

UC Berkeley

UC Berkeley Previously Published Works

Title

The inflated Chern-Simons number in spectator chromo-natural inflation

Permalink

<https://escholarship.org/uc/item/6m5039zf>

Journal

Journal of High Energy Physics, 2023(1)

ISSN

1126-6708

Authors

Bagherian, Hengameh

Reece, Matthew

Xu, Weishuang Linda

Publication Date

2023

DOI

10.1007/jhep01(2023)099

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

The inflated Chern-Simons number in spectator chromo-natural inflation

Hengameh Bagherian,^a Matthew Reece^a and Weishuang Linda Xu^{b,c}

^a*Department of Physics, Harvard University,
Cambridge, MA 02138, U.S.A.*

^b*Berkeley Center for Theoretical Physics, University of California,
Berkeley, CA 94720, U.S.A.*

^c*Theoretical Physics Group, Lawrence Berkeley National Laboratory,
Berkeley, CA 94720, U.S.A.*

E-mail: hengameh@g.harvard.edu, mreece@g.harvard.edu, wlxu@lbl.gov

ABSTRACT: The chromo-natural inflation (CNI) scenario predicts a potentially detectable chiral gravitational wave signal, generated by a Chern-Simons coupling between a rolling scalar axion field and an $SU(2)$ gauge field with an isotropy-preserving classical background during inflation. However, the generation of this signal requires a very large integer Chern-Simons level, which can be challenging to explain or embed in a UV-complete model. We show that this challenge persists in the phenomenologically viable spectator field CNI (S-CNI) model. Furthermore, we show that a clockwork scenario giving rise to a large integer as a product of small integers can never produce a Chern-Simons level large enough to have successful S-CNI phenomenology. We briefly discuss other constraints on the model, both in effective field theory based on partial-wave unitarity bounds and in quantum gravity based on the Weak Gravity Conjecture, which may be relevant for further explorations of alternative UV completions.

KEYWORDS: Axions and ALPs, Chern-Simons Theories, Cosmology of Theories BSM, Early Universe Particle Physics

ARXIV EPRINT: [2207.11262](https://arxiv.org/abs/2207.11262)

Contents

1	Introduction	1
2	The Spectator Chromo-Natural Inflation (S-CNI) model	4
2.1	S-CNI ingredients	4
2.2	S-CNI tensor perturbations	5
2.3	Axion periodicity	5
2.4	Equations of motion	7
3	Constraints from the EFT	9
3.1	Agreement with observational data	9
3.2	Physical control	12
3.3	Low energy parameter space exploration	13
4	Assessing the Λ range in S-CNI	15
5	The S-CNI model and clockwork	18
5.1	Clockwork constraints	18
5.2	A lower bound on f from clockwork	19
5.3	Restricting the Λ range with clockwork	19
5.4	Numerical results	20
6	Additional constraints	21
6.1	Perturbative unitarity bounds	21
6.2	The weak gravity conjecture constraint	24
7	Conclusion	26
A	Deriving the ϵ_ϕ equation of motion	28

1 Introduction

Inflation is the leading paradigm for the origin of the nearly scale-invariant primordial density perturbations that seeded the formation of structure in the universe. As such, it is potentially a powerful window to physics at energies far above what can be explored in terrestrial experiments. However, we have no clear information on the energy scale at which inflation took place. In conventional single field slow-roll inflation models, a measurement of primordial gravitational waves (tensor modes) would provide such information: both the Hubble expansion rate during inflation and the inflaton field excursion during inflation are proportional to the square root of the tensor-to-scalar ratio r [1]. A wide range of

inflation models share these properties [2, 3]. Thus, a near-future measurement of primordial gravitational waves would lead to the conclusion that the inflaton traversed a roughly Planckian range of field values during inflation, and that the Hubble scale during inflation was near $5 \cdot 10^{13}$ GeV. It is important to assess the robustness of these conclusions, in order to firmly anchor our knowledge of high-scale inflationary physics relative to the lower energies explored in particle physics.

A leading contender for a qualitatively different source of inflationary gravitational waves arises from models with large, *classical* non-abelian gauge field backgrounds during inflation. The original incarnation of such a scenario, “gauge-flation,” made the key observation that a classical background for an SU(2) gauge field could spontaneously break the product of (internal) gauge rotations and spatial rotations to the diagonal, preserving isotropy [4, 5]. This was shortly followed by a variant model, chromo-natural inflation (CNI), which replaced dimension-eight gauge field interactions in gauge-flation with a dimension-five coupling of an axion χ to SU(2) gauge fields [6] (see [7–10] for the relationship between the CNI model and the original gauge-flation). Ordinarily, tensor perturbations sourced by gauge field perturbations would be *quadratic*, $h_{ij} \sim \delta A_i \cdot \delta A_j$, and hence highly suppressed. The classical gauge field backgrounds in these models, however, allow gauge perturbations to source tensor perturbations *linearly*: $h_{ij} \sim A_i^{\text{cl}} \cdot \delta A_j$. Furthermore, because the axion coupling to gauge fields leads to tachyonic amplification of one helicity of the gauge field [11, 12], this interaction sources *chiral* gravitational waves [13], a possibility that received intense attention (e.g., [14–20]). At first glance, then, the CNI model seems very appealing: it relies on ingredients (axions coupled to gauge fields through a Chern-Simons interaction) that are ubiquitous in UV completions [21]; it provides a novel physical mechanism for sourcing gravitational waves, deviating from the standard logic relating a tensor signal to the scale of inflation; and, fortunately, it can also be distinguished observationally through the intrinsic chirality of its tensor modes. Furthermore, the isotropic SU(2) gauge field background was shown to be an attractor solution, at least beginning from certain anisotropic classical backgrounds [22].

Subsequent investigation of chromo-natural inflation revealed three potentially problematic aspects of the model: inconsistency of the minimal model with observations; back-reaction from large perturbations on the dynamics; and a required axion coupling to gauge fields that is large enough to pose a difficulty for UV completions. Our focus in this paper is the third aspect, though to assess its importance we must take into account the others as well. We will now briefly summarize the first two aspects, before explaining the third in more detail.

The first challenge to the original CNI model was a detailed observational one: the parameter space could not simultaneously accommodate the measured value of the spectral index n_s and the upper bound on r [13, 14, 23]. This spurred the development of modifications to the underlying model, e.g., higgsed chromo-natural inflation, in which the SU(2) gauge fields acquire mass through the Higgs mechanism [24]. Another, more well-studied, modification is the spectator chromo-natural inflation (S-CNI) model, introduced in [20]. This scenario, which will be our focus in this paper, assumes the existence of a scalar inflaton field in addition to the axion and SU(2) gauge fields of the original CNI

model. By decoupling the scale of the inflationary potential from the axion dynamics, this opens up a larger parameter space and relaxes the observational constraints. The S-CNI model was argued to allow for a large gravitational wave signal even for low-scale inflation [25]. It was also argued to be an attractor solution beginning from anisotropic initial conditions [26, 27].

A further challenge to the CNI model arises from backreaction of perturbations. The large gravitational wave signals of interest involve very large occupation numbers of gauge field modes, which source tensor perturbations. The gauge modes can backreact on the classical evolution of the fields, at some point invalidating the perturbation theory, a constraint that removes portions of the parameter space where the sourced contribution to r is much larger than the vacuum contribution to r [28]. Furthermore, the tensor perturbations can also source *scalar* perturbations via a one-loop diagram, first studied in the original CNI model in [29] and later in the S-CNI model in [30]. Accounting for this effect is important to accurately estimate the tensor-to-scalar ratio. Working in a reliably calculable regime with small backreaction, these studies (contrary to the initial ones like [25]) find severe limitations on the prospects for obtaining a dominantly chiral gravitational wave signal orders of magnitude larger than what one would expect in single field slow-roll inflation with similar Hubble scale. On the other hand, they leave open the possibility of a tensor signal that has an observable chirality, which would be a dramatic signal of dynamics beyond standard single field slow-roll inflation.

In this paper, we assess whether the surviving S-CNI parameter space is feasible from the viewpoint of quantum field theory. Although the basic ingredients of the CNI model are appealingly simple and expected to arise in many UV completions, the detailed parameter choices are more problematic. A key feature of the CNI model, in common with other models like [31], is the need for a large coupling of axions to gauge fields, of the form $\frac{\lambda}{f}\chi F\tilde{F}$ where χ is an axion with a potential of period $2\pi f$ and $\lambda \sim 100$ for phenomenological consistency of the model. As originally pointed out in [32], and elaborated on in [33], this poses a puzzle. An axion is not just any scalar field: it is a periodic scalar, and this implies that such a coupling is in fact a Chern-Simons term with coefficient an *integer* multiple of $g^2/(32\pi^2)$. We review this below in section 2.3. When g is perturbatively small (as it must be, for consistency of the model), large λ can require that the integer coefficient be enormously large. This requires explanation. In this paper, we will see that this issue is also present in the Spectator CNI model, and again, the large integer cannot be explained with the clockwork mechanism [34–37]. A related investigation of challenges involved in obtaining a sufficiently large Chern-Simons term, together with an observationally consistent cosmology with enhanced tensor modes, in the context of a controlled calculation within a string-theoretic UV completion of S-CNI has appeared in ref. [38].

In section 2 we will review the basic setup of the S-CNI model and establish our notation. In section 3, we discuss a number of constraints on the model, arising from compatibility with observations and the existence and control of a slow-roll solution within the effective theory. These constraints have been derived in earlier literature (e.g., [20, 28, 30]), although in some cases our treatment is slightly different. After presenting these constraints, we summarize the relatively small viable parameter space that remains. In section 4, we provide

some useful intermediate technical results to better understand the parameter space and facilitate the subsequent discussion. In section 5, we show that the parameters of the S-CNI model cannot be explained with the clockwork mechanism. Section 6 discusses other constraints on the model, both EFT constraints from perturbative unitarity and (conjectural) UV constraints from embedding in quantum gravity. In section 7, we offer concluding remarks.

2 The Spectator Chromo-Natural Inflation (S-CNI) model

In this section, we review the basic structure of the spectator chromo-natural inflation (S-CNI) model.

2.1 S-CNI ingredients

Let us first review the ingredients of the S-CNI model [20]. First is the inflaton ϕ , with potential $V(\phi)$ that is assumed to dominate the energy density of the universe. In particular, ϕ should slowly roll for at least 50 to 60 e-folds, and we expect inflation to end when ϕ transitions away from slowly rolling, just as in conventional inflation models. Second, we have the chromo-natural sector [6], which consists of an axion field χ interacting with SU(2) gauge fields A_μ^a via a Chern-Simons term.¹ This sector is assumed to interact with ϕ only through gravity. During the observable era of inflation (i.e., those e-folds that sourced the modes that we measure in the CMB), the field χ is assumed to be rolling in a potential $U(\chi)$ while the gauge field has a classical background that preserves isotropy [4], which we parametrize through a function of time, $Q(t)$:

$$\langle A_0^a(t) \rangle = 0, \quad \langle A_i^a(t) \rangle = \delta_i^a a(t) Q(t). \quad (2.1)$$

Here $a(t)$ is the scale factor, a is an SU(2) adjoint index, and i is a spatial Lorentz index. As emphasized in the S-CNI context in [20, 30], following similar work in other inflation scenarios [39, 40], χ need not roll for the full duration of ϕ -driven inflation. The generation of an interesting gravitational wave signal could occur during a limited interval when χ is rolling. We denote the number of e-folds of axion rolling by N_χ , which we take to be $N_\chi \gtrsim 10$ [30] in order to cover the observationally relevant scales, and assume this epoch was followed by roughly $60 - N_\chi$ additional e-folds of ϕ -driven inflation. Note that smaller values of $N_\chi = 2, 5$ have been studied in slightly different variations of axion gauge field inflation [40]. We will not assume any complete model to explain what triggers the axion's rolling at a particular time, but simply focus on the observable signals generated from an EFT capturing the interval in which ϕ and χ both roll. As we will see, this is already highly constrained.

The EFT describing the evolution of ϕ , χ , and A_μ^a is assumed to be [20]

$$\frac{1}{\sqrt{|\det g|}} \mathcal{L} = -\frac{1}{2}(\partial_\mu \phi)^2 - V(\phi) - \frac{1}{2}(\partial_\mu \chi)^2 - U(\chi) - \frac{1}{4} F_{\mu\nu}^a F^{a\mu\nu} - \frac{\lambda}{4f} \chi F_{\mu\nu}^a \tilde{F}^{a\mu\nu}, \quad (2.2)$$

¹For the purposes of the model, we could equally well (up to a relative factor of 2 in the allowed normalization of the Chern-Simons term) assume the gauge group to be SO(3).

where $\tilde{F}^{a\mu\nu} \equiv \frac{1}{2\sqrt{|\det g|}} \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma}^a$ and $a \in \{1, 2, 3\}$ is the SU(2) adjoint index. We denote the SU(2) gauge coupling as g , not to be confused with the metric $g_{\mu\nu}$. We work in a mostly-plus metric signature, and assume a homogeneous, isotropic, background described by the Friedmann-Robertson-Walker (FRW) prescription. Throughout this paper, we will assume a simple periodic axion potential,

$$U(\chi) = \mu^4 [1 + \cos(\chi/f)]. \tag{2.3}$$

The scale μ could arise from confinement in an additional non-abelian gauge sector coupled to χ . Because the axion periodicity will play a key role in our discussion below, we will shortly comment on it in more detail, in section 2.3.

We adopt the following notation, widely used in the literature, for frequently occurring dimensionless combinations of parameters and field values:

$$m_Q \equiv \frac{gQ}{H}, \quad \xi \equiv \frac{\lambda}{2fH} \dot{\chi}, \quad \Lambda \equiv \frac{\lambda Q}{f}. \tag{2.4}$$

Note that m_Q can be thought of as the mass of the gauge field in Hubble units. These are all time-dependent quantities, but can be treated as approximately constant due to slow-roll conditions (to be discussed below in section 2.4).

2.2 S-CNI tensor perturbations

In order to make predictions regarding the tensor to scalar ratio, we will also need to define the tensor perturbation modes. We are following the notation from ref. [30], to write down the gauge field and metric perturbations:

$$\delta A_\mu^1 = a(x)[0, T_+(x, z), T_\times(x, z), 0], \quad \delta A_\mu^2 = a(x)[0, T_\times(x, z), -T_+(x, z), 0], \tag{2.5}$$

$$\delta g_{12} = -\delta g_{22} = a^2(x)h_+(x, z)/\sqrt{2}, \quad \delta g_{12} = \delta g_{21} = a^2(x)h_\times(x, z)/\sqrt{2}, \tag{2.6}$$

where modes are propagating along z and $x \equiv ka/H$. We can now define the left-handed and right-handed helicity tensors as

$$\hat{h}^\pm \equiv aM_{\text{Pl}}/2 [h_+ \mp ih_\times], \quad \hat{t}^\pm \equiv a(x) [T_+ \mp iT_\times]. \tag{2.7}$$

2.3 Axion periodicity

We assume that χ is a periodic scalar field, i.e., that there is an identification

$$\chi \cong \chi + 2\pi n f_\chi, \quad n \in \mathbb{Z}, \tag{2.8}$$

for some constant f_χ . This is a *gauge redundancy*, i.e., these different values of χ are different ways of referring to literally the same physical field configuration. This is a common feature of (pseudo-)Goldstone bosons, including axions. The periodicity of χ has important implications for both the potential $U(\chi)$ and the Chern-Simons term $\chi F\tilde{F}$, previously discussed in the context of the CNI model in [33].

The Chern-Simons term $\chi F\tilde{F}$ is not gauge-invariant: shifting χ by $2\pi f_\chi$ adds an effective θ -term to the SU(2) gauge field action. However, all physical quantities are invariant

if this added term is an integer multiple of $\frac{g^2}{16\pi} F_{\mu\nu}^a \tilde{F}^{a\mu\nu}$. The reason is that the integral of this term over spacetime, for any gauge field configuration, is itself 2π times an integer (the “instanton number”), so the path integral weight $\exp(iS[\phi, \chi, A])$ is invariant even though the action itself is not. As a result, the coefficient λ/f is actually *quantized*, i.e., it must be an integer multiple of a base unit:

$$\frac{\lambda}{4f} = k \frac{g^2}{32\pi^2} \frac{1}{f_\chi}, \quad k \in \mathbb{Z}. \tag{2.9}$$

The integer k is referred to as the *level* of the Chern-Simons term.

At this stage it is tempting to identify f_χ with f , but we should not do so: there is another subtlety to confront first. If we demand that $U(\chi)$ respects the periodicity (2.8), we would conclude that $1/f$ must be an integer multiple of $1/f_\chi$. However, there is a well-known loophole in this argument. In fact, we will allow it to be an *inverse* integer multiple:

$$\frac{1}{f} = \frac{1}{j} \frac{1}{f_\chi}, \quad j \in \mathbb{Z}. \tag{2.10}$$

As written, this appears to be a violation of gauge invariance, but in a full model it can be realized as a form of higgsing (spontaneously breaking the gauge invariance). There are many models that exhibit a *monodromy*, where as the axion field winds around its circle, another parameter of the theory (e.g., some flux) changes by an integer value. These effects can change the *effective* periodicity of the axion potential from $2\pi f_\chi$ to $2\pi j f_\chi$ for an integer j , allowing for j different “branches” of the potential to restore the true $2\pi f_\chi$ period [41, 42].²

Putting together (2.9) and (2.10), we learn that

$$\lambda = jk \frac{g^2}{8\pi^2}, \quad j, k \in \mathbb{Z}. \tag{2.11}$$

This suggests that the natural value of λ is very small, at weak gauge coupling g . Only models with a large integer jk can produce a large λ . This is a fundamental challenge for building a complete CNI or S-CNI model [32, 33].

Here we have assumed that the potential remains periodic, with a finite number j of branches. A qualitatively different scenario can arise in which there are *infinitely* many branches, leading to an effectively non-periodic axion potential [43, 44]. We do not consider this case here, sticking with the cosine potential studied in most of the CNI and S-CNI literature. However, more general monodromy potentials have been invoked in similar contexts [45–47], and could be an interesting target for further study.

One might also ask: why assume that χ is periodic at all? Perhaps it is a generic scalar with a coupling to $F^a \tilde{F}^a$, free from the constraints imposed by periodicity. However, in this case, without the protection afforded χ by being a pseudo-Nambu-Goldstone boson, it would be difficult to explain the flatness of its potential. Furthermore, we are aware of only two basic mechanisms for generating $\chi F^a \tilde{F}^a$ couplings: one is to integrate out SU(2)-charged fermions whose mass matrix depends on the value of χ , which would lead to similar

²In fact, this is essentially the same mechanism that allows an effectively fractional Chern-Simons level to appear in the fractional quantum Hall effect.

challenges whether or not χ is periodic; the second is to obtain χ as a zero mode of a higher-dimensional gauge field participating in a Chern-Simons interaction with the SU(2) gauge fields, in which case we expect χ to automatically be periodic. Thus, we do not expect abandoning the periodicity assumption to make the task of UV completing the model any easier.

2.4 Equations of motion

We can define a set of slow-roll parameters in the S-CNI model, all of which should be much smaller than one during inflation:

$$\epsilon_H \equiv -\frac{\dot{H}}{H^2}, \quad \epsilon_\phi \equiv \frac{\dot{\phi}^2}{2M_{\text{Pl}}^2 H^2}, \quad \epsilon_\chi \equiv \frac{\dot{\chi}^2}{2M_{\text{Pl}}^2 H^2}, \quad \epsilon_B \equiv \frac{g^2 Q^4}{M_{\text{Pl}}^2 H^2}, \quad \epsilon_E \equiv \frac{(HQ + \dot{Q})^2}{H^2 M_{\text{Pl}}^2}. \quad (2.12)$$

Here we use a dot to denote a derivative with respect to the FRW time coordinate t , H is the Hubble parameter $H(t) \equiv \dot{a}(t)/a(t)$, and $M_{\text{Pl}} \equiv \frac{1}{\sqrt{8\pi G_N}} \approx 2.4 \cdot 10^{18}$ GeV is the reduced Planck scale. We also define $\epsilon_A \equiv \epsilon_B + \epsilon_E$. There are also slow-roll conditions on second derivatives:

$$|\ddot{\phi}| \ll |H\dot{\phi}|, \quad |\ddot{\chi}| \ll |H\dot{\chi}|, \quad |\ddot{Q}| \ll |H\dot{Q}|. \quad (2.13)$$

Before using any approximations, the classical background equations of motion for the fields ϕ , χ , and Q are

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0, \quad (2.14)$$

$$\ddot{\chi} + 3H\dot{\chi} + U'(\chi) = -\frac{3g\lambda}{f} Q^2 (\dot{Q} + HQ), \quad (2.15)$$

$$\ddot{Q} + 3H(\dot{Q} + HQ) - H^2 Q (1 - \dot{H}/H^2) + 2g^2 Q^3 = \frac{g\lambda}{f} \dot{\chi} Q^2. \quad (2.16)$$

In these equations, a prime denotes a derivative of a function with respect to its argument. These must be supplemented with one linear combination of the Einstein equations; a useful choice can be expressed in terms of the slow-roll parameters (2.12):

$$\epsilon_H = \epsilon_\chi + \epsilon_\phi + \epsilon_A, \quad (2.17)$$

which also allows for straightforward assessment of the relative importance of the inflaton, axion, and gauge field contributions on the dynamics of inflation. The four equations (2.14), (2.15), (2.16), and (2.17) fully determine the evolution of the classical background, assuming an FRW metric and the isotropic SU(2) ansatz (2.1).

These equations depend on a choice of the inflaton potential $V(\phi)$. However, we will follow [20] in not choosing an explicit inflaton potential $V(\phi)$. Instead, we will replace the equation of motion (2.14) with an approximate equation of motion for the slow-roll parameter ϵ_ϕ . The dynamics of ϵ_ϕ are fixed by observed properties of the primordial density fluctuations, which we assume are sourced dominantly by ϕ , allowing us to remain agnostic to the underlying potential. The amplitude of the scalar power spectrum can be written in terms of ϵ_ϕ and ϵ_H , derived in [30] as

$$\mathcal{P}_\zeta = \frac{H^2}{8\pi^2 M_{\text{Pl}}^2} \frac{\epsilon_\phi}{\epsilon_H^2}, \quad (2.18)$$

while the time-evolution of ϵ_ϕ may be inferred from the spectral tilt of the spectrum, taken as approximately constant, $n_s \simeq 0.965 \pm 0.004$ [48]. This is given explicitly by

$$\dot{\epsilon}_\phi \approx H\epsilon_\phi \left[-6\dot{H}/H^2 + (n_s - 1) + 2\ddot{H}/(\dot{H}H) \right], \quad (2.19)$$

and we refer the reader to appendix A for details of this derivation. Note that underlying this entire prescription is the assumption that slow-roll conditions are met. We treat (2.15), (2.16), (2.17), and (2.19) (rather than (2.14)) as the four equations specifying the background evolution.

After using the slow-roll assumptions to drop the second derivatives, the two equations (2.15) and (2.16) can be separated into an equation for $\dot{\chi}$ and one for \dot{Q} . It is useful to think of the equation of motion for \dot{Q} as an equation for gradient flow in an effective potential $W_{\text{eff}}(Q)$ (with χ and H treated, for this purpose, as approximately constant) [6, 23]. Specifically,

$$\dot{Q} \approx -\frac{\partial W_{\text{eff}}(Q)}{\partial Q} = -\frac{3HQ(2f^2H^2 + 2g^2f^2Q^2 + g^2\lambda^2Q^4) + g\lambda Q^2 fU'(\chi)}{3(3f^2H^2 + g^2\lambda^2Q^4)}. \quad (2.20)$$

The solutions of interest for us have Q approximately sitting at a nonzero minimum of this effective potential.

The χ equation of motion leads to an important conclusion about the necessary size of the parameter λ . Using the slow-roll equation for $\dot{\chi}$ and rewriting $U'(\chi)$ in terms of other parameters by assuming that Q sits at a value where the right-hand side of (2.20) is zero, we find:

$$\dot{\chi} \approx \frac{2Hf}{\lambda} (m_Q + m_Q^{-1}). \quad (2.21)$$

From the definition of ξ in (2.4), this is equivalent to the claim

$$\xi \approx m_Q + m_Q^{-1}. \quad (2.22)$$

The expression (2.21) makes it clear that large λ suppresses the evolution of $\dot{\chi}$; this is the key friction effect that causes the time evolution of the axion to differ in the presence of the SU(2) gauge fields compared to an ordinary axion, allowing the CNI model to have sub-Planckian axion inflation (in contrast to natural inflation [49]). On the other hand, we also see from (2.21) that in the *small* λ limit, $\dot{\chi}$ becomes large: the axion evolves even faster than it otherwise would.

Ignoring the time dependence of Q , we can roughly estimate the number of e-folds of CNI dynamics from (2.21) simply by asking how many Hubble times it takes for χ to change value by f . We see that

$$N_\chi \sim \frac{\lambda}{2(m_Q + m_Q^{-1})}. \quad (2.23)$$

This derivation is valid in the S-CNI model, but leads to similar conclusions as in the original CNI model (e.g., [6] states that $N_\chi \lesssim 0.6\lambda$). Requiring a number of e-folds $N_\chi \gtrsim 10$ translates into the demand that

$$\lambda \sim 2(m_Q + m_Q^{-1}) N_\chi \gtrsim 60, \quad (2.24)$$

where we have used that (as explained below) $m_Q \sim 2.7$ in the viable parameter space. Thus, we are only interested in parameter space with large coupling λ . In terms of the parametrization (2.11), this means that

$$jk = \frac{8\pi^2\lambda}{g^2} \gtrsim 5 \cdot 10^7 \left(\frac{10^{-2}}{g}\right)^2. \quad (2.25)$$

This large integer proves to be a significant challenge for UV completions of the model, as we will see below.

3 Constraints from the EFT

As explored in previous work [7, 20, 28, 30], multiple constraints and inequalities split and restrain the available space of parameters for the S-CNI model. In this section we will discuss and expand on some of these bounds. We put these constraints in two broad categories: compatibility with observational bounds and physical control over the behavior of solutions.

The set of key parameters that we will mainly use to dictate the observable behavior of S-CNI are H , g , and m_Q . These govern the energy content stored in the gauge field and (via subtraction from Hubble) in the inflaton. The prominence of the axion as a player in inflationary phenomenology, however, is governed by an additional parameter Λ (defined in (2.4)), which encodes dependence on the axion-related parameters: λ , f , and (by way of equations of motion) μ . For sufficiently small Λ , the axion becomes dynamic enough to impact the evolution of density perturbations.

This parameter Λ has been traditionally taken to its large limit $\Lambda \gg \sqrt{2}$ and decoupled from the analysis, as analytically solving the equations of motion for Q significantly simplifies in this limit. In addition, in the original CNI model having a large Λ was necessary to drive a sufficiently long phase of slow-roll inflation [20, 23, 30]. Thanks to the addition of the new spectator scalar, however, S-CNI doesn't require large values of Λ to last for enough e-folds. This opens up an unexplored region for CNI dynamics. However, because of the aforementioned analytic appeal of the $\sqrt{2} \ll \Lambda$ regime and the similarities in the behavior of the axion-gauge sector to that of the original model, most of the S-CNI literature has been focused on large Λ regimes.

In this work, we analyze the full range of values for Λ , and in later sections explore the limitations of the clockwork mechanism in this space. Hence, it is important to keep track of the Λ dependence of each of our bounds as we come across them in this section.

3.1 Agreement with observational data

In this section, we discuss constraints that need to be met if the model is to comply with experimental data. There are, generically speaking, three cosmological observables that inflation-era physics must confer with: the ratio between primordial tensor and scalar fluctuations r , the amplitude of adiabatic scalar modes \mathcal{P}_ζ , and its spectral tilt n_s . We use the measured value of n_s to fix the evolution of the spectator inflaton ϕ (see appendix A for details), and so a spectral index consistent with the data is a presupposed part of our

S-CNI phenomenology. In the text following, we discuss constraints from the other two measurements.

Recent data from the BICEP/Keck collaboration have placed stringent limits on the tensor-to-scalar ratio, $r < 0.036$ [50] at the 95% CL. Here, we summarize the derivation of r under the S-CNI framework, following closely the work of ref. [20].

The metric perturbation modes \hat{h}^\pm are sourced in part by the tensor perturbations of the gauge field \hat{t}^\pm , both defined in section 2. An instability of the $+$ -helicity gauge field tensor mode during horizon crossing induces a large enhancement in the metric fluctuations of the same helicity. As a result, the observable tensor modes are likewise highly enhanced, and the enhanced component is entirely chiral.

The power spectrum of these sourced gravitational modes at the super-horizon limit is

$$\mathcal{P}_h^s = \frac{\mathcal{P}_h^v}{2} |\sqrt{2kx} \lim_{x \rightarrow 0} \hat{h}^+(x, k)|^2, \quad (3.1)$$

where $\mathcal{P}_h^v = \frac{2}{\pi^2} \frac{H^2}{M_{\text{Pl}}^2}$ is the amplitude of the gravitational power spectrum produced through vacuum fluctuation; the measurement of r probes the sum total amplitude of both the vacuum and sourced contributions.

The tensor fluctuations of the gauge field evolve as $\hat{t}^+ = i^\beta (2k)^{-1/2} W_{\beta, \alpha}(-2ix)$. The $W_{\beta, \alpha}(z)$ are Whittaker functions with arguments $\alpha \equiv -i\sqrt{2m_Q\xi} - 1/4$ and $\beta \equiv -i(m_Q + \xi)$. Using Bunch-Davies vacuum solutions to construct the Green's function for \hat{h}^+ and integrating it over the Whittaker source terms (see appendix E of [28] and appendix B of [18]), we arrive at an analytic expression

$$|\sqrt{2kx} \lim_{x \rightarrow 0} \hat{h}^+(x, k)|^2 = |\mathcal{F}_E \sqrt{\epsilon_E} + \mathcal{F}_B \sqrt{\epsilon_B}|^2 \equiv \epsilon_B \mathcal{F}^2, \quad (3.2)$$

with

$$\begin{aligned} \mathcal{F}_E = & \frac{\pi^{i\beta+1} \sec(\pi\alpha) [16\alpha^4 - 40\alpha^2 + 9]^{-1}}{\Gamma(1-\beta)\Gamma(-\alpha-\beta+\frac{1}{2})\Gamma(\alpha-\beta+\frac{1}{2})} \left[16(4\alpha^2 - 8\beta - 1)\Gamma(1-\beta)^2 \right. \\ & + [-16\alpha^4 + 8\alpha^2(5-8\beta) + 16\beta(1-8\beta) - 9]\Gamma(-\alpha-\beta+1/2)\Gamma(\alpha-\beta+1/2) \\ & \left. + (16\alpha^4 - 40\alpha^2 + 9)\Gamma(-\beta)\Gamma(1-\beta) \right], \end{aligned} \quad (3.3)$$

and

$$\mathcal{F}_B = \frac{\pi^{i\beta} \sec(\pi\alpha) [-4\alpha^2 - 8i\beta m_Q + 9] \left[\frac{4\alpha^2 + 8\beta - 1}{\Gamma(1-\beta)} + \frac{(4\alpha^2 - 8\beta - 1)\Gamma(-\beta)}{\Gamma(-\alpha-\beta+\frac{1}{2})\Gamma(\alpha-\beta+\frac{1}{2})} \right]}{16\alpha^4 - 40\alpha^2 + 9}. \quad (3.4)$$

We refer the reader to [20] for further details, though note that our analytical expression \mathcal{F}_B is slightly different due to an assumed typo in their text.³

³Our result closely follows their numerical fit $\mathcal{F}^2 \sim \exp(3.6m_Q)$ (also in [30]) and reproduces their figures, so the disagreement must be a spurious one.

These expressions, along with eqs. (3.1), (3.2), allow us a direct determination of r for a given g, m_Q , and H . The bound on tensor-to-scalar ratio thus translates to a bound on this parameter space,

$$r = \frac{\mathcal{P}_h^v + \mathcal{P}_h^s}{\mathcal{P}_\zeta} = r_{\text{vac}} \left(1 + \frac{\epsilon_B}{2} \mathcal{F}^2 \right) < 0.036. \quad (\text{CON.1})$$

As the characteristic signature of CNI scenarios are these chiral tensor modes, model realizations where they constitute a larger fraction of the allowed total are more phenomenologically interesting to pursue. We highlight regions of parameter space where the contributions of the sourced gravitational waves dominates over vacuum contributions,

$$\mathcal{R}_{GW} = \frac{r_s}{r_{\text{vac}}} = \frac{\epsilon_B}{2} \mathcal{F}^2 > 1. \quad (\text{CON.2})$$

The derivation of (CON.1) and (CON.2) relied on slow-roll assumptions but not on the large- Λ approximation, and are sensitive only to the choice of g, m_Q and H (or equivalently r_{vac}). To see this dependence explicitly, we can use definitions (2.4) and (2.12) to write (CON.2) as

$$\mathcal{R}_{GW} = \frac{\epsilon_B}{2} \mathcal{F}^2 = \frac{m_Q^4}{g^2} \frac{H^2}{M_{\text{Pl}}^2} \frac{\mathcal{F}^2}{2} = \frac{m_Q^4}{g^2} \frac{\pi^2}{2} r_{\text{vac}} \mathcal{P}_\zeta \frac{\mathcal{F}^2}{2} > 1, \quad (\text{CON.2}^*)$$

with $\mathcal{P}_\zeta^{\text{obs}} \simeq 2.1 \cdot 10^{-9}$ [48] being the amplitude of the scalar power spectrum.

To this point, our model must also produce the correct amount of scalar fluctuations \mathcal{P}_ζ . The spectator inflation scenario assumes that both the scale of inflation and size of scalar fluctuations are set dominantly by the inflaton ϕ . However, the evolution of these perturbations are governed also by the *dynamics* of Hubble, $\epsilon_H = -\dot{H}/H^2$, to which ϵ_ϕ need not be the dominant contributor. The expression for the scalar power spectrum was generalized along precisely these lines by ref. [30], giving

$$\mathcal{P}_\zeta = \frac{H^2}{8\pi^2 M_{\text{Pl}}^2} \frac{\epsilon_\phi}{\epsilon_H^2}. \quad (3.5)$$

Substituting $\epsilon_H = \epsilon_\chi + \epsilon_A + \epsilon_\phi$, we obtain a requirement for ϵ_ϕ that is entirely agnostic of underlying inflaton model,

$$\epsilon_\phi = \frac{\epsilon_A + \epsilon_\chi}{2} \left[A - 2 \pm \sqrt{A(A-4)} \right], \quad \text{with } 4 \leq A \equiv \frac{H^2}{8\pi^2 M_{\text{Pl}}^2 \mathcal{P}_\zeta} \frac{1}{\epsilon_A + \epsilon_\chi}. \quad (\text{CON.3})$$

Given a fixed inflation scale H and \mathcal{P}_ζ , the analytic constraint $4 \leq A$ puts an upper bound on $\epsilon_A + \epsilon_\chi$ to ensure the existence of a slow-roll solution to our equations. In other words, (CON.3) ensures the predicted value for the scalar power spectrum is correctly normalized with respect to the observed $\mathcal{P}_\zeta^{\text{obs}} \simeq 2.1 \cdot 10^{-9}$ for given slow-roll parameters. Intuitively, if the evolution of Hubble is too much faster than the rolling of the inflaton, the inflaton perturbations are too washed-out to constitute our observed primordial inhomogeneities.

The two solutions for ϵ_ϕ correspond to regimes $\frac{\epsilon_\phi}{\epsilon_A + \epsilon_\chi} > 1$ and $\frac{\epsilon_\phi}{\epsilon_A + \epsilon_\chi} < 1$, which split the parameter space of the S-CNI model [30]. In this paper we call the former branch *classic*

S-CNI while the latter is named *mixed* S-CNI. In classic S-CNI, the inflaton both dominates the energy density of the universe and is responsible for driving inflation forward. In mixed S-CNI, the inflaton still dominates the energy density, but the dynamics of inflation are dictated by the gauge field. Both branches show attractor solution behavior.

The Λ -dependence of (CON.3) plays a significant role in our future discussions. Taking $\dot{Q} \approx 0$, and using definitions (2.4) and (2.12), we can write

$$\epsilon_A + \epsilon_\chi = \frac{H^2/M_{\text{Pl}}^2}{g^2} \left(m_Q^4 + m_Q^2 + \frac{2(1+m_Q^2)^2}{\Lambda^2} \right). \quad (3.6)$$

As we discussed in section 2.4, χ rolls quickly when λ is small, as reflected here in the growth of ϵ_χ at small $\Lambda \propto \lambda$. In the opposite limit, $\Lambda \gg \sqrt{2}$, ϵ_χ is suppressed and setting $\epsilon_H \simeq \epsilon_\phi + \epsilon_B$ is a good approximation (as one notes that $\epsilon_E/\epsilon_B \simeq m_Q^{-2}$, and $m_Q > \sqrt{2}$ as discussed below). We can then replace (CON.3) with

$$4 \lesssim \frac{H^2}{8\pi^2 M_{\text{Pl}}^2 \mathcal{P}_\zeta \epsilon_B} \frac{1}{\epsilon_B} = \frac{g^2}{8\pi^2 \mathcal{P}_\zeta m_Q^4} \quad (\text{CON.3}^*)$$

This approximation breaks down when ϵ_χ can no longer be ignored in the sum $\epsilon_H = \epsilon_\phi + \epsilon_A + \epsilon_\chi$ at $\Lambda \sim \sqrt{2}$.

3.2 Physical control

Our second category is physical consistency; these constraints restrict us to regimes where our control over the theory is robust and our solutions self-consistent.

First and foremost, we assume that inflation is well-described by slow-roll, in the regime where eqs. (2.12) and (2.13) are valid.

Next, we avoid tachyonic *scalar* perturbations on sub-horizon scales by putting a lower bound on m_Q (defined in (2.4)) [13],

$$\sqrt{2} < m_Q. \quad (\text{CON.4})$$

At all values of Λ , this condition is necessary to have a theory that is well-described by an approximate classical solution with a nearly scale-invariant spectrum of density perturbations.

A major approximation taken for feasibility of computation is to neglect the backreaction of enhanced tensor perturbations on the background gauge fields. As derived in refs. [20, 30], the backreaction term altering the right hand side of eq. (2.16), is

$$\mathcal{T}_{BR}^Q = gH^3\xi/(12\pi^2) [\mathcal{B}(m_Q) - \tilde{\mathcal{B}}(m_Q)/\xi], \quad (3.7)$$

where ξ was defined in (2.4). In the above equation

$$\mathcal{B} \equiv \int_{x_{\min}}^{x_{\max}} dx x |i^{-\alpha} W_{\alpha,\beta}(2ix)|^2, \quad \tilde{\mathcal{B}} \equiv \int_{x_{\min}}^{x_{\max}} dx x^2 |i^{-\alpha} W_{\alpha,\beta}(2ix)|^2, \quad (3.8)$$

where $x_{\min}/_{\max} = m_Q + \xi \pm \sqrt{m_Q^2 + \xi^2}$ is the interval in which the tensor mode undergoes tachyonic enhancement. We will use the numerical fit $\mathcal{B}(m_Q) - \tilde{\mathcal{B}}(m_Q)/\xi \simeq 2.3 \cdot e^{3.9m_Q}$ to facilitate our computations throughout this paper [30].

To quantify the effects of backreaction, we rewrite (2.16) in terms of $\partial V_{\text{eff}}(Q)/\partial Q \equiv 2H^2Q(1 + m_Q^2) - \frac{g\lambda}{f}\dot{\chi}Q^2$ [28]:

$$\ddot{Q} + 3H\dot{Q} + \dot{H}Q + \partial V_{\text{eff}}/\partial Q \simeq -\mathcal{T}_{BR}^Q. \tag{3.9}$$

Under slow-roll assumptions, we ignore $|\ddot{Q}/H\dot{Q}|$ and set $\partial V_{\text{eff}}/\partial Q \simeq 0$. If we further assume $\epsilon_H = -\dot{H}/H^2 \simeq \text{const}$, eq. (3.9) can be solved as a standard non-homogeneous first order differential equation. The homogeneous solution to (3.9) is $Q^{(0)}(t) = Q(1 - \dot{H}/3Ht)$ which leads to the full solution

$$Q^{(1)}(t) - Q \simeq -\frac{\mathcal{T}_{BR}^Q}{3H} t \left[1 + \frac{\dot{H}}{\mathcal{T}_{BR}^Q} Q - (2m_Q\xi^{-1} + 3.9m_Q)\frac{\dot{H}}{3H}t \right] \simeq -\frac{\mathcal{T}_{BR}^Q}{3H}t, \tag{3.10}$$

where Q sets the right hand side of (2.20) to zero. Now, keeping the backreaction small corresponds to

$$|Q^{(1)}(t) - Q| \cdot Q^{-1} \simeq \left| -\frac{\mathcal{T}_{BR}^Q}{3H} \frac{t}{Q} \right| = \frac{2}{3}N_\chi \cdot \left| \frac{g^2}{24\pi^2} \frac{\mathcal{B}(m_Q) - \tilde{\mathcal{B}}(m_Q)/\xi}{m_Q/\xi} \right| \ll 1, \tag{CON.5}$$

where $N_\chi \sim 10$ is the number of e-folds that the axion rolls. At first glance, (CON.5) differs from that of ref. [30] by a factor of ~ 6 . However, note that their expression is time dependent, whereas (CON.5) has been integrated over the period of axion rolling. The extra factor keeps track of backreaction that accumulates over time. In this paper, we will be using this slightly modified backreaction constraint with the extra factor included. Note that no Λ assumptions were invoked during this discussion, and so (CON.5) holds for the full range of Λ .

Our last consistency constraint comes directly from ref. [30] and ensures quantum effects can be safely ignored in our classical field backgrounds:

$$\mathcal{R}_{\delta\phi} \equiv \frac{\langle \delta\phi^{(s)}\delta\phi^{(s)} \rangle}{\langle \delta\phi^{(v)}\delta\phi^{(v)} \rangle} \simeq \frac{5 \cdot 10^{-12}}{\left(1 + \frac{\epsilon_\Lambda}{\epsilon_\phi}\right)^2} e^{7m_Q} m_Q^{11} N_\chi^2 r_{\text{vac}}^2 < 1. \tag{CON.6}$$

The Λ dependence of (CON.6) is more subtle. In deriving (CON.6), taking $\Lambda \gg 1$ is a built-in assumption because the non-linear contributions of the axion to the scalar power spectrum only dominate over linear contributions when $\Lambda^2 \gg 1$ [30]. This leads us to believe the contributions of (CON.6) are negligible when $\Lambda < 1$. Incidentally, (CON.6) is the only constraint that behaves differently in the classic S-CNI vs. mixed S-CNI branches, because the expression substituted for ϵ_ϕ is different for two branches (see (CON.3)). We, therefore, expect the two branches of S-CNI to show similar behaviors in the $\Lambda < \sqrt{2}$ regime where (CON.6) is irrelevant.

3.3 Low energy parameter space exploration

Now, armed with observational bounds and physical constraints, ensuring well-behaved solutions, we begin to explore the space that meets all of these criteria.

We start analyzing the effects of our combined constraints in the $\Lambda \gg \sqrt{2}$ region first. The approximate constraint given in (CON.3*), valid in this regime, is only a function

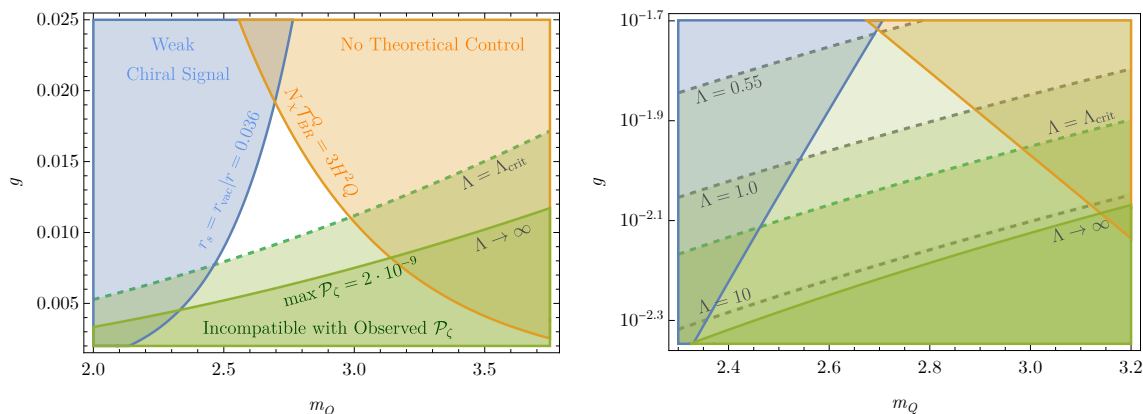


Figure 1. The various constraints from the low energy EFT are superposed to determine the allowed $m_Q - g$ parameter space. (CON.4), $\sqrt{2} < m_Q$, is satisfied everywhere on both panels. The blue shaded region doesn't provide a strong chiral gravitational signal and is in violation of (CON.2*). The orange shaded region experiences too much contamination from the backreaction (CON.5). The green shaded region with a solid border is incompatible with (CON.3*) and only predicts scalar power spectrum values that are smaller than $\mathcal{P}_\zeta^{\text{obs}} = 2.1 \cdot 10^{-9}$. Its solid green border corresponds to $\Lambda \rightarrow \infty$. (Left Panel) Only a small white region doesn't violate any of the constraints. As we lower Λ the solid green border moves higher and reduces the available parameter space (see (3.6)). The green shaded region with a dashed border breaks (CON.3) for values of $\Lambda \geq \Lambda_{\text{crit}} \sim \sqrt{2}$ (see (4.6)). (Right Panel) The viable region for $m_Q - g$ values (see left panel) superposed with dashed gray $\Lambda_{\text{min}} = \text{const.}$ contours that saturate (4.8). For any pair (m_Q, g) , there is a limit on how low Λ can go and still accommodate the correct value for \mathcal{P}_ζ . No choice of (m_Q, g) allows for $\Lambda \lesssim 0.55$, whereas $\Lambda \rightarrow \infty$ represents the irreducible, conservative, ruled-out region of parameter space where no value of Λ is viable.

of variables g and m_Q . We thus obtain the relation $m_Q \leq (32\pi^2 \mathcal{P}_\zeta)^{-1/4} \sqrt{g} \simeq 35\sqrt{g}$. This, in conjunction with (CON.4), fixes the value of viable m_Q within the g -dependent region $\sqrt{2} < m_Q \lesssim 35\sqrt{g}$ [30], as well as providing a lower-bound on the value for g , $(\frac{\sqrt{2}}{35})^2 \simeq 1.6 \cdot 10^{-3} \lesssim g$. Along the same slice of parameter space we may impose (CON.5) to restrict to regions with controlled backreaction.

Finally, we consider (CON.1) and (CON.2) along these axes: for fixed g and m_Q the value of ϵ_B is determined by the scale of inflation $\epsilon_B \sim H^2$. Thus the sourced contribution to r inherits an even steeper dependence on H than its vacuum counterpart, $r_{\text{vac}} \sim H^2$ vs. $r_s \sim H^4$ — the strength of distinctive S-CNI signatures is enhanced significantly for higher scales of inflation. H is, however, forbidden from becoming arbitrarily large by the observational bound (CON.1), or equivalently the saturation of this bound sets a maximum scale of inflation for each $\{g, m_Q\}$. This in turn sets a maximally viable signal ratio \mathcal{R}_{GW} , to which we impose (CON.2) for a conservative constraint on this space. This method of obtaining a maximally viable (CON.2) is equivalent to setting $r_{\text{vac}} = 0.018$ in (CON.2*). The synthesis of these various constraints, explored in figure 1 (left panel), delineates a highly confined region of g and m_Q where observable, predictive, and viable S-CNI theories may reside.

We can read off the small allowed range of g and m_Q as

$$4.5 \cdot 10^{-3} \lesssim g \lesssim 1.9 \cdot 10^{-2}, \tag{3.11}$$

$$2.3 \lesssim m_Q \lesssim 3.1. \tag{3.12}$$

These ranges are similar to those quoted in [30], but differ slightly because of the updated observational bound on r , our new expression for \mathcal{F}^2 , and our inclusion of the time integration factor in the backreaction constraint. Intuitively speaking, too large of a coupling generates too much backreaction or too little tensor perturbation. On the other hand, too small of a coupling causes the gauge field to be too light to source \mathcal{P}_ζ .

As we move away from the large Λ limit, the only constraint that changes is (CON.3). Increasingly smaller values of Λ puts even more stringent bounds on the $g - m_Q$ parameter space; however, the allowed region is centered on more or less the same values. The effect of decreasing Λ is illustrated in figure 1 (right panel), and we find that no viable parameter space exists for $\Lambda \lesssim 0.55$.

In figure 2, we take $g = 0.01$ (which lies almost in the center of the narrow interval (3.11)), and show the overlap of all constraints discussed in this section in the $H - m_Q$ plane. Note that (CON.6) is only valid for $\Lambda^2 \gg 1$ and has different expressions depending on which branch of the S-CNI model we pick; it is the only curve that changes between the left (classic S-CNI) and right (mixed S-CNI) panels. It is clear from the figure that the classic S-CNI branch is unable to satisfy all the constraints at the same time when $\Lambda \gg 1$, but the mixed S-CNI branch does so in a small region of available values for H/M_{Pl} :

$$6.2 \cdot 10^{-6} \lesssim H/M_{\text{Pl}} \lesssim 1.0 \cdot 10^{-5}, \tag{3.13}$$

and it follows that the same parameter space is available to the classic S-CNI case in the limit of small Λ where (CON.6) vanishes. Notice that the above argument assumes a benchmark value of $N_\chi \sim 10$, whereas for $N_\chi \lesssim 5$, (CON.6) relaxes and allows a small viable space for the classic branch as well as the mixed one. Much like our procedure for figure 1, we have drawn a dashed green line to show how (CON.3) would change as one decreases Λ . The available values for H/M_{Pl} only become stricter for smaller $\Lambda \sim \sqrt{2}$. This indicates that although (CON.6) relaxes in this regime, the parameter space newly opened up in the classic S-CNI scenario may yet be challenged by the more severe (CON.3).

To summarize, the phenomenological constraints confine three of the parameters in the model to narrow ranges around $g \simeq 0.01$, $H \simeq 8.2 \cdot 10^{-6} M_{\text{Pl}}$, and $m_Q \simeq 2.7$. These further imply that $Q \simeq 1.6 \cdot 10^{-3} M_{\text{Pl}}$. We also show that for $\Lambda \gg \sqrt{2}$, the only viable branch is the mixed S-CNI scenario. To give some sense of the possible values of other parameters: $\lambda = 60$ and $\Lambda = 10$ would imply $\mu \approx 5.4 \cdot 10^{-4} M_{\text{Pl}}$ (which is relatively insensitive to Λ provided $\Lambda \gg 1$) and $f \approx 1.4 \cdot 10^{-2} M_{\text{Pl}}$.

4 Assessing the Λ range in S-CNI

From (3.11), (3.12), and (3.13), we have fairly narrow constraints on the parameters g , m_Q , and H . Thus, it is useful to rewrite the equation of motion involving Q in terms of

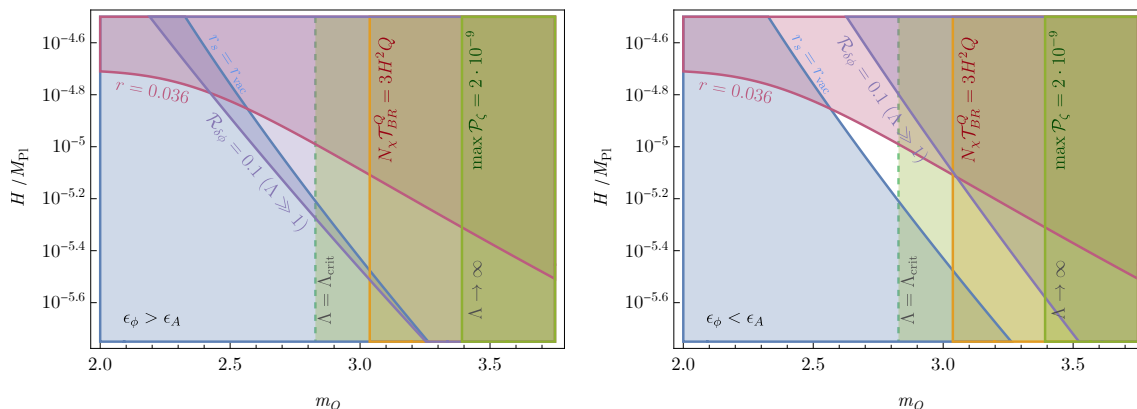


Figure 2. Allowed parameter space of S-CNI in the plane of $m_Q - H/M_{\text{Pl}}$. $\sqrt{2} < m_Q$ is satisfied and $g = 0.01$ is fixed everywhere on both panels. The panels show all constraints discussed in this section in classic (left) and mixed (right) S-CNI scenarios: the green shaded region with solid border breaks (CON.3*), the dark yellow region is contaminated by (CON.5), the green shaded region with dashed border violates (CON.3) for $\Lambda \sim \sqrt{2}$, the dark pink region breaks the observational bound $r < 0.036$ (CON.1), the chiral signature in the blue region is too weak (see (CON.2)), and the purple region is contaminated with too large a non-linear contribution to the scalar power spectrum, (CON.6). At small Λ this constraint becomes irrelevant. (Left Panel) For large $\Lambda^2 \gg 1$ values, the $\mathcal{R}_{\delta\phi} > 0.1$ constraint kills the available parameter space for enhanced chiral signatures of the classic S-CNI branch. (Right Panel) The mixed S-CNI branch satisfies all constraints for $6.2 \cdot 10^{-6} \lesssim H/M_{\text{Pl}} \lesssim 1.0 \cdot 10^{-5}$.

these essentially fixed parameters and Λ , following definitions in (2.4). In this section, we will discuss the allowed values of Λ , showing that they fall into two regimes, one of which is more promising than the other. The results of this section will play an important role in the constraints that we present in subsequent sections.

Slow-roll solutions have an approximately constant, nonzero, value of Q during the CNI phase, determined by (2.20). Such a value must be a solution to the polynomial equation

$$g^2 \lambda^2 Q^4 + 2g^2 f^2 Q^2 + 2f^2 H^2 + \frac{g\lambda f}{3H} U'(\chi) Q = 0, \quad (4.1)$$

where we view χ as being approximately constant when solving for Q . For concreteness, here and in subsequent sections we will take $\chi/f \approx \pi/2$, which maximizes the size of the gravitational wave perturbations (see, e.g., appendix A of [51]) and sets $U'(\chi) \approx -\mu^4/f$. Viewed as an equation for Q , (4.1) is a messy quartic equation. However, we can rewrite it as a quadratic equation for Λ , with $g, \mu^4/f, H$, and m_Q held fixed:

$$\Lambda^2 - \frac{\mu^4/f}{3H^3 m_Q^2/g} \Lambda + 2(1 + m_Q^{-2}) = 0. \quad (4.2)$$

This equation has the two solutions

$$\Lambda_{1,2} = \frac{\mu^4/f}{6m_Q^2 H^3/g} \left[1 \pm \sqrt{1 - 2(\mu^4/f)^{-2} (6m_Q^2 H^3/g)^2 (1 + m_Q^{-2})} \right]. \quad (4.3)$$

Because Λ is real by definition, we immediately see that a viable parameter space requires that the argument of the square root is positive, i.e., that

$$\frac{\mu^4}{f} \geq \frac{6\sqrt{2}m_Q H^3}{g} \sqrt{m_Q^2 + 1} \equiv \eta H^3, \quad (4.4)$$

where we have defined

$$\eta \equiv \frac{6\sqrt{2}m_Q}{g} \sqrt{m_Q^2 + 1} \approx 4.7 \cdot 10^3, \quad (4.5)$$

for benchmark values $g = 0.014$ and $m_Q = 2.7$. In particular, this gives a lower bound on the larger root Λ_1 , saturated when the square root is zero, and a similar upper bound on the smaller root Λ_2 ,

$$\begin{aligned} \Lambda_1 &\geq \Lambda_{\text{crit}} \equiv \sqrt{2} [1 + m_Q^{-2}]^{1/2} > \sqrt{2}, \\ \Lambda_2 &\leq \Lambda_{\text{crit}}. \end{aligned} \quad (4.6)$$

The first inequality selects Λ_1 as the solution to the region of the parameter space where $\Lambda \gg \sqrt{2}$ is satisfied. If we are to explore regions of the parameter space where $\Lambda \leq \sqrt{2}$, we need to pick Λ_2 as the solution relating λ to f and μ . In the S-CNI literature, most of the attention has been directed toward the Λ_1 solution; as already noted in section 3, assuming $\Lambda \gg \sqrt{2}$ simplifies the analytic expressions.

A good solution of (4.1) should be an attractor of the Q equation of motion, i.e., a value at which $W''_{\text{eff}}(Q) > 0$ (see (2.20)). Taking a derivative of the right-hand side of (2.20) and substituting the value of μ^4/f determined by (4.1), one finds that a given solution is a minimum provided that

$$\Lambda^2 \geq \frac{2}{3} (m_Q^{-2} - 1), \quad (4.7)$$

a condition that is always satisfied in our parameter space of interest (CON.4). Hence the solutions (4.3) are always good minima of the effective potential.

The solution Λ_2 can correspond to small values of λ , which may not provide sufficiently many e-folds of S-CNI physics (2.24) because χ rolls too quickly. Another consequence of fast-rolling χ is a modification of the scalar power spectrum. In fact, combining (CON.3) and (3.6), we find a constraint

$$\Lambda \geq \Lambda_{\text{min}} \equiv \frac{1 + m_Q^2}{g} 8\pi \sqrt{P_\zeta} \left[1 - \frac{32\pi^2 P_\zeta}{g^2} m_Q^2 (1 + m_Q^2) \right]^{-1/2} \simeq 0.74, \quad (4.8)$$

using benchmark values ($g = 0.014$, $m_Q = 2.7$) and the measured value P_ζ^{obs} in the last step. Thus we see that, given the requirement of a valid solution with a scalar power spectrum matching observational data, the Λ_2 regime offers only limited room below $\Lambda = \sqrt{2}$.

Note that figure 1 (right panel) has made clear that the constraints (CON.2*), (CON.5), and (CON.3) represent severe restrictions on g , m_Q , and indeed Λ . We see that the lowest viable value of Λ is $\simeq 0.55$, where the parameter space vanishes to a point; said point is extremal in the $g - m_Q$ plane but the most accommodating in Λ , allowing for all $\Lambda \geq \Lambda_{\text{min}} \simeq 0.55$. Our benchmark values are chosen such that they are comparatively

central in the allowed $g - m_Q$ region, but a relatively large range of $\Lambda \lesssim \Lambda_{\text{crit}}$ can be accommodated therein.

This is a significant result which is independent of any UV completion: even though, unlike the original CNI, $\Lambda \lesssim \sqrt{2}$ was hypothesized as a viable parameter space for S-CNI, not a lot of freedom has been found upon exploration.

It is worth noting, however, that the blue and orange regions carved out in figure 1 are not strictly speaking “no-go” zones, and thus smaller values of Λ are likewise in principle possible to achieve. However, more work must be done in order to predict or discover S-CNI in our universe: if the chiral tensor modes are too small (violating (CON.2*)), then precise data and analysis techniques are required to extract any CNI signal. If they are too large (violating (CON.5)), then a more robust treatment of these equations must be developed before predictions can be made.

5 The S-CNI model and clockwork

In section 3, we detailed constraints on the S-CNI model based primarily on physical consistency with our observed universe. Having cut out much of the naïve parameter choices, in this section we examine whether the remaining parameter space admits a UV completion in terms of the clockwork scenario [34–36]. The analogous question in the original CNI model was studied in [33], which provides the basis for our analysis.

5.1 Clockwork constraints

As discussed in section 2.3, the axion periodicity $\chi \cong \chi + 2\pi f_\chi$ imposes a quantization condition on the coupling: $\lambda = \frac{j \cdot k \cdot g^2}{8\pi^2}$, where k and j are integers generated from different physics in the UV completion. The integer k can arise, for example, from integrating out fermions carrying SU(2) gauge charge that have χ -dependent masses. In order for such an effective theory of fermions and gauge fields to be valid, we require $k < 4\pi/g^2$ [33].

The clockwork mechanism [34–36] is a way to obtain a large integer j as a product of smaller integers. It is, essentially, an iterated version of an older idea of “axion alignment” [42]. One has a set of n axions χ_n with a potential of the form

$$\sum_{i=1}^{n-1} \mu_{i+1}^4 \cos \left[\frac{m_i \chi_i}{f_i} + \frac{\chi_{i+1}}{f_{i+1}} \right] + \mu_1^4 \cos \frac{\chi_1}{f_1}, \tag{5.1}$$

with m_i a small integer. Integrating out the heaviest modes leads to one parametrically light mode χ with an effective potential having $j = \prod m_i$. In this way, one can obtain an exponentially large enhancement of λ through a product of small integers.

In this setting, a bound can be derived using the fact that each cosine term in (5.1) mediates axion scattering, and perturbative unitarity provides an upper bound on the corresponding ratios like μ_{i+1}^4/f_i^4 . This was shown to imply $j < f/\mu$ [33], leading to the conclusion

$$\lambda \leq \frac{f}{2\pi\mu}. \tag{5.2}$$

5.2 A lower bound on f from clockwork

The clockwork constraint (5.2) gives an upper bound on μ , for fixed f and λ . We previously derived a lower bound on μ , (4.4), simply from the existence of a stationary value of Q in the equation of motion. We can combine these two constraints to obtain a range of μ ,

$$(\eta H^3 f)^{1/4} \leq \mu \leq \frac{f}{2\pi\lambda} \lesssim \frac{f}{4\pi\xi N_\chi}, \quad (5.3)$$

where in the last step we used the requirement (2.24) that λ be large enough to allow for N_χ e-folds of χ rolling. Requiring that the range (5.3) is not empty now translates into a *lower bound* on f ,

$$f \gtrsim \eta^{1/3} (4\pi\xi N_\chi)^{4/3} H \simeq 4.7 \cdot 10^4 H, \quad (5.4)$$

where we substituted $m_Q = 2.7$, $g = 0.014$, and $N_\chi = 10$ in the last step for a numerical estimate. Together with the range (3.13) of allowed Hubble values, this already shows that achieving a viable clockwork scenario would require f to be very near the Planck scale.

5.3 Restricting the Λ range with clockwork

We can also rearrange our existing constraints to obtain an *upper* bound on f , assuming a lower bound on Λ . In the Λ_1 regime, we have such a bound from the basic constraint (4.6). In the Λ_2 regime, we have the lower bound (4.8) that originated from the power spectrum constraint (CON.3). By rewriting the clockwork constraint (5.2) in terms of Λ , we can express this as a range of allowed Λ values:

$$\Lambda_{\min} \leq \Lambda \leq \Lambda_{\max}^{\text{CW}}(\mu), \quad (5.5)$$

where

$$\Lambda_{\max}^{\text{CW}}(\mu) = \frac{m_Q H}{2\pi g \mu}. \quad (5.6)$$

Imposing $\Lambda_{\min} \leq \Lambda_{\max}^{\text{CW}}(\mu)$ gives an upper bound on μ , which we can combine with the lower bound (4.4) $\mu^4 \geq \eta H^3 f$ to obtain

$$(\eta H^3 f)^{1/4} \leq \mu \leq \frac{m_Q H}{2\pi g \Lambda_{\min}}. \quad (5.7)$$

Because we have eliminated the f -dependence on the right-hand side by rewriting λ/f in terms of Λ , we thus learn that the existence of a clockworkable range of Λ implies an *upper bound* on f :

$$f \leq H \frac{m_Q^2}{6\sqrt{2}(2\pi)^4 g^3 (1 + m_Q^{-2})^{1/2} \Lambda_{\min}^4} \simeq 620H, \quad (5.8)$$

again using $m_Q = 2.7$ and $g = 0.014$ in the last step for a numerical estimate and taking $\Lambda_{\min} \simeq 0.74$ from (4.8). In the Λ_1 branch, a stronger constraint (by about a factor of 17) can be obtained using Λ_{crit} for Λ_{\min} .

The upper bound (5.8) directly clashes with the lower bound (5.4) that we derived above, showing that there is no clockworkable parameter range.

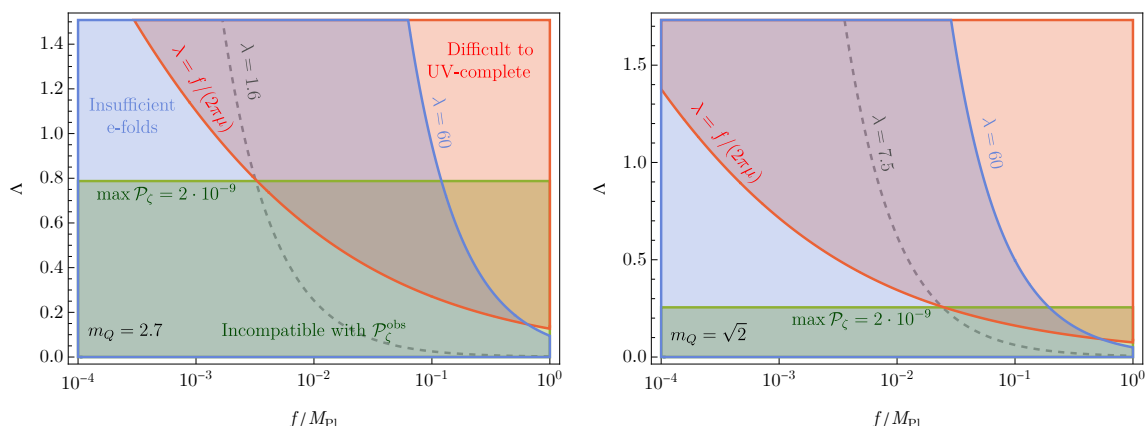


Figure 3. Exploring the $\Lambda_2 \leq \Lambda_{\text{crit}} \sim \sqrt{2}$ (see (4.6)) solutions with various constraints. $g = 0.014$ and $H = 8.2 \cdot 10^{-6}$ for both panels. The red shaded region is forbidden by clockwork (5.9). The blue shaded region excludes point with $N_\chi < 10$ (5.10). The green shaded area doesn't produce correct values for \mathcal{P}_ζ . The small area at the bottom right of both panels, trapped between the red and blue curves, corresponds to f values from (5.4). Not that all regions outside of the blue shaded area correspond to $f \sim 10^{-1} M_{\text{Pl}}$. (Left Panel) Choosing the benchmark value $m_Q = 2.7$ allows us to lower Λ_{min} down to 0.74 (see (4.8)). No clockworkable solution with $\lambda \geq 60$ and correct \mathcal{P}_ζ exists. (Right Panel) Lowering benchmark values to $m_Q = \sqrt{2}$ doesn't resolve the UV completion challenge.

5.4 Numerical results

To better understand how the clockwork upper bound on f clashes with the lower bound on Λ , we present a graphical rendition of our equations here. In figure 3 (left panel), the red shaded region is prohibited by the clockwork constraint. We can find the inequality that describes this region by eliminating μ from the right hand side of (5.5). We replace μ in terms of f and Λ by solving (4.2) and find

$$f \leq \frac{H m_Q^2 / g^3}{48 \pi^4 \Lambda^3 [\Lambda^2 + 2(1 + m_Q^{-2})]}, \quad (5.9)$$

as the expression for the clockwork region in the $f - \Lambda$ plane. Notice that this is a Λ -dependent upper bound on f , which becomes arbitrarily weak at small Λ ; this is in contrast to (5.8) above, which used the minimum Λ from the power spectrum to give a Λ -independent bound on f . If we meet the condition (5.9) while staying above contours of $\lambda = 60$, we have found the region which satisfies the lower bound (5.4). The $\lambda = \text{const.}$ contours linearly decay as f gets closer to M_{Pl} . We exclude contours with smaller values for λ in our plot by

$$\Lambda f = \lambda m_Q H / g \gtrsim 60 \cdot m_Q H / g. \quad (5.10)$$

The only region meeting all these requirements resides at the very bottom right of the left panel. The f values residing in this region are extremely close to M_{Pl} , and, indeed, coincide with the same ones that satisfy our clockwork lower bound on f .

Both panels in figure 3 span the range $0 \leq \Lambda < \Lambda_{\text{crit}}$. Consequently, $\Lambda_1 \geq \Lambda_{\text{crit}}$ solutions have no chance of being clockworked.

Now, remember the green shaded region is forbidden for our Λ_2 solutions because it fails to produce the correct \mathcal{P}_ζ value; hence, Λ_2 solutions are also kept far from the bottom right region. Notice that even if we had picked different g and m_Q values from (3.11) and (3.12) to minimize the solid green border to $\Lambda_{\min} \simeq 0.55$ (see figure 1 right), we would still be too far above the desired values.

It's worth noting that even without clockwork, taking $\Lambda \lesssim \sqrt{2}$ puts f within an order of magnitude of M_{Pl} . This only happens if one abides by $\lambda \gtrsim 60$ and the observed value of \mathcal{P}_ζ . So, in addition to accounting for a large coefficient of the Chern-Simons term, justifying this near Planck scale value of f is a second concern that alternative UV completions in this regime have to address.

To provide yet another point of view, observe that for our choice of benchmark values, the gray λ contour indicates S-CNI doesn't enter the clockworkable regime until $\lambda \lesssim 1.6$. This value for λ corresponds to $N_\chi \sim 0.25$ which doesn't make for enough e-folds to observe any interesting phenomenology. We have also checked that numerical solutions of the classical equations of motion (2.15), (2.16), (2.17), and (2.19) in this parameter regime are inconsistent with slow-roll evolution of χ and Q .

Breaking the g upper bound from (3.11) results in losing control over the backreaction effects (see figure 1 left), but one may break the m_Q lower bound (3.12) with the aim of obtaining even smaller values for Λ_{\min} . Taking $m_Q \lesssim 2.7$ violates the $\mathcal{R}_{GW} > 1$ bound and results in a less chiral signal. We lowered the bound to $m_Q = \sqrt{2}$ (see figure 3 right) below which we face tachyonic instabilities in scalar perturbations. Even at this new minimum, S-CNI doesn't become clockworkable until $\lambda \lesssim 7.5$ with a corresponding $N_\chi \sim 1.2$ (gray dashed contour). As mentioned earlier, this λ is still too small to account for enough e-folds of interesting phenomenology.

To summarize, even venturing beyond the $\Lambda > \sqrt{2}$ limit and localizing to small Λ_2 solutions still results in a large Chern-Simons coefficient that cannot be clockworked. We conclude that S-CNI is not compatible with a clockwork UV completion in any region of its available parameter space.

6 Additional constraints

In this section, we put clockwork aside and consider whether the remaining parameter space from 3.3 is well-behaved as a quantum field theory, and challenge this phenomenologically-viable region with more fundamental (albeit sometimes conjectural) criteria. First, we require the couplings and scales that describe the effective theory of S-CNI to respect partial wave unitarity bounds. Next, we consider what portions of the remaining parameter space are consistent with the weak gravity conjecture.

6.1 Perturbative unitarity bounds

The partial wave unitarity bound states that EFT amplitudes cannot be arbitrarily large, if we want our perturbative approach to remain reliable. The perturbative unitarity bound on the $\chi\chi \rightarrow \chi\chi$ scattering amplitude arising from the cosine potential for χ was previously discussed in [33], and played a role in the derivation of the bound (5.2) on λ in the clockwork

scenario. In this section we will give a more *direct* bound on λ , based on the scattering of gauge fields and axions mediated by the $\lambda\chi F\tilde{F}$ coupling. Similar perturbative unitarity bounds on such a coupling have recently been derived (with very different applications in mind) in refs. [52, 53]. The basic idea of such bounds is to use the unitarity of the S -matrix, together with a decomposition of scattering states into angular momentum modes, to bound partial-wave amplitudes which are functions of center-of-mass energy \sqrt{s} alone. Specifically, for a $2 \rightarrow 2$ scattering process of states with momentum along the z axis and incoming helicities λ_1, λ_2 into outgoing states at angles (θ, ϕ) with outgoing helicities $-\lambda_3, -\lambda_4$, the amplitude decomposes as

$$\mathcal{M}_{\{\lambda_i\}}(s, \theta, \phi) = 16\pi \sum_j (2j+1) \sqrt{(1 + \delta_{\lambda_1\lambda_2})(1 + \delta_{\lambda_3\lambda_4})} e^{i(\lambda+\mu)\phi} d_{\lambda, -\mu}^j(\theta) \mathcal{T}_{\{\lambda_i\}}^j(s). \quad (6.1)$$

Here $\lambda = \lambda_1 - \lambda_2$ and $\mu = -(\lambda_3 - \lambda_4)$. The combination $e^{i(\lambda+\mu)\phi} d_{\lambda, -\mu}^j(\theta)$ corresponds to a Wigner D -function evaluated on this specific kinematic configuration. The D -functions have simple orthogonality properties, so it is possible to read off a given partial-wave amplitude $\mathcal{T}_{\{\lambda_i\}}^j(s)$ by integrating the full amplitude \mathcal{M} against a D -function. In the special case $\lambda = \mu = 0$, these reduce to the familiar Legendre polynomials. The partial wave unitarity bound constrains the size of $\mathcal{T}_{\{\lambda_i\}}^j(s)$ for an *elastic* scattering process, at any s :

$$|\mathcal{T}^j(s) - i/2| \lesssim 1/2. \quad (6.2)$$

In particular, for the amplitudes we will consider where the leading tree-level contribution to $\mathcal{T}^j(s)$ is real, we will demand that this contribution be smaller than $1/2$. For more details, a textbook discussion appears in [54].

To generalize the partial wave unitarity bound for the S-CNI model, we computed the tree level scattering amplitudes of gauge field axion scattering from

$$, \text{ and } \quad (6.3)$$

We have computed these amplitudes using the spinor helicity formalism; in this case, we chose the external χ to be massless to facilitate the computation (expecting that perturbative unitarity will break down only at energies $\sqrt{s} \gg m_\chi$). We find:

$$\mathcal{M}_{+-}^{A\chi} = \delta^{ab} \frac{u+s}{su} \left(\frac{\lambda}{2f}\right)^2 \langle 14 \rangle^2 [34]^2 \propto \delta^{ac} \left(\frac{\lambda}{4f}\right)^2 2s(1 - \cos\theta), \quad (6.4)$$

$$\mathcal{M}_{++}^{A\chi} = -2\delta^{ac} \left(\frac{\lambda}{2f}\right)^2 \langle 13 \rangle^2 \propto \delta^{ab} \left(\frac{\lambda}{4f}\right)^2 4s(1 - \cos\theta). \quad (6.5)$$

These calculations are presented in the all-incoming convention for momenta and helicities, i.e., the subscripts $+$ and $-$ correspond to the incoming helicity of the two gauge fields (particles 1 and 3). The δ^{ac} factor ensures that the color of the outgoing gauge field is the

same as that of the incoming gauge field. We use \propto here to indicate that we have dropped *phases*, that is, we have carried out replacements like $\langle 14 \rangle \rightarrow \sqrt{u}$. Constant prefactors have not been dropped. The angular dependence of the amplitude $\mathcal{M}_{++}^{A\chi}$ is exactly that of the Wigner D -function $D_{1,-1}^1$, whereas the amplitude $\mathcal{M}_{+-}^{A\chi}$ has nontrivial overlap with a large set of D -functions $D_{1,1}^j$.

A second unitarity constraint shows up when considering the gauge boson scatterings that are mediated by the axion,

$$\begin{array}{c}
 \epsilon_1 \quad \epsilon_3 \\
 \diagdown \quad \diagup \\
 \text{---} \text{---} \text{---} \\
 \diagup \quad \diagdown \\
 \epsilon_2 \quad \epsilon_4
 \end{array},
 \quad
 \begin{array}{c}
 \epsilon_1 \quad \epsilon_3 \\
 \diagdown \quad \diagup \\
 \text{---} \text{---} \text{---} \\
 \diagup \quad \diagdown \\
 \epsilon_2 \quad \epsilon_4
 \end{array},
 \quad \text{and} \quad
 \begin{array}{c}
 \epsilon_1 \quad \epsilon_3 \\
 \diagdown \quad \diagup \\
 \text{---} \text{---} \text{---} \\
 \diagup \quad \diagdown \\
 p_2 \quad p_4
 \end{array}.
 \tag{6.6}$$

In this case we use spinor helicity for the external gauge bosons while keeping the mass in the internal propagator, finding:

$$\mathcal{M}_{++--}^{AA} = \left(\frac{\lambda}{2f}\right)^2 \frac{\delta^{ab}\delta^{cd}}{s - m_\chi^2} \langle 12 \rangle^2 [34]^2 \propto \left(\frac{\lambda}{4f}\right)^2 \frac{4s^2 \delta^{ab}\delta^{cd}}{s - m_\chi^2}, \tag{6.7}$$

$$\begin{aligned}
 \mathcal{M}_{+-+-}^{AA} &= \left(\frac{\lambda}{2f}\right)^2 \frac{\delta^{ac}\delta^{bd}}{t - m_\chi^2} \langle 24 \rangle^2 [13]^2 \propto \left(\frac{\lambda}{4f}\right)^2 \frac{4t^2 \delta^{ac}\delta^{bd}}{t - m_\chi^2} \\
 &\approx -\left(\frac{\lambda}{4f}\right)^2 \delta^{ac}\delta^{bd} 2s(1 - \cos\theta), \tag{6.8}
 \end{aligned}$$

$$\begin{aligned}
 \mathcal{M}_{+---}^{AA} &= \left(\frac{\lambda}{2f}\right)^2 \frac{\delta^{ad}\delta^{bc}}{u - m_\chi^2} \langle 23 \rangle^2 [14]^2 \propto \left(\frac{\lambda}{4f}\right)^2 \frac{4u^2 \delta^{ad}\delta^{bc}}{u - m_\chi^2} \\
 &\approx -\left(\frac{\lambda}{4f}\right)^2 \delta^{ad}\delta^{bc} 2s(1 + \cos\theta), \tag{6.9}
 \end{aligned}$$

$$\mathcal{M}_{++++}^{AA} = -\left(\frac{\lambda}{4f}\right)^2 \left[\frac{\delta^{ab}\delta^{cd}}{s - m_\chi^2} [12]^2 [34]^2 + \frac{\delta^{ac}\delta^{bd}}{t - m_\chi^2} [34]^2 [13]^2 + \frac{\delta^{ad}\delta^{bc}}{u - m_\chi^2} [14]^2 [23]^2 \right]. \tag{6.10}$$

Again, we have used the all-incoming convention for momenta and helicities; $a, b, c,$ and d are the colors of the gauge fields 1, 2, 3, and 4. Expressions following \propto have dropped phases, while expressions following \approx have been taken to the high-energy limit and hence set $m_\chi \rightarrow 0$.

Simply by dimensional analysis, all of these amplitudes grow with energy roughly as $\lambda^2 E^2 / f^2$. However, the precise bound on a partial wave amplitude depends on the constant coefficient as well as the overlap of the angular structure in \mathcal{M} with the appropriate Wigner D -function for the partial wave. From inspecting the leading growth with s of the elastic scattering processes we have computed, we find that the amplitude \mathcal{M}_{++--}^{AA} with all four gauge fields of the same color leads to a relatively stringent bound:

$$\sqrt{s} \leq 8\sqrt{\pi} \frac{f}{\lambda}. \tag{6.11}$$

This is a bound on the center-of-mass energy at which the $2 \rightarrow 2$ scattering process can be approximately described by the EFT. If we further suppose that the physics describing

the origin of the scale μ should be describable within an EFT containing the $\lambda\chi F\tilde{F}$ term, we would conclude that

$$\mu \lesssim 8\sqrt{\pi}\frac{f}{\lambda}. \tag{6.12}$$

It is intriguing that this bound has the same parametric form as the constraint (5.2) arising in clockwork models. However, it is quantitatively weaker; the coefficient $8\sqrt{\pi}$ is larger than the coefficient $1/(2\pi)$ by about two orders of magnitude. By scattering more general states which are superpositions of different colors and helicities and finding eigenvalues of the \mathcal{T} matrix, it is possible that we could refine the perturbative unitarity bound by an $O(1)$ factor. However, we do not expect that a bound along these lines can be nearly as strong as the clockwork bound. The prefactor in the bound on μ appeared to the $4/3$ power in the lower bound on f (5.4) and to the *fourth* power in the upper bound on f (5.8). Thus, unlike the clash between these bounds that arose in the clockwork-based argument, there are many orders of magnitude of f , from $\sim 100H$ to $\sim 10^{10}H$, allowed by the perturbative unitarity bound (6.12).

Although the perturbative unitarity constraint on the scattering of axions and gauge bosons is not strong enough by itself to eliminate the interesting S-CNI parameter space, we present the bounds here in case they could form a useful starting point for further exploration. In recent years, the general logic of unitarity bounds has been extended in the direction of quite powerful S -matrix bootstrap and positivity bound techniques, which might be interesting to apply to the S-CNI model in the future (see, e.g., [55, 56] for recent surveys).

6.2 The weak gravity conjecture constraint

Theories of quantum gravity are expected to lack any global symmetries (see, e.g., [57, 58]). From this general idea, a number of quantitative (but conjectural) constraints on EFTs coupled to gravity have emerged in recent years. These potentially lead to important constraints on variations of chromo-natural inflation, though as we will see, obtaining strong constraints requires going beyond the most minimal conjectures.

The Weak Gravity Conjecture (WGC) provides a quantitative explanation of what goes wrong in a theory that has an approximate global symmetry due to a small gauge coupling. In particular, the magnetic WGC holds that an effective gauge theory with coupling g must break down at energies $\Lambda_{UV} \lesssim gM_{\text{Pl}}$ [59]. Although the original form of this conjecture applied only to U(1) gauge theories, subsequent work has argued that it should apply to nonabelian gauge theories as well [32, 60] (see also [61] and the review [62]). Because the original (non-spectator) CNI model required an extremely small gauge coupling, $g \sim 10^{-6}$, it ran into immediate tension with the WGC [32].

The spectator CNI models, preferring $g \sim 10^{-2}$ (3.11), are somewhat less constrained, but we still conclude from this that the EFT should break down at or below the scale

$$\Lambda_{UV} \lesssim gM_{\text{Pl}} \lesssim 4 \cdot 10^{16} \text{ GeV}. \tag{6.13}$$

While the minimal WGC arguments do not specify precisely what happens at this scale, in practice it is generally a radical departure from 4d local EFT (e.g., extra dimensions or

high-spin string states appear), so we should require 4d mass scales like μ and H to be below this bound. The parameter space of interest satisfies this constraint with room to spare.

The WGC also constrains theories of axions. The minimal axion WGC requires that, for an axion χ of decay constant f_χ , there must be instantons with action $S_\chi \lesssim n_{\text{inst}} M_{\text{Pl}}/f_\chi$. Checking this constraint on our model is not completely straightforward, without knowing the microscopic origin of the μ^4 factor in the axion potential. However, it is reasonable to expect that $\mu^4 \sim \Lambda_{\text{UV}}^4 \exp(-S_\chi/b)$ where S_χ is an instanton action, for some $O(1)$ coefficient b , and thus $S_\chi \sim O(1) \log(\Lambda_{\text{UV}}/\mu)$. In the parameter space of interest (see the end of section 3.3), $\log(M_{\text{Pl}}/\mu) \sim 7$, which is smaller than $M_{\text{Pl}}/f \sim 10^2$. Thus, the axion WGC bound is expected to be satisfied by whatever UV physics generates the axion potential.⁴ There is also a *magnetic* axion WGC, which requires the existence of an axion string (around which the χ field winds) with tension $T_\chi \lesssim 2\pi f_\chi M_{\text{Pl}} = 2\pi f M_{\text{Pl}}/j$. Perturbativity of the gauge theory requires that $kg^2/(4\pi) \lesssim 1$, which requires that $j \gtrsim 2\pi\lambda \gtrsim 4 \cdot 10^2$. If $f \sim 10^{-2} M_{\text{Pl}}$, this implies that the string mass scale is $\sqrt{2\pi T_\chi} \lesssim 10^{-2} M_{\text{Pl}}$. This is safely larger than μ or H , so there is no reason to expect the axion string modes to alter the dynamics of the model. Thus, we conclude that the S-CNI model can survive both the electric and magnetic forms of the axion WGC.

A potentially more interesting constraint arises from the conjecture that quantum gravity requires axions to eliminate a (-1) -form Chern-Weil symmetry associated with the $SU(2)$ instanton number [63]. The axion χ couples to two different kinds of instantons: those responsible for generating the $U(\chi)$ term, and the $SU(2)$ instantons. Unless these two instantons are related in the UV, e.g., through unification (unlikely, since they have very different actions), this means that χ gauges only one linear combination of the two (-1) -form symmetries. Another axion θ , then, would be required to eliminate the remaining $SU(2)$ instanton number symmetry. The axion WGC applied to this other axion would imply that it has a decay constant $f_\theta \lesssim \frac{g^2}{8\pi^2} M_{\text{Pl}} \sim 3 \cdot 10^{12} \text{ GeV} \left(\frac{g}{10^{-2}}\right)^2$. This is a more interesting constraint, because it is close to the scale $H/(2\pi)$ inferred from (3.13). Hence, this hypothetical second axion would fluctuate significantly during inflation, leading to the production of strings around which θ winds; call these θ -strings. The magnetic axion WGC for θ indicates that these would have tension

$$T_\theta \lesssim 2\pi f_\theta M_{\text{Pl}} \lesssim \frac{g^2}{4\pi} M_{\text{Pl}}^2. \tag{6.14}$$

The production of loops of θ -string has an energy cost, which should not form a significant source of backreaction on inflation. We can estimate this as follows: in a region of size L , we expect the typical fluctuation of the field θ over one Hubble time to be of order $\frac{H}{2\pi}(HL)$. We form a string when this is of order $2\pi f_\theta$, which occurs in a region of length $L_* \sim \frac{(2\pi)^2 f_\theta}{H^2}$, with energy $T_\theta L_*$. This should be compared to the rate at which energy is ordinarily depleted from a region of volume L_*^3 during one Hubble time, $\sim \epsilon_H (H^2 M_{\text{Pl}}^2) \frac{L_*^3}{H}$. Demanding that

⁴We have $1/f_\chi = j/f$; naively the cosine potential appears to have $n_{\text{inst}} = 1/j$, but this can arise from fundamental instantons of charge $(0, 1)$ and $(1, j)$ in a two-axion model [42]. These considerations do not seem to alter the conclusion, at least without a detailed model in which the $O(1)$ coefficients in the bound can be checked.

the string cost is small compared to the ordinary depletion of energy leads to the bound

$$T_\theta \ll \epsilon_H \frac{(4\pi^2 f_\theta)^2}{H^2} M_{\text{Pl}}^2. \tag{6.15}$$

We expect the θ -string mass scale $M_\theta \equiv \sqrt{2\pi T_\theta}$ to be a strong UV cutoff, in the sense that local QFT breaks down at this scale due to the appearance of fundamental high-spin states (see [64]). In the preferred parameter space of the S-CNI model, $\epsilon_H \sim 10^{-5}$, and the bounds (6.14) and (6.15) are numerically similar, with the former implying $M_\theta \lesssim 10^{16}$ GeV and the latter implying a somewhat weaker bound $M_\theta \ll 10^{17}$ GeV. Thus, this constraint is not so different from the gM_{Pl} WGC bound. However, it does have potentially important consequences: if the theory approximately saturates these bounds, then θ -strings potentially play a role in the inflationary dynamics. These strings would likely persist after inflation, because the SU(2) confinement scale is far below our current Hubble scale, so confinement does not occur to generate string-destroying domain walls. Such strings could be detected through their gravitational signatures; they would also emit SU(2) dark radiation, potentially detectable through N_{eff} measurements. (Other sources of SU(2) dark radiation are discussed in [65].)

Since the additional axion θ relies on a substantially stronger conjecture than the original WGC, we will not pursue a more detailed exploration of its consequences here. The implications for chromo-natural inflation of a completely different set of conjectural Swampland constraints are discussed in [66].

7 Conclusion

The chromo-natural inflation model and its relatives offer an appealing scenario that generates a distinctive chiral primordial gravitational wave signal. It is important to understand whether such models can actually be realized as consistent quantum field theories, and if so, whether these can be embedded in quantum gravity. We have built on earlier results [33] to argue that there is a substantial challenge to realizing a UV completion of the Spectator CNI model of ref. [20]. The basic problem is that the effective coupling λ of the axion to SU(2) gauge fields must be large (see eq. (2.24)), in order to support several e-folds in which the CNI phenomenology is operative. On the other hand, this coupling must be an integer multiple of a perturbatively small loop factor (see (2.9)). One attempt to realize the necessary large integer coefficient is for it to arise as a product of smaller integers, through the clockwork mechanism [34–36]. We have shown, in section 5, that this mechanism cannot generate a sufficiently large coupling to satisfy all of the phenomenological and consistency requirements on the spectator CNI model: one can derive two bounds on the axion scale f , (5.4) and (5.8), that are mutually inconsistent by two orders of magnitude.

All of the details of the clockwork scenario, in the end, are subsumed in the statement that $\lambda/f \lesssim \mu^{-1}/(2\pi)$. This excludes the parameter space of the model. It is intriguing that bounds on λ/f can also be deduced purely within effective field theory: as we argued in section 6, perturbative unitarity alone implies that $\lambda/f \lesssim 8\sqrt{\pi}\mu^{-1}$. This has the same form as the clockwork constraint, but is numerically weaker, and is insufficient to fully exclude

the parameter space of the model. We have also briefly discussed possible constraints from variations on the Weak Gravity Conjecture. As is often the case in such applications, there are potentially interesting bounds but they require stronger conjectures beyond the minimal and most well-established ones.

In light of our results, there are some further potential avenues to explore. One would be to investigate models with a more general potential $U(\chi)$ rather than a simple cosine potential. Axion monodromy [43, 44] could allow for an effectively non-periodic potential, an idea that has been invoked, but not thoroughly explored, in the CNI context [45–47]. Structurally, these models have a similar challenge in explaining the origin of the large coupling of axions to gauge fields, but the details are different enough that the constraints discussed here should be re-evaluated.

Another direction would be to attempt to find a set of assumptions weaker than a clockwork UV completion, but stronger than perturbative unitarity in the EFT, that is sufficient to rule out the model. For example, one could focus on models in which the $\chi F\tilde{F}$ coupling is generated by integrating out 4d fermions in various representations of the $SU(2)$ gauge group, and study additional higher-dimension operators arising in the EFT and relative constraints on their coefficients. In recent years, various positivity and bootstrap bounds have also been used to strengthen conclusions beyond those available from partial-wave unitarity bounds. These methods could be applied in this context. A quite different possible origin for the $\chi F\tilde{F}$ coupling is from a higher-dimensional Chern-Simons term, with a large level, potentially arising from a flux through still other extra dimensions. Given the phenomenological requirements of the model, the scales H , μ , and f must be sufficiently close to the Planck scale that there is relatively little room for engineering hierarchies of scales involving multiple extra dimensions. Nonetheless, it could be interesting to assess the possibility more quantitatively (see [38] for an exploration along these lines in the context of Type IIB string theory).

Using the clockwork scenario as an example, we have highlighted the challenges involved in finding a UV-complete theory that explains the origin of parameters in the S-CNI model. We expect that similar challenges would arise in any attempt at building a UV completion. However, we do not have a sharp no-go theorem. Given the potential importance of a chiral gravitational wave signal, we hope to see further creative ideas for new models and mechanisms that could convincingly embed such a signal in a consistent quantum theory.

Acknowledgments

HB thanks A. Bedroya and E. Sussman for helpful discussions. HB and MR are supported in part by the NASA Grant 80NSSC20K0506 and the DOE Grant DE-SC0013607. W.L.X. thanks the Mainz Institute of Theoretical Physics of the Cluster of Excellence PRISMA+ (Project ID 39083149) for its hospitality during completion of part of this work, and is supported by the U.S. Department of Energy under Contract DE-AC02-05CH11231.

A Deriving the ϵ_ϕ equation of motion

In this appendix, we present the derivation of the equation (2.19) for the time dependence of ϵ_ϕ without reference to the form of $V(\phi)$. First, we present some definitions. We will use the differential number of e-folds, $dN = -Hdt$, and the slow-roll parameters, both defined in terms of $H(t)$ (with subscript H) and in terms of $V(\phi)$ (with subscript ϕ) (see, e.g., chapter 18 of [67]):

$$\epsilon_H \equiv -\frac{\dot{H}}{H^2} = \frac{d \ln H}{dN}, \quad (\text{A.1})$$

$$\eta_H \equiv \epsilon_H - \frac{1}{2} \frac{\dot{\epsilon}_H}{H\epsilon_H} = \epsilon_H + \frac{1}{2} \frac{d \ln \epsilon_H}{dN} \quad (\text{A.2})$$

$$\epsilon_\phi \equiv \frac{\dot{\phi}^2}{2H^2 M_{\text{Pl}}^2} \approx \frac{1}{2} M_{\text{Pl}}^2 \left(\frac{V'(\phi)}{V(\phi)} \right)^2, \quad (\text{A.3})$$

$$\eta_\phi \equiv M_{\text{Pl}}^2 \frac{V''(\phi)}{V(\phi)}. \quad (\text{A.4})$$

In the main text, we defined ϵ_ϕ in (2.12) (as above) in terms of $\dot{\phi}$, because this is exactly what appears in the classical equation of motion (2.17). The conventional definition, however, is the final term in (A.3) in terms of V'/V . These are not equivalent, but they are approximately equal, assuming that ϕ satisfies the slow-roll condition (2.13) and that $V(\phi)$ dominates the energy density of the universe, so that $V(\phi) \approx 3H^2 M_{\text{Pl}}^2$.

We would like to find an equation for the time derivative $\dot{\epsilon}_\phi$ that does not involve the exact form of the potential $V(\phi)$. We will do this in two steps: first, we will find an expression for $\dot{\epsilon}_\phi$ that depends on V only implicitly through η_ϕ . Then, we will relate the spectral index $n_s - 1$ to η_ϕ , allowing us to finally obtain (2.19).

Taking a time derivative of (A.3) and using the time derivative of the slow-roll approximation $3H\dot{\phi} = -V'(\phi)$, we find that

$$\dot{\epsilon}_\phi \approx 2H\epsilon_\phi(2\epsilon_H - \eta_\phi). \quad (\text{A.5})$$

This involves $V(\phi)$ only through the slow-roll parameter η_ϕ . Our next goal is to re-express this in terms of n_s .

Consider the spectral tilt

$$n_s - 1 = \frac{d \ln \mathcal{P}_\zeta}{d \ln k} = \frac{d \ln \mathcal{P}_\zeta}{dN} \cdot \frac{dN}{d \ln k}. \quad (\text{A.6})$$

Using the equation $\mathcal{P}_\zeta = \frac{H^2}{8\pi^2 M_{\text{Pl}}^2 \epsilon_H^2} \epsilon_\phi$ from [30], we have

$$\frac{d \ln \mathcal{P}_\zeta}{dN} = 2 \frac{d \ln H}{dN} + \frac{d \ln \epsilon_\phi}{dN} - 2 \frac{d \ln \epsilon_H}{dN}. \quad (\text{A.7})$$

From the horizon crossing relation $k = aH$, we have

$$d \ln k = -dN + d \ln H \quad \Rightarrow \quad \frac{dN}{d \ln k} = \left(-1 + \frac{d \ln H}{dN} \right)^{-1} = -(1 - \epsilon_H)^{-1}. \quad (\text{A.8})$$

From the definition (A.2), we have

$$\frac{d \ln \epsilon_H}{dN} = 2(\eta_H - \epsilon_H). \quad (\text{A.9})$$

Finally, we can rewrite (A.5) as

$$\frac{d \ln \epsilon_\phi}{dN} = -2(2\epsilon_H - \eta_\phi). \quad (\text{A.10})$$

Plugging in (A.1), (A.10), and (A.9) into (A.7), and the result together with (A.8) into (A.6), we find

$$n_s - 1 \approx (4\eta_H - 2\eta_\phi - 2\epsilon_H) \cdot (1 - \epsilon_H)^{-1} \approx 4\eta_H - 2\eta_\phi - 2\epsilon_H, \quad (\text{A.11})$$

up to higher order terms in the slow-roll expansion and corrections due to subdominant contributions to the energy density of the universe.

Finally, we substitute (A.11) into (A.5) to obtain

$$\dot{\epsilon}_\phi \approx H\epsilon_\phi(6\epsilon_H - 4\eta_H + n_s - 1), \quad (\text{A.12})$$

which we quoted in the main text as (2.19), substituting explicit expressions for ϵ_H and η_H in terms of $H(t)$.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited. SCOAP³ supports the goals of the International Year of Basic Sciences for Sustainable Development.

References

- [1] D.H. Lyth, *What would we learn by detecting a gravitational wave signal in the cosmic microwave background anisotropy?*, *Phys. Rev. Lett.* **78** (1997) 1861 [[hep-ph/9606387](#)] [[INSPIRE](#)].
- [2] D. Baumann and D. Green, *A field range bound for general single-field inflation*, *JCAP* **05** (2012) 017 [[arXiv:1111.3040](#)] [[INSPIRE](#)].
- [3] M. Mirbabayi, L. Senatore, E. Silverstein and M. Zaldarriaga, *Gravitational waves and the scale of inflation*, *Phys. Rev. D* **91** (2015) 063518 [[arXiv:1412.0665](#)] [[INSPIRE](#)].
- [4] A. Maleknejad and M.M. Sheikh-Jabbari, *Gauge-fflation: inflation from non-Abelian gauge fields*, *Phys. Lett. B* **723** (2013) 224 [[arXiv:1102.1513](#)] [[INSPIRE](#)].
- [5] A. Maleknejad and M.M. Sheikh-Jabbari, *Non-Abelian gauge field inflation*, *Phys. Rev. D* **84** (2011) 043515 [[arXiv:1102.1932](#)] [[INSPIRE](#)].
- [6] P. Adshead and M. Wyman, *Chromo-natural inflation: natural inflation on a steep potential with classical non-Abelian gauge fields*, *Phys. Rev. Lett.* **108** (2012) 261302 [[arXiv:1202.2366](#)] [[INSPIRE](#)].
- [7] P. Adshead and M. Wyman, *Gauge-fflation trajectories in chromo-natural inflation*, *Phys. Rev. D* **86** (2012) 043530 [[arXiv:1203.2264](#)] [[INSPIRE](#)].

- [8] M.M. Sheikh-Jabbari, *Gauge-flaton vs chromo-natural inflation*, *Phys. Lett. B* **717** (2012) 6 [[arXiv:1203.2265](#)] [[INSPIRE](#)].
- [9] A. Maleknejad and M. Zarei, *Slow-roll trajectories in chromo-natural and gauge-flaton models, an exhaustive analysis*, *Phys. Rev. D* **88** (2013) 043509 [[arXiv:1212.6760](#)] [[INSPIRE](#)].
- [10] A. Maleknejad, M.M. Sheikh-Jabbari and J. Soda, *Gauge fields and inflation*, *Phys. Rept.* **528** (2013) 161 [[arXiv:1212.2921](#)] [[INSPIRE](#)].
- [11] M.S. Turner and L.M. Widrow, *Inflation produced, large scale magnetic fields*, *Phys. Rev. D* **37** (1988) 2743 [[INSPIRE](#)].
- [12] W.D. Garretson, G.B. Field and S.M. Carroll, *Primordial magnetic fields from pseudo-Goldstone bosons*, *Phys. Rev. D* **46** (1992) 5346 [[hep-ph/9209238](#)] [[INSPIRE](#)].
- [13] P. Adshead, E. Martinec and M. Wyman, *Gauge fields and inflation: chiral gravitational waves, fluctuations, and the Lyth bound*, *Phys. Rev. D* **88** (2013) 021302 [[arXiv:1301.2598](#)] [[INSPIRE](#)].
- [14] P. Adshead, E. Martinec and M. Wyman, *Perturbations in chromo-natural inflation*, *JHEP* **09** (2013) 087 [[arXiv:1305.2930](#)] [[INSPIRE](#)].
- [15] A. Maleknejad, M. Noorbala and M.M. Sheikh-Jabbari, *Leptogenesis in inflationary models with non-Abelian gauge fields*, *Gen. Rel. Grav.* **50** (2018) 110 [[arXiv:1208.2807](#)] [[INSPIRE](#)].
- [16] A. Maleknejad, *Chiral gravity waves and leptogenesis in inflationary models with non-Abelian gauge fields*, *Phys. Rev. D* **90** (2014) 023542 [[arXiv:1401.7628](#)] [[INSPIRE](#)].
- [17] I. Obata and J. Soda, *Chiral primordial chiral primordial gravitational waves from dilaton induced delayed chromonatural inflation*, *Phys. Rev. D* **93** (2016) 123502 [*Addendum ibid.* **95** (2017) 109903] [[arXiv:1602.06024](#)] [[INSPIRE](#)].
- [18] A. Maleknejad, *Axion inflation with an SU(2) gauge field: detectable chiral gravity waves*, *JHEP* **07** (2016) 104 [[arXiv:1604.03327](#)] [[INSPIRE](#)].
- [19] I. Obata and J. Soda, *Oscillating chiral tensor spectrum from axionic inflation*, *Phys. Rev. D* **94** (2016) 044062 [[arXiv:1607.01847](#)] [[INSPIRE](#)].
- [20] E. Dimastrogiovanni, M. Fasiello and T. Fujita, *Primordial gravitational waves from axion-gauge fields dynamics*, *JCAP* **01** (2017) 019 [[arXiv:1608.04216](#)] [[INSPIRE](#)].
- [21] E. Martinec, P. Adshead and M. Wyman, *Chern-Simons EM-flaton*, *JHEP* **02** (2013) 027 [[arXiv:1206.2889](#)] [[INSPIRE](#)].
- [22] A. Maleknejad and E. Erfani, *Chromo-natural model in anisotropic background*, *JCAP* **03** (2014) 016 [[arXiv:1311.3361](#)] [[INSPIRE](#)].
- [23] E. Dimastrogiovanni and M. Peloso, *Stability analysis of chromo-natural inflation and possible evasion of Lyth's bound*, *Phys. Rev. D* **87** (2013) 103501 [[arXiv:1212.5184](#)] [[INSPIRE](#)].
- [24] P. Adshead, E. Martinec, E.I. Sfakianakis and M. Wyman, *Higgsed chromo-natural inflation*, *JHEP* **12** (2016) 137 [[arXiv:1609.04025](#)] [[INSPIRE](#)].
- [25] T. Fujita, R. Namba and Y. Tada, *Does the detection of primordial gravitational waves exclude low energy inflation?*, *Phys. Lett. B* **778** (2018) 17 [[arXiv:1705.01533](#)] [[INSPIRE](#)].
- [26] I. Wolfson, A. Maleknejad and E. Komatsu, *How attractive is the isotropic attractor solution of axion-SU(2) inflation?*, *JCAP* **09** (2020) 047 [[arXiv:2003.01617](#)] [[INSPIRE](#)].

- [27] I. Wolfson, A. Maleknejad, T. Murata, E. Komatsu and T. Kobayashi, *The isotropic attractor solution of axion-SU(2) inflation: universal isotropization in Bianchi type-I geometry*, *JCAP* **09** (2021) 031 [[arXiv:2105.06259](#)] [[INSPIRE](#)].
- [28] A. Maleknejad and E. Komatsu, *Production and backreaction of spin-2 particles of SU(2) gauge field during inflation*, *JHEP* **05** (2019) 174 [[arXiv:1808.09076](#)] [[INSPIRE](#)].
- [29] A. Papageorgiou, M. Peloso and C. Unal, *Nonlinear perturbations from the coupling of the inflaton to a non-Abelian gauge field, with a focus on chromo-natural inflation*, *JCAP* **09** (2018) 030 [[arXiv:1806.08313](#)] [[INSPIRE](#)].
- [30] A. Papageorgiou, M. Peloso and C. Unal, *Nonlinear perturbations from axion-gauge fields dynamics during inflation*, *JCAP* **07** (2019) 004 [[arXiv:1904.01488](#)] [[INSPIRE](#)].
- [31] M.M. Anber and L. Sorbo, *Naturally inflating on steep potentials through electromagnetic dissipation*, *Phys. Rev. D* **81** (2010) 043534 [[arXiv:0908.4089](#)] [[INSPIRE](#)].
- [32] B. Heidenreich, M. Reece and T. Rudelius, *The weak gravity conjecture and emergence from an ultraviolet cutoff*, *Eur. Phys. J. C* **78** (2018) 337 [[arXiv:1712.01868](#)] [[INSPIRE](#)].
- [33] P. Agrawal, J. Fan and M. Reece, *Clockwork axions in cosmology: is chromonatural inflation chrononatural?*, *JHEP* **10** (2018) 193 [[arXiv:1806.09621](#)] [[INSPIRE](#)].
- [34] K. Choi, H. Kim and S. Yun, *Natural inflation with multiple sub-Planckian axions*, *Phys. Rev. D* **90** (2014) 023545 [[arXiv:1404.6209](#)] [[INSPIRE](#)].
- [35] K. Choi and S.H. Im, *Realizing the relaxation from multiple axions and its UV completion with high scale supersymmetry*, *JHEP* **01** (2016) 149 [[arXiv:1511.00132](#)] [[INSPIRE](#)].
- [36] D.E. Kaplan and R. Rattazzi, *Large field excursions and approximate discrete symmetries from a clockwork axion*, *Phys. Rev. D* **93** (2016) 085007 [[arXiv:1511.01827](#)] [[INSPIRE](#)].
- [37] S.-L. Cheng, W. Lee and K.-W. Ng, *Primordial black holes and associated gravitational waves in axion monodromy inflation*, *JCAP* **07** (2018) 001 [[arXiv:1801.09050](#)] [[INSPIRE](#)].
- [38] J. Holland, I. Zavala and G. Tasinato, *On chromonatural inflation in string theory*, *JCAP* **12** (2020) 026 [[arXiv:2009.00653](#)] [[INSPIRE](#)].
- [39] N. Barnaby, J. Moxon, R. Namba, M. Peloso, G. Shiu and P. Zhou, *Gravity waves and non-Gaussian features from particle production in a sector gravitationally coupled to the inflaton*, *Phys. Rev. D* **86** (2012) 103508 [[arXiv:1206.6117](#)] [[INSPIRE](#)].
- [40] R. Namba, M. Peloso, M. Shiraishi, L. Sorbo and C. Unal, *Scale-dependent gravitational waves from a rolling axion*, *JCAP* **01** (2016) 041 [[arXiv:1509.07521](#)] [[INSPIRE](#)].
- [41] E. Witten, *Large N chiral dynamics*, *Annals of Physics* **128** (1980) 363.
- [42] J.E. Kim, H.P. Nilles and M. Peloso, *Completing natural inflation*, *JCAP* **01** (2005) 005 [[hep-ph/0409138](#)] [[INSPIRE](#)].
- [43] E. Silverstein and A. Westphal, *Monodromy in the CMB: gravity waves and string inflation*, *Phys. Rev. D* **78** (2008) 106003 [[arXiv:0803.3085](#)] [[INSPIRE](#)].
- [44] N. Kaloper and L. Sorbo, *A natural framework for chaotic inflation*, *Phys. Rev. Lett.* **102** (2009) 121301 [[arXiv:0811.1989](#)] [[INSPIRE](#)].
- [45] A. Maleknejad, *Gravitational leptogenesis in axion inflation with SU(2) gauge field*, *JCAP* **12** (2016) 027 [[arXiv:1604.06520](#)] [[INSPIRE](#)].

- [46] A. Maleknejad, $SU(2)_R$ and its axion in cosmology: a common origin for inflation, cold sterile neutrinos, and baryogenesis, *Phys. Rev. D* **104** (2021) 083518 [[arXiv:2012.11516](#)] [[INSPIRE](#)].
- [47] A. Maleknejad, Chiral anomaly in $SU(2)_R$ -axion inflation and the new prediction for particle cosmology, *JHEP* **21** (2020) 113 [[arXiv:2103.14611](#)] [[INSPIRE](#)].
- [48] PLANCK collaboration, Planck 2018 results. VI. Cosmological parameters, *Astron. Astrophys.* **641** (2020) A6 [Erratum *ibid.* **652** (2021) C4] [[arXiv:1807.06209](#)] [[INSPIRE](#)].
- [49] K. Freese, J.A. Frieman and A.V. Olinto, Natural inflation with pseudo Nambu-Goldstone bosons, *Phys. Rev. Lett.* **65** (1990) 3233.
- [50] BICEP and KECK collaborations, Improved constraints on primordial gravitational waves using Planck, WMAP, and BICEP/Keck observations through the 2018 observing season, *Phys. Rev. Lett.* **127** (2021) 151301 [[arXiv:2110.00483](#)] [[INSPIRE](#)].
- [51] B. Thorne, T. Fujita, M. Hazumi, N. Katayama, E. Komatsu and M. Shiraishi, Finding the chiral gravitational wave background of an axion- $SU(2)$ inflationary model using CMB observations and laser interferometers, *Phys. Rev. D* **97** (2018) 043506 [[arXiv:1707.03240](#)] [[INSPIRE](#)].
- [52] I. Brivio, O.J.P. Éboli and M.C. Gonzalez-Garcia, Unitarity constraints on ALP interactions, *Phys. Rev. D* **104** (2021) 035027 [[arXiv:2106.05977](#)] [[INSPIRE](#)].
- [53] S.C. İnan and A.V. Kisselev, Probe of axion-like particles in vector boson scattering at a muon collider, [arXiv:2207.03325](#) [[INSPIRE](#)].
- [54] C. Itzykson and J.B. Zuber, *Quantum field theory*, McGraw-Hill, New York, NY, U.S.A. (1980).
- [55] M. Kruczenski, J. Penedones and B.C. van Rees, Snowmass white paper: S-matrix bootstrap, [arXiv:2203.02421](#) [[INSPIRE](#)].
- [56] C. de Rham, S. Kundu, M. Reece, A.J. Tolley and S.-Y. Zhou, Snowmass white paper: UV constraints on IR physics, in *2022 Snowmass summer study*, (2022) [[arXiv:2203.06805](#)] [[INSPIRE](#)].
- [57] T. Banks and N. Seiberg, Symmetries and strings in field theory and gravity, *Phys. Rev. D* **83** (2011) 084019 [[arXiv:1011.5120](#)] [[INSPIRE](#)].
- [58] D. Harlow and H. Ooguri, Symmetries in quantum field theory and quantum gravity, *Commun. Math. Phys.* **383** (2021) 1669 [[arXiv:1810.05338](#)] [[INSPIRE](#)].
- [59] N. Arkani-Hamed, L. Motl, A. Nicolis and C. Vafa, The string landscape, black holes and gravity as the weakest force, *JHEP* **06** (2007) 060 [[hep-th/0601001](#)] [[INSPIRE](#)].
- [60] B. Heidenreich, M. Reece and T. Rudelius, Sharpening the weak gravity conjecture with dimensional reduction, *JHEP* **02** (2016) 140 [[arXiv:1509.06374](#)] [[INSPIRE](#)].
- [61] C.F. Cota, A. Klemm and T. Schimannek, State counting on fibered CY-3 folds and the non-Abelian weak gravity conjecture, *JHEP* **05** (2021) 030 [[arXiv:2012.09836](#)] [[INSPIRE](#)].
- [62] D. Harlow, B. Heidenreich, M. Reece and T. Rudelius, The weak gravity conjecture: a review, [arXiv:2201.08380](#) [[INSPIRE](#)].
- [63] B. Heidenreich, J. McNamara, M. Montero, M. Reece, T. Rudelius and I. Valenzuela, Chern-Weil global symmetries and how quantum gravity avoids them, *JHEP* **11** (2021) 053 [[arXiv:2012.00009](#)] [[INSPIRE](#)].

- [64] B. Heidenreich, M. Reece and T. Rudelius, *The weak gravity conjecture and axion strings*, *JHEP* **11** (2021) 004 [[arXiv:2108.11383](#)] [[INSPIRE](#)].
- [65] M. Kakizaki, M. Ogata and O. Seto, *Dark radiation in spectator axion-gauge models*, *PTEP* **2022** (2022) 033E02 [[arXiv:2110.12936](#)] [[INSPIRE](#)].
- [66] M. Montero, J.B. Muñoz and G. Obied, *Swampland bounds on dark sectors*, *JHEP* **11** (2022) 121 [[arXiv:2207.09448](#)] [[INSPIRE](#)].
- [67] D.H. Lyth and A.R. Liddle, *The primordial density perturbation: cosmology, inflation and the origin of structure*, Cambridge University Press (2009).