present fast oscillating phases; as we shall see, these terms give a negligible contribution to the final result once the time integrals are performed. This is analogous to $|\mathcal{T}| \gg |\mathcal{E}|$ in eq. (294). Squaring the expressions (276) and (278) we obtain

$$\begin{aligned} \mathcal{D}_{(0,0)} [\tau_1,\, \tau_2;\, p]^2 &= \frac{|\beta_p|^2 \left(|\alpha_p|^2 + |\beta_p|^2\right)}{2\omega_p(\tau_1)\,\omega_p(\tau_2)} + \text{oscillatory phases} \\ \mathcal{D}_{(1,0)} [\tau_1,\, \tau_2;\, p]^2 &= \omega_p^2(\tau_1)\, \mathcal{D}_{(0,0)} [\tau_1,\, \tau_2;\, p]^2 + \text{oscillatory phases} \\ \mathcal{D}_{(0,1)} [\tau_1,\, \tau_2;\, p]^2 &= \omega_p^2(\tau_2)\, \mathcal{D}_{(0,0)} [\tau_1,\, \tau_2;\, p]^2 + \text{oscillatory phases} \\ \mathcal{D}_{(1,1)} [\tau_1,\, \tau_2;\, p]^2 &= \omega_p^2(\tau_1)\,\omega_p^2(\tau_2)\, \mathcal{D}_{(0,0)} [\tau_1,\, \tau_2;\, p]^2 + \text{oscillatory phases} \end{aligned} \quad (305)$$

We recall that $\omega_p(\tau_i) = \sqrt{M(\tau_i)^2 + p^2}$, and that momentum integrand has its support in the region $p \ll M(\tau_i)$. We see, however, that the O $\left(M^4\right)$ and O $\left(M^2\right)$ parts of the curly parenthesis in (304) cancel for the non-oscillatory contributions (305).

[49] We remind that the tensor power is obtained from $\langle h_\lambda h_\lambda \rangle$. We have verified that $\langle h_\lambda h_{\lambda'} \rangle \propto \delta_{\lambda\lambda'}$.

Using the full expressions (276) and (278) for the correlators, we obtain

$$\begin{aligned}P_\lambda^{\rm sourced} = \frac{2k^3}{3\pi^4 a^2 M_p^4}\int_{\tau_*}^{\tau} d\tau_1 \frac{G_k(\tau,\,\tau_1)}{a(\tau_1)}\int_{\tau_*}^{\tau} d\tau_2 \frac{G_k(\tau,\,\tau_2)}{a(\tau_2)}\int dp\, \frac{p^2}{\omega_p(\tau_1)\,\omega_p(\tau_2)} \\ \times\Bigg\{\frac{2p^4}{5}|\beta_p|^2\left(|\alpha_p|^2+|\beta_p|^2\right) - \frac{4}{5}p^2|\beta_p|^2\left[M^2(\tau_1)\ \mathrm{Re}\left(\alpha_p^*\beta_p\,\mathrm{e}^{i\gamma(\tau_1)}\right)+\tau_1\leftrightarrow\tau_2\right] \\ +\frac{2}{5}\left[7M^2(\tau_1)\,M^2(\tau_2)+6p^2\left[M^2(\tau_1)\,M^2(\tau_2)\right]+6p^4\right] \\ \times\left[\mathrm{Re}\left(\alpha_p^{*2}\beta_p^2\mathrm{e}^{i[\gamma(\tau_1)+\gamma(\tau_2)]}\right)+|\beta_p|^4\mathrm{Re}\left(\mathrm{e}^{i[\gamma(\tau_1)-\gamma(\tau_2)]}\right)\right]\Bigg\}\end{aligned} \tag{306}$$

where $\gamma(\tau_i)$ are the oscillatory phases defined immediately after (292). The first term in the curly parenthesis is obtained from the non-oscillatory parts (305), and we see that it is indeed of $\mathcal{O}\left(p^4\right)$. As we shall see, this is the term that dominates the final result.

Eq. (306) is therefore characterized by a part without oscillatory phase plus a part with one oscillatory phase plus a part with two oscillatory phases. For each part we only take the leading prefactor in the $p \ll M(\tau_i)$ regime. This expression then rewrites

$$\begin{aligned}P_\lambda^{\rm sourced} \simeq \frac{2k^3}{3\pi^4 a^2 M_p^4}\int dp\, p^2\Bigg\{\frac{2}{5}p^4|\beta_p|^2\left(|\alpha_p|^2+|\beta_p|^2\right)\tilde{\mathcal{T}}_k^2 \\ -\frac{8}{5}p^2|\beta_p|^2\tilde{\mathcal{T}}_k\mathrm{Re}\left(\alpha_p^*\beta_p\tilde{\mathcal{E}}_k\right)+\frac{14}{5}\mathrm{Re}\left(\alpha_p^{*2}\beta_p^2\tilde{\mathcal{E}}_k^2\right)+\frac{14}{5}|\beta_p|^4|\tilde{\mathcal{E}}_k|^2\Bigg\}\end{aligned} \tag{307}$$

where

$$\tilde{\mathcal{T}}_k \equiv \int_{\tau_{\rm min}}^{\tau} d\tau' \frac{G_k(\tau,\,\tau')}{a(\tau')\,\omega_p(\tau')}\ , \quad \tilde{\mathcal{E}}_k \equiv \int_{\tau_{\rm min}}^{\tau} d\tau' \frac{G_k(\tau,\,\tau')}{a(\tau')\,\omega_p(\tau')}\,M^2\left(\tau'\right)\ \mathrm{e}^{2i\int_{\tau_*}^{\tau'}d\tau'' M(\tau'')} \tag{308}$$

We evaluate the time integrals after the particle production ($\tau > \tau_{\rm min}$), so that α_p and β_p can indeed be taken as constant. Particle production is completed when $\omega_p^2 > \dot{\omega}_p$, which, in the support region $p \lesssim \sqrt{\dot{m}_*}$ of the momentum integral, gives $t_{\rm min} \simeq t_* + \frac{1}{\sqrt{\dot{m}_*}}$. This corresponds to the conformal time $\tau_{\rm min} \simeq -\frac{1}{H}\exp\left(-\frac{H}{\sqrt{\dot{m}_*}}\right)$. For $\tau > \tau_{\rm min}$, we can approximate $\omega_p \simeq M \simeq a\, m$, with m given in (295). We then obtain the expression

$$\begin{aligned}
\tilde{\mathcal{T}}_k &\simeq \frac{H^2 a(\tau)}{\dot{m}_* k^3} \int_{-H\tau}^{\exp\left(-\frac{H}{\sqrt{\dot{m}_*}}\right)} dy' \left[\frac{k}{H} y' \cos\left(\frac{k}{H} y'\right) - \sin\left(\frac{k}{H} y'\right)\right] \frac{y'}{\ln(y')} \\
&\simeq \frac{H^2 a(\tau)}{\dot{m}_* k^3} \left[\frac{k}{H} \cos\left(\frac{k}{H}\right) - \sin\left(\frac{k}{H}\right)\right] \int_0^{\exp\left(-\frac{H}{\sqrt{\dot{m}_*}}\right)} \frac{dy'}{\ln(y')} \\
&\simeq \frac{H^2 a(\tau)}{\dot{m}_* k^3} \left[\sin\left(\frac{k}{H}\right) - \frac{k}{H} \cos\left(\frac{k}{H}\right)\right] \ln\left(\frac{\sqrt{\dot{m}_*}}{H}\right)
\end{aligned} \tag{309}$$

where the integration variable in the first expression is $y' = -H\tau'$. In going from the first to the second line we have used the fact that the integrand is peaked at the asymptotically early times, while in the final expression we have used $H \ll \sqrt{\dot{m}_*}$. We note that the result is only logarithmically sensitive to the difference between $\tau_{\min}$ and τ_*.

Under the approximation $\omega \simeq M$, the integral $\tilde{\mathcal{E}}_k$ coincides with the integral $\mathcal{E}_k$ defined in (294) (we actually notice that $|\tilde{\mathcal{E}}_k| \lesssim |\mathcal{E}_k|$). We obtained an upper bound for this quantity in (298). The relative contribution of the two time integrals to (307) can therefore be estimated as

$$\left|\frac{p^2\, \tilde{\mathcal{T}}_k}{\tilde{\mathcal{E}}_k}\right|_{p \simeq \sqrt{\dot{m}_*}} > \ln\left(\frac{\sqrt{\dot{m}_*}}{H}\right) \gg 1 \tag{310}$$

This shows that the oscillatory integral $\tilde{\mathcal{E}}_k$ can be disregarded in our estimate of $P_\lambda^{\rm sourced}$,

$$P_\lambda^{\rm sourced}\big|_{\text{from } A_T} \simeq \frac{4k^3}{15\pi^4 a^2 M_p^4} \tilde{\mathcal{T}}_k^2 \int dp\, p^6 |\beta_p|^2 \left(|\alpha_p|^2 + |\beta_p|^2\right) \tag{311}$$

Using the results (271), we finally obtain

$$P_\lambda^{\rm sourced}\big|_{\text{from } A_T} \simeq \frac{\left(8+\sqrt{2}\right) H^4\, \dot{m}_*^{3/2}}{32\,\pi^7\, M_p^4\, k^3} \left[\sin\left(\frac{k}{H}\right) - \frac{k}{H} \cos\left(\frac{k}{H}\right)\right]^2 \ln^2\left(\frac{\sqrt{\dot{m}_*}}{H}\right) \tag{312}$$

where we have $k \ll \sqrt{\dot{m}_*}$. while the result is exponentially suppressed and uninteresting at larger momenta. We discuss this result in Subsection VI 1.3.

In these last two expressions we have emphasized that in this computation we have considered only the GWs sourced by the transverse modes of the gauge field. In Appendix E we present the full computation, including also the longitudinal vector polarization. We find that the expression (311) is replaced by

$$P_\lambda^{\text{sourced}}\big|_{\text{from } A_T \text{ \& } A_L} \simeq \frac{2k^3}{15\pi^4 a^2 M_p^4}\tilde{\mathcal{T}}_k^2 \int dp\, p^6 \left[2|\beta_p|^2\left(|\alpha_p|^2+|\beta_p|^2\right)+|\beta_p^L|^2\left(|\alpha_p^L|^2+|\beta_p^L|^2\right)\right] \tag{313}$$

where α^L and β^L are the Bogolyubov coefficients of the longitudinal mode. Namely, we see that, if they are produced in the same amount, the longitudinal quanta contribute to the GWs as one transverse polarization. This results in a factor $\frac{3}{2}$ multiplying the final result (312).

1.3 Phenomenology

The observable spectrum of curvature fluctuations in the model (267) is given by the sum of the sourced contribution and the usual nearly scale-invariant contribution from the vacuum fluctuations; see Part III. The sourced part of the spectrum in the model (267) leads to a localized "bump" feature in the primordial power spectrum for scales leaving the horizon at the moment $t = t_*$, when the gauge fields are produced. We have illustrated this feature in the left panel of Figure 14 for an arbitrary choice of parameters. The most stringent observational constraints on the model (267) come from non-observation of such localized bump features. These observational constraints are the subject of this subsection. [50]

We can write the total observable power spectrum in the following form:

$$P_\zeta(k) = \underbrace{\mathcal{P}\left(\frac{k}{k_*}\right)^{n_s-1}}_{=P_\zeta^{\text{vac}}(k)} + \underbrace{A_b\, S_b\left[\frac{k}{k_b}\right]}_{=P_\zeta^{\text{sourced}}(k)}\,. \tag{314}$$

The spectrum of the vacuum fluctuations are characterized as usual: k_* is the pivot scale (taken to be $0.002\,\text{Mpc}^{-1}$, consistent with [166]) while the amplitude and tilt are given by

$$\mathcal{P} \equiv \frac{H^2}{8\pi^2\epsilon M_p^2}\,, \qquad n_s = 1 + 2\eta - 6\epsilon\,. \tag{315}$$

Here all quantities are understood to be evaluated at the moment when the pivot scale left

[50] In this discussion, we disregard the contribution from the longitudinal vector mode to the source of tensor and scalar perturbations. As discussed at the end of the previous Subsection, the longitudinal mode changes the result for the GW spectrum by at most a factor $\frac{3}{2}$ (if it is produced in equal amount to each tensor mode). We expect an analogous enhancement for ζ. The current discussion can be easily modified to account for these additional $\mathcal{O}(1)$ factors; all of our conclusions would be unchanged.

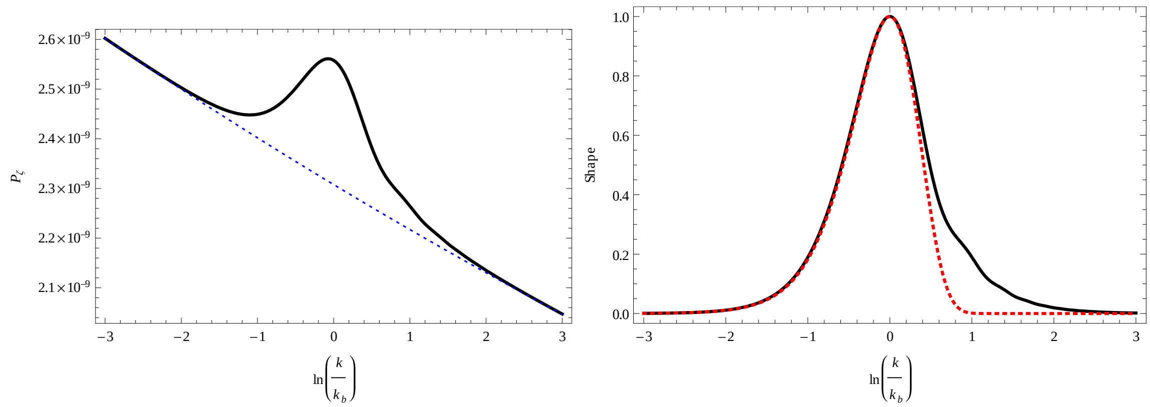

Figure 14: The localized feature in the power spectrum and its shape function.

the horizon. The sourced contribution in (314) describes a localized bump-like feature that we characterize by an amplitude A_b, a location k_b, and a "shape function" S_b. The shape and amplitude are given by:

$$A_b \equiv 3.2 \cdot 10^{-2}\, \mathcal{P}^2\, \epsilon^2\, \frac{\dot{m}_*^{7/2}}{H^7}\,, \qquad S_b[x] = 4.7\, x^3\, \left[{}_2F_3\left(\frac{3}{2}, \frac{3}{2}; \frac{5}{2}, \frac{5}{2}, \frac{5}{2}; -5.5\, x^2\right)\right]^2 . \tag{316}$$

The shape function has been constructed so that the global maximum is $S_b(x = 1) = 1$, meaning that the feature in (314) reaches its maximum value, A_b, when $k = k_b$. We assume that particle production could have taken place at any moment during inflation, so the location of the feature is treated as arbitrary.[51]

The bump-like feature (316) is very similar to the one that would be generated due to instantaneous production during inflation of *scalar* particles directly coupled to the inflaton. The signatures of scalar particle production during inflation, including important rescattering effects, were fully derived in [43]. The phenomenology of the bump-like feature in the curvature spectrum was subsequently studied in [42]. (See [35] for a discussion of non-Gaussian signatures and [34] for a review.) To derive observational constraints we follow closely the analysis in [42] and replace the somewhat complicated shape function (316) by the following simple fitting function

$$S_{\rm fit}(x) = 4.5\, x^3\, e^{-1.5x^2}\,, \tag{317}$$

[51] Concretely $k_b \approx 4.67H$ if the scale factor is normalized to one at $t = t_*$.

which, again, is normalized to have maximal value unity at $x = 1$. In the right panel of Figure 14 we show that the simple formula (317) provides an adequate description of the feature.

In [42], a variety of data sets were used to perform a detailed analysis of the observational constraints on localized features with shape (317); we refer the reader to that paper for a discussion of the methodology. See also [87] for forecast constraints from Large Scale Structure data, and see [93] for a discussion of both current and forecasted constraints from measurements of the CMB energy spectrum. In the current Part of the thesis, we are mostly interested in a feature localized on CMB scales, in which case the likelihood contours presented in [42] are approximately flat and the observational bound can be roughly summarized as:

$$\frac{A_b}{\mathcal{P}} \lesssim 0.1 \qquad \Rightarrow \qquad \frac{\dot{m}_*^{1/2}}{H} \lesssim 1.2\,\mathcal{P}^{-1/7}\,\epsilon^{-2/7} \tag{318}$$

Since the sourced part needs to be subdominant in the power spectrum, $\mathcal{P} = 2.5 \cdot 10^{-9}$; then, the bound (318) can be expressed as

$$\frac{\dot{m}_*^{1/2}}{H} \lesssim 20\,\epsilon^{-2/7}\,. \tag{319}$$

(Recall that $\dot{m}_*^{1/2} \gg H$ and $\epsilon \ll 1$ are both required for theoretical consistency.) We will see shortly that the observational bound (319) excludes the possibility of having any interesting effect from particle production in the tensor spectrum.

So far we have discussed the bump-like feature in the spectrum of curvature perturbations. However, there will also be a corresponding localized feature in the *bispectrum*. This kind of localized non-Gaussianity is very far from scale invariant and hence quite different from the bispectrum templates that are most often used in data analysis. To get a very rough sense of the amplitude of non-Gaussianity in our model, we have computed the effective (k-dependent) $f_{\rm NL}^{\rm equil}(k)$ parameter, which exhibits a bump-like structure. This parameter assumes the maximal value:

$$f_{\rm NL}^{\rm equil.eff.}(k)\Big|_{\rm max} \approx 1.9 \cdot 10^{-10}\,\epsilon^3\,\frac{\dot{m}_*^{9/2}}{H^9} \lesssim 92\,\epsilon^{3/7} \quad , \quad \text{at } k \simeq 1.2\,k_b \tag{320}$$

where $\mathcal{P} = 2.5 \cdot 10^{-9}$ has been used in the first equality and the upper bound comes from imposing (319). It should be emphasized that existing observational constraints (or forecasts) on $f_{\rm NL}^{\rm equil}$ *cannot* be applied directly to (320). A dedicated search for this type of localized non-Gaussianity would be interesting; however, it is beyond the scope of the present consideration. (See also [35, 56] for a discussion about localized non-Gaussianities from scalar particle production during inflation.) In the event of a detection of a bump in the power spectrum, localized non-Gaussian features could play an important role to falsify (or support) models of particle production during inflation.

Using our result (312), the total primordial tensor spectrum in the model (267) is given by

$$P_{\rm GW}(k) = \sum_\lambda \left[P_\lambda^{\rm vac} + P_\lambda^{\rm sourced} \right] = \frac{2H^2}{\pi^2 M_p^2} \left[1 + 4.17 \cdot 10^{-4} \, \frac{H^2}{M_p^2} \, \frac{\dot{m}_*^{3/2}}{H^3} \ln^2 \left(\frac{\dot{m}_*^{1/2}}{H} \right) S_{\rm GW}\left[x\right] \right] . \tag{321}$$

where in the shape function

$$S_{\rm GW}\left[x\right] \equiv \frac{0.0226}{x^3} \left[\sin\left(4.67\, x\right) - 4.67\, x\, \cos\left(4.67\, x\right) \right]^2 , \tag{322}$$

the scale x is normalized as in the scalar shape function (316). The shape function is maximized at $x = \frac{k}{k_b} \simeq 0.53$, where it evaluates to $\simeq 1$.

To estimate the amplitude of the GW signal that can be obtained from particle production, we evaluate the tensor-to-scalar ratio on scales where the second term in (321) is maximized. Assuming the observational constraint (319) is satisfied we can approximate $P_\zeta \approx P_\zeta^{\rm vac} = 2.5 \cdot 10^{-9}$ and we find:

$$r(k)\Big|_{\rm max} \approx 16\epsilon \left[1 + 8.2 \cdot 10^{-11} \, \epsilon \, \frac{\dot{m}_*^{3/2}}{H^3} \ln^2 \left(\frac{\dot{m}_*^{1/2}}{H} \right) \right] , \quad \text{at } \; k \simeq 0.53 \; k_b \tag{323}$$

Using (319) and $\epsilon \lesssim 0.006$ (corresponding to $r_{\rm vac} \lesssim 0.1$) it is straightforward to show that the sourced contribution to (323) is always $\lesssim 10^{-6}$, which is undetectably small. We conclude that an observationally interesting signature in GWs cannot be obtained in the model (267).

Before concluding, we note that the estimates presented here should be interpreted as lower bounds on the efficiency of particle production effects. We have considered only a single instance where $\psi = 0$, leading to a single burst of vector particle production. However, in a concrete model one could expect ψ to undergo damped oscillations about the minimum of its potential, passing through zero several times before its kinetic energy is dissipated due to Hubble friction or backreaction effects. In such a scenario resonance effects would be expected to enhance the occupation number of the produced gauge fields by some factor E that could, in principle, be $\gg 1$. The enhancement $n_k \to E n_k$ leads to a factor of E^2 in the sourced contribution to the tensor spectrum, so that the relevant term in (323) becomes

$$r(k)\Big|_{\text{max,sourced}} \approx 1.3 \cdot 10^{-9}\, E^2 \epsilon^2 \, \frac{\dot{m}_*^{3/2}}{H^3} \ln^2\left(\frac{\dot{m}_*^{1/2}}{H}\right) , \tag{324}$$

where we still assume that the scalar spectrum is dominated by the vacuum fluctuations. The sourced part of the scalar spectrum also gets enhanced by a factor of E^2, so the bound (319) becomes stronger:

$$\frac{\dot{m}_*^{1/2}}{H} \lesssim 20(\epsilon E)^{-2/7} . \tag{325}$$

Combining these results we find:

$$r(k)\Big|_{\text{max,sourced}} \lesssim 10^{-5}\ (E\epsilon)^{8/7}\ \ln^2\left[20(E\epsilon)^{-2/7}\right] . \tag{326}$$

The value $E \gtrsim 200\,\epsilon^{-1}$ gives $r_{\text{sourced}} \sim 0.01$ at the bump. It would be interesting to study under which conditions this value can be reached in a concrete model.

1.4 Comparison with GW Sourced by Modes of Different Spins

In Subsection VI 1.2, we have computed the power in GWs sourced by vector fields produced in the model (267). Schematically, the sourced part of the power spectrum is $\mathcal{P}_\lambda \propto \int d^3p\, \langle TT \rangle$, where T is the traceless-transverse spatial part of the energy-momentum tensor of source (the gauge fields, in this case) with appropriate contraction. The spatial part of the energy momentum tensor contain dominant terms that scale as $T_{ij} \sim M^2 A_i A_j$ in the $M \gg p$ limit (we recall that p and M are, respectively, the momentum and mass of the quanta sourcing the GWs). One could therefore conclude that $\mathcal{P}_\lambda \propto \int dp\, p^2\, M^4$. However, we showed that the dominant terms cancel against each other. Also the next to leading term in a $\frac{p^2}{M^2}$ Taylor expansion of the integral cancel, and one is left with $\mathcal{P}_\lambda \propto \int dp\, p^6$, see eq. (311).

In Appendix F we performed the analogous computation using fermion fields rather than vector fields as sources. In this case the spatial part of the energy-momentum tensor has terms of the type $T_{ij} \sim \bar{\chi}\gamma_i p_j \chi$, and one may conclude that $\mathcal{P}_\lambda \propto \int dp\, p^4\, M^2$ (the factor M^2 coming from the different normalization of the fermion wave function with respect to the vector one, compare the function f in (270) and in (F.5)). Also in this case there is however a cancellation, resulting in $\langle \chi^2 \rangle \propto p^2$, see eq. (F.14), and in $\mathcal{P}_\lambda \propto \int dp\, p^6$, see eq. (F.15).

These two scalings agree with that obtained if the source is a scalar particle. In this case, $T_{ij} \propto p_i p_j \phi^2$, and one immediately has $\mathcal{P}_\lambda \propto \int dp\, p^6$ without any cancellation. In fact, from our results (311) and (F.15), and from the result for the analogous computation with a scalar source given in [99], we obtain a very general expression for the power spectrum:

$$P_\lambda^{\rm sourced} \simeq \frac{2\, g_s\, k^3}{15\pi^4 a^2 M_p^4} \tilde{\mathcal{T}}_k^2 \int dp\, p^6 |\beta_p|^2 \left(|\alpha_p|^2 + (-1)^{2s}\; |\beta_p|^2 \right) \tag{327}$$

where s is the spin of the sourcing field, and g_s is the number of degrees of freedom of that field: $g_s = 1$ for a scalar, $g_s = 2$ for a vector if the longitudinal mode is produced in a negligible amount ($g_s = 3$ if it is produced in the same amount as each transverse mode), and $g_s = 4$ for a Dirac fermion.

We see that, apart from the difference in the number of degrees of freedom, and a small difference due to the spin statistics, the different fields in the nonrelativistic regime $M \gg p$ have a comparable quadrupole moment (transverse and traceless projection of T_{ij}) and generate a comparable amount of GWs.

2 Model II: Vector Produced by a Pseudo-Scalar Interaction

In this section we consider the following model

$$S = \int d^4x \sqrt{-g} \left[\frac{M_p^2}{2} R - \underbrace{\frac{1}{2}(\partial\varphi)^2 - V(\varphi)}_{\text{inflaton sector}} - \underbrace{\frac{1}{2}(\partial\psi)^2 - U(\psi) - \frac{1}{4}F^2 - \frac{\psi}{4f} F\tilde{F}}_{\text{hidden sector}} \right] . \tag{328}$$

In addition to a standard inflationary sector, we have introduced a "hidden" sector consisting of a light pseudoscalar, ψ, and a U(1) gauge field, A_μ, whose energy density is small as compared to that of the inflaton (so that the Friedmann equation takes the usual form $3H^2 M_p^2 \approx V(\phi)$). As in Section VI 1, the hidden sector in (328) has been introduced so that the production of gauge field fluctuations can provide a new source of inflationary GWs, complementary to the usual quantum vacuum fluctuations of the tensor part of the metric. Unlike the model of Section VI 1, however, we will see that particle production in the theory (328) occurs continuously during inflation, leading to broad-band signatures rather than localized features in the scalar and tensor n-point correlation functions.

The coupling $\psi^{(0)}(t) F\tilde{F}$ of the gauge field to the time-dependent pseudoscalar condensate leads to an exponential production of fluctuations A_μ. This effect has already been discussed at length in Part IV – see also Refs. [22,45,48], for example – and here we only review the key features that will be necessary for our analysis. Employing the decomposition (269) we find the following linearized equation of motion of the gauge field mode functions

$$\left[\partial_\tau^2 + k^2 \pm \frac{2k\xi}{\tau}\right] A_\pm(\tau, k) = 0\,, \qquad \xi \equiv \frac{\dot{\psi}^{(0)}}{2Hf}\,. \tag{329}$$

If the pseudoscalar is in an overdamped regime then the parameter ξ can be treated as a constant. Moreover, we assume that $\dot{\psi}^{(0)} > 0$ so that the "+" helicity state of the gauge field gets copiously produced while the "−" state remains in the vacuum and its effect is renormalized away.[52] The properly normalized solutions of (329) can be written as [22],[53]

$$A_+(\tau, k) \approx \left(\frac{-\tau}{8\xi k}\right)^{1/4} e^{\pi\xi - \sqrt{-2\xi k\tau}}\,, \qquad A'_+(\tau, k) \approx \left(\frac{2\xi k}{-\tau}\right)^{1/2} A_+(\tau, k)\,. \tag{330}$$

This solution is valid only in the phase space interval $\frac{1}{8\xi} \ll -k\tau \ll 2\xi$, where the production of gauge fluctuations is most important. By restricting ourselves to this regime we effectively cut-off an ultra-violet divergence associated with the usual quantum vacuum fluctuations of the gauge field on sub-horizon scales, as discussed in Part IV. We have also assumed that $\xi \gtrsim \mathcal{O}(1)$, so that the phase space of produced fluctuations is non-trivial and each mode experiences a significant exponential enhancement, $e^{\pi\xi} \gg 1$, near horizon crossing. (For $\xi < 1$ there is no interesting particle production in the model.)

We are interested in a scenario where inflation is driven by the potential energy of the φ field, so that $3H^2M_p^2 \approx V(\varphi)$; the "hidden" sector in (328) instead should give a small contribution to the total energy density of the universe. This requirement imposes several constraints on the model parameters, which we now discuss. We must first require that the energy density in the produced gauge field fluctuations is smaller than the kinetic energy of ψ,

$$\frac{1}{2}\langle \vec{E}^2 + \vec{B}^2 \rangle \ll \frac{\dot{\psi}^{(0)2}}{2}\,, \tag{331}$$

where the "electric" and "magnetic" fields are $E_i \equiv -\frac{1}{a^2}A'_i$, $B_i \equiv \frac{1}{a^2}\epsilon_{ijl}\partial_j A_l$. Using the

[52] None of our result for the scalar or tensor correlation functions will depend on the choice $\dot{\psi}^{(0)} > 0$.

[53] A detailed derivation is shown in Appendix B.

solution (330) to evaluate the expectation value, the condition (331) can be written as:

$$\frac{H^2}{\dot{\psi}^{(0)}} \ll 60\,\xi^{3/2} e^{-\pi\xi} \,. \tag{332}$$

We also require that the energy density of the rolling pseudoscalar can be neglected with respect to that of the inflaton:

$$\frac{1}{2}\left(\dot{\psi}^{(0)}\right)^2 + U\left(\psi^{(0)}\right) \ll 3H^2 M_p^2 \,. \tag{333}$$

Finally, we require that ξ is adiabatically evolving, $\frac{\dot{\xi}}{H\,\xi} = \frac{\ddot{\psi}^{(0)}}{H\,\dot{\psi}^{(0)}} - \frac{\dot{H}}{H^2} \ll 1$, so that it is appropriate to treat it as nearly constant during the time interval in which each mode of A is relevant (namely, close to horizon crossing, when the mode is produced, and affect cosmological perturbations). As $|\frac{\dot{H}}{H^2}| \ll 1$ during inflation, we need to require

$$\frac{\ddot{\psi}^{(0)}}{H\,\dot{\psi}^{(0)}} \ll 1 \quad \Leftrightarrow \quad m_\psi \ll \frac{3H}{2} \tag{334}$$

where in the last condition we have approximated $U(\psi)$ as a quadratic potential, and we have required the evolution of $\psi^{(0)}$ to be in the overdamped regime. Throughout our analysis we will require that the conditions (332), (333), and (334) are simultaneously satisfied (these conditions are discussed at the end of Subsection 2.3).

The background equation for the pseudoscalar reads:

$$\ddot{\psi}^{(0)} + 3H\dot{\psi}^{(0)} + U'\left(\psi^{(0)}\right) = \frac{1}{f}\langle \vec{E}\cdot\vec{B}\rangle \,. \tag{335}$$

It is interesting to note that the condition (332) guarantees right-hand side of this equation can be disregarded. Indeed,

$$|U'\left(\psi^{(0)}\right)| \gg \frac{1}{f}|\langle \vec{E}\cdot\vec{B}\rangle| \quad \Leftrightarrow \quad \frac{H^2}{\dot{\psi}^{(0)}} \ll 82\xi^{3/2} e^{-\pi\xi} \,. \tag{336}$$

which is implied by (332).

The quantity $\mathcal{D}^{(\lambda)}$ introduced in (65) characterizes the two-point function of the produced gauge field fluctuations. This function, and its derivative, can be written explicitly in terms of the c-number mode functions as

$$\mathcal{D}^{(\lambda)}_{(0,0)}\left[\tau_1,\tau_2;q\right] \equiv A_\lambda(\tau_1,q)A^*_\lambda(\tau_2,q)\,, \quad \mathcal{D}^{(\lambda)}_{(1,1)}\left[\tau_1,\tau_2;q\right] \equiv A'_\lambda(\tau_1,q)A^{*\prime}_\lambda(\tau_2,q)\,. \tag{337}$$

Considering only the "+" helicity state and using the approximate solution (330) we have

$$\mathcal{D}^{(+)}_{(0,0)}\left[\tau_1,\tau_2;q\right] \approx \frac{(\tau_1\tau_2)^{1/4}}{\sqrt{8\xi q}}e^{2\pi\xi-\sqrt{2\xi q}\left[\sqrt{-\tau_1}+\sqrt{-\tau_2}\right]}\,, \quad \mathcal{D}^{(+)}_{(1,1)}\left[\tau_1,\tau_2;q\right] \approx \frac{2\xi q}{\sqrt{\tau_1\tau_2}}\mathcal{D}^{(+)}_{(0,0)}\left[\tau_1,\tau_2;q\right]\,. \tag{338}$$

2.1 Scalar Perturbations Sourced by the Vector Modes

2.1.1 The Master Equation

The hidden sector in (328) decouples from the inflaton in the limit $M_p \to \infty$. However, at finite M_p, gravitational couplings will transmit the effects of particle production in the hidden sector to the inflaton perturbations, modifying the usual predictions for the observable curvature fluctuations. To see this effect we derive a master equation for the inflaton perturbations in the model (328) – see Appendix A. At linear order in perturbation theory we have recovered the standard result

$$\left(\partial_\tau^2 - \nabla^2\right)\delta_1\varphi + 2\mathcal{H}\delta_1\varphi' \simeq 0\,, \tag{339}$$

where we work to leading order in slow roll parameters. At second order, instead, in Appendix A.3 we find the following equation of motion

$$\left(\partial_\tau^2 - \nabla^2\right)\delta_2\varphi + 2\mathcal{H}\delta_2\varphi' \simeq -\frac{\varphi^{(0)\prime}a^2}{2M_p^2\mathcal{H}}\left\{\underbrace{\frac{E^2+B^2}{2}}_{\text{gives } J_1} + \frac{1}{a^4}\underbrace{\nabla^{-2}\partial_\tau\left[a^4\vec{\nabla}\cdot\left(\vec{E}\times\vec{B}\right)\right]}_{\text{gives } J_2}\right\} + \ldots \tag{340}$$

which is formally equivalent to (283) with $m^2 \to 0$. In (340) the $\cdots$ schematically denotes terms involving $\delta_1\varphi^2$, ψ_1^2 and $h^2_{1,ij}$ which do not involve the exponential factors $e^{\pi\xi}$ that characterize the gauge field modes (330) and may therefore be neglected.

As we did in Subsection VI 1.1, both the first and second order equations can be combined into the single master equation (50). The first-order mode $\delta_1\varphi$ is the homogeneous solution of this master equation, while the second-order $\delta_2\varphi$ is the particular solution of this master equation. The source in the master equation can be written as a sum of two terms

$$J_\varphi\left(\tau,\vec{k}\right)=J_1\left(\tau,\vec{k}\right)+J_2\left(\tau,\vec{k}\right)\,. \tag{341}$$

The source J_2 is associated with the second term on the right hand side of (340), which is non-local in position space. We show in Appendix A.3 that this term is actually infra-red finite, and that the full source can be cast in the form

$$\begin{aligned} J_\varphi\left(\tau,\vec{k}\right) &= \frac{\varphi'^{(0)}a^3}{4M_p^2\mathcal{H}}\int\frac{d^3p}{(2\pi)^{3/2}}\left[-1+\frac{(p-|\vec{k}-\vec{p}|)^2}{k^2}\right] \\ &\quad\times\left[\tilde{E}_i\left(\tau,\vec{p}\right)\tilde{E}_i\left(\tau,\vec{k}-\vec{p}\right)+\tilde{B}_i\left(\tau,\vec{p}\right)\tilde{B}_i\left(\tau,\vec{k}-\vec{p}\right)\right]\,. \end{aligned} \tag{342}$$

where we have defined the "electric" and "magnetic" field operators as $\tilde{E}_i(\tau,\vec{k})\equiv-\frac{1}{a^2}\tilde{A}_i'\left(\tau,\vec{k}\right)$ and $\tilde{B}_i\left(\tau,\vec{k}\right)\equiv\frac{i}{a^2}\epsilon_{ijl}k_j\tilde{A}_l\left(\tau,\vec{k}\right)$. This expression appears simpler than the corresponding source (288) in the previous model, due to the fact that only the modes A_+ are relevant here.

In this model the energy density in the "electric" field dominates over that in the "magnetic" field [45]. Dropping terms involving $\tilde{B}_i$ we have a source term of the form (51) where the model-dependent operator can be approximated by

$$\hat{O}_{\varphi,ij}\left(\tau,\vec{k},\vec{p}\right)\approx\frac{\varphi'^{(0)}}{4M_p^2\mathcal{H}a}\left[-1+\frac{(p-|\vec{k}-\vec{p}|)^2}{k^2}\right]\,\delta_{ij}\,\partial_\tau^{(1)}\partial_\tau^{(2)}\,, \tag{343}$$

where we recall our notation (289).

2.1.2 Two-point and Three-point Correlation Functions

We now proceed to compute the two-point and three-point correlation functions of the gauge invariant curvature perturbation, ζ. We are only interested in contributions that are sourced by particle production effects, since the vacuum fluctuations in the model (328) are standard. The sourced contribution to the power spectrum of ζ is given by (71). As discussed above, we only consider the $\sigma=+$ contributions in the sum over helicity states. Using the explicit expression (343) for the operator $\hat{O}$, along with the identity (66), we find that (71) can be written as

$$P_\zeta^{\rm sourced}(k) \simeq \frac{k^3}{64\pi^2 M_p^4 a^2} \int \frac{d\tau_1}{a(\tau_1)}\, G_k(\tau,\tau_1) \int \frac{d\tau_2}{a(\tau_2)}\, G_k(\tau,\tau_2)$$
$$\times \int \frac{d^3p}{(2\pi)^3} \left[1 + \frac{p^2 - \vec{k}\cdot\vec{p}}{p|\vec{k}-\vec{p}|}\right]^2 \left[1 - \frac{(p-|\vec{k}-\vec{p}|)^2}{k^2}\right]^2 \mathcal{D}^{(+)}_{(1,1)}\left[\tau_1,\tau_2;p\right] \mathcal{D}^{(+)}_{(1,1)}\left[\tau_1,\tau_2;|\vec{k}-\vec{p}|\right] . \tag{344}$$

Next we insert our previous result (338) for $\mathcal{D}^{(+)}_{(1,1)}$. Since we are interested in computing the spectrum at late times, the explicit expression (61) for the Green function can be employed. The sourced power spectrum takes the form

$$\begin{aligned} P_\zeta^{\rm sourced}(k) &= \frac{\xi e^{4\pi\xi} H^4}{128\pi^2 M_p^4} \int \frac{d^3q}{(2\pi)^3} q^{1/2} |\hat{k}-\vec{q}|^{1/2} \left[1 - (q - |\hat{k}-\vec{q}|)^2\right]^2 \\ &\qquad \times \left[1 - \frac{\vec{q}\cdot(\hat{k}-\vec{q})}{q|\hat{k}-\vec{q}|}\right]^2 \mathcal{I}^2\left[q, |\hat{k}-\vec{q}|\right] \end{aligned} \tag{345}$$

where we have introduced dimensionless variables $\vec{q} \equiv \vec{p}/k$ and $\hat{k} \equiv \vec{k}/k$. The dimensionless time integral is defined as

$$\mathcal{I}\left[a,b\right] \equiv \int_{-k\tau}^{\infty} dz\, \frac{\sin z - z\cos z}{z^{1/2}}\, e^{-2\sqrt{2\xi z}\left[\sqrt{a}+\sqrt{b}\right]} , \tag{346}$$

(notice that $z \equiv -k\tau'$). In the super-horizon regime, $-k\tau \ll 1$, we can set the lower bound of integration to zero. Moreover, particle production effects are most interesting in the regime $\xi \gtrsim \mathcal{O}(1)$, in which case the integral (346) has most of its support in the region $z \ll 1$ where we can approximate $\sin z - z\cos z \approx \frac{z^3}{3}$. Hence, we have the following analytical approximation

$$\mathcal{I}\left[a,b\right] \approx \int_0^\infty dz\, \frac{z^{5/2}}{3}\, e^{-2\sqrt{2\xi z}\left[\sqrt{a}+\sqrt{b}\right]} = \frac{15}{32\sqrt{2}\left(\sqrt{a}+\sqrt{b}\right)^7 \xi^{7/2}} . \tag{347}$$

Finally, the momentum integral in (345) must be performed numerically, giving the final result

$$P_\zeta^{\rm sourced}(k) \approx 4\cdot 10^{-10} \frac{H^4}{M_p^4} \frac{e^{4\pi\xi}}{\xi^6} . \tag{348}$$

The computation of the three-point correlation function is completely analogous to the derivation that we have outlined for the source power spectrum. Here we simply state the final result for the bispectrum in the equilateral configuration:

$$B_\zeta\left(k_1 = k_2 = k_3 \equiv k\right) \approx 2.6 \cdot 10^{-13} \frac{H^6}{M_p^6} \frac{e^{6\pi\xi}}{\xi^9} \frac{1}{k^6} \,. \tag{349}$$

The reason for considering the equilateral configuration is that, in this model, the source at any moment is dominated by modes with wavelength comparable to the horizon at that moment. This generates mostly correlations between scalar perturbations of comparable size [45, 48].

2.2 Gravity Waves Sourced by the Vector Modes

The produced gauge field fluctuations that are described by the mode solution (330) carry anisotropic stress/energy and provide a source of GW fluctuations that is complimentary to the standard quantum vacuum fluctuations from inflation.[54] The computation of the GW spectrum in the model (328) follows closely what we had for the the curvature perturbation. In fact, the identical computation was done in Subsection IV 2.5. Therefore, we here simply state the final result:

$$\begin{aligned} P_L &= P_L^{\rm vac} + P_L^{\rm sourced} \simeq \frac{H^2}{\pi^2 M_p^2}\left[1 + 8.6 \cdot 10^{-7}\,\frac{H^2}{M_p^2}\frac{e^{4\pi\xi}}{\xi^6}\right] \\ P_R &= P_R^{\rm vac} + P_R^{\rm sourced} \simeq \frac{H^2}{\pi^2 M_p^2}\left[1 + 1.8 \cdot 10^{-9}\,\frac{H^2}{M_p^2}\frac{e^{4\pi\xi}}{\xi^6}\right] \end{aligned} \tag{350}$$

The first term in the square braces corresponds to the usual contribution from the quantum vacuum fluctuations of the graviton, and the two helicities have equal power. The second term, on the other hand, corresponds to GW perturbations that have been sourced by particle production effects in the hidden sector. The production is much more significant for the h_R mode, due to the fact that the source consists of only A_+ modes.

[54] At second order in cosmological perturbation theory tensor fluctuations can also be sourced by bi-linear combinations of the first order scalar fluctuations, $\delta\varphi$ and $\delta\psi$. Neither of these exhibit the exponential enhancement that characterizes the linear gauge field perturbations – see equation (330) – therefore we can safely neglect this effect in what follows.

2.3 Phenomenology

In this subsection we explore the phenomenology of the scalar and tensor cosmological fluctuations in the model (328). We first consider the spectrum of curvature fluctuations. The total observable power spectrum is the sum of (348) and the standard result for the vacuum fluctuations (see the general result in eq. (71)). Explicitly we have

$$P_\zeta \approx \mathcal{P}\left[1 + 2.5\cdot 10^{-6}\,\epsilon^2\mathcal{P}\frac{e^{4\pi\xi}}{\xi^6}\right]\,, \qquad \mathcal{P} \equiv \frac{H^2}{8\pi^2\epsilon M_p^2}\,, \tag{351}$$

where the first term in the square braces is the usual spectrum from quantum vacuum fluctuations, while the second term is the sourced contribution coming from gauge field production in the model (328).

Next, we consider the spectrum GW fluctuations. From (350) and (351) we can write the tensor-to-scalar ratio as

$$r \equiv \frac{\sum_\lambda P_\lambda}{P_\zeta} \approx 16\epsilon\frac{1 + 3.4\cdot 10^{-5}\epsilon\mathcal{P}\frac{e^{4\pi\xi}}{\xi^6}}{1 + 2.5\cdot 10^{-6}\epsilon^2\mathcal{P}\frac{e^{4\pi\xi}}{\xi^6}}\,. \tag{352}$$

When both the tensor and scalar spectra are dominated by the vacuum fluctuations we recover the standard result, $r \approx 16\epsilon$. On the other hand, if the sourced contributions to (350) and (351) dominate then we have $r \approx 218$, independently of model parameters. The current observational limit is $r \lesssim 0.11$ [12], while $r \sim 0.01$ might be detectable with future missions [58]. Since the expression (352) interpolates between 16ϵ and 218, it follows that an observable signal can be obtained for *any* value of ϵ.

Let us now discuss the phenomenology of the scalar perturbations. From (351) and (352) we get

$$\frac{P_\zeta^{\rm sourced}}{P_\zeta^{\rm vac}} \simeq \frac{r - 16\,\epsilon}{218} \tag{353}$$

We therefore see that the scalar power spectrum is dominated by the vacuum part ($P_\zeta \simeq \mathcal{P}$). Concerning the scalar bispectrum, the effective nonlinearity parameter is immediately obtained from (77) and (349):

$$f_{\rm NL}^{\rm equil.eff.} \approx 1.5\cdot 10^{-9}\epsilon^3\frac{\mathcal{P}^3}{P_\zeta^2}\frac{e^{6\pi\xi}}{\xi^9}\,. \tag{354}$$

The current CMB bound $f_{\rm NL}^{\rm equil} = -42 \pm 75$ at 68% CL [13],[55] while $f_{\rm NL} \sim \mathcal{O}(10)$ might be accessible in the near future.

To obtain the correct amplitude of density fluctuations, we must impose the normalization condition $P_\zeta = 2.2 \cdot 10^{-9}$ on CMB scales. We can use this condition to eliminate the parameter $\mathcal{P}$ in favour of ξ and ϵ. Having done so, the key observables r and $f_{\rm NL}^{\rm equil.eff}$ then depend only on the model-dependent quantities ξ and ϵ, which in turn depend on the inflationary potential and the dynamics of the hidden sector fields. In Figure 15 we plot our results for r and $f_{\rm NL}$ as a function of ξ, for various representative choices of ϵ.

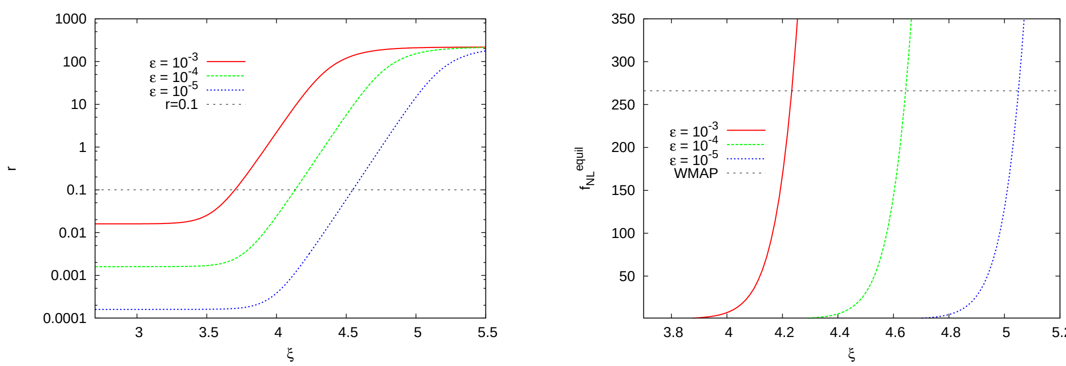

Figure 15: The tensor-to-scalar ratio and non-linearity parameter as functions of ξ for several illustrative choices of ϵ.

The observational bound on the tensor-to-scalar ratio forces us into a region of parameter space where non-Gaussianity is undetectably small. Therefore GW fluctuations constitute the most interesting phenomenology associated with the model (328). This is shown in Figure 16 where we plot contours in the $\xi - \epsilon$ plane leading to various phenomenologically interesting scenarios. We note that our findings are relevant also for values of ϵ smaller than those shown in the figure.

As discussed in [23, 45, 284], the sourced contribution to the tensor spectrum is chiral; only one helicity state is efficiently sourced by the gauge field fluctuations (330). This effect may be detected through TB and EB correlations in the CMB [147, 262]. This was first

[55] The bound from 7-year WMAP results was $-114 < f_{\rm NL}^{\rm equil} < 266$ at 68% CL [192], which is used for our present analysis.

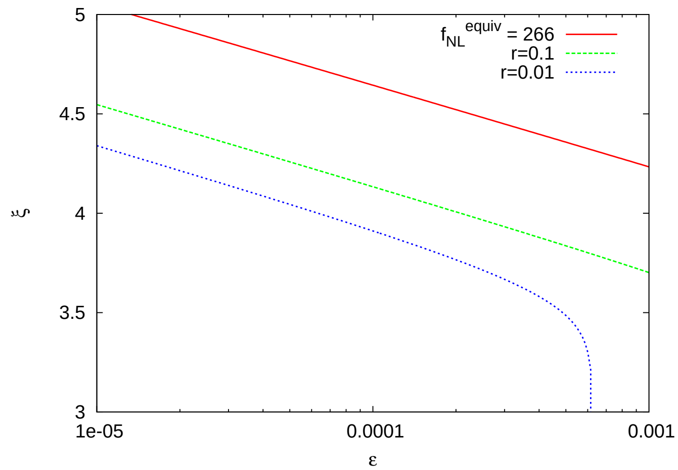

Figure 16: Exclusion curves for parameter space from the WMAP bound on non-Gaussianity, and the WMAP and prospective bounds on tensor-to-scalar ratio r.

explored by [284] in the case in which the inflaton is the pseudo-scalar sourcing the vector modes; in this case, the direct inflaton-gauge field coupling is so strong that, typically, the main bound on the gauge field production is given by the sourced scalar perturbations (non-Gaussianity [45, 48] and, depending on the inflaton potential, increased power at small scales [47, 232]). To overcome this, [284] assumed the presence of ~ 1000 sourcing gauge fields (this decreases the amount of non-Gaussianity), or the curvaton mechanism for the generation of the scalar perturbations. For some values of parameters, the signal can be above the 1σ detection line for a cosmic-variance limited experiment [284]. As we shall now discuss, a more optimistic conclusion is reached if one assumes that the gauge field production occurs in a sector only gravitationally coupled to the inflaton, as we have studied here.

A measure of the net handedness of the tensor modes is the following quantity:

$$|\Delta\chi| \equiv \left|\frac{P_+ - P_-}{P_+ + P_-}\right| = \frac{3.4 \cdot 10^{-5} \epsilon \mathcal{P} \frac{e^{4\pi\xi}}{\xi^6}}{1 + 3.4 \cdot 10^{-5} \epsilon \mathcal{P} \frac{e^{4\pi\xi}}{\xi^6}} \simeq 1 - \frac{16\,\epsilon}{r}\,, \tag{355}$$

which interpolates between zero (at small ξ, when the vacuum fluctuations dominate the tensor mode spectrum) and unity (at large ξ when the sourced GW dominate the tensor mode spectrum). In the final approximation we have used the fact that, for $r < 0.1$, the scalar power spectrum in this model is dominated by the vacuum modes.

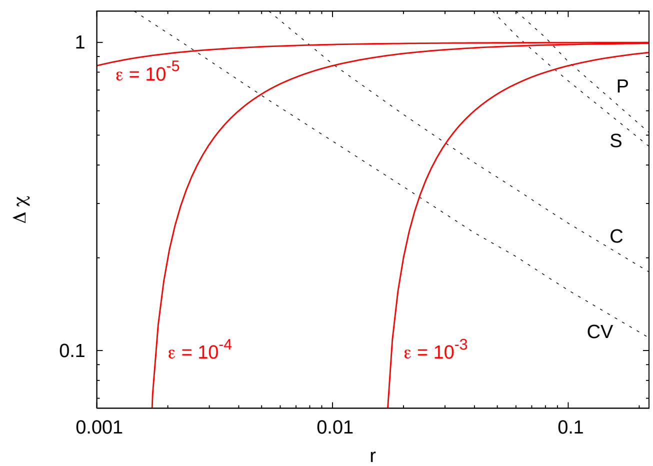

Figure 17: Predictions for the chirality vs. tensor-to-scalar ratio, as compared to 1σ detection curves for the Planck (P), SPIDER (S), CMB-Pol (C), and a cosmic-variance limited (CV) experiment.

In Figure 17 we plot the relation (355) in the r vs $\Delta\chi$ plane, for a few representative values of ϵ; each of the red/solid lines is characterized by a given ϵ, and by varying ξ (growing ξ leads to more GW production, and therefore greater values of r and ξ). We stress that arbitrary large values of r in the range shown in the figure can be reached for any value of ϵ. As ϵ decreases, this requires a greater and greater amount of sourced modes, which in turn leads to a greater and greater $\Delta\chi$. This explain why, for any given obtained r, greater $\Delta\chi$ correspond to smaller ϵ. These predictions are superimposed in the figure to 1σ detection lines from various experiments; from top to bottom, the lines shown are for the ongoing and forthcoming Planck (P) [4] and SPIDER (S) [100] experiments, for the suggested CMB-Pol experiment (C) [58], and for a hypothetical cosmic-variance limited experiment (CV). The signal needs to be above a line to be detectable at 1σ by that experiment. These lines are taken by Figure 2 of [147]. We observe that, for some values of parameters, the parity-violation could be detected (at least at 1σ) already by the ongoing / forthcoming Planck and SPIDER experiments.

Before concluding this section, we comment on the constraints (332) and (333) which are necessary for the consistency of our calculation. We find:

$$0.074 \frac{\sqrt{\epsilon \mathcal{P}} e^{\pi \xi}}{\xi^{5/2}} \ll \frac{f}{M_p} \ll \frac{1.2}{\xi} \sqrt{1 - \frac{U(\psi)}{V(\varphi)}} \tag{356}$$

where the first condition, obtained from (332), ensures that the energy density of the produced gauge quanta is smaller than the kinetic energy of ψ, while the second condition, obtained from (333), ensures that the energy density of ψ is smaller than that of the inflaton.

The interval in (356) exists for any reasonable choice of model parameters, therefore we can always choose f/M_p such that that various backreaction constraints are satisfied. The lower bound on the inflationary scale becomes simpler in the limit in which the sourced part of the GW signal dominates over the vacuum one (which is the regime of our most interest). Using (352) in this regime, the lower limit in (356) can be expressed as:

$$\frac{f}{M_p} \gg \frac{3.7 \cdot 10^{-3} \, r^{1/4}}{\xi} \sim \mathrm{O}\left(10^{-4}\right) \quad , \quad r_{\rm vac} < r_{\rm sourced} \sim 0.01 - 0.1 \tag{357}$$

where r has been chosen so that the GW signal is observable in the near future (we note that this requires $\xi \sim 4 - 5$).

Finally, using (315), the condition (334) for the adiabatic evolution of ξ can be cast in the form

$$\frac{m_\psi}{M_p} \ll 6.6 \cdot 10^{-4} \sqrt{\epsilon} \tag{358}$$

For a quadratic inflaton potential, $m_\varphi \simeq 6.4 \cdot 10^{-6} \, M_p$ and $\sqrt{\epsilon} \simeq 0.09$ (for 60 e-folds of inflation). The condition (358) then rewrites $m_\psi \ll 9 m_\varphi$.

3 Summary

A large experimental effort is currently taking place to detect GWs from inflation. The conventional vacuum signal will be detectable only if the scale of inflation is sufficiently high, $V^{1/4} \gtrsim 10^{16}$ GeV (corresponding to $r \gtrsim 0.01$, or $\epsilon \gtrsim 6 \cdot 10^{-4}$ in single field slow roll inflation). At face value, this seems to imply that such an experimental effort will be unsuccessful for a lower scale inflation. In this Part of the thesis, we have challenged to this conventional lore and discussed ways to distinguish new sources of GWs from the standard vacuum ones.

Let us denote by X the field that is produced during inflation, and that sources the GWs. In the models considered in [99, 271] quanta of X are produced by the motion of the inflaton, which we denote by φ. This implies a direct coupling between the inflaton and the produced quanta; as a consequence, if the inflaton is the source of cosmological perturbations ζ, quanta of X will source ζ with a stronger than gravitational interaction. On the contrary, the source of GWs from X is of gravitational strength. To minimize the relative amount of produced ζ vs. produced GWs, in this Part we made the opposite assumption, considering the weakest possible coupling (in standard gravitational theory) between X and the inflaton: namely, we assumed that φ and X are coupled only gravitationally. We therefore assumed that particle production occurs in a "hidden sector" and computed the amount of scalar perturbations ζ induced by X through a purely gravitational interaction. We then computed the amounts of GWs produced by X in the two models and compared the two effects.

Clearly, there is a large arbitrariness in the choice of the model for particle production, and we do not claim our findings to be exhaustive; in particular, we did not study here the analogous of all the scenarios considered in [271], where for instance multiple bursts of particle production and production of strings were also studied. We study two models in which X is a vector field, which is produced by the motion of a field $\psi \neq \varphi$. In Model I, the vector field has a mass term $\psi(t)^2 A^2$, and quanta of A are produced when the classical value of ψ crosses zero. In Model II the vector is continuously sourced by a pseudo-scalar $\frac{\psi}{f} F\tilde{F}$ interaction. The reasons for considering these two models is that, after the particle production, the vector quanta are highly massive in Model I, while massless in Model II. We showed that in the first case this gives rise to a strong suppression of the GW signal with respect to the amount of scalar perturbations. In the remainder of this concluding section, we summarize our findings in these two models, together with some discussion.

1. **Model I:** The main signature of particle production in this model is a bump in the scalar power spectrum, at the scales that exited the horizon when the gauge quanta were produced.[56] If this bump will be observed, this mechanism can be supported / disproved by the presence/ absence of an analogous bump in the bispectrum, which we also computed here. The spectrum of GWs produced by the gauge quanta also presents a peak at the same scales, which – if sufficiently high – could distinguish them from

[56] This is qualitatively identical to the findings of [43], where the sourced field is a scalar with mass depending on the inflaton.

> the vacuum GWs. However, we found that, once the bound from not having observed a large bump in the scalar spectrum is respected, the amount of GWs produced in this model is completely unobservable ($r \lesssim 10^{-6}$). The relative smallness of the GWs vs scalar perturbations produced in the model may come as a surprise, due to the fact that the gauge quanta are coupled gravitationally to both these quantities. The reason for the suppression is due to the fact that the quanta are highly non-relativistic after they are produced. This highly suppresses their quadrupole moment, and the amount of GWs that they generate.

To verify this, we also computed the amount of GWs produced if the vector field is replaced by a fermion, with mass $\propto \psi(t)$. Ref. [99] computed the amount of GWs produced by a scalar with mass $\propto \psi(t)$. From our two results, and from the result of [99] for the scalar case, we actually obtained the very general formula (327) for the amount of GWs produced in all these cases. We see that there is no large enhancement between the different spins, apart from the proportionality of the final result to the number of degrees of freedom in each case.

We did not compute the amount of scalar perturbations ζ sourced in the fermionic case. However, we believe that also in this case the result will be analogous to the one that we have computed, and that therefore our conclusions on the relative importance of GWs vs scalar perturbations production apply for produced particles of any spin.

In this model, the vector field is massive after the production, and it therefore also possesses a longitudinal component. The results for this component are more model dependent than those of the transverse components: they depend on the specific choice of the potential U for ψ (the results for the transverse component are independent of U provided that it is smaller than the kinetic energy of ψ during particle production). The longitudinal mode may drive the theory out of perturbative regime when $m \propto \psi(t) \to 0$. We showed that this is avoided if the ratio $\frac{1}{\psi}\frac{dU}{d\psi}$ remains finite as $\psi \to 0$. This is, for instance, the case if $U \approx \frac{1}{2}m_\psi^2\psi^2$ at the origin. These considerations can be relevant for all the models in which symmetries are enhanced at some point during the cosmological evolution. For instance, we expect massive gauge modes to become massless when different branes move to the same bulk location as in the trapping mechanism of [188]. We also found that, for $U \approx \frac{1}{2}m_\psi^2\psi^2$, the longitudinal component is produced as much as each transverse component, and sources the same amount of GWs, for the most reasonable values of the mass m_ψ.

It is worth pointing out that our study applies to a single instance of particle production. Things may be different in cases of multiple bursts of particle production; for instance, if the potential U is not flat at large ψ, so that ψ performs oscillations about its minimum, gauge quanta will be produced at each oscillation, in a regime of parametric resonance [190]. We have found that the GW signal can reach an interesting level if the parametric resonance enhances the amount of produced quanta by a factor of $\gtrsim 200\,\epsilon^{-1}$ with respect to the single episode of particle production (ϵ being the slow roll parameter). It would be interesting to study under which conditions this value can be reached in a concrete model.

2. **Model II:** The amount of GWs produced in this model was already computed in [45, 48, 284]. The novel computation performed in this thesis is the amount of scalar perturbations produced in this model under the assumption that the inflaton is only gravitationally coupled to the gauge field, and the comparison of the two effects. We found that coupling the inflaton only gravitationally sufficiently suppresses the amount of scalar modes generated in this model, so that the limits from non-Gaussianity are irrelevant when compared to those from GWs. Therefore, particle production in this model can lead to GWs observable at the CMB scales. Differently from Model I, the GWs in this model are produced at all scales, without a localized bump. However, this signal may be distinguishable from the vacuum one since one GW helicity is produced in a much stronger amount than the other one, and this can lead [284] to observable nonvanishing TB and EB correlations in the CMB [147, 262].

Ref. [284] already studied whether the parity violation in the sourced GWs produced by this model can be observed. Also in that case, the problem was to suppress the non-Gaussianity of the scalar perturbations produced by the gauge field [48]. This was overcome in [284] by assuming the presence of ~ 1000 sourcing gauge fields, or the curvaton mechanism for the generation of the scalar perturbations. It was shown in [284] that, under these assumptions, and for some values of parameters, the parity-violation signal can be above the 1σ detection line for a cosmic-variance limited experiment [284]. We have seen that in our implementation of the mechanism (namely, by assuming that the gauge field is only gravitationally coupled to the inflaton) the parity violation can, for some choice of parameters, be observed already by the ongoing / forthcoming experiments.

We have seen that backreaction bounds from this mechanism are under control for an axion scale f in the interval $10^{-4}M_p \lesssim f \lesssim M_p$. Interestingly, the axion decay constant one typically finds in string theory is of the order of the GUT scale $f \sim 10^{16}$ GeV (see, e.g., [30, 291]) which fits comfortably within this window. Indeed, given the UV sensitivity of inflation, it is natural to ask whether one can realize our model in string theory. The low energy spectrum of string theory contains generically many axion-like particles, which arise from the reduction of antisymmetric p-form fields on p-cycles of the internal space.[57] These closed string axions have a pseudoscalar coupling $\psi F\tilde{F}$ to U(1) gauge fields on the worldvolume of D-branes.[58] In such string theory setting, it is not difficult to find inflaton candidates with no direct coupling to the axion-gauge field sector. Explicit model building possibilities (with the inflaton being an axion or not) remain to be explored.

[57] In addition to these closed string axion-like particles, there are also open string axions but their presence is more model-dependent. For this discussion, we shall focus on closed string axions. Moreover, we consider only those axions that are not projected out by discrete symmetries (e.g., orientifolding), and do not receive a Stuckelberg mass.

[58] Such couplings arise from the reduction of Chern-Simons terms in the worldvolume action of D-branes, e.g., $\int_{D_{p+4}} C_p \wedge F \wedge \tilde{F}$.

Part VII
Conclusions

In this thesis, we have considered the phenomenological signature of particle physics interactions that may have been present during primordial inflation. For definiteness, we have mostly studied the cases in which a slowly rolling field (the inflaton, in most cases) is coupled through some characteristic gauge invariant interaction to a vector field. A free gauge field with the standard $-\frac{1}{4}F^2$ kinetic term is conformally coupled to the FRW metric and its perturbations are not excited by the expansion of the universe. The couplings we have studied, however, break this conformal invariance, modifying the dispersion relation of the gauge field. This leads to non-trivial production of the gauge quanta, continuously or at a specific time, during inflation.

The produced quanta have an energy significantly smaller than that of the unperturbed inflaton for all or most of the duration of inflation. However, they can strongly affect the cosmological perturbations. Specifically, they act as a "source" of the inflaton perturbation, which in turn imprints in the curvature perturbation. The sourced contribution is uncorrelated to the contribution from the pure inflaton vacuum modes, and thus these two contributions simply add up in any statistical quantity with vanishing cross correlations. The standard vacuum contribution is nearly Gaussian. However, this is not the case for the sourced perturbations. We have shown in this thesis that the level and the shape of non-Gaussianity in these models provide surprisingly rich phenomenology. If detected, this would serve as a powerful tool to probe the details of the physics during inflation and to discriminate between a plethora of inflationary models.

In most of the cases we have studied, the gauge quanta are produced by the motion of the inflation, and therefore there is a direct, stronger than gravitational, coupling between the inflaton and the gauge field. This is the reason why these quanta produce much more inflaton perturbations than gravitational waves (GWs). These models result in a standard GW signal at the CMB scales, once the limits resulting from the non-observation of non-Gaussianity are imposed. There are however also interesting cases in which the GW production can be relevant. In this thesis we studied two such possibilities. Firstly, we studied the GW signal at much smaller scales (those produced towards the end of inflation), relevant for terrestrial GW interferometers; at such scales there is more room for a significant GW production, since the limits on the scalar perturbations are less strong than those at CMB scales, and more uncertain. Secondly, we studied the possibility that the vector field has no direct coupling to the inflaton, and it is produced by a different field. This case represents a non-trivial dynamics occurring in a "hidden sector." No sector is however completely hidden, since gravitational interactions are always present, and we saw that such particle production can be constrained by the amount of inflaton perturbations or GWs that it would produce gravitationally.

To be concrete, let us summarize the different cases that we explored in this thesis.

1. **Direct coupling to the inflaton**

   (a) Pseudo-scalar $\varphi F\tilde{F}$ interaction – Part IV

   (b) Dilaton-like $I^2(\varphi)\,F^2$ interaction – Part V

2. **No direct coupling** – Part VI

   (a) Instantaneous vector production from a sudden variation of its mass – Section VI 1

   (b) Continuous production from a rolling pseudo scalar (other than the inflaton) – Section VI 2

Case 1(a) is typical of models of axion inflation, which are highly motivated particle physics models of inflation, since they offer a mechanism to protect the flatness of the inflaton potential. We found a large, nearly equilateral non-Gaussianity. The non-detection of non-Gaussianity with Planck constrains the axion decay constant $f \gtrsim 10^{16}$ GeV, which is in the range that is naturally obtained in the UV completed theories of axion inflation. Future GW interferometer experiments, which probe much smaller scales than the CMB scales, can improve the bound over the CMB non-Gaussianity limits.

Case 1(b) applies to models that have been proposed as a mechanism of primordial magnetogenesis and anisotropic inflation. We saw that the resulting non-Gaussianity has a shape peaked at squeezed triangles with a unique angular dependence, which, if detected, would provide evidence for interaction processes that involve higher-spin fields.

Case 2(a) was studied in [99], whose original claim was that it could lead to a significant GW production. We found that this is not the case: the resultant GW signals are unlikely to be detected in the current and future CMB missions. This is because the quadrupole moments of the non-relativistic particles are highly suppressed. The scalar perturbations exhibit a bump in the power spectrum, and once the bound on this localized feature is respected, the amount of GWs produced in this model is completely unobservable ($r \lesssim 10^{-6}$), unless some additional mechanism is involved.

Case 2(b) was "constructed" by us as a prototype of a concrete example that can produce an interesting GW signal at CMB scales without conflicting with the CMB bounds on non-Gaussianity (we are not aware of any other working example). The GW signals from this model can be large regardless of the energy scale of inflation, and if they are large, one helicity state of the GWs is produced in a much stronger amount than the other state. This "chiral" GW signature is potentially detectable by the ongoing and forthcoming missions.

In summary, we showed that particle production would indeed provide novel predictions that were absent in the simplest inflationary models. The non-Gaussianity and GWs can be at the detectable level by the ongoing or forthcoming observational missions, with the features unique to the types of interactions that characterize the models. The Planck satellite mission has started releasing the results (see e.g. [11–13]) with more data to come, as well as there are several other ongoing/planned missions, such as SPIDER [100], SPT [85], ACT [194], EBEX [247], Advanced LIGO [162], Advanced Virgo [3], GEO-HF [309], and KAGRA [201]. Their improved sensitivities will constrain the inflationary parameters at the unprecedented level. Particle production can provide the mechanism that opens interesting observational windows of inflationary cosmology in prospect of these experiments.

References

[1] http://ligo.org/.

[2] http://lisa.nasa.gov/.

[3] https://wwwcascina.virgo.infn.it/advirgo.

[4] http://www.rssd.esa.int/index.php?project=PLANCK.

[5] The Scientific programme of planck. 2006.

[6] B. Abbott et al. An Upper Limit on the Stochastic Gravitational-Wave Background of Cosmological Origin. *Nature*, 460:990, 2009.

[7] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.

[8] V. Acquaviva, N. Bartolo, S. Matarrese, and A. Riotto. Second order cosmological perturbations from inflation. *Nucl.Phys.*, B667:119–148, 2003.

[9] F. C. Adams, J. R. Bond, K. Freese, J. A. Frieman, and A. V. Olinto. Natural inflation: Particle physics models, power law spectra for large scale structure, and constraints from COBE. *Phys.Rev.*, D47:426–455, 1993.

[10] P. Ade et al. Planck 2013 results. XV. CMB power spectra and likelihood. 2013.

[11] P. Ade et al. Planck 2013 results. XVI. Cosmological parameters. 2013.

[12] P. Ade et al. Planck 2013 results. XXII. Constraints on inflation. 2013.

[13] P. Ade et al. Planck 2013 Results. XXIV. Constraints on primordial non-Gaussianity. 2013.

[14] P. Adshead and M. Wyman. Chromo-Natural Inflation: Natural inflation on a steep potential with classical non-Abelian gauge fields. *Phys.Rev.Lett.*, 108:261302, 2012.

[15] N. Aghanim et al. Planck 2013 results. XXVII. Doppler boosting of the CMB: Eppur si muove. 2013.

[16] A. Albrecht and P. J. Steinhardt. Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking. *Phys.Rev.Lett.*, 48:1220–1223, 1982.

[17] S. H.-S. Alexander, M. E. Peskin, and M. M. Sheikh-Jabbari. Leptogenesis from gravity waves in models of inflation. *Phys.Rev.Lett.*, 96:081301, 2006.

[18] M. Alishahiha, E. Silverstein, and D. Tong. DBI in the sky. *Phys.Rev.*, D70:123505, 2004.

[19] R. Allahverdi, R. Brandenberger, F.-Y. Cyr-Racine, and A. Mazumdar. Reheating in Inflationary Cosmology: Theory and Applications. *Ann.Rev.Nucl.Part.Sci.*, 60:27–51, 2010.

[20] B. Allen and J. D. Romano. Detecting a stochastic background of gravitational radiation: Signal processing strategies and sensitivities. *Phys.Rev.*, D59:102001, 1999.

[21] M. M. Anber and L. Sorbo. N-flationary magnetic fields. *JCAP*, 0610:018, 2006.

[22] M. M. Anber and L. Sorbo. Naturally inflating on steep potentials through electromagnetic dissipation. *Phys.Rev.*, D81:043534, 2010.

[23] M. M. Anber and L. Sorbo. Non-Gaussianities and chiral gravitational waves in natural steep inflation. *Phys.Rev.*, D85:123537, 2012.

[24] R. Apreda, M. Maggiore, A. Nicolis, and A. Riotto. Gravitational waves from electroweak phase transitions. *Nucl.Phys.*, B631:342–368, 2002.

[25] N. Arkani-Hamed, H.-C. Cheng, P. Creminelli, and L. Randall. Extra natural inflation. *Phys.Rev.Lett.*, 90:221302, 2003.

[26] D. Babich, P. Creminelli, and M. Zaldarriaga. The Shape of non-Gaussianities. *JCAP*, 0408:009, 2004.

[27] V. Balasubramanian and P. Berglund. Stringy corrections to Kahler potentials, SUSY breaking, and the cosmological constant problem. *JHEP*, 0411:085, 2004.

[28] V. Balasubramanian, P. Berglund, J. P. Conlon, and F. Quevedo. Systematics of moduli stabilisation in Calabi-Yau flux compactifications. *JHEP*, 0503:007, 2005.

[29] K. Bamba and J. Yokoyama. Large scale magnetic fields from inflation in dilaton electromagnetism. *Phys.Rev.*, D69:043507, 2004.

[30] T. Banks, M. Dine, P. J. Fox, and E. Gorbatov. On the possibility of large axion decay constants. *JCAP*, 0306:001, 2003.

[31] T. Banks, D. B. Kaplan, and A. E. Nelson. Cosmological implications of dynamical supersymmetry breaking. *Phys.Rev.*, D49:779–787, 1994.

[32] J. M. Bardeen. Gauge Invariant Cosmological Perturbations. *Phys.Rev.*, D22:1882–1905, 1980.

[33] J. M. Bardeen, P. J. Steinhardt, and M. S. Turner. Spontaneous Creation of Almost Scale - Free Density Perturbations in an Inflationary Universe. *Phys.Rev.*, D28:679, 1983.

[34] N. Barnaby. Nongaussianity from Particle Production During Inflation. *Adv.Astron.*, 2010:156180, 2010.

[35] N. Barnaby. On Features and Nongaussianity from Inflationary Particle Production. *Phys.Rev.*, D82:106009, 2010.

[36] N. Barnaby, T. Biswas, and J. M. Cline. p-adic Inflation. *JHEP*, 0704:056, 2007.

[37] N. Barnaby, J. R. Bond, Z. Huang, and L. Kofman. Preheating After Modular Inflation. *JCAP*, 0912:021, 2009.

[38] N. Barnaby and J. M. Cline. Nongaussian and nonscale-invariant perturbations from tachyonic preheating in hybrid inflation. *Phys.Rev.*, D73:106012, 2006.

[39] N. Barnaby and J. M. Cline. Large Nongaussianity from Nonlocal Inflation. *JCAP*, 0707:017, 2007.

[40] N. Barnaby and J. M. Cline. Nongaussianity from Tachyonic Preheating in Hybrid Inflation. *Phys.Rev.*, D75:086004, 2007.

[41] N. Barnaby and J. M. Cline. Predictions for Nongaussianity from Nonlocal Inflation. *JCAP*, 0806:030, 2008.

[42] N. Barnaby and Z. Huang. Particle Production During Inflation: Observational Constraints and Signatures. *Phys.Rev.*, D80:126018, 2009.

[43] N. Barnaby, Z. Huang, L. Kofman, and D. Pogosyan. Cosmological Fluctuations from Infra-Red Cascading During Inflation. *Phys.Rev.*, D80:043501, 2009.

[44] N. Barnaby, J. Moxon, R. Namba, M. Peloso, G. Shiu, and P. Zhou. Gravity waves and non-Gaussian features from particle production in a sector gravitationally coupled to the inflaton. *Phys.Rev.*, D86:103508, 2012.

[45] N. Barnaby, R. Namba, and M. Peloso. Phenomenology of a Pseudo-Scalar Inflaton: Naturally Large Nongaussianity. *JCAP*, 1104:009, 2011.

[46] N. Barnaby, R. Namba, and M. Peloso. Observable non-gaussianity from gauge field production in slow roll inflation, and a challenging connection with magnetogenesis. *Phys.Rev.*, D85:123523, 2012.

[47] N. Barnaby, E. Pajer, and M. Peloso. Gauge Field Production in Axion Inflation: Consequences for Monodromy, non-Gaussianity in the CMB, and Gravitational Waves at Interferometers. *Phys.Rev.*, D85:023525, 2012.

[48] N. Barnaby and M. Peloso. Large Nongaussianity in Axion Inflation. *Phys.Rev.Lett.*, 106:181301, 2011.

[49] N. Barnaby and S. Shandera. Feeding your Inflaton: Non-Gaussian Signatures of Interaction Structure. *JCAP*, 1201:034, 2012.

[50] N. Bartolo, E. Dimastrogiovanni, S. Matarrese, and A. Riotto. Anisotropic bispectrum of curvature perturbations from primordial non-Abelian vector fields. *JCAP*, 0910:015, 2009.

[51] N. Bartolo, E. Dimastrogiovanni, S. Matarrese, and A. Riotto. Anisotropic Trispectrum of Curvature Perturbations Induced by Primordial Non-Abelian Vector Fields. *JCAP*, 0911:028, 2009.

[52] N. Bartolo, E. Komatsu, S. Matarrese, and A. Riotto. Non-Gaussianity from inflation: Theory and observations. *Phys.Rept.*, 402:103–266, 2004.

[53] N. Bartolo, S. Matarrese, M. Peloso, and A. Ricciardone. The anisotropic power spectrum and bispectrum in the $f(\phi)F^2$ mechanism. *Phys.Rev.*, D87:023504, 2013.

[54] N. Bartolo, S. Matarrese, A. Riotto, and A. Vaihkonen. The Maximal Amount of Gravitational Waves in the Curvaton Scenario. *Phys.Rev.*, D76:061302, 2007.

[55] B. A. Bassett, S. Tsujikawa, and D. Wands. Inflation dynamics and reheating. *Rev.Mod.Phys.*, 78:537–589, 2006.

[56] T. Battefeld and J. Grieb. Anatomy of bispectra in general single-field inflation – modal expansions. *JCAP*, 1112:003, 2011.

[57] D. Baumann, A. Dymarsky, I. R. Klebanov, L. McAllister, and P. J. Steinhardt. A Delicate universe. *Phys.Rev.Lett.*, 99:141601, 2007.

[58] D. Baumann et al. CMBPol Mission Concept Study: Probing Inflation with CMB Polarization. *AIP Conf.Proc.*, 1141:10–120, 2009.

[59] D. Baumann and L. McAllister. Advances in Inflation in String Theory. *Ann.Rev.Nucl.Part.Sci.*, 59:67–94, 2009.

[60] C. Bennett et al. First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Preliminary maps and basic results. *Astrophys.J.Suppl.*, 148:1, 2003.

[61] C. Bennett et al. Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results. 2012.

[62] A. Berera. Warm inflation. *Phys.Rev.Lett.*, 75:3218–3221, 1995.

[63] M. Berg, E. Pajer, and S. Sjors. Dante's Inferno. *Phys.Rev.*, D81:103535, 2010.

[64] P. Binetruy, A. Bohe, C. Caprini, and J.-F. Dufaux. Cosmological Backgrounds of Gravitational Waves and eLISA/NGO: Phase Transitions, Cosmic Strings and Other Sources. *JCAP*, 1206:027, 2012.

[65] J. Blanco-Pillado et al. Racetrack inflation. *JHEP*, 0411:063, 2004.

[66] J. Blanco-Pillado et al. Inflating in a better racetrack. *JHEP*, 0609:002, 2006.

[67] J. Bock et al. Study of the Experimental Probe of Inflationary Cosmology (EPIC)-Intemediate Mission for NASA's Einstein Inflation Probe. 2009.

[68] J. R. Bond, A. V. Frolov, Z. Huang, and L. Kofman. Non-Gaussian Spikes from Chaotic Billiards in Inflation Preheating. *Phys.Rev.Lett.*, 103:071301, 2009.

[69] J. R. Bond, L. Kofman, S. Prokushkin, and P. M. Vaudrevange. Roulette inflation with Kahler moduli and their axions. *Phys.Rev.*, D75:123511, 2007.

[70] L. A. Boyle and P. J. Steinhardt. Probing the early universe with inflationary gravitational waves. *Phys.Rev.*, D77:063504, 2008.

[71] J. Braden, L. Kofman, and N. Barnaby. Reheating the Universe After Multi-Field Inflation. *JCAP*, 1007:016, 2010.

[72] I. Brown and R. Crittenden. Non-Gaussianity from cosmic magnetic fields. *Phys.Rev.*, D72:063002, 2005.

[73] C. Burgess, J. M. Cline, and M. Postma. Axionic D3-D7 Inflation. *JHEP*, 0903:058, 2009.

[74] C. Burgess, L. Leblond, R. Holman, and S. Shandera. Super-Hubble de Sitter Fluctuations and the Dynamical RG. *JCAP*, 1003:033, 2010.

[75] C. T. Byrnes, K.-Y. Choi, and L. M. Hall. Conditions for large non-Gaussianity in two-field slow-roll inflation. *JCAP*, 0810:008, 2008.

[76] C. T. Byrnes, L. Hollenstein, R. K. Jain, and F. R. Urban. Resonant magnetic fields from inflation. *JCAP*, 1203:009, 2012.

[77] C. T. Byrnes and G. Tasinato. Non-Gaussianity beyond slow roll in multi-field inflation. *JCAP*, 0908:016, 2009.

[78] R.-G. Cai, B. Hu, and H.-B. Zhang. Acoustic signatures in the Cosmic Microwave Background bispectrum from primordial magnetic fields. *JCAP*, 1008:025, 2010.

[79] R. R. Caldwell, L. Motta, and M. Kamionkowski. Correlation of inflation-produced magnetic fields with scalar fluctuations. *Phys.Rev.*, D84:123525, 2011.

[80] P. Candelas, E. Perevalov, and G. Rajesh. Toric geometry and enhanced gauge symmetry of F theory / heterotic vacua. *Nucl.Phys.*, B507:445–474, 1997.

[81] C. Caprini and R. Durrer. Gravitational waves from stochastic relativistic sources: Primordial turbulence and magnetic fields. *Phys.Rev.*, D74:063521, 2006.

[82] C. Caprini, R. Durrer, and G. Servant. Gravitational wave generation from bubble collisions in first-order phase transitions: An analytic approach. *Phys.Rev.*, D77:124015, 2008.

[83] C. Caprini, R. Durrer, and G. Servant. The stochastic gravitational wave background from turbulence and magnetic fields generated by a first-order phase transition. *JCAP*, 0912:024, 2009.

[84] C. Caprini, F. Finelli, D. Paoletti, and A. Riotto. The cosmic microwave background temperature bispectrum from scalar perturbations induced by primordial magnetic fields. *JCAP*, 0906:021, 2009.

[85] J. Carlstrom et al. The 10 Meter South Pole Telescope. *Publ.Astron.Soc.Pac.*, 123:568–581, 2011.

[86] A. Chambers and A. Rajantie. Lattice calculation of non-Gaussianity from preheating. *Phys.Rev.Lett.*, 100:041302, 2008.

[87] T. Chantavat, C. Gordon, and J. Silk. Large Scale Structure Forecast Constraints on Particle Production During Inflation. *Phys.Rev.*, D83:103501, 2011.

[88] X. Chen, R. Easther, and E. A. Lim. Large Non-Gaussianities in Single Field Inflation. *JCAP*, 0706:023, 2007.

[89] X. Chen, R. Easther, and E. A. Lim. Generation and Characterization of Large Non-Gaussianities in Single Field Inflation. *JCAP*, 0804:010, 2008.

[90] X. Chen, M.-x. Huang, S. Kachru, and G. Shiu. Observational signatures and non-Gaussianities of general single field inflation. *JCAP*, 0701:002, 2007.

[91] X. Chen and Y. Wang. Quasi-Single Field Inflation and Non-Gaussianities. *JCAP*, 1004:027, 2010.

[92] D. Chialva. Gravitational waves from first order phase transitions during inflation. *Phys.Rev.*, D83:023512, 2011.

[93] J. Chluba, A. L. Erickcek, and I. Ben-Dayan. Probing the inflaton: Small-scale power spectrum constraints from measurements of the CMB energy spectrum. *Astrophys.J.*, 758:76, 2012.

[94] D. J. Chung, E. W. Kolb, A. Riotto, and I. I. Tkachev. Probing Planckian physics: Resonant production of particles during inflation and features in the primordial power spectrum. *Phys.Rev.*, D62:043508, 2000.

[95] A. Coc, N. J. Nunes, K. A. Olive, J.-P. Uzan, and E. Vangioni. Coupled Variations of Fundamental Couplings and Primordial Nucleosynthesis. *Phys.Rev.*, D76:023511, 2007.

[96] J. P. Conlon and F. Quevedo. Kahler moduli inflation. *JHEP*, 0601:146, 2006.

[97] J. P. Conlon and F. Quevedo. Astrophysical and cosmological implications of large volume string compactifications. *JCAP*, 0708:019, 2007.

[98] J. P. Conlon, F. Quevedo, and K. Suruliz. Large-volume flux compactifications: Moduli spectrum and D3/D7 soft supersymmetry breaking. *JHEP*, 0508:007, 2005.

[99] J. L. Cook and L. Sorbo. Particle production during inflation and gravitational waves detectable by ground-based interferometers. *Phys.Rev.*, D85:023534, 2012.

[100] B. Crill et al. SPIDER: A Balloon-borne Large-scale CMB Polarimeter. 2008.

[101] S. Crowder, R. Namba, V. Mandic, S. Mukohyama, and M. Peloso. Measurement of Parity Violation in the Early Universe using Gravitational-wave Detectors. 2012.

[102] C. Cutler. Gravitational waves from neutron stars with large toroidal B fields. *Phys.Rev.*, D66:084025, 2002.

[103] R. H. Cyburt, B. D. Fields, K. A. Olive, and E. Skillman. New BBN limits on physics beyond the standard model from He-4. *Astropart.Phys.*, 23:313–323, 2005.

[104] T. Damour and A. Vilenkin. Cosmic strings and the string dilaton. *Phys.Rev.Lett.*, 78:2288–2291, 1997.

[105] T. Damour and A. Vilenkin. Gravitational wave bursts from cosmic strings. *Phys.Rev.Lett.*, 85:3761–3764, 2000.

[106] B. de Carlos, J. Casas, F. Quevedo, and E. Roulet. Model independent properties and cosmological implications of the dilaton and moduli sectors of 4-d strings. *Phys.Lett.*, B318:447–456, 1993.

[107] V. Demozzi, V. Mukhanov, and H. Rubinstein. Magnetic fields from inflation? *JCAP*, 0908:025, 2009.

[108] R. Dicke, P. Peebles, P. Roll, and D. Wilkinson. Cosmic Black-Body Radiation. *Astrophys.J.*, 142:414–419, 1965.

[109] K. Dimopoulos. Can a vector field be responsible for the curvature perturbation in the Universe? *Phys.Rev.*, D74:083502, 2006.

[110] K. Dimopoulos. Correlated curvature perturbations and magnetogenesis from the GUT gauge bosons. *Astropart.Phys.*, 42:86–89, 2013.

[111] K. Dimopoulos and M. Karciauskas. Parity Violating Statistical Anisotropy. *JHEP*, 1206:040, 2012.

[112] K. Dimopoulos, M. Karciauskas, D. H. Lyth, and Y. Rodriguez. Statistical anisotropy of the curvature perturbation from vector field perturbations. *JCAP*, 0905:013, 2009.

[113] K. Dimopoulos, M. Karciauskas, and J. M. Wagstaff. Vector Curvaton with varying Kinetic Function. *Phys.Rev.*, D81:023522, 2010.

[114] K. Dimopoulos, M. Karciauskas, and J. M. Wagstaff. Vector Curvaton without Instabilities. *Phys.Lett.*, B683:298–301, 2010.

[115] K. Dimopoulos, D. Wills, and I. Zavala. Statistical Anisotropy from Vector Curvaton in D-brane Inflation. *Nucl.Phys.*, B868:120–155, 2013.

[116] S. Dimopoulos, S. Kachru, J. McGreevy, and J. G. Wacker. N-flation. *JCAP*, 0808:003, 2008.

[117] A. D. Dolgov, D. Grasso, and A. Nicolis. Relic backgrounds of gravitational waves from cosmic turbulence. *Phys.Rev.*, D66:103505, 2002.

[118] A. G. Doroshkevich and I. D. Novikov. *Dokl. Akad. Nauk SSSR*, 154:809, 1964.

[119] J. F. Dufaux, A. Bergman, G. N. Felder, L. Kofman, and J.-P. Uzan. Theory and Numerics of Gravitational Waves from Preheating after Inflation. *Phys.Rev.*, D76:123517, 2007.

[120] J.-F. Dufaux, G. Felder, L. Kofman, and O. Navros. Gravity Waves from Tachyonic Preheating after Hybrid Inflation. *JCAP*, 0903:001, 2009.

[121] J.-F. Dufaux, D. G. Figueroa, and J. Garcia-Bellido. Gravitational Waves from Abelian Gauge Fields and Cosmic Strings at Preheating. *Phys.Rev.*, D82:083518, 2010.

[122] J. Dunkley et al. The Atacama Cosmology Telescope: Cosmological Parameters from the 2008 Power Spectra. *Astrophys.J.*, 739:52, 2011.

[123] R. Durrer, L. Hollenstein, and R. K. Jain. Can slow roll inflation induce relevant helical magnetic fields? *JCAP*, 1103:037, 2011.

[124] R. Easther, J. Giblin, John T., and E. A. Lim. Gravitational Wave Production At The End Of Inflation. *Phys.Rev.Lett.*, 99:221301, 2007.

[125] R. Easther and L. McAllister. Random matrices and the spectrum of N-flation. *JCAP*, 0605:018, 2006.

[126] M. Farhang, J. R. Bond, O. Dore, and C. B. Netterfield. Primordial Gravitational Wave Detectability with Deep Small-Sky CMB Experiments. 2011.

[127] E. Fenu, D. G. Figueroa, R. Durrer, and J. Garcia-Bellido. Gravitational waves from self-ordering scalar fields. *JCAP*, 0910:005, 2009.

[128] S. Ferrara and R. Kallosh. Creation of Matter in the Universe and Groups of Type E7. *JHEP*, 1112:096, 2011.

[129] R. Flauger, L. McAllister, E. Pajer, A. Westphal, and G. Xu. Oscillations in the CMB from Axion Monodromy Inflation. *JCAP*, 1006:009, 2010.

[130] R. Flauger and E. Pajer. Resonant Non-Gaussianity. *JCAP*, 1101:017, 2011.

[131] J. Fonseca and D. Wands. Non-Gaussianity and Gravitational Waves from Quadratic and Self-interacting Curvaton. *Phys.Rev.*, D83:064025, 2011.

[132] K. Freese, J. A. Frieman, and A. V. Olinto. Natural inflation with pseudo - Nambu-Goldstone bosons. *Phys.Rev.Lett.*, 65:3233–3236, 1990.

[133] A. Friedman. On the Curvature of space. *Z.Phys.*, 10:377–386, 1922.

[134] G. Gamow. The origin of elements and the separation of galaxies. *Phys. Rev.*, 74:505–506, Aug 1948.

[135] A. Gangui, F. Lucchin, S. Matarrese, and S. Mollerach. The Three point correlation function of the cosmic microwave background in inflationary models. *Astrophys.J.*, 430:447–457, 1994.

[136] J. Garcia-Bellido and D. G. Figueroa. A stochastic background of gravitational waves from hybrid preheating. *Phys.Rev.Lett.*, 98:061302, 2007.

[137] J. Garcia-Bellido, D. G. Figueroa, and A. Sastre. A Gravitational Wave Background from Reheating after Hybrid Inflation. *Phys.Rev.*, D77:043517, 2008.

[138] M. Gasperini, M. Giovannini, and G. Veneziano. Primordial magnetic fields from string cosmology. *Phys.Rev.Lett.*, 75:3796–3799, 1995.

[139] M. Gasperini and G. Veneziano. Pre - big bang in string cosmology. *Astropart.Phys.*, 1:317–339, 1993.

[140] C. Germani and A. Kehagias. UV-Protected Inflation. *Phys.Rev.Lett.*, 106:161302, 2011.

[141] S. Ghigna, M. Lusignoli, and M. Roncadelli. Instability of the invisible axion. *Phys.Lett.*, B283:278–281, 1992.

[142] S. B. Giddings and A. Strominger. Axion Induced Topology Change in Quantum Gravity and String Theory. *Nucl.Phys.*, B306:890, 1988.

[143] S. B. Giddings and A. Strominger. STRING WORMHOLES. *Phys.Lett.*, B230:46, 1989.

[144] M. Giovannini. The Magnetized universe. *Int.J.Mod.Phys.*, D13:391–502, 2004.

[145] M. Giovannini. Electric-magnetic duality and the conditions of inflationary magnetogenesis. *JCAP*, 1004:003, 2010.

[146] G. Giudice, M. Peloso, A. Riotto, and I. Tkachev. Production of massive fermions at preheating and leptogenesis. *JHEP*, 9908:014, 1999.

[147] V. Gluscevic and M. Kamionkowski. Testing Parity-Violating Mechanisms with Cosmic Microwave Background Experiments. *Phys.Rev.*, D81:123529, 2010.

[148] A. Golovnev, V. Mukhanov, and V. Vanchurin. Vector Inflation. *JCAP*, 0806:009, 2008.

[149] D. Grasso and H. R. Rubinstein. Magnetic fields in the early universe. *Phys.Rept.*, 348:163–266, 2001.

[150] D. Green, B. Horn, L. Senatore, and E. Silverstein. Trapped Inflation. *Phys.Rev.*, D80:063533, 2009.

[151] D. R. Green. Reheating Closed String Inflation. *Phys.Rev.*, D76:103504, 2007.

[152] P. B. Greene and L. Kofman. Preheating of fermions. *Phys.Lett.*, B448:6–12, 1999.

[153] P. B. Greene and L. Kofman. On the theory of fermionic preheating. *Phys.Rev.*, D62:123516, 2000.

[154] T. W. Grimm. Axion inflation in type II string theory. *Phys.Rev.*, D77:126007, 2008.

[155] L. Grishchuk. Amplification of gravitational waves in an istropic universe. *Sov.Phys.JETP*, 40:409–415, 1975.

[156] C. Grojean and G. Servant. Gravitational Waves from Phase Transitions at the Electroweak Scale and Beyond. *Phys.Rev.*, D75:043507, 2007.

[157] A. Gumrukcuoglu, L. Kofman, and M. Peloso. Gravity Waves Signatures from Anisotropic pre-Inflation. *Phys.Rev.*, D78:103525, 2008.

[158] S. Gupta, A. Berera, A. Heavens, and S. Matarrese. Non-Gaussian signatures in the cosmic background radiation from warm inflation. *Phys.Rev.*, D66:043510, 2002.

[159] A. H. Guth. The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems. *Phys.Rev.*, D23:347–356, 1981.

[160] A. H. Guth and E. J. Weinberg. Could the Universe Have Recovered from a Slow First Order Phase Transition? *Nucl.Phys.*, B212:321, 1983.

[161] S. Hannestad, T. Haugbolle, P. R. Jarnhus, and M. S. Sloth. Non-Gaussianity from Axion Monodromy Inflation. *JCAP*, 1006:001, 2010.

[162] G. M. Harry. Advanced LIGO: The next generation of gravitational wave detectors. *Class.Quant.Grav.*, 27:084006, 2010.

[163] S. Hawking, I. Moss, and J. Stewart. Bubble Collisions in the Very Early Universe. *Phys.Rev.*, D26:2681, 1982.

[164] B. Himmetoglu, C. R. Contaldi, and M. Peloso. Ghost instabilities of cosmological models with vector fields nonminimally coupled to the curvature. *Phys.Rev.*, D80:123530, 2009.

[165] B. Himmetoglu, C. R. Contaldi, and M. Peloso. Instability of anisotropic cosmological solutions supported by vector fields. *Phys.Rev.Lett.*, 102:111301, 2009.

[166] G. Hinshaw et al. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Data Processing, Sky Maps, and Basic Results. *Astrophys.J.Suppl.*, 180:225–245, 2009.

[167] G. Hinshaw et al. Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. 2012.

[168] B. Holdom. Two U(1)'s and Epsilon Charge Shifts. *Phys.Lett.*, B166:196, 1986.

[169] R. Holman et al. Solutions to the strong CP problem in a world with gravity. *Phys.Lett.*, B282:132–136, 1992.

[170] R. Holman and A. J. Tolley. Enhanced Non-Gaussianity from Excited Initial States. *JCAP*, 0805:001, 2008.

[171] Y. Hu, M. S. Turner, and E. J. Weinberg. Dynamical solutions to the horizon and flatness problems. *Phys.Rev.*, D49:3830–3836, 1994.

[172] S. J. Huber and T. Konstandin. Gravitational Wave Production by Collisions: More Bubbles. *JCAP*, 0809:022, 2008.

[173] K. Jones-Smith, L. M. Krauss, and H. Mathur. A Nearly Scale Invariant Spectrum of Gravitational Radiation from Global Phase Transitions. *Phys.Rev.Lett.*, 100:131302, 2008.

[174] R. Kallosh. On inflation in string theory. *Lect.Notes Phys.*, 738:119–156, 2008.

[175] R. Kallosh, A. D. Linde, D. A. Linde, and L. Susskind. Gravity and global symmetries. *Phys.Rev.*, D52:912–935, 1995.

[176] R. Kallosh, N. Sivanandam, and M. Soroush. Axion Inflation and Gravity Waves in String Theory. *Phys.Rev.*, D77:043501, 2008.

[177] N. Kaloper, A. Lawrence, and L. Sorbo. An Ignoble Approach to Large Field Inflation. *JCAP*, 1103:023, 2011.

[178] N. Kaloper and L. Sorbo. A Natural Framework for Chaotic Inflation. *Phys.Rev.Lett.*, 102:121301, 2009.

[179] M. Kamionkowski, A. Kosowsky, and M. S. Turner. Gravitational radiation from first order phase transitions. *Phys.Rev.*, D49:2837–2851, 1994.

[180] M. Kamionkowski and J. March-Russell. Planck scale physics and the Peccei-Quinn mechanism. *Phys.Lett.*, B282:137–141, 1992.

[181] A. Kandus, K. E. Kunze, and C. G. Tsagas. Primordial magnetogenesis. *Phys.Rept.*, 505:1–58, 2011.

[182] S. Kanno, J. Soda, and M.-a. Watanabe. Cosmological Magnetic Fields from Inflation and Backreaction. *JCAP*, 0912:009, 2009.

[183] M. Karciauskas. The Primordial Curvature Perturbation from Vector Fields of General non-Abelian Groups. *JCAP*, 1201:014, 2012.

[184] M. Karciauskas, K. Dimopoulos, and D. H. Lyth. Anisotropic non-Gaussianity from vector field perturbations. *Phys.Rev.*, D80:023509, 2009.

[185] S. Kawamura et al. The Japanese space gravitational wave antenna: DECIGO. *Class.Quant.Grav.*, 28:094011, 2011.

[186] M. Kawasaki, K. Kohri, T. Moroi, and A. Yotsuyanagi. Big-Bang Nucleosynthesis and Gravitino. *Phys.Rev.*, D78:065011, 2008.

[187] J. E. Kim, H. P. Nilles, and M. Peloso. Completing natural inflation. *JCAP*, 0501:005, 2005.

[188] L. Kofman, A. D. Linde, X. Liu, A. Maloney, L. McAllister, et al. Beauty is attractive: Moduli trapping at enhanced symmetry points. *JHEP*, 0405:030, 2004.

[189] L. Kofman, A. D. Linde, and A. A. Starobinsky. Reheating after inflation. *Phys.Rev.Lett.*, 73:3195–3198, 1994.

[190] L. Kofman, A. D. Linde, and A. A. Starobinsky. Towards the theory of reheating after inflation. *Phys.Rev.*, D56:3258–3295, 1997.

[191] E. Komatsu et al. Non-Gaussianity as a Probe of the Physics of the Primordial Universe and the Astrophysics of the Low Redshift Universe. 2009.

[192] E. Komatsu et al. Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *Astrophys.J.Suppl.*, 192:18, 2011.

[193] E. Komatsu and D. N. Spergel. Acoustic signatures in the primary microwave background bispectrum. *Phys.Rev.*, D63:063002, 2001.

[194] A. Kosowsky. The Atacama Cosmology Telescope Project: A Progress Report. *New Astron.Rev.*, 50:969–976, 2006.

[195] A. Kosowsky, A. Mack, and T. Kahniashvili. Gravitational radiation from cosmological turbulence. *Phys.Rev.*, D66:024030, 2002.

[196] A. Kosowsky, M. S. Turner, and R. Watkins. Gravitational radiation from colliding vacuum bubbles. *Phys.Rev.*, D45:4514–4535, 1992.

[197] A. Kosowsky, M. S. Turner, and R. Watkins. Gravitational waves from first order cosmological phase transitions. *Phys.Rev.Lett.*, 69:2026–2029, 1992.

[198] L. M. Krauss. Gravitational waves from global phase transitions. *Phys.Lett.*, B284:229–233, 1992.

[199] L. M. Krauss, K. Jones-Smith, H. Mathur, and J. Dent. Probing the Gravitational Wave Signature from Cosmic Phase Transitions at Different Scales. *Phys.Rev.*, D82:044001, 2010.

[200] J. Kumar, L. Leblond, and A. Rajaraman. Scale Dependent Local Non-Gaussianity from Loops. *JCAP*, 1004:024, 2010.

[201] K. Kuroda. Status of LCGT. *Class.Quant.Grav.*, 27:084004, 2010.

[202] L. Leblond and E. Pajer. Resonant Trispectrum and a Dozen More Primordial N-point functions. *JCAP*, 1101:035, 2011.

[203] G. Lemaitre. A homogeneous Universe of constant mass and growing radius accounting for the radial velocity of extragalactic nebulae. *Annales Soc.Sci.Brux.Ser.I Sci.Math.Astron.Phys.*, A47:49–59, 1927.

[204] A. Lewis. The real shape of non-Gaussianities. *JCAP*, 1110:026, 2011.

[205] A. R. Liddle. Inflation as the unique causal mechanism for generating density perturbations on scales well above the Hubble radius. *Phys.Rev.*, D51:5347–5351, 1995.

[206] A. R. Liddle and D. Lyth. *Cosmological inflation and large scale structure.* 2000.

[207] A. R. Liddle, A. Mazumdar, and F. E. Schunck. Assisted inflation. *Phys.Rev.*, D58:061301, 1998.

[208] A. Linde, S. Mooij, and E. Pajer. Gauge field production in SUGRA inflation: local non-Gaussianity and primordial black holes. 2012.

[209] A. D. Linde. A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems. *Phys.Lett.*, B108:389–393, 1982.

[210] A. D. Linde. Coleman-Weinberg Theory and a New Inflationary Universe Scenario. *Phys.Lett.*, B114:431, 1982.

[211] A. D. Linde. *Particle physics and inflationary cosmology*, volume 5. 1990.

[212] A. D. Linde. Hybrid inflation. *Phys.Rev.*, D49:748–754, 1994.

[213] A. D. Linde. Inflationary Cosmology. *Lect.Notes Phys.*, 738:1–54, 2008.

[214] D. Lopez Nacir, R. A. Porto, L. Senatore, and M. Zaldarriaga. Dissipative effects in the Effective Field Theory of Inflation. *JHEP*, 1201:075, 2012.

[215] M. LoVerde, A. Miller, S. Shandera, and L. Verde. Effects of Scale-Dependent Non-Gaussianity on Cosmological Structures. *JCAP*, 0804:014, 2008.

[216] A. Lue, L.-M. Wang, and M. Kamionkowski. Cosmological signature of new parity violating interactions. *Phys.Rev.Lett.*, 83:1506–1509, 1999.

[217] D. Lyth. Large Scale Energy Density Perturbations and Inflation. *Phys.Rev.*, D31:1792–1798, 1985.

[218] D. H. Lyth. What would we learn by detecting a gravitational wave signal in the cosmic microwave background anisotropy? *Phys.Rev.Lett.*, 78:1861–1863, 1997.

[219] D. H. Lyth and A. Riotto. Particle physics models of inflation and the cosmological density perturbation. *Phys.Rept.*, 314:1–146, 1999.

[220] D. H. Lyth, C. Ungarelli, and D. Wands. The Primordial density perturbation in the curvaton scenario. *Phys.Rev.*, D67:023503, 2003.

[221] J. Magueijo and D. M. Benincasa. Chiral vacuum fluctuations in quantum gravity. *Phys.Rev.Lett.*, 106:121302, 2011.

[222] D. Maity. Kinetic Gravity Braiding and axion inflation. *Phys.Lett.*, B720:389–392, 2013.

[223] J. M. Maldacena. Non-Gaussian features of primordial fluctuations in single field inflationary models. *JHEP*, 0305:013, 2003.

[224] A. Maleknejad and M. Sheikh-Jabbari. Gauge-flation: Inflation From Non-Abelian Gauge Fields. 2011.

[225] K. A. Malik. A not so short note on the Klein-Gordon equation at second order. *JCAP*, 0703:004, 2007.

[226] K. A. Malik and D. Wands. Cosmological perturbations. *Phys.Rept.*, 475:1–51, 2009.

[227] V. Mandic, E. Thrane, S. Giampanis, and T. Regimbau. Parameter Estimation in Searches for the Stochastic Gravitational-Wave Background. *Phys.Rev.Lett.*, 109:171102, 2012.

[228] S. Marassi, R. Ciolfi, R. Schneider, L. Stella, and V. Ferrari. Stochastic background of gravitational waves emitted by magnetars. 2010.

[229] J. Martin and J. Yokoyama. Generation of Large-Scale Magnetic Fields in Single-Field Inflation. *JCAP*, 0801:025, 2008.

[230] X. Martin and A. Vilenkin. Gravitational radiation from monopoles connected by strings. *Phys.Rev.*, D55:6054–6060, 1997.

[231] L. McAllister, E. Silverstein, and A. Westphal. Gravity Waves and Linear Inflation from Axion Monodromy. *Phys.Rev.*, D82:046003, 2010.

[232] P. D. Meerburg and E. Pajer. Observational Constraints on Gauge Field Production in Axion Inflation. *JCAP*, 1302:017, 2013.

[233] P. D. Meerburg, J. P. van der Schaar, and P. S. Corasaniti. Signatures of Initial State Modifications on Bispectrum Statistics. *JCAP*, 0905:018, 2009.

[234] P. D. Meerburg, J. P. van der Schaar, and M. G. Jackson. Bispectrum signatures of a modified vacuum in single field inflation with a small speed of sound. *JCAP*, 1002:001, 2010.

[235] A. Megevand. Gravitational waves from deflagration bubbles in first-order phase transitions. *Phys.Rev.*, D78:084003, 2008.

[236] A. Misra and P. Shukla. Large Volume Axionic Swiss-Cheese Inflation. *Nucl.Phys.*, B800:384–400, 2008.

[237] S. Mohanty and A. Nautiyal. Natural inflation at the GUT scale. *Phys.Rev.*, D78:123515, 2008.

[238] I. G. Moss and C. Xiong. Non-Gaussianity in fluctuations from warm inflation. *JCAP*, 0704:007, 2007.

[239] V. Mukhanov. *Physical foundations of cosmology.* 2005.

[240] V. F. Mukhanov. Gravitational Instability of the Universe Filled with a Scalar Field. *JETP Lett.*, 41:493–496, 1985.

[241] K. Nakamura et al. Review of particle physics. *J.Phys.*, G37:075021, 2010.

[242] K. Nakayama and J. Yokoyama. Gravitational Wave Background and Non-Gaussianity as a Probe of the Curvaton Scenario. *JCAP*, 1001:010, 2010.

[243] A. Nicolis. Relic gravitational waves from colliding bubbles and cosmic turbulence. *Class.Quant.Grav.*, 21:L27, 2004.

[244] H. P. Nilles, M. Peloso, and L. Sorbo. Coupled fields in external background with application to nonthermal production of gravitinos. *JHEP*, 0104:004, 2001.

[245] J. Ohashi and S. Tsujikawa. Potential-driven Galileon inflation. *JCAP*, 1210:035, 2012.

[246] S. Olmez, V. Mandic, and X. Siemens. Gravitational-Wave Stochastic Background from Kinks and Cusps on Cosmic Strings. *Phys.Rev.*, D81:104028, 2010.

[247] P. Oxley et al. The EBEX experiment. *Proc.SPIE Int.Soc.Opt.Eng.*, 5543:320–331, 2004.

[248] E. Pajer and M. Peloso. A review of Axion Inflation in the era of Planck. 2013.

[249] M. Peloso and L. Sorbo. Preheating of massive fermions after inflation: Analytical results. *JHEP*, 0005:016, 2000.

[250] A. A. Penzias and R. W. Wilson. A Measurement of excess antenna temperature at 4080-Mc/s. *Astrophys.J.*, 142:419–421, 1965.

[251] E. Phinney. The Rate of neutron star binary mergers in the universe: Minimal predictions for gravity wave detector. *Astrophys.J.*, 380:L17–L21, 1991.

[252] D. I. Podolsky, G. N. Felder, L. Kofman, and M. Peloso. Equation of state and beginning of thermalization after preheating. *Phys.Rev.*, D73:023501, 2006.

[253] D. Polarski and A. A. Starobinsky. Semiclassicality and decoherence of cosmological perturbations. *Class.Quant.Grav.*, 13:377–392, 1996.

[254] J. Preskill. MAGNETIC MONOPOLES. *Ann.Rev.Nucl.Part.Sci.*, 34:461–530, 1984.

[255] B. Ratra. Cosmological 'seed' magnetic field from inflation. *Astrophys.J.*, 391:L1–L4, 1992.

[256] T. Regimbau and J. A. de Freitas Pacheco. Cosmic background of gravitational waves from rotating neutron stars. *Astron. and Astrophys.*, 376:381, 2001.

[257] C. Reichardt et al. A measurement of secondary cosmic microwave background anisotropies with two years of South Pole Telescope observations. *Astrophys.J.*, 755:70, 2012.

[258] G. Rigopoulos, E. Shellard, and B. van Tent. Large non-Gaussianity in multiple-field inflation. *Phys.Rev.*, D73:083522, 2006.

[259] H. Robertson. Kinematics and World-Structure. *Astrophys.J.*, 82:284–301, 1935.

[260] A. E. Romano and M. Sasaki. Effects of particle production during inflation. *Phys.Rev.*, D78:103522, 2008.

[261] D. Ryu, D. R. Schleicher, R. A. Treumann, C. G. Tsagas, and L. M. Widrow. Magnetic fields in the Large-Scale Structure of the Universe. *Space Sci.Rev.*, 166:1–35, 2012.

[262] S. Saito, K. Ichiki, and A. Taruya. Probing polarization states of primordial gravitational waves with CMB anisotropies. *JCAP*, 0709:002, 2007.

[263] D. Salopek and J. Bond. Nonlinear evolution of long wavelength metric fluctuations in inflationary models. *Phys.Rev.*, D42:3936–3962, 1990.

[264] M. Sasaki. Large Scale Quantum Fluctuations in the Inflationary Universe. *Prog.Theor.Phys.*, 76:1036, 1986.

[265] C. Savage, K. Freese, and W. H. Kinney. Natural Inflation: Status after WMAP 3-year data. *Phys.Rev.*, D74:123511, 2006.

[266] T. E. science team. "einstein gravitational wave telescope conceptual design study". *https://tds.ego-gw.it/ql/?c=7954*, 2011.

[267] D. Seery. One-loop corrections to a scalar field during inflation. *JCAP*, 0711:025, 2007.

[268] D. Seery. Magnetogenesis and the primordial non-gaussianity. *JCAP*, 0908:018, 2009.

[269] D. Seery and J. E. Lidsey. Primordial non-Gaussianities in single field inflation. *JCAP*, 0506:003, 2005.

[270] D. Seery, K. A. Malik, and D. H. Lyth. Non-gaussianity of inflationary field perturbations from the field equation. *JCAP*, 0803:014, 2008.

[271] L. Senatore, E. Silverstein, and M. Zaldarriaga. New Sources of Gravitational Waves during Inflation. 2011.

[272] L. Senatore, K. M. Smith, and M. Zaldarriaga. Non-Gaussianities in Single Field Inflation and their Optimal Limits from the WMAP 5-year Data. *JCAP*, 1001:028, 2010.

[273] L. Senatore and M. Zaldarriaga. On Loops in Inflation. *JHEP*, 1012:008, 2010.

[274] T. Seshadri and K. Subramanian. CMB bispectrum from primordial magnetic fields on large angular scales. *Phys.Rev.Lett.*, 103:081303, 2009.

[275] N. Seto and A. Taruya. Measuring a Parity Violation Signature in the Early Universe via Ground-based Laser Interferometers. *Phys.Rev.Lett.*, 99:121101, 2007.

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

[277] J. R. Shaw and A. Lewis. Constraining Primordial Magnetism. *Phys.Rev.*, D86:043510, 2012.

[278] M. Shiraishi, D. Nitta, S. Yokoyama, K. Ichiki, and K. Takahashi. Cosmic microwave background bispectrum of vector modes induced from primordial magnetic fields. *Phys.Rev.*, D82:121302, 2010.

[279] X. Siemens, V. Mandic, and J. Creighton. Gravitational wave stochastic background from cosmic (super)strings. *Phys.Rev.Lett.*, 98:111101, 2007.

[280] J. L. Sievers et al. The Atacama Cosmology Telescope: Cosmological parameters from three seasons of data. 2013.

[281] E. Silverstein and A. Westphal. Monodromy in the CMB: Gravity Waves and String Inflation. *Phys.Rev.*, D78:106003, 2008.

[282] G. F. Smoot et al. Structure in the COBE differential microwave radiometer first year maps. *Astrophys.J.*, 396:L1–L5, 1992.

[283] J. Soda. Statistical Anisotropy from Anisotropic Inflation. *Class.Quant.Grav.*, 29:083001, 2012.

[284] L. Sorbo. Parity violation in the Cosmic Microwave Background from a pseudoscalar inflaton. *JCAP*, 1106:003, 2011.

[285] D. Spergel et al. First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Determination of cosmological parameters. *Astrophys.J.Suppl.*, 148:175–194, 2003.

[286] A. A. Starobinsky. Relict Gravitation Radiation Spectrum and Initial State of the Universe. (In Russian). *JETP Lett.*, 30:682–685, 1979.

[287] A. A. Starobinsky. Dynamics of Phase Transition in the New Inflationary Universe Scenario and Generation of Perturbations. *Phys.Lett.*, B117:175–178, 1982.

[288] P. J. Steinhardt. RELATIVISTIC DETONATION WAVES AND BUBBLE GROWTH IN FALSE VACUUM DECAY. *Phys.Rev.*, D25:2074, 1982.

[289] K. Story et al. A Measurement of the Cosmic Microwave Background Damping Tail from the 2500-square-degree SPT-SZ survey. 2012.

[290] K. Subramanian. Magnetic fields in the early universe. *Astron.Nachr.*, 331:110–120, 2010.

[291] P. Svrcek and E. Witten. Axions In String Theory. *JHEP*, 0606:051, 2006.

[292] T. Takahashi and J. Soda. Chiral Primordial Gravitational Waves from a Lifshitz Point. *Phys.Rev.Lett.*, 102:231301, 2009.

[293] P. Trivedi, T. Seshadri, and K. Subramanian. Cosmic Microwave Background Trispectrum and Primordial Magnetic Field Limits. *Phys.Rev.Lett.*, 108:231301, 2012.

[294] P. Trivedi, K. Subramanian, and T. Seshadri. Primordial Magnetic Field Limits from Cosmic Microwave Background Bispectrum of Magnetic Passive Scalar Modes. *Phys.Rev.*, D82:123006, 2010.

[295] M. S. Turner and L. M. Widrow. Inflation Produced, Large Scale Magnetic Fields. *Phys.Rev.*, D37:2743, 1988.

[296] T. Uchiyama et al. Present status of large-scale cryogenic gravitational wave telescope. *Class.Quant.Grav.*, 21:S1161–S1172, 2004.

[297] F. Vernizzi and D. Wands. Non-gaussianities in two-field inflation. *JCAP*, 0605:019, 2006.

[298] A. Vilenkin and E. Shellard. *Cosmic Strings and Other Topological Defects*. Cambridge University Press, 1994.

[299] L. Visinelli. Natural Warm Inflation. *JCAP*, 1109:013, 2011.

[300] J. M. Wagstaff and K. Dimopoulos. Particle Production of Vector Fields: Scale Invariance is Attractive. *Phys.Rev.*, D83:023523, 2011.

[301] A. Walker. On Milne's Theory of World-Structure. *Proc.Lond.Math.Soc.*, s2-42 (1):90–127, 1937.

[302] D. Wands, K. A. Malik, D. H. Lyth, and A. R. Liddle. A New approach to the evolution of cosmological perturbations on large scales. *Phys.Rev.*, D62:043527, 2000.

[303] M.-a. Watanabe, S. Kanno, and J. Soda. Inflationary Universe with Anisotropic Hair. *Phys.Rev.Lett.*, 102:191302, 2009.

[304] S. Weinberg. The quantum theory of fields. Vol. 2: Modern applications. 1996.

[305] S. Weinberg. Quantum contributions to cosmological correlations. *Phys.Rev.*, D72:043514, 2005.

[306] S. Weinberg. *Cosmology.* 2008.

[307] L. M. Widrow. Origin of galactic and extragalactic magnetic fields. *Rev.Mod.Phys.*, 74:775–823, 2002.

[308] L. M. Widrow, D. Ryu, D. R. Schleicher, K. Subramanian, C. G. Tsagas, et al. The First Magnetic Fields. *Space Sci.Rev.*, 166:37–70, 2012.

[309] B. Willke et al. The GEO-HF project. *Class.Quant.Grav.*, 23:S207–S214, 2006.

[310] E. Witten. Cosmic Separation of Phases. *Phys.Rev.*, D30:272–285, 1984.

[311] C. Wu, V. Mandic, and T. Regimbau. Accessibility of the Gravitational-Wave Background due to Binary Coalescences to Second and Third Generation Gravitational-Wave Detectors. *Phys.Rev.*, D85:104024, 2012.

[312] K. K. Wu, O. Lahav, and M. J. Rees. The large-scale smoothness of the Universe. *Nature*, 397:225–230, 1999.

[313] J. Yadav, S. Bharadwaj, B. Pandey, and T. Seshadri. Testing homogeneity on large scales in the Sloan Digital Sky Survey Data Release One. *Mon.Not.Roy.Astron.Soc.*, 364:601–606, 2005.

[314] S. Yokoyama and J. Soda. Primordial statistical anisotropy generated at the end of inflation. *JCAP*, 0808:005, 2008.

[315] Y. Zeldovich, I. Y. Kobzarev, and L. Okun. Cosmological Consequences of the Spontaneous Breakdown of Discrete Symmetry. *Zh.Eksp.Teor.Fiz.*, 67:3–11, 1974.

Appendices

A Cosmological Perturbation Theory

This Appendix is devoted to some technical details on second-order cosmological perturbation theory in the spatially flat gauge, deriving the full equations for the inflaton perturbation $\delta\varphi$. We first derive the master equation for $\delta\varphi$, including all the terms that are relevant for the models we study in this thesis. Then, in the following Subsections we discuss each model separately: A.1 for axion inflation (Part IV), A.2 for the dilaton-like coupling model (Part V), A.3 for the model of a hidden-sector rolling pseudo scalar (Model II of Part VI), and A.4 for the model of burst vector production due to variation of its mass (Model I of Part VI).

For this purpose, we consider the action with all the relevant ingredients,

$$S = \int d^4x \left[\frac{M_p^2}{2} R - \frac{1}{2} \partial_\mu \varphi \, \partial^\mu \varphi - V(\varphi) - \frac{I^2(\varphi)}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} m^2 A_\mu A^\mu - \frac{\alpha}{4f} \varphi F_{\mu\nu} \tilde{F}^{\mu\nu} \right] \tag{A.1}$$

where m is the vector mass of which time variation is controlled by some field other than the inflaton.[59] We perform computations in the spatially flat gauge $\delta g_{ij,\mathrm{scalar}} = 0$ (the subscript "scalar" denotes the scalar sector of the metric perturbations). We expand the perturbations up to second order in perturbation theory. The fields φ and A_μ (in the Coulomb gauge) are expanded as

$$\varphi(t, \vec{x}) = \varphi^{(0)}(t) + \delta_1 \varphi(t, \vec{x}) + \delta_2 \varphi(t, \vec{x}) \,, \tag{A.2}$$

$$A_\mu(t, \vec{x}) = (0, A_i(t, \vec{x})) \,. \tag{A.3}$$

In fact it suffices to include only the first order for A_μ, and we do not need to evaluate it at second order. In all of our considerations in this thesis, the gauge field has no zero-order part. Consequently, the metric and inflaton perturbations do not mix with the gauge field modes at linear order; this is because the gauge field enters already quadratically in the action for the perturbations. (This is the same reason that in the main text allowed us to compute A_μ disregarding inflaton and metric perturbations.) We note that the gauge field can still affect the first order metric/inflaton perturbations through its backreaction on the background evolution. However, this can be disregarded as long as the backreaction bounds are satisfied.

The curvature perturbations ζ in this gauge is

$$\zeta = -\frac{H}{\dot{\varphi}^{(0)}} \left(\delta_1 \varphi + \delta_2 \varphi \right) \,. \tag{A.4}$$

[59] The vector mass term $m^2 A^2$ is relevant for Model I of Part VI, and there it is related to the vev of a hidden-sector scalar field by $m = e\psi^{(0)}$, where e is the "charge" of the scalar field.

We are interested in a master equation for ζ. We start from the inflaton equation of motion. Varying the action (A.1) with respect to φ, we obtain

$$-\frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}\,g^{\mu\nu}\,\partial_\nu\,\varphi\right)+V_\varphi=-\frac{II_\varphi}{2}F_{\mu\nu}F^{\mu\nu}-\frac{\alpha}{4f}F_{\mu\nu}\tilde{F}^{\mu\nu}\,. \tag{A.5}$$

Formally, not only the φ and A_μ perturbations, but also the metric perturbations enter in this equation. However, one can use the Einstein constraint equations to express the modes δg_{00} and δg_{0i} in terms of the other perturbations,

$$\delta g_{00}=\delta g_{00}\left[\delta_{1,2}\varphi,\,A_\mu,\,h_{ij}\right]\;,\quad \delta g_{0i}=\delta g_{0i}\left[\delta_{1,2}\varphi,\,A_\mu,\,h_{ij}\right]\;\;, \tag{A.6}$$

where h_{ij} are the tensor modes of the metric perturbations. Then we insert these expressions back into (A.5) to close the system.[60] The explicit expressions for (A.6) can be most directly obtained from the Einstein equations.[61] This procedure has already been detailed in [225]; here we generalize those results to include also the presence of the gauge field. The vector field modes enter in these equations through their contribution to the energy momentum tensor

$$T_{\mu\nu,\mathrm{gauge}}=I^2F_{\mu\alpha}\,F_\nu{}^{\alpha}+m^2\,A_\mu A_\nu+g_{\mu\nu}\left(-\frac{I^2}{4}F_{\mu\nu}F^{\mu\nu}-\frac{m^2}{2}A_\mu A^\mu\right)\;. \tag{A.7}$$

Note that the axion-gauge coupling $\varphi F\tilde{F}$ does not couple to gravity and has no contribution to the energy-momentum tensor.

As the energy-momentum tensor (A.7) is quadratic in the gauge field modes, the first-order metric and inflaton perturbations are not affected by the gauge field. Therefore the master equation at first order reproduces the standard first-order expression

$$\left[\partial_\tau^2+2\mathcal{H}\partial_\tau-\nabla^2+\left(a^2V_{\varphi\varphi}-3\,\frac{\varphi^{(0)\prime\,2}}{M_p^2}\right)\right]\delta_1\varphi=0 \tag{A.8}$$

where we remind $\nabla^2\equiv\partial_i\partial_i$ and $\mathcal{H}\equiv\frac{a'}{a}$. At the linear-order perturbation $\delta_1\varphi$, eq. (A.8) applies to all the models we consider in this thesis.

[60] Equivalently, we can eliminate the δg_{00} and δg_{0i} modes from the action using the so-called energy and momentum constraints in the ADM formalism. Clearly, this is equivalent to using (A.6) at the level of the equations of motion. The master equation, in absence of gauge fields, was first derived in [225].

[61] We note that only the tensor modes h_{ij} enter in the spatial perturbations δg_{ij}, since $\delta g_{ij,\mathrm{scalar}}=0$ in our gauge, and since we can disregard vector metric perturbations as in the standard case.

At next order in perturbation theory we instead find

$$\left[\partial_\tau^2 + 2\mathcal{H}\partial_\tau - \nabla^2 + \left(a^2 V_{\varphi\varphi} - 3\,\frac{\varphi^{(0)\prime\,2}}{M_p^2}\right)\right]\delta_2\varphi = a^2\,\frac{I_\varphi}{I}\left(\vec{E}^2 - \vec{B}^2\right) + a^2\,\frac{\alpha}{f}\vec{E}\cdot\vec{B}$$
$$-\frac{a^2\varphi^{(0)\prime}}{2M_p^2\mathcal{H}}\left\{\frac{\vec{E}^2+\vec{B}^2}{2} + \frac{1}{a^4}\partial^{-2}\partial_\tau\left[a^4\vec{\nabla}\cdot\left(\vec{E}\times\vec{B}\right)\right] + \frac{m^2}{2a^2}\,A_i A_i\right\} + \cdots \quad \text{(A.9)}$$

Dots denote terms which are proportional to squares of the first-order inflaton and tensor metric perturbations, namely they are of $\mathcal{O}\left(\delta_1\varphi^2\right)$ and of $\mathcal{O}\left(h_{1,ij}^2\right)$; these terms, which are slow-roll suppressed, are explicitly given in [225] and are known to yield an undetectable contribution to $f_{\rm NL}$ [270], and thus we disregard them. Here, the "electric" and "magnetic" fields are defined in analogy to the electromagnetic expressions, namely

$$E_i = -\frac{\langle I\rangle}{a^2}A_i' \;\;,\;\; B_i = \frac{\langle I\rangle}{a^2}\epsilon_{ijk}\partial_j A_k \qquad \text{(A.10)}$$

Note that the terms in the second line of (A.9) arise from the consistent inclusion of the metric perturbations, while the terms in the right-hand side of the first line of (A.9) are present even without them. Also note that the terms $a^2 V_{\varphi\varphi}$ and $3\frac{\varphi^{(0)\prime\,2}}{M_p^2}$ on the left-hand sides of (A.8) and (A.9) are proportional to the slow-roll parameters. Since we are not interested in slow-roll corrections to our leading results, we disregard these terms in the main text.

Eq. (A.9) contains all the possible source terms for the models we study in this thesis; not all of them are relevant simultaneously for any one of the models. In the following Subsections, we deduce (A.9) to the form that is applicable to each model we consider.

A.1 Massless Gauge Field with Direct Axion-Gauge Coupling

Here we consider the case where the gauge field is massless and is coupled to the pseudo-scalar inflaton, which represents the model of axion inflation studied in Part IV. Thus, we set $m = 0$ and $I = 1$ in (A.9).

Moreover, the Planck-suppressed source terms, coming from the consistent inclusion of the metric perturbations (namely the second line of (A.9)), are completely negligible as compared to the pseudo-scalar coupling in the first line of (A.9). To see this, note that the strength of the axion-gauge direct interaction is controlled by the axion decay constant which is

$$\frac{\alpha}{f} \gtrsim 10^2\,\frac{1}{M_p} \qquad \text{(A.11)}$$

whenever the non-Gaussianity is observationally interesting (see Section IV 3). The analogous coupling strength associated with the Planck-suppressed interaction in the second line of (A.9) is instead

$$\frac{\varphi^{(0)\prime}}{\mathcal{H}M_p^2} \sim \frac{\sqrt{\epsilon}}{M_p} \sim 10^{-1}\,\frac{1}{M_p} \tag{A.12}$$

where in the last approximation we have assumed the slow-roll parameter $\epsilon \sim 10^{-2}$, which is true in most axion inflation models. This analysis indicates that the inclusion of metric perturbations has no significant impact, and can therefore be neglected, in the interesting regime where $f_{\rm NL} \gtrsim \mathcal{O}(1)$.[62]

We thus arrive at the master equation for $\delta_2\varphi$ for the axion inflation model:

$$\left[\partial_\tau^2 + 2\mathcal{H}\partial_\tau - \nabla^2 + \left(a^2 V_{\varphi\varphi} - 3\,\frac{\varphi^{(0)\prime\,2}}{M_p^2}\right)\right]\delta_2\varphi \simeq a^2\frac{\alpha}{f}\vec{E}\cdot\vec{B} \tag{A.13}$$

which is identified, up to slow-roll suppressed terms (which we are indifferent to), with eq. (95) in the main text.

A.2 Massless Gauge Field with Direct Dilaton-like Coupling

In this Subsection, we consider the massless gauge field with a dilaton-like $I^2(\varphi)F^2$ coupling, which represents the model of Part V. Hence we set $m = 0$ and $\alpha/f = 0$ in (A.9). As in A.1, the Planck suppressed source terms in (A.9) are negligible compared to the direct coupling term. To see this, let us recall our functional forms (193), $V(\varphi) \propto \varphi^r$ and $I \propto \exp\left(-\frac{\varphi^2}{rM_p^2}\right)$, and $\epsilon \equiv \frac{M_p^2}{2}\left(\frac{V_\varphi}{V}\right)^2$. We then see

$$\frac{I_\varphi}{I} \sim \frac{1}{\sqrt{\epsilon}\,M_p}\ , \qquad \frac{\varphi^{(0)\prime}}{M_p^2\mathcal{H}} \sim \frac{\sqrt{\epsilon}}{M_p}\ . \tag{A.14}$$

Due to the slow-roll condition $\epsilon \ll 1$, we conclude that the Planck suppressed source terms are completely negligible compared to the direct coupling term. Therefore we obtain the master equation for $\delta_2\varphi$

$$\left[\partial_\tau^2 + 2\mathcal{H}\partial_\tau - \nabla^2 + \left(a^2 V_{\varphi\varphi} - 3\,\frac{\varphi^{(0)\prime\,2}}{M_p^2}\right)\right]\delta_2\varphi \simeq a^2\frac{I_\varphi}{I}\left(\vec{E}^2 - \vec{B}^2\right) \tag{A.15}$$

which is eq. (214), reported in the main text.

[62] This intuitive idea has been exploited on numerous occasions in the literature and was recently conjectured as a general decoupling principle in [202].

A.3 Massless Gauge Field with No Direct Coupling

In this Subsection, we now consider Model II of Part VI, in which a massless gauge field has no direct coupling to the inflaton. We thus set $I = 1$, $\alpha/f = 0$, and $m = 0$ in (A.9). We see that all the source terms arise due to the consistent inclusion of the metric perturbations and are obtained by integrating out the non-dynamical modes. The master equation for $\delta_2\varphi$ in this case reads

$$\left[\partial_\tau^2 + 2\mathcal{H}\partial_\tau - \nabla^2 + \left(a^2 V_{\varphi\varphi} - 3\,\frac{\varphi^{(0)\prime\,2}}{M_p^2}\right)\right]\delta_2\varphi \simeq -\frac{a^2\varphi^{(0)\prime}}{2M_p^2\mathcal{H}}\left\{\frac{\vec{E}^2+\vec{B}^2}{2} + \frac{1}{a^4}\partial^{-2}\partial_\tau\left[a^4\vec{\nabla}\cdot\left(\vec{E}\times\vec{B}\right)\right]\right\} . \quad \text{(A.16)}$$

Now we want to derive the expression (342) for the scalar source in Model II of Part VI. In eq. (340) we separated the source in the two parts $J_1 + J_2$, where J_1 corresponds to the $\vec{E}^2 + \vec{B}^2$ term in (A.16) and J_2 to the $\vec{E}\times\vec{B}$ term. In Fourier space we have the relatively simple expression

$$J_1\left(\tau,\vec{k}\right) = -\frac{a^3\varphi^{(0)\prime}}{4M_p^2\mathcal{H}}\int\frac{d^3p}{(2\pi)^{3/2}}\left[\,\tilde{E}_i\left(\tau,\vec{p}\right)\tilde{E}_i\left(\tau,\vec{k}-\vec{p}\right) + \tilde{B}_i\left(\tau,\vec{p}\right)\tilde{B}_i\left(\tau,\vec{k}-\vec{p}\right)\right] , \quad \text{(A.17)}$$

for the first part. For the second part, we have instead

$$J_2\left(\tau,\vec{k}\right) = \frac{\varphi^{(0)\prime}}{2M_p^2\mathcal{H}a}\frac{\partial_\tau}{k^2}\int\frac{d^3p}{(2\pi)^{3/2}}\left[p_i k_i\tilde{A}_j\left(\tau,\vec{p}\right)\tilde{A}_j'\left(\tau,\vec{k}-\vec{p}\right) - k_i k_j\tilde{A}_i\left(\tau,\vec{p}\right)\,\tilde{A}_j'\left(\tau,\vec{k}-\vec{p}\right)\right] . \quad \text{(A.18)}$$

Some manipulations are instead necessary to put the non-local source term J_2 in a form that is amenable to computations and is manifestly infra-red finite. We begin by noting that $\tilde{A}_i$ effectively only contains the "+" helicity state in this model. This allows us to use the identity

$$k_i\,k_j\,\epsilon_i^{(+)}\left(\vec{p}\right)\epsilon_j^{(+)}\left(\vec{k}-\vec{p}\right) = \left[p|\vec{k}-\vec{p}| + (\vec{k}-\vec{p})\cdot\vec{p}\right]\,\epsilon_i^{(+)}\left(\vec{p}\right)\,\epsilon_i^{(+)}\left(\vec{k}-\vec{p}\right) , \quad \text{(A.19)}$$

to simplify the tensor structure in the second term of (A.18) and re-write J_2 as

$$J_2\left(\tau,\vec{k}\right) = \frac{\varphi'^{(0)}}{2M_p^2\mathcal{H}a}\frac{1}{k^2}\int\frac{d^3p}{(2\pi)^{3/2}}\;p\left[p - |\vec{k}-\vec{p}|\right]\partial_\tau\left[\,\tilde{A}_i\left(\tau,\vec{p}\right)\tilde{A}_i'\left(\tau,\vec{k}-\vec{p}\right)\,\right] . \quad \text{(A.20)}$$

Next we perform the time derivative and use the equation of motion (329) to eliminate $\tilde{A}''$

which gives the result

$$J_2\left(\tau,\vec{k}\right)=\frac{\varphi'^{(0)}a^3}{2M_p^2\mathcal{H}}\frac{1}{k^2}\int\frac{d^3p}{(2\pi)^{3/2}}\left(|\vec{k}-\vec{p}|-p\right)\Bigg[-p\tilde{E}_i\left(\tau,\vec{p}\right)\tilde{E}_i\left(\tau,\vec{k}-\vec{p}\right)$$
$$+\left(|\vec{k}-\vec{p}|+\frac{2\xi}{\tau}\right)\tilde{B}_i\left(\tau,\vec{p}\right)\tilde{B}_i\left(\tau,\vec{k}-\vec{p}\right)\Bigg]. \tag{A.21}$$

We rewrite (A.21) changing the integration variable $\vec{p}\to\vec{k}-\vec{p}$. We add the resulting expression to (A.21), and divide by two. We obtain:

$$J_2\left(\tau,\vec{k}\right)=\frac{\varphi'^{(0)}a^3}{4M_p^2\mathcal{H}}\int\frac{d^3p}{(2\pi)^{3/2}}\frac{(p-|\vec{k}-\vec{p}|)^2}{k^2}\left[\tilde{E}_i\left(\tau,\vec{p}\right)\tilde{E}_i\left(\tau,\vec{k}-\vec{p}\right)+\tilde{B}_i\left(\tau,\vec{p}\right)\tilde{B}_i\left(\tau,\vec{k}-\vec{p}\right)\right]. \tag{A.22}$$

We note that this expression is manifestly finite in the limit $k^2\to 0$, proving that the non-local source term in (340) does not lead to any spurious effects in the infra-red.

Adding (A.17) and (A.22) leads to the expression (342) reported in the main text.

A.4 Massive Gauge Field with No Direct Coupling

In this Subsection, we again consider no direct coupling between the inflaton and gauge field, but in this case the gauge field is massive, i.e. $m\neq 0$. Setting $I=1$ and $\alpha/f=0$ leads (A.9) to the equation for $\delta_2\varphi$ in this case,

$$\left[\partial_\tau^2+2\mathcal{H}\partial_\tau-\nabla^2+\left(a^2V_{\varphi\varphi}-3\,\frac{\varphi^{(0)\prime\,2}}{M_p^2}\right)\right]\delta_2\varphi$$
$$\simeq-\frac{a^2\varphi^{(0)\prime}}{2M_p^2\mathcal{H}}\left\{\frac{\vec{E}^2+\vec{B}^2}{2}+\frac{1}{a^4}\partial^{-2}\partial_\tau\left[a^4\vec{\nabla}\cdot\left(\vec{E}\times\vec{B}\right)\right]+\frac{m^2}{2a^2}\,A_iA_i\right\} \tag{A.23}$$

As in A.3, all the source terms are only gravitational and arise through integrating out the non-dynamical modes of the metric perturbations (with the nonzero mass term this time).

Now we derive the approximated expression (288) for the scalar source in Model I of

Part VI. From the three terms at the right-hand side of (A.23) we obtain, respectively,

$$J\left[A_\mu^2\right] = J_1\left(\tau,\vec{k}\right) + J_2\left(\tau,\vec{k}\right) + J_3\left(\tau,\vec{k}\right)$$

$$J_1\left(\tau,\vec{k}\right) = -\frac{\varphi^{(0)\prime}}{4M_p^2\mathcal{H}a}\int\frac{d^3p}{(2\pi)^{3/2}}\tilde{A}_i'\left(\tau,\vec{p}\right)\tilde{A}_i'\left(\tau,\vec{k}-\vec{p}\right) \tag{A.24}$$

$$J_2\left(\tau,\vec{k}\right) = \frac{\varphi^{(0)\prime}}{2M_p^2\mathcal{H}a}\frac{1}{k^2}\int\frac{d^3p}{(2\pi)^{3/2}}\partial_\tau\Big[p_i\,k_i\,\tilde{A}_j\left(\tau,\vec{p}\right)\tilde{A}_j'\left(\tau,\vec{k}-\vec{p}\right) - k_i\,k_j\,\tilde{A}_i\left(\tau,\vec{p}\right)\tilde{A}_j'\left(\tau,\vec{k}-\vec{p}\right)\Big] \tag{A.25}$$

$$J_3\left(\tau,\vec{k}\right) = -\frac{a\,m^2\varphi^{(0)\prime}}{4M_p^2\mathcal{H}}\int\frac{d^3p}{(2\pi)^{3/2}}\tilde{A}_i\left(\tau,\vec{p}\right)\tilde{A}_i\left(\tau,\vec{k}-\vec{p}\right). \tag{A.26}$$

In the part J_1 we have actually disregarded the "magnetic" contribution with respect to the "electric" one.The relative contribution between the two terms in the integrand is $\frac{B^2}{E^2} \sim \frac{\tilde{A}'^2}{\tilde{A}^2} \sim \frac{k^2,\,k\,p,\,p^2}{a^2m^2}$. The quantity k is the momentum of the cosmological perturbations that we are computing; we show in the main text that the signal from particle production is maximal at $k \sim H$ (we recall that the scale factor is normalized to one at the moment of particle production). We also show in the main text that the source integrand is peaked at $p \sim \sqrt{\dot{m}_*}$. We therefore have $k \ll p$ due to (272). Finally $M \equiv am \geq m \gg \sqrt{\dot{m}_*}$ after the particle production has taken place. As a consequence $p \ll M$ and the "magnetic" contribution can indeed be disregarded.

We perform the time derivative in J_2, and eliminate the second derivative through the equation of motion $\tilde{A}_i'' \simeq -M^2\,\tilde{A}_i$. We obtain

$$J_2 \simeq \frac{\dot{\varphi}^{(0)}}{2M_p^2Ha}\frac{1}{k^2}\int\frac{d^3p}{(2\pi)^{3/2}}\left(\vec{k}\cdot\vec{p}\,\delta_{ij} - k_i\,k_j\right)\left[\tilde{A}_i'\left(\vec{p}\right)\tilde{A}_j'\left(\vec{k}-\vec{p}\right) - M^2\tilde{A}_i\left(\vec{p}\right)\tilde{A}_j\left(\vec{k}-\vec{p}\right)\right]. \tag{A.27}$$

We rewrite (A.27) changing the integration variable $\vec{p} \rightarrow \vec{k}-\vec{p}$. We add the resulting expression to (A.27), and divide by two. We obtain

$$J_2 \simeq \frac{\dot{\varphi}^{(0)}}{2M_p^2Ha}\left(\frac{1}{2}\delta_{ij} - \hat{k}_i\hat{k}_j\right)\int\frac{d^3p}{(2\pi)^{3/2}}\left[\tilde{A}_i'\left(\vec{p}\right)\tilde{A}_j'\left(\vec{k}-\vec{p}\right) - M^2\tilde{A}_i\left(\vec{p}\right)\tilde{A}_j\left(\vec{k}-\vec{p}\right)\right] \tag{A.28}$$

It is worth noting that this expression explicitly shows that J_2 does not diverge in the $k \rightarrow 0$ limit. This is not immediately obvious from the original expression, since J_2 originates from the term with an inverse Laplacian in (283).

Adding this expression for J_2 to the expression for J_1 and J_3 given above, one readily obtains the result (288), given in the main text.

B Gauge Field Mode Functions in Axion Inflation

In this Appendix we discuss the solutions of equation (87), describing the tachyonic growth of gauge field fluctuations in the background of the slowly rolling inflaton (hence, ξ can be taken constant at leading order in the slow roll parameters). We require that the gauge field is initially in the adiabatic vacuum: $A_\pm(\tau,k) = e^{-ik\tau}/\sqrt{2k}$ for $k\tau \to -\infty$. The mathematical properties used here can be found in Chapter 14 of [7]. The solution of (87) which satisfies this condition may be expressed in terms of Coulomb functions:

$$A_+(\tau,k) = \frac{1}{\sqrt{2k}}\left[G_0(\xi,-k\tau) + iF_0(\xi,-k\tau)\right] \tag{B.1}$$

The production of gauge fluctuations is only interesting in the region of phase space $-k\tau \ll 2\xi$ and when $e^{\pi\xi} \gg 1$ (see the discussion in Subsection IV 1.2). In this regime equation (B.1) may be very well approximated in terms of the modified Bessel function of the second kind, $K_\nu(z)$, as

$$A_+(\tau,k) \cong \sqrt{\frac{-2\tau}{\pi}}\, e^{\pi\xi}\, K_1\left[2\sqrt{-2\xi k\tau}\right] \tag{B.2}$$

The growth of the modes (B.2) saturates deep in the IR: for $-k\tau \to 0$ we have $A_+ \to e^{\pi\xi}/(2\sqrt{\pi k\xi})$ so that the physical electric and magnetic field vectors (85) decay sufficiently far outside the horizon. An inspection of the solutions shows that the interesting physical effects (for instance, the production of ζ^2 and ζ^3 correlators) take place in the region $(8\xi)^{-1} \lesssim -k\tau \lesssim 2\xi$ of phase space. In this regime we can take the large argument asymptotics of the Bessel function in (B.2) to obtain a very simple representation of the modes:

$$\begin{aligned} A_+(\tau,k) &\cong \frac{1}{\sqrt{2k}}\left(\frac{-k\tau}{2\xi}\right)^{1/4} e^{\pi\xi - 2\sqrt{-2\xi k\tau}} \\ A'_+(\tau,k) &\cong \sqrt{\frac{2k\xi}{-\tau}} A_+(\tau,k) \end{aligned} \tag{B.3}$$

Throughout the majority of this Part we employ the representation (B.3) of the modes, for brevity of exposition. However, we have verified that none of our results changes significantly if we use the more accurate expression (B.2). Formally, the only effect of using (B.2) rather than (B.3) is that, for any two (rescaled) momenta $\vec{q}_1$ and $\vec{q}_2$, the quantity

$$\mathcal{I}\left[c\left(|\vec{q}_1|^{1/2} + |\vec{q}_2|^{1/2}\right)\right] \simeq \int_0^\infty dx\,(\sin x - x\cos x)\,\mathrm{e}^{-c\left(|\vec{q}_1|^{1/2}+|\vec{q}_2|^{1/2}\right)\sqrt{x}} \tag{B.4}$$

entering in eqs. (120) and (132) gets replaced by

$$\mathcal{I}_B \equiv \frac{2c}{\pi}\frac{|\vec{q}_1|^{1/4}\,|\vec{q}_2|^{1/4}}{|\vec{q}_1|^{1/2}+|\vec{q}_2|^{1/2}}\int_0^\infty dx\,\sqrt{x}\,[\sin x - x\cos x] \times \left[|\vec{q}_1|^{1/2}\,K_1\left(c\,|\vec{q}_1\,x|^{1/2}\right)\,K_0\left(c\,|\vec{q}_2\,x|^{1/2}\right)+|\vec{q}_2|^{1/2}\,K_1\left(c\,|\vec{q}_2,x|^{1/2}\right)\,K_0\left(c\,|\vec{q}_1\,x|^{1/2}\right)\right] \tag{B.5}$$

We verified that the results obtained with this replacement are in excellent agreement with those presented in the main text. Specifically, we obtained identical values to those presented in Table 1 for the non-Gaussianity shape; for the non-Gaussianity parameter $f_{\rm NL}^{\rm equil}$ we obtained values consistent within a few percent with those shown in Figure 8. This explicitly confirms that the approximated expressions (B.3) are adequate for our analysis.

C Exact Definition of ζ with a Scalar Inflaton

In this Appendix we compute the exact expression for the gauge invariant curvature perturbation in the model studied in Part V. We denote the exact expression by $\zeta_{\rm exact}$. We instead denote by ζ the combination $\zeta = -\frac{H}{\dot{\varphi}^{(0)}}\,\delta\varphi$, as we do everywhere in the main text of this thesis. In this Appendix we show that the difference between ζ and $\zeta_{\rm exact}$ is completely negligible for all our purposes. We decompose

$$\begin{aligned}\zeta_{\rm exact} &= \zeta_{\rm exact,1}+\zeta_{\rm exact,2}\\ &= \left(\zeta_{\rm exact,1}+\zeta_{\rm exact,2}\big|_{\text{sourced by }\delta_1\varphi\text{ and }\delta_1 g_{\mu\nu}}\right)+\zeta_{\rm exact,2}\big|_{\text{sourced by }\delta_1 A_\mu}\end{aligned} \tag{C.1}$$

where the number in the suffix denotes the order in perturbation theory, and where, in total analogy with what we did in (216), we have separated the part of $\zeta_{\rm exact,2}$ sourced by the first order vacuum fluctuations of the inflaton and the metric from the part sourced by the vector field ($\delta_1 A_\mu$ coincides with the quantity denoted by A_μ in the main text, since the gauge field has no vacuum expectation value). The two terms in the round parenthesis are uncorrelated with the last term. These two terms coincide with those computed without gauge field (more precisely, the gauge field affects them only due to the backreaction on the background dynamics, which we impose to be subdominant). As in the standard case, for these terms we have at super-horizon scales

$$\zeta_{\rm exact,1}+\zeta_{\rm exact,2}\big|_{\text{sourced by }\delta_1\varphi\text{ and }\delta_1 g_{\mu\nu}} = -\frac{H}{\dot{\varphi}^{(0)}}\delta_1\varphi+\mathcal{O}\left[(\delta_1\varphi)^2\,,\,(\delta_1 g_{\mu\nu})^2\,,\,(\delta_1\varphi\times\delta_1 g_{\mu\nu})\right] \tag{C.2}$$

The precise expression for the second term is given in eq. (7.71) of [226]. Since the standard single scalar field inflation results apply for these terms, we know that the quadratic term in this expression gives a negligible contribution to the power spectrum and leads to unobservable non-Gaussianity. Therefore, we disregard it in our computations.

The formal expression for $\zeta_{\rm exact,2}\big|_{\text{sourced by }\delta_1 A_\mu}$ can be immediately obtained from eq. (7.71) of [226]. This expression is written before fixing any gauge, and reads

$$\zeta_{\rm exact,2} = -\psi_2 - \frac{\mathcal{H}}{\rho^{(0)\prime}}\delta_2\rho + \mathcal{O}\left[\left(\delta_1\varphi\right)^2 ,\, \left(\delta_1 g_{\mu\nu}\right)^2 ,\, \left(\delta_1\varphi\times\delta_1 g_{\mu\nu}\right)\right] \tag{C.3}$$

where ψ_2 is a second order perturbation entering in the spatial part of the metric, $\delta_2 g_{ij,\rm scalar} = a^2\left[-2\psi_2\delta_{ij} + 2E_{2,ij}\right]$. In the (spatially) flat gauge that we are using ($\psi_2 = E_2 = 0$) this first term is absent. The third term in (C.3) is the term $\zeta_{\rm exact,2}|_{\text{sourced by } \delta_1\varphi \text{ and } \delta_1 g_{\mu\nu}}$ that we are disregarding. Therefore

$$\zeta_{\rm exact,2}|_{\text{sourced by } \delta_1 A_\mu} = -\frac{\mathcal{H}}{\rho^{(0)\prime}}\delta_2\rho|_{\text{sourced by } \delta_1 A_\mu} \tag{C.4}$$

We stress that this is an exact relation.

From now on, when we write a second order quantity we only mean the part sourced by $\delta_1 A_\mu$, without indicating it explicitly. The quantity $\delta_2\rho$ in (C.4) is the second order perturbation of $-T^0_{\ 0}$. By evaluating it, we have

$$\zeta_{\rm exact,2} = -\frac{\mathcal{H}}{\rho^{(0)\prime}}\left[\frac{1}{a^2}\,\varphi^{(0)\prime}\,\delta_2\varphi' - \frac{1}{a^2}\varphi^{(0)\prime\,2}\,\phi_2 + V_\varphi\,\delta_2\varphi + \frac{1}{2}\left(\vec{E}^2 + \vec{B}^2\right)\right] \tag{C.5}$$

where $\phi_2 \equiv -\left(2a^2\right)^{-1}\delta_2 g_{00}$, while E and B are the electric and the magnetic field modes, respectively. Eliminating the non-dynamical mode ϕ_2 from this expression, we find

$$\zeta_{\rm exact,2} = \zeta_2 - \frac{\zeta_2'}{3\mathcal{H}} - \frac{S^{(3)}}{3\,\mathcal{H}} + \frac{\vec{E}^2 + \vec{B}^2}{6\left(\rho^{(0)} + p^{(0)}\right)} \tag{C.6}$$

where $\rho^{(0)}$ and $p^{(0)}$ are the background energy density and pressure, respectively, and

$$S^{(3)} = \frac{I^2}{2a^2 M_p^2}\nabla^{-2}\left[\partial_i\delta_1 A_j'(\partial_i\delta_1 A_j - \partial_j\delta_1 A_i) + \delta_1 A_i'\nabla^2\delta_1 A_i\right]\ . \tag{C.7}$$

Also this relation is exact.

We now show that the last three terms on the right hand side of this expression can be completely disregarded with compared to the first one. We do so by showing that (i) they are already subdominant during inflation, and (ii) they decrease relatively to the first term by many orders of magnitude during reheating. Already the statement (i) would be sufficient to disregard them.

To verify the statement (i), we compare the r.m.s. of the various terms during inflation, when the modes are on super horizon scales. We have

$$\sqrt{\langle\zeta_2^2\rangle} \simeq 4\,\mathcal{P}\,\ln^2\left(\frac{a\left(\tau\right)}{a_{\rm in}}\right) \tag{C.8}$$

where we remind that $\mathcal{P}$, defined in eq. (205), is the standard result for the first order power spectrum. We then have

$$\frac{\zeta_2'}{3\mathcal{H}\zeta_2}|_{\text{r.m.s.}} \simeq \frac{2}{3\,\ln\frac{a(\tau)}{a_{\text{in}}}} \qquad \text{during inflation} \tag{C.9}$$

This ratio evaluates to $2/\left(3N_{\text{tot}}\right) < 0.01$ at the end of inflation (we recall that N_{tot} denotes the total number of e-folds of inflation).

For the third term in (C.6), we see that the quantity $S^{(3)}$ defined in (C.7) has three terms that are parametrically of the same order. Therefore we can estimate

$$S^{(3)}|_{\text{r.m.s.}} \lesssim \frac{3\,I^2}{2a^2M_p^2}\,\left\langle \delta A_i'\,\delta A_i\right\rangle \tag{C.10}$$

where the factor of 3 accounts for the possibility that the contributions from the three terms add up in magnitude, although there may actually be cancellations (in this way we obtain a safe upper bound for this third term). Inserting (48) in this expression, and using (197) in the super horizon regime (we recall that $n = 2$), we obtain

$$\frac{S^{(3)}}{3\mathcal{H}\zeta_2}|_{\text{r.m.s.}} \lesssim \frac{3\,\epsilon}{2\,\ln\frac{a(\tau)}{a_{\text{in}}}} \qquad \text{during inflation} \tag{C.11}$$

So we see that the contribution of the third term in (C.6) is suppressed by an ϵ factor with respect to the already negligible contribution from the second term.

For the last term in (C.6), using eq. (200), we have instead

$$\frac{\vec{E}^2+\vec{B}^2}{6\left(\rho^{(0)}+p^{(0)}\right)\zeta_2}|_{\text{r.m.s.}} \simeq \frac{3}{4\,\ln\frac{a(\tau)}{a_{\text{in}}}} \qquad \text{during inflation} \tag{C.12}$$

which again evaluates to < 0.01 at the end of inflation.

We see that the corrections to $\zeta_{\text{exact},2} - \zeta_2$ can be disregarded already at the end of inflation (they are smaller than the accuracy with which we have evaluated ζ_2). Although this is not needed, we can actually verify that these corrections even decrease by several orders of magnitude during reheating.

For a massive inflaton potential, $|\varphi^{(0)}| \propto a^{-3/2}$ during reheating. Therefore, the energy density and pressure of the inflaton behave as those of non-relativistic matter. The energy density in the gauge field instead decreases as a^{-4} at the super-horizon scales of our interest. Therefore, the system rapidly approaches the single fluid regime, with a frozen $\zeta_{\text{exact},2} \simeq \zeta_2$. One can easily verify that the last two terms in (C.6) also rapidly decrease; specifically, they scale as $a^{-3/2}$ and a^{-1}, respectively. Assuming an instantaneous inflaton decay at $t = t_{\text{reh}}$, the ratio of the scale factor between the end of inflation and reheating is

$$\frac{a_{\text{end inf}}}{a_{\text{reh}}} \simeq 10^{-10} \left(\frac{T_{\text{rh}}}{10^9\,\text{GeV}}\right)^{4/3} \left(\frac{10^{15}\,\text{GeV}}{H_{\text{inf}}}\right)^{2/3} \tag{C.13}$$

where T_{rh} is the temperature of the bath formed by the inflaton decay products.

Therefore, we have explicitly verified that $\zeta = -\frac{H}{\dot{\varphi}^{(0)}}\,\delta\varphi$ is a perfectly good expression for the gauge invariant curvature perturbation in this model.

D The In-In Formalism

In this Appendix we verify that the Green's function method employed in the text is equivalent to the "in-in" formalism at leading order, using the model in Part V as an example. To compute correlation functions using the in-in method, we must first identify the interaction Hamiltonian. To this end, we employ the Arnowitt-Deser-Misner (ADM) form of the metric and integrate out the lapse function and shift vector. This procedure has been described in [268] and also [45], so we do not reproduce the details of the calculation here. The quadratic action for the inflaton perturbation is

$$S_2 = \frac{1}{2}\int d\tau d^3x a^2 \left[(\partial_\tau\delta\varphi)^2 - \partial_i\delta\varphi\partial_i\delta\varphi - \left(a^2 V_{\varphi\varphi} - 3\frac{\varphi^{(0)\prime\,2}}{M_p^2}\right)\delta\varphi^2\right]. \tag{D.1}$$

The cubic interaction terms in the Lagrangian are

$$\begin{aligned} S_3 \;=\;& \int d\tau d^3x a^4 \frac{I_\varphi}{I}\delta\varphi\left[\vec{E}^2 - \vec{B}^2\right] \\ &+ \int d\tau d^3x a^4 \frac{\varphi^{(0)\prime}}{\mathcal{H}M_p^2}\delta\varphi\left[-\frac{1}{4}(\vec{B}^2+\vec{E}^2) - \frac{1}{2a^4}\nabla^{-2}\partial_\tau\left(a^4\vec{\nabla}\cdot(\vec{E}\times\vec{B})\right)\right] \end{aligned} \tag{D.2}$$

Here we have suppressed terms of the form $(\delta\varphi)^3$, which are irrelevant for our calculation. Varying (D.1) and (D.2) and expanding $\delta\varphi = \delta_1\varphi + \delta_2\varphi$ reproduces exactly the master equation (A.9) with $\alpha/f = 0$ and $m = 0$.

For the choice of $I(\varphi)$ under consideration the interactions on the first line of (D.2) are controlled by the dimensionful coupling $\left|\frac{I_\varphi}{I}\right| \sim \sqrt{\frac{1}{\epsilon}}\frac{1}{M_p}$. In contrast, the interactions in the second line of (D.2) are controlled by a coupling $\left|\frac{\varphi^{(0)\prime}}{\mathcal{H}M_p^2}\right| \sim \frac{\sqrt{\epsilon}}{M_p}$. Therefore in the slow roll limit, $\epsilon \ll 1$, the interactions on the first line of (D.2) are the dominant ones.

At leading order in a slow roll expansion, the interaction Hamiltonian $H_I(t) = -\int d^3x a^3\mathcal{L}_3$ can be written as

$$H_I(t) = \frac{\dot{\varphi}^{(0)}}{H}\int d^3q\, J_{\vec{q}}(\tau)\,\zeta_{-\vec{q}}(\tau)\,, \tag{D.3}$$

where $\zeta = -\frac{H}{\dot{\varphi}^{(0)}}\delta\varphi$ is the curvature perturbation and the source $J_k(\tau)$ was defined in (223). The in-in formula is

$$\begin{aligned}\left\langle \zeta_{\vec{k}_1}\zeta_{\vec{k}_2}\cdots\zeta_{\vec{k}_n}(t)\right\rangle &= \sum_{N=0}^{\infty}(-i)^N\int^t dt_1\int^{t_1}dt_2\cdots\int^{t_{N-1}}dt_N \\ &\quad\times\left\langle\left[\left[\left[\zeta_{\vec{k}_1}\zeta_{\vec{k}_2}\cdots\zeta_{\vec{k}_n}(t),H_I(t_1)\right],H_I(t_2)\right]\cdots,H_I(t_N)\right]\right\rangle . \end{aligned}\tag{D.4}$$

A key simplification arises from noting that the mode functions of the produced gauge fluctuations are real-valued, up to an irrelevant constant phase. This implies that the produced gauge field fluctuations are commuting variables, to a very good approximation. We have

$$[\partial_\tau A_i, A_j]\approx 0\,,\tag{D.5}$$

where it is understood that only superhorizon modes are relevant. Consequently, the source terms $J_q(t)$ in (D.4) are mutually commuting and they may be pulled out of the nested commutator. The remaining commutators are easily evaluated using the formula

$$\left[\zeta_{\vec{k}_1}(\tau_1),\zeta_{\vec{k}_2}(\tau_2)\right]\cong -i\frac{H^2}{\dot{\varphi}^{(0)\,2}}\frac{G_{k_1}(\tau_1,\tau_2)}{a(\tau_1)a(\tau_2)}\,\delta^{(3)}\left(\vec{k}_1+\vec{k}_2\right)\,,\tag{D.6}$$

where the Green function was defined in (227). This formula is valid only for $\tau_1\geq\tau_2$.

Using the commutativity of the source terms and the formula (D.6) it is straightforward to evaluate the sourced contribution to the n-point correlation functions of ζ. For the 2-point and 3-point we have

$$\begin{aligned}\left\langle\zeta_{\vec{k}_1}\zeta_{\vec{k}_2}(\tau)\right\rangle\Big|_{\text{sourced}} &\approx\left(-\frac{H}{a\dot{\varphi}^{(0)}}\right)^2\int_{-\infty}^{\tau}d\tau_1 d\tau_2\,G_{k_1}(\tau,\tau_1)G_{k_2}(\tau,\tau_2)\,\left\langle J_{\vec{k}_1}(\tau_1)J_{\vec{k}_2}(\tau_2)\right\rangle\,,\\ \left\langle\zeta_{\vec{k}_1}\zeta_{\vec{k}_2}\zeta_{\vec{k}_3}(\tau)\right\rangle\Big|_{\text{sourced}} &\approx\left(-\frac{H}{a\dot{\varphi}^{(0)}}\right)^3\int_{-\infty}^{\tau}d\tau_1 d\tau_2 d\tau_3\,G_{k_1}(\tau,\tau_1)G_{k_2}(\tau,\tau_2)G_{k_3}(\tau,\tau_3)\\ &\qquad\times\left\langle J_{\vec{k}_1}(\tau_1)J_{\vec{k}_2}(\tau_2)J_{\vec{k}_3}(\tau_3)\right\rangle\,.\end{aligned}$$

These coincide exactly with what we obtained previously using the Green function method. We have also verified that the cross-correlation of gauge field fluctuations with the curvature perturbation agrees with what was presented in [79], at leading order.

E Longitudinal Mode of the Massive Gauge Field

In the model (267) of Part VI, the vector has also a longitudinal mode, which was disregarded in the computations performed in the main text. The longitudinal sector is actually more subtle than the transverse one, as perturbation theory may break down in the $m \to 0$ limit. This can be seen, for example, from the action of the longitudinal modes, as we now show.

Let us start by disregarding the perturbations of Ψ, as we do in the main text. Also, for simplicity, we disregard the expansion of the universe. As we have discussed in the main text, this is a good approximation as long as $H \ll \sqrt{m_*}$. Then the mass of the vector field is a classical function of time, and the longitudinal mode is encoded in the decomposed modes

$$A_\mu\left(x\right) = \int \frac{d^3k}{\left(2\pi\right)^{3/2}} \mathrm{e}^{i\vec{k}\cdot\vec{x}} \left(\tilde{A}_0,\, i\, k_i\, \tilde{\chi}\right) \tag{E.1}$$

which leads to the longitudinal sector of Lagrangian

$$\mathcal{L} \supset -\frac{1}{4}F^2 - \frac{m^2}{2}A^2 = \frac{1}{2}\left[k^2|\tilde{\chi}'|^2 - k^2\tilde{A}_0^\dagger\tilde{\chi}' - k^2\tilde{\chi}'^\dagger\tilde{A}_0 + k^2|\tilde{A}_0|^2 - k^2m^2|\tilde{\chi}|^2 + m^2|\tilde{A}_0|^2\right] \tag{E.2}$$

where $m = e\psi^{(0)}$. The component $\tilde{A}_0$ is non-dynamical (it enters in the action without time derivatives) and it can be integrated out. Namely, from (E.2) we obtain the equation

$$\tilde{A}_0 = \frac{k^2}{k^2+m^2}\,\tilde{\chi}' \tag{E.3}$$

Plugging this solution back into (E.2) leads to the action of the longitudinal mode

$$S_{\rm long} = \frac{1}{2}\int dtd^3k\left[|\tilde{L}'|^2 - \left(k^2+m^2+\frac{3k^2m'^2}{\left(k^2+m^2\right)^2} - \frac{m''}{m}\,\frac{k^2}{k^2+m^2}\right)|\tilde{L}|^2\right] \tag{E.4}$$

where $\tilde{L} \equiv \frac{km}{\sqrt{k^2+m^2}}\,\tilde{\chi}$. The field $\tilde{L}$ is the canonical field associated with the longitudinal mode. We see that its equation of motion has a term that formally diverges as $m \to 0$. [63] The formal reason for the divergence is that the kinetic term for the original longitudinal mode vanishes in this limit, as it appears from the relation between $\tilde{L}$ and $\tilde{\chi}$ (E.4). This is not surprising, since a massless vector has only transverse modes.

When, as in the present case, the U(1) symmetry is broken spontaneously, there is actually not a decrease of the number of degrees of freedom when the classical background part $\Psi^{(0)}$ vanishes. The physical mass term in the original action is obtained from $\mathcal{L} \supset e^2|\Psi^{(0)}\left(t\right) + \delta\Psi|^2\, A^2$, and the quantity that we have denoted by m is only related to the classical part, $m^2 \equiv 2e^2|\Psi^{(0)}\left(t\right)|^2$. However, when $\Psi^{(0)}\left(t\right) = 0$, the fluctuations of Ψ cannot be disregarded, and one does not obtain a truly massless vector mode at this point (compare this with what would happen if m was a hard and time dependent mass in the original theory; in this case there would be instead a discontinuity in the number of degrees of freedom at $m = 0$).

In short, a full study of the longitudinal sector would require going beyond the linearized theory; this may also affect the transverse sector, since all the sectors are coupled to each

[63] Clearly, the same equation of motion can also be obtained by writing out the equations for the system in terms of the modes $\tilde{A}_0$ and $\tilde{\chi}$ and by eliminating $\tilde{A}_0$ from these equations.

other beyond the linearized level. Here we use a simpler approach, and discuss under which conditions the longitudinal mode does not blow up in the linearized theory; if this is the case, then one can expect that dealing with the full theory is unnecessary. We note that these considerations apply for a general class of model in which symmetries are enhanced at some point during the cosmological evolution. For instance we expect massive gauge modes to become massless when different branes move to the same bulk location as in the trapping mechanism of [188].

Using the background equations of motion, the dangerous factor m''/m in the linearized computation (E.4) rewrites

$$\frac{m''}{m} = \frac{(a\,\psi)''}{a\,\psi} = \frac{a''}{a} - \frac{a^2\,U_{,\psi}}{\psi} \simeq -\frac{U_{,\psi}}{\psi} \tag{E.5}$$

where the first two equations are exact, and in the final step we instead disregard the expansion of the universe.

Therefore, the equation of motion of $\tilde{L}$ remain finite provided that $U_{,\psi}/\psi$ does not diverge. This could be for instance the case for a quadratic potential $U \sim \frac{1}{2} m_\psi^2 \psi^2$ for $\psi \sim 0$. In our computations, we assumed a constant $\dot{\psi}^{(0)} \equiv \frac{\dot{m}_*}{e}$, corresponding to $\psi^{(0)} = \frac{\dot{m}_*}{e}\,(t - t_*)$. As we discussed in the main text, this implies that the gauge fields are produced during the time interval $\Delta t \sim \frac{1}{\sqrt{\dot{m}_*}}$ around $t = t_*$. Then, for a quadratic potential, imposing that the potential energy gained by $\psi^{(0)}$ during this interval is smaller than its kinetic energy at t_* (so that $\dot{\psi}^{(0)}$ can indeed be taken as constant) amounts in requiring $m_\psi \ll \sqrt{\dot{m}_*}$. This also ensures that the period of oscillations of $\psi^{(0)}$ is much greater than the time in which particle production takes place, so that one can indeed treat α and β as constant when computing the amount of perturbations sourced by the gauge modes.

Using (E.4), we computed the occupation number of longitudinal vector modes with $U \sim \frac{1}{2} m_\psi^2 \psi^2$, and $m_\psi \ll \sqrt{\dot{m}_*}$. We found that in this regime the longitudinal mode is produced in essentially the same amount as each transverse vector mode (for instance, we found that for $m_\psi = 0.1\,\sqrt{\dot{m}_*}$ the total number densities of the longitudinal and of one transverse polarization differ from each other by less than 1%). We stress however that this conclusion is model dependent, as the precise evolution of the effective frequency in (E.4) depends on the details of $U\,(\psi)$. On the contrary, the amounts of the transverse modes produced is independent of U, provided that it remains sufficiently smaller than the kinetic energy during the time of production.

In the remainder of this Appendix we study the spectrum of gravitational waves produced in this model, once also the longitudinal modes are taken into account. When all modes are taken into account, we formally separate the energy-momentum tensor (302) of the vector field into $T_{\mu\nu}^{TT} + T_{\mu\nu}^{LT} + T_{\mu\nu}^{LL}$, where the first term is quadratic in the transverse polarizations, and it is the only one used in the main text, while the third term is quadratic in the longitudinal polarization, and the second term is the “mixed term”. Inserting the last two terms into (54) we obtain the two contributions

$$J_\lambda^{LT}\left(\vec{k}\right)=\frac{\Pi_{ij}^{(\lambda)*}}{aM_p}\left(\hat{k}\right)\int\frac{d^3p}{(2\pi)^{3/2}}2ip_iM^2\left[-\frac{\partial_\tau^{(1)}\,\partial_\tau^{(2)}}{p^2+M^2}+1\right]\tilde{\chi}\left(\vec{p}\right)\tilde{A}_j\left(\vec{k}-\vec{p}\right)$$
$$J_\lambda^{LL}\left(\vec{k}\right)=\frac{\Pi_{ij}^{(\lambda)*}}{aM_p}\left(\hat{k}\right)\int\frac{d^3p}{(2\pi)^{3/2}}M^2p_ip_j\left[-\frac{M^2\partial_\tau^{(1)}\,\partial_\tau^{(2)}}{\left[p^2+M^2\right]\left[(k-p)^2+M^2\right]}+1\right]\tilde{\chi}\left(\vec{p}\right)\tilde{\chi}\left(\vec{k}-\vec{p}\right) \tag{E.6}$$

to the gravitational wave source in (53), where we remind $M = a\,m = a\,e\,\psi^{(0)}$. These add up to the term J_λ^{TT}, which is the only one studied in Subsection VI 1.2 of the main text (where it is denoted simply by J_λ). To obtain the total power spectrum we need to evaluate the correlator $\left\langle (J_\lambda^{TT}+J_\lambda^{LT}+J_\lambda^{LL})_{\tau_1,\vec{k}_1}(J_{\lambda'}^{TT}+J_{\lambda'}^{LT}+J_{\lambda'}^{LT})_{\tau_2,\vec{k}_2}\right\rangle$ and insert it in (63). Each piece in J_λ is correlated only with the corresponding piece in $J_{\lambda'}$.

Proceeding as in the main text, we obtain

$$\left\langle J_\lambda^{LT}\left(\tau_1,\,\vec{k}\right)J_\lambda^{LT}\left(\tau_2,\,\vec{k}'\right)\right\rangle\simeq\frac{2}{5\pi^2M_p^2}\delta_{\lambda\lambda'}\delta^{(3)}\left(\vec{k}+\vec{k}'\right)\frac{M\left(\tau_1\right)^2}{a\left(\tau_1\right)}\frac{M\left(\tau_2\right)^2}{a\left(\tau_2\right)}\int dpp^4$$
$$\times\left[-\frac{\delta_{a1}}{p^2+M^2}+\delta_{a0}\right]_{\tau_1}\left[-\frac{\delta_{b1}}{p^2+M^2}+\delta_{b0}\right]_{\tau_2}\mathcal{C}_{(a,b)}\left[\tau_1,\,\tau_2;\,p\right]\mathcal{D}_{(a,b)}\left[\tau_1,\,\tau_2;\,p\right] \tag{E.7}$$

and

$$\left\langle J_\lambda^{LL}\left(\tau_1,\,\vec{k}\right)J_\lambda^{LL}\left(\tau_2,\,\vec{k}'\right)\right\rangle\simeq\frac{2}{15\pi^2M_p^2}\delta_{\lambda\lambda'}\delta^{(3)}\left(\vec{k}+\vec{k}'\right)\frac{M\left(\tau_1\right)^2}{a\left(\tau_1\right)}\frac{M\left(\tau_2\right)^2}{a\left(\tau_2\right)}\int dpp^6$$
$$\times\left[-\frac{M^2\,\delta_{a1}}{\left(p^2+M^2\right)^2}+\delta_{a0}\right]_{\tau_1}\left[-\frac{M^2\,\delta_{b1}}{\left(p^2+M^2\right)^2}+\delta_{b0}\right]_{\tau_2}\mathcal{C}^2_{(a,b)}\left[\tau_1,\,\tau_2;\,p\right] \tag{E.8}$$

In these expressions we have introduced the correlators

$$\left\langle\tilde{\chi}\left(\tau_1,\,\vec{p}_1\right)\tilde{\chi}\left(\tau_2,\,\vec{p}_2\right)\right\rangle=\mathcal{C}_{(0,0)}\left[\tau_1,\,\tau_2;\,p_1\right]\,\delta^{(3)}\left(\vec{p}_1+\vec{p}_2\right) \tag{E.9}$$

which evaluate to

$$\mathcal{C}_{(0,0)}\left[\tau_1,\,\tau_2;\,p\right]=\frac{\sqrt{p^2+M^2\left(\tau_1\right)}\sqrt{p^2+M^2\left(\tau_2\right)}}{p^2M\left(\tau_1\right)M\left(\tau_2\right)}\mathcal{D}^L_{(0,0)}\left[\tau_1,\,\tau_2;\,p\right] \tag{E.10}$$

where $\mathcal{D}^L_{(a,b)}$ are formally identical to the $\mathcal{D}^{(\sigma)}_{(a,b)}$ quantities given in eqs. (276) and (278), with the only difference that in $\mathcal{D}^L_{(a,b)}$ we use the Bogolyubov coefficients and the mode functions of the longitudinal mode:

$$\tilde{L}\left(\vec{k}\right) = \left[\alpha_k^L \, g_k + \beta_k^L \, g_k^*\right] a_{\vec{k}}^L + \text{h.c.} \tag{E.11}$$

where a_k^L is an annihilation operator of the longitudinal vector mode, and

$$g_k \equiv \frac{\mathrm{e}^{-i\int^\tau d\tau' \, \omega_L(\tau')}}{\sqrt{2\omega_L\left(\tau\right)}} \;, \quad \omega_L \equiv \left(k^2 + M^2 + \frac{3k^2 M'^2}{\left(k^2+M^2\right)^2} - \frac{M''}{M}\frac{k^2}{k^2+M^2}\right)^{1/2} \tag{E.12}$$

As in the computation of $\langle J_\lambda^{TT} J_{\lambda'}^{TT}\rangle$ presented in the main text, the square of the correlators appearing in (E.7) and (E.8) contain terms proportional to fast oscillating phases, which give a negligible contribution to the final result. Disregarding these terms, we obtain

$$\begin{aligned}\left\langle J_\lambda^{LT}\left(\tau_1, \vec{k}\right) J_\lambda^{LT}\left(\tau_2, \vec{k}'\right)\right\rangle &\simeq \frac{-\delta_{\lambda\lambda'}}{20\pi^2 M_p^2}\delta^{(3)}\left(\vec{k}+\vec{k}'\right) \\ &\times \int dp p^6 \mathrm{Re}\left[\alpha_p \alpha_p^{L*}\beta_p^*\beta_p^L Q\left(\tau_1\right) Q\left(\tau_2\right) - \left|\beta_p\right|^2 \left|\beta_p\right|^2 Q\left(\tau_1\right) Q^*\left(\tau_1\right)\right]\end{aligned} \tag{E.13}$$

where

$$Q\left(\tau\right) \equiv \frac{1}{aM}\left[\frac{2M'}{M^2} - \frac{3iM'^2}{M^4} + \frac{iM''}{M^3}\right] \tag{E.14}$$

and

$$\begin{aligned}\left\langle J_\lambda^{LL}\left(\tau_1, \vec{k}\right) J_\lambda^{LL}\left(\tau_2, \vec{k}'\right)\right\rangle &\simeq \frac{\delta_{\lambda\lambda'}}{15\pi^2 M_p^2}\,\delta^{(3)}\left(\vec{k}+\vec{k}'\right)\frac{1}{aM\left(\tau_1\right)}\frac{1}{aM\left(\tau_2\right)} \\ &\times \int dp p^6 |\beta_p^L|^2 \left(|\alpha_p^L|^2 + |\beta_p^L|^2\right)\end{aligned} \tag{E.15}$$

In these expressions, we have retained only the dominant contributions to the ω_L in the adiabatic regime (namely, we have used $M' \ll M^2$ and $M'' \ll M^3$). As in the computation of the main text, the terms without oscillatory phases are of $\mathrm{O}\left(p^6\right)$, as a consequence to a cancellation of the terms that would be dominant terms in the transverse and traceless projection of the energy-momentum tensor. This cancellation is already visible in (E.7) and (E.8), see the discussion after the analogous expression (304) presented in the main text.

We see that the mixed term (E.13) is suppressed in the adiabatic regime, and can be disregarded. From the (E.15) we instead obtain

$$P_\lambda^{\text{sourced}}\Big|_{\text{from } A_L} \simeq \frac{2k^3}{15\pi^4 a^2 M_p^4}\tilde{\mathcal{T}}_k^2 \int dp \, p^6 |\beta_p^L|^2 \left(|\alpha_p^L|^2 + |\beta_p^L|^2\right) \tag{E.16}$$

where $\tilde{\mathcal{T}}_k$ is defined in (308). This contribution adds up to the one of the transverse modes leading to the result (313) given in the main text.

F Fermionic Production and Gravitational Waves

In this Appendix, we outline the computation of gravitational waves produced by a fermion with a mass varying as in Model I of Part VI, studied in the main text. If we denote by X the fermion in the original action, and we rescale, $\chi = X\, a^{3/2}$, the action for the fermion field becomes identical to the one of a massive fermion in Minkowski spacetime, whose mass is multiplied by the scale factor

$$S_f = \int d^4x\, \bar{\chi} \left[i\, \gamma^\mu \partial_\mu - M\left(t\right) \right] \chi \quad , \quad M = g\, a\left(t\right) \left[\psi^{(0)}\left(t\right) - \psi_* \right] \tag{F.1}$$

As in the main text, $\psi^{(0)}$ is a homogeneous classical field that evaluates to ψ_* at some given moment t_* during inflation, while g is the coupling of the Yukawa interaction in the original action. Without loss of generality we can choose $\psi_* = t_* = 0$. As in the main text, we consider a regime in which the expansion of the universe can be disregarded during the particle production, and we expand

$$g\, \psi^{(0)}\left(t\right) \equiv \dot{m}_*\, t \quad , \quad t \simeq 0 \tag{F.2}$$

We instead include the expansion of the universe when we study the amount of gravitational waves sourced by the fermionic quanta produced at $t \simeq 0$.

The fermionic production in this model was studied in [146, 152, 153, 249] for reheating after inflation and leptogenesis, and in [94] for the imprint on the scalar power spectrum. Here we follow the computations of [244, 249], skipping some intermediate steps. We refer the reader to those works for details. One decomposes[64]

$$\chi = \int \frac{d^3k}{\left(2\pi\right)^{3/2}}\, \mathrm{e}^{i\vec{k}\cdot\vec{x}} \tilde{\chi}\left(\tau,\, \vec{k}\right) \;,\quad \tilde{\chi}\left(\tau,\, \vec{k}\right) = \sum_r \left[\mathcal{U}_r\left(\tau,\, \vec{k}\right)\, a_r\left(\vec{k}\right) + \mathcal{V}_r\left(\tau,\, -\vec{k}\right)\, b_r^\dagger\left(-\vec{k}\right) \right] \tag{F.3}$$

where, for $\vec{k}$ aligned along the $z-$axis,

$$\mathcal{U}_r\left(\tau,\, k_z\right) = \frac{1}{\sqrt{2}} \left(\begin{array}{c} U_+\left(\tau,\, k\right)\, \psi_r \\ U_-\left(\tau,\, k\right)\, r\, \psi_r \end{array} \right) \;,\quad \mathcal{V}_r\left(\tau,\, -k_z\right) = \frac{1}{\sqrt{2}} \left(\begin{array}{c} -V_+\left(\tau,\, k\right)\, \psi_{-r} \\ -V_-\left(\tau,\, k\right)\, r\, \psi_{-r} \end{array} \right) \tag{F.4}$$

where $\psi_+ = \left(\begin{array}{c} 1 \\ 0 \end{array} \right)$ and $\psi_- = \left(\begin{array}{c} 0 \\ 1 \end{array} \right)$. The spinor $\mathcal{V}$ is related to $\mathcal{U}$ by charge conjugation, giving $V_\pm\left(\tau,\, k\right) = U_\mp^*\left(\tau,\, k\right)$.

It is convenient to decompose the spinors in terms of the Minkowski solutions (more

[64] We use

$$\gamma^0 = \left(\begin{array}{cc} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{array} \right) \quad , \quad \gamma^i = \left(\begin{array}{cc} 0 & \sigma^i \\ -\sigma^i & 0 \end{array} \right)$$

and, only in this Appendix, we switch to $+, -, -, -$ signature for the metric.

precisely, the adiabatic solution, since in the current case M is not constant)

$$U_{\pm}(\tau,\, k) = \alpha\,(t)\, f_{\pm}(\tau,\, k) \mp \beta\,(t)\, f^*_{\mp}(\tau,\, k) \quad , \quad f_{\pm}(\tau,\, k) \equiv \sqrt{1 \pm \frac{M}{E}}\, \mathrm{e}^{-i\int_{\tau_*}^{\tau} d\tau'\, E} \tag{F.5}$$

and $E \equiv \sqrt{k^2 + M^2}$. Starting from $\beta = 0$ at asymptotically early times, and using (F.2) one finds, at late times (see, for instance, [249] for the computation)

$$\alpha\,(t \gg 0) \simeq \frac{2\sqrt{\pi}}{q} \mathrm{e}^{iq^2/2} \mathrm{e}^{i\pi/4} \left(\frac{q}{\sqrt{2}}\right)^{-iq^2} \frac{\mathrm{e}^{-\pi q^2/4}}{\Gamma\left(\frac{-iq^2}{2}\right)} \;, \quad \beta\,(t \gg 0) \simeq -\mathrm{e}^{-\pi q^2/2} \;, \quad q \equiv \frac{k}{\sqrt{\dot{m}_*}} \tag{F.6}$$

As in the bosonic case, $|\beta|^2$ is the occupation number, and Pauli blocking is ensured by the fact that, in the fermionic case, the Bogolyubov coefficients satisfy $|\alpha|^2 + |\beta|^2 = 1$.

To compute the gravitational waves produced by these quanta, we proceed as in Part III, and obtain the formal solution (62) for the sourced part $Q_\lambda^{\rm sourced}$ of the canonical gravitational wave modes introduced in (52). Using the fermionic energy momentum tensor

$$T_{\mu\nu} = \frac{i}{4a^2} \left[\bar{\chi}\gamma_\mu\partial_\nu\chi + \bar{\chi}\gamma_\nu\partial_\mu\chi - (\partial_\mu\bar{\chi})\,\gamma_\nu\chi - (\partial_\nu\bar{\chi})\,\gamma_\mu\chi\right] \tag{F.7}$$

we can cast the source appearing in (62) in the form

$$J_\lambda\left(\tau,\, \vec{k}\right) = \frac{1}{2aM_p}\, \Pi_{ij}^{(\lambda)*}\left(\hat{k}\right) \int \frac{d^3p}{(2\pi)^{3/2}} \bar{\tilde{\chi}}\,(\tau,\, \vec{p})\, (\gamma_i\, p_j + \gamma_j\, p_i)\; \tilde{\chi}\left(\tau,\, \vec{k} + \vec{p}\right) \tag{F.8}$$

After Wick contraction, the correlator of two sources acquires the form

$$\begin{aligned} \left\langle J_\lambda\left(\tau,\, \vec{k}\right) J_{\lambda'}\left(\tau',\, \vec{k}'\right)\right\rangle \;&=\; \frac{1}{M_p^2 a\,(\tau)\, a\,(\tau')}\, \Pi_{ij}^{(\lambda)*}\left(\hat{k}\right) \Pi_{lm}^{(\lambda')*}\left(\hat{k}'\right) \int \frac{d^3p\, d^3p'}{(2\pi)^3} \\ &\times \mathrm{Tr}\left[\gamma^i p^j \left\langle \tilde{\chi}\left(\tau,\, \vec{k}+\vec{p}\right)\, \bar{\tilde{\chi}}\,(\tau',\, \vec{p}')\right\rangle \gamma^l p'^m \left\langle \bar{\tilde{\chi}}\,(\tau,\, \vec{p})\; \tilde{\chi}\left(\tau',\, \vec{k}'+\vec{p}'\right)\right\rangle^T\right] \end{aligned} \tag{F.9}$$

where the transposition acts on the spinor indices. The correlators appearing in this expression need to be regularized. We adopt the same prescription adopted in Section VI 1 for the vector field correlators. Namely, we normal order the fields appearing in the correlators with respect to the time dependent operators

$$\begin{aligned} \hat{a}\left(\vec{k}\right) \;&\equiv\; \alpha\, a\left(\vec{k}\right) - \beta^*\, b^\dagger\left(-\vec{k}\right) \\ \hat{b}^\dagger\left(-\vec{k}\right) \;&\equiv\; \beta\, a\left(\vec{k}\right) + \alpha^*\, b^\dagger\left(-\vec{k}\right) \end{aligned} \tag{F.10}$$

that diagonalize the Hamiltonian at any given time. We then recall that the vacuum of the theory is annihilated by the original time-independent a and b operators. After some algebra, we obtain

$$\left\langle : \tilde{\chi}\left(\tau,\,\vec{k}+\vec{p}\right)\,\bar{\tilde{\chi}}\left(\tau',\,\vec{p}'\right) : \right\rangle = \delta^{(3)}\left(\vec{p}'-\vec{p}-\vec{k}\right)\,\begin{pmatrix} C_{11}\left[\tau,\,\tau';\,p'\right]\,\mathbb{1} & C_{12}\left[\tau,\,\tau';\,p'\right]\,\vec{\sigma}\cdot\hat{p}' \\ C_{21}\left[\tau,\,\tau';\,p'\right]\,\vec{\sigma}\cdot\hat{p}' & C_{22}\,\left[\tau,\,\tau';\,p'\right]\,\mathbb{1} \end{pmatrix}$$
$$= \delta^{(3)}\left(\vec{p}'-\vec{p}-\vec{k}\right)\,\left(\frac{\mathcal{C}_{11}+\mathcal{C}_{22}}{2}\,\mathbb{1}+\frac{\mathcal{C}_{11}-\mathcal{C}_{22}}{2}\,\gamma^0 + \frac{\mathcal{C}_{12}+\mathcal{C}_{21}}{2}\,\gamma^0\vec{\gamma}\cdot\hat{p}'+\frac{\mathcal{C}_{12}-\mathcal{C}_{21}}{2}\vec{\gamma}\cdot\hat{p}'\right)$$
$$\left\langle \bar{\tilde{\chi}}\left(\tau,\,\vec{p}\right)\,\tilde{\chi}\left(\tau',\,\vec{k}'+\vec{p}'\right)\right\rangle^T = \delta^{(3)}\left(\vec{p}-\vec{p}'-\vec{k}'\right)\,\begin{pmatrix} -\mathcal{C}_{22}\left[\tau,\,\tau';\,p\right]\,\mathbb{1} & \mathcal{C}_{21}^*\left[\tau,\,\tau';\,p\right]\,\vec{\sigma}\cdot\hat{p} \\ \mathcal{C}_{12}^*\left[\tau,\,\tau';\,p\right]\,\vec{\sigma}\cdot\hat{p} & -\mathcal{C}_{11}\left[\tau,\,\tau';\,p\right]\,\mathbb{1} \end{pmatrix} \tag{F.11}$$

where, when particle creation has completed, and (F.6) hold,

$$\begin{aligned}
\mathcal{C}_{11}\left[\tau,\,\tau';\,p'\right] &= \frac{1}{2}\Big\{-|\beta|^2\left[f_+\left(\tau,p'\right)f_+^*\left(\tau',\,p'\right)-f_-^*\left(\tau,p'\right)f_-\left(\tau',\,p'\right)\right] \\
&\qquad -\alpha\beta^* f_+\left(\tau,p'\right)f_-\left(\tau',\,p'\right)-\alpha^*\beta f_-^*\left(\tau,p'\right)f_+^*\left(\tau',\,p'\right)\Big\} \\
\mathcal{C}_{12}\left[\tau,\,\tau';\,p'\right] &= \frac{1}{2}\Big\{|\beta|^2\left[f_+\left(\tau,p'\right)f_-^*\left(\tau',\,p'\right)+f_-^*\left(\tau,p'\right)f_+\left(\tau',\,p'\right)\right] \\
&\qquad -\alpha\beta^* f_+\left(\tau,p'\right)f_+\left(\tau',\,p'\right)+\alpha^*\beta f_-^*\left(\tau,p'\right)f_-^*\left(\tau',\,p'\right)\Big\} \\
\mathcal{C}_{21}\left[\tau,\,\tau';\,p'\right] &= -\mathcal{C}_{12}^*\left[\tau,\,\tau';\,p'\right] \\
\mathcal{C}_{22}\left[\tau,\,\tau';\,p'\right] &= \mathcal{C}_{11}^*\left[\tau,\,\tau';\,p'\right]
\end{aligned} \tag{F.12}$$

We insert these expressions into (F.9). We evaluate the trace, and (as in the analogous computation of the main text) disregard $\vec{k}$ and $\vec{k}'$ in comparison to $\vec{p}$ and $\vec{p}'$. The d^3p' integral can be performed using one delta function. We perform the angular part of the remaining d^3p integral. Finally, we use the property of the polarization operators Π_{ij}. We obtain

$$\left\langle J_\lambda\left(\tau,\,\vec{k}\right)J_{\lambda'}\left(\tau',\,\vec{k}'\right)\right\rangle = \frac{2}{15\,\pi^2}\frac{\delta^{(3)}\left(\vec{k}+\vec{k}'\right)\delta_{\lambda\lambda'}}{M_p^2 a\left(\tau\right)a\left(\tau'\right)}\int dp\,p^4\,\mathrm{Re}\left[5\,\mathcal{C}_{11}^2+\mathcal{C}_{12}^2\right] \tag{F.13}$$

Squaring the correlators, we have

$$\begin{aligned}
\mathcal{C}_{11}^2 &= |\beta|^2\left(|\alpha|^2-|\beta|^2\right)\,\frac{p^2}{2\,E\left(\tau\right)\,E\left(\tau'\right)}+\text{oscillatory phases} \\
\mathcal{C}_{12}^2 &= -\mathcal{C}_{11}^2+\text{oscillatory phases}
\end{aligned} \tag{F.14}$$

and we see that the non oscillatory part of the integrand in (F.13) is of $\mathcal{O}\left(p^6\right)$ as in the vector case studied in the main text. The oscillatory part gives a negligible contribution to the tensor power. Combining this result with (63) and (72) we obtain

$$P_\lambda^{\rm sourced} \simeq \frac{8k^3}{15\pi^4 a^2 M_p^4} \tilde{\mathcal{T}}_k^2 \int dp\, p^6 |\beta|^2 \left(|\alpha|^2 - |\beta|^2 \right) \tag{F.15}$$

where $\tilde{\mathcal{T}}_k$ is defined in (308).

Inserting the result (309) for this quantity, and performing the momentum integral, gives

$$P_\lambda^{\rm sourced} \simeq \frac{\left(8-\sqrt{2}\right)\, H^4\, \dot{m}_*^{3/2}}{16\,\pi^7\, M_p^4\, k^3} \left[\sin\left(\frac{k}{H}\right) - \frac{k}{H}\, \cos\left(\frac{k}{H}\right) \right]^2 \ln^2\left(\frac{\sqrt{\dot{m}_*}}{H}\right) \;, \quad k \ll \sqrt{\dot{m}_*} \tag{F.16}$$

while the result is exponentially suppressed at higher momenta.