



Axion quality straight from the GUT

Luca Vecchi^a

Istituto Nazionale di Fisica Nucleare (INFN), Sezione di Padova, Padova, Italy

Received: 1 September 2021 / Accepted: 11 October 2021 / Published online: 23 October 2021
© The Author(s) 2021

Abstract Composite axion scenarios offer a robust field theoretic justification for the existence of a Peccei–Quinn symmetry of high quality. We present a new class of realizations that are naturally embedded in Grand-Unified Theories, retain asymptotic freedom for all gauge groups, and protect the axion shift symmetry up to operators of dimension 12. Our setup leads to a number of distinctive signatures at low energies. First, additional composite scalars are predicted; some of these are viable dark matter candidates for values of the decay constant that are likely too low for the QCD axion abundance to be relevant. Second, an approximate unification of the Standard Model gauge couplings takes place at the axion scale, while leaving the actual quark-lepton unification at much higher energies as usual. This suggests the existence of GUT relics with Standard Model gauge quantum numbers at potentially accessible scales.

Contents

1 Introduction	1
2 Warming up with $SU(N_c) \times SU(5)^2$	2
3 The $SU(N_c) \times SO(10)^4$ model	5
3.1 Model setup and symmetry breaking	5
3.2 UV contributions to the axion potential	6
3.3 Alternative realizations	7
4 Signatures	7
4.1 Gauge-coupling unification at f_a ?	7
4.2 Unstable ALPs at the TeV	8
4.3 Long-lived relics	8
Heavy NGBs	9
Axion density	9
4.3.1 Glueball dark matter	9
4.3.2 NGB dark matter	10
5 Conclusions	11
References	12

^ae-mail: luca.vecchi@pd.infn.it (corresponding author)

1 Introduction

The non-observation of the electric dipole moment of the neutron sets an upper bound on the QCD topological angle:

$$|\bar{\theta}| \lesssim 10^{-10}. \quad (1)$$

The smallness of $|\bar{\theta}|$ can be explained by postulating non-generic properties of the UV completion of the Standard Model (SM). The identification of such properties is a challenging theoretical and experimental task, called Strong CP Problem. Among the various solutions of the Strong CP Problem are UV completions that invoke a global $U(1)_{\text{PQ}}$ anomalous under QCD.

Within the new physics sector the symmetry $U(1)_{\text{PQ}}$ may be linearly realized, explicitly broken, or spontaneously broken. When the $U(1)_{\text{PQ}}$ is linearly realized, $\bar{\theta}$ can be rotated away becoming unphysical. By anomaly matching, once the massive new physics states are integrated out the low energy spectrum necessarily includes (composite or fundamental) massless colored fermions. Unfortunately, experiments firmly preclude the possibility of exotic massless quarks, and also the massless up-quark hypothesis [1–3] appears to be rather strongly rejected (see, e.g., [4] and more recently [5]). If instead the $U(1)_{\text{PQ}}$ is explicitly broken by the new physics, the full theory has a CP-odd angle $\bar{\theta}_{\text{full}} = \bar{\theta} - \bar{\theta}'$ constructed as a linear combination of the QCD topological angle and some other anomalous parameter $\bar{\theta}'$ of the new physics. Postulating a softly-broken Z_2 mirror symmetry such that $\bar{\theta}' = \bar{\theta}$ makes the overall phase $\bar{\theta}_{\text{full}}$ vanish [6]. Because this is the quantity we actually measure below the mirror sector scale, the strong CP problem is solved. To be compatible with collider searches, explicit realizations of this idea must ensure the scale of soft Z_2 breaking lies within a certain range.

In the most popular $U(1)_{\text{PQ}}$ -based approach to the Strong CP Problem, the anomalous Peccei–Quinn symmetry is spontaneously broken and an approximate Nambu–Goldstone boson, the QCD axion a , is predicted [7–9]. An anomalous

rotation of the colored fermions does not remove the topological angle; rather, it shifts the axion in such a way as to keep the dynamical field $\bar{a}/f_a = \bar{\theta} + a/f_a$ invariant. Remarkably, the vacuum energy $V_{\text{QCD}}(\bar{a})$ of QCD gets minimized at $\langle \bar{a} \rangle = 0$. This means that the value of the observable topological angle gets automatically relaxed to zero, $\bar{\theta}_{\text{obs}} \equiv \langle \bar{a} \rangle = 0$, provided the axion potential is entirely dominated by non-perturbative QCD effects. The QCD axion has been sought after for many years. No evidence has ever been found and the experimental searches are interpreted as constraints on the parameter space. Because all axion couplings are controlled by the scale f_a of spontaneous $U(1)_{\text{PQ}}$ breaking, data translate into lower bounds on this quantity. For the invisible axion models considered here, the most stringent constraints read $f_a \gtrsim 10^8$ GeV. Such large scales make the axion potential extremely sensitive to tiny effects from unknown physics at the shortest imaginable distances [10–12].

Being an anomalous symmetry, $U(1)_{\text{PQ}}$ cannot be exact in the far UV. At best, we may expect it to be broken solely by quantum gravity effects parametrized at low energies via Planck-scale suppressed interactions. These induce corrections ΔV to the axion potential that are generically not aligned with $V_{\text{QCD}}(\bar{a})$, the origin of field space is destabilized and a non-vanishing observable $\bar{\theta}_{\text{obs}} \equiv \langle \bar{a} \rangle/f_a$ is re-introduced. The full potential is now the sum of a term due entirely to QCD and one from tiny effects at the Planck scale:

$$V(\bar{a}) = V_{\text{QCD}}(\bar{a}) + \Delta V(\bar{a}). \quad (2)$$

At leading order in ΔV we find

$$\bar{\theta}_{\text{obs}} = - \left. \frac{\Delta V'}{f_a V''_{\text{QCD}}} \right|_{\langle \bar{a} \rangle = 0}. \quad (3)$$

The question of why $|\bar{\theta}_{\text{obs}}|$ is so small becomes, in axion models, the question of why the $U(1)_{\text{PQ}}$ is so accurate despite it being only an approximate symmetry of the UV dynamics. This “ $U(1)_{\text{PQ}}$ quality problem” is so significant because the QCD potential is controlled by physics at the GeV scale, $V_{\text{QCD}} \sim m_\pi^2 f_\pi^2$, whereas ΔV is a function of the large f_a and the Planck scale $f_{\text{Pl}} = 2.4 \times 10^{18}$ GeV. Using naive dimensional analysis we expect $\Delta V \sim c 16\pi^2 f_a^4 (f_a/f_{\text{Pl}})^{d-4}$, with c some coefficient and d the dimension of the associated $U(1)_{\text{PQ}}$ -violating operator. For a generic, maximally strong UV dynamics $c \sim 1$. As long as $f_a \ll f_{\text{Pl}}$ one can hope to suppress the destabilizing effect by making sure the operator is sufficiently irrelevant. Plugging these estimates in (3), the experimental bound (1) translates into $d \gtrsim 9$ for $f_a \sim 10^8$ GeV. A scale of order, say, $f_a \sim 10^{11}$ GeV would instead require $d \gtrsim 12$. As f_a approaches the Planck scale the only way to prevent $|\bar{\theta}_{\text{obs}}|$ from exceeding the experimental limit is to have $|c| \lesssim 10^{-10}$. Unfortunately, to a low energy observer such a condition is no different than the experimental input (1) we started with. The necessary non-generic UV condi-

tions to ensure $|c| \ll 1$ may be somehow realized in the correct theory of quantum gravity, for example in String theory. However, in this paper we take the point of view of the low energy observer and look for explanations within the realm of quantum field theory.

Robust realizations of the QCD axion within a field theory approach require two fundamental ingredients: an axion decay constant parametrically smaller than the Planck scale and an accurate $U(1)_{\text{PQ}}$. Composite axion models can feature both ingredients. In such a framework the axion arises as a Nambu-Goldstone boson (NGB) of some exotic strong dynamics that undergoes chiral symmetry breaking at a scale $\sim f_a$ [13]. As long as the axion constituents are chiral under the new gauge interactions, a number of accidental chiral symmetries, among which $U(1)_{\text{PQ}}$, may be very accurate. Furthermore, a value of f_a that is simultaneously parametrically larger than the weak scale and lower than the Planck scale is completely natural in composite axion models, where f_a gets generated via dimensional transmutation. In non-supersymmetric models with fundamental scalars, on the other hand, the value of f_a results from a fine-tuning that is much more severe than just setting $|\bar{\theta}| < 10^{-10}$ by hand in a given UV completion of the SM.

The first composite axion model with chiral constituents was proposed in [14]. In those types of models the fermion constituents are charged under three non-abelian gauge symmetries and to have a sufficiently accurate $U(1)_{\text{PQ}}$ some of the gauge groups must have rank bigger than a certain lower bound. As a consequence there is a structural tendency to developing low energy Landau poles [15]. One can evade this problem by replacing one or more of these non-abelian factors with abelian groups [16] (one should also worry about cubic anomalies in this case). The price to pay is however a loss of asymptotic freedom. Alternatively, one can construct asymptotically-free models using a “moose” gauge structure [17] that resembles axion scenarios in a local extra dimension [18]. The purpose of the present paper is proposing a new class of asymptotically-free composite axion models with a natural embedding in some Grand-Unified Theory (GUT) and a $U(1)_{\text{PQ}}$ protected up to operators of dimension $d = 12$.

2 Warming up with $SU(N_c) \times SU(5)^2$

We begin our discussion with a very simple model. Unfortunately its protection of the $U(1)_{\text{PQ}}$ is at most barely sufficient. Yet, this toy model allows us to illustrate some of the key features that characterize the more efficient scenarios discussed in Sect. 3.

Our toy-model is based on an $SU(N_c) \times SU(5)_1 \times SU(5)_2$ gauge symmetry. The axion constituents are shown in Table 1. They are all Weyl fermions, chiral under $SU(5)^2$. The $SU(N_c)$ is the dynamics that becomes strong and generates

Table 1 Axion constituents, and their accidental $U(1)$ symmetries, in our toy-model

	$SU(N_c)$	$SU(5)_1$	$SU(5)_2$	$U(1)_B$	$U(1)_V$	$U(1)_{PQ}$
ψ	\mathbf{N}_c	$\mathbf{10}$	$\mathbf{1}$	1	1	1
$\tilde{\psi}$	\mathbf{N}_c	$\bar{\mathbf{5}}$	$\mathbf{1}$	1	-2	-2
ψ^c	$\bar{\mathbf{N}}_c$	$\mathbf{1}$	$\bar{\mathbf{10}}$	-1	-1	1
$\tilde{\psi}^c$	$\bar{\mathbf{N}}_c$	$\mathbf{1}$	$\mathbf{5}$	-1	2	-2

a composite axion at scales of order f . The $SU(5)_1$ is to be identified with the GUT, with the SM fermions belonging to three families in the $\mathbf{10} \oplus \bar{\mathbf{5}} \in SU(5)_1$ and being singlets of $SU(N_c) \times SU(5)_2$. As usual, scalar fields with large masses should be included to obtain a realistic model at the GUT scale. The axion physics is not going to be affected by them provided they do not carry $SU(N_c) \times SU(5)_2$ charges, which is what we assume. No other fundamental scalar is introduced.¹

At scales of order f_{GUT} the GUT is broken into the SM as usual:

$$SU(5)_1 \rightarrow G_{\text{SM},1} \equiv SU(3)_{G,1} \times SU(2)_{W,1} \times U(1)_{Y,1}. \quad (4)$$

At $f \ll f_{\text{GUT}}$, when the strong dynamics of the axion becomes important, it is appropriate to decompose $\psi, \tilde{\psi}$ in irreducible $G_{\text{SM},1}$ representations: $\psi = \psi_q \oplus \psi_u \oplus \psi_e, \tilde{\psi} = \psi_d \oplus \psi_\ell$, where $\psi_{q,u,e,d,\ell}$ are in the fundamental of $SU(N_c)$, singlets under $SU(5)_2$ and transform under $G_{\text{SM},1}$ as the SM quarks and leptons q, u, e, d, ℓ .

Let us next discuss under which conditions our gauge interactions remain asymptotically free. The requirement of asymptotic freedom for $SU(5)_2$ is satisfied for any $N_c \leq 13$. This is a very mild constraint. A much more significant upper bound on N_c is found demanding that $SU(5)_1$ remains UV-free. This bound is model-dependent because it depends on the additional GUT fields necessary to explain (4) and the SM Yukawa couplings. For definiteness we consider a minimal (non-supersymmetric) scenario in which the scalar responsible for breaking $SU(5)_1$ into the SM is in the representation **24** whereas the observed pattern of Yukawa couplings for the SM fermions may be reproduced with a Higgs in the **5**. Rather than introducing additional fundamental fields, which would quickly take our model away from asymptotic freedom, the necessary quark-lepton splitting is achieved via effective dimension-5 operators with two SM fermions and two powers of the scalars $\phi_H \sim \mathbf{24} \oplus \mathbf{5}$. These interactions

¹ We have nothing new to say about the generation and the stabilization of the hierarchy between the weak and the GUT scales. The hierarchy problem is assumed to be addressed in some way. On our part, the axion models proposed here do not introduce new fine-tunings; in particular, the new mass scale f_a is dynamically generated and can be naturally smaller than the Planck mass.

are large enough provided

$$f_{\text{GUT}} \gtrsim 10^{16} \text{ GeV}. \quad (5)$$

With this minimal GUT-scale field content the upper bound reads $N_c \leq 9$. The key assumption that $SU(N_c)$ confines and breaks its chiral symmetries provides instead a lower bound on N_c . Ignoring the weak gauging of $SU(5)^2$, the $SU(N_c)$ dynamics behaves as if it was a QCD-like theory with N_c colors and $N_f = 15$ massless flavors. According to a number of independent arguments (for an incomplete list see, e.g., [19–24]), confinement and dynamical chiral symmetry breaking occurs when $N_f/N_c \lesssim 3 - 4$. We will work in the regime $5 \leq N_c \leq 9$.

Under our assumptions, at $f \sim 10^8 - 10^9$ GeV the $SU(N_c)$ becomes strong and non-trivial vacuum condensates form:

$$\frac{1}{10} \langle \psi \psi^c \rangle = \frac{1}{5} \langle \tilde{\psi} \tilde{\psi}^c \rangle \sim 4\pi f^3 \sqrt{\frac{3}{N_c}}, \quad (6)$$

where we borrowed the factors of 4π from QCD and the powers of N_c from large N_c scaling.² The particle content of Table 1 features a number of accidental symmetries. We spare the reader of the details, which will instead be presented for the models of Sect. 3. The bottom line is that (6) break the gauge symmetry $G_{\text{SM},1} \times SU(5)_2$ into the diagonal G_{SM} , which is to be identified with the SM gauge symmetry we experimentally observe. In the process, many Nambu-Goldstone bosons (NGBs) are generated. Some are eaten via the Higgs mechanism, some acquire large masses squared $\sim g_{\text{SM}}^2 f^2 / N_c$, and finally four SM-neutral NGBs remain massless at the renormalizable level. Among these, three are associated to chiral symmetries that appear below the GUT scale but are in fact explicitly broken by the weak gauging of $SU(5)^2$, which explains why these are not shown in Table 1. As a consequence, these three NGBs acquire masses from dimension-6 interactions below the GUT scale (we will study them in Sect. 4.2). The actual QCD axion a is associated to the breaking of the anomalous $U(1)_{\text{PQ}}$ of Table 1.

The lowest dimensional operators that violate $U(1)_{\text{PQ}}$ are formed by the product of $\psi^\dagger \bar{\sigma}^\mu \tilde{\psi}$ and appropriate SM fermion currents, and have dimension $d = 6$. Yet, such interactions are innocuous. Indeed, to generate an axion potential ΔV in our QCD-like models one must break *both* $U(1)_{\text{PQ}} \subset SU(N_f)_L$ as well as $U(1)_{\text{PQ}} \subset SU(N_f)_R$. In the presence of explicit breaking of $SU(N_f)_L$ only, as it is the case for the dimension-6 interactions just mentioned, one could remove completely from the Lagrangian the non-derivative couplings of the axion performing a global $SU(N_f)_R$ rotation. This demonstrates that the above dimension-6 interactions are not

² For QCD in the chiral limit we have $\langle \psi \psi^c \rangle \simeq (240 \text{ MeV})^3$ at $\mu \sim 1$ GeV (no summation over flavor). This agrees with the naive dimensional analysis estimate $\langle \psi \psi^c \rangle \sim 4\pi f_\pi^3$ up to a number close to 1.

harmful unless they are accompanied by operators that also violate the $U(1)_{PQ} \subset SU(N_f)_R$ carried by $\psi^c, \tilde{\psi}^c$. Fortunately, the couplings of the latter fields are more severely constrained by invariance under $SU(5)_2$ and $SU(N_c)$, as we now argue.

The leading $SU(5)_2$ -invariant combinations that can break $U(1)_{PQ} \subset SU(N_f)_R$ are $\psi^c \tilde{\psi}^c \tilde{\psi}^c$ and $\psi^c \psi^c (\tilde{\psi}^c)^\dagger$. The former requires three more powers of $\psi, \tilde{\psi}$ to make an $SU(N_c)$ -singlet, leading us to operators of at least dimension $d = 9$. The other option, i.e. $\psi^c \psi^c (\tilde{\psi}^c)^\dagger$, may be combined with $\psi \psi (\tilde{\psi})^\dagger$, which again gives $d = 9$ operators. Hence, we conclude that ΔV is first induced by operators of the type

$$(\psi \psi^c)(\tilde{\psi} \tilde{\psi}^c)^2, \quad (\psi \psi^c)^2(\tilde{\psi} \tilde{\psi}^c)^\dagger. \tag{7}$$

Strictly speaking, one may also build an $SU(N_c)$ singlet from $\psi^c \psi^c (\tilde{\psi}^c)^\dagger$ by adding either ψ or $\tilde{\psi}$ (or even their conjugates if we are willing to pay the additional price of a gauge field strength), but the resulting operator would not be invariant under $G_{SM,1} \subset SU(5)_1$ unless at least an additional $SU(N_c)$ -singlet, say a Higgs and/or an even number of SM fermions, are added. One can verify that the effect of these operators on the axion potential is negligible compared to the one coming directly from (7).

Having established that the dominant contribution to ΔV arises from the set of operators in eq. (7), we can finally estimate the expected size of the observed topological angle (3). For later convenience we analyze the impact of a general operator of dimension d of the type $(\Psi \Psi^c)^{d/3}$, with Ψ, Ψ^c the axion constituents. We expect the operator to be suppressed by the ultimate cutoff, say the Planck scale $f_{Pl} \simeq 2.4 \times 10^{18}$. Naive dimensional analysis is used to estimate the Wilson coefficient. Denoting by g_{Pl} the typical coupling at the cutoff, e.g. the string coupling, the operator appears in the effective Lagrangian as

$$\mathcal{L}_{PQ} \supset \frac{c_{Pl}}{g_{Pl}^{d/3-2} f_{Pl}^{d-4}} (\Psi \Psi^c)^{d/3} + \dots, \tag{8}$$

where c_{Pl} is of order unity if we conservatively assume the operator is generated at tree-level by exotic states of mass $\sim g_{Pl} f_{Pl}$. For any $d > 6$ the Wilson coefficient gets smaller as the coupling increases (we have fixed f_{Pl}). Large- N counting suggests that the maximal allowed value satisfies $g_{Pl}^2 N_{dof} \sim 16\pi^2$, with $N_{dof} \gg 1$ the number of degrees of freedom. In our estimates we will saturate the bound with $N_{dof} \sim N_f N_c$. Fortunately, our numerical results are modestly affected by the unknown parameter g_{Pl} .

One should then run the Wilson coefficient down to energies of order the confinement scale, i.e. $c_{Pl} \rightarrow c_{IR}$. We may estimate this effect assuming (as in the large N_c approximation) the operator anomalous dimension factorizes as $d/3$ times the anomalous dimension of the quark bilinear $\Psi \Psi^c$. Limiting our analysis to a domain $g_{IR}^2 N_c / (4\pi) < 0.5 - 1$ in which a perturbative calculation is expected to be reliable,

we observe an enhancement of order $c_{IR}/c_{Pl} \sim 10 - 10^2$. The contribution to the axion potential due to (8) can thus be written as

$$\Delta V = \frac{c_{IR}}{g_{Pl}^{d/3-2} f_{Pl}^{d-4}} n_{\mathcal{O}} \left(4\pi f^3 \sqrt{\frac{3}{N_c}} \right)^{d/3} e^{2iq_a a/f} + \text{hc} \tag{9}$$

where q_a is the $U(1)_{PQ}$ charge of (8) and $n_{\mathcal{O}} \propto N_f^{d/3}$ is a numerical factor that depends on the number of condensates (6) that appear in the operator (8). For example, the second operator in (7) contributes with $n_{\mathcal{O}} \sim 10^2 \times 5^1$ and $q_a = 4/\sqrt{15}$. For generic complex coefficients c_{IR} , uncorrelated with the bare $\bar{\theta}$, ΔV implies a contribution to the QCD topological angle (see Eq. (3))

$$|\bar{\theta}_{\text{obs}}| = \frac{(m_u + m_d)^2 (4\pi)^2 f_a^4}{m_u m_d m_\pi^2 f_\pi^2} \left(\frac{f_a}{f_{Pl}} \right)^{d-4} \times 4n_{\mathcal{O}} q_a |c_{IR}| |\sin(\text{Arg}(c_{IR}) - \bar{\theta})| \times \left(\frac{4\pi}{g_{Pl}} \right)^{d/3-2} (\mathcal{A}_a^G)^{d-1} \left(\frac{3}{N_c} \right)^{d/6}. \tag{10}$$

Here we used the leading order expression for the axion mass $m_a^2 = V''_{\text{QCD}} = [m_u m_d / (m_u + m_d)^2] m_\pi^2 f_\pi^2 / f_a^2$, with $m_u/m_d \simeq 0.48$, and used the definition $f = \mathcal{A}_a^G f_a$, where \mathcal{A}_a^G is the QCD anomaly coefficient determining the axion coupling to gluons (see Sect. 3.2 for details). Note that the powers of $\mathcal{A}_a^G = N_c/\sqrt{15} \propto N_c$ and N_f (implicitly included in $n_{\mathcal{O}}$), as well as the RG effects, go in the direction of enhancing $|\bar{\theta}_{\text{obs}}|$. Other than that, the estimate (10) agrees very well with the one given in Sect. 1.

Employing typical values of $n_{\mathcal{O}}$ and q_a (such as those that have been mentioned above (10)), the constraint (1) reads $f_a \lesssim 10^8$ GeV, which falls too short to convince us this is a viable model. One can certainly hope that the actual value of the unknown coefficient c_{Pl} may be suppressed in a realistic theory of gravity. But this is not what we do here. Our point of view is the one of an agnostic low energy observer who would like to find an answer to a puzzling result, i.e. (1), without having to resort to unjustified assumptions about the UV. In other words, we are interested in finding scenarios that work with coefficients c_{Pl} of order unity. Such realizations are presented in Sect. 3.

Before concluding this discussion we would like to stress that our toy model differs significantly from the one of Ref. [25], where the axion constituents are also in a chiral $\mathbf{10} \oplus \bar{\mathbf{5}}$ representation of some $SU(5)$. In that paper the gauge group $SU(5)$ is identified with the strong axion dynamics; the constituents are therefore chiral under the confining force but vector-like under the SM. Here the situation is completely reversed. Our axion constituents carry vector-like representations under the confining $SU(N_c)$ dynamics and are instead chiral under the weakly-gauged, grand-unified $SU(5)^2$. This

Table 2 Axion constituents and the abelian (continuous) accidental symmetries of the theory

	$SU(N_c)$	$SO(10)_1$	$SO(10)_2$	$SO(10)_3$	$SO(10)_4$	$U(1)_B$	$U(1)_V$	$U(1)_{PQ}$
ψ	\mathbf{N}_c	$\mathbf{16}$	$\mathbf{1}$	$\mathbf{1}$	$\mathbf{1}$	1	1	1
$\tilde{\psi}$	\mathbf{N}_c	$\mathbf{1}$	$\mathbf{1}$	$\mathbf{16}$	$\mathbf{1}$	1	-1	-1
ψ^c	$\overline{\mathbf{N}}_c$	$\mathbf{1}$	$\overline{\mathbf{16}}$	$\mathbf{1}$	$\mathbf{1}$	-1	-1	1
$\tilde{\psi}^c$	$\overline{\mathbf{N}}_c$	$\mathbf{1}$	$\mathbf{1}$	$\mathbf{1}$	$\overline{\mathbf{16}}$	-1	1	-1

latter feature characterizes all our constructions and has interesting phenomenological implications discussed in Sect. 4.1. The dimensionality of the leading operators that violate $U(1)_{PQ}$ is inherited from the chiral $\mathbf{10} \oplus \overline{\mathbf{5}} \in SU(5)$, and is thus the same here as well as [25], though.

3 The $SU(N_c) \times SO(10)^4$ model

The previous model does not quite reach the desired level of protection of the axion $U(1)_{PQ}$. The ultimate reason is that it is sufficient to have three powers of $\mathbf{10} \oplus \overline{\mathbf{5}} \in SU(5)$ to build a complete gauge singlet under the chiral $SU(5)^2$. Our next models prevent this from happening by embedding the axion constituents in the spinorial representation of $SO(10)$. Apart from a more significant suppression of ΔV , in many respects the new models are qualitatively similar to our earlier toy scenario.

3.1 Model setup and symmetry breaking

Consider then a model based on the gauge symmetry $SU(N_c) \times SO(10)^4$. The representations of the axion constituents are shown in Table 2, and are all chiral under the weakly-gauged $SO(10)^4$.

$SO(10)_1$ is the GUT and the SM fermions (plus the right-handed neutrinos) are as usual embedded in three generations of $\mathbf{16}_i$. They are singlets of $SU(N_c)$ and $SO(10)_{2,3,4}$. As in Sect. 2, other fields charged under $SO(10)_1$, but not under $SU(N_c)$ and $SO(10)_{2,3,4}$, can and must be included to obtain a realistic GUT. Specifically, the minimal (non-supersymmetric) set of Higgses belong to the representation $\phi_H \sim \mathbf{10} \oplus \mathbf{16} \oplus \mathbf{45} \in SO(10)_1$. These trigger the symmetry breaking pattern $SO(10)_1 \rightarrow G_{SM,1}$ at f_{GUT} . As long as (5) holds, the observed pattern of Yukawa couplings for the SM fermions may be reproduced with a renormalizable coupling to the scalar $\mathbf{10}$ and effective dimension-5 operators $\phi_H^2 \mathbf{16}_i \mathbf{16}_j / f_{Pl}$. Under these assumptions asymptotic freedom for $SO(10)^4$ implies an upper bound $N_c \leq 16$. Because in this new scenario the number of $SU(N_c)$ flavors has grown to 32, the lower bound on N_c from requiring chiral symmetry breaking becomes $N_c \geq 11$. In the end our model lives in the regime

$$11 \leq N_c \leq 16. \tag{11}$$

Table 3 Accidental $U(1)$ global symmetries of the fields of Table 2 arising below the GUT scale

	$U(1)_{S,L}$	$U(1)_{A1}$	$U(1)_{A2}$	$U(1)_{A3}$
ψ_q	-1	1	0	-1
ψ_u	-1	1	-1	1
ψ_e	-1	-3	-3	1
ψ_d	3	1	1	1
ψ_ℓ	3	-3	0	-1
ψ_N	-5	-3	3	1

As in the toy model, at $f \ll f_{GUT}$ the $SU(N_c)$ becomes strong and a non-trivial vacuum for the fermion bilinears form:

$$\frac{1}{16} \langle \psi \psi^c \rangle = \frac{1}{16} \langle \tilde{\psi} \tilde{\psi}^c \rangle \sim 4\pi f^3 \sqrt{\frac{3}{N_c}}. \tag{12}$$

Switching off the gauge $SO(10)^4$ couplings the particle content of Table 2 enjoys a large global $G_{\text{global}} \equiv SU(32)_L \times SU(32)_R \times U(1)_B$ of independent (non-anomalous) unitary rotations of $(\psi, \tilde{\psi})$ and $(\psi^c, \tilde{\psi}^c)$. The weak gauging of $SO(10)^4$ breaks G_{global} explicitly, leaving the global subgroup $SO(10)^4 \times U(1)_B \times U(1)_V \times U(1)_{PQ}$ intact, see Table 2. Below f_{GUT} we can decompose ψ in irreducible representations of $G_{SM,1}$: $\psi = [\psi_q \oplus \psi_u \oplus \psi_e] \oplus [\psi_d \oplus \psi_\ell] \oplus [\psi_N]$, with ψ_N a singlet of $G_{SM,1}$ and $SO(10)_{2,3,4}$. The global $SO(10)_1$ is thus further broken into $G_{SM,1} \times U(1)^4$. Without loss of generality the abelian factor $U(1)^4$ can be chosen to act only on the components $\psi_{q,u,e,d,\ell,N}$. It is very simple to identify it. Of the five independent non-anomalous rotations that these fields can be subject to, one is hyper-charge and is already included in $G_{SM,1}$; the remaining four are shown explicitly in Table 3.

The symmetry breaking pattern induced by the condensates (12) turns out to be

$$\begin{aligned} & [G_{SM,1} \times SO(10)_2 \times SO(10)_3 \times SO(10)_4] \\ & \times U(1)_B \times U(1)_V \times U(1)_{S,L} \times [U(1)^4]_{\text{axial}} \\ & \rightarrow [G_{SM} \times SO(10)_{3+4}] \\ & \times U(1)_B \times U(1)_V \times U(1)_{S,V}, \end{aligned} \tag{13}$$

where for brevity we introduced the notation

$$[U(1)^4]_{\text{axial}} \equiv U(1)_{\text{PQ}} \times U(1)_{A1} \times U(1)_{A2} \times U(1)_{A3} \tag{14}$$

for the broken axial charges. The condensate $\langle \psi \psi^c \rangle$ breaks $G_{\text{SM},1} \times SO(10)_2$ into the vectorial gauge subgroup G_{SM} , which plays the role of the observed SM, whereas $\langle \tilde{\psi} \tilde{\psi}^c \rangle$ preserves the diagonal subgroup of $SO(10)_3 \times SO(10)_4$. The unbroken abelian group is a linear combination of the global $U(1)_B \times U(1)_V \times U(1)_{S,L}$ and a gauged $U(1)_{S,R} \subset SO(10)_2$. Indeed, recall that the maximal subgroup of $SO(10)_2$ is an $SU(5) \times U(1)_{S,R}$ defined by $\mathbf{16} = \mathbf{10}_{-1} \oplus \bar{\mathbf{5}}_3 \oplus \mathbf{1}_{-5}$. Hence, there is a local $U(1)_{S,R} \subset SO(10)_2$ that acts on ψ^c exactly as $U(1)_{S,L}$ acts on ψ^\dagger , see Table 3. The vectorial combination of $U(1)_{S,L} \times U(1)_{S,R}$ is left unbroken by the condensates (12). This is what we denoted by $U(1)_{S,V}$ in (13).

Without the gauged $SO(10)^4$, the approximate global symmetry breaking pattern $G_{\text{global}} \rightarrow H_{\text{global}} \equiv SU(32)_V \times U(1)_B$ produces 1023 would-be NGBs in the $\mathbf{32} \otimes \bar{\mathbf{32}} - \mathbf{1} \in SU(32)_V$. We may find their gauge quantum numbers by decomposing the broken generators in irreducible representations of the unbroken gauge group. But it is as instructive to write a more compact expression showing the decomposition under the larger $SO(10)_{1+2} \times SO(10)_{3+4} \subset SU(32)_V$. The result is as follows

$$\begin{aligned} \text{NGBs} \sim & (\mathbf{45}, \mathbf{1}) \oplus (\mathbf{1}, \mathbf{45}) \\ & \oplus (\mathbf{210}, \mathbf{1}) \oplus (\mathbf{1}, \mathbf{210}) \\ & \oplus (\mathbf{16}, \bar{\mathbf{16}}) \oplus (\bar{\mathbf{16}}, \mathbf{16}) \\ & \oplus (\mathbf{1}, \mathbf{1}). \end{aligned} \tag{15}$$

The NGBs in the first line of (15) are unphysical and provide the longitudinal components of the broken vectors. The remaining NGBs are all physical. Most of them are charged under $G_{\text{SM}} \times SO(10)_{3+4}$ and get a positive mass squared $\sim g_{\text{unbroken}}^2 f^2 / N_c$. There are then four complete singlets that remain massless at the renormalizable level. As we anticipated for the model in Sect. 2, these arise from the breaking of (14). One is obviously the axion in the last line of (15), and is associated to the breaking of $U(1)_{\text{PQ}}$. The other three are contained in the SM-singlet components of $(\mathbf{210}, \mathbf{1})$. They arise from the breaking of the accidental $U(1)_{A1} \times U(1)_{A2} \times U(1)_{A3}$ appearing below the GUT scale. A completely analogous thing happens in the toy model of the previous section, as the reader can verify.

The four singlet NGBs will be denoted by a, a_1, a_2, a_3 , respectively. They can be described by a 32-dimensional special unitary matrix $\Sigma = e^{2ia_\alpha T_\alpha / f}$ transforming as $\Sigma \rightarrow L \Sigma R^*$ under $SU(32)_L \times SU(32)_R$, where the generators, normalized such that $\text{Tr}[T_\alpha T_\beta] = \delta_{\alpha\beta} / 2$, are diagonal and can be readily seen from Tables 2 and 3. In particular $T_{\text{PQ}} = \frac{1}{8} \text{diag}(\mathbf{1}_{16 \times 16}, -\mathbf{1}_{16 \times 16})$ and similarly for $T_{A1, A2, A3}$.

The factor of 2 in Σ is chosen to conform to the standard definition $\langle 0 | J_{\text{PQ}}^\mu(0) | a(p) \rangle = i f p^\mu$.

Some of these pseudo-scalars have anomalous couplings to the SM gauge fields proportional to the anomaly coefficients³ $\mathcal{A}_\alpha^{\text{gauge}} \delta^{AB} = 4N_c \text{Tr}[T_\alpha T_{\text{gauge}}^A T_{\text{gauge}}^B]$, where T_{gauge}^A are the SM generators, with A, B indices in the adjoint, and the trace runs over all the flavors of the axion constituents. For example, those of a read

$$\begin{aligned} -\delta \mathcal{L}_{\text{WZW}} \supset & \frac{g_G^2}{64\pi^2} \mathcal{A}_a^G \frac{a}{f} \epsilon^{\mu\nu\alpha\beta} G_{\mu\nu}^A G_{\alpha\beta}^A \\ & + \frac{g_W^2}{64\pi^2} \mathcal{A}_a^W \frac{a}{f} \epsilon^{\mu\nu\alpha\beta} W_{\mu\nu}^i W_{\alpha\beta}^i \\ & + \frac{g_Y^2}{64\pi^2} \mathcal{A}_a^Y \frac{a}{f} \epsilon^{\mu\nu\alpha\beta} B_{\mu\nu} B_{\alpha\beta}. \end{aligned} \tag{16}$$

with $\mathcal{A}_a^G = \mathcal{A}_a^W = N_c, \mathcal{A}_a^Y = 5N_c/3$. Similar considerations apply to $a_{1,2,3}$. Despite a unique linear combination of $a, a_{1,2,3}$ can be chosen to couple to the SM gluons, the actual QCD axion is to be identified with a . Indeed, as anticipated before and shown explicitly in Sect. 4.2, the scalars $a_{1,2,3}$ acquire masses above the TeV from dimension-6 interactions. Hence they cannot be relevant to the Strong CP Problem. The only potential impact they can have is to shift $\bar{\theta}$ above the QCD scale, an effect which is completely irrelevant to the QCD axion solution. On the other hand the QCD axion a remains massless up to a much higher degree, see Sect. 3.2. Our focus for the rest of the section will therefore be on a . Conventionally the scale f_a is defined as

$$f_a = \frac{f}{\mathcal{A}_a^G} = \frac{f}{N_c}. \tag{17}$$

Before moving on let us note we can verify a posteriori the vectorial gauge group $G_{\text{SM}} \times SO(10)_{3+4}$ remains unbroken. All the physical NGBs charged under it have large positive masses squared and, at the renormalizable level, vanishing vacuum expectation values [26,27]. The (misaligning) effect induced by unavoidable higher-dimensional operators is suppressed by at least powers of $f^2 / f_{\text{GUT}}^2 \ll g_{\text{unbroken}}^2 / 16\pi^2$ and hence is much smaller than the (aligning) contributions due to the gauge couplings.

3.2 UV contributions to the axion potential

To relax $\bar{\theta}$ to zero, and thus solve the Strong CP Problem, we have to make sure that the NGB associated to the breaking of $U(1)_{\text{PQ}}$, namely the QCD axion a , has a potential dominated by non-perturbative QCD effects. The UV correction ΔV of eq. (2) originates from local operators that explicitly break

³ The additional factor of 2 in this expression comes from our conventional factor of 2 in Σ .

the chiral symmetry $U(1)_{\text{PQ}}$. In these models such effects are severely constrained by the local $SO(10)$'s.

The most minimal gauge singlets built out of a $\mathbf{16} \in SO(10)$ are either $\mathbf{16} \otimes \mathbf{16} \otimes \mathbf{16} \otimes \mathbf{16}$ or $\mathbf{16} \otimes \overline{\mathbf{16}}$. With the particle content of Table 2 each fermion is charged under a single $SO(10)$, so the combination $\mathbf{16} \otimes \overline{\mathbf{16}}$ is automatically invariant under $U(1)_{\text{PQ}}$. We are thus left with the former structure, which by construction cannot be $SU(N_c)$ invariant, see (11). To form a complete gauge and Lorentz singlet, four powers of $\psi, \tilde{\psi}$ should be accompanied by at least four powers of $\psi^c, \tilde{\psi}^c$, and vice-versa. It follows that the most relevant operators that violate $U(1)_{\text{PQ}}$ are of the type

$$(\psi \psi^c)^4, \quad (\tilde{\psi} \tilde{\psi}^c)^4, \tag{18}$$

and thus belong to the class in Eq. (8) with dimension $d = 12$. This is a qualitative improvement compared to the model of Sect. 2. We can next employ the general expression (10) to estimate the observable vacuum angle in this model. One can check that the operator of type (18) with the largest $n_{\mathcal{O}q_a}$ is $(\tilde{\psi} \tilde{\psi}^c)^4$, which contributes with $n_{\mathcal{O}} \sim 16^4 = 65\,536$ (!) and $|q_a| = 1$, see Table 3. Plugging these values in (10), together with $A_a^G = N_c$ from (17) and $|c_{\text{PI}}| = 1$ (the RG effect again leads to an enhancement of a factor of 10^2), the experimental constraint (1) becomes $f_a \lesssim (3 - 4) \times 10^9$ GeV, with N_c varying in the range (11). Had we not included the large factor of $n_{\mathcal{O}} \sim 16^4$, nor the enhancement $(A_a^G)^{d-1} (3/N_c)^{d/6} \sim N_c^9$, nor the RG enhancement, we would have approached the optimistic $f_a \lesssim 10^{11}$ GeV mentioned in the introduction. Unfortunately such enhancements are there.

We can finally compare the above ‘‘stability’’ bound on f_a to the lower bound from experimental searches. The low energy phenomenology of our models is in the KSVZ class [28,29]. The effective coupling of the axion to photons, after the mixing with π_0 is taken into account, is given by $g_{a\gamma\gamma} = \alpha_{\text{em}}/(2\pi)[\mathcal{A}_a^Y/\mathcal{A}_a^G - 1.92(4)]/f_a$, where $\mathcal{A}_a^Y \equiv \mathcal{A}_a^W + \mathcal{A}_a^Z = \mathcal{A}_a^G(8/3)$ [30]. The coupling to neutrons and protons is $g_{ann} = -0.02(3)m_n/f_a$ and $g_{app} = -0.47(3)m_p/f_a$ (see, e.g., [31]). In these scenarios the astrophysics bounds are currently dominated by the supernova constraint on the proton coupling (see e.g. [32]) and read $f_a \gtrsim 1.3 \times 10^8$ GeV. We conclude that these scenarios have an allowed region of parameter space of one to two orders of magnitude in which the Strong CP Problem is robustly solved.

3.3 Alternative realizations

One can obtain alternative models adding some or removing one of the weakly gauged $SO(10)$ groups. Consider for instance removing one of the gauge factors in Table 2, say $SO(10)_4$. The resulting scenario has a smaller $SU(N_c) \times$

$SO(10)^3$ gauge and the axion constituents are:

$$\begin{aligned} \psi &\sim (\mathbf{N}_c, \mathbf{16}, \mathbf{1}, \mathbf{1}) \\ \tilde{\psi} &\sim (\mathbf{N}_c, \mathbf{1}, \mathbf{1}, \mathbf{16}) \\ \psi^c &\sim (\overline{\mathbf{N}}_c, \mathbf{1}, \overline{\mathbf{16}}, \mathbf{1}) \\ \tilde{\psi}^c &\sim 16(\overline{\mathbf{N}}_c, \mathbf{1}, \mathbf{1}, \mathbf{1}). \end{aligned} \tag{19}$$

The physics of the QCD axion is virtually unchanged compared to that of Sect. 3.2. In particular, also here the leading operators violating $U(1)_{\text{PQ}}$ have dimension 12.

The main difference is that now no additional gauge symmetry survives after chiral symmetry breaking besides the SM, i.e. $SO(10)_3$ is completely broken. Furthermore, as opposed to the model of Sect. 3 where only four neutral NGBs arise, here many more are present. These reside in the SM-singlets in $(\mathbf{16}, \overline{\mathbf{16}}) \oplus (\overline{\mathbf{16}}, \mathbf{16})$ and of course in all of the $(\mathbf{1}, \mathbf{210})$ of (15). Interestingly, among these are potential dark matter candidates, as discussed in Sect. 4.3.2.

4 Signatures

Our models predict a number of generic signatures: an approximate gauge-coupling unification, heavy unstable ALPs, and dark matter candidates. The first two are essentially the same in the models of Sects. 2 and 3 and will be studied in Sects. 4.1 and 4.2. Which stable relic can play the role of dark matter depends on the model, though. In the scenario of Sect. 3 natural candidates are the lightest glueball of the unbroken $SO(10)_{3+4}$ and/or heavy SM-neutral NGBs; in scenarios of the type of Sect. 3.3 dark matter could be in the form of a neutral GeV-scale NGB and/or a heavy hadron again. We will discuss these in Sects. 4.3.1 and 4.3.2 respectively.

4.1 Gauge-coupling unification at f_a ?

A distinctive feature of our models is that the axion constituents are chiral under the weakly-coupled gauge groups. The SM gauge symmetry observed at low energy, G_{SM} , is always in the unbroken vectorial subgroup of the chiral symmetry. An approximate unification of the SM gauge couplings $g_G, g_W, g_{Y'} \equiv \sqrt{5/3}g_Y$ at the axion scale follows as a generic prediction. Let us discuss this feature in the context of the $SU(N_c) \times SO(10)^4$ model.

Below the GUT scale the three $G_{\text{SM},1}$ couplings $g_{G,1}, g_{W,1}$, and $g_{Y',1} \equiv \sqrt{5/3}g_{Y,1}$ start to depart from each other. When we reach $\mu \sim \mu_{\text{gauge}} \equiv g_{\text{GUT}}f/\sqrt{N_c} = g_{\text{GUT}}\sqrt{N_c}f_a$ we find that the observed SM gauge symmetry is in the vectorial subgroup of $G_{\text{SM},1} \times SO(10)_2$. Equation (13) induces

a new threshold, which at tree-level (13) reads

$$\frac{1}{g_i^2} = \frac{1}{g_{i,1}^2} + \frac{1}{g_2^2}, \quad i = G, W, Y', \tag{20}$$

with g_2 the $SO(10)_2$ and all couplings are renormalized at $\mu \sim \mu_{\text{gauge}}$. Now, for generic values $g_2^2 \gtrsim g_{i,1}^2$ the threshold correction (20) is very mild, if relevant at all. In the complementary regime $g_2^2 \lesssim g_{i,1}^2$ (for earlier speculations on this possibility see [33])⁴ the linear combination of vector fields that describes the low energy SM gauge bosons is mostly oriented along the $SO(10)_2$ vectors and the couplings observed at low energy thus have a comparable value $g_i^2 = g_2^2[1 + \mathcal{O}(g_2^2/g_{i,1}^2)]$. This results in an approximate gauge coupling unification at scales $\mu \sim \mu_{\text{gauge}} \ll g_{\text{GUT}}f_{\text{GUT}}$ parametrically smaller than those at which true unification of quarks and leptons actually occurs. Gauge coupling unification is only *approximate* because an accurate one would require $g_2^2 \ll g_{i,1}^2$, which can only be realized if the couplings $g_{i,1}^2$ are non-perturbative at threshold.

Physics in the approximately unified regime $g_2^2 \lesssim g_{i,1}^2$ is quite interesting. The gauge couplings almost unify at the axion threshold, and yet proton decay is not induced because the $SO(10)_2$ gauge bosons do not carry baryon nor lepton charges. One has to probe physics at the GUT scale in order to discover that the SM fermions fill complete $SO(10)_1$ GUT representations. We have thus separated gauge-coupling unification from proton decay. As an extreme limit, one can even push f_{GUT} to the Planck scale, thus suppressing proton decay beyond any foreseeable experimental reach, while preserving an approximate gauge-coupling unification at relatively low scales [33]. Looking from bottom-up we find that approximate gauge coupling unification at the axion scale leads to a clear low-energy prediction: there must exist new light particles charged under the SM that make the gauge couplings g_i unify at much lower scales than in usual GUTs. It is not hard to find relics of grand-unified multiplets that can serve our purpose. For example, four Dirac electroweak doublets with hypercharge $\pm 1/2$ plus two Dirac singlets with hypercharge ± 1 , all with masses of order the TeV, would be enough to reach SM gauge coupling unification at 10^{11} GeV.

4.2 Unstable ALPs at the TeV

Our models feature a unique model-independent axial symmetry at the renormalizable level, namely $U(1)_{\text{PQ}}$. Yet, below the GUT scale an additional approximate one arises. For definiteness we focus on the model of Sect. 3, where the extra broken generators are those of $U(1)_{A1} \times U(1)_{A2} \times U(1)_{A3}$. Completely analogous considerations apply to the scenario

⁴ Our gauge coupling unification has also something in common with the string-inspired work [34] and the “fake GUT” scenario of [35], though important differences exist.

of Sect. 2. We are thus interested in the physics of the axion-like particles (ALPs) $a_{\text{ALP}} = \{a_1, a_2, a_3\}$, which we recall are the SM-singlets in (210, 1) of (15).

The key observation is that $U(1)_{A1} \times U(1)_{A2} \times U(1)_{A3}$ is violated by the gauge couplings. Indeed, as opposed to T_{PQ} , the corresponding generators do not commute with the generators of the heavy vectors, $T_L \in SO(10)_1/G_{\text{SM},1}$. This latter fact explains why that symmetry did not appear above the GUT scale in Table 2. The tree exchange of the heavy vectors therefore leads to dimension-6 interactions of the type $(\psi^\dagger T_L \psi)^2/f_{\text{GUT}}^2$ that break $U(1)_{A1} \times U(1)_{A2} \times U(1)_{A3}$ completely. As we emphasized in Sect. 2, this is not sufficient to generate a potential for a_{ALP} , though, because such interactions only break $SU(N_f)_L$. The necessary breaking of $SU(N_f)_R$ comes from the gauging of $T_R \in SO(10)_{3,4}$. Combining the dimension-6 operators $(\psi^\dagger T_L \psi)^2/f_{\text{GUT}}^2$ with a loop of $SO(10)_{3,4}$ vectors we get a contribution to the potential of a_{ALP} , which technically can be parametrized by chiral invariants of the form $\text{Tr}[T_L^J \Sigma T_R^I \Sigma^\dagger T_L^J \Sigma T_R^I \Sigma^\dagger]$. This translates into a mass matrix (including mass mixing terms) of the order

$$m_{\text{ALP}}^2 \sim \frac{g^2}{N_c} \frac{f^4}{f_{\text{GUT}}^2} = (2.5 \text{ TeV})^2 \left(\frac{g}{0.6}\right)^2 \left(\frac{N_c}{12}\right)^3 \times \left(\frac{f_a}{10^9 \text{ GeV}}\right)^4 \left(\frac{10^{16} \text{ GeV}}{f_{\text{GUT}}}\right)^4. \tag{21}$$

Nothing like this can be written down for a , since $U(1)_{\text{PQ}}$ is not broken by the gauge interactions.

One can readily see from Table 3 that these extra NGBs have anomalous couplings $\propto \mathcal{A}_{\text{ALP}}^{\text{SM}}$ to the SM gauge fields (B_μ, W_μ, G_μ), see above (16) for an explicit expression of $\mathcal{A}_{\text{ALP}}^{\text{SM}}$, and are therefore unstable. Up to unknown mixing angles, their lifetimes are

$$\frac{1}{\tau_{\text{ALP}}} = \left(\frac{\mathcal{A}_{\text{ALP}}^{\text{SM}}}{\mathcal{A}_a^G}\right)^2 \left(\frac{g_{\text{SM}}^2}{64\pi^2 f_a}\right)^2 \frac{m_{\text{ALP}}^3}{\pi} n_{\text{SM}}, \tag{22}$$

with $n_{\text{SM}} = 1, 3, 8$ for $\text{SM} = Y, W, G$ a multiplicity factor. Given typical values of the parameters the decay safely takes place at temperatures $T \gtrsim 10^{1-2}$ GeV. Hence, the heavy partners of the QCD axion disappear well before nucleosynthesis and lead to no obvious signatures.

4.3 Long-lived relics

A generic problem of axion models is the production of topological defects via the Kibble–Zurek mechanism. To avoid conflicts with cosmology we assume the highest temperature reached by the Universe, T_{max} , never exceeded the critical temperature of the $SU(N_c)$ dynamics, expected to be of order $\Lambda \equiv 4\pi f/\sqrt{N_c} = 4\pi f_a \sqrt{N_c}$ by analogy with large N

QCD:

$$T_{\max} \lesssim 4 \times 10^9 \text{ GeV} \left(\frac{N_c}{12}\right)^{1/2} \left(\frac{f_a}{10^9 \text{ GeV}}\right). \tag{23}$$

Under this condition the $SU(N_c)$ dynamics underwent its phase transition during or before inflation (we will see below that also $H_I \ll 4\pi f_a$ holds) and topological defects were subsequently inflated away.

Composite axion models also predict a whole tower of heavy hadrons. The lightest baryons have mass $\sim N_c \Lambda$ and are cosmologically stable. Because of (23) their abundance is completely negligible throughout the thermal history of the Universe. Heavy mesons with masses of order $\sim \Lambda$ quickly decay into NGBs as in ordinary QCD. The physics of the NGBs is more model-dependent.

Heavy NGBs

In the scenario of Sect. 3 all NGBs besides the QCD axion and the three unstable singlets of Sect. 4.2 are heavy. They decay into SM fermions, gauge bosons, or into fermions and other NGBs.⁵ The lightest NGB carrying $SO(10)_{3+4}$ charge, we call it Π_s , is stable. We expect it to be the SM-neutral component of $(\mathbf{16}, \mathbf{16}) \oplus (\overline{\mathbf{16}}, \mathbf{16})$ in the third line of (15), say $\Pi_s \sim \psi_N \tilde{\psi}^c$, since the masses of all its charged partners receive additional positive contributions from the SM gauge loops. A robust way to ensure its abundance does not exceed the density of the observed dark matter, in conflict with the well established hot big-bang theory, is to postulate that the reheating temperature $T_{RH} \leq T_{\max}$ is sufficiently smaller than its mass. This constraint roughly reads $T_{RH} < m_{\Pi_s}/30$, where $m_{\Pi_s}^2 \sim (g_{\text{unbroken}}^2/16\pi^2)C_2\Lambda^2$. Hence we demand

$$T_{RH} \lesssim 10^8 \text{ GeV} \left(\frac{g_{\text{unbroken}}}{0.5}\right) \left(\frac{N_c}{12}\right)^{1/2} \left(\frac{f_a}{10^9 \text{ GeV}}\right), \tag{24}$$

with g_{unbroken} the coupling of $SO(10)_{3+4}$ and $C_2 = 45/8$ the quadratic Casimir of the spinorial of $SO(10)_{3+4}$, the representation carried by Π_s . With (24) satisfied, the model of Sect. 3 has a negligible abundance of $SU(N_c)$ hadrons. Potential cosmological signatures may come from the $SO(10)_{3+4}$ gauge bosons, as discussed in Sect. 4.3.1.

The situation is a bit different in the scenario of Sect. 3.3. In that case the NGBs in the $(\mathbf{210}, \mathbf{1})$ of (15) again decay into SM fermions, and those in the third line of (15) again decay into the lightest of them, i.e. Π_s . However, Π_s and all the states in $(\mathbf{1}, \mathbf{210})$ are now massless at tree-level, and couple very weakly to the bath via higher-dimensional operators.

⁵ For a heavy fermion singlet N the decay of some of the SM-charged NGBs may take place between BBN and matter domination and destroy the primordial abundance of ${}^4\text{He}$ (see e.g. [36]). To avoid affecting appreciably physics at BBN we assume at least one of the singlets has $m_N \lesssim 10^{11}$ GeV.

Still, Π_s may be abundantly produced via the decay of its heavy partners. To avoid this we impose the same condition as (24), with g_{unbroken} replaced by the SM couplings. As a consequence of this assumption, also in the scenario of Sect. 3.3 the present-day population of heavy hadrons is negligible. The physics of the light, decoupled singlets Π_s and $(\mathbf{1}, \mathbf{210})$ will be analyzed in Sect. 4.3.2.

Axion density

Equation (23) implies the axion abundance is set by vacuum misalignment. Assuming for definiteness that the temperature-dependence of the axion mass is approximately given by $m_a \propto T^{-\delta}$ for temperatures above the critical one, the fraction of dark matter in the form of cold QCD axions turns out to scale as

$$\frac{\rho_a}{\rho_{\text{DM}}} \sim \theta_{\text{init},a}^2 \left(\frac{f_a}{10^{12} \text{ GeV}}\right)^{\frac{\delta+3}{\delta+2}}, \tag{25}$$

where $\theta_{\text{init},a}$ is the initial misalignment angle. The result (25) depends uniquely on the initial condition and the (temperature-dependent) QCD potential, and is therefore governed by $f_a = f/N_c$ rather than f , see (17). The main unknown are $\theta_{\text{init},a}$ and δ . Taking as a reference the result of a dilute instanton gas approximation, where $\delta \sim 4$, and the natural value $\theta_{\text{init},a} \lesssim 1$ we see that for the relatively low values $f_a \sim 10^{8-10}$ GeV considered here the axion constitutes at most a small fraction of the observed dark matter. With f_a fixed, the abundance decreases as δ decreases.

During inflation the axion field undergoes quantum fluctuations of typical size of the Hubble scale H_I in that epoch. The resulting isocurvature perturbations are constrained by cosmological observations. This latter constraint can be expressed as an upper bound on H_I [37]: $H_I \lesssim 10^7 \text{ GeV} (f_a/10^{11} \text{ GeV})^{(1+\delta)/(4+2\delta)} \sqrt{\rho_{\text{DM}}/\rho_a}$, which gets weaker as the QCD axion abundance decreases. Combining with (25) we obtain

$$H_I \lesssim \frac{10^8}{\theta_{\text{init},a}} \text{ GeV} \left(\frac{10^9 \text{ GeV}}{f_a}\right)^{\frac{1}{2+\delta}}. \tag{26}$$

This bound applies to all pre-inflationary QCD axion scenarios. In our case it is a bit relaxed compared to what is often quoted because the axion is likely a small component of the dark matter.

4.3.1 Glueball dark matter

The model of Sect. 3 predicts an unbroken hidden non-abelian gauge group, i.e. $SO(10)_{3+4}$. Below the scale $\Lambda = 4\pi f_a \sqrt{N_c}$ the associated vectors are coupled to the SM very weakly, mainly via the ordinary gluons via higher dimen-

sional operators of the form⁶

$$c \frac{N_c g_3^2 g_G^2}{16\pi^2 \Lambda^4} G_{\mu\nu} G^{\mu\nu} G'_{\rho\sigma} G'^{\rho\sigma}, \tag{27}$$

and similarly for the electroweak bosons. The interaction is so irrelevant that they never equilibrate with the plasma. The hidden gluons form a gas of self-interacting particles living with its own entropy $s'(T')$ at its own temperature T' , which differs from the temperature T of radiation. At the end of inflation the hidden gluons are populated dominantly by the annihilation of SM gluons according to $dn'/dt + 3Hn' = \Gamma$, with H the Hubble scale and Γ the Boltzmann collision term. (By the assumption (24) the amount of vectors produced by the annihilation of the hadrons at the scale $\sim \Lambda$ is instead completely negligible.) After being produced, the hidden gas thermalizes right away with entropy $s'(T') = (2\pi^4/45\zeta_3)n'(T')$.

The non-renormalizable nature of (27) implies $\Gamma \propto T^{12}$, so that the production is mostly active at the highest accessible temperatures $T \sim T_{RH}$. The entropy of the hidden gluons over the radiation entropy may be estimated to be of order

$$\xi \equiv \frac{s'}{s} \sim \frac{2\pi^4}{45\zeta_3} \frac{\Gamma}{sH} \Big|_{T=T_{RH}} \ll 1, \tag{28}$$

where we find

$$\Gamma \sim \frac{288}{\pi^5} n_G \left(\frac{8\pi^2 T^6}{63} \right)^2 \left(c \frac{N_c g_{Hid}^2 g_G^2}{16\pi^2 \Lambda^4} \right)^2. \tag{29}$$

The quantity ξ then remains constant at all $T \ll T_{RH}$. At later times the $SO(10)_{3+4}$ force confines and the entropy carried by the relativistic hidden gluons gets re-distributed among non-relativistic glueballs. The heavy $SO(10)_{3+4}$ glueballs quickly decay into the lightest, Φ , of mass m_Φ . A simple estimate shows that Φ has a lifetime longer than the age of the Universe for $m_\Phi \lesssim 5 \times 10^5$ GeV, $c \sim 1$, N_c in the range (11), and $f_a = 10^{8-10}$ GeV.

The present-day abundance of Φ can be calculated as follows (see Ref. [38]). After confinement the new gas of non-relativistic Φ 's remains self-interacting for a while until the number-changing processes go out of equilibrium at $T'_d \lesssim m_\Phi$. We can relate the temperatures of the hidden gas to that of ordinary radiation via $s_\Phi(T'_d) = s'(T'_d) = \xi s(T_d)$. For the range of parameters we are interested in, T_d is always many orders of magnitude above the keV and the energy density ρ_Φ of the glueballs at decoupling is much smaller than

that of radiation. For any $T < T_d$ the non-interacting glueball gas behaves as ordinary cold relics: $\rho_\Phi/s = \xi \rho_\Phi/s'$ stays constant and progressively becomes more and more relevant than radiation. The present-day glueball energy density is of order

$$\frac{\rho_\Phi(T_0)}{s(T_0)} = \xi \frac{\rho_\Phi(T'_d)}{s'(T'_d)} \sim \xi m_\Phi. \tag{30}$$

Using (29) and saturating the upper bound (24) the abundance of the stable glueballs reaches a value comparable to the observed dark matter. This conclusion is however extremely sensitive to T_{RH} because of the high temperature-dependence in Γ . A dedicated study of the reheating process would be needed to clarify if Φ can be a realistic dark matter candidate in a regime in which the heavy NGBs have a negligible relic density.

4.3.2 NGB dark matter

In the scenario of Sect. 3.3 there is no hidden gauge symmetry but there is a whole host of light SM-neutral NGBs, as we have seen at the end of Sect. 4.3. In particular, Π_s and those in the representation (1, 210) of (15) are SM singlets, massless at the renormalizable level, and have no anomalous couplings to the SM gauge fields. Some of these are exactly stable in the renormalizable Lagrangian, namely those that are protected by an ‘‘isospin’’ symmetry under which the $SO(10)_3$ gauge bosons transform in the adjoint.

A potential is generated by a combination of dimension-6 operators $((\tilde{\psi}^c)^\dagger \tilde{\sigma}^\mu \tilde{\psi}^c)^2/f_{\text{Pl}}^2$, and loops of the massive $SO(10)_3$ gauge bosons. We can recycle (21) to get the typical parametric dependence of the masses:

$$\begin{aligned} m_{\text{NGB}}^2 &= c_{\text{NGB}} \frac{g^2}{N_c} \frac{f^4}{f_{\text{Pl}}^2} \\ &= (10 \text{ GeV})^2 \left(\frac{c_{\text{NGB}}}{1} \right) \left(\frac{g}{0.6} \right)^2 \left(\frac{N_c}{12} \right)^3 \left(\frac{f_a}{10^9 \text{ GeV}} \right)^4. \end{aligned} \tag{31}$$

A possible decay into SM fermions and gauge bosons requires breaking the ‘‘isospin’’, triggered by operators suppressed by the Planck scale. Yet, m_{NGB} is so small compared to $4\pi f_{\text{Pl}}$ that the resulting lifetime is much longer than the age of the Universe. These NGBs are therefore cosmologically stable and may contribute to the dark matter.

Because of (23) and (24) the stable GeV-scale NGBs have no thermal population and their energy density is set by vacuum misalignment. The standard argument goes as follows. The random initial energy density stayed approximately constant for large T . As the temperature dropped below $\sim T_{\text{osc}}$, the latter being defined by the condition $3H(T_{\text{osc}}) = m_{\text{NGB}}$, the NGB zero mode started to oscillate and its energy density to decrease as ordinary matter. For typ-

⁶ Dimension-6 operators involving the Higgs doublet are also generated at higher loop order. Despite the lower dimensionality, however, the UV diagrams responsible for inducing them are highly convergent. As a consequence their Wilson coefficients are suppressed by the weak scale, i.e. $m_{\text{weak}}^2/\Lambda^2$, and their effect turn out to be subleading compared to (27) in the interesting region of parameter space where $m_\Phi \gg m_{\text{weak}}$. We thank R. Contino for emphasizing this to us.

ical parameters of our models it turns out that $T_{\text{osc}} \gg T_{\text{RH}}$. This implies that the NGB field started to oscillate during inflation. Simplifying our calculation assuming a sudden transition at T_{RH} , the energy density at later times is given by $\rho_{\text{NGB}} = \theta_{\text{init}}^2 m_{\text{NGB}}^2 f^2 (T/T_{\text{RH}})^3 / 2$. Hence the present-day value is

$$\frac{\rho_{\text{NGB}}(T_0)}{\rho_{\text{DM}}(T_0)} \simeq \left(\frac{\theta_{\text{init}}}{10^{-3}} \right)^2 \left(\frac{c_{\text{NGB}}}{1} \right) \left(\frac{g}{0.6} \right)^2 \left(\frac{N_c}{12} \right)^5 \times \left(\frac{f_a}{10^9 \text{ GeV}} \right)^6 \left(\frac{10^8 \text{ GeV}}{T_{\text{RH}}} \right)^3, \quad (32)$$

where the mass was taken from (31). The bottom line is that the GeV-scale stable NGBs carry too much energy unless their initial displacement is somewhat small. This is reminiscent of the large f_a regime of the ordinary QCD axion, though here f_a is so small that the energy of the QCD axion field is completely negligible. This conclusion is rather conservative, though, because the NGB population could well be depleted significantly during the temperature range $T_{\text{RH}} \leq T \leq T_{\text{max}}$. In such a case even $\theta_{\text{init}} \sim 1$ could be acceptable.

5 Conclusions

In building composite axion models one knows from the start that the symmetry breaking scale f_a must be very large. Physics at such energies might be more appropriately described in terms of complete multiplets of some GUT. It is therefore only natural to ask whether one can build realistic, asymptotically-free scenarios based on popular grand-unified groups like $SU(5)$ and $SO(10)$. We found that the answer is in the affirmative.

Both $SU(5)$ and $SO(10)$ have chiral representations, the basic ingredient of a working composite axion model. We then looked for scenarios with minimal field content, chiral under $SU(5)$ or $SO(10)$. To avoid a loss of asymptotic freedom the axion constituents were taken to be charged under at most two gauge groups. The existence of an anomalous $U(1)_{\text{PQ}}$ of high quality imposed further constraints on the type and number of representations. The simplest model satisfying these requests is based on the $\mathbf{10} \oplus \bar{\mathbf{5}}$ chiral multiplet of $SU(5)$. Unfortunately, despite it being quite minimal and elegant, such a scenario allows for UV contributions to the axion potential from operators of dimension 9, which may be too little to obtain a realistic axion model.

An axion symmetry with higher quality is obtained exploiting the spinorial $\mathbf{16}$ of $SO(10)$. In models of this type all $SO(10)$ singlets with $U(1)_{\text{PQ}}$ charge contain at least four powers of the axion constituents. As a result the minimal Lagrangian operator that can contribute to the axion potential has dimension 12. We analyzed two models of this class, but many others can be built along these lines.

A rough estimate naively implies that Planck-scale suppressed operators of dimension 12 could be compatible with $f_a \sim 10^{11}$ GeV. However, composite axion models tend to have large color and flavor multiplicities – a must if we want to embed our scenarios in GUTs in which all gauge groups are asymptotically free – as well as renormalization group effects which go in the direction of enhancing the sensitivity to UV contributions to the potential. Employing naive dimensional analysis, the best tool the effective field theorist has at her or his disposal to estimate the Wilson coefficient of the Planck-scale suppressed operator, we find that our $SO(10)$ models are left with a window of about one to two orders of magnitude $10^8 \text{ GeV} \lesssim f_a \lesssim 10^{9-10} \text{ GeV}$ in which the Strong CP Problem is robustly solved consistently with all experimental data. Our conservative estimate seems to us more justified than the naive one.

Interestingly, the very structure of our scenarios implies a number of very distinctive predictions at low energies. The *chiral embedding* of the axion constituents with respect to the weakly coupled gauge groups, a characteristic feature of our constructions, implies an approximate gauge coupling unification of (all or just two of) the SM gauge couplings generically occurs at around the axion scale. From a bottom-up perspective this requires new light particles, potentially accessible to colliders, that make the SM gauge couplings unify at 10^{10-11} GeV. Above that scale one discovers that the SM gauge fields are mostly the vectors of a grand-unified group, and yet the SM fermions do not fill complete GUT multiplets, so baryon number violating operators do not appear at that threshold. Rather, the actual lepton-quark unification occurs at much larger scales, as usual.

Another generic prediction of our scenarios is the existence of several light SM-neutral bosonic states beyond the QCD axion. Some of these can be potential dark matter candidates precisely in the regime $f_a \sim 10^{8-10}$ GeV expected in our models, i.e. when the QCD axion typically contributes negligibly. In the models of Sect. 3 these dark matter candidates are in the form of long-lived glueballs of an unbroken hidden gauge sector populated at re-heating, whereas in the models of Sect. 3.3 they correspond to GeV-scale NGBs with a cold abundance set by vacuum misalignment. In addition we always find unstable ALPs with anomalous couplings to the SM gauge bosons and masses of order a few TeV generated at the GUT scale.

In one way or another, all field theoretic explanations of the axion quality problem look a bit ad hoc. Our $SO(10)$ models are no exception. Yet, taking into account the good quality of their $U(1)_{\text{PQ}}$, the natural predisposition to accommodate grand-unified scenarios, and their distinctive phenomenological signatures, we think that our models do not compare too bad with the competition either. Perhaps the ideas developed in this paper will find a more compelling incarnation in the future.

Acknowledgements We would like to thank F. D'Eramo for comments on the cosmological signatures of our models, L. Di Luzio and R. Ziegler for discussions on $SO(10)$ and grand-unification, L. Martucci for conversations on axions in string theory, M. Redi for sharing some of his expertise in composite axion models, I. Shoemaker for suggestions, and especially R. Contino, A. Podo, and F. Revello for comments on the draft and discussions about their independent forthcoming paper on axions. This research was partly supported by the Italian MIUR under contract 2017FMJFMW (PRIN2017).

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors' comment: The results of this study are analytical and summarised by the formulas presented in the text.]

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. Funded by SCOAP³.

References

1. H. Georgi, I.N. McArthur, HUTP-81/A011
2. K. Choi, C.W. Kim, W.K. Sze, Phys. Rev. Lett. **61**, 794 (1988). <https://doi.org/10.1103/PhysRevLett.61.794>
3. T. Banks, Y. Nir, N. Seiberg, [arXiv:hep-ph/9403203](https://arxiv.org/abs/hep-ph/9403203)
4. S. Aoki, Y. Aoki, C. Bernard, T. Blum, G. Colangelo, M. Della Morte, S. Dürr, A.X. El Khadra, H. Fukaya, R. Horsley et al., Eur. Phys. J. C **74**, 2890 (2014). <https://doi.org/10.1140/epjc/s10052-014-2890-7>. [arXiv:1310.8555](https://arxiv.org/abs/1310.8555) [hep-lat]
5. C. Alexandrou, J. Finkenrath, L. Funcke, K. Jansen, B. Kostrzewa, F. Pittler, C. Urbach, Phys. Rev. Lett. **125**(23), 232001 (2020). <https://doi.org/10.1103/PhysRevLett.125.232001>. [arXiv:2002.07802](https://arxiv.org/abs/2002.07802) [hep-lat]
6. A. Hook, Phys. Rev. Lett. **114**(14), 141801 (2015). <https://doi.org/10.1103/PhysRevLett.114.141801>. [arXiv:1411.3325](https://arxiv.org/abs/1411.3325) [hep-ph]
7. R.D. Peccei, H.R. Quinn, Phys. Rev. Lett. **38**, 1440–1443 (1977). <https://doi.org/10.1103/PhysRevLett.38.1440>
8. S. Weinberg, Phys. Rev. Lett. **40**, 223–226 (1978). <https://doi.org/10.1103/PhysRevLett.40.223>
9. F. Wilczek, Phys. Rev. Lett. **40**, 279–282 (1978). <https://doi.org/10.1103/PhysRevLett.40.279>
10. R. Holman, S.D.H. Hsu, T.W. Kephart, E.W. Kolb, R. Watkins, L.M. Widrow, Phys. Lett. B **282**, 132–136 (1992). [https://doi.org/10.1016/0370-2693\(92\)90491-L](https://doi.org/10.1016/0370-2693(92)90491-L). [arXiv:hep-ph/9203206](https://arxiv.org/abs/hep-ph/9203206)
11. S.M. Barr, D. Seckel, Phys. Rev. D **46**, 539–549 (1992). <https://doi.org/10.1103/PhysRevD.46.539>
12. M. Kamionkowski, J. March-Russell, Phys. Lett. B **282**, 137–141 (1992). [https://doi.org/10.1016/0370-2693\(92\)90492-M](https://doi.org/10.1016/0370-2693(92)90492-M). [arXiv:hep-th/9202003](https://arxiv.org/abs/hep-th/9202003)
13. J.E. Kim, Phys. Rev. D **31**, 1733 (1985). <https://doi.org/10.1103/PhysRevD.31.1733>
14. L. Randall, Phys. Lett. B **284**, 77–80 (1992). [https://doi.org/10.1016/0370-2693\(92\)91928-3](https://doi.org/10.1016/0370-2693(92)91928-3)
15. B.A. Dobrescu, Phys. Rev. D **55**, 5826–5833 (1997). <https://doi.org/10.1103/PhysRevD.55.5826>. [arXiv:hep-ph/9609221](https://arxiv.org/abs/hep-ph/9609221)
16. H. Fukuda, M. Ibe, M. Suzuki, T.T. Yanagida, Phys. Lett. B **771**, 327–331 (2017). <https://doi.org/10.1016/j.physletb.2017.05.071>. [arXiv:1703.01112](https://arxiv.org/abs/1703.01112) [hep-ph]
17. M. Redi, R. Sato, JHEP **05**, 104 (2016). [https://doi.org/10.1007/JHEP05\(2016\)104](https://doi.org/10.1007/JHEP05(2016)104). [arXiv:1602.05427](https://arxiv.org/abs/1602.05427) [hep-ph]
18. H.C. Cheng, D.E. Kaplan, [arXiv:hep-ph/0103346](https://arxiv.org/abs/hep-ph/0103346)
19. T. Appelquist, J. Terning, L.C.R. Wijewardhana, Phys. Rev. Lett. **77**, 1214–1217 (1996). <https://doi.org/10.1103/PhysRevLett.77.1214>. [arXiv:hep-ph/9602385](https://arxiv.org/abs/hep-ph/9602385)
20. T. Appelquist, A.G. Cohen, M. Schmaltz, Phys. Rev. D **60**, 045003 (1999). <https://doi.org/10.1103/PhysRevD.60.045003>. [arXiv:hep-th/9901109](https://arxiv.org/abs/hep-th/9901109)
21. E. Poppitz, M. Unsal, JHEP **09**, 050 (2009). <https://doi.org/10.1088/1126-6708/2009/09/050>. [arXiv:0906.5156](https://arxiv.org/abs/0906.5156)
22. E. Poppitz, M. Unsal, JHEP **12**, 011 (2009). <https://doi.org/10.1088/1126-6708/2009/12/011>. [arXiv:0910.1245](https://arxiv.org/abs/0910.1245) [hep-th]
23. T.A. Ryttov, R. Shrock, Phys. Rev. D **97**(2), 025004 (2018). <https://doi.org/10.1103/PhysRevD.97.025004>. [arXiv:1710.06944](https://arxiv.org/abs/1710.06944) [hep-th]
24. E. Gardi, G. Grunberg, JHEP **03**, 024 (1999). <https://doi.org/10.1088/1126-6708/1999/03/024>. [arXiv:hep-th/9810192](https://arxiv.org/abs/hep-th/9810192)
25. M.B. Gavela, M. Ibe, P. Quilez, T.T. Yanagida, Eur. Phys. J. C **79**(6), 542 (2019). <https://doi.org/10.1140/epjc/s10052-019-7046-3>. [arXiv:1812.08174](https://arxiv.org/abs/1812.08174) [hep-ph]
26. M.E. Peskin, Nucl. Phys. B **175**, 197–233 (1980). [https://doi.org/10.1016/0550-3213\(80\)90051-6](https://doi.org/10.1016/0550-3213(80)90051-6)
27. J. Preskill, Nucl. Phys. B **177**, 21–59 (1981). [https://doi.org/10.1016/0550-3213\(81\)90265-0](https://doi.org/10.1016/0550-3213(81)90265-0)
28. J.E. Kim, Phys. Rev. Lett. **43**, 103 (1979). <https://doi.org/10.1103/PhysRevLett.43.103>
29. M.A. Shifman, A.I. Vainshtein, V.I. Zakharov, Nucl. Phys. B **166**, 493–506 (1980). [https://doi.org/10.1016/0550-3213\(80\)90209-6](https://doi.org/10.1016/0550-3213(80)90209-6)
30. M. Srednicki, Nucl. Phys. B **260**, 689–700 (1985). [https://doi.org/10.1016/0550-3213\(85\)90054-9](https://doi.org/10.1016/0550-3213(85)90054-9)
31. G. Grilli di Cortona, E. Hardy, J. Pardo Vega, G. Villadoro, JHEP **01**, 034 (2016). [https://doi.org/10.1007/JHEP01\(2016\)034](https://doi.org/10.1007/JHEP01(2016)034). [arXiv:1511.02867](https://arxiv.org/abs/1511.02867) [hep-ph]
32. L. Di Luzio, M. Giannotti, E. Nardi, L. Visinelli, Phys. Rept. **870**, 1–117 (2020). <https://doi.org/10.1016/j.physrep.2020.06.002>. [arXiv:2003.01100](https://arxiv.org/abs/2003.01100) [hep-ph]
33. N. Weiner, [arXiv:hep-ph/0106097](https://arxiv.org/abs/hep-ph/0106097)
34. R. Barbieri, G.R. Dvali, A. Strumia, Phys. Lett. B **333**, 79–82 (1994). [https://doi.org/10.1016/0370-2693\(94\)91010-3](https://doi.org/10.1016/0370-2693(94)91010-3). [arXiv:hep-ph/9404278](https://arxiv.org/abs/hep-ph/9404278)
35. M. Ibe, S. Shirai, M. Suzuki, T.T. Yanagida, Phys. Rev. D **100**(5), 055024 (2019). <https://doi.org/10.1103/PhysRevD.100.055024>. [arXiv:1906.02977](https://arxiv.org/abs/1906.02977) [hep-ph]
36. M. Kawasaki, K. Kohri, T. Moroi, Y. Takaesu, Phys. Rev. D **97**(2), 023502 (2018). <https://doi.org/10.1103/PhysRevD.97.023502>. [arXiv:1709.01211](https://arxiv.org/abs/1709.01211) [hep-ph]
37. P.A.R. Ade et al. [Planck], Astron. Astrophys. **594**, A20 (2016). <https://doi.org/10.1051/0004-6361/201525898>. [arXiv:1502.02114](https://arxiv.org/abs/1502.02114) [astro-ph.CO]
38. E.D. Carlson, M.E. Machacek, L.J. Hall, Astrophys. J. **398**, 43–52 (1992). <https://doi.org/10.1086/171833>