Reduced axion abundance from an extended symmetry

Itamar J. Allali,[1,*] Mark P. Hertzberg,[1,2,3,4,†] and Yi Lyu[5,‡]

[1]*Institute of Cosmology, Department of Physics and Astronomy, Tufts University, Medford, Massachusetts 02155, USA*
[2]*Center for Astrophysics, Harvard-Smithsonian, Cambridge, Massachusetts 02138, USA*
[3]*Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*
[4]*Theoretical Physics Center, Department of Physics, Brown University, Providence, Rhode Island 02912, USA*
[5]*University of California Santa Cruz, Santa Cruz, California 95064, USA*

(Received 27 November 2022; accepted 14 February 2023; published 24 February 2023)

In recent work we showed that the relic dark matter abundance of QCD axions can be altered when the Peccei-Quinn (PQ) field is coupled to very light scalar/s, rendering the effective axion mass dynamical in the early universe. In this work we develop this framework further, by introducing a new extended symmetry group to protect the new particles' mass. We find that with a new global $SO(N)$ symmetry, with large N, we can not only account for the lightness of the new scalars, but we can reduce the axion relic abundance in a technically natural way. This opens up the possibility of large PQ scales, including approaching the grand unified theory (GUT) scale, and still naturally producing the correct relic abundance of axions. Also, in these models the effective PQ scale is relatively small in the very early universe, and so this can help toward alleviating the isocurvature problem from inflation. Furthermore, instead of possible overclosure from cosmic strings, the extended symmetry implies the formation of nontopological textures which provide a relatively small abundance.

DOI: 10.1103/PhysRevD.107.043535

I. INTRODUCTION

The QCD axion continues to be one of the most well-motivated particles beyond the Standard Model (SM). It first emerged as a dynamical solution to the strong CP problem of QCD [1], explaining the smallness of the otherwise allowed $\theta_0 G\tilde{G}$ term in the SM action. This occurs by postulating a new singlet scalar (axion θ) with an approximate shift symmetry. Such a scalar is allowed to couple to the SM by the dimension 5 operator $\theta G\tilde{G}$. By carefully studying the ensuing dynamics one finds the effective θ is driven toward zero. Because such a scalar is electrically neutral and long-lived [2,3], it has been identified as a viable dark matter (DM) candidate, possibly explaining the missing ∼85% of the mass of the universe.

In its basic UV completion, this scalar, axion, arises through the spontaneous breakdown of a new postulated global Peccei-Quinn (PQ) symmetry [1] at some high scale f_a. The axion acquires a small potential and mass from QCD instantons. For generic initial conditions, the axion will be displaced away from the minimum of its potential in the early universe, then later rolling, oscillating, and redshifting toward zero at late times. This dynamical relaxation toward zero resolves the strong CP problem, while the oscillations themselves behave as nonrelativistic matter in the late universe [4–6]. In standard cosmologies, the abundance of axions at late times is determined by the symmetry breaking scale, f_a as

$$\Omega_a \sim \left(\frac{f_a}{10^{12}\ \mathrm{GeV}}\right)^{7/6}\langle\theta_i^2\rangle \tag{1}$$

with $\langle\theta_i^2\rangle$ denoting the spatial average of the square of the initial displacement angle. For the axion to account for the observed DM ($\Omega_a \approx 0.25$), one needs f_a around 2×10^{11} GeV (unless $\langle\theta_i^2\rangle$ is fine-tuned to a small value), giving a mass

$$m_a = \frac{\Lambda_0^2}{f_a} \approx 28\ \mu\mathrm{eV}\left(\frac{2\times 10^{11}\ \mathrm{GeV}}{f_a}\right) \tag{2}$$

(where $\Lambda_0 \approx 90$ MeV is related to the QCD scale and quark masses). In the so-called "standard axion window," the

*itamar.allali@tufts.edu
†mark.hertzberg@tufts.edu
‡ylyu11@ucsc.edu

upper bound on the breaking scale is $f_a \lesssim 10^{12}$ GeV. This follows from taking $\langle\theta_i^2\rangle \sim 1$ and then toward the end of this window, the axion will make up all (or most) of the dark matter of the universe, and beyond this upper bound the axion would be too abundant and would overclose the universe. The lower bound of the standard window is $f_a \gtrsim 10^9$ GeV, from demanding that the axion-photon coupling and axion-nucleon couplings in standard constructions do not produce large effects in stellar physics (e.g., stellar cooling, supernovae cooling, etc.).

There exist theoretical motivations, however, to prefer values for the PQ symmetry-breaking scale much higher than this standard axion window. It may be more plausible for the physics associated with PQ symmetry breaking to occur at a higher scale suggested by fundamental physics, for instance in grand unified theories (GUT) or string theory [7]. However, this seems to significantly overproduce the dark matter. There are ways around it by invoking unnaturally small initial misalignment angles θ_i during inflation; although such models tend to in turn overproduce isocurvature modes in the CMB (see, e.g., [8–13]).

In our previous work [14], we discussed a new framework in which to expand the window of viable f_a; both to higher and lower values. We proposed a dynamical PQ scale mechanism, wherein the PQ field Φ is coupled to an additional scalar field χ, which allows the effective PQ scale to evolve with time. This evolution alters the predicted abundance of axions and consequently widens the axion window. However, the lightness of the χ field was left unexplained.

In this work, we develop this framework further, by addressing the issue of the lightness of the χ field. We do so by introducing an extended symmetry group, involving the PQ field Φ and an additional N_χ scalars χ_j. This new enlarged PQ sector has a total of $N_\chi + 2$ scalar degrees of freedom (two from Φ and N_χ from the χ_j), which organize into an approximate $SO(N_\chi + 2)$ symmetry. As is standard, QCD instantons explicitly break the $U(1)$ subgroup involving the complex Φ, leaving behind a $SO(N_\chi)$ symmetry. With typical initial conditions, this model has only one free parameter, the number of additional fields N_χ, which determines to which degree the abundance of the axion is suppressed as compared to the standard QCD axion. Finally, there is also the introduction of a mass of χ, which is allowed to be small as it represents a small breaking of the original $SO(N_\chi + 2)$ symmetry; so its lightness is technically natural. By making the mass nonzero, the χ fields eventually relax to the bottom of their potential, and if their mass is small, their abundance is small too. As we will show, for a sufficiently large symmetry group, this scenario can accommodate $f_a \gg 10^{12}$ GeV while keeping the axion abundance below the upper bound $\Omega_a \lesssim 0.25$ to avoid overclosing the universe. For some other interesting mechanisms that alter the axion's dynamics and abundance, see, e.g., [15–27].

For such high f_a, one is normally concerned about the symmetry being broken before the end of inflation, which could generate appreciable isocurvature modes. But in the presence of many fields, the symmetry can be restored, avoiding this problem.

The outline of our paper is as follows: In Sec. II we briefly review the standard QCD axion and our previous dynamical PQ model. In Sec. III we introduce the model with an extended symmetry group. In Sec. IV we discuss the basics of the cosmic evolution. In Sec. V we perform a numerical analysis of the equations of motion in the homogenous approximation. In Sec. VI we discuss constraints from isocurvature bounds during inflation, defects, unitarity bounds, and the plausibility of this construction. Finally, in Appendix A we give some more details of the effective action, in Appendix B we describe the eigenmodes from inhomogeneities, and in Appendix C we study possible resonance into such inhomogeneities.

II. STATIC AND DYNAMIC PECCEI-QUINN MECHANISMS

A. Static recap

The starting point for the canonical Peccei-Quinn (PQ) mechanism involves a complex PQ scalar field $\Phi = \rho e^{i\theta}$, which enjoys a global $U(1)_{\rm PQ}$ symmetry. The axion is the angular degree of freedom θ which becomes a (pseudo-) Goldstone boson when the symmetry is spontaneously broken. The effective Lagrangian density for this axion is given by

$$\mathcal{L} = \sqrt{-g}\left[\frac{1}{2}|\partial\Phi|^2 - \frac{\lambda}{4}(|\Phi|^2 - f_a^2)^2 - V(\theta, T)\right]. \quad (3)$$

Interactions with the Standard Model give the axion a potential, $V(\theta, T) \approx \Lambda(T)^4(1 - \cos\theta)$ (in fact there are $\mathcal{O}(1)$ corrections to this shape, but they will not be important here). Consequently the axion acquires a mass, which depends on temperature. Its low temperature value m_a is related to $\Lambda_0^4 = f_a^2 m_a^2$, where f_a, the PQ scale, is the vacuum expectation value (VEV) of the radial PQ field ρ after symmetry breaking. The choice of the scale f_a for the symmetry breaking dictates the quantity of axions produced by the misalignment mechanism in the early universe according to Eq. (1).

Rather than discussing Ω_a, which is time dependent, it is convenient to use a more fundamental abundance parameter [13]

$$\xi_a(T) \equiv \frac{\rho_a(T)}{n_\gamma(T)} \quad (4)$$

with T the temperature of the universe. At late times, this tends to a constant as both the numerator and denominator redshift together as $1/a^3$. The observed DM abundance ($\Omega_a \approx 0.25$) at late times is the value

$$\xi_{\rm obs} \approx 2.9 \text{ eV}. \tag{5}$$

For further details on the standard axion setup and evolution, see Ref. [14] and others (e.g., [26,28] and references therein).

B. Dynamic recap

In our recent work [14], we proposed a mechanism to alter the abundance of axions produced by the misalignment mechanism. We introduced a new scalar degree of freedom χ which couples to the axion in a way that makes the PQ scale f_a effectively dynamical. This can result in viable abundance predictions for axions outside of the standard allowed window for f_a. For instance, in the unaltered misalignment mechanism, one can place an upper bound on $f_a \lesssim 10^{12}$ GeV to avoid over-closing the universe with too large a density of axions. However, with the increasing-PQ-scale model discussed in [14], we can accommodate $f_a \sim 10^{16}$ GeV, motivated by physics at the GUT scale.

The action of Eq. (3) was modified to be

$$\mathcal{L} = \sqrt{-g}\left[\frac{1}{2}|\partial\Phi|^2 - \frac{\lambda}{4}(|\Phi|^2 - f(\chi)^2)^2 - V(\theta, T) + \frac{1}{2}(\partial\chi)^2 - \frac{1}{2}m_\chi^2\chi^2\right] \tag{6}$$

where the function $f(\chi)$ determines how the new field χ couples and thus the time-dependent behavior of the now dynamical PQ scale. As in Eq. (1), the energy density of the axions $\rho_a \propto f_a^{7/6}$, and it is also proportional to the axion mass. Since the mass is inversely proportional to f_a, as the effective f_a increases with time, the effective m_a decreases, and the abundance is suppressed by an overall factor of $1 + 7/6$ powers of the effective initial $f_i = f(\chi_i)$

$$\frac{\xi_a}{\xi_{a,\rm std}} = \left(\frac{f_i}{f_a}\right)^{13/6} \tag{7}$$

where $\xi_{a,\rm std}$ is the abundance of the standard QCD axion and ξ_a is that of the axion in the dynamical PQ scale model.

III. EXTENDED SYMMETRY

Although our work [14] had nice phenomenological success, it left the question of the lightness of χ unexplained. In particular, in the above action there is no symmetry protecting the χ from being heavy, and so its lightness appears tuned.

A. New class of models

In this work, we wish to stay within this overarching framework, but provide a concrete example in which the mass of χ is protected by a new global symmetry. It will turn out that in order to alter the axion abundance appreciably, we will need many new scalar fields. Correspondingly, we will need to appeal to an extended symmetry group involving all $N_\chi + 2$ degrees of freedom in the new enlarged axion sector (N_χ new fields and 2 components of the complex PQ field).

We will consider the following updated action, involving N_χ new scalars χ_j

$$\mathcal{L} = \sqrt{-g}\left[\frac{1}{2}(|\partial\Phi|^2 + \sum_{j=1}^{N_\chi}(\partial\chi_j)^2) - \frac{\lambda}{4}\left(|\Phi|^2 + \sum_{j=1}^{N_\chi}\chi_j^2 - f_a^2\right)^2 - V(\theta,\chi_j,T)\right]. \tag{8}$$

Apart from the potential V, which will be self-consistently taken to be very small, this action is invariant under an $SO(N_\chi + 2)$ transformation between the two degrees of freedom of $\Phi = \rho e^{i\theta}$ (real and imaginary) and the N_χ new scalars.

In the very early universe, all of the scalars participate in the $SO(N_\chi + 2)$ symmetry. At low energies, the potential term in eq. (8) proportional to $\lambda/4$ spontaneously breaks the symmetry to a residual $SO(N_\chi + 1)$ group, leaving $N_\chi + 1$ Goldstone bosons. This results in the field ρ (the radial mode of Φ) acquiring a vacuum expectation value (VEV) $\langle\rho\rangle = f_i$, while the remaining scalars θ and χ_j obtain random initial values θ_i and $\chi_{j,i}$ and are prevented from evolving by Hubble friction. The initial value of the "effective" PQ scale

$$f_i^2 = f_a^2 - \sum_j^{N_\chi}\chi_{j,i}^2 < f_a^2 \tag{9}$$

is thus smaller at early times than the vacuum value f_a.

B. Induced masses

As the temperature of the universe begins to approach the QCD phase transition, QCD instantons induce a potential for the axion, explicitly breaking the $U(1)$ subgroup of the above extended symmetry group, leaving a residual unbroken $SO(N_\chi)$. This leaves N_χ Goldstone bosons, which we can identify as the χ particles. With this mass for the axion, it begins to roll down its potential when $3H \sim m_a$, at a temperature T_{osc} as is usual. Its oscillations behave as cold dark matter (CDM).

Also, the residual $SO(N_\chi)$ symmetry can allow for a mass term for χ. We can write this as (note that in Eq. (8), both of these potentials are represented by V)

$$V(\theta,\chi_j,T) \approx \frac{1}{2}m_\chi^2\sum_j^{N_\chi}\chi_j^2 + \Lambda(T)^4(1-\cos\theta) \tag{10}$$

where m_χ is the same mass for each χ_j. The presence of this mass, means that eventually the χ will relax to zero. However, we would like to assume that $m_\chi < m_a(T_{osc})$ such that the axion begins its oscillations first, and the presence of χ alters in a crucial way the axion evolution. (If the χ were very heavy, we could just integrate it out, and it would play no important role for the axion).

Note that since our theory began with an $SO(N_\chi + 2)$ symmetry at leading approximation, which prevents a χ mass, it is technically natural for χ to be light as it represents a small explicit breaking of this extended symmetry. Since the axion's small mass explicitly breaks the extended symmetry, then so too it will generate a small mass for χ. By considering a one-loop diagram provided by the interaction $\Delta\mathcal{L} = -\sum_j \chi_j^2(\partial\theta)^2/2$ (see Appendix A for the relevant action of the low energy theory), one can compute this mass. It can be readily estimated as $m_\chi^2 \sim m_a^2\Lambda_{\rm UV}^2/((4\pi)^2 f_a^2)$, where $\Lambda_{\rm UV}$ is a cutoff on the loop integral; this should be taken to be of the order the mass of the radial PQ mode $\Lambda_{\rm UV} \sim m_{\rm PQ} = \sqrt{2\lambda} f_a$. So the induced mass is $m_\chi^2 \sim \lambda m_a^2/(4\pi)^2$. As discussed in Sec. VI C we already know that we need to take λ somewhat small to maintain perturbative unitarity, so this self-consistently implies $m_\chi \ll m_a$, as we assume in this work.

The above QCD axion potential is taken to be of the standard form $\Delta V = \Lambda(T)^4(1 - \cos\theta)$. However, we note that this should not be used when the magnitude of the PQ field ρ happens to be very small or vanish, since in that regime θ is not well defined. In this paper, we will generally assume that ρ is not particularly small. In fact as we will explain in the next section, the initial condition is naturally on the order $\sqrt{2} f_a/\sqrt{N_\chi + 2}$ and its late time value is f_a. However, one can consider the case in which ρ is accidentally much smaller. Here one should alter the potential ΔV accordingly. A parametrization of an anticipated potential that incorporates both the angular dependence and the radial dependence is of the form $\Delta V = (\rho^2/(M^2 + \rho^2))\Lambda(T)^4(1 - \cos\theta)$. In this parametrization, when $\rho \to 0$ then indeed $\Delta V \to 0$, and the dependence on θ disappears. On the other hand, for large values of $\rho \gg M$ the effective potential becomes nonzero and the generation of the potential for θ from QCD instantons becomes standard. Hence as long as M, the characteristic cross-over scale, is somewhat smaller than f_a (depending on N_χ), then one expects our upcoming primary results to be unaltered.

We also note that the construction in this paper has some overlap with the very interesting Ref. [29]. As is seen in that paper, the final abundance is reduced for many fields; we shall find compatible results here.

An important question is whether this all has an embedding within a UV complete model, with heavy fermions, etc. (requiring a significant extension beyond the minimal models [30–33]); but we leave this for future consideration.

IV. COSMOLOGICAL EVOLUTION

After the spontaneous breaking of the extended symmetry group (occurring well before the QCD phase transition since f_a is very large), the value of ρ is frozen as

$$\rho = \sqrt{f_a^2 - \sum_{j=1}^{N_\chi} \chi_j^2}. \tag{11}$$

Using $|\Phi| = \rho$ and $|\partial\Phi|^2 = (\partial\rho)^2 + \rho^2(\partial\theta)^2$, we insert this into the above action and obtain a kind of nonlinear σ model for the remaining $N_\chi + 1$ light degrees of freedom; see Appendix A for this action.

It is straightforward to vary the action and obtain its classical equations of motion (since these light fields are typically at very high occupancy, the classical field approximation should suffice here). To write down the equations of motion, we work with dimensionless variables defined by

$$\tau \equiv m_a t, \quad Y_j \equiv \frac{\chi_j}{f_a}, \quad L(T) \equiv \frac{\Lambda(T)^4}{\Lambda_0^4}, \quad \mu \equiv \frac{m_\chi}{m_a}. \tag{12}$$

In the equations of motion, we can, to first approximation, ignore spatial variations and focus on the zero modes of the fields. A study of inhomogeneities in the fields is given in Appendices B and C, where we check for possible instabilities or resonance in the system. Also, a discussion of defects is given in Sec. VI B.

It suffices to ignore anharmonicity of the potential and write $\sin\theta \approx \theta$. Also, we write $H = 1/(2t)$ in the radiation era treated in the Friedmann-Robertson-Walker (FRW) approximation. The temperature dependence of the axion potential can be estimated as $L(T) \sim (T_{qcd}/T)^8$ for $T \gg T_{qcd}$ and $L(T) \approx 1$ for $T \ll T_{qcd}$ where $T_{qcd} \sim 100$ MeV is of the order of the temperature of the QCD phase transition. While this temperature dependence has been confirmed by recent lattice studies, the $n = 8$ exponent is only approximate and may take a slightly different value.

In general there are N_χ independent new fields. However, due to the residual $SO(N_\chi)$ symmetry, they all evolve in a similar way, only possibly differing by their initial conditions. For simplicity, here we mention the case in which their initial conditions are all equal, giving a set of identical equations of motion $Y \equiv Y_1 = Y_2 = \ldots Y_{N_\chi}$ as

$$\theta_{\tau\tau} + \left(\frac{3}{2\tau} - \frac{2N_\chi Y Y_\tau}{1 - N_\chi Y^2}\right)\theta_\tau + \frac{L(T)}{1 - N_\chi Y^2}\theta = 0 \tag{13}$$

$$Y_{\tau\tau} + \left(\frac{3}{2\tau} + \frac{N_\chi Y Y_\tau}{1 - N_\chi Y^2}\right) Y_\tau + (1 - N_\chi Y^2)(\mu^2 + \theta_\tau^2) Y = 0 \tag{14}$$

where each subscript τ corresponds to a derivative in τ (i.e., $\theta_\tau \equiv d\theta/d\tau, Y_\tau \equiv dY/d\tau$, etc.). At early times, the scalars are Hubble friction dominated, so we pick initial conditions for their velocities to be $\theta_{\tau,i} = Y_{\tau,i} = 0$, which is consistent with the underlying $SO(N_\chi)$ symmetry and the discrete $Y_j \to -Y_j$ symmetry.

The extension to random initials conditions is analytically simple, though numerically much more complicated. So we shall use this special case, which we believe suffices to illustrate the essential behavior. Nevertheless, it would be useful to extend our analysis to more generic initial conditions.

A. Initial conditions

For the canonical axion, involving only the complex field Φ, the initial condition for θ is typically taken to be $\mathcal{O}(1)$. In the scenario when PQ symmetry is broken before the end of inflation, this can be understood as a random typical angle between $[-\pi, \pi]$. In the case of PQ symmetry breaking after the end of inflation, our observable universe is a huge collection of regions acquiring different θ values at symmetry breaking; this necessarily means that the average of $\langle \theta_i^2 \rangle = \pi^2/3$ is $\mathcal{O}(1)$.

With the new extended symmetry group, one anticipates that the symmetry is broken in a random way; the $N_\chi + 2$ degrees of freedom each have an equal chance of taking on some value, but the sum of their squares is constrained by the symmetry breaking potential. Let us briefly recast the complex PQ field as $\Phi = \phi_1 + i\phi_2$, such that the symmetry breaking results in the constraint

$$\phi_1^2 + \phi_2^2 + \sum_j^{N_\chi} \chi_j^2 = f_a^2. \tag{15}$$

The symmetry ensures that each of the $N_\chi + 2$ scalars takes on a random initial value with equal probability distribution. The unconditional probability distribution for each individual random variable ψ (where ψ can be $\phi_{1,i}$ or $\phi_{2,i}$ or $\chi_{j,i}$) can be shown to be

$$p(\psi) = \frac{1}{\mathcal{M}} (f_a^2 - \psi^2)^{(N_\chi - 1)/2} \tag{16}$$

where $\mathcal{M}$ is a normalization factor; whose value is $\mathcal{M} = \sqrt{\pi}\Gamma((1 + N_\chi)/2) f_a^{N_\chi}/\Gamma(1 + N_\chi/2)$. For large N_χ this becomes a Gaussian distribution with vanishing mean. Its variance is [which can also be inferred from Eq. (15)]

$$\langle \psi^2 \rangle = \langle \phi_{1,i}^2 \rangle = \langle \phi_{2,i}^2 \rangle = \langle \chi_{j,i}^2 \rangle = \frac{f_a^2}{N_\chi + 2}. \tag{17}$$

We choose the initial conditions for each field to be exactly the root mean square for simplicity. This is equivalent to fixing $\rho_i^2 = \phi_{1,i}^2 + \phi_{2,i}^2 = 2f_a^2/(N_\chi + 2)$, and taking $\theta_i = \pi/4$ and $Y_{j,i} = 1/\sqrt{N_\chi + 2}$. Then, as the χ_j evolve once Hubble drops below their mass, the effective VEV of ρ $(= \sqrt{\phi_1^2 + \phi_2^2})$ shifts according to Eq. (15).

One may have noticed that the equations of motion (13) and (14) may be rendered independent of N_χ under the change of variables $\tilde{Y} \equiv \sqrt{N_\chi} Y$. Then, for $N_\chi \gg 1$, the initial conditions become $\tilde{Y}_{j,i} = \sqrt{N_\chi}/\sqrt{N_\chi + 2} \approx 1$ and the parameter N_χ naively appears to drop out of the model altogether. However, to properly analyze the large N_χ regime, one should Taylor expand $\tilde{Y}_{j,i} = (1 + 2/N_\chi)^{-1/2} \approx (1 - 1/N_\chi)$, meaning terms involving $(1 - N_\chi Y^2)$ are initially $(1 - \tilde{Y}_{j,i}^2) \approx 2/N_\chi$, which is clearly sensitive to N_χ.

The key phases of the evolution are as follows: first, the effective friction terms proportional to θ_τ in Eq. (13) and Y_τ in Eq. (14), respectively, become canonical (i.e., $3/2\tau = 3H$) at early times. Next, the effective mass term for the axion is approximately $N_\chi L(\tau)/2$ which becomes very large with large N_χ; this implies an earlier onset of oscillations and suppressed abundance of axions. Further, the effective mass of Y becomes proportional to $2/N_\chi$, and thus shrinks with growing N_χ, delaying the onset of oscillations for the Y field with larger N_χ. This predicted behavior from examining the equations of motion is verified in the numerical solutions presented in Figs. 2–5.

B. Analytical estimates for relic abundance

These initial conditions result in the initial effective PQ scale from Eq. (9) to be

$$f_i = f_a \left(\frac{2}{N_\chi + 2}\right)^{1/2} \tag{18}$$

Making the same estimate as in Eq. (7), we make the following analytical prediction $\xi_{a,an}$ for the suppression of the axion abundance at late times

$$\frac{\xi_{a,an}}{\xi_{a,\mathrm{std}}} = \left(\frac{f_i}{f_a}\right)^{13/6} = \left(\frac{2}{N_\chi + 2}\right)^{13/12}. \tag{19}$$

Thus, the fractional change in the axion abundance compared to the standard theory should depend solely on the number of additional fields N_χ.

The relic abundance of χ is more involved to estimate. However, as we showed in our previous paper [14], the case of a single χ field has an estimated relic abundance $\xi_{\chi,an}$ of

$$\frac{\xi_{\chi,an}}{\xi_{a,\mathrm{std}}} = \frac{g_\chi^{3/4} g_{s*}}{g_*^{7/12} g_{s\chi}} \frac{T_{\mathrm{QCD}}^{2/3}}{\Lambda_0^{2/3}} \frac{f_a^{1/3}}{M_{\mathrm{Pl}}^{1/3}} \frac{\sqrt{\mu}\langle Y_i^2\rangle}{\langle\theta_i^2\rangle} \tag{20}$$

(where we have taken Eq. (41) of Ref. [14] and replaced $F \to 1$, as is appropriate to match onto the model here). Now in order to account for the total energy density stored in all of our N_χ fields, we simply sum over N_χ. And to account for our initial conditions, we choose $\langle Y_i^2\rangle = 1/(N_\chi + 2)$. Hence the total is

$$\frac{\xi_{\chi,an,\mathrm{total}}}{\xi_{a,\mathrm{std}}} = \frac{g_\chi^{3/4} g_{s*}}{g_*^{7/12} g_{s\chi}} \frac{T_{\mathrm{QCD}}^{2/3}}{\Lambda_0^{2/3}} \frac{f_a^{1/3}}{M_{\mathrm{Pl}}^{1/3}} \frac{\sqrt{\mu}}{\langle\theta_i^2\rangle} \frac{N_\chi}{N_\chi + 2}. \tag{21}$$

Interestingly, this shows that for large N_χ, the abundance is fixed. Hence in order to ensure the χ abundance is small, we only need to assume a small χ mass, which suppresses this through the factor $\sqrt{\mu} = \sqrt{m_\chi/m_a}$. Since the χ mass is protected by our original symmetry, a small χ abundance is plausible within the effective theory.

We note that this simple analytical estimate for ξ_χ tends to overpredict the abundance of χ. This is because the estimate relies on the mass term μ^2 in Eq. (14) controlling the onset of oscillations of χ. However, in the regime when $\theta_\tau^2 > \mu^2$, this hierarchy causes the χ field to oscillate a bit earlier, and therefore have a smaller abundance than predicted here. These formulas for ξ_χ become more accurate when $\theta_\tau^2 < \mu^2$ when $H \sim m_{\chi,\mathrm{eff}}$. Crucially, this ambiguity does not influence the prediction for the abundance of the axion because, regardless of the time of χ oscillations, the axion always oscillates first and thus its abundance is determined appropriately.

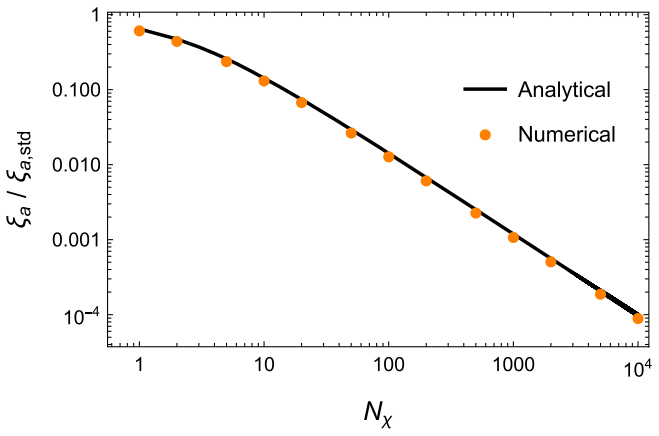

FIG. 1. The numerically computed ratio $\xi_a/\xi_{a,\mathrm{std}}$ is shown for various values of N_χ in the orange points, indicating that the abundance is suppressed with growing N_χ. It can be seen that the analytical prediction (solid black) for the abundance suppression is accurately reproduced by the numerical solution. In this plot, the numerical solutions correspond to $f_a \sim 10^{14}$ GeV, but the behavior is expected to be the same for any large f_a.

V. NUMERICAL ANALYSIS

The dynamics of the model can be solved for numerically. Recall that the N_χ additional fields are treated as having identical histories, and therefore one needs only to solve the two coupled differential equations (13) and (14). Below, we present the time-evolution and the effects of choosing different numbers N_χ. We show first the behavior of the altered abundance compared to a standard axion $\xi_a/\xi_{a,\mathrm{std}}$ in Fig. 1. We see good agreement between our analytical prediction for the axion relic abundance (solid line) and the numerically obtained abundance.

Next, we display the detailed numerical results for the case $f_a \sim 10^{14}$ GeV as an example, choosing $N_\chi = 1$, 10, 100, 1000 in Figs. 2–5. We see in these examples too that the analytical estimate (dashed line) matches well with the numerical solution.

To raise toward simple estimates for the grand unified scale, one could strive for $f_a \sim 10^{15}$–10^{16} GeV. However, numerically solving for the dynamics of these choices for large N_χ presents challenges. To capture the behavior of the light χ_j fields, one needs to integrate through large amounts of time, at which point the rapid oscillations of the axion become difficult to handle numerically. Though we are confident that the results presented for $f_a \sim 10^{14}$ GeV can be used to infer results for larger f_a, it remains important to

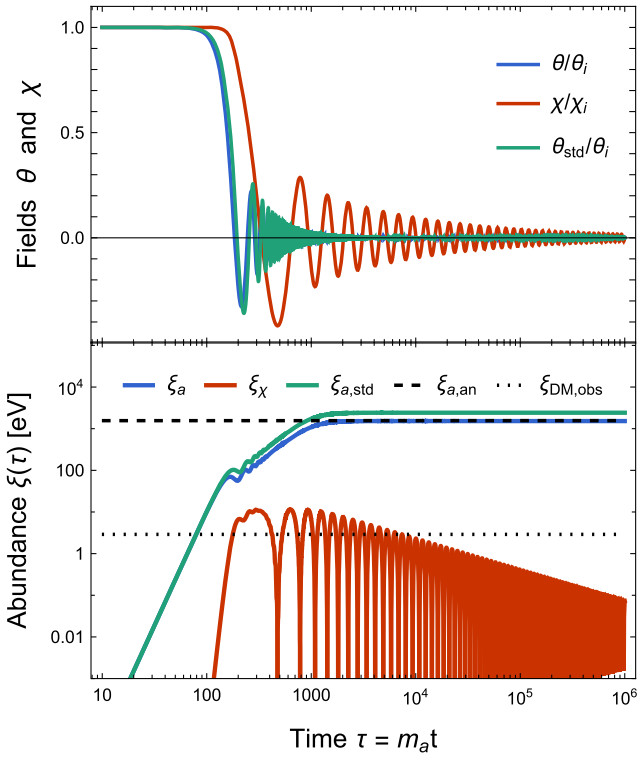

FIG. 2. The results of numerical solutions of the field equations are shown for $f_a \sim 10^{14}$ GeV with $N_\chi = 1$ and $m_\chi/m_a = 10^{-5}$. Top: the time evolution of the θ and χ_j fields. Bottom: the abundances of θ and χ_j.

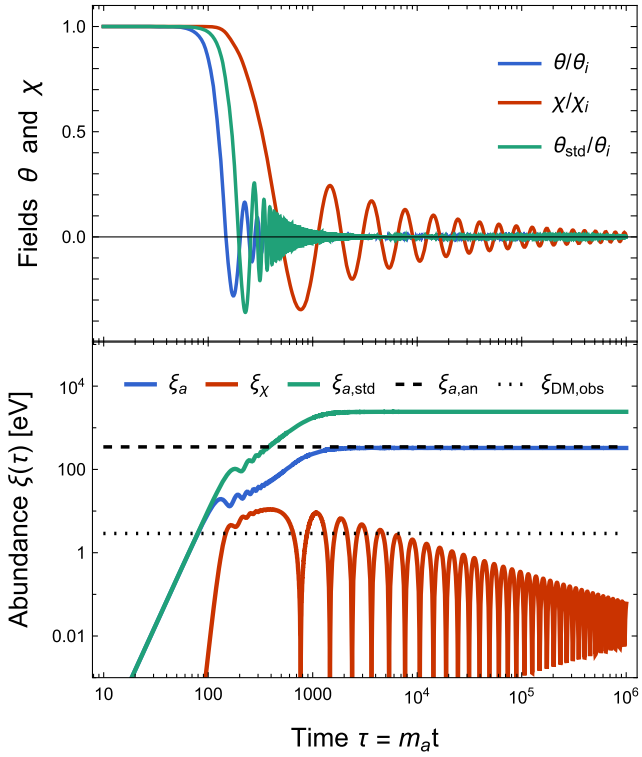

FIG. 3. The results of numerical solutions of the field equations are shown for $f_a \sim 10^{14}$ GeV with $N_\chi = 10$ and $m_\chi/m_a = 10^{-5}$. Top: the time evolution of the θ and χ_j fields. Bottom: the abundances of θ and χ_j.

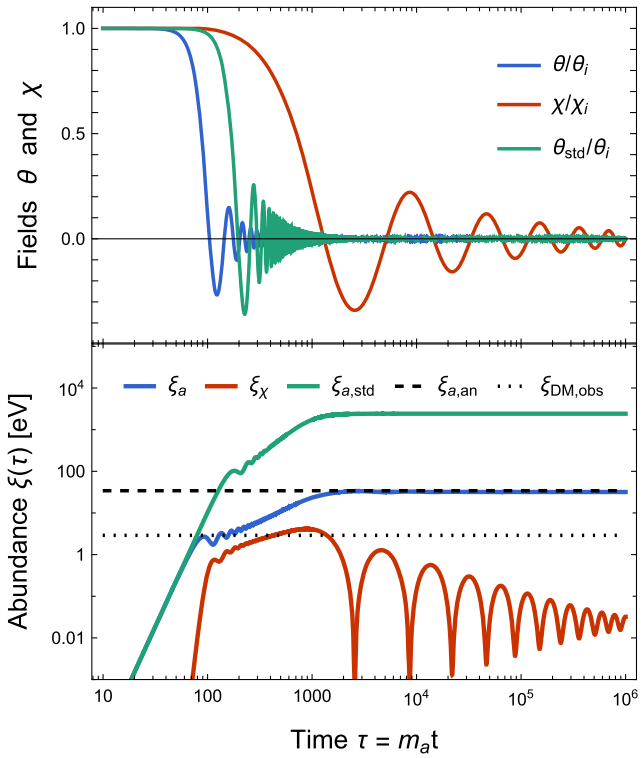

FIG. 4. The results of numerical solutions of the field equations are shown for $f_a \sim 10^{14}$ GeV with $N_\chi = 100$ and $m_\chi/m_a = 10^{-5}$. Top: the time evolution of the θ and χ_j fields. Bottom: the abundances of θ and χ_j.

verify this explicitly with more sophisticated numerical methods. Furthermore, the effects of inhomogeneities may be important. We discuss these effects in Appendix B, arguing that the homogeneous analysis is sufficient to capture the behavior of the model.

In any case, our existing numerics shows that for large N_χ, the final axion abundance is reduced as anticipated. Extending to larger f_a, we need larger N_χ; $f_a \sim 10^{15}$ and 10^{16} GeV require $N_\chi \sim 10^4$ and 10^5 new species, respectively, with corresponding huge symmetry groups.

VI. DISCUSSION

There are many interesting points to discuss within this framework. We shall discuss several key points in this section.

A. Isocurvature fluctuations

As is well known, if there are light fields present during cosmic inflation, they acquire a de Sitter temperature and fluctuate as $\delta\phi \sim T_{dS} = H_{\rm inf}/(2\pi)$ per Hubble time. If such fields go on to provide a significant fraction of the dark matter, then this translates into significant isocurvature fluctuations at early times, which leave an imprint on the CMB (e.g., see Refs. [12,13].) Then if the Hubble scale of inflation is large, the amplitude of these isocurvature modes is large, and ruled out by current constraints.

In standard axion models with very high f_a, the PQ symmetry is spontaneously broken during inflation, as $T_{dS} = H_{\rm inf}/(2\pi) \ll f_a$. This results in the light axion forming during inflation and giving rise to large isocurvature fluctuations. However, in the presence of our new class of models the situation is altered. In the presence of many fields, the condition for symmetry breakdown depends on the combined fluctuations of our $N_\chi + 2$ fields

$$\Delta^2 = (\delta\phi_1)^2 + (\delta\phi_2)^2 + \sum_{j=1}^{N_\chi} (\delta\chi_j)^2. \tag{22}$$

Symmetry breakdown occurs when these fluctuations are smaller than the PQ sale f_a, i.e., $\Delta \lesssim f_a$. Since each field acquires the de Sitter fluctuations $\delta\phi \sim T_{dS}$, the condition is

$$T_{dS} \lesssim \frac{f_a}{\sqrt{N_\chi + 2}}. \tag{23}$$

This condition is much more difficult to satisfy when N_χ is large, as we are assuming here. Therefore it is much more plausible that the symmetry remains unbroken during inflation. Currently inflation has its Hubble scale bounded

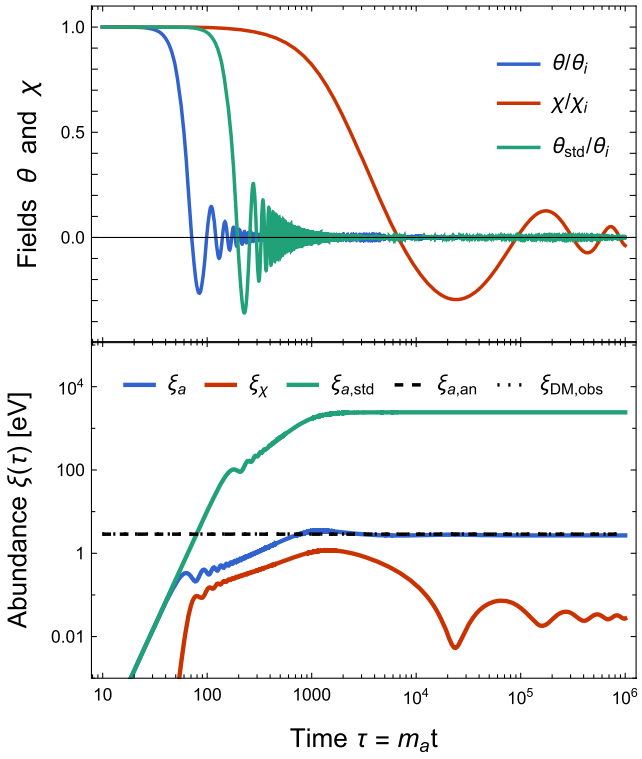

FIG. 5. The results of numerical solutions of the field equations are shown for $f_a \sim 10^{14}$ GeV with $N_\chi = 1000$ and $m_\chi/m_a = 10^{-5}$. Top: the time evolution of the θ and χ_j fields. Bottom: the abundances of θ and χ_j.

by the lack of detection of B-modes in the CMB. This corresponds to a bound on the tensor to scalar ratio of $r \lesssim 0.04$ [34,35]. Converting this to a Hubble scale and in turn a de Sitter temperature, we have $T_{dS} \lesssim 10^{13}$ GeV. Hence even if we push f_a toward $\sim 10^{14}$ GeV, we can plausibly violate inequality (23) with $N_\chi \sim 10^3$. In this case, there are no appreciable isocurvature modes generated, which is compatible with current data. Then, pushing to even higher f_a requires larger N_χ to suppress its abundance, and this larger N_χ in turn alters the condition in Eq. (23). For $f_a \sim 10^{15}$ GeV, the required $N_\chi \sim 10^4$ is sufficient to obtain symmetry-breaking after inflation, while for $f_a \sim 10^{16}$ GeV, one needs more than the required $N_\chi \sim 10^5$ to avoid isocurvature bounds.

On the other hand, as the scale of inflation (and the de Sitter temperature) are lowered, the condition for symmetry breakdown is easier to obey and appreciable isocurvature modes can arise. We note that in our setup here, this is primarily carried by the axion, as the abundance of the χ fields are small. Other ideas to suppress the isocurvature modes include [36].

B. Defects

When spontaneous symmetry-breaking occurs after the end of inflation, as is argued in the previous section, there can be the creation of topological defects. For the standard PQ models, cosmic strings arise from the Kibble mechanism where super horizon regions of space acquire a different $\langle \theta_i \rangle$, connecting in configurations with nontrivial winding. When the $U(1)_{\rm PQ}$ is explicitly broken by QCD instantons, the number of degenerate vacua in the periodicity of $\theta \in [0, 2\pi)$ dictate the number of domain walls attached to each string (N_{DW}) in the string-domain-wall network of topological defects.

If $N_{DW} > 1$, the network is stable and dominates the universe, over-closing it. If $N_{DW} = 1$, there are no stable domain walls, only cosmic strings. After the QCD phase transition, the would-be domain walls are pulled together and decay, producing an additional source of axions. It has been debated in the literature whether this enhances the final abundance compared to the misalignment value by a factor of a few or dozens (see, e.g., Refs. [37,38]). Regardless, the standard axion requires the consideration of this topological defect network to accurately predict the axion abundance.

In contrast, when the classical symmetry of the scalar sector has been promoted from $U(1) = SO(2)$ to $SO(N_\chi + 2)$, the story is quite different. When the spontaneous symmetry breaking occurs, a string network will not form via the Kibble mechanism, since there is now multiple Goldstone bosons. For example, if $N_\chi = 1$, the spontaneously broken $SO(3)$ leaves behind two Goldstones, leading to the creation of monopoles. For larger N_χ, textures can be formed. For $N_\chi \geq 3$ these are nontopological and collapse when entering the horizon and so they have a small relic abundance. Outside the horizon, they enter a kind of scaling solution for large N_χ [39–42]. Any residual imprints from textures, would be a signature of this construction. However, the presence of the axion and χ masses will suppress textures at later times by making $\theta = \chi_j = 0$ a preferred value. In any case, for $N_{DW} = 1$, the theory is safe from over-closure from defects.

For $N_{DW} > 1$, there are still potentially dangerous domain walls from the axion's N_{DW} discrete minima. However, now the remaining $SO(N_\chi)$ symmetry, which is unbroken by QCD instantons, enhances the vacuum manifold to a N_χ-sphere. This prevents the stability of such domain walls. This is because even if locally the axion is trapped in one of its discrete minima, it can shed energy into this degenerate N_χ-sphere until it reaches another discrete vacuum, removing a domain wall. However, eventually the small but nonzero χ mass becomes relevant and suppresses this process. A full analysis of this process is beyond the scope of the current work.

C. Unitarity bounds

A concern in the model is that with a very large number of scalars, one should check that the theory remains unitary. Our fields have quartic interactions with one another of the form

$$\Delta\mathcal{L} = -\sum_{i,j}\frac{\lambda}{4}\chi_i^2\chi_j^2. \tag{24}$$

So for example, if we compute the annihilation cross section of a pair of particles "1" into any final state at energies above the PQ radial mass $m_{\rm PQ} = \sqrt{2\lambda}f_a$, we have

$$\sigma_{1\to\rm All} \approx \frac{\lambda^2(N_\chi+10)}{8\pi E_{\rm cm}^2}. \tag{25}$$

For large N_χ we risk violating the unitarity bound $\sigma \le (2\pi)/E_{\rm cm}^2$. Thus, we need to scale down $\lambda \le (4\pi)/\sqrt{N_\chi+10}$ to avoid this problem. So if $N_\chi \sim 10^4$, we need to impose $\lambda < 10^{-1}$, or so, to maintain perturbative unitarity. However, this does not appear to be a huge problem. In fact it self-consistently reinforces the lightness of χ, as discussed in Sec. III B.

D. Future work and plausibility

A very important question for future consideration is the plausibility of this new (large) symmetry group. Ideas within unification often involve large groups, such as $SO(10)$, etc., but we are making a case for potentially even much larger groups. Can this fit in and improve ideas within unification, or does it make the situation more difficult?

Relatedly, it is important to develop microscopic models with fermions. In standard axion models, there are additional heavy fermions that are charged under the $U(1)_{\rm PQ}$ symmetry. Naively this breaks our starting $SO(N_\chi+2)$ symmetry already. So it remains an open question to develop alternate models with fermions that may account for this altered symmetry structure. As it stands, we have a consistent effective field theory for a collection of scalars, one of which—the axion—is assumed to couple to gluons with a dimension 5 coupling $\theta G\tilde{G}$; QCD instantons still generate a potential for this and it still solves the strong CP problem, as the standard axion models do. A full UV completion though is an important direction for future work.

ACKNOWLEDGMENTS

We thank Jose Blanco-Pillado for helpful comments. M. P. H and Y.L acknowledge support from the Tufts Visiting and Early Research Scholars' Experiences Program (VERSE). M. P. H is supported in part by National Science Foundation grant No. PHY-2013953. I. J. A. acknowledges support from the John F. Burlingame Graduate Fellowship in Physics at Tufts.

APPENDIX A: REDUCED ACTION

After eliminating the radial mode $\rho = |\Phi|$, we obtain a reduced action for the remaining $N_\chi + 1$ light degrees of freedom. Let us organize the χ_j into a vector $\vec{\chi}$ to express it. We obtain

$$\mathcal{L} = \sqrt{-g}\left[\frac{1}{2}|\partial\vec{\chi}|^2 + \frac{1}{2}\frac{(\vec{\chi}\cdot\partial\vec{\chi})^2}{f_a^2-|\vec{\chi}|^2} + \frac{1}{2}(f_a^2-|\vec{\chi}|^2)(\partial\theta)^2 - \Lambda(T)^4(1-\cos\theta) - \frac{1}{2}m_\chi^2|\vec{\chi}|^2\right]. \tag{A1}$$

From this low energy action, all the results of the paper can be derived.

APPENDIX B: INHOMOGENEITIES

Let us expand around our homogeneous background fields as

$$\theta = \theta_0(t) + \varepsilon(t,\mathbf{x}), \qquad \chi_j = \chi_0(t) + \gamma_j(t,\mathbf{x}). \tag{B1}$$

We wish to work to quadratic order in the action. The zeroth order terms were already solved for numerically in this paper; the first order terms will vanish by the Euler-Lagrange equations. The second order Lagrangian for the perturbations is found to be

$$\begin{aligned}\frac{\mathcal{L}}{a^3} = {}& \frac{1}{2}|\partial\vec{\gamma}|^2 + \frac{\chi_0^2(\sum_j\partial\gamma_j)^2}{2(f_a^2-N_\chi\chi_0^2)} + \frac{1}{2}(f_a^2-N_\chi\chi_0^2)(\partial\varepsilon)^2 - 2\chi_0\dot{\theta}_0\sum_j\gamma_j\dot{\varepsilon} - \frac{1}{2}(m_\chi^2+\dot{\theta}_0^2)|\vec{\gamma}|^2 - \frac{1}{2}\Lambda^4(T)\varepsilon^2 \\ & + \frac{\dot{\chi}_0^2/f_a^2}{2(1-N_\chi\chi_0^2/f_a^2)^3}\left[(1+N_\chi\chi_0^2/f_a^2)^2\left(\sum_j\gamma_j\right)^2 + N_\chi^2\chi_0^2/f_a^2(1-N_\chi\chi_0^2/f_a^2)|\vec{\gamma}|^2\right] \\ & + \frac{\chi_0\dot{\chi}_0}{(f_a^2-N_\chi\chi_0^2)^2}\left[N_\chi(f_a^2-N_\chi\chi_0^2)\vec{\gamma}\cdot\dot{\vec{\gamma}} + (f_a^2+N_\chi\chi_0^2)\left(\sum_{i,j}\gamma_i\dot{\gamma}_j\right)\right].\end{aligned} \tag{B2}$$

While this action looks somewhat complicated, things are significantly simplified by identifying the normal modes of the system, which we can decompose into an adiabatic mode and a collection of isocurvature modes.

1. Adiabatic mode

The adiabatic mode has

$$\gamma \equiv \gamma_1 = \gamma_2 = \ldots = \gamma_{N_\chi} \quad \text{and} \quad \varepsilon \neq 0. \tag{B3}$$

The corresponding equations of motion in this case are

$$\ddot{\gamma} - \frac{\nabla^2}{a^2}\gamma + \left(3H + \frac{2N_\chi \chi_0 \dot{\chi}_0}{f_a^2 - N_\chi \chi_0^2}\right)\dot{\gamma} + \left((1 - 3N_\chi \chi_0^2/f_a^2)(m_\chi^2 + \dot{\theta}_0^2) + \frac{N_\chi (f_a^2 + N_\chi \chi_0^2)\dot{\chi}_0^2}{(f_a^2 - N_\chi \chi_0^2)^2}\right)\gamma + 2\chi_0(1 - N_\chi \chi_0^2/f_a^2)\dot{\theta}_0 \dot{\varepsilon} = 0 \quad \text{(B4)}$$

$$\ddot{\varepsilon} - \frac{\nabla^2}{a^2}\varepsilon + \left(3H - \frac{2N_\chi \chi_0 \dot{\chi}_0}{f_a^2 - N_\chi \chi_0^2}\right)\dot{\varepsilon} + \frac{\Lambda^4(T)}{f_a^2 - N_\chi \chi_0^2}\varepsilon - \frac{2N_\chi \partial_t (a^3 \chi_0 \dot{\theta}_0 \gamma)}{a^3 (f_a^2 - N_\chi \chi_0^2)} = 0. \quad \text{(B5)}$$

In the equation for γ, we see that at early times when χ_0 takes its initial value of $\chi_i = f_a/\sqrt{N_x + 2}$, the coefficient of $(m_\chi^2 + \dot{\theta}_0^2)\gamma$ is negative. This represents a tachyonic instability. The instability is tempered by the presence of Hubble friction and the other effective mass term, however there is still an instability. This is seen in Fig. 6. This affects the *homogeneous* mode for γ the most, as any derivatives only enhance the mass by k^2/a^2. Hence this means that the system may shift slightly to a different homogeneous mode. Plausibly, this instability will lead to a small shift in the final relic abundance. This has been verified numerically in the case worked out in Fig. 6, but this deserves further consideration.

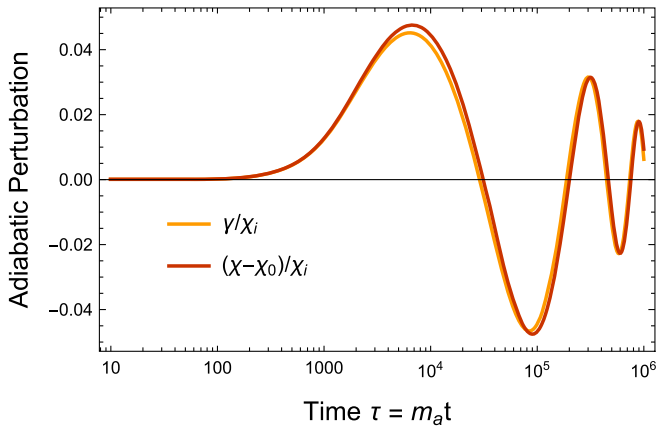

FIG. 6. A homogeneous adiabatic perturbation in the χ_j fields by $\chi_j = \chi_0 + \gamma$ for $f_a \sim 10^{14}$ GeV and $m_\chi/m_a = 10^{-5}$. We take $N_\chi = 1000$ and the initial perturbation value $\gamma_i = 10^{-4}\chi_i$ with $\chi_i = f_a/\sqrt{N_\chi + 2}$. Note this γ_i is only 1 order of magnitude smaller than the maximum value allowed for an adiabatic perturbation to be well defined given the constraint $\rho^2 + \sum_j \chi_j^2 = f_a^2$. The yellow curve is the result from the linear theory, while the red curve is from solving the full equations.

2. Isocurvature modes

For new qualitative behavior, we need to explore the remaining set of possible perturbations. The remaining $N_\chi - 1$ eigenmodes are all of the isocurvature form in which every fluctuation in a γ_j is compensated by an equal and opposite fluctuation in another $\gamma_{i\neq j}$, i.e.,

$$\sum_j \gamma_j = 0 \quad \text{and} \quad \varepsilon = 0. \quad \text{(B6)}$$

As an example, we can have

$$\gamma \equiv \gamma_1 = -\gamma_2 \quad \text{and} \quad \gamma_3 = \cdots = \gamma_{N_\chi} = \varepsilon = 0. \quad \text{(B7)}$$

The equation for this γ is of the relatively simple form

$$\ddot{\gamma} - \frac{\nabla^2}{a^2}\gamma + 3H\dot{\gamma} + \left(\left(1 - \frac{N_\chi \chi_0^2}{f_a^2}\right)(m_\chi^2 + \dot{\theta}_0^2) + \frac{N_\chi \dot{\chi}_0^2}{f_a^2 - N_\chi \chi_0^2}\right)\gamma = 0. \quad \text{(B8)}$$

Here we see that there is no tachyonic instability. Here the coefficient of $m_\chi^2 + \dot{\theta}_0^2$ is the positive factor

$$\alpha \equiv 1 - \frac{N_\chi \chi_0^2}{f_a^2}. \quad \text{(B9)}$$

At early times with $\chi_0 = \chi_i = f_a/\sqrt{N_\chi + 2}$ this takes on the small value $\alpha = 1/(1 + N_\chi/2)$. At later times, once χ_0 has rolled down, then one has $\alpha \approx 1$.

Since there is no tachyonic instability here, the homogenous mode is essentially stable. However, as θ_0 oscillates there could be resonance into γ modes of finite wave number. We now turn to study this possibility.

APPENDIX C: RESONANCE

1. Parametric resonance

The driving term $\dot{\theta}_0^2\gamma$ can potentially give rise to parametric resonance. As a starting point, let us ignore Hubble expansion for the moment (we will compare to it soon). Once the axion is oscillating, we can approximate its background behavior as

$$\theta_0 = \theta_a \sin(m_{\text{eff}} t), \quad \text{(C1)}$$

where the axion's effective mass $m_{\text{eff}} = m_a/\sqrt{\alpha}$ is controlled by α. The amplitude of oscillation θ_a is initially $\theta_a \sim 1$, and we shall treat its decreasing over time adiabatically.

With χ_0 either frozen at χ_i at early times or negligible at late times, the k-space isocurvature mode equation (B8) is

$$\ddot{\tilde{\gamma}} + (m_\chi^2 \alpha + k^2 + m_a^2\theta_a^2/2 + m_a^2\theta_a^2 \cos(2m_{\text{eff}} t)/2)\tilde{\gamma} = 0. \quad \text{(C2)}$$

This is of the form of the Mathieu equation. For $\theta_a < 1$ and $m_\chi \ll m_a$, we can focus on the narrow resonance regime, in which the maximum Floquet exponent can be shown to be given by (e.g., see Ref. [43])

$$\mu_{\max} = \frac{m_a^2 \theta_a^2}{8 m_{\rm eff}} \tag{C3}$$

near the wave number $k_{\max} = \sqrt{m_{\rm eff}^2 - m_\chi^2 \alpha - m_a^2 \theta_a^2/2}$.

Now we would like to include expansion in a simple adiabatic way. Here things appear complicated since the amplitude of oscillations θ_a is decreasing, while the axion mass increases due to its temperature dependence, until well after the QCD phase transition. However, there is a nice simplification: the axion number density is $n_a = \frac{1}{2} m_{\rm eff} f_{\rm eff}^2 \theta_a^2$ and due to entropy conservation, it redshifts in the usual way as $n_a \propto 1/a^3$. Since $m_{\rm eff} f_{\rm eff}^2 \propto \sqrt{\alpha}$ this implies that $\mu \propto 1/a^3$ also. We can compare this to Hubble; in a radiation era we have $H \propto 1/a^2$. Recall that oscillations begin when $3H \approx m_{\rm eff}$. So the ratio of the Floquet exponent to Hubble can be expressed over time as

$$\frac{\mu_{\max}}{H} \approx \alpha \frac{3\theta_i^2}{8} \frac{a_{osc}}{a} \tag{C4}$$

where a_{osc} is the scale factor when the axion starts oscillating. This means that when oscillations begin and $\theta_i \sim 1$, the ratio $\mu_{\max}/H \sim \alpha_i = 1/(1 + N_\chi/2)$. So for large N_χ the condition for resonance $\mu_{\max}/H \gtrsim 1$ is not satisfied. At later times, when χ_0 has itself decreased and $\alpha \approx 1$, the ratio a_{osc}/a is small and the resonance condition $\mu_{\max}/H \gtrsim 1$ is still not satisfied.

2. Forced resonance

By causality, fields will tend to be inhomogeneous on superhorizon scales. As modes enter the horizon, they will acquire a wavelength of the order Hubble momentarily. This provides a type of inhomogeneous background of χ waves $\chi_s = \chi_s(t, \mathbf{x})$. These can be inserted back into the χ_j equation to act as a source for itself

$$\chi_j = \chi_s + \gamma_j. \tag{C5}$$

By subtracting out the static piece of the axion $\langle \dot\theta_0^2 \rangle$ to remove possible secular growth, we have

$$\ddot\gamma_j - \frac{\nabla^2}{a^2}\gamma_j + 3H\dot\gamma_j + \tilde m_\chi^2 \gamma_j = -\alpha(\dot\theta_0^2 - \langle \dot\theta_0^2 \rangle)\chi_s \tag{C6}$$

where $\tilde m_\chi^2 \equiv m_\chi^2 \alpha + \alpha \langle \dot\theta_0^2 \rangle$. Let us again ignore expansion and write the axion as $\theta_0 = \theta_a \sin(m_{\rm eff} t)$. Then we have

$$\ddot\gamma_j - \nabla^2 \gamma_j + \tilde m_\chi^2 \gamma_j = -\theta_a^2 m_a^2 \cos(2 m_{\rm eff} t)\chi_s/2. \tag{C7}$$

We can expand the waves in terms of its Fourier transform as

$$\chi_s = \int \frac{d^3k}{(2\pi)^3}(c_k e^{-i\omega_k t} + c_{-k}^* e^{i\omega_k t}) e^{ikx} \tag{C8}$$

with $\omega_k = \sqrt{\tilde m_\chi^2 + k^2}$. For $\omega_k \approx m_{\rm eff}$ this gives rise to resonant behavior from a forced oscillator. By imposing the initial condition $\gamma_j(0) = \gamma_j'(0) = 0$, we can readily solve this equation. By passing to Fourier space and focusing on the near resonance wave numbers we find

$$\tilde\gamma_j = \frac{i m_a^2 \theta_a^2 (c_k e^{i m_{\rm eff} t} - c_{-k}^* e^{-i m_{\rm eff} t}) \sin(t(m_{\rm eff} - \omega_k))}{8 m_{\rm eff}(m_{\rm eff} - \omega_k)}. \tag{C9}$$

Now the energy can be written as an integral over a kind of k-space density as

$$E_\gamma = \int \frac{d^3k}{(2\pi)^3} \rho_k \tag{C10}$$

with

$$\rho_k = \sum_j \left(\frac{1}{2}|\dot{\tilde\gamma}_j|^2 + \frac{1}{2}\omega_k^2 |\tilde\gamma_j|^2 \right). \tag{C11}$$

By inserting the above solution, and again staying near resonance, we have

$$\rho_k = \sum_j \frac{(|c_k|^2 + |c_{-k}|^2) m_a^4 \theta_a^4 \sin^2(t(m_{\rm eff} - \omega_k))}{64 (m_{\rm eff} - \omega_k)^2}. \tag{C12}$$

At late times we can simplify the time dependence by using the identity

$$\frac{\sin^2(t(m_{\rm eff} - \omega_k))}{(m_{\rm eff} - \omega_k)^2} \to \pi t \delta(m_{\rm eff} - \omega_k) \tag{C13}$$

which is the standard simplification that leads to Fermi's golden rule. This shows that the energy is growing linearly in time, as anticipated from a forced oscillator near resonance. We can insert into the energy integral, immediately carry out the radial integral using the delta function, giving the rate of change of energy into γ as

$$\dot E_\gamma = \sum_j \frac{V(\mathcal{N}_{k_r} + \mathcal{N}_{-k_r}) m_a^4 \theta_a^4 k_r}{256\pi} \tag{C14}$$

where $k_r = \sqrt{m_{\rm eff}^2 - \tilde m_\chi^2}$ is again the resonant wave number and V is the volume of some large box in which we perform this computation. Here we have defined the occupancy number

$$\mathcal{N}_{k_r} \equiv \frac{2|c_{k_r}|^2\omega_{k_r}}{V}. \tag{C15}$$

It can be readily checked that the total number of particles of each species is given by

$$N_j = \sum_k \mathcal{N}_k = V\int \frac{d^3k}{(2\pi)^3}\mathcal{N}_k \tag{C16}$$

so this definition of $\mathcal{N}_k$ makes sense.

Now we are in a regime in which the energy density of axions is dominant and approximated as

$$\rho_a = \frac{1}{2}m_a^2 f_a^2\theta_a^2 \tag{C17}$$

with total energy $E_a = \rho_a V$. By energy conservation, the energy that is going into γ must come from the axion; so we have $\dot{E}_a = -\dot{E}_\gamma$. The corresponding relative rate of change of energy is

$$\Gamma_a = \frac{|\dot{E}_a|}{E_a} = \sum_j \frac{(\mathcal{N}_{k_r} + \mathcal{N}_{-k_r})m_a^2\theta_a^2 k_r}{128\pi f_a^2}. \tag{C18}$$

On the other hand, we can compare this to the perturbative annihilation rate of 2 axions into 2χ particles via the quartic coupling $\Delta\mathcal{L} = -|\vec{\chi}|^2(\partial\phi)^2/(2f_{\rm eff}^2)$. For nonrelativistic axions, this can be readily shown to be

$$\Gamma_{2\to 2} = n_a\sum_j \sigma_j v = n_a\sum_j \frac{m_{\rm eff}\hat{k}_r}{16\pi f_{\rm eff}^4} \tag{C19}$$

where $\hat{k}_r = \sqrt{m_{\rm eff}^2 - m_\chi^2}$ is the on-shell wave number in vacuum.

Then noting that the axion number density is

$$n_a = \frac{1}{2}m_{\rm eff}f_{\rm eff}^2\theta_a^2 \tag{C20}$$

we see that the energy rate of change is (ignoring the tiny difference between k_r and $\hat{k}_r$)

$$\Gamma_a \sim \alpha^2\sum_j (\mathcal{N}_{k_r} + \mathcal{N}_{-k_r})n_a\sigma_j v. \tag{C21}$$

So we get the standard annihilation rate, but enhanced in the classical field regime by the χ_j occupancy numbers of the resonant modes $\mathcal{N}_{k_r}$ and $\mathcal{N}_{-k_r}$ (reflecting the fact that the particles are pair produced back to back) and suppressed by the factor α^2 due to the dynamics being noncanonical.

At the time at which the axion begins oscillating $m_{\rm eff} \sim 3H$, one can anticipate the relevant χ modes have typically wave number $k \sim H$ due to causality; this means they are near resonance. The occupancy is then $\mathcal{N}_{k_r} \sim \chi_s^2/k_r^2 \sim f_{\rm eff}^2/m_{\rm eff}^2 = \alpha^2 f_a^2/m_a^2$. This gives an initial rate

$$\Gamma_{a,i} \sim \alpha^2 N_\chi m_{\rm eff}\theta_i^2. \tag{C22}$$

Since $\alpha^2 N_\chi \sim 1/N_\chi$ for large N_χ, then this is much smaller than Hubble $H \sim m_{\rm eff}$ at this time.

Furthermore, this redshifts very quickly, because not only is there a factor of $1/a^3$ from the axion's number density. But also the resonant occupancy numbers are depleting due to redshifting. At later times, it is difficult to find χ waves with high occupancy at the axion's mass.

[1] R. D. Peccei and H. R. Quinn, Phys. Rev. Lett. **38**, 1440 (1977).
[2] S. Weinberg, Phys. Rev. Lett. **40**, 223 (1978).
[3] F. Wilczek, Phys. Rev. Lett. **40**, 279 (1978).
[4] J. Preskill, M. B. Wise, and F. Wilczek, Phys. Lett. **120B**, 127 (1983).
[5] L. F. Abbott and P. Sikivie, Phys. Lett. **120B**, 133 (1983).
[6] M. Dine and W. Fischler, Phys. Lett. **120B**, 137 (1983).
[7] P. Svrček and E. Witten, J. High Energy Phys. 06 (2006) 051.
[8] S. Y. Pi, Phys. Rev. Lett. **52**, 1725 (1984).
[9] A. Linde, Phys. Lett. B **259**, 38 (1991).
[10] F. Wilczek, arXiv:hep-ph/0408167.
[11] M. Tegmark, A. Aguirre, M. J. Rees, and F. Wilczek, Phys. Rev. D **73**, 023505 (2006).
[12] P. Fox, A. Pierce, and S. D. Thomas, arXiv:hep-th/0409059.
[13] M. P. Hertzberg, M. Tegmark, and F. Wilczek, Phys. Rev. D **78**, 083507 (2008).
[14] I. J. Allali, M. P. Hertzberg, and Y. Lyu, Phys. Rev. D **105**, 123517 (2022).
[15] G. Dvali, arXiv:hep-ph/9505253.
[16] L. Heurtier, F. Huang, and T. M. Tait, J. High Energy Phys. 12 (2021) 216.
[17] R. T. Co, L. J. Hall, and K. Harigaya, Phys. Rev. Lett. **120**, 211602 (2018).
[18] R. T. Co, E. Gonzalez, and K. Harigaya, J. High Energy Phys. 05 (2019) 163.
[19] R. T. Co, E. Gonzalez, and K. Harigaya, J. High Energy Phys. 05 (2019) 162.
[20] R. T. Co, L. J. Hall, and K. Harigaya, Phys. Rev. Lett. **124**, 251802 (2020).

[21] R. Kitano and W. Yin, J. High Energy Phys. 07 (2021) 078.
[22] K. Nakayama and W. Yin, J. High Energy Phys. 10 (2021) 026.
[23] T. Kobayashi and R. K. Jain, J. Cosmol. Astropart. Phys. 03 (2021) 025.
[24] K. R. Dienes, J. Kost, and B. Thomas, Phys. Rev. D **93**, 043540 (2016).
[25] K. R. Dienes, J. Kost, and B. Thomas, Phys. Rev. D **95**, 123539 (2017).
[26] L. Di Luzio, M. Giannotti, E. Nardi, and L. Visinelli, Phys. Rep. **870**, 1 (2020).
[27] S. Ramazanov and R. Samanta, arXiv:2210.08407.
[28] G. G. di Cortona, E. Hardy, J. P. Vega, and G. Villadoro, J. High Energy Phys. 01 (**2016**) 034.
[29] A. Chaumet and G. D. Moore, arXiv:2108.06203.
[30] J. E. Kim, Phys. Rev. Lett. **43**, 103 (1979).
[31] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Nucl. Phys. **B166**, 493 (1980).
[32] A. R. Zhitnitsky, Sov. J. Nucl. Phys. **31**, 260 (1980).
[33] M. Dine, W. Fischler, and M. Srednicki, Phys. Lett. **104B**, 199 (1981).
[34] P. A. Ade *et al.*, Phys. Rev. Lett. **127**, 151301 (2021).
[35] Y. Akrami *et al.*, Astron. Astrophys. **641**, A10 (2020).
[36] Y. Bao, J. Fan, and L. Li, arXiv:2209.09908.
[37] M. Hindmarsh, J. Lizarraga, A. Lopez-Eiguren, and J. Urrestilla, Phys. Rev. Lett. **124**, 021301 (2020).
[38] M. Gorghetto, E. Hardy, and G. Villadoro, SciPost Phys. **10**, 050 (2021).
[39] D. N. Spergel, N. Turok, W. H. Press, and B. S. Ryden, Phys. Rev. D **43**, 1038 (1991).
[40] N. Turok and D. N. Spergel, Phys. Rev. Lett. **66**, 3093 (1991).
[41] W. Hu, D. N. Spergel, and M. J. White, Phys. Rev. D **55**, 3288 (1997).
[42] M. A. Amin and D. Grin, Phys. Rev. D **90**, 083529 (2014).
[43] M. P. Hertzberg, J. Karouby, W. G. Spitzer, J. C. Becerra, and L. Li, Phys. Rev. D **90**, 123529 (2014).