Master Thesis

Pseudo Nambu-Goldstone Boson Inflation

prepared by

Stefan Marinus Rodrigues Sandner

at the DESY Theory Group &
II. Institute of Theoretical Physics

April 2019

First referee: Dr. Valerie Domcke
Second referee: Prof. Dr. Géraldine Servant
Supervisor: Dr. Valerie Domcke
Abstract

We investigate an early Universe inflationary phase driven by a pseudo Nambu-Goldstone boson. The setup is not only naturally protected in the sense of t’Hooft, but such particles are also ubiquitous in quantum and string theories. In particular, we may adopt the interpretation of an axion. We derive phenomenological interesting consequences by couplings to (non-) Abelian gauge fields. The only interaction term which respects the axion shift symmetry in this setup is of Chern-Simons form. Starting from the Bunch-Davies vacuum, we show that tachyonic enhanced gauge field modes source gravitational waves strong enough to be detectable at various stages of the inflationary phase. However, agreement with cosmic microwave background (CMB) observations require a super-Planckian periodicity scale of the potential. We try to recover a viable effective field theory (EFT) description by distinguishing between the axion decay constant and the periodicity scale. This may be achieved by e.g. aligning multiple axion fields. Nonetheless, we present a bottom-up EFT approach which puts stringent bounds on the non-Abelian model — almost entirely closing the parameter space. Additionally, we show that field excursions in the remaining viable window of large gauge couplings are of order $10^4$ times larger than the axion decay constant — indicating a difficult string theoretical realization. We propose a novel more realistic model of simultaneous axion couplings to both Abelian and non-Abelian gauge fields. In the case of a coupling hierarchy, such a degenerated gauge field unification gracefully circumvents the EFT bound. This complex gauge group structure is indeed desirable to provide a natural Standard Model connection to reheat the Universe through the Adler-Bell-Jackiv anomaly. We discuss that only gauge field self-interactions may lead to a formation of a thermal bath. Moreover, we introduce distortions of the CMB as a new strong constraint and simultaneously show that tensor contributions to them are generically highly suppressed.
Contents

1 Introduction 1

2 Cosmology and Inflation Basics 3

3 Generic Pseudo Nambu-Goldstone Boson Inflation 9
  3.1 Setup & Coupling to Abelian Gauge Fields .......................... 10
  3.2 Numerical Treatment & Phenomenology .............................. 14
  3.3 CMB Distortion as a New Strong Constraint ....................... 19

4 Natural and Chromo-Natural Inflation 23
  4.1 Framework .......................................................... 24
     4.1.1 Homogeneous Background ...................................... 25
     4.1.2 Gauge Field Fluctuations ...................................... 27
     4.1.3 Background Transition ........................................ 30
     4.1.4 Gravitational Wave Fluctuation ............................... 33
  4.2 Numerical Treatment & Phenomenology ............................ 34
  4.3 Model Constraints .................................................. 40
     4.3.1 Numerical Limit ............................................... 40
     4.3.2 Effective Field Theory Bounds ............................... 41
     4.3.3 Parameter Scan ................................................. 45

5 Towards a Completion of Natural Inflation 48
  5.1 Gauge Field Unification ............................................ 48
     5.1.1 Guideline & Phenomenology ................................... 49
     5.1.2 Numerical & Theoretical Bounds ............................... 55
  5.2 Standard Model Connection ......................................... 57
     5.2.1 Coupling to Fermions .......................................... 58
     5.2.2 Thermalization .................................................. 63
     5.2.3 Phenomenology .................................................. 66

6 Discussion 68

A Window Function 72

B Comments on the Planck Cosmological Parameters 74
  B.1 Energy Densities .................................................... 74
  B.2 Uncertainty in the required Inflation Time ...................... 75
Notation

We use the natural unit system $\hbar = c = k_B = 1$. This implies the dimensional relation $[\text{Energy}] = [\text{Mass}] = [\text{Temperature}] = [\text{Time}^{-1}] = [\text{Length}^{-1}]$. Additionally we set the reduced Planck mass to unity, i.e. $M_p^2 = \hbar c/(8\pi G_{\text{Newton}}) = 1$.

We work within the framework of Einstein general relativity with the metric described by a flat Friedman-Lemaître-Robertson-Walker space-time $ds^2 = -dt^2 + a(t)dx^2$. This implies the metric signature $(-,+,+,+)$. The scale factor $a$ encodes the expansion of the Universe. On any quantity a subscript zero, i.e. $X_0$, denotes that it must be evaluated today.

We define the inflationary e-folding number as $dN \equiv -H dt$. This implies $N_i < N_j$ for $t_i > t_j$.

A four tensor is denoted with Greek indices $(\alpha, \beta, \ldots)$ running over $(0, 1, 2, 3)$, such that a contravariant four vector is $X^\alpha \equiv (X^0, X)$. So, a three vector is denoted by $\mathbf{X}$ and a three tuple is labeled with Latin indices starting from $i$ running over $(1, 2, 3)$, i.e. $X_i$. Indices are raised and lowered by means of the FLRW metric. Additionally we label the gauge indices with Latin indices starting from $a$. We sum over repeated indices.

The totally anti-symmetric rank-4 Levi-Civita tensor is used with the convention $\epsilon^{0123} = -\epsilon_{0123} = +1$.

Partial derivatives are indicated as $\frac{\partial}{\partial X} \equiv \partial_X$. The derivative with respect to cosmic time $t$ is denoted as $\partial_t X \equiv \dot{X}$ and with respect to conformal time as $\partial_{\tau} X \equiv X'$. We take for the Einstein Telescope a signal-to-noise ratio of $\text{SNR} = 10$ and a run time of $t = 10$ years. For LISA we take the same.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΛCDM</td>
<td>Standard model of cosmology</td>
</tr>
<tr>
<td>SM</td>
<td>Standard Model of particle physics</td>
</tr>
<tr>
<td>GW</td>
<td>Gravitational wave</td>
</tr>
<tr>
<td>CMB</td>
<td>Cosmic Microwave Background</td>
</tr>
<tr>
<td>pNGB</td>
<td>Pseudo Nambu-Goldstone boson</td>
</tr>
<tr>
<td>CS</td>
<td>Chern-Simons</td>
</tr>
<tr>
<td>EOM</td>
<td>Equation of motion</td>
</tr>
<tr>
<td>CNI</td>
<td>Chromo-natural inflation</td>
</tr>
<tr>
<td>EFT</td>
<td>Effective field theory</td>
</tr>
<tr>
<td>FLRW</td>
<td>Friedman-Lemaître-Robertson-Walker</td>
</tr>
<tr>
<td>BBN</td>
<td>Big Bang nucleosynthesis</td>
</tr>
<tr>
<td>GUT</td>
<td>Grand Unified Theory</td>
</tr>
<tr>
<td>vev</td>
<td>Vacuum expectation value</td>
</tr>
<tr>
<td>PQ</td>
<td>Peccei &amp; Quinn</td>
</tr>
<tr>
<td>ODE</td>
<td>Ordinary differential equation</td>
</tr>
<tr>
<td>GR</td>
<td>Einstein general relativity</td>
</tr>
<tr>
<td>QFT</td>
<td>Quantum field theory</td>
</tr>
<tr>
<td>ET</td>
<td>Einstein Telescope</td>
</tr>
<tr>
<td>LISA</td>
<td>Laser Interferometer Space Antenna</td>
</tr>
<tr>
<td>ABJ</td>
<td>Adler-Bell-Jackiv</td>
</tr>
<tr>
<td>HLL</td>
<td>Higher Landau level</td>
</tr>
<tr>
<td>LLL</td>
<td>Lowest Landau level</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Experimental progress in the recent years pushed theoretical cosmology to a high precision field. Especially the measurement of the cosmic microwave background (CMB) [1–5] and of the first gravitational waves (GWs) [6–13] mark the opening of new windows to test cosmological models. The standard model of cosmology (ΛCDM) — containing a cosmological constant Λ and cold dark matter — is very successful in describing the evolution of our Universe from the Big Bang until today. The model passes stringent tests, which reach from the large scale structure distribution over the cosmic abundance of the light elements to the existence of the CMB. However, some of the measurements can only be explained by highly fine tuned initial conditions of the model. For example the CMB contains within the ΛCDM framework many causally disconnected regions. This does not fit the measured homogeneity naturally. The idea of cosmic inflation [14, 15] proposes that an exponential expansion of the early Universe can naturally explain the CMB — and other here not so important shortcomings [16–19]. Interestingly, such a setup builds the connection between early Universe quantum fluctuations to the large scale structure observed today. A promising realization of this phase is through a pseudo Nambu-Goldstone boson (pNGB). This setup is protected by technical naturalness in the sense of t’Hooft [20], such that (s.t.) the initial condition problem of ΛCDM is solved and not shifted as in some other inflation models [21, 22]. Notably, it may also lead to interesting phenomenological implications, if we couple the pNGB to (non-) Abelian gauge fields through a derivative coupling, namely a Chern-Simons (CS) term. This does not break the naturalness of the setup, since the shift-symmetry of the pNGB is preserved. Rather, it is enhances by providing a simple way to couple the dark inflation sector to
the Standard Model of particle physics (SM). Much effort has been put into the study of possible gravitational wave signals [23–25], primordial black hole production [26–28] and the formation of non-Gaussian contributions to scalar and tensor power spectra [29–32]. Widely unnoticed in this inflationary context, however, is the possibility that the energy release at high redshift may lead to a development of a chemical potential [33–35]. The result is a distortion of the CMB black body spectrum at high frequencies, thereby further constraining the window for viable models. We see that although we cannot directly look further back than the CMB release, the astonishing sensitivity of experiments put tight constraints on model building of early Universe physics.

In this thesis we study the inflationary phase, starting from its general conditions up to the physical plausible scenario of pNGB inflation in the presence of Abelian and non-Abelian gauge groups. The final connection to the SM is also studied. Therefore, we structure the thesis as follows. In chapter 2 we review the required cosmological basics to derive conditions each inflation model has to fulfill. We then focus in chapter 3 on generic pNGB inflationary models. We derive the equations of motions (EOMs) in the vacuum case as well as in the presence of an Abelian CS interaction. The phenomenological implication of such a setup is studied, focusing on the formation of CMB distortions. In the subsequent chapter 4 we investigate the natural interpretation that the pNGB is an axion-like particle. In the presence of non-Abelian gauge fields this is known under the name chromo-natural inflation [36]. The evolution of the gauge field is discussed, starting from the quantum mechanical Bunch-Davies vacuum. This is contrary to what is frequently studied in literature [24, 36–40]. We provide a detailed parameter scan by numerically solving the EOMs. A bottom-up discussion of the viability of the effective field theory (EFT) description is done. In chapter 5 we demonstrate a model of axion-like inflation in the presence of coupling to both, Abelian and non-Abelian fields. The necessity of this unified approach is discussed in the context of required SM coupling. We summarize and discuss our results in chapter 6.
Chapter 2

Cosmology and Inflation Basics

We give a short introduction to aspects of cosmology important within our context. The focus will be on the vanishing curvature of our Universe and the homogeneity of the CMB, both measured to high precision. We show how an inflationary phase can naturally explain them. The main point is the avoidance of fine tuned initial conditions.

The Copernican principle [41–43] assumes homogeneity and isotropy for our Universe. The resulting line element for the space-time connection is then described with the Friedmann-Lemaître-Robertson-Walker (FLRW) metric [44].

$$d s^2 = -d t^2 + a^2(t) \left( \frac{d r^2}{1 - k r^2} + r^2 (d \theta^2 + \sin^2 \theta d \phi^2) \right).$$ \hspace{1cm} (2.1)

Here and throughout the notes we use a 4D metric with signature diag$(-1, +1, +1, +1)$ and the system of natural units, i.e. $c = \hbar = k_B = 1$. This implies the dimensional relation [Energy] = [Mass] = [Temperature] = [Time$^{-1}$] = [Length$^{-1}$]. The scale factor $a(t)$ characterize the evolution of the size of a spacelike hypersurface with physical time $t$. The normalization is chosen to be $a(t_{\text{today}}) \equiv a_0 \equiv 1$. On any used quantity a subscript zero will indicate that it has to be evaluated today, unless otherwise mentioned. To account for the observed flatness of the Universe [2] we will set the curvature parameter $k$ equal to zero, s.t.

$$d s^2 = -d t^2 + a^2(t) d x^2.$$ \hspace{1cm} (2.2)
Thus, the evolution of the Universe is only characterized by the scale factor \( a(t) \) and its form can be obtained by solving the Einstein field equation

\[
G_{\mu\nu} = T_{\mu\nu}.
\]

(2.3)

The left hand side defines the Einstein tensor and the right hand side the stress-energy tensor, connecting space-time and matter. We additionally set the reduced Planck mass to unity, i.e. \( M_p^2 = \frac{\hbar c}{8\pi G_{\text{Newton}}} = 1 \). The Einstein tensor is trivially fixed once the metric is defined, which in our case is the FLRW. So, to proceed we have to find a general form for the stress-energy tensor. By noticing that \( T_{\mu\nu} \) must fulfill the Bianchi identity (since \( G_{\mu\nu} \) does) it must be symmetric and by isotropy the spatial components must be equal. The simplest realization is that of a perfect fluid

\[
T_{\mu\nu} = \text{diag}(\rho, -p, -p, -p),
\]

(2.4)

where \( \rho \) denotes the energy density and \( p \) the pressure. Then, the Einstein field equation reduces to the two coupled Friedmann equations

\[
H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{\rho}{3}
\]

(2.5)

\[
\dot{H} + H^2 = \frac{\ddot{a}}{a} = -\frac{1}{6}(\rho + 3p).
\]

(2.6)

Here we introduced the Hubble parameter \( H = \dot{a}/a \) with unit \([\text{Time}^{-1}]\). If more than one matter species contribute to the total energy density or pressure — as it is the case in our Universe — the total contribution is obtained by simply summing over all constituents, i.e. \( \rho = \sum_i \rho_i \) and \( p = \sum_i p_i \). For each species we can define the contribution to the energy budget of the Universe today as

\[
\Omega_i = \frac{\rho_{0,i}}{\rho_{\text{crit}}} = \frac{\rho_{0,i}}{3H_0^2},
\]

(2.7)

where the critical density \( \rho_{\text{crit}} \) is given by the today’s density for exactly zero curvature contribution. Here we already encounter the first ΛCDM puzzle: Why does the energy density of our Universe coincide with the critical one, implying the consistency relation \( \sum_i \Omega_i = 1 \)?

Let us make the arising puzzle clearer by rewriting the first Friedmann eqs. into
the most convenient form

\[ H^2 = H_0^2 \sum_i \Omega_i a^{-3(1+\omega_i)}, \]  

(2.8)

where we introduced the equation of state parameter \( \omega \equiv p/\rho \). The values for \( H_0 \) and \( \Omega_i \) are measured by the Planck Collaboration to very high accuracy [2], s.t. the only free parameter is \( \omega \). For non-relativistic matter we have \( \omega = 0 \) and for relativistic matter/radiation we have \( \omega = 1/3 \). The cosmological constant is defined to obey \( p = -\rho \), leading to \( \omega = -1 \), which makes the Hubble parameter constant. The Friedmann equations indicate that for an expanding Universe, \( \dot{a} > 0 \), we have \( \ddot{a} < 0 \) if the right hand side of eq. (2.6) is negative. This holds if the Universe only contains ordinary matter which satisfy the strong energy condition \( (\rho + 3p) \geq 0 \). This in turns means, that for such a Universe there has been a singularity in the finite past s.t. \( a(t = 0) = 0 \), known as the Big Bang \(^1\). Separating the variables and then integrate equation (2.8), leads to (for \( \omega \neq -1 \))

\[ \int dt \propto \int a^{\frac{1}{2}(1+3\omega)} \quad \Rightarrow \quad t \propto a^{\frac{2}{3}(1+\omega)} \quad \Leftrightarrow \quad a(t) \propto t^{\frac{2}{2(1+\omega)}}. \]  

(2.9)

This also makes clear that for \( t = 0 \) the scale factor also will be zero, confirming the above qualitatively derived singularity. Let us now define another very important quantity in cosmology, namely the conformal time

\[ d\tau = \frac{dt}{a(t)}. \]  

(2.10)

One of the reasons to introduce the conformal time is, that the trajectory of radial propagating massless particles can be represented as 45° lines in the \( \tau - r \) plane. Plugging eq. (2.9) into eq. (2.10) yields (for \( \omega \neq -1 \))

\[ \tau = a^{\frac{1}{2}(1+3\omega)} \quad \Rightarrow \quad a(\tau) = \tau^{\frac{2}{1+3\omega}}. \]  

(2.11)

For the case of a cosmological constant \( \omega = 1 \) we derive the important de-Sitter approximation

\[ \tau = \int dt e^{-Ht} = \frac{-1}{aH}. \]  

(2.12)

\(^1\)One may argue that this conclusion only holds if general relativity (GR) is valid for all energies. Since we lack a UV completion of GR until today, the Big Bang may reduce to the breakdown of GR.
We will make use of this approximation frequently, since inflation is defined as a stage of \( \omega \sim -1 \). Now we have all the essential tools together to understand exemplary ΛCDM problems.

**Horizon Problem**

With the above defined conformal time, it is naturally to define the co-moving particle horizon\(^2\)

\[
\chi = \int_{t_e}^{t} \frac{dt}{a(t)}. \tag{2.13}
\]

It measures the maximal distance a light ray can travel between the time it has been emitted \( t_e \) and the present time \( t \). We may rewrite the integral in terms of the co-moving Hubble radius \( 1/(aH) \)

\[
\chi = \int \frac{d \ln a}{(aH)}. \tag{2.14}
\]

What does this imply? As long as the strong energy condition is satisfied — which is true for the ΛCDM — the co-moving horizon grows monotonically with time. Thus, the fraction of causally connected parts of the Universe grows with time. To put this the other way around, we can say that the parts of the Universe which have been in causal contact with each other decrease when going back in time. In particular this means that co-moving scales entering the horizon today must have been outside the horizon at CMB decoupling. This, however, seems unnatural if we compare to the scale invariant measured homogeneity of the CMB \(^2\). Such a homogeneity only can be achieved if many causally disconnected patches of the CMB would have started with the same initial conditions \(^3\).

**Flatness Problem**

As mentioned before, the Universe is measured to be flat, i.e. \( k = 0 \). To see why this measurement is very puzzling within the framework of standard Big Bang cosmology, we will now consider the Friedmann equation (2.8) with a non-vanishing curvature. It then takes the form

\[
H^2 = H_0^2 \sum_i \Omega_i a^{-3(1+\omega_i)} + H_0^2 \Omega_k/a^2. \tag{2.15}
\]

The energy budget of the curvature is defined as \( \Omega_k = -k/(a_0 H_0)^2 \), which modifies also the consistency relation to \( \sum_i \Omega_i + \Omega_k = 1 \). We can divide eq. (2.15) by \( H^2 \) to bring it

\(^2\)It coincides with the conformal time only because we set \( c = 1 \). In SI units we would have \( \chi = c \tau \).

\(^3\)In fact, any two points on the CMB sky map with an angular separation bigger than \( \sim 2.3^\circ \) could not have been in causal contact in standard Big Bang cosmology. This means that the CMB consists of roughly \( 10^4 \) causally disconnected parts in this theory.
into the form
\[
\sum_i \frac{\rho_i \, a^{-3(1+\omega_i)}}{3H^2} - 1 \equiv \sum_i \frac{\Omega_i(a)}{(aH)^2} = 1 = \frac{k}{(aH)^2}.
\] (2.16)

Remember that the co-moving Hubble radius $1/(aH)$ increases with time in ΛCDM cosmology. Consequently the left hand side of the equation has to increase in the same manner. Thus, the measured value of \(\sum_i \Omega_i(a_0) \simeq 1\) requires enormous fine tuning at early times. To make this point clear, one can easily calculate back in time e.g. to the Big Bang Nucleosynthesis (BBN). BBN roughly took place at a temperature \(T = 100\) keV. Converting this into the scale factor at this time accordingly to \(a = T_{\gamma,0}/T\), with \(T_{\gamma,0} = 0.235\) meV denoting the Photon temperature today [45], we get \(a_{BBN} = 2.35 \times 10^{-9}\).

This implies the following deviation from flatness to guarantee the measured value of today 4
\[
\sum_i \Omega_i(a_{BBN}) - 1 \leq \mathcal{O}(10^{-10}).
\] (2.17)

Going further back in time leads to an even smaller deviation to account for the flatness measured today. This extreme fine tuning of initial conditions is very puzzling.

To emphasize it again, both mentioned problems are really only problems of fine tuned initial conditions and not a failure of the theory. However, this necessary extreme fine tuning is in some sense nonphysical. We saw that, both problems arise due to the strictly increasing co-moving Hubble horizon. An elegant way out may be given by the idea of modifying the behaviour of the Hubble horizon at early times. Instead of the monotonically increasing behaviour, one could add a decreasing phase. This simple modification does not only solve both problems, but can also explain e.g. the small CMB inhomogeneities [46] and the large scale structure formation and distribution [47].

Exactly this behaviour of the horizon defines cosmological inflation. We will now discuss how the horizon and flatness problem are solved within the framework of inflation.

**Horizon Problem:** If the co-moving Hubble horizon had a decreasing phase, causally disconnected regions could have been in contact at earlier times. Hence, the observed CMB homogeneity may be a consequence of causal physics at very early times.

**Flatness Problem:** If the co-moving Hubble horizon had a decreasing phase, the observed energy budget of \(\sum_i \Omega_i(a) = 1\) is an attractor solution of eq. (2.16).

Let us not turn to the question on how to achieve an inflationary phase. As pointed out above, the crucial ingredient is the decreasing Hubble horizon \((aH)^{-1}\). But what

\[\text{Here we included the measured curvature } \Omega_k = 0.0008 \text{ [3].}\]
exactly does this mean? Putting this in a mathematical way it reads
\[
\frac{d}{dt}\left(\frac{1}{aH}\right) = -\frac{\ddot{a}}{(aH)^2} \frac{1}{H} < 0. \tag{2.18}
\]

Since the co-moving Hubble horizon is strictly positive, the only way to guarantee this inequality is to require \(\ddot{a} > 0\). Hence, a decreasing Hubble horizon is equivalent to an accelerated expansion of the Universe. We can express this also as a time derivative of the Hubble parameter (since it already contains the first time derivative of the scale factor \(a\)) and get
\[
\frac{\ddot{a}}{a} = H^2 + \dot{H} = H^2(1 - \epsilon). \tag{2.19}
\]

Here, we introduced the slow-roll parameter \(\epsilon \equiv -\frac{\dot{H}}{H^2} = \frac{3}{2} \times (\omega + 1)\). This parameter is of enormous importance for inflation and the reader should keep it in mind. Especially the requirement for a inflationary phase to occur is given by \(\epsilon < 1\), c.f. eqs. (2.18) and (2.19). We can connect the \(\epsilon\) parameter to the inflationary e-fold number via \(dN \equiv -Hdt\). This yields \(\epsilon dN = d\ln H\). Note that \(N\) is counted backwards in our definition, s.t. when \(N_i < N_j\) the earlier time corresponds to \(N_i\).

But how to achieve an accelerated expansion? Recall that we discussed that the Friedmann equations lead to \(\ddot{a} < 0\), if the strong energy condition is fulfilled. Thus, to get the reverse behaviour of \(\ddot{a}\), we have to break the strong energy condition, which means
\[
p < -\frac{1}{3} \rho. \tag{2.20}
\]

In appendix B.2 we discuss the required time scale of the breaking of the strong energy condition. For the rest of the work we use the usual minimum time of \(N_{CMB} \sim 50 - 60\). At this time during inflation the observed CMB scales have exited the horizon. How this long lasting phase is achieved will be discussed in the next chapters, since the exact form of the equation of state is model dependent.
Chapter 3

Generic Pseudo
Nambu-Goldstone Boson Inflation

In this chapter we will discuss how inflation can be achieved with a generic pseudo
Nambu-Goldstone boson, referred to as the inflaton. This means that we leave the exact
potential arbitrary until we study the phenomenology. We will derive the equations
of motion and show how they can be solved analytically to good approximation. Our
focus however will not be the standard vacuum dynamic. Rather, we investigate the
phenomenological rich possibility of an additional derivative coupling of the inflaton to
$N$ Abelian gauge fields. We show how the dynamic gets strong modification due to the
presence of an CP-odd term in the Lagrangian. This leads to the well known result of
a tachyonic enhanced gauge field mode [48]. Its implication on stochastic gravitational
wave production and other phenomenological aspects will be studied. In particular, we
show how this paradigm can dynamically generate a chemical potential which in turn
leads to a deviation from the observed nearly perfect temperature black body spectrum
at high frequencies. These so called $\mu$ distortions are in general sensitive to the integrated
power spectra, both for the scalar and tensor. We will prove, however, that the tensor
mode is sub-dominant and can be safely neglected. Providing an accurate parameter
scan, we demonstrate how future experiments can tighten the parameter space for such
models. Thereby we show that the constraint may be stronger than the small non-
Gaussianity constraint, lifting the $\mu$ distortions to a stringent test which these models
have to pass.
3.1 Setup & Coupling to Abelian Gauge Fields

Let us start by considering the inflaton $\phi$ to be interacting with $\mathcal{N} U(1)$ gauge fields $A_\mu$ through a derivative coupling, i.e. Chern-Simons coupling. The gauge fields can be interpreted for a first analysis as dark photons, making the setup closed and independent from the Standard Model. Additionally, we consider the case of $\mathcal{N} = 1$, but a generalization is easy since it is only a constant re-scaling of the gauge fields. The interaction term with effective coupling $\alpha/f_c$ is then given by

$$L_{\text{int}} = -\frac{\alpha}{4f_c} \phi F_{\mu\nu} \tilde{F}^{\mu\nu},$$

where the electromagnetic field strength tensor is $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ with the four-potential $A^\mu \equiv (\chi, A)$ and its dual is defined as $\tilde{F}^{\mu\nu} \equiv 1/2\sqrt{-g}\epsilon^{\mu\nu\alpha\beta} F_{\alpha\beta}$. The appearing metric scalar $g$ is given by the determinant of $g_{\mu\nu}$. We work withing the framework of Einstein gravity and take the metric to be described by a flat FLRW metric $ds^2 = -dt^2 + a^2(t)dx^2 = a^2(\tau)(-d\tau^2 + dx^2)$. For $\epsilon$ we take the rank-4 Levi-Civita tensor with convention $\epsilon^{0123} \equiv +1 = -\epsilon_{0123}$. The total action is given by

$$S = \int d^4x \sqrt{-g} \left( \frac{R}{2} - \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{\alpha}{4f_c} \phi F_{\mu\nu} \tilde{F}^{\mu\nu} - V(\phi) \right) \equiv S_{EH} + S_{L}.$$

For now, the potential $V(\phi)$ is arbitrary, but sufficiently flat to successfully support inflation. It shall be induced by some additional symmetry breaking process leading to a pNGB. During this phase the potential has to guarantee the derived constraint

$$\epsilon \equiv \frac{3}{2}(\omega_\phi + 1) \simeq \frac{1}{\epsilon} \left( \frac{\partial_\mu V}{V} \right)^2 \equiv \epsilon_V < 1.$$

To obtain the EOMs, we basically only have to use the Euler-Lagrange formalism in an expanding background, since the action does not contain higher derivatives. This leads for the gauge fields to

$$\partial_\mu F^{\mu\nu} + \frac{\alpha}{f_c} \partial_\mu \left( \phi \sqrt{-g} \tilde{F}^{\mu\nu} \right) + \partial^\rho \partial_\mu A^\mu = 0.$$
But note that the action exhibits an gauge symmetry of the form \( A_\mu \to A_\mu + \partial_\mu \lambda, \lambda \in \mathbb{R} \). This enables us to fix a gauge \( A^0 = 0 \) and set \( \nabla \cdot \mathbf{A} = 0 \), leading to

\[
\mathbf{A}'' - \nabla^2 \mathbf{A} - \frac{\alpha}{f_c} \phi' \nabla \times \mathbf{A} = 0.
\]  

(3.6)

Here and in the following we denote \( \partial_\tau x \equiv x' \) as the derivative with respect to (w.r.t.) the conformal time \( \tau \). The derivative w.r.t. the cosmic time \( t \) is denoted by \( \partial_t x \equiv \dot{x} \). The EOM for the inflaton is trivially given by

\[
\ddot{\phi} - \nabla^2 \phi + 3H \dot{\phi} + \partial_\phi V(\phi) = -\frac{\alpha \sqrt{-g}}{4f_c} F_{\mu \nu} \tilde{F}^{\mu \nu}.
\]  

(3.7)

We can compute the scalar product between the field strength tensor and its (Hodge) dual in terms of the physical electric and magnetic fields

\[
\mathbf{B} = \frac{1}{a^2(\tau)} \nabla \times \mathbf{A}(\tau, \mathbf{x}), \quad \mathbf{E} = -\frac{1}{a^2(\tau)} \mathbf{A}'(\tau, \mathbf{x}).
\]  

(3.8)

So we arrive at the inflaton equation

\[
\ddot{\phi} - \nabla^2 \phi + 3H \dot{\phi} + \partial_\phi V(\phi) = \frac{\alpha}{f_c} \mathbf{E} \cdot \mathbf{B},
\]  

(3.9)

The appearing Hubble parameter can be obtained from the first Friedmann equation. Note, that the second Friedmann equation is – contrary to the first – of differential form for the Hubble parameter. This ’dynamic’ Hubble equation however has basically no other impact than the ’non-dynamic’ one in our particular case, since inflation is described by a nearly de-Sitter space-time, meaning an approximate constant Hubble parameter. The first Friedmann equation is, as shown in the previous chapter, given by \( 3H^2 = \rho \). The zero-zero component of the energy momentum tensor defines the energy density \( \rho \) and is obtained from the particle physics dynamics

\[
T_{\mu \nu} \equiv -\frac{2}{\sqrt{-g}} \frac{\delta S_L}{\delta g^{\mu \nu}}
\]  

(3.10)
It splits up into the part arising from the inflaton and from the electromagnetic stress-energy. They are respectively given by\(^1\)

\[
T_{\mu\nu}^\phi = \frac{2}{\sqrt{-g}} \frac{\delta}{\delta g^{\mu\nu}} \left[ \sqrt{-g} \left( \frac{1}{2} \partial_\alpha \phi g^{\alpha\beta} \partial_\beta \phi + V(\phi) \right) \right] = \partial_\mu \phi \partial_\nu \phi - g_{\mu\nu} \left( \frac{1}{2} \partial^\alpha \phi \partial_\alpha \phi + V(\phi) \right),
\]

\[
T_{\mu\nu}^{\text{em}} = \frac{2}{\sqrt{-g}} \frac{\delta}{\delta g^{\mu\nu}} \left[ \frac{1}{4} \sqrt{-g} F_{\alpha\beta} F^{\alpha\beta} \right] = F_{\mu\alpha} F_{\nu}^{\alpha} - \frac{1}{4} g_{\mu\nu} F^{\alpha\beta} F_{\alpha\beta}.
\]

Note that the inflaton–gauge field interaction term in the Lagrangian does not contribute to the energy momentum tensor, since the term decouples from gravity by definition of the dual field strength tensor. Thus, for the zero-zero component of the full energy momentum tensor we obtain

\[
T_{00} = \frac{1}{2} \dot{\phi}^2 + \frac{1}{2} \nabla^2 \phi + V(\phi) + \frac{1}{2} E^2 + B^2.
\] (3.11)

Hence, the set of coupled EOMs which describes the system of eq. (3.2) is

\[
0 = A'' - \nabla^2 A - \frac{\alpha}{f_c} \phi' \nabla \times A,
\] (3.12)

\[
\frac{\alpha}{f_c} \langle E \cdot B \rangle = \ddot{\phi} + 3H \dot{\phi} + \partial_\phi V(\phi),
\] (3.13)

\[
H^2 = \frac{1}{3} \left( \frac{1}{2} \dot{\phi}^2 + V(\phi) + \frac{1}{2} (E^2 + B^2) \right),
\] (3.14)

where we consider the inflaton to be homogeneous \(\phi = \phi(t)\). The expectation value \(\langle ... \rangle\) for the physical fields encodes the back-reaction of the produced gauge quanta on the homogeneous dynamics of \(\phi(t)\) and \(a(t)\) [29]. This will become important, because we may use the slow roll limit [49] to decouple the EOMs, as we will show in the following.

Before we proceed, we may consider the inflaton vacuum motion, i.e. setting \(\alpha/f_c = 0\). In this case, the EOMs trivially decouple. In the typical slow roll approximation [49], where we say that the vacuum energy dominates over the inflaton kinetic energy and the inflaton acceleration may be safely neglected \(|\ddot{\phi}| \ll H|\dot{\phi}|, |\partial_\phi V|\), we can even solve the inflaton evolution analytically. We show this explicitly in appendix C.3 for a fixed potential for illustrative purpose. This will be extremely helpful later on, since this roughly sets the initial conditions for the full set of coupled EOMs, in particular when \(\alpha/f_c \neq 0\). This is because upon expecting eq. (3.13) we may interpret the inflaton–gauge field interaction as an additional friction for the inflaton evolution.

\(^1\)One may use the relation \(2\sqrt{-g} \delta = \sqrt{-g} g^{\mu\nu} \delta g_{\mu\nu}\).
Therefore we investigate the gauge field evolution and demonstrate that it has only significant impact on the dynamic towards the end of inflation. This makes the slow roll approximated initial conditions well justified.

Decomposing the gauge field $A(\tau, x)$ into its Fourier modes leads to

$$\left( \partial_{\tau}^2 + k^2 \pm \frac{2k \xi}{-\tau} \right) A_\pm(\tau, k) = 0. \quad (3.15)$$

Note, in (quasi) de-Sitter space we have $\tau \in (-\infty, 0]$ and thus $-\tau \in \mathbb{R}^+$. Here, we introduced the (by definition positive) parameter

$$\xi \equiv \frac{\alpha |\dot{\phi}|}{2f_c H}. \quad (3.16)$$

In appendix C.1 we derive this EOM in more detail.

A qualitative inspection of the equation reveals that the mode $A_-$ experiences a tachyonic instability, if we assume without loss of generality $\dot{\phi} < 0$ (and $\phi > 0, \partial_\phi V(\phi) > 0$). This must lead to an exponential growth of low momentum modes triggered by the mass term $m^2 = k^2 - 2k \xi / (-\tau) < 0$. This reveals the importance of the $\xi$ parameter as it controls the gauge field growth. In the slow roll regime ($\xi \sim \text{const.}$) this equation decouples from the coupled set of the EOM. We then can bring it into the form of a radial Schrödinger equation in a Coulomb potential. This allows an analytic solution in terms of Whittaker functions. Requiring the gauge field to be initially in the adiabatic vacuum (Bunch-Davies vacuum), the boundary condition is set to be of the form $A_\lambda(\tau, k) = e^{-ik\tau}/\sqrt{2k}$ for $k\tau \to -\infty$. We rigorously derive the analytic solution in appendix C.1 and only quote the final result for clarity

$$A_-(\tau, k) = \frac{1}{\sqrt{2k}} e^{\pi \xi/2} W_{-\xi, 1/2}(2ik\tau). \quad (3.17)$$

This enables us to simplify the EOM for the inflaton enormously. Ignoring the polarization mode $A_+$ we get with eq. (3.8)

$$\langle E \cdot B \rangle = -\frac{1}{4\pi^2 a^4} \int dk k^3 |A_-|^2, \quad (3.18)$$

$$\frac{1}{2} \langle E^2 + B^2 \rangle = \frac{1}{4\pi^2 a^4} \int dk k^2 \left( |A_-|^2 + k^2 |A_-|^2 \right). \quad (3.19)$$
Using the analytic expression for the $A_-$ mode, we can obtain [29]

\[
\langle E \cdot B \rangle \simeq 2.4 \times 10^{-4} \frac{H^4}{\xi^4} e^{2\pi \xi} \equiv H^4 f_1(\xi),
\]

\[
\frac{1}{2} (E^2 + B^2) \simeq 1.4 \times 10^{-4} \frac{H^4}{\xi^3} e^{2\pi \xi} \equiv H^4 f_2(\xi).
\]

We will work with this expression in the rest of the text and also used them for numerical calculations. That is why we show the derivation and discuss the validation of these approximations explicitly in appendix C.2. Important for the reader, however, is only that they approximate the setup for $\xi \gtrsim 3$ very good.

To simplify the numerical calculation, it is convenient to change from (conformal) time to e-folds. Recall that the relation is given by $dN = -H dt$. Then, the remaining equations read

\[
\frac{\alpha}{f_c} f_1(\xi) H^2 = \partial_N^2 \phi - \partial_N \phi \left( 3 - \frac{\partial_N H}{H} \right) + \frac{1}{H^2} \partial_\phi V(\phi),
\]

\[
H = \left( q(\xi) - \sqrt{q(\xi)^2 - V(\phi) / f_2(\xi)} \right)^{1/2},
\]

where we defined

\[
q(\xi) \equiv \frac{6 - (\partial_N \phi)^2}{4 f_2(\xi)}.
\]

These EOMs lay the foundation for our numerical analysis and the investigation of the phenomenology, which we will discuss in the next section.

### 3.2 Numerical Treatment & Phenomenology

In this section we will show how to solve the equation for $\phi$ and thus for $H$. Since these equations only can be solved numerically one has to provide initial conditions. The inhomogeneity acts additional to $H$ like a friction term, which is dominant for large $\xi$ but can be neglected for small $\xi$. Thus, the solution of $\phi$ should not differ (significantly) from the homogeneous slow-roll approximation $\phi_{sr}$ at early stages of inflation, in particular at $N_{\text{CMB}}$. So the early time slow-roll solution sets the initial condition for the full EOM.

We may emphasize our viewpoint as follows:
1. The homogeneous equation

Consider first the homogeneous equation for $\phi_{sr}$ in the slow-roll limit, i.e. set $\partial_N^2 \phi_{sr} = 0$ and $3H^2 = V$. Once the potential is fixed, this equation may be solved analytically, where the violation of the slow-roll condition $\epsilon = 1$ fixes the integration constant. The end of inflation can be defined to happen at $N = 0$.

2. Solving the full equation

The full equation for $\phi$ now can be solved with the help of the $\phi_{sr}$. The initial conditions for $\phi$ can be extracted as

$$\phi(N_{CMB}) = \phi_{sr}(N_{CMB}) \quad \wedge \quad \partial_N \phi(N_{CMB}) = \partial_N \phi_{sr}(N_{CMB}). \quad (3.25)$$

Solve until $\epsilon = 1$. This will always happen for $N_{\text{end}} < 0$ if $\alpha/f_c > 0$.

3. Shift of the solution

We like to set the end of inflation to be at $N = 0$. We thus shift the solution like $\phi(N) \rightarrow \phi(N - N_{\text{end}})$. The same shift may also be applied to $\phi_{sr}$ to depict both functions in one plot.

To this end we have the fix the potential. We study the dynamic in the case of a Starobinsky potential and a quadratic potential, which is the first order approximation of an axion-like potential. They are respectively given by

$$V_s(\phi) = V_{0,s} \left( 1 - e^{-\gamma \phi} \right)^2, \quad (3.26)$$
$$V_q(\phi) = V_{0,q} \phi^2. \quad (3.27)$$

The constants $V_{0,(s,q)}$ and $\gamma$ should be chosen s.t. the solutions do not violate any constraints from the CMB measurements [3], which we will discuss later on. The exact values we used for them are given in the caption of fig. 3.1 where we show the solution for $\phi$. We can clearly see that the Starobinsky potential induces a much higher friction term than the quadratic potential. This is because the shape of the potential is much flatter. In figure 3.2 we depict the $\xi$ parameter as well the Hubble parameter for the Starobinsky potential. Note that super-Planckian field values do not point in the direction of a quantum gravity, i.e. the breakdown of Einstein gravity. In fact, this is only reached if the energy density exceeds the Planck boundary and not the field alone [19].
Pseudo Nambu-Goldstone Boson Inflation

Figure 3.1: Solution for the inflationary field $\phi$ in axion inflation and for $\phi_{sr}$ in the homogeneous slow-roll approximation. In the left panel we depict the evolution with the Starobinsky potential with constants $V_{0,s} = 1.52 \times 10^{-9}, \gamma = 0.3, \alpha/f_c = 75$. And in the right panel we took $V_{0,q} = 2.65 \times 10^{-11}, \alpha/f_c = 35$ for the quadratic potential. In both figures we give the numerical result for the additional e-folds of the inflationary phase due to the gauge field friction $\Delta N$.

Figure 3.2: **Left:** Evolution of the $\xi$ parameter. The effective gauge field friction induces an upper bound contrary to the standard slow roll case. **Right:** The Hubble parameter and its simplest slow roll approximation.

Trivially this is not the case in our models due to the high suppression factor arising from the potential.

Let us now turn to the constraints the CMB puts on these models of pNGB inflation. We list the most important ones

1. **COBE Normalization:** $\Delta^2_{\xi} = (2.142 \pm 0.049) \times 10^{-9}$ [3].
2. **Spectral Index:** $n_s \equiv \frac{d \ln \Delta^2_{\xi}}{d \ln k} + 1 = 0.9667 \pm 0.0040$ [3].
3. **Running of the Spectral Index:**
   - $\alpha_s \equiv \frac{dn_s}{d \ln k} = -0.0057 \pm 0.0071$. 68 % CL [5].
   - Or: $\alpha_s \equiv \frac{dn_s}{d \ln k} = -0.002 \pm 0.013$. 95 % CL [3]
4. Small non-Gaussianity: $f_{NL}^{\text{equil}} < 39 \implies \xi \lesssim 2.5$ [4, 29]. See the following main text for details of how $f_{NL}^{\text{equil}}$ constraints $\xi$.

Note that all the constraints should be fulfilled at horizon crossing, i.e. evaluated at $N_{\text{CMB}}$. The CMB measurements provide very tight constraints to the physics of the inflationary phase. Any model should obey these measurements. Clearly we now need a conversion to switch from the e-folds $N$ to the co-moving wave number $k$

$$k(N) = \frac{2\pi k_*}{0.002 \text{ Mpc}^{-1}} 10^2 \text{ Hz} c e^{N_{\text{CMB}}-44.9} e^{-N},$$  

(3.28)

which we derive in appendix B.2. The pivot scale of the Planck 2015 results is $k_* = 0.05 \text{ Mpc}^{-1}$. Let us now calculate the constraint quantities from the model which we are studying.

To get the scalar power spectrum it is not sufficient to rely on the former homogeneous approximation $\phi = \phi(t)$. Rather, we split the field into a homogeneous part and a fluctuation, i.e. $\phi(t, x) = \phi(t) + \delta \phi(t, x)$. A common approach is then to choose the spatially-flat gauge $\zeta = H/|\dot{\phi}| \delta \phi$. With this ansatz one can obtain [27]

$$\Delta_\zeta^2 = \Delta_{\zeta,\text{vac}}^2 + \Delta_{\zeta,\text{axion}}^2 = \left(\frac{H}{2\pi \partial_N \phi}\right)^2 + \left(\frac{\alpha H^2 f_1(\xi)}{3 f_c \partial_N \phi \left(1 + 2\pi \xi \alpha H^2 f_1(\xi) \frac{\partial N \phi}{\dot{\phi} N \phi}\right)}\right)^2,$$

(3.29)

where we used the definition of $f_1(\xi)$ as in eq. (3.20). The tensor power spectrum can be obtained similarly by looking at the metric perturbation and one gets after adding up both polarization modes [29]

$$\Delta_t^2 = \Delta_{t,\text{vac}}^2 + \Delta_{t,\text{axion}}^2 = 2 \times \frac{H^2}{\pi^2} \left(1 + 4.3 \times 10^{-7} H^2 \frac{e^{4n \xi}}{\xi^6}\right).$$

(3.30)

In figure 3.3 we depict both power spectra as a function of $N$. Also we indicate the important range for our application to calculate CMB distortions arising through energy release at this range, see next section.

Equipped with the formula for the power spectra – in particular for $\Delta_\zeta$ which is basically the two point correlation function of the density perturbation – one can calculate the corresponding three point correlation function. This quantity directly
Pseudo Nambu-Goldstone Boson Inflation

\[ f_{\text{NL}}^{\text{equil}} = 7.4 \times 10^{-8} \frac{\Delta_5^{6} \zeta_{\text{vac}} \xi^{6\pi\xi}}{\Delta_\xi^2 \xi^{8.1}}. \]  

Figure 3.3: We depict the scalar- and tensor power spectrum, respectively. For the Starobinsky- and quadratic potential we used the constants as given in the main text. All CMB constraints are fulfilled. The dashed lines in the same color correspond to the respective vacuum contribution to the power spectrum. Additionally, we indicate in both plots with a vertical dashed line the region of interest for investigating \( \mu \)-type distortions, see next chapter for more details.

relates to the bispectrum, which in turn defines the non-linearity parameter \( f_{\text{NL}} \) as [29]

We plot this quantity as a function of \( \xi \) along the COBE Normalization curve in figure 3.4. Solving the equation for the maximal allowed non-Gaussianity, i.e. \( f_{\text{NL}}^{\text{equil}} = 39 \), we obtain \( \xi \sim 2.5 \). We note that the curve \( f_{\text{NL}}^{\text{equil}}(\xi) \) increases extremely fast in this regime s.t. improvements in the measurement have no net effect. Although the conversion between \( f_{\text{NL}}^{\text{equil}} \) and \( \xi \) is model dependent — since \( f_{\text{NL}}^{\text{equil}} \) is directly dependent on \( \partial N \phi \) — we checked that we get the same result for both potentials we study.

It is well known that it is a very tight constraint that the primordial non-Gaussianity has to be small. Especially many non-standard inflationary models fail to obey the constraint set by \( f_{\text{NL}} \). However, another constraint coming from the anisotropic heating due to inflation can lead to observable \( \mu \)-type distortions. This constraint applies to the both the scalar and tensor power spectrum and it is sensitive to the integrated spectra. Usually neglected in literature, we will investigate the strength of this constraint in detail in the next chapter. We will provide the first ever calculation of the \( \mu \)-type distortions due to the tensor modes. We then may compare it to the other constraints, especially to \( f_{\text{NL}} \).
3.3 CMB Distortion as a New Strong Constraint

We study the application of the power spectra in the case of CMB distortions. We will investigate how measured constraints on the chemical potential $\mu$ may tighten the allowed parameter space for our considered inflationary model. What is behind $\mu$ distortions, is that the CMB may have tiny deviations from a perfect black body. Still assuming that the Universe started from a pure black body distribution, a phase of (quasi) exponential expansion – as it happens in inflation – leads to a superposition of (many) different black bodies. Each patch for itself is now a pure black body with a given temperature. But it is easy to see that a superposition of black bodies at different temperatures is not itself a black body. Due to usual Thomson scattering the mixing of the patches sources an average distortion in the correlation function between Compton-$y$ and the CMB temperature $C_{l}^{T,y}$. This in turn evolves into the $\mu$ distortions via Compton scattering. These distortions are then nothing else but the integrated anisotropic heating rate – caused by the damping of primordial temperature fluctuations – weighted with a distortion visibility function capturing the thermodynamics.
We investigate the $\mu$ distortion with the use of a so called Window function. This has the advantage that it is model independent. Practically, it separates the dissipation and thermalization physics from the model dependent power spectra. For clarity we show the window function only in the appendix A. Of importance here is only the fact that the window function for scalar $W_\zeta$ and tensor $W_t$ distortion may be expressed to good approximation analytically $[33, 34]$. The $\mu$ distortions are then calculated as

$$\mu_i = \int_0^\infty d\ln(k) \Delta_i^2(k) W_i(k),$$

where $i = \zeta, t$ represent the scalar and tensor contributions. In figure 3.5 we depict both window functions and $d\mu_i$ for both studied potentials. This figure clearly indicates that we have to integrate over a much bigger time interval to obtain the tensor contributions to the $\mu$–type distortions as it is for the scalar case. With the model constants as before we find for the Starobinsky and Quadratic potential respectively

$$\mu_\zeta^s = 3.5 \times 10^{-6}, \quad \mu_t^s = 1.6 \times 10^{-12}, \quad \mu_\zeta^q = 2 \times 10^{-8}, \quad \mu_t^q = 3 \times 10^{-14}.$$  

The current constraint from the COBE/FIRAS measurement is clearly fulfilled, i.e. $\mu \leq 9 \times 10^{-5}$ $[45]$. The tensor contribution is in both cases suppressed by six orders of magnitude. Thus, it is a valid approximation to only consider the scalar contribution. If we compare this result now with the constraint on the small non-Gaussianity, we find that the latter one constraints the parameter space for allowed potential configurations stronger than the $\mu$ constraint. However, this situation changes if future experiments like
PIXIE [50] provide more accurate measurement for $\mu$. The sensitivity to detect $\mu$-type distortions will then nearly reach the vacuum contribution, constraining it to $\mu \leq 5 \times 10^{-8}$ at $5\sigma$ confidence level [50]. Adopting this constraint, the before used constants for the Starobinsky potential would be excluded. Therefore, the PIXIE constraint tightens the parameter space more than the non-Gaussianity constraint. We also can provide an exemplary parameter point for the Starobinsky potential fulfilling this new PIXIE constraint: $V_0 = 1.0 \times 10^{-9}, \gamma = 0.3, \alpha/f_c = 57$. This implies $\mu = 1.5 \times 10^{-8}$ but only $\xi = 1.6$. In figure 3.6 we show a detailed phase space scan and plot the final result in the $\mu - \xi$ plane. We restrict our scan to the Starobinsky potential. The main result of this plot is that it reveals the full power of the new PIXIE constraint compared to the FIRAS and the non-Gaussianity constraint. We filtered the results to fulfill the quoted Planck constraints [3].

This should be regarded as an additional motivation for future experiments to measure $\mu$-type distortions with higher accuracy. The insight that these measurements could give, goes far beyond the insights (realistic) improvements on the measurement of the non-Gaussianity could provide. To understand this argument, we may consider

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig3_6.png}
\caption{We show for the Starobinsky potential points in the $\mu$-$\xi$ plane which fulfill $\Delta_c, n_s$ and $\alpha_s$ [3]. Only $\sim 3.6\%$ of the studied points survived these bounds. Additional $\sim 21.5\%$ of the points could be excluded with the PIXIE constraint, whereas the non-Gaussianity only rule out $\sim 0.5\%$. The colored horizontal lines indicate the FIRAS (red) and PIXIE (blue) limit on $\mu$. The black line represents the limit on small non-Gaussianity.}
\end{figure}
an hypothetical improvement of the $f_{\text{NL}}^{\text{equil}}$ parameter by a factor of ten, i.e. $f_{\text{NL}}^{\text{equil}} < 39 \rightarrow f_{\text{NL}}^{\text{equil}} < 4$. Along the COBE normalization curve, this would mean that this very good improvement on $f_{\text{NL}}^{\text{equil}}$ not really touches the $\xi$ constraint. Rather, we would have $\xi < 2.5 \rightarrow \xi < 2.37$. The question now should be, by how much one would have to optimize the $f_{\text{NL}}^{\text{equil}}$ measurement to catch up with the PIXIE constraint. For the above found parameter point, we have $\xi = 1.6$, implying $f_{\text{NL}}^{\text{equil}} \leq 10^{-5}$. Thus, only an improvement of six orders of magnitude for the non-Gaussianity can catch up with the PIXIE constraint. But this is also what we found with the help of figure 3.6. So, with future experiments the widely unnoticed $\mu$ constraint should be considered as a strong bound general inflationary models have to pass, not only pNGB models.
Chapter 4

Natural and Chromo-Natural Inflation

In this chapter we study the natural interpretation that the pNGB inflaton is an axion-like particle. We focus on a model where the inflaton is coupled to non-Abelian gauge fields through a Chern-Simons term. The simplest realization of such a setup is with an $SU(2)$ gauge group, called chromo-natural inflation. We circumvent the highly complex non-linear dynamics by restricting ourselves to the linearized regime. We discuss the validity of such an approximation with a detailed parameter scan. This reveals that models with realistic gauge coupling near the GUT scale do not undergo an inherent non-Abelian evolution which is calculable within the linearized approach. We also study possible phenomenological implications. In particular we show that there is only a small window for sufficient gravitational wave production which can be assigned to the $SU(2)$ dynamics. Most notably we find that the magnetic drift regime [24, 36–40] does not fall in this window. This is because we track the gauge field evolution from the quantum mechanical Bunch-Davies vacuum on, contrary to what is frequently done in literature [24, 36–40]. We show how fluctuations trigger a non-zero vacuum expectation value (vev) for the background field only at late times. In this spirit we show that chromo-natural inflation requires a super-Planckian periodicity for the potential to be in agreement with Planck data [1]. We try to recover the setup by choosing a sub-Planckian EFT cut-off scale. However, we find that for nearly the complete parameter space the de-Sitter Gibbons-Hawking temperature is too high to break the global Peccei & Quinn symmetry. Further requiring the normal mass hierarchy, only a tiny window
survives these theoretical constraints. We additionally discuss restriction within this window coming from string theoretical model building.

4.1 Framework

The general idea will be the same as in the abelian case and it will be important to keep the results in mind. The action thus modifies to

$$S = \int d^4 x \sqrt{-g} \left( \frac{R}{2} - \frac{1}{2} \partial \mu \phi \partial^\mu \phi - V(\phi) - \frac{1}{4} F_{\mu \nu}^a F_{\mu \nu}^a - \frac{\alpha}{4 f_c} \phi F_{\mu \nu}^a \tilde{F}_{\mu \nu}^a \right),$$

(4.1)\equiv S_{EH} + S_\phi + S_{YM} + S_{CS}.

(4.2)

where the field strength tensor encodes the three $SU(2)$ gauge fields and their coupling $g$ through the definition $F_{\mu \nu}^a = \partial_\mu A_\nu^a(\tau, x) - \partial_\nu A_\mu^a(\tau, x) + g \epsilon^{abc} A_\mu^b(\tau, x) A_\nu^c(\tau, x)$. The dual is defined as before. Varying the action leads to the following coupled EOMs in case of a homogeneous inflaton field

$$\ddot{\phi} + 3 H \dot{\phi} + \partial_\phi V = - \frac{\alpha}{4 f_c} F_{\mu \nu}^a \tilde{F}_{\mu \nu}^a,$$

(4.3)

where $\epsilon_{\mu \nu \rho \sigma}$ encodes the three $SU(2)$ gauge fields and their coupling $g$ through the definition $F_{\mu \nu}^a = \partial_\mu A_\nu^a(\tau, x) - \partial_\nu A_\mu^a(\tau, x) + g \epsilon^{abc} A_\mu^b(\tau, x) A_\nu^c(\tau, x)$. The dual is defined as before. Varying the action leads to the following coupled EOMs in case of a homogeneous inflaton field

$$\ddot{\phi} + 3 H \dot{\phi} + \partial_\phi V = - \frac{\alpha}{4 f_c} F_{\mu \nu}^a \tilde{F}_{\mu \nu}^a,$$

(4.3)

Evidently, these non-linear EOMs cannot be solved easily. In fact they are so strongly coupled that the dynamics is very sensitive to the gauge field background. For a clear investigation of this system one could for example use lattice simulations. However, the exponentially expanding background is only one of many difficulties of the method [51]. This very challenging task has to be postponed to future work, since it goes far beyond the scope of this work. To study the system anyway, one may expand the gauge fields around an isotropic and homogeneous background like

$$A(\tau, x) = A(0)(\tau) + \delta A(\tau, x).$$

(4.5)

This linearization of course sets limits to the analysis. It is only valid as long as the classical background dominates over the quantum fluctuation. In the following two sub-sections we will discuss the EOM for the homogeneous background as well as the gauge
field fluctuations.

### 4.1.1 Homogeneous Background

Any homogeneous and isotropic $SU(2)$ gauge field can be transformed into the form

\[ A^a_0(\tau) = 0, \quad A^a_i(\tau) = \delta^a_i f(\tau). \]  

(4.6)

Inserting this form for the background field into the EOM (4.4), one obtains

\[ \frac{d^2}{d\tau^2} g f(\tau) + 2(g f(\tau))^3 - \frac{2\xi}{-\tau} (gf(\tau))^2 = 0, \]  

(4.7)

where, remember, $\xi$ is a function of $d\phi$. This ODE comes with some symmetries which facilitate the analysis. One can observe, that the coupling $g$ acts as a scale factor for $f(\tau)$. Additionally, the transformation $gf(\tau) \rightarrow -gf(\tau)$ and $\xi \rightarrow -\xi$ deliver the same solution modulo $\pm$. Related to the first observation, the ODE generically can be expressed as an autonomous ODE, meaning that the dependent variable does not explicitly appear. It was shown in ref. [52], that there exist three distinct type of solutions with properties:

1. For the $c_0$ - type solution the function remains bounded and the physical background field approaches zero. At all times it is sub-dominant to the (tachyonic) fluctuations.

2. For the $c_2$ - type solution the function grows proportional to $(-\tau)^{-1}$.

3. Only for finely tuned initial conditions the $c_1$ - type solution originates, but then also behave qualitatively like the $c_2$ - type solution. However, once fluctuations are present, this solution cannot be stable.

For our analysis it will be important to know for which initial condition which solution will emerge. We can only proceed with $c_{1,2}$ - type solutions, since only for them there is a given chance that the background dominates over the fluctuations. For the $c_0$ - type solution this cannot be the case and thus it is a priori impractical for our analysis. It was demonstrated in ref. [52], that for large $\xi$, i.e. $\xi \geq 2$, the $c_2$ - type solution has an overwhelming higher probability to be present than the other
two. The explicit late time solution in the slow roll limit of constant $\xi$ is given by

$$gf(\tau) = \frac{c_1 \xi}{-\tau},$$

with the three different types of solutions (for completeness)

$$c_0 = 0, \quad c_1 = \frac{1}{2} \left(1 - \sqrt{1 - (2/\xi)^2}\right), \quad c_2 = \frac{1}{2} \left(1 + \sqrt{1 - (2/\xi)^2}\right).$$

In order to have $f(\tau) \in \mathbb{R}$ we see that the $c_2$ - type solution can only appear for $\xi \geq 2$. Thus, when this solution is allowed, it is immediately more likely than the others. Inserting the $c_2$ - type solution back into the inflation EOM (4.3), one gets

$$\frac{\partial^2 N}{\partial \phi} - 3 \frac{\partial N}{\partial \phi} \left(1 - \frac{\partial N}{3H}\right) + \frac{\partial_y V(\phi)}{H^2} - \frac{3\alpha}{f_c g^2} H^2 (c_2 \xi)^3 = 0.$$  

(4.10)

The Hubble parameter is given by [37]

$$3H^2 = \frac{3}{2g^2} (H^2 c_2 \xi - H \partial N (H c_2 \xi))^2 + \frac{3}{2g^2} (H c_2 \xi)^4 + \frac{1}{2} H^2 (\partial N \phi)^2 + V(\phi).$$

(4.11)

We found that it can be well approximated by neglecting the the second term in the first bracket and it is also not very different from the vacuum Hubble parameter

$$H \simeq \sqrt{\frac{3 - 1/2(\partial N \phi)^2}{2s(N)}} - \sqrt{\left(\frac{3 - 1/2(\partial N \phi)^2}{2s(N)}\right)^2 - \frac{V(\phi)}{s(N)}} \simeq \sqrt{\frac{V(\phi)}{3 - 1/2(\partial N \phi)^2}},$$

(4.12)

where we defined

$$s(N) = \frac{3c_2^2 \xi^2}{2g^2} \left(1 + c_2^2 \xi^2\right).$$

(4.13)

The at the beginning mentioned magnetic drift regime is reached, when the gauge field induced friction-like term is bigger than the Hubble friction, i.e. $\alpha \xi H/(g f_c) \gg 1$ [37].

In particular we found that the changeover happens at $\alpha / f_c \sim \sqrt{g/H}$. Our analysis, however, will be universal and not restricted to this case.

Therefore, we will now also show the complete coupled set of EOMs which we used to verify the analytic solution for the gauge field background and thus eq. (4.10). For numerical reasons we express the background as $f(\tau) \equiv a(t) \gamma(t)$, as it was originally
proposed in [36]. Then we have:

\[
\frac{\partial^2 N}{\partial N^2} - 3 \frac{\partial N}{\partial N} \left( 1 - \frac{\partial N H}{3 H} \right) + \frac{\partial \phi V(\phi)}{H^2} = - \frac{3 g \alpha}{f_c} \gamma^2 \left( \frac{\gamma}{H} - \frac{\partial N \gamma}{H} \right),
\]

(4.14)

\[
\frac{\partial^2 \gamma}{\partial N^2} - 3 \frac{\partial N}{\partial N} \left( 1 - \frac{\partial N H}{3 H} \right) + \gamma \left( 2 - \frac{\partial N H}{H} \right) + 2 \frac{g}{f_c} t^2 \gamma^3 = - \frac{g \alpha}{f_c} \gamma^2 \left( \frac{\partial N \phi}{H} \right),
\]

(4.15)

\[
\frac{3}{2} H \left( \gamma - \partial N \gamma \right)^2 + \frac{3}{2} \frac{g^2 \gamma^4}{H^2} + V(\phi) = 3 H^2.
\]

(4.16)

We could confirm that for all parameter points we study, the deviation from eqs. (4.8) and (4.10) is less than a few percents, see therefore section 4.2 and especially appendix D.

Thus, we will work with the analytic solution for \( f(\tau) \) in the following. However, we restore the time evolution of \( \xi \) in the analytic solution. Although this might seem inconsistent, since a constant \( \xi \) is assumed to arrive at the analytic solution, this is what we find to restore the full dynamics of eqs. (4.14)-(4.16) in the best way.

### 4.1.2 Gauge Field Fluctuations

Generically, fluctuations are obtained by expanding around a homogeneous part. So, the basis for the analysis is the gauge field decomposition

\[
A_i^a(\tau, x) = f(\tau) \delta_i^a \phi + \delta A_i^a(\tau, x).
\]

(4.17)

To obtain the linearized EOM for the fluctuations, the action has to be varied quadratically in all field fluctuations working within the ADM formalism [53]. The mathematical prescription is very lengthy and not of importance in this work and thus will be skipped. We will only quote the final result for the physical linear EOM in the gauge field fluctuation [52]

\[
- \frac{\partial^2}{\partial N^2} \delta A_i^a - 2 g e^{a b i} f^b \delta \phi + \frac{\partial^2}{\partial N^2} \delta A_i^a - \partial_i \partial_j \delta A_j^a - g^2 f^2 \left( 2 \frac{\partial^2}{\partial N^2} A_k^a + \delta A_i^a - \delta A_k^a \right)
\]

\[
+ 2 g f e^{a b c} \partial_b \partial A_i^c + g f e^{a b c} \partial_i \partial A_b^c + g f e^{a b i} \delta \partial A_i^a + \frac{\alpha}{f_c} \left( f_l e^{j i j} \partial_j \partial \phi + g f^2 \delta^a \delta \phi \right)
\]

\[
+ \alpha \left( f^l e^{j k l} \partial_j \partial \phi + g f^2 \delta^a \delta \phi \right) - f^l \partial_i N^a + f^l \delta^a \partial_i N^l - g e^{a i l} \left( 3 f f^l N^a + f^2 N^{a l} \right)
\]

\[
+ \delta^a \left( f^l \partial_i N + f^l \delta N^l \right) + g f^2 e^{a b i} \partial_b \delta N - 2 g^2 f^3 \delta^a \delta N = 0,
\]

(4.18)
where in the ADM formalism $N$ represents the lapse, $N^i$ the shift and $\gamma_{ij}$ is the transverse and traceless quantum fluctuation of the three-dimensional metric s.t. $\gamma_{ii} = \partial_i \gamma_{ij} = 0$. To proceed, one has to fix the gauge and then decompose the gauge fields into (canonically normalized) helicity eigenstates

$$\delta \tilde{A}_\mu(\tau, k) = \sum_\lambda \left[ (\hat{g}_\lambda)_\mu^\omega(x) + (\hat{f}_\lambda)_\mu^\omega(x) \right],$$

where $x \equiv -k\tau$ is dimensionless and $\lambda = (\pm 0, \pm 1, \pm 2)$. Inserting this into the EOM, this leads to a system of equations for the coefficients $\omega_\lambda^i$ with $i \in g, f$ denoting the physical and constraint degree of freedom, respectively. We only care about the physical helicity modes. In fact, we are particular interested in growing modes which have the biggest impact on the dynamics. The only mode which can exhibit a tachyonic instability is the $-2$ mode and the equation reads [52]

$$\frac{d^2}{dx^2} \omega_{-2}^{(g)}(x) + \left( 1 - \frac{2\xi(x/k)}{x} + 2 \left( \frac{\xi(x/k)}{x} - 1 \right) y_k(x) \right) \omega_{-2}^{(g)}(x) = 0,$$  \hspace{1cm} (4.20)

where we defined

$$y_k(x) \equiv \frac{g f(\tau)}{k} = \frac{c_i(x/k)\xi(x/k)}{x}. \hspace{1cm} (4.21)$$

Modes gets generally frozen when they leave the horizon. Thus we have to evaluate $\omega_{-2}$ for modes exiting the horizon at different e-folds. Necessary to this end is the fixing relation between wave number and e-fold for horizon crossing $k(N) = e^{-N/H(N)}$.\footnote{Also proven to be useful is the relation $-\tau(N) = e^N/H(N)$, which implies $x = -k\tau = 1$ for horizon exit.}

Let us take a step back for the rest of the section and do the slow roll approximation $\xi \sim \text{const.}$ for a first qualitative analysis.\footnote{So when we type $y(x)$ we mean that the functional dependence in $\xi$ should be dropped, thus also in the background field.} This allows an analytic solution in terms of the Whittaker function normalized by the Bunch-Davies vacuum ($\omega^{(g)}_{-2}(x) = \exp(ix)$ for $x \to \infty$)

$$\omega^{(g)}_{-2}(x) = e^{(1+c_i)\pi x/2} W_{(1+c_i)\xi,x}^{2c_i,1/4}(-2ix).$$  \hspace{1cm} (4.22)

\footnote{Note that we already took the analytic approximation for the gauge field background, what we found to be in perfect agreement with the full numerical solution.}
Interestingly, in a background described by a $c_0$ solution this coincides with the Abelian solution, c.f. eq. (3.17). This is because for $x < 2\xi$ the mode is tachyonically enhanced due to the same negative effective mass squared term as in the Abelian case, i.e. $m^2 = 1 - \xi/x$. So we find that CNI reduces to natural inflation [48] in the case of a vanishing background field. This will always be the case when starting from the Bunch-Davies vacuum. Thus, we will refer to this stage as the 'Abelian regime'. A completely different dynamics will be obtained when for large $\xi$ the $c_2$ solution emerges. The presence of a background with non-zero vev drastically modifies the behaviour of the enhanced mode. Key feature is the stopping of the exponential growth of the mode already shortly after it begins. The reason lies within the modification of the mass term

$$m^2 = 1 - \frac{2\xi}{x} + \frac{\xi(x-x)}{x^2} \left( 1 + \sqrt{1 - 4/\xi^2} \right), \quad (4.23)$$

yielding a two sided bounded region for a possible tachyonic instability, see fig. 4.1. The bounds are given by

$$x_{\text{min}} \equiv \xi \left( 1 + c_2 - \sqrt{1 + c_2^2} \right) < x < \xi \left( 1 + c_2 + \sqrt{1 + c_2^2} \right) \equiv x_{\text{max}}. \quad (4.24)$$

So, the time of the growth can be estimated with $\Delta x \simeq \xi \sqrt{8}$. The effects arising in this regime can undoubtedly be assigned to the $SU(2)$ dynamics.

We may emphasize our viewpoint as follows: As long as we are in the limit of
small gauge coupling and/or small gauge field amplitudes, the SU(2) group will act as 
\(2^2 - 1 = 3\) copies of an Abelian group. For a vanishing background, i.e. **c_0 solution**, 
the linearized SU(2) dynamics can be described in the Abelian limit, meaning that 3 
modes will be enhanced for \(x < 2\xi\). The natural growth of \(\xi\) will eventually trigger 
the transition to the fundamentally non-Abelian regime and manifest itself through 
the **c_2 background**. This in turn modifies the effective mass terms in the ODE for 
the helicity modes, leaving only one single tachyonic enhanced mode. Although this 
enhancement occurs earlier, the amplitude will eventually saturate and start to slowly 
decrease with rapid oscillations.

### 4.1.3 Background Transition

The remaining question now is, how the background evolves over the course of infla-
tion, where \(\xi\) typically increases. In particular, we would like to know for which initial 
conditions the **c_2 solution** will be present and especially at what time this will be the 
case. In a realistic scenario, however, the background field would dynamically evolve in 
the presence of the fluctuations. Since we are working in the homogeneous field limit, 
we can only approximate this behaviour. So, we neglect the fluctuations during the 
discussion and include them argumentative in the final result. It was shown in ref. [52] 
that the background field undergoes an oscillatory phase in the far past, for general ini-
tial conditions\(^4\). Thus, the task is to estimate when the oscillatory phase ends and the 
background approaches the solution given in eq. (4.8). One can show that the solution 
at early times are characterized by two constants, an amplitude \(A \geq 0\) and a phase 
\(u_0 \in [0, 4K(-1))\), where \(K(-1) \simeq 1.3\) denotes the complete elliptical integral of the 
first kind. They are uniquely determined once the initial conditions are fixed. In the 
infinite past the background field can be well approximated with the Jacobian elliptic 
function \(sn\) to be \(f \sim A \text{sn}(u_0 + Ag \tau, -1) + \xi/(−\tau g)\), see ref. [52] for details. The regime 
of oscillatory domination then corresponds to \(\xi/(−\tau g) \ll A\). This leads to a suitable 
criterion for which times the background field strongly oscillates\(^5\)

\[-\tau A g \gg \xi. \quad (4.25)\]

\(^4\)We note, that this is not true in the presence of fluctuations, since the growing super-horizon 
fluctuations make the classical approach invalid.

\(^5\)This criterion is only true for \(A > 0\). For the case \(A = 0\) however there are trivially no oscillations.
Seeking for a better intuitive understanding what the transition regime actually means, we may bring the background field EOM into an autonomous form. This gives us the opportunity to visualize the dynamics in a simple 2d phase space. We can choose the transformation 

\[ q = -g f e^{-N} \quad \text{and} \quad p = g(\partial_N f)e^{-N}, \]

bringing it to the desired form

\[ \frac{dq}{dN} = q - p, \quad \frac{dp}{dN} = 2(q^3 - \xi q^2 + p). \]  

Thus, the solution can be visualized by displaying the flow of the vector field \((q - p, 2(q^3 - \xi q^2 + p))\). We show the result in figure 4.2. The zeros of the vector field are given by \( z_i = c_i(\xi, \xi) \). The figure demonstrates that all the flow lines fall into the different \( c_i \) solutions. At early times the lines circuit the zeros of the vector field, marking the oscillatory regime. Once they cross the border regime of \(-\tau A g \sim \xi\), they very quick approach one of the \( z_i \) attractors. We especially highlight the two allowed \( c_1 \) flow lines. There are only exactly two such lines, since the \( c_1 \) solution forms a one-parameter family, thus being unstable under perturbations. This is contrary to the \( c_{0,2} \) solution which form a two-parameter family, giving rise to an infinite set of stable trajectories. The phase space study reveals a neat symmetry of the trajectories.

---

6By (a particular \( c_i \)) flow line we mean the line of the vector field falling into one of the \( z_i \) zeros of the vector field. No two of such flow lines are allowed to cross, making any flow line path unique.

7The approximate amplitude of the fluctuations \( \mathcal{A} \) is after the transformation given by

\[ D = \left( p^2 + q^4 - 4\xi q^3/3 \right)^{1/4}, \]

see ref. [52].
In particular, the two $c_1$ flow lines form the border for which trajectories evolve into either the $z_0$ or $z_2$ valley. We can easily see that the region of $c_2$ - type solutions occupy the majority of the phase space, hence they are very stable under perturbations. Given the boundary we derived for the end of the oscillatory phase, we can legitimately assume that trajectories outside this boundary will evolve to high probability into the $z_2$ valley by following the $c_2$ flow lines.

With this understanding, let us now include the background field fluctuations which we have neglected so far. Figuratively speaking, the fluctuations act like perturbations for the field moving along a certain flow line. We start in the Bunch-Davies vacuum in the infinite past, which is equivalent to a zero gauge field background vev. Hence, we are lead consider the $z_0$ zero to be the starting point. It is stable under small perturbations, but when the threshold of $-\tau A g \sim \xi$ is reached, the background field 'falls' into the $z_2$ zero. Hence, the background acquires a non-zero vev given by the $c_2$ solution.

In a more complete picture, we would expect that this transition is sourced by the growing quantum fluctuations of the background field. It is thus natural to replace in first approximation the classical amplitude $A$ with its quantum analogue, leading to the transition time

$$-\tau g \sqrt{\langle A^2 \rangle} \sim \xi.$$  

We discussed that fluctuations are described by the Abelian limit for a background with vanishing vev, i.e. $c_0$ solution. Therefore, we should evaluate the transition time with

$$\langle A_{ab}^2 \rangle^{1/2} = \langle A(\tau, x)A(\tau, x) \rangle^{1/2} = \left( \int \frac{d^3 k}{(2\pi)^3} \left( \tilde{A}_+(k) \tilde{A}_+^*(k) + \tilde{A}_-(k) \tilde{A}_-^*(k) \right) \right)^{1/2}$$

$$\simeq e^{\pi \xi/2} \left( \int \frac{dk}{4\pi^2} k |W_{-i\xi,1/2}(2ik\tau)| \right)^{1/2}$$

$$= \frac{aH}{2\pi} e^{\pi \xi/2} \left( \int dx |W_{-i\xi,1/2}(-2ix)| \right)^{1/2}$$

$$\sim 8 \times 10^{-3} \frac{e^{2.8\xi}}{-\tau}.$$  

---

8 We note that the probability is dependent on the only free parameter $\xi$ and it increases with increasing $\xi$.

9 The last approximation starts to significantly deviate from the actual quantity the smaller $\xi$ gets. However, even for the detailed parameter space scan - which we will provide at the end of this chapter - the important transition regime will be at $N \lesssim 30$. This directly translates to sufficiently large $\xi$ such that the approximation holds very good. Thus the reader may unworried use it to reproduce the results.
So, the time when the background field acquires a non-zero vev by changing from the $c_0$ to the $c_2$ solution is determined by\(^\text{10}\)

$$
\frac{1}{\xi} e^{2.88} \sim \frac{10^3}{8g}.
$$

### 4.1.4 Gravitational Wave Fluctuation

Of particular phenomenological interest is the coupling between the metric tensor fluctuations and the gauge field mode, i.e. helicity $-2$ mode. This directly relates to the amplitude of gravitational waves. As before, we will only quote the final result for the EOM — now for the gravitational waves \([52]\)

$$
\frac{a}{4} \left( (a\gamma_{ij})'' - \left( \partial^2 + \frac{\Omega''}{\alpha} (a\gamma_{ij}) \right) \right) = \frac{1}{2a} \left( f'^2 - g^2 f^4 \right) (a\gamma_{ij}) - f' \partial_0 \delta A^{(i)}_j
$$

$$
+ f' \partial_{(i} \delta A_{j)}^{(\ell)} - 2g f^2 \epsilon^{(i(\ell} \partial_{)j) \delta A^{a}_{\ell)} + g^2 f^3 \delta A^{(i)}_{j) },
$$

where (..) and [...] means that the indices should be symmetrized and anti-symmetrized respectively. Expressing the metric tensor perturbations in the helicity basis as

$$
a_{ij} = \frac{1}{2} \sum_{\lambda = \pm 2} \omega_{\lambda}^{(\gamma)} \begin{pmatrix} 0 & 0 & 0 \\ 0 & \mp i & 1 \\ 0 & 1 & \pm i \end{pmatrix},
$$

the EOM for the $\pm 2$ coefficients are given by \([52]\)

$$
\frac{d^2}{dx^2} \begin{pmatrix} \omega_{\pm 2}^{(g)}(x) \\ \omega_{\pm 2}^{(\gamma)}(x) \end{pmatrix} + N_{\pm 2}(x) \frac{d}{dx} \begin{pmatrix} \omega_{\pm 2}^{(g)}(x) \\ \omega_{\pm 2}^{(\gamma)}(x) \end{pmatrix} + M_{\pm 2}(x) \begin{pmatrix} \omega_{\pm 2}^{(g)}(x) \\ \omega_{\pm 2}^{(\gamma)}(x) \end{pmatrix} = 0.
$$

The appearing matrices are defined like

$$
N_{\pm 2}(x) = \frac{dy(x)}{dx} \frac{Hx}{g} \begin{pmatrix} 0 & -1 \\ 4 & 0 \end{pmatrix},
$$

$$
M_{\pm 2}(x) = \begin{pmatrix} 1 & 2 & \left( \frac{\xi}{x} + 1 \right) y(x) H \left( -\frac{dy(x)}{dx} - (2\xi + x) y(x)^2 + xy(x)^3 \right) \\ -4H x y(x) \xi^2 (\pm x - \xi) & 1 - \frac{2\xi}{x^2} + \frac{2H^2 \xi^2}{y^2 x^2} \left( y(x)^4 - \left( \frac{dy(x)}{dx} \right)^2 \right) \end{pmatrix}.
$$

\(^{10}\)We comment on alternative definitions of the transition time in appendix D.
We may use that the $+2$ mode is not enhanced and is thus given by the known solution of the Mukhanov-Sasaki equation $\omega^{(\gamma)}_{+2}(x < 1) = 1$ [54–56]. Then, the energy in GW per logarithmic frequency interval normalized to the critical energy density is given by [52]

$$\Omega_{\text{GW}} = \frac{\Omega_r}{24} \left( \frac{2H}{\pi} \right)^2 \left( x + \left( x|\omega^{(\gamma)}_{-2}(x)| \right)^2 \right) \bigg|_{x<1}. \quad (4.38)$$

The evaluation at $x < 1$ ensures that the amplitude is frozen, since modes cross the horizon at $x = -k\tau \sim 1$. The stochastic GW background is already at linear level enhanced at high frequencies due to the mode couplings. This is contrary to the Abelian case where such a coupling only appears at non-linear level. This is because there is no background gauge field present and the stress energy tensor is bi-linear in the gauge fields.

For the numerical analysis which we will present in the following, we restore the time evolution, i.e. $y(x) \to y_k(x), \xi \to \xi(x/k), H \to H(x/k)$.

### 4.2 Numerical Treatment & Phenomenology

In the previous section we provided the necessary framework to study the dynamics. Here, we will give a manual on how to solve them. Recap, that in the linearized non-Abelian theory there are essentially two distinct types of gauge field backgrounds, denoted by $c_{0,2} - \text{type solution}$. We discussed that for sufficiently small $\xi$ the only stable background is of the $c_0 - \text{type}$, i.e. vanishing. The tachyonically enhanced perturbation evolve in the same way as in the Abelian case. Thus, we refer to the phase of a zero background vev as the 'Abelian Regime'. The accelerated motion of the inflaton drives the background with high probability to the $c_2 - \text{type solution}$. This is basically because the large fluctuations from the tachyonically enhanced mode push the background away from the $c_0 - \text{type solution}$. The changeover in the background approximately happens when the fluctuations reach $\langle A_{ab}^2 \rangle^{1/2} \sim \xi/(-g\tau)$, c.f. eq. (4.27). In the non-Abelian regime the dynamics completely changes, since tachyonic mode is damped by the background.

We can use our qualitative understanding to formulate an easy to follow users manual for the numerical study:
1. Slow Roll

To find the initial conditions, we match to the slow roll approximated solution in the far past. Once the potential is fixed this can be done analytically and for an axion-like potential as given in eq. (4.44) we have

$$\phi_{sr} = 2 f_V \arccos \left( \sqrt{1 + \frac{1}{(2 f_V^2)^2} \times e^{-\frac{N}{2 f_V}}} \right).$$

(4.39)

2. The Abelian Regime

Solve the Abelian equations (3.13) and (3.14) until

$$\langle A_{ab}^2 \rangle^{1/2} = \xi / (-\tau).$$

(4.40)

We denote this as the matching point.

3. The Non-Abelian Regime

Switch from the vanishing background to the $c_2$ solution for the background and solve eq. (4.10). As initial condition we take the value of $\phi$ from the Abelian limit and its derivative at the matching point. If one solves the full coupled set of EOMs (4.14)-(4.16), the additional initial conditions for the background field are given by the value of $f$ and its derivative at the matching point. Solve until $\epsilon = 1$. This will always happen for $N_{\text{end}} < 0$ if $\alpha/f_c > 0$.

4. Shift of the Solution

We like to set the end of inflation to be at $N = 0$. We thus shift the solution like $\phi(N) \to \phi(N - N_{\text{end}})$. The same shift may also be applied to $\phi_{sr}$ to depict both functions in one plot.

Although this procedure is quite rough, it should give a good first approximation to the complete dynamics. Physically there should not be a discrete point where the dynamics suddenly changes. In a more realistic description the transition from the Abelian to non-Abelian regime should be smooth. However, this uncertainty should only touch a

\[\text{footnote text}\]
few e-folds around the matching point. Outside this regime we expect a washout of the rough matching procedure due to a rapid approach to an equilibrium dynamic. But note, that we can only make reliable predictions in the non-Abelian regime as long as the linearization of the EOM is justified. This is

$$\langle \delta A_{\text{NA}}^2 \rangle^{1/2} \ll f, \quad (4.41)$$

which we will denote as the strong criterion. Another criterion is the weak criterion [43]

$$-\partial_N \langle \delta A_{\text{NA}}^2 \rangle^{1/2} \ll f. \quad (4.42)$$

The non-Abelian variance is given by

$$\langle \delta A_{\text{NA}}^2 \rangle \bigg|_{N=N_h} \simeq \frac{1}{-\tau(N)^2} \int \frac{dx}{2\pi^2} \left| \omega_{-2}(x) \right|^2 \bigg|_{N=N_h}. \quad (4.43)$$

The $N_h$ indicates the e-fold of horizon exit. Requiring to track the inflaton dynamics safely until $N \sim 10$, we will use both criteria to quantify the maximal numerical allowed parameter space\footnote{We generically take $N_{\text{max}} = 10$, but the exact choice has no deeper meaning.}.

To get a better feeling for the setup, we will study the dynamic with an example. The potential we study here will be the standard periodic axion potential

$$V = \Lambda^4 \left( 1 - \cos \left( \frac{\phi}{f_V} \right) \right), \quad (4.44)$$

introduced in the context of inflation as natural inflation (Abelian) [57] and chromo-natural inflation (non-Abelian) [36]. The constants must of course be chosen such that the constraints discussed in the previous chapter are fulfilled. This can be approximately done analytically in the slow roll limit, shown in appendix C.3 explicitly for this potential. So, we fix $\Lambda^4 = 4.74 \times 10^{-9}$ and $f_V = 9.2$. Furthermore, the two couplings will be fixed to $\alpha/f_c = 35$ and $g = 5 \times 10^{-3}$. Note, that contrary to the in literature treated case of $f_c \equiv f_V$, we take them not necessarily to coincide in order to have an EFT description in accordance with general relativity (GR) and standard quantum field theory (QFT). More discussion about this important step and its consequences can be found in the next section. Following the numerical guideline, we obtain the results for $\phi$ and $\xi$ shown in figure 4.3. The instant transition from the Abelian to the non-Abelian regime manifests
Figure 4.3: **Left**: We show the evolution of the inflaton for all three possible cases, i.e. i) vacuum dynamics and in the presence of ii) Abelian and iii) non-Abelian gauge groups. The parameters are given in the main text. **Right**: Evolution of the $\xi$ parameter which encodes the whole information for the tachyonic instability in the gauge field for the ii) Abelian and iii) non-Abelian case. In both plots we additionally indicate with the black vertical line the matching point, obtained by solving eq. (4.40).

itself as a discontinuity in $d^2\phi$. Thus, our matching procedure is clearly visible in the $\xi$ evolution. Furthermore, the gauge friction is revealed to be sub-dominant, which is contrary to the pure Abelian case. Although the $\xi$ parameter drops rapidly at the matching point — which is caused by instantly turning on a non-zero background vev — the gauge friction cannot keep it low. This is because the $SU(2)$ induced friction has a power law form ($\propto \xi^3$) whereas it is exponential sensitive in the Abelian regime, c.f. eqs. (3.22) and (4.10). In a more complete theory, the transition between these two regimes should be smooth. Apparently, we cannot provide this within the linearized framework. Thus, we will avoid to make any statement for quantities which are somehow near the matching point. In particular this implies that we exclude $\sim \pm 2$ e-folds around the transition regime when making contact to observable quantities. This conservative approach substantiate the validity of the calculations especially for the phenomenology.

However, this is not the only apparent bound we have to face. To obtain this plot, we ignored the limit of the linearized prescription. This means that we calculated the whole dynamics with the a priori assumption that the $c_2$ background dominates over its fluctuations. As discussed before, reliable predictions can only be made if this is indeed true and we have to check this a posteriori. Before presenting the result for the enhanced non-Abelian fluctuation, we noticed during the numerical calculations, that
the full solution of eq. (4.35) with a \(c_2\) background can be very well approximated by

\[
\omega^{(g)}_{-2}(x) \bigg|_{N=N_h} = e^{\kappa(N)\pi/2}W_{-i\kappa(N),-i\mu(N)}(-2ix) \bigg|_{N=N_h+3},
\]

with \(\kappa(N) \equiv (1 + c_2)\xi \simeq 2\xi, \mu(N) \equiv \xi \sqrt{2c_2 - 1/(4\xi^2)} \simeq \sqrt{2}\xi\). Note that this is of very similar form as the slow roll solution for the uncoupled system of eq. (4.22). The \(N\) evolution is only included with a +3 shift. Now recap that we only need the mode as an intermediate result to calculate the variance. The variance gets its main contribution only from the time where the mode is tachyonically enhanced. That is, we integrate from \(x_{\text{min}}\big|_{N=N_h}\) to \(x_{\text{max}}\big|_{N=N_h}\), c.f. eq. (4.24). So it suffices to guarantee that the full numerical solution is well approximated within this regime. And in fact we checked that the result only differs for \(x \lesssim 0.1 \ll x_{\text{min}}\). In figure 4.4 we show the evolution for modes exiting the horizon at different \(N_h\) and the resulting variance. Modes exiting the horizon

**Figure 4.4:** **Left:** We show the evolution of the tachyonic amplified +2 mode for different \(N_h\), obtained by numerically solving eq. (4.20). **Right:** The resulting variance and the \(c_2\) - type background in the non-Abelian regime (yellow). We also include our analytic approximation for the variance as well as for the background (black), which fits the numerical result very well. We see how the implementation of the weak criterion allows to extend the linearized analysis for a few e-folds.

at later times, i.e. smaller \(N\), have a bigger amplitude, which is physically plausible, since there is basically more time for the amplification due to large \(\xi\). When turning to the fluctuations, we see that the growth eventually gets so strong, such that the fluctuations begin to dominate over the background at \(N \sim 10\). The exact time when this happens is sensitive to the criterion one chooses. We included the weak and the strong criterion, as defined above. We find that in both criteria the analysis finds its limit. However, the weak criterion pushes the bounds of the linearized analysis evidently to smaller e-folds.
the failing of the linearized analysis happens at such small $N \sim 10$, the uncertainty is only the exact end of inflation, but this does not spoil the phenomenological predictions.

So, let us calculate the energy stored as GWs with the help of eq. (4.38). To this end, we need to solve the coupled system of ODEs given in eq. (4.35). The result for the $\omega_{-2}^{(\gamma)}$ mode together with the resulting GW spectrum is shown in figure 4.5. Although

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.5.png}
\caption{Left: The evolution of the $-2$ tensor mode fluctuation which is coupled to the tachyonically enhanced $-2$ gauge field mode. The black vertical line represents the horizon crossing time. The fluctuations get frozen as expected. Right: The resulting stochastic GW background. The colored region represent the weak and strong criterion for which the linearized analysis fails. The vertical line indicates the matching point. An inherent non-Abelian signal can be detected even when excluding a few e-folds around the transition regime. Compared to the pure Abelian signal, the signal is suppressed by $\sim 2$ orders of magnitude. For reference we also depict the vacuum contribution as a dashed curve.}
\end{figure}

the last few e-folds come with uncertainties and we exclude a few e-folds around the matching point, one can expect a detectable GW signal from the inherent non-Abelian regime. It is different from the signal one would expect from a pure Abelian setup in that sense that the non-Abelian amplitude is highly suppressed. Both signals fall into the detectable range of Einstein Telescope (ET). However, we avoid making any predictions for GW detection with the Laser Interferometer Space Antenna (LISA), because the matching point is around the sensitivity maximum at $N \sim 20$.

Additionally, we may shortly discuss the $\mu$ distortions. As we showed in section 3.3 they are only sensitive to the integrated scalar power spectrum. As discussed in ref. [52], the non-Abelian scalar power spectrum is in general enhanced compared to the Abelian case. However, since the window function (shown in figure 3.5) has a steep cut-off at $N \simeq 40$, we do not expect non-Abelian effects to contribute to the distortion. This is because the matching point will always happen at later time, such that the CMB
distortions get sourced by the Abelian limit of the theory. So we do not expect a
significant different result than for the pure Abelian case.

4.3 Model Constraints

In this section we address numerical and theoretical bounds which the setup of chromo-
natural inflation has to face. We separate the discussion for both types of the constraints,
since they are qualitatively completely different. On the one hand, the numerical con-
straint is due to our restriction to linearized EOMs. We provide an accurate parameter
scan indicating the numerical reliability and show possible GW signals detectable with
ET. On the other hand, the theoretical constraints are axiomatic, undoubtedly restrict-
ing the model in the context of an effective field theory. With the final parameter scan
over the $(\alpha/f_c, g)$-plane, we find that nearly the complete model is ruled out theoret-
ically. And even in the allowed region we find that the model building is desperately
hard to realize. Notably, when imposing the natural scale ordering, as it is known
from the Peccei & Quinn original axion proposal [58] and the original natural inflation
proposal [57], we find that chromo-natural inflation is totally excluded.

4.3.1 Numerical Limit

In the previous section we demonstrated for exemplary parameter points where the lin-
earized analysis finds its limit. We denoted the violation of the strong criterion as the
domination of the fluctuations over the background at a given time. And the weak
criterion where the fluctuations per e-fold are compared to the background. The expo-
nential growth of the Abelian fluctuations will certainly overcome the nearly constant
background field amplitude. However, if the criterion is only broken at $N < 10$, basi-
cally the only uncertainty is the exact end of inflation. There, however, the numerical
uncertainties are anyway high due to possible strong inflaton–gauge field back reactions.
But since GW detectors are sensitive to GW releases at $N > 10$ and CMB scales are
measured at $N \sim 55$, we take the analysis to be trustworthy in the case of a violation
only at $N < 10$. Concerning the matching point, we can calculate most of the dynamics
within the Abelian regime if it is at $N < 10$. Evidently, then the non-Abelian fluctua-
tions will dominate only for $N < 10$ (if they ever will). This goes hand in hand with the
first point, thus making the calculation to good approximation also reliable. So we can emphasize the numerical limit as follows:

- **Weakly Allowed:** The matching point is at \( N < 10 \) or the weak criterion is only broken for \( N < 10 \), c.f. eq. (4.42).

- **Strongly Allowed:** The matching point is at \( N < 10 \) or the strong criterion is only broken for \( N < 10 \), c.f. eq. (4.41).

Naturally, a point which is strongly allowed is also weakly allowed — but not the other way around.

### 4.3.2 Effective Field Theory Bounds

We already indicated previously that we work work with three different scales, namely \( f_c, f_V, \Lambda \). The potential height is given by \( \Lambda^4 \) and simply represents the confinement scale of the theory. In usual axion models we have \( f_c = f_V \) given by the decay constant. The reason why we like to choose \( f_c \neq f_V \) is simply because (chromo-) natural inflation requires a super-Planckian periodicity \( f_V \), see appendix C.3. When both scales are taken to be equal, this means that we would have a super-Planckian EFT cut-off. But, such a high cut-off breaks down any viable EFT description. So, in order to obey classical GR and standard QFT, we have to demand \( f_c \ll 1 \), which implies \( f_c \ll f_V \). It turns out, that this requirement is not strict enough. Rather, we should additionally demand

\[
T_{\text{PQ}} = C \times f_c \gg T_{\text{dS}} = H/(2\pi), \tag{4.46}
\]

where \( C \) is a constant which we have to determine now. We impose this requirement, since the potential breaking temperature \( T_{\text{PQ}} \) should be above the to de-Sitter space inherent Gibbons-Hawking temperature \( T_{\text{dS}} \) \([59]\). We note, that this effective horizon temperature prescription gives the same numerical bound as to consider that the vacuum fluctuation — also given by \( H/(2\pi) \) \([60]\) — should break the potential. In general, the EFT cut-off scale is related to the symmetry breaking temperature through a constant. Clearly, when the temperature is high enough, the symmetry is restored and thus no particle will be present. Only for temperatures below \( T_{\text{PQ}} \) the potential gets broken. The resulting 'mexican hat' potential induces through its angular degree of freedom a
Nambu-Goldstone boson, which acquires through non-perturbative instanton effects a periodic potential at the confinement scale. This actually leads approximately to the development of the typical quoted cosine potential, breaking the shift symmetry into a discrete subgroup. Thus, it is additionally natural to assume that this processes should happen in the typical order, namely $C \times f_c \gg \Lambda$. It is the same order which one is familiar with in the case of the axion originally imposed by Peccei & Quinn [58] and in the case of the original natural inflation proposal [57]. So, finally we have $f_V \gg C \times f_c \gg \Lambda \gg T_{dS}$.

Note, however, that the condition $C \times f_c \gg \Lambda$ is not a stringent theoretical requirement, since both enter the Lagrangian as free parameter. But, it is not clear how such a theory would behave, when the order is reversed. Let us assume $C \times f_c \ll \Lambda$ for a moment and look what we get. This order implies, that typically the radial component of the field acquires a mass which is lighter than that of the axion. For a few reasons this is dubious and should be avoided. First, due to the naturalness argument proposed by t’Hooft [20]. It states that the symmetry of the system should be enhanced if the symmetry breaking parameter goes to zero and thus the parameter should be small.

The axion mass is clearly natural small as we would restore the shift symmetry of the system when $m_\phi = \sqrt{2\Lambda^2/f_V} \to 0$. So, the mass of the axion should naturally be the lightest in the setup, also because quantum corrections to the parameter are proportional to the parameter itself. But since the radial component is not protected by a natural symmetry, we expect significant loop contributions to the bare mass, which are difficult to quantify. Second, we are working in an EFT framework where the radial component is integrated out. This, however, is in contradiction to the EFT approach if we have $m_\rho < m_\phi$. Third, when we consider the two-field model of radial part plus axion, it is not clear which one of these particles will play the role of the inflaton. So, we are led to consider $m_\rho > m_\phi$ as a theoretical bound to protect the EFT description. To make this more explicit, let us consider the original mexican hat potential of Peccei & Quinn

$$V_{PQ}(\Phi) = \lambda \left( |\Phi|^2 + \frac{v_{PQ}^2}{2} \right)^2, \quad (4.47)$$

where the axion is the imaginary part of the complex scalar field $\Phi = \sqrt{1/2} \rho \exp(i\phi)$ and the Peccei & Quinn scale is given by $v_{PQ} = C' f_c \simeq C f_c$. The coupling parameter $\lambda$ is dimensionless and must satisfy $\lambda \leq 1$ to have a reasonable mass term. To obtain the mass of the radial part $\rho$ we Taylor expand to second order in $\phi = 0$ and $\rho = v_{PQ}$. This
leads to

$$\frac{1}{2}m_{\rho}^2 = \lambda v_{\text{PQ}}^2 = C^2 \lambda f_c^2. \quad (4.48)$$

By comparing the axion mass obtained from the explicit symmetry breaking with the radial mass obtained by the global symmetry breaking, we need

$$f_c \gg \Lambda \times \left( \frac{\Lambda}{f_V} \right) \times \frac{C}{\sqrt{\lambda}} \quad (4.49)$$

for a viable EFT description. The setup is qualitatively shown in figure 4.6, which should make the discussion intuitively clear. The complete order of bounds thus is

---

**Figure 4.6:** **Top: Left:** Symmetry of the PQ potential for different temperatures $T$. The broken symmetry with the angular degree of freedom (axion) is obtained for $T < T_{\text{PQ}}$. **Right:** The resulting periodic potential above the confinement scale $\Lambda$ breaking the continuous shift symmetry around the bottom of the potential into a discrete subgroup $\phi \rightarrow \phi + 2\pi f_V$. **Bottom:** Examples of the globally broken Peccei & Quinn potential together with the explicit periodic symmetry breaking around the bottom of the potential. The left plot shows the natural case of $m_\phi < m_\rho$, whereas in the right we indicate how the contrary case would look like.
\[ T_{\text{PQ}} = C \times f_c (\gg \Lambda) \gg \Lambda \times \left( \frac{\Lambda}{f_V} \right) \times \frac{C}{\sqrt{\Lambda}} \gg \frac{H}{2\pi} = T_{\text{dS}} (\equiv \Delta_{\delta\phi}). \]  

(4.50)

Let us now turn to the quantification of the constant \( C \). In general, \( C \) depends on the Peccei & Quinn charge \( c_{\text{PQ}} \) assigned to (SM) fermions. For the two well known axion-photon coupling models KSVZ \[61, 62\] and DFSZ \[63\], one obtains \( C = c_{\text{PQ}}/\pi = \mathcal{O}(1) \).

So, we may be safe to conclude that also for our case an order one number is sufficient to assume. This leads us to the conclusion \( f_c \simeq v_{\text{PQ}} = T_{\text{PQ}} \), setting \( C = 1 \).

Let us emphasize the so obtained EFT constraint. The requirement of sub-Planckian EFT cut-off and the agreement with Planck cosmological data \[1\], forces us to consider \( f_V \gg f_c \). Together with the need of a broken PQ symmetry (= existence of the axion), we obtain \( f_V \gg f_c \gg T_{\text{dS}} = H/(2\pi) \). This can be directly translated into a bound in the \((\alpha/f_c, g)\)-plane, since \( \alpha \) is related to the gauge group coupling through \( 4\pi\alpha = g^2 \).

We thus have to impose the consistency relation

\[ \frac{\alpha}{f_c} \ll \frac{g^2}{2H}. \]  

(4.51)

The other direct EFT bound arises from the requirement of natural mass ordering \( m_\phi < m_\rho \), leading to

\[ \frac{g^2}{4\pi(\alpha/f_c)} \gg \Lambda \times \frac{\Lambda}{f_V} \times \frac{1}{\sqrt{\Lambda}} \sim 10^{-3}\Lambda, \]  

(4.52)

where we minimized the right hand side by setting \( \lambda = 1 \). Note that more stringent bounds can be derived with smaller values for \( \lambda \).\(^{13}\) We however would like to keep the analysis as broad as possible and do not rely on a specific model. Furthermore, we may indicate for which region in the parameter space we would have the natural scale ordering

\[ f_c = \frac{g^2}{4\pi(\alpha/f_c)} \gg \Lambda \gg T_{\text{dS}} = \frac{H}{2\pi}. \]  

(4.53)

Also very interesting to point out is the ratio between the two scales \( f_V \) and \( f_c \), since this may be important for string theoretical model building. A realistic region which can be reached is up to \( f_V/f_c \sim \mathcal{O}(10^3) \). Ratios (far) above this threshold probably cannot be reached or require desperate model building. Thus, we may impose the by

\(^{13}\)One may notice that for \( \lambda \to 0 \) we need \( f_c \to \infty \). This is totally reasonable as the global symmetry breaking in this limit is always softer as the explicit one, making the radial mass lighter than the angular one.
string theory motivated loosely potential building bound

\[ f_V/f_c \lesssim \mathcal{O}(10^3). \quad (4.54) \]

Moreover, we note that the region \( \Lambda > f_c > \Lambda^2/f_V \) is also hard to reach in string theory, since \( \Lambda \) is typically exponentially suppressed compared to the string scale approximately given by \( f_c \). Thus also the region between eq. (4.52) and (4.53) may be considered with caution.

### 4.3.3 Parameter Scan

In figure 4.7 we constraint the parameter points in the \((\alpha/f_c, g)\)-plane with a detailed scan. We varied \( \alpha/f_c \) with a linear step size of 0.1 up to a maximum of \( \alpha/f_c = 35 \),

\[ 0.7 \simeq 1/\sqrt{2} \leq g \leq 0.7 \times 10^{-4}. \]

We take the GUT scale coupling as the upper bound of our analysis. Let us discuss the numerical result and the theoretical bounds separately.
We could show numerically that only a small region of the parameter space undergoes a non-trivial non-Abelian evolution and are simultaneously reliable to calculate within our linearized approach. A vast region is excluded since the long-running dominance of the fluctuation inhibits any reliable prediction of the non-Abelian dynamic. The limits for the numerical breakdown are given in section 4.3.1. Interestingly, the in literature studied magnetic drift regime — where the non-Abelian gauge field induced friction dominates over the Hubble friction \(\alpha \xi H/(gf_c) \gg 1\) — only occupies a small window of the parameter space. And even within this window, the expected signals from sufficient GW production arise more or less only due to the Abelian limit of the theory. Hence, they are not distinguishable from a pure Abelian theory. This is a somewhat surprising result, since this very regime is studied because of the expectation of interesting phenomenological implications. The other motivation was a sub-Planckian EFT cut-off when an initially non-zero background field vev is present. We also could not verify this result in the context of a dynamical vev generation by the fluctuations starting from the Bunch-Davies vacuum. So, both motivations are pointless in a physical realistic scenario. Note, however, that in our conservative approach we avoided any prediction for GW signals if the matching point lies between \(10 < N_{\text{match}} < 15\). This should minimize the expected error caused by the rough background matching prescription. Also, we emphasize again that in our linearized prescription we neglected many contribution arising from non-linear terms. The reader thus should be aware, that the marked contour lines for the violation of the strong and weak criteria are obtained by setting \(\langle \delta A_{N\Lambda}^2 \rangle = f\) and \(-\partial_N \langle \delta A_{N\Lambda}^2 \rangle = f\) respectively. The negligence of the many non-linear terms will probably play a role a few e-folds before this border line is reached. This may be of importance, when we like to calculate the expected inherent non-Abelian GW signal in more detail. Anyway, we only find a small region of such GW signals.

Of more importance, however, are the derived theoretical bounds for the EFT description of chromo-natural inflation. We can show that the majority of parameter points are theoretical excluded, also depicted in figure 4.7. This means that the model constellation is a priori disallowed. Especially, it has nothing to do with the numerical evaluation and uncertainties. Thus, also a more complete numerical analysis cannot help here. Interestingly, not only the complete range of the magnetic drift regime violates the bounds, but also nearly the complete parameter space. Regions which are theoretical allowed can only be found at large \(g\), i.e. \(g \gtrsim 0.1\). Perturbative control requires as usual \(g < 1\). However, for \(0.1 < g < 1\) we find tight constraints coming from the
numerical evaluation within the linearized framework. Noticeably, no parameter point which is both numerical and theoretical allowed would leave any detectable phenomenological imprint. Neither can we calculate the non-trivial non-Abelian dynamic safely. The situation gets even worse, if we impose the natural order condition $f_c \gg \Lambda$. Then, the parameter space is completely closed and chromo-natural inflation is excluded. And even if we not take this to be a strict bound we run into troubles due to the extremely high ratio $f_V/f_c$. In the tiny region where $m_\phi < m_r$, we need $f_V/f_c \sim 10^4$ to obey CMB constraints [1]. Such a high ratio is however only very hard to realize in string theoretical axion models, if at all possible.
Chapter 5

Towards a Completion of Natural Inflation

A physical more realistic model is that both Abelian and non-Abelian gauge groups are present during inflation. As in the Standard Model of particle physics, we first start with unifying $U(1) \times SU(2)$. We are still working with an axion-like particle driving inflation – but now in the presence of a bigger particle content. We discuss the gauge field unification and show its phenomenological implications with a broad parameter scan. Further, we discuss the coupling to the Standard Model of particle physics through the Adler-Bell-Jackiv anomaly. A SM coupling is anyway needed to provide reheating, underpinning a posteriori the need of the non-Abelian gauge fields. We discuss that in the pure Abelian case we do not expect thermalization of the fermion sector, whereas different results may be expected for the non-Abelian sector.

5.1 Gauge Field Unification

We start the gauge field unification with its simplest realization, namely $U(1) \times SU(2)$. Basically this means that all (non-) Abelian field modes back react on the inflaton and thereby dictating its dynamic. We will show that in the linearized limit the inflaton evolution is dominated by the friction of the tachyonic enhanced Abelian gauge field mode. This means that the $SU(2)$ gauge fields are approximately undergoing an isolated evolution in the background. However, the effective reduction to a pure $U(1)$ gauge
group is not fully justified. The gaining of a non-zero vev for the non-Abelian gauge field background leaves unique phenomenological imprints, making the unified setup well distinguishable from the pure (non-) Abelian one. Note that this result is not restricted to a linearized analysis, although we will working within this approximation. We show the limit of the approximation by calculating the fluctuations in the non-Abelian sector. Our results are only reliable as long as the $SU(2)$ background field dominates over the fluctuations. Therefore, we provide a detailed parameter scan to accurately define the numerical limits of the linearized framework. However, we point out that our conservative approach may be further extendable for some e-folds. This is because we found that the energy density stored in the non-Abelian sector is sub-dominant compared to the Abelian one. In any case, we show that most of the parameter points are disallowed due to too low cut-off scales of the effective field theory. Additionally we comment on a viable possibility to circumvent the strict EFT cut-off. A window of possible GW detection is presented, in accordance with the EFT description.

5.1.1 Guideline & Phenomenology

Let us start by writing down the action for the setup

$$S = \int d^4x \sqrt{-g} \left( \frac{R}{2} - \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - V(\phi) - \frac{1}{4} F^{(a)\mu\nu} F^{\mu\nu}_{(a)} - \frac{\alpha_1(2)}{4f_c} \phi F^{(a)\mu\nu}_{(a)} \tilde{F}^{\mu\nu}_{(a)} \right),$$

(5.1)

where we have to sum the Abelian and non-Abelian contribution, referred to with the index 1 and 2 respectively. In principle, the resulting EOMs have been already derived in the previous chapters. We now have to combine them properly. Physically speaking, the fields from both gauge groups act like friction terms for $\phi$ and modify its vacuum dynamic. Thus, we only have to add them up to get

$$0 = \partial^2_N \phi - 3 \partial_N \phi + \frac{\partial_\phi V(\phi)}{H^2} - \frac{\alpha}{f_c} \left( \frac{3H^2}{g^2} (c_i \xi)^3 + f_1(\xi) H^2 \right),$$

(5.2)

$$0 \simeq H^4 (f_2(\xi) + \tilde{s}) + H^2 \left( \frac{1}{2} (\partial_N \phi)^2 - 3 \right) + V(\phi),$$

(5.3)

$$\tilde{s} \equiv \frac{3c_i^2 \xi^2}{2g^2} (1 + c_i^2 \xi^2),$$

(5.4)

where we already approximated $\partial_N H \simeq 0$ and $\partial_N \xi \simeq 0$ and took $\alpha_1 = \alpha_2$ for simplicity. The functions $f_{1,2}$ and $c_i$ have been defined in the chapters 3 and 4 respectively. Also
we directly use the analytic approximation for the non-Abelian gauge field background, as discussed in the previous chapter. At this point, we have to be careful what exactly the first term of the gauge field friction means. Remember, that the $SU(2)$ gauge group acts like $\sim 3$ copies of the Abelian $U(1)$ at the start of inflation. We discussed, that the non-Abelian dynamics is pushed away from this $c_0$ background limit to the inherently non-Abelian $c_2$ background dynamic if $\langle A^2 \rangle^{1/2} = \xi/(-g\tau)$ is reached. Thus, the non-Abelian friction term has two different properties at early and late times. So, as long as $c_i = c_0$, we replace the non-Abelian friction with its Abelian limit, i.e. $\frac{3H^2}{g^2}(c_i\xi)^3 \to f_1H^2$, and the Hubble parameter reduces to the one derived for the Abelian case, i.e. $\dot{s} \to 0$. As the matching point is reached, we switch to the $c_2$ background, meaning that we now take $c_i = c_2$. This consideration results in

$$
\partial_N^2 \phi - 3\partial_N \phi + \frac{\partial_\phi V(\phi)}{H^2} = \begin{cases} 
2 \times \frac{g}{f_c} f_1(\xi)H^2 & \text{for } \langle A^2 \rangle^{1/2} < \xi/(-g\tau) \\
\frac{\alpha}{f_c} H^2 \left( \frac{3}{g^2}(c_2\xi)^3 + f_1(\xi) \right) & \text{else}
\end{cases}. \quad (5.5)
$$

We take generically the factor of two for the same reason as discussed in the previous chapter.

The numerical strategy to solve this EOM will be the same as discussed in the previous two chapters. We show the result for $\phi$ and $\xi$ in figure 5.1, where we exemplary choose $g = 5 \times 10^{-3}$ and $\alpha/f_c = 35$, as in our studies of the non-Abelian case. It

![Figure 5.1: In the left panel we show the evolution the inflaton in the presence of $U(1) \times SU(2)$ gauge groups and compare it to the standard slow roll inflation, the $U(1)$ and $SU(2)$ limit. The right panel basically depicts $d\phi$, where we also include the (non)-Abelian limit. We see that the an effective $U(1)$ description is valid for this parameter choice.](image)

shows that the complex case of $U(1) \times SU(2)$ effectively reduces to an $U(1)$ theory at the level of the $\phi$ evolution. So the production of the non-Abelian gauge fields have no
significant effect, meaning that its friction contribution is vacuum-like. Thus, the non-Abelian gauge fields undergo an isolated evolution. This could have been anticipated also from the studies we did for the pure non-Abelian case, since \( \partial_N \xi_{SU(2)} \simeq \partial_N \xi_{sr} \). If we compare the evolution of \( \xi \) in the \( U(1) \times SU(2) \) with the \( U(1) \) limit, we notice that they to good approximation coincide, see also figure 5.2. The only significant deviation is around the matching point, but this is due to our linearized analysis, where we have to change in eq. (5.5) from one case to the other one. However, especially in the non-Abelian magnetic drift regime the deviation due to this transition becomes significant. The discontinuity at the matching point is so large, that in the most cases even the rapid increase of \( \xi_{SU(2)} \) cannot reach the level of \( \xi_{U(1)} \) until the end of Inflation. Thus, the non-Abelian gauge fields have a profound impact on the inflaton dynamics in this case.

Let us go back to the study of the above chosen exemplary parameter point. Another argument that we can effectively deal with an \( U(1) \) in this case comes from the consideration of the distribution of the gauge field energy density. The energy density in the Abelian case is given in eq. (3.21). For the non-Abelian limit one obtains

\[
E_i^a = \delta_i^a \partial_0 f^a + \partial_0 \delta A_i^a, \\
B_i^a = \epsilon_{ijk} \partial_j \delta A_k^a + gf^2 \delta_i^a - gf \delta A_i^a + g \frac{\epsilon_{ijk} \epsilon^{abc}}{2} \delta A_j^b \delta A_k^c.
\]

In the limit of \( \langle \delta A^2 \rangle \ll f \), which we anyhow need for a viable linearized analysis, this reduces to

\[
\frac{1}{2} \langle E^2 + B^2 \rangle \simeq \frac{\xi^2 H^4}{2g^2} (1 + \xi^2).
\]

The resulting distribution of the energy density is shown in figure 5.2. Around the matching point, the non-Abelian contribution dominates due to the \( d^2 \phi \) discontinuity. This changes after a few e-folds, since in the Abelian case the energy density shows an exponential dependence on \( \xi \), whereas in the non-Abelian case it is only proportional to \( \xi^4 \). That the energy density in both cases drops for the last few e-folds is due to the rapidly decreasing Hubble parameter at this time, indicating the violation of the slow roll parameter. However, at these times the analysis is not necessarily trustworthy. We have to check this explicitly by calculating the variance of the non-Abelian sector. This can be done in the same manner as described in the previous chapter. We show the result in figure 5.3. Again, we find that the maximum of our analysis is reached at
Figure 5.2: **Left:** We show ratios of $d\xi$ in different limits. It demonstrates the impact of the gauge fields on the inflaton dynamics. The dark cyan curve reveals that the non-Abelian background field only has minor impact, since it is roughly constant around $\sim 2$ (black dotted line). Whereas the brown curve indicates the major contribution from the Abelian gauge field. The magenta line than shows the quotient between the non-Abelian and Abelian limit. Since the curve is always above the black dotted line except for a few e-folds around the matching point, the dynamics can be safely approximated by an $U(1)$.

**Right:** Distribution of the energy density. The Abelian contribution dominates over the non-Abelian. This supports our statement that we can approximate the dynamics as $U(1) \times SU(2) \simeq U(1)$ for parameter points outside the magnetic drift regime. Interestingly, the main component of the energy arises from the magnetic field in the non-Abelian case and from the electric field in the Abelian case.

$N \gtrsim 10$ with the strong criterion and at $N \lesssim 10$ with the weak criterion. This might be a surprising result, since the evolution of $\xi$ is completely different compared to the non-Abelian case, c.f. figure 5.1. With the bound on $\xi$ (roughly) set by the Abelian limit, one expects a slower growth of the fluctuations after the matching point. So one might conclude that the linearized analysis is safely executable for more e-folds than in the non-Abelian case. However, this does not necessarily lead to a later violation of the linearized prescription. This is because the matching point is pushed to higher $N$ compared to the pure non-Abelian case – also caused by the bound on $\xi$. So basically there is more time until $N = 10$ for the fluctuation to overcome the background. And since the fluctuation are proportional to $\exp(\xi)$ and the background only to $\xi$, this will eventually happen not many e-folds after the matching point, c.f. figure 4.7. In figure 5.3 one can nicely see around $N \sim 17$ how the evolution of $\xi$ receive the bound of the Abelian limit and does not follow the non-Abelian evolution anymore. This causes a drop in the growth of the fluctuation, leading that with the weak criterion the analysis can be pushed much farther towards the end of inflation than in the pure non-Abelian case. In this example, we obtain a difference between these two criteria of $\Delta N \simeq 10$ – compared to $\Delta N \simeq 3$ in the pure non-Abelian case.
Pseudo Nambu-Goldstone Boson Inflation

Figure 5.3: Here we compare the variance with the background field of the non-Abelian sector. We see that around $N \sim 17$ the growth of the variance changes qualitatively. This is because $\xi$ reaches the bound set by the dominant Abelian gauge friction, after rapidly increasing around the matching point discontinuity. This drop in the growth of the fluctuation leads to the fact that the weak criterion is only violated much later than it would have been in the pure non-Abelian case.

In the next section we will provide a detailed scan over the parameter space to obtain the limits of the theory, i.e. the region in the non-Abelian sector for which the linearized analysis is justified. But first, let us proceed with calculating the GW background produced within the $U(1) \times SU(2)$ framework. Since both gauge groups are to good approximation decoupled, the expected GW signal is the sum of both production sources. Thus, the resulting amplitude is given by the right combination of eqs. (3.30) and (4.38), where one has to pay attention not to count the vacuum contribution twice. Then, the total GW amplitude is for $N < N_{\text{match}}$ and for $N > N_{\text{match}}$ respectively given by

$$\Omega_{\text{gw}} h^2 = \frac{\Omega_r}{24} \left( \frac{H}{\pi} \right)^2 \times \left( 4(x|\omega^{(5)}_{<2}(x)|)^2 + 2 + 8.6 \times 10^{-7} H^2 e^{4\pi \xi} \right) \bigg|_{x<1} h^2, \quad (5.9)$$

$$\Omega_{\text{gw}} h^2 = 2\frac{\Omega_r}{24} \left( \frac{H}{\pi} \right)^2 \times \left( 1 + 8.6 \times 10^{-7} H^2 e^{4\pi \xi} \right) h^2. \quad (5.10)$$
The result is shown in figure 5.4. We also separate the result from its Abelian and non-Abelian contribution. In addition we show the GW strength in the case of only Abelian gauge fields present during inflation. Recalling the evolution of the $\xi$ parameter from figure 5.1, one could naively expect that both models, pure $U(1)$ and $U(1) \times SU(2)$, would end in the same GW signal. However, as we see from eqs. (5.9) and (5.10), the GW production is exponentially sensitive to the evolution of the $\xi$ parameter. Thus, although the discontinuity in $d^2\phi$ due to the matching procedure (and hence the drop in $\xi$) is in this case small, it has a big impact on $\Omega_{gw}h^2$. This leads to a completely different expected GW signal distribution, which can be used to distinguish between both models. Note that we are not referring to the drop in the spectrum, which is in...
any case non-physical. We would never expect a peaked GW signal at intermediate scales. Rather, the transition should be smooth, such that the peak is washed out. But still the signals can be safely distinguished. As we will see in the next section, there will be also another source which leads to an unique GW signal.

Let us here also mention the restriction due to the linearized analysis for this parameter point. We have a special situation, where the sensitivity curve of ET falls into the region of the theory which is numerical disallowed by the strong criterion but nevertheless fully allowed by the weak criterion. It is not so easy to decide if we should trust the analysis in this region, especially when making contact to observable quantities. In a conservative approach which we followed in our studies, we take a possible signal in this constellation of model parameters only with reservation as a support of the $U(1) \times SU(2)$ theory. For this reason we indicate in the detailed parameter scan, which we provide in the next section, in which part of the 'allowed region' a GW signal for ET falls.

5.1.2 Numerical & Theoretical Bounds

We presented our setup of a unified axion-like inflation in its simplest realization of a $U(1) \times SU(2)$ in the previous section. The illustration of the dynamics with a concrete parameter point study helped us to understand the phenomenological implications. The naively expected reduction to an effective Abelian theory could not be sustained. The set up brings along its very unique phenomenology which can be very well distinguished from a pure Abelian or non-Abelian theory of axion-like inflation. We will now study the numerical reliability of the linearized approach we used to describe the dynamics and indicate the resulting uncertainties resulting from this in a broad parameter scan. In addition we will focus on GW signals which can be possibly detected by ET.

Recall that the numerical bounds for the non-Abelian sector can be summarized as follows

- **Weakly Allowed**: The matching point is at $N < 10$ or the weak criterion is only broken for $N < 10$, c.f. eq. (4.42).

- **Strongly Allowed**: The matching point is at $N < 10$ or the strong criterion is only broken for $N < 10$, c.f. eq. (4.41).
Although we showed in the previous section that the energy density of the non-Abelian sector is sub-dominant to the Abelian one, we take these bounds in our conservative approach for the numerical constraint. However, one could argue that the sub-dominance of the non-Abelian sector allows to extend the numerical viability even after both criteria are violated. We note that the energy density will however grow with the fluctuations – which means exponentially – such that this consideration would maximally allow an extension of the numerical analysis for a few e-folds. We neglect this possible extension in our conservative approach. The theoretical bound is essentially the same as in the pure non-Abelian case, namely the requirement of sufficient high cut-off of the effective field theory. Here, one might vary the situation a bit by considering a degeneration of the form $\alpha_1 \neq \alpha_2$. We restricted ourselves however to the case where both couplings and cut-offs are the same. In a more dedicated study one could do the parameter scan with this additional degree of freedom. What is expected is that the situation of the EFT cut-off can be circumvented (a bit) for a coupling hierarchy $\alpha_1 \gg \alpha_2$. This is because the cut-off scale $f_c$ can be linearly increased with $\alpha_1$. The stringent coupling relation in the non-Abelian sector then can be easily accounted for in the limit of small $\alpha_2$. However, one still has to obey the bound $\alpha/f_c \lesssim 35$ to be in agreement with CMB observations. So, for our case of equal effective couplings, the theoretical constraints will be the same as in the previous chapter. In figure 5.5 we show the parameter scan. Due to the constantly present Abelian contribution to the dynamics, we expect a much bigger range of possible GW detection. This is confirmed when comparing figure 5.5 with figure 4.7. The GW signal for small gauge group coupling $g \lesssim 7 \times 10^{-4}$ gets pushed to lower values of $\alpha/f_c$ compared to the regime of $g \gtrsim 2 \times 10^{-2}$, because we count the dominant Abelian contribution twice due to the late matching point in the $SU(2)$ sector. The naively expected independence of the gravitational wave signal from the non-Abelian gauge group coupling $g$ cannot be verified. This is for the same reason as in the pure non-Abelian case, namely that the $g$ dependent shifting of the end of inflation results in different values of $d\phi$ at the ET peak sensitivity. Since $\Omega_{GW}$ is highly sensitive to this very quantity, even the 'Abelian spectrum' gets in some sense a dependence on the non-Abelian gauge group coupling. Note that this is not the case for large $g$, since the matching point is pushed to earlier times. This results in the same $d\phi$ for all parameter points in the ET range. The reason can be found in the tiny drop in

\footnote{If, however, the non-linear dynamics reduce the fluctuation strength the analysis can be pushed even farther.}
Figure 5.5: **Left:** The detailed parameter scan for the $U(1) \times SU(2)$ theory. We highlight the vast region of possible GW detection with black. The green and orange region indicate parameter points which are strongly and weakly allowed, respectively. Parameter points under the lowest red contour line are in both cases allowed due to the matching point being smaller than 10. Thus, only parameter points in the green or orange region support the manifest $U(1) \times SU(2)$ dynamic for a few $e$-folds. For details see text. **Right:** For completeness the EFT constraints of the non-Abelian sector.

$\xi$ and the negligible non-Abelian gauge field friction contribution, rapidly pushing the system to its equilibrium dynamics given by the Abelian limit. A new feature of the expected GW signals is that it may produce GW signals strong enough to be detected at lower $\alpha/f_c$ compare to the pure non-Abelian case. So, as it can be extracted from the figure, the dynamic has an unique $U(1) \times SU(2)$ part and thus can be distinguished from a pure Abelian theory – even in our linearized approach. This surprisingly positive result may motivate future studies to investigate this theory in more details. But the phenomenological implication is not the only reason to continue the study. We will point out in the next section the physical naturalness of this unification and why it is actually required for a realistic theory of inflation.

### 5.2 Standard Model Connection

Any theory of inflation essentially needs to provide a stage of reheating afterwards in order to generate the standard hot Big Bang paradigm. The most natural way to achieve this is by direct coupling to the SM. Hence, we interpret the gauge field as a SM gauge field and couple it to massless Dirac fermions. We study the resulting back reaction and show that the fermions do not thermalize in the Abelian case. Other results are expected for non-Abelian fields, as we will point out.
5.2.1 Coupling to Fermions

In the discussion and calculations so far, we assumed that the exponential expansion of the Universe during inflation inhibits any formation of a thermal bath of produced particles. However, due to the tachyonic instability of one of the gauge field modes, this assumption is not trivial. The basic idea of thermalized axion inflation is that the exponential gauge field production enhances the cross section due to high occupation numbers, eventually overcoming the dilution and thus forming a thermal bath [64]. This would mean that the energy distribution is modified, which in turn changes the prediction of the setup. We will justify our non-thermal assumption in the case that the $A_\mu$ belongs to the SM $U(1)$, following ref. [65]. This in turn naturally completes our setup by providing reheating at the end of inflation. The additional ingredient thus will be a massless fermion, which is non-trivially related to the gauge field due to the Adler-Bell-Jackiv anomaly [66, 67]. This is, that the classical conservation of the axial current $J^\mu_5$ is broken by non-perturbative quantum effects leading to $\partial_\mu J^\mu_5 \sim \tilde{F}_{\mu\nu} F^{\mu\nu}$. A chiral rotation of the Lagrangian then demonstrates the necessity of an unified axion inflation model when we want to naturally couple it to the SM anomaly. This requires also the investigation of thermalization in the non-Abelian sector. However, the allowed gauge boson self interactions make this a highly non-trivial task and make the analysis a priori completely different to the pure Abelian case. We leave this as a task for future research and focus here on the Abelian sector and discuss the non-Abelian case briefly at the end.

In general, thermalization is approximately reached when the scattering rate is comparable to or higher than the Hubble rate. In our case the Hubble rate can be simply approximated by $H \sim 10^{-5}$. More complicated to obtain is the scattering rate $\Gamma = n \times \sigma$, where $n$ is the number density and $\sigma$ the interaction cross section. In the following we will discuss how to obtain both quantities for our particular setup.

Let us therefore start by writing down the modified action, including the fermion–gauge field interaction$^2$

$$S = \int d^4 x \sqrt{-g} \left( \frac{R}{2} - \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - V(\phi) - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{\alpha}{4 f_c} \phi F_{\mu\nu} \tilde{F}^{\mu\nu} - i \bar{\psi} i D\psi \right). \quad (5.11)$$

$^2$One may use the conformal invariance of massless fermions and gauge fields to re-define the fields with proper scaling of $a$. These co-moving fields can then be described within Minkowski space-time.
The massless Dirac fermion $\psi$ has the charge $Q$ under the Abelian gauge group with coupling $4\pi\alpha = g^2$. The gauge covariant derivative is defined as $D_\mu \equiv (\partial_\mu - igQA_\mu + \omega^{ab}_\mu \gamma^{ab})$ with the spin connection $\omega$ accounting for the space-time geometry. The Feynman slash for a quantity is defined by $\mathcal{D} \equiv \gamma^\mu D_\mu$. The resulting EOM for $\phi$ is trivially of the same form as discussed before, but the gauge field EOM gets an extra term accounting for the fermion interaction $\delta S \supset -gQ\bar{\psi}\gamma^\mu\psi \delta A_\mu$. This modification of the gauge field dynamic accounts for the back reaction of the fermions. Already at this stage the impact of the fermions on the inflaton dynamics can be understood, even though there is no direct interaction between these two sectors at first glance. The fermion–gauge field interaction modifies the formation of the strong EM field induced by the tachyonic enhanced gauge field mode in the absence of fermions. We may expect already here a (significant) modification of this behaviour due to the non-trivial fermion contribution to the EM field through the back reaction and the acceleration of the charged fermions in the field itself. This can be done more explicit by deriving the EOM for the fermion vector current and axial current respectively given by

$$0 = \partial_\mu (\sqrt{-g}\bar{\psi}\gamma^\mu \psi) \equiv \partial_\mu (\sqrt{-g}J^\mu_\psi),$$
$$0 = \partial_\mu (\sqrt{-g}\bar{\psi}\gamma^\mu \gamma^5 \psi) + \sqrt{-g}Q^2\alpha F\tilde{F} \equiv \partial_\mu (\sqrt{-g}J^\mu_5) + \sqrt{-g}Q^2\alpha F\tilde{F},$$

where we have by definition $J^\mu_\psi/5 = J^\mu_R \pm J^\mu_L$ with the current handedness given by $J_{R/L} = \bar{\psi}\gamma^\mu \mathcal{P}_{R/L}\psi$. The latter axial current equation is in fact a generalization of the ABJ anomaly within an expanding space-time. By conformal re-scaling the dilution factor of $\sqrt{-g}$ can be cancelled, leading to the well known ABJ anomaly [66, 67]. Thus, the axial current equation relates two classically non-equivalent theories in a non-trivial way, such that we may do the replacement

$$S \supset -\int d^4x \sqrt{-g} \frac{\alpha}{\sqrt{f_c}} \phi F\tilde{F} \quad \rightarrow \quad S \supset \int d^4x \frac{\alpha}{2Q^2f_c} \partial_\mu (\sqrt{-g}J^\mu_5).$$

This manifestly connects the inflaton with the fermion through a derivative coupling, which for example allows the process $\phi \rightarrow A_\mu A_\nu$ with an intermediate fermion loop. Note that both expressions can be analysed in Minkowski space after proper conformal re-scaling, which may facilitate the analysis.

We may illustrate the qualitative behaviour of the system by a quick analysis of
the energy density of each component

\[
\dot{\rho}_\phi + 3H\langle \dot{\phi}^2 \rangle = -\frac{\alpha \dot{\phi}}{f_c} \langle E \cdot B \rangle, \tag{5.15}
\]

\[
\dot{\rho}_A + 4H\rho_A = \frac{\alpha \phi}{f_c} \langle E \cdot B \rangle - \langle E \cdot (gQJ_\psi) \rangle, \tag{5.16}
\]

\[
\dot{\rho}_\psi + 4H\rho_\psi = \langle E \cdot (gQJ_\psi) \rangle. \tag{5.17}
\]

Obviously, we have an energy transfer from the inflaton into the gauge field in the same way as discussed in the previous chapters. However, the additional term arising in the gauge field sector, indicates a negative energy transfer from the gauge field to the fermion. The fermion itself trivially gains this amount of energy. But note, that in order not to run into any contradictions, we must strictly demand \(\dot{\phi}\langle E \cdot B \rangle > 0\).

So, we found a mechanism that produces fermions in an anti-symmetric way due to the anomaly of the axial current. Additionally to this process, we generally have to expect efficient fermion pair production in the presence of (strong) EM background fields, known as Schwinger effect [68]. This does not contribute to the anomaly relation, since it implies equal creation of left- and right-handed fermions and hence \(\partial_\mu J^\mu_5 = 0\) for this process. However, both processes may be related as one can understand when studying the EOM for the left- and right-handed fermions

\[
0 = (i\partial_r \pm i\partial_i \cdot \sigma \cdot gQA_0 \pm gQA_i \cdot \sigma^i)\psi_{R/L}, \tag{5.18}
\]

where \(\sigma\) denote the Pauli matrices. A detailed investigation of the fermion production, however, requires exact information of all correlators over the complete course of inflation. This complex situation may be approximated by considering one Hubble patch at a given time \(t\) together with the standard EM field correlators used throughout this work. Then one can study the fermion production inside each patch. This non-trivial viewpoint can be justified a posteriori when we find that i) the production rate is faster than the Hubble rate and ii) the resulting fermions do not thermalize. Note that the second point implies the first one but not vice versa. However, it is much easier to show the first one, which we will do now. The second one will be discussed in the next chapter. Anticipating the results, we find both assumptions to be justified, thereby confirming the validity of the analysis.

With this ansatz, one can find that that the EOM (5.18) leads to a discretization
of the dispersion relation in the presence of EM fields [65], known as Landau levels. The higher Landau levels (HLLs) account for the pair creation of particles through quantum tunneling between different levels. This is because the structure of the HLLs are independent from the handedness of the fermion, leading to a symmetric production of both left- and right-handed fermions. Thus, the fermion production process related to the HLLs cannot account for the axial current anomaly. Rather, it can be interpreted as the aforementioned Schwinger effect. Contrary, the lowest Landau level (LLL) exhibits a sign change in the dispersion relation for left- and right-handed fermion such that $\omega_{\psi_R} = -\omega_{\psi_L}$. This is because one may regard the fermions in the LLL to move along with the magnetic field. So, the fermion spin must be parallel (anti-parallel) to the field line for the left-handed (right-handed) fermions, accounting for the sign change in the dispersion relation. This difference in the fermion spectrum has important implications. Calculating the induced left- and right-handed currents, one can directly (re-) derive the equation for the axial current non-conservation. Thus, the LLL accounts for the production of the chiral charge. Reference [65] explicitly derived the resulting fermion production rates for the $n$ Landau levels to be

$$
\dot{n}_\psi^{\text{LLL}} = 2 \times \frac{g^2 Q^2}{4\pi^2} EB, \quad (5.19)
$$

$$
\dot{n}_\psi^{\text{HLL}} = 4 \times \sum_{n=1}^{\infty} \frac{g^2 Q^2}{4\pi^2} EBe^{-\frac{2\pi n B}{E}} = 4 \times \frac{g^2 Q^2}{4\pi^2} EB \frac{1}{e^{2\pi B/E} - 1}. \quad (5.20)
$$

This makes the relation between axial current conserving and non-conserving process clear. The latter can be obtained as the case of $n = 0$ from the first one. A first evidence that the approximation used in the discussion and derivation of the equation (see below eq. (5.18)) can be prompt delivered by comparing it to $H^4$. For the natural values of the gauge coupling at SM GUT-scale $g = 1/\sqrt{2}$ and $Q = 1$, we find the a priori assumption i) to be violated only for values of $\xi \lesssim 2.5$. This is below our scale of interest (recall that $\xi$ grows over the course of inflation and is at CMB scales around 2). The justification of point ii) will be delivered in the next section. But to this end, we need to investigate the fermion–gauge field back reaction in more detail.

It is now important to clarify the exact value of the EM fields, taking into account the discussed fermion–gauge field interaction. As we already roughly qualitatively analyzed, we expect the induced fermion current to play an important role for the dynamics
through its back reaction and acceleration in the EM field itself. We can make this statement more explicit by imposing a fundamental consistency condition which the setup has to fulfill at all times. Let us consider the energy transfer from the gauge fields to the fermions explicitly, given by eq. (5.16). Since the fermion term comes with a minus sign, the energy of the gauge field sector gets reduced. However, the gauge fields will receive simultaneously an energy boost coming from the inflaton sector. The situation may be understood intuitively, when we investigate two limits of the configuration. The first one is trivial, if we assume that the fermion do not pull significant energy out of the gauge field sector. Then the natural growth of the EM fields is not disturbed, and there will be no further bound worth to consider. Thinking the other way around, namely a significant energy extraction of the fermions out of the gauge field sector, leads to dramatic modifications. In this case, the EM fields will be reduced in their strength due to the non-neglectable energy loss to the fermions. However, this necessarily leads at the same time to a decreasing fermion production. One can understand, that the direct coupling of these two systems than must lead to an equilibrium configuration where both, fermion and gauge field production, reach a stable attractor solution. So this qualitative analysis reveals, that we cannot circumvent the direct calculation of the induced fermion current as a function of the EM fields, resulting in

\[
gQ \langle J_\psi \rangle \simeq (g|Q|)^3 6\pi^2 \coth \left( \frac{\pi B}{E} \right) (E \cdot B) H^{-1}. \tag{5.21}
\]

Let us analyse how this result affects the EM field. By plugging this expression into the energy density evolution equation for the gauge fields from eq. (5.16), we obtain an upper bound on the EM field strength given by the solution of

\[
0 = -2H(E^2 + B^2) + 2\xi_{\text{eff}} HEB, \tag{5.22}
\]

where the new appearing effective \(\xi\) parameter is given by

\[
\xi_{\text{eff}} \equiv \xi - \frac{(g|Q|)^3}{12\pi^2} \coth \left( \frac{\pi B}{E} \right) EH^{-2}. \tag{5.23}
\]

In figure 5.6 we show solutions to this equation for different values of \(\xi\). The exponential

---

\[\text{Footnote:} \quad \text{Another consistency condition is to postulate that the gauge field cannot acquire more energy than the inflaton itself has, c.f. eq. (5.15). Anticipating the result of the consistent EM fields, we find only for } \xi \gtrsim 20 \text{ a violation of this condition. Hence, the energy conservation does not significantly restrict the formation of EM fields for configurations of the system which obey current CMB bounds.} \]
growth of the free EM fields are strongly damped due to the presence of the fermions. So, in practical terms, we would like to extract the maximally allowed values for the EM field components as a function of $\xi$ and compare it to the free EM value. This is equivalent to finding for fixed $\xi$ the maximal value of the electric and magnetic field, such that the consistency relation (5.22) still holds. We denote this as the consistent EM field. The result is also shown in figure 5.6, for the independent maximum of electric and magnetic field.

However, one may notice that we also neglected the to fermions inherent property that any two of them cannot posses the exact same quantum characteristics. That is, the possibility of Pauli blocking in the HLLs, whereas we do not have to worry about the LLL due to its asymmetry. With the obtained EM field we can safely show that the criterion for efficient Pauli blocking $g Q E t_p \gg \lambda_c$, with the characteristic fermion production time $t_p = \dot{n}_\psi^{-1} \lambda_c^{-3}$ and the Compton wavelength $\lambda_c$, is never reached.

5.2.2 Thermalization

We are now equipped with all necessary tools to tackle the question of thermalization in the fermion sector. We assumed throughout the analysis that the scattering rate is suppressed and we will now justify this. Recall, that the scattering rate among particles is generally given by $\Gamma_{sc} = \tau_{sc}^{-1} = n \times \sigma_{sc}$, where we introduced the time between two
scattering events $\tau_{sc}$. The scattering rate for two fermions exchanging a massless gauge boson is well known and given by $\sigma_{sc} \sim 2\alpha^2 / \pi s$ [69]. The center of mass energy can be obtained by the acceleration of the fermions within the EM field and their transverse energy, leading to $\sqrt{s} = 2(p_T^2 + gQEB\tau_{sc})$. It is clear that we have to distinguish the cases of LLL and HLL, since the transverse energy is only acquired through the level splitting, giving by the discretization of the fermion energy dispersion. So, in the LLL case we have $p_T^2 = 0$ and for the HLL the dispersion relation yields $p_T^2 = 2ng|Q|B$ [65]. On the other hand, the number density of the fermions has been lengthy discussed and is given by $n_\psi = \hat{n}_\psi \times T$, where $T$ denotes the time of continuous fermion production. We may put an upper bound on this by assuming $T = H^{-1}$. Putting all the pieces together we first estimate the scattering rate in the LLL to be

$$\Gamma_{sc}^{-1} = \tau_{sc} \simeq \frac{\alpha^2 B}{2\pi E} H^{-1}. \quad (5.24)$$

Extracting from fig. 5.6 a rough estimate of $B/E \lesssim 10$, we obtain $\Gamma_{sc} \sim H \times (10/g^2)^2$. So, for all values of $g$ which are in accordance with the requirement of perturbative control we see that the LLL fermions thermalize. This indicates that the analysis of the LLL is not justified.

Let us turn to the HLL. The situation here gets modified due to the discussed level splitting. This means, that we have to take into account the transverse energy of the fermions. Since it is level dependent, we have to work with the number density specific to each of the Landau levels, given by the first equality in eq. (5.20). Then we have to sum over all levels, which qualitatively looks like

$$\Gamma_{sc} \supset \sum_{n=1}^{\infty} \left[ (An + B\Gamma_{sc}^2)^{-1} \times e^{-2\pi n B/E} \right], \quad (5.25)$$

where $A, B$ are dimension fixing variables. We see that higher levels are coming with an exponential suppression, which allows us to estimate the scattering rate roughly only with the lowest level $n = 1$. This gives us $\Gamma_{sc} \sim H \times 10^{-5}$, indicating that HLL fermions indeed do not thermalize. In figure 5.7 we plot the (nearly) full HLL scattering rate after summing over the first 50 levels. For completeness we also show the LLL rate.

Apparently, both fermion production rates lead to different results concerning the thermalization. To decide if the fermion sector indeed thermalizes or not, we have to check the distribution of the energy density among the LLL and HLLs. The result can
Figure 5.7: **Left:** Scattering rates for the LLL and the first fifty HLLs. Thermalization is reached in the LLL, but not in the HLLs. For completeness we also show the result which one would obtain without the consistency relation imposed for the gauge field sector. **Right:** Distribution of the energy density among the different Landau levels. The majority of the energy density is stored within the HLLs. Thus, fermions which obey the consistency relation do not thermalize.

In case of consistent EM fields the HLLs dominates the energy distribution within the fermion sector. Thus, only a minority of the fermion phase space thermalizes. The overwhelming main contribution to the fermion current comes from the HLLs, which are dominated by the acceleration in the EM field and not by thermal motion. Hence, it is valid to conclude that the fermion sector does not form a thermal bath, which justifies the assumptions made to reach the result a posteriori. Note that the gauge field sector has to satisfy the imposed consistency condition, which basically damp the formation of strong EM fields. Without setting this upper bound, we would arrive at a contrary conclusion, namely that fermions start to thermalize towards the end of inflation.

In the case of non-Abelian gauge groups present during the inflationary evolution, the situation is yet again different. The gauge boson self-interaction drastically modifies the scenario and it is not trivial to clarify the net impact for the thermalization process. But what we expect is that the scattering efficiency is high enough to lead to a complete thermalization. On the other hand, we discussed around figure 5.2 that the non-Abelian gauge fields do not have a significant energy contribution compared to the vacuum. So, probably there will be no considerable fermion production within this model. In a realistic scenario, where we couple to the ABJ anomaly, we need the presence of both Abelian and non-Abelian gauge groups. Within the model of $U(1) \times SU(2)$ coupled to the fermions it is thus not clear how the energy is distributed among the different
5.2.3 Phenomenology

Let us briefly analyse the phenomenological implications of the setup with the Abelian gauge field coupled to fermions. The main change in the inflaton evolution arises due to the fermion induced upper bound on the EM fields. This modification reduces the effective friction term in the inflaton motion. So, we do expect the delay of the end of inflation compared to the vacuum dynamics to be relaxed compared to the case without fermions. Also, we expect that the predictions for e.g. the scalar power spectrum and gravitational wave production significantly change.

In principle, we already derived the EOM for the inflaton lengthy in the previous chapters. We only have to do the proper replacement for the EM field configuration, now in the presence of fermions. Thus, the only change in the EOM will be the change of $\langle E \cdot B \rangle$ from the free field to the consistent field with the fermion induced upper bound in eq. (3.13). This is valid since we work in the decoupling limit between inflaton and gauge field sector. We show the results for $\phi$ and $\xi$ in figure 5.8, for the case of natural inflation. Model constants in accordance with Planck data are given in the figure caption. We clearly see that our prediction of a slow roll alike inflaton dynamics is verified. The main difference is that the $\xi$ parameter does not receive an upper bound induced by the gauge field friction as in the pure Abelian case, c.f. also figure 3.2. This is because in the presence of fermions the exponential friction term induced by the EM

![Figure 5.8: Inflaton evolution for the three here relevant cases of i) vacuum, ii) presence of Abelian gauge fields and iii) presence of Abelian gauge fields & Dirac fermions. Model constants are in accordance with Planck data, $\alpha/f_c = 35$, $\Lambda^4 = 4.75 \times 10^{-9}$ and $f_V = 9.2$.](image-url)
field is regulated to a power-law dependence. This also implies that we escape the strong back reaction regime between the gauge field and the inflaton, which is present towards the end of inflation in the setup without fermions [51]. Hence, we expect less theoretical uncertainties for the inflaton calculation in the case with fermions and gauge fields than with only gauge fields.

The next step is to make contact to direct observational quantities like the GW amplitude or the scalar power spectrum, which is related to the discussed $\mu$-type distortions. The latter can be well approximated with the formula given in eq. (3.29), whereas the former one requires more detailed knowledge about the contribution from the fermion sector via gravitational anomaly [70]. Note again, that for using the formula of eq. (3.29) we have to replace the EM field with its consistent upper bound. One finds an enhanced scalar power spectrum at small to intermediate scales. However, the inefficient friction towards the end of inflation causes a drop in the power spectrum as $N \to 0$. This peaked structure of the power spectrum may lead to a relaxation of the light primordial black hole bound, while heavier ones still may populate the phase space efficiently, contribution thereby significantly to the dark matter density [65]. So, a more detailed phenomenological study of this setup may reveal very interesting features.
Chapter 6

Discussion

Summary

We investigated an inflationary phase in the early Universe driven by a pseudo Nambu-Goldstone boson (pNGB). We first studied the setup for generic pNGB models and derived the equations of motion (EOMs) in the case for its standard vacuum dynamics as well as in the presence of an additional coupling to an Abelian Chern-Simons (CS) term. The latter one is known to leave interesting phenomenological imprints \cite{25, 29, 48}, as for example sufficient gravitational wave (GW) production detectable by future experiments like the Einstein Telescope (ET). Therefore, any non-detection puts strong constraints on the setup. We, however, put our focus on the mostly unnoticed constraint on distortions of the measured nearly perfect temperature black body distribution of the cosmic microwave background (CMB).

These so called $\mu$-type distortions may arise due to energy release in the early Universe at redshift $10^4 \lesssim z \lesssim 10^6$ \cite{33, 34}, naturally achieved within the last 60 e-folds by a pNGB inflaton coupled to gauge fields. The so developed chemical potential $\mu$ leads to deviations from the perfect black body distribution. In general, this is sensitive to the integrated power spectrum of both, scalar and tensor perturbations. We used an ansatz which allows us to separate the thermalization physics from the model dependent power spectra \cite{33, 34}. Equipped with this, our results are twofold. First, we showed that the tensor contribution is suppressed by some six orders of magnitude. Second, the scalar contribution may play an important role in constraining inflationary models. Future experiments like the Primordial Inflation Explorer (PIXIE) \cite{50} may put more stringent
bounds than the famous frequently studied small non-Gaussianity constraint \([46, 71–80]\). We demonstrated this with a detailed parameter scan in the case that the pNGB acquires a Starobinsky potential and is coupled to Abelian gauge fields.

Then we moved on to study the theory of a pNGB coupled to non-Abelian gauge fields, again through a CS term. Here, we focused on the natural interpretation that the pNGB is an axion-like particle and studied the implications of this interpretation. This setup is known as chromo-natural inflation (CNI) \([36–40]\). We discussed that when we start from the quantum mechanical Bunch-Davies vacuum, it reduces to the Abelian case in the far past, which we call Abelian natural inflation (ANI) \([48]\). This is because the inherent non-Abelian background field dynamically evolves over the course of inflation, starting from a zero vacuum expectation value (vev). The transition to a non-zero vev is discussed in the linearized approach, a posteriori connecting it to the gauge field fluctuations.

Within this linearized framework the background vev transition is instantaneous. We determined the time when this should happen, following ref. \([52]\). In the slow roll limit the EOMs for the gauge field and the inflaton completely decouple and the gauge field background may even be solved analytically. We verified the analytical approximation by solving the coupled EOMs directly. The linearized approach we followed puts bounds on the reliability of the obtained results, especially when making contact to observational quantities. Therefore, we provided a detailed parameter scan, demonstrating the numerical limits. This revealed that only a small region with inherent non-Abelian dynamics produce GWs strong enough to be detectable with ET. Interestingly, the magnetic drift regime studied in literature \([24, 36–38, 40]\) where the non-Abelian induced gauge field friction dominates over the Hubble friction can practically only be related to GW signal arising from the Abelian limit of the theory. Thus, in this regime the theory of CNI cannot be distinguished from ANI. However, far more important are the derived theoretical bounds.

We were able to derive very tight constraints for the parameter space within the prescription of an effective field theory (EFT). To be in agreement with Planck data \([1]\), we verified that the periodicity of the potential must be super-Planckian. The super-Planckian periodicity leads us to the necessity to distinguish it from the EFT cut-off scale entering through the CS coupling. We find that the de-Sitter Gibbons-Hawking temperature is not always low enough to guarantee the existence of the axion-like particle in such a setup. However, the main constraint arises due to the requirement that
the axion is the lightest particle in order to integrate our other degrees of freedoms. This leads to an enforcement of large gauge couplings, almost entirely excluding the parameter space.

Interestingly, the tiny window which is untouched by these constraints never undergoes a non-trivial inherent non-Abelian evolution within the linearized approach. Also we do not expect detectable GW signals from this region of parameter space. Additionally, we point out that the ratio between the EFT cut-off and the periodicity within this window is around $10^4$, which may be beyond string theoretical realizations [81]. If we on top require the natural scale ordering from the original axion model of Peccei & Quinn [58], we find that practically no window of CNI survives the EFT bounds.

For completing the (chromo)-natural inflation setup, we considered the physical more realistic scenario of the presence of both, Abelian and non-Abelian gauge groups. Within this unified framework the aforementioned EFT bounds may be avoided if the axion-like particle couples to the Abelian fields (much) stronger than to the non-Abelian ones. We, however, studied exemplary the case of equal coupling and motivate future studies for a more detailed analysis. We showed that the setup may leave distinct GW signatures, although the window is somewhat tiny.

If we want to successfully couple the dark inflation sector with the Standard Model (SM) through the Adler-Bell-Jackiv (ABJ) anomaly [66, 67], the introduced unification is actually needed. This will also naturally provide re-heating of the Universe after the end of inflation. We discussed the discretization of the fermion energy into Landau levels for the Abelian case, leading to Schwinger pair production for the higher levels and fermion production related to the ABJ anomaly [65]. This efficient fermion production avoids any formation of a thermal bath due to damping of the natural gauge field growth. The phenomenology of the setup is shortly discussed and its interesting features are emphasized.

**Outlook**

We hope that the work motivates further studies on some of the topics in more detail. Altogether, there are many questions unanswered and things we pointed out that deserve a careful treatment. This begins with the task to control the strong back reaction regime of the gauge field—pNGB model, present when the slow roll condition
is (strongly) violated. This typically happens towards the end of inflation. Recently a work on this topic has been published, considering the coupling to Abelian gauge fields [51]. Therein, the authors concentrated on the last e-folds of inflation, focusing on the reheating efficiency of the setup. One may use this as a starting point to also study a wider range of inflationary e-folds and the phenomenology. Then, the next natural step would be the extension to non-Abelian gauge groups, which is however highly non-trivial. But it would be at some point necessary to keep track of the non-linear contribution to the gauge fields in the CNI model. This would allow to extend the numerical viability to the whole studied parameter space, thereby alleviating the numerical restriction we had to face within this work. Also, we highly encourage to study the EFT description problem of (chromo-) natural inflation. In particular how to realize super-Planckian periodicity for the potential. We note that this is already a field of very active research in the context of string theory [82–94]. Our derivation of the theoretical bounds on the EFT setup definitely deserves some attention to make it more explicit.
Appendix A

Window Function

For the scalar window function we took the definitions from [33], whereas for the tensor window function we worked with [34]. They are respectively given by

\begin{align}
W_\zeta(k) &= 1.4 \times \frac{64}{45} c^2 D^2 k^2 \int_{z_{\text{rec}}}^\infty dz \frac{(1 + z)}{d(z) H(z)} J_\mu(z) \sin(\kappa r_s(z)) e^{-2k^2/k_D^2(z)} , \\
W_t(k) &= 1.4 \times \frac{4}{45} \int_{z_{\text{rec}}}^\infty dz \frac{H(z)}{d(z)(1 + z)} T_h(k\chi) T_\theta \left(\frac{(1 + z)ck}{d}\right) e^{-\Gamma\gamma\chi} J_\mu(z).
\end{align}

For the scalar and tensor window function, we integrate from the Big Bang to the time of recombination \( z_{\text{rec}} = 10^3 \). The first constant factor arises due to the mode specific efficiency, approximated by \( D = (1 + 4R_\nu/15)^{-1} \), where the fractional contribution of massless neutrinos to the energy density of relativistic species is \( R_\nu = \rho_\nu/\rho_\gamma + \rho_\nu \simeq 0.41 \). The time derivative of the Thompson optical depth is given by \( \dot{d}(z) = 4.4 \times 10^{-21}(1 + z)^2 \frac{1}{s} \). And the sound horizon is \( r_s(z) = \frac{\chi}{\sqrt{3}} \). If we neglect baryon loading, the cut-off scale of the scalar window function is determined by the photon damping scale in the radiation dominated epoch by

\begin{align}
k_D(z) &= \int_{z_{\text{rec}}}^\infty dz \frac{8c^2(1 + z)}{45d(z)H(z)} \sim 4 \times 10^{-6}(1 + z)^{3/2} \text{ Mpc}^{-1}.
\end{align}

Approximating further, that the distortion shape only depends on the total number of Compton scattering events but not on the anisotropies itself, one gets for the distortion...
visibility function

\[ J_{\mu}(z) = e^{-(z/z_{dc})^{5/2}} \times \left( 1 - e^{-\left(\frac{1.88}{z_{dc}}\times 10^5\right)} \right). \]  \hfill (A.4)

It accounts for the effects and efficiency of thermalization at early times, erasing the distortions for large enough redshift \( z \geq z_{dc} = 2 \times 10^6 \). For the tensor window function we have the tensor transfer function

\[ T_h(x) = 2 \left( \sum_{n=0}^{6} a_n (n j_n(x) - x j_{n+1}(x)) \right)^2, \]  \hfill (A.5)

where \( j_n(x) \) denotes the spherical Bessel function and the prefactors are given by \( a_0 = 1 \), \( a_2 = 0.243807 \), \( a_4 = 5.28424 \times 10^{-2} \), \( a_6 = 6.13545 \times 10^{-3} \). However, this function does not account for polarization states or higher \( l \) modes, and one has to introduce an improved tight coupling approximation for \( l_{\text{max}} = 20 \)

\[ T_{\theta}(x) = \frac{91x^2 + 4.48x + 1}{55x^4 + 100x^3 + 90.2x^2 + 4.64x + 1}. \]  \hfill (A.6)

The numerical most demanding part is the scale dependent damping coefficient, determined by

\[ \partial_t \left( \Gamma_{\gamma}^\ast \chi \right) = \frac{32H^2(1 - R_{\nu})T_{\theta}(k/d')}{15d'}, \]  \hfill (A.7)

with \( d' = \dot{d}a/c \). This factor is for our purpose of a relatively fast calculation not feasible and we approximate it. We fix \( k \equiv 1 \), s.t. the function looses its dependence on \( k \).

Due to the exponential suppression of this function in the tensor window function, the deviation should be small for the \( k \) range of interest. Thus we did our calculation with

\[ \Gamma_{\gamma}^\ast \chi(z, k) \rightarrow \Gamma_{\gamma}^\ast \chi(z, k = 1) \equiv \Gamma_{\gamma} \chi(z) \]  \hfill (A.8)

We checked the deviation for some random \( k \) values and found practically no difference, but never ran our code for the whole range with the full function.

\(^1\)Note here, that we took this definition from paper \([33]\). It changed w.r.t. \([34]\) by adding the second term.
Appendix B

Comments on the Planck Cosmological Parameters

B.1 Energy Densities

For the numerical calculations in chapter 3 we used the Planck 2015 results [3]. Note, however, that the change in the 2018 results [1] is somewhat tiny such that our results are not affected. The matter density and dark energy density is given by

\[ \Omega_m = 0.3089 \pm 0.0062, \quad \Omega_\Lambda = 0.6911 \pm 0.0062. \]  

(B.1)

We set the curvature contribution to zero, i.e. \( \Omega_k = 0 \), which is in perfect agreement with the measurement. The Hubble parameter today is given by

\[ H_0 = 67.74 \pm 0.46 \text{ km/s/Mpc}. \]  

(B.2)

The relativistic energy density is not explicitly given, but has to be derived starting from Planck’s law for the Photon number density

\[ n_0(\nu)d\nu = \frac{8\pi\nu^2d\nu}{\exp\left(\frac{h\nu}{k_B T_0}\right) - 1}, \]  

(B.3)
where \( \nu \) denotes the frequency and we changed back to SI units. The Photon energy density is then given by

\[
\rho_{\gamma,0} = \int_0^\infty d\nu \, n(\nu) \nu h/c^2 = \frac{8\pi^5 k_B^4}{15c^5 h^3} \times T_0^4. \tag{B.4}
\]

With the COBE/FIRAS measurement of the today’s Photon temperature \( T_0 = 2.72549 \pm 0.00057 \) K \(^\text{[95]}\), we get

\[
\rho_{\gamma,0} = 4.64511 \times 10^{-31} \, \text{kg/m}^3. \tag{B.5}
\]

Also Neutrinos and other dark radiation - we will denote them both together as \( \nu \) - contribute to the total relativistic energy density. Its impact can be calculated from the Photon energy density as

\[
\rho_{\nu,0} = N_{\text{eff}} \left( \frac{7}{8} \right) \left( \frac{4}{11} \right)^{4/3} \times \rho_{\gamma,0}, \tag{B.6}
\]

where \( N_{\text{eff}} = 3.04 \pm 0.33. \) Adding them both up and express the result in terms of the critical energy density \( \rho_{\text{crit}} = 3H_0^2/(8\pi G_{\text{Newton}}) \), we end up with the mean value

\[
\Omega_{r,0} = 9.12 \times 10^{-5}. \tag{B.7}
\]

### B.2 Uncertainty in the required Inflation Time

The general picture of inflation is that co-moving scales exit the horizon during the inflationary phase of our Universe due to the decreasing co-moving horizon. After the end of inflation, the strong energy condition is no longer violated and we have \( \omega > -1/3. \)

This in turn means that the co-moving horizon is described by its standard description, meaning that it monotonically increases over time. So, scales which exited the horizon at inflation will eventually re-enter the horizon\(^1\).

However, there is an uncertainty regarding the exact time when the scales of the CMB measurement left the horizon during inflation. This is because we cannot exactly track the evolution of the horizon, since it is generally model dependent. An approximate formula for the number of e-folds needed to explain the CMB scale measurement is given

\(^1\)Note that the presence of the measured dark energy content of our present Universe reverses the horizon behaviour again. But here we are interest in times before the CMB, so this does not play a role for our analysis.
by [1, 96]

\[ N_* \simeq 67 - \ln \left( \frac{k_*}{H_0} \right) + \frac{1}{4} \ln \left( \frac{V_*^2}{\rho_{\text{end}}} \right) + \frac{1}{12} \left( 1 - 3 \omega_{\text{int}} \right) \ln \left( \frac{\rho_{\text{th}}}{\rho_{\text{end}}} \right) - \frac{1}{12} \ln (g_{\text{th}}), \]  

(B.8)

where the subscript \( \star \) denotes that the corresponding quantity has to be evaluated at horizon exit. Furthermore, we directly see the complication to obtain the exact CMB horizon re-entry time due to the necessary ingredient of the energy density at reheating/thermalization \( \rho_{\text{th}} \), together with the effective equation of state between the end of inflation and reheating \( \omega_{\text{int}} \), which is usually approximated like \( 0 \leq \omega_{\text{int}} \leq 1/6 \). Also we see a dependence on the effective number of bosonic degrees of freedom at the reheating time \( g_{\text{th}} \) and the energy density at the end of inflation, denoted by \( \rho_{\text{end}} \). However, since all those quantities are logarithmically weighted, we do not expect a deviation from the reference number 67 more than of \( \mathcal{O}(1) \). And indeed, reference [96] derived a maximum number of e-folds under the assumption that there is no (significant) drop in the energy density towards the end of inflation

\[ N_{\text{CMB}}^{\max} \simeq 63.3 + \frac{1}{4} \ln(8\pi\epsilon_V). \]  

(B.9)

The Planck collaboration measured [1]

\[ \epsilon_1 < 0.0063, \quad \epsilon_2 = 0.030^{+0.007}_{-0.005}. \]  

(B.10)

The horizon flow functions are defined as \( \epsilon_1 = -\dot{H}/H^2 \) and the higher orders \( \epsilon_{n+1} = \epsilon_n/(He_n) \). For the slow roll potential parameter it is sufficient to restrict the analysis to linear level (in case of a scalar inflaton), and the general dependency is given by [97]

\[ \epsilon_V = \epsilon_1 \frac{(1 - \epsilon_1/3 + \epsilon_2/6)^2}{(1 - \epsilon_1/3)^2}. \]  

(B.11)

So at the end one obtains a maximum for the slow roll potential parameter for the scales exiting the horizon during inflation which correspond to the measured CMB scales \( \epsilon_V < 0.0063 \). This translates into an (double) upper bound\(^2\) for the e-fold number

\(^2\)The terminology of a double upper bound comes from the case that the e-fold number \( N_{\text{CMB}} \) already is a theoretical upper bound on \( N \). Together with the measured bound on \( \epsilon_V \) the bound comes from experiment and theory.
corresponding to the CMB scale

\[ N_{\text{CMB}}^{\text{max}} < 62.84. \]  

We note that this upper bound is not really strong and can be easily violated by different mechanisms, like energy scale reduction during inflation or non-instantaneous phase transitions. However, one naturally assumes \( 50 \lesssim N_{\text{CMB}} \lesssim 60 \). In figure B.1 we illustrate the described setup and additionally show how this translates into an uncertainty for the inflationary e-fold for which the GW production should be evaluated. Necessary to understand this is the fixing relation between frequency today and e-fold of horizon exit at inflation, quoted in the main text. Generally the dependence is

\[ e^{N_{\text{CMB}} - N} = \frac{f_{\text{CMB}}}{f}. \]  

By choosing a particular reference value for \( f_{\text{CMB}} \) and demanding that the prior equation holds for the pivot scale \( k_{\text{CMB}} = 0.05 \text{ Mpc}^{-1} \) of the Planck collaboration [1, 5] we arrive

\[ \text{Figure B.1: Left:} \text{ We indicate the uncertainty for the exact time when density fluctuations re-enter the Hubble horizon after inflation. This translates into an uncertainty in terms of } N_\star \text{ for co-moving scales. This arises mainly due to the unknown physics of pre- and reheating after the inflationary condition } \epsilon < 1 \text{ is violated. The red dot correspond the the horizon exit of a particular fluctuation and both green dots show the model dependence on the re-entry time. Here we refer with co-moving scales to the co-moving wavelength. Right: The aforementioned uncertainty directly translates into an uncertainty in terms of which time phase of the inflationary dynamic can be probed by (future) GW experiments. This is because the relation between the e-fold of horizon exit during inflation and the GW frequency corresponding to this today is linearly dependent on the required CMB e-fold time } N_{\text{CMB}}. \text{ We particular highlight the e-fold which the ET would test, for which we know its peak sensitivity to be around } f = 10 \text{ Hz. The plot is made with the pivot scale of } k_{\text{CMB}} = 0.05 \text{ Mpc}^{-1}. \]
at

\[ N = N_{\text{CMB}} - \log \left( \frac{f}{\text{Hz}} \right) + \log \left( \frac{0.05 \text{ Mpc}^{-1}}{2\pi \text{ Hz}} \times c \right), \quad (B.14) \]

where we restored the \( c \) factor to make the logarithm manifest dimensionless. The last log can be approximated by \(-37.10\). So we see that for a given frequency \( f \) there is a linear uncertainty in the corresponding e-fold due to the \( N_{\text{CMB}} \) uncertainty. Thus, the usual requirement of \( 50 \lesssim N_{\text{CMB}} \lesssim 60 \) translates into an e-fold range for fixed frequency. In the case of the ET sensitivity maximum this corresponds to a range of \( 10 \lesssim N_{\text{ET}} \lesssim 20 \). We also note that it is not clear which is the 'right' pivot scale to use. Part of the measurements of the Planck collaboration and especially the WMAP measurements have been made at the pivot scale \( k_{\text{CMB}} = 0.002 \text{ Mpc}^{-1} [2, 98, 99] \), such that the log factor changes to \(-40.32\). Including the ambiguousness for the pivot scale then results in the total uncertainty for the ET e-fold range

\[ 7 \lesssim N_{\text{ET}} \lesssim 20. \quad (B.15) \]

We then decided to fix \( N_{\text{ET}} = 11 \) for the calculations in the chapters 4 and 5.

Finally, note that we work within the standard framework of Einstein’s general theory of relativity. In other models, e.g. Palatini gravity – where the metric is completely unrelated to the affine connection – one obtains a completely different pre- and reheating dynamics for the same model [100]. This contributes also to the completely unknown dynamics after the violation of the inflationary slow roll parameter. This uncertainty can be covered in the cloak of the uncertainty of the exact \( N_{\text{CMB}} \) value.
Appendix C

Analytic Approximations

In this appendix we review the most important analytical approximations we used in the main text. This concerns the solution for gauge field fluctuations in the Abelian as well as non-Abelian case. Also we comment on the expectation values for the electric and magnetic fields induced in the Abelian dynamic. We especially highlight the numerical uncertainty due to the needed integral cut-off caused by the infinite vacuum energy contribution. Finally, we also show analytically that the required periodicity scale of the axion potential has to be super-Planckian to be in accordance with the Planck data.

C.1 The Gauge Field

Let us focus on the Abelian case of axion inflation, since we rigorously showed in the main text that even in the dynamical evolving non-Abelian case, the Abelian limit is accurate for nearly the entire inflationary phase. Also, we note that the given inherently non-Abelian gauge field fluctuation differential equation is of the same form and thus can easily be solved with the same method.

Given the real space ordinary differential equation (ODE) (3.12) we may proceed by decomposing the gauge field into its Fourier modes

\[ \mathbf{A}(\tau, \mathbf{x}) = \sum_{\lambda = \pm} \int \frac{d^3 k}{(2\pi)^{3/2}} \left( a_\lambda(k) \epsilon_\lambda(k) A_\lambda(\tau, \mathbf{k}) e^{i k \mathbf{x}} + \text{h.c.} \right). \]  

\[ \text{(C.1)} \]
We impose the standard commutation rule for the creation and annihilation operators

\[ [a_\lambda(k), a^\dagger_{\lambda'}(k')] = \delta_{\lambda,\lambda'} \delta(k - k'). \]  

(C.2)

The polarization vectors satisfy the following relations

\[ k \cdot \epsilon_\lambda(k) = 0, \quad k \times \epsilon_\lambda(k) = (-\lambda) i k \epsilon_\lambda(k), \quad \epsilon_\lambda(-k) = \epsilon_\lambda^*(k), \quad \epsilon_\lambda(k) \cdot \epsilon_{\lambda'}(k) = \delta_{\lambda,\lambda'}. \]

With this we arrive at the Fourier space ODE for the gauge field

\[ \left( \partial^2_x + k^2 \pm k \alpha \frac{\dot{\phi}}{H f_c} \right) A_{\pm}(\tau, k) = 0. \]  

(C.3)

Introducing the dimensionless variable \( x \equiv -k\tau \) and the gauge field coefficient \( \omega(x) = \sqrt{2kA} \) we can rewrite the ODE as

\[ \left( \partial^2_x + 1 \pm \frac{2\xi}{x} \right) \omega_{\pm}(x) = 0, \]  

(C.4)

where we used the manifest positive quantity \( 2H \xi = |\dot{\phi}|/f_c \simeq \text{const.} \) and de-Sitter space approximation. The minus mode gets tachyonically enhanced when \( m^2_- = 1 - 2\xi/x < 0 \), whereas this will never be the case for the plus mode (by definition of the manifest positive \( \xi \) parameter). Thus, we expect an exponential growth of low momentum modes. Furthermore, one may now notice that this is in the form of the radial Schrödinger ODE with a Coulomb potential. This in turn is just a special case of the general confluent hypergeometric ODE, which can be solved in terms of the confluent hypergeometric functions \( F_1 \) and \( U \). Our particular interest in growing modes for the gauge field allows us to neglect \( F_1 \). So, the the solution is given by (up to complex conjugation)

\[ \omega_-(x) = ce^{ix} U(1 + i\xi, 2, -2ix). \]  

(C.5)

Matching the gauge field value in the infinite past to the Bunch-Davies vacuum

\[ \lim_{x \to \infty} \omega(x) = \exp(ix), \]  

(C.6)
will fix the integration constant $c$. Let us use the asymptotic series of $U$ to show this

$$U(a, b, x) \sim \left( \frac{1}{x} \right)^a (1 + a(b - a)x^{-1} + 1/2a(a + 1)(a + b - 1)(2 + b - a)x^{-2} + \mathcal{O}(x^{-3})) .$$

Since the limit must be finite, we also may calculate the limit of $\omega \omega^*$ to get rid of the complex exponential phase\(^1\). This leads to

$$\omega \omega^* = c^2 x^2 U(1 + i\xi, 2, -2ix) U(1 - i\xi, 2, 2ix) \quad (C.7)$$

$$= c^2 x^2 \left( \frac{1}{2ix} \right)^{1+i\xi} \left( 1 + (1 + i\xi)(2 - 1 - i\xi)(-1) \right) \times \quad (C.8)$$

$$\times \left( \frac{1}{2ix} \right)^{1-i\xi} \left( 1 + (1 - i\xi)(2 + 1 + i\xi) \right) + \mathcal{O}(x^{-2}) \quad (C.9)$$

$$= c^2 x^2 \left( \frac{1}{2ix} \right)^{1+i\xi} \left( \frac{1}{2ix} \right)^{1-i\xi} + \mathcal{O}(x^{-1}) \quad (C.10)$$

$$= -c^2 x^2 e^{-\pi\xi} \left( \frac{1}{2ix} \right)^2 + \mathcal{O}(x^{-1}). \quad (C.11)$$

The matching to the Bunch-Davies vacuum in the infinite past $x \to \infty$ yields

$$-c^2 x^2 e^{-\pi\xi} \left( \frac{1}{2ix} \right)^2 + \mathcal{O}(x^{-1}) \overset{!}{=} 1 \implies c = 2e^{\pi\xi/2}. \quad (C.12)$$

Further, we may reduce our $\omega$ function with the relationship between the confluent hypergeometric function of the second kind and the Whittaker function, given by

$$U(a, b, x) = \frac{e^{\pi/2} W_{1/2(-2a+b), 1/2(b-1)}(x)}{x^{b/2}}. \quad (C.13)$$

This leads to

$$\omega_-(x) = ie^{\pi\xi/2} W_{-i\xi, 1/2}(-2ix). \quad (C.14)$$

After dropping the irrelevant complex prefactor of $i$, we get the result used in the main text.

\(^1\)Another approach is simply to show that in the limit of large $x$ we get a unit circle. In particular we have $c x U \sim c/(2 \exp (\pi\xi/2)) i \exp (-i\xi \ln 2x)$, where the last part defines a complex cycle. So we also obtain $c = 2 \exp (\pi\xi/2)$.\)
C.2 The electric and magnetic field

The four vector gauge field potential relates in general to electric and magnetic fields. The physical fields are defined as

\[ B = \frac{1}{a^2} \nabla \times A(\tau, x), \quad E = -\frac{1}{a^2} A'(\tau, x). \]  

(C.15)

When taking again only the growing mode into account, this leads to

\[ \langle E \cdot B \rangle = -\frac{1}{4\pi^2 a^4} \int dk k^3 |A_+|^2, \]  

(C.16)

\[ \frac{1}{2} \langle E^2 + B^2 \rangle = \frac{1}{4\pi^2 a^4} \int dk k^2 \left( |A'_+|^2 + k^2 |A_+|^2 \right). \]  

(C.17)

However, we now have to specify the integration range, since modes with large momenta remain in their vacuum state and thus contribute to the \( U(1) \) vacuum energy. First note, that the integration can be done over the variable \( x \equiv -k\tau \). Upon expecting eq. (C.4), we see that the tachyonic enhancement occurs when \( 1 - 2\xi/x < 0 \). Thus, an upper integration bound of \( x_{\text{max}} \sim 2\xi \) should capture the main integral contribution. Additionally, a lower bound \( x_{\text{min}} = 1/(8\xi) \) was found by the authors of ref. [29], but the implementation of this has no effect than just setting \( x_{\text{min}} = 0 \). Let us first focus on the product of electric and magnetic field and than adopt the same method for the second quantity.

To simplify eq. (C.16) we may use the derived analytic approximation for the enhanced gauge field mode from the previous section, namely

\[ \sqrt{2k} A_-(x) = e^{\pi \xi/2} W_{-\xi,1/2}(-2ix). \]  

(C.18)

So, after transforming the integration measure from \( k \) to \( x \), we arrive at

\[ \langle E \cdot B \rangle = \frac{H^4}{8\pi^2} e^{\pi\xi} \int_0^{x_{\text{max}}} x^3 \partial_x |W_{-\xi,1/2}(-2ix)|^2. \]  

(C.19)

We immediately see that \( \langle E \cdot B \rangle \) is a function of \( \xi \) only in de-Sitter space. When approximating the whole set up further with

\[ \sqrt{2k} A_-(x) \sim \left( \frac{x}{2\xi} \right)^{1/4} e^{\pi \xi - 2\sqrt{2\xi}}, \]  

(C.20)
the integral may be solved analytically \cite{48}. This gives the value

\[
\langle E \cdot B \rangle \sim 2.4 \times 10^{-4} H^4 e^{2\pi \xi} \frac{\xi^4}{\xi^4}.
\]  

(C.21)

In figure C.1 we demonstrate the deviation from the numerical evaluation of the integral and how it depends on the chosen cut-off scale \(x_{\text{max}}\). It reveals how sensitive the result is on the cut-off choice. A deviation is only apparent at small \(\xi\) and for \(\xi \gtrsim 3\) we find a good coincidence. Since the range of interest is exactly this, i.e. \(\xi \gtrsim 3\), the analytic approximation is a reasonable choice. We only demand high accuracy for this region because for smaller \(\xi\) the inflaton follows the vacuum slow roll evolution, meaning that the electric and magnetic fields are (super) weak. This is physical intuitive since the (tachyonic) growth of the gauge field is dependent on \(\dot{\phi}\). In slow roll we have small \(\dot{\phi}\) which translates into the absence of a tachyonic gauge field mode. Only for increasing \(\dot{\phi} \propto \xi\) the gauge field becomes important. Thus, a deviation from the analytic electric and magnetic field can be neglect for small \(\xi\). This is in accordance of what ref. \cite{101} found.

### C.3 Planck Constraints on Axion Inflation in the Slow Roll Limit

We will use this section to provide an analytic derivation of the slow roll dynamics, focusing on the periodic axion potential. We show that after applying the Planck data
for the spectral tilt $n_s$ and the scalar-to-tensor ratio $r$ this model requires a super-Planckian periodicity. Further, we use this derived periodicity to estimate the potential height by matching it to the scalar power spectrum $\Delta^2_\zeta$. Since CMB measurement are taken at $\sim 50-60$ e-folds before the end of inflation, we do not expect relevant deviations when including gauge field (friction), because they only become important for increasing $\dot{\phi}$, i.e. late times. An exception can be found if the delay of inflation is so huge that the shift of the end of inflation between slow roll and non slow roll is $O(10)$. Then, we take the slow roll parameters as a starting point to numerically find the parameter points fulfilling the CMB constraints. The small non-Gaussianity constraint sets anyway an upper bound on the interaction coupling, such we do not have to worry much about this.

Let us start with the general vacuum inflaton dynamics described by

$$\ddot{\phi} + 3H \dot{\phi} + \partial_\phi V(\phi) = 0,$$

$$\frac{1}{3} \left( \frac{1}{2} \dot{\phi}^2 + V(\phi) \right) = H^2,$$

where we take the potential to be $V(\phi) \equiv \Lambda^4 (1 - \cos(\phi/f_V))$. The aim will be to find an analytic approximation for the inflaton value at CMB scales. Therefore, we start by integrating the standard definition of the number of e-folds

$$|N| = \int_{t}^{t_{\text{end}}} H dt = \int_{\phi}^{\phi_{\text{end}}} \frac{H}{\dot{\phi}} d\phi \sim \int_{\phi}^{\phi_{\text{end}}} \frac{V(\phi)}{\partial_\phi V(\phi)} d\phi,$$

where we used the slow roll approximation of the EOM $\dot{\phi} \sim -\partial_\phi V/(3H)$ and of the Hubble parameter $H^2 \sim V/3$. The value of $\phi_{\text{end}}$ can be well approximated by the violation of the potential slow roll parameter $\epsilon_V = 1/2((\partial_\phi V)/V)^2 = 1$, leading to

$$\phi_{\text{end}} = 2f_V \arctan \left( \sqrt{\frac{1}{2f_V^2}} \right).$$

The analytic expression for the inflaton value in the slow roll limit can be calculated to

$$\phi(N) = 2f_V \arccos \left( \sqrt{\frac{1}{1 + 1/(2f_V^2)}} \times e^{-\frac{N}{2f_V^2}} \right).$$
Continuing using the slow roll limit, we can express $\Delta^2$, $n_s$ and $r$ purely in terms of the potential

$$\Delta^2 \sim \frac{1}{24\pi^2} \frac{V^*}{\epsilon^*}, \quad n_s \sim 2 \frac{\partial_{\phi\phi} V^*}{V^*} - 3 \left( \frac{\partial_{\phi} V^*}{V^*} \right)^2 + 1, \quad r \sim 16\epsilon^*.$$  \hspace{1cm} (C.27)

Note that all quantities have to be evaluated at horizon crossing, denoted with the $\star$. With our analytic formula from eq. (C.26), we can see that $n_s$ and $r$ are functions of $f_V$ only, after fixing the horizon exit time $N$ of course. Let us use this property to plot the value of $n_s$ and $r$ for different periodicity $f_V$ at CMB scales $N_s = N_{\text{CMB}} = (50, 60)$. The measured values by Planck [1], i.e. $n_s = 0.9665 \pm 0.0038$ and $r \lesssim 0.07$, then constraint the periodicity of the potential [60]. The result is shown in figure C.2. Apparently, the model is slightly disfavored by the current data, but with $f_V \sim 7M_p$ we are still in the Planck confidence region. Fixing this value as the periodicity, we may further use the measurement of the scalar power spectrum [2] to estimate the height of the potential. The resulting height of $\Lambda^4 \sim 6.3 \times 10^{-9}$ can be extracted from the same figure. This is in agreement with the derived result in [1].

**Figure C.2:** **Left:** The theoretical predicted correlated $n_s$ and $r$ values for natural inflation as a function of $f_V$ only. We display the curve for different CMB scales horizon exit times. To be in full accordance with the Planck data (green shaded region) the model requires a super-Planckian periodicity of $f_V \sim 7M_p$. **Right:** The fixed periodicity can be used to estimate the potential height $\Lambda^4$ which matches the scalar power spectrum normalization of $\Delta^2 \times 10^9 = 2.105 \pm 0.03$. The experimental results are taken from [1, 2] respectively.
Appendix D

Details on the Non-Abelian Gauge Field Background

In the linearized analysis for the $SU(2)$ gauge group one expands the gauge fields around an isotropic and homogeneous background like

$$A(\tau, \bar{x}) = A^{(0)}(\tau) + \delta A(\tau, \bar{x}). \quad (D.1)$$

Then, we can transform any $SU(2)$ gauge field into the form

$$A_0^a(\tau) = 0, \quad A_i^a(\tau) = \delta_i^a f(\tau), \quad (D.2)$$

where $f(\tau)$ must be real valued. When taking $f(\tau) \equiv a(t)\gamma(t)$ for numerical reasons, one obtains the following coupled ODEs for CNI [36]

$$\partial_N^2 \phi - 3 \partial_N \phi \left(1 - \frac{\partial_N H}{3H}\right) + \frac{\partial_\phi V(\phi)}{H^2} = -\frac{3g\alpha}{f_c} \gamma^2 \left(\frac{\gamma}{H} - \frac{\partial_N \gamma}{H}\right), \quad (D.3)$$

$$\partial_N^2 \gamma - 3 \partial_N \gamma \left(1 - \frac{\partial_N H}{3H}\right) + \gamma \left(2 - \frac{\partial_N H}{H}\right) + 2g^2 \gamma^3 = -\frac{g\alpha}{f_c} \gamma^2 \partial_N \phi, \quad (D.4)$$

$$\frac{3}{2} H (\gamma - \partial_N \gamma)^2 + \frac{3}{2} g^2 \gamma^4 + H^2 \frac{(\partial_N \phi)^2}{2} + V(\phi) = 3H^2. \quad (D.5)$$

For eq. (D.4) an analytic solution in the slow roll limit was found [36, 52]

$$gHq = c_i \xi, \quad (D.6)$$
with the three different types of solutions

\[ c_0 = 0, \quad c_1 = \frac{1}{2} \left( 1 - \sqrt{1 - \left(\frac{2}{\xi}\right)^2} \right), \quad c_2 = \frac{1}{2} \left( 1 + \sqrt{1 - \left(\frac{2}{\xi}\right)^2} \right). \]  

(D.7)

Here we introduce the notation for the 'analytic background field' \( q \) and for the 'numerical background field' \( \gamma \). As rigorously shown in ref. [52], the \( c_2 \) - type solution is overwhelming likely to be present once it is possible to be present, i.e. for \( \xi \geq 2 \). We now show for the first time that the analytic solution approximates the full solution of the homogeneous background to a satisfying level even for \( \xi \neq \text{constant} \). Recall, that the background field in this form is only present in the non-Abelian case due to the gauge field decomposition from eq. (D.2). So, it can be inherently assigned to the non-Abelian dynamics. In the numerical calculation we switch from the Abelian limit of the theory, where no background field is present, to the non-Abelian regime, when

\[ \langle A_{ab}^2 \rangle^{1/2} = \xi c_2 / (-g \tau). \]  

(D.8)

From this point on, we did the calculation always with the analytic \( c_2 \) solution. However, there is a non-negligible uncertainty in the matching point. The matching point was derived as the point when the analytic solution leaves its oscillatory behaviour and approaches the \( c_2 \) solution. This however, may be too conservative since also smaller fluctuations may reach the \( c_1 \) solution (saddle point) and then eventually evolve into the \( c_2 \) solution. The uncertainty in the matching point is shown in figure D.1. It shows, how the matching point may happen at smaller \( \xi \) – which means earlier time and bigger \( N \) – with decreasing gauge coupling \( g \) compared to our approach with eq. (D.8). And in fact the magnetic drift regime is reached only for (very) small \( g \). So the uncertainty where the matching point has to be increases in this case. This reveals that the in the literature studied regime has also in the linearized limit (big) numerical problems. We compare in figure D.2 the numerical solution of the coupled ODEs with the analytic \( c_2 \) solution. The numerical solution of the background field oscillates for a few e-folds around the \( c_2 \) solution. The smaller the non-Abelian gauge coupling is, the bigger the oscillatory amplitude gets. This is because the discontinuity at the matching point gets larger and thus the uncertainty increases. But then, both results are in good agreement. This indicates that the \( c_2 \) solution is indeed an attractor solution for the background. Deep inside the magnetic drift regime this is no longer the case. What happens there,
Figure D.1: We plot the matching point in $\xi$ as a function of the non-Abelian gauge coupling $g$. The red line represents our used matching procedure $\langle A_{AB}^2 \rangle^{1/2} = n \times \xi c_2 / (-g \tau)$ with $n = 1$. While the black line indicate how the matching point varies for $n = \{1/3, 3\}$. The green dotted line shows how the matching point changes when matching to the $c_1$ solution. The difference between these two different matching procedures increase with decreasing $g$. Additionally we show the boundary where no $c_2$ solution is allowed.

is that the background field approaches the $c_0$ solution after highly oscillating around the $c_2$ solution. This unlikely case arises due the increasing instability of the initial conditions at the matching point. So, what might seem as a solution to the problem is a variation of the initial conditions. Explicitly one might decrease the change of the background at the matching point in order to reduce/control the oscillation amplitude. We find already for small variations that the background is pushed to the $c_2$ solution. However, only for a $\gtrsim 90\%$ decrease the solutions start to overlap to good approximation. The same is true if one varies the initial value of the background field itself or both. But this seems a bit tricky and in no sense very reliable. Additionally, the evolution of the inflaton in this case does not seem very physical around the matching point. We observe a sudden increase in $\phi$ which is meaningless in our context. The background field effectively acts like a friction term to the inflaton dynamics but this cannot increase the inflaton value, only slow down the decrease. So, for parameter points deep inside the magnetic drift regime the matching procedure of eq. (D.8) fails. However, the steep increase of $\phi$ may point towards the point as discussed in figure D.1. Namely that the matching point should be chosen at earlier times with the boundary conditions given by the unstable $c_1$ solution. We show the results in figure D.3. Since we match earlier and
Figure D.2: We compare the evolution of the inflaton and the background obtained from numerically solving the coupled EOMs with the uncoupled EOM using the analytic $c_2$ solution for the background field. We used the matching point as given in eq. (D.8). Both results are in good agreement except deep inside the magnetic drift regime, where the matching uncertainty is increased. There the background approaches the zero solution. Additionally, the upwards pointing solution for the inflaton field is non-physical. So this matching procedure fails for the numerical analysis of the magnetic drift regime. In the main text we discuss how to circumvent this problem.

the gauge field friction term is (by definition of the magnetic drift regime) dominant, clearly the inflationary phase last longer than with the matching as before. This point can be added to the uncertainty we have anyway when inflation exactly ends. But the
important point is, that we could numerically verify that the background approaches the analytic $c_2$ solution very well, even deep inside the magnetic drift regime. For the scan we did in section 4.3 we however took always the conservative matching procedure of eq. (D.8), since it is not clear for which parameter point which matching should be done. Because when matching outside of the magnetic drift regime with the $c_1$ solution, the background numerically approaches zero. Only if one varies the boundary conditions, one gets the desired $c_2$ solution. For the parameter point studied this is only the case for e.g. $\gamma' = q' \rightarrow \gamma' \gtrsim 80 \times q'$. A possible way out is offered by the idea of matching with the $c_1$ solution, but take the $c_2$ solution as boundary conditions. However, this is not consistent and we only include this idea for completeness.

So in the end we see that we always reach numerically the $c_2$ analytic solution if the background has a non-zero value. But the numerical instability in the boundary conditions lead us to the conclusion to work with the analytic solution $q(t)$ and $f(\tau)$. 

**Figure D.3:** Upper panel: Matching with the $c_1$ solution when the saddle point is reached. Lower panel: The same plot as before, but now included the analytic solution $q$ when we match with our standard criterion (D.8). We see an uncertainty of around three e-folds between our standard matching and the numerical solution with the saddle point matching ($c_1$ solution). Else, they coincide to very good accuracy.
Bibliography


Acknowledgement

I Like to express my deepest gratitude to Valerie Domcke. Her supervision could not have been better. This made the stay at DESY from day one very pleasant. The many discussions we had were very insightful, especially because of her outstanding physical intuition and fundamental understanding. Also I am very happy to have shared a phenomenally productive office with my friends Felix Giese, Henrique Rubira and Peera Simakachorn. Furthermore, I appreciate a lot the help of Filippo Sala, Géraldine Servant and Thomas Konstandin in the search for a PhD position.
Ich versichere, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe angefertigt habe. Ich versichere, diese Arbeit nicht bereits in einem anderen Prüfungsverfahren eingereicht zu haben.

Hamburg,