Crossing into the $m_a > f_a$ region for leptophilic ALPs

Marta F. Zamoro,[*] Álvaro Lozano-Onrubia,[†] Luca Merlo,[‡] and Samuel Rosende Herrero[§]

Departamento de Física Teórica and Instituto de Física Teórica UAM/CSIC, Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain

(Received 12 June 2026; accepted 26 August 2026; published 2 November 2026)

Axion-like particles (ALPs) are typically identified as pseudoscalars whose couplings are shift-symmetry invariant with the exception of their couplings to gauge bosons and their mass term. Additionally, the ALP mass m_a is usually assumed to be (much) smaller than the ALP decay constant f_a. The latter condition is conservative, at best, and excludes part of the ALP parameter space that is presently viable. We revisit the criterion of the $m_a \ll f_a$ and perform an analysis focusing on leptophilic ALPs. In particular, we explore regions of the parameter space still uninvestigated, where $m_a > f_a$, thus providing a phenomenological study of the ALP-lepton couplings complementary to the existing literature. We point out that a leptophilic ALP may explain the -3.8σ tension in the anomalous magnetic dipole moment of the electron for the cesium determination in a large region of the $m_a \times f_a$ parameter space, testable in the near future through studies on $\mu \to e$ conversion in nuclei.

DOI: 10.1103/xf8c-823x

I. INTRODUCTION

Despite its overwhelming success in describing fundamental particles and their interactions, the Standard Model (SM) of particle physics presents a number of striking shortcomings. One of the most prominent is the so called strong CP problem, namely, the absence of CP violating (CPV) sources in the strong sector of the SM, independent of the experimentally established CPV sources in the weak sector, i.e., those encountered in fermion mixing and embedded in the Cabibbo-Kobayashi-Maskawa (CKM) and Pontecorvo-Maki-Nakagawa-Sakata (PMNS) mixing matrices. One of the most accepted solutions to this problem is the QCD axion model [1–7]. In short, the particle spectrum is extended beyond the SM to include new scalar and/or fermionic fields, and the SM symmetry group is supplemented by a global $U(1)_{\rm PQ}$ factor, the so-called Peccei-Quinn (PQ) symmetry, which by construction shall be anomalous with respect to QCD. Once the PQ symmetry is spontaneously broken, a new Nambu-Goldstone boson (NGB) termed the axion arises. This NGB can absorb any additional CPV sources beyond the CKM and PMNS ones, representing a dynamical solution to the strong CP problem. In all existing realizations of this solution, the axion turns out to be a pseudo-NGB (pNGB) due to nonperturbative QCD effects that give rise to an axion mass m_a. This QCD axion mass satisfies a strict inverse proportionality relation with its characteristic scale f_a, which suppresses all its interactions:

$$m_a f_a \approx m_\pi f_\pi, \tag{1.1}$$

where m_π and f_π are the pion mass and its decay constant, respectively [for a formal derivation, see, for instance, Ref. [8]; also, recently in Ref. [9], the axion-pion mixing contribution to the photon coupling has been derived using lattice methods].

The low-energy phenomenology of the different QCD axion models can be described in full generality in terms of an effective field theory (EFT). Following the pioneering article Ref. [10], the axion is described as a pNGB whose interactions respect a shift symmetry with the exception of its couplings to the gauge bosons, known as anomalous terms, and its mass term that follows the relation in Eq. (1.1), in accordance with the chiral perturbation theory (χPT) treatment. The coetaneous idea of a particle with properties closely resembling those of the QCD axion, with the crucial exception of the axion mass relation, had already surfaced in other contexts. Hence, the effective description of Ref. [10] was assimilated elsewhere too. The axion-like particle (ALP) emerged to denominate any particle with interactions described by shift-symmetric couplings, anomalous terms, and a free mass that specifically does not need to satisfy the relation in Eq. (1.1).

[*]Contact author: marta.zamoro@uam.es
[†]Contact author: alvaro.lozano.onrubia@csic.es
[‡]Contact author: luca.merlo@uam.es
[§]Contact author: samuel.rosende@uam.es

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

The literature presents a rich variety of ALPs: they show up in composite Higgs models [11–14], in supersymmetric constructions [15], and in the context of string theory [16–20], among others. ALPs have also been proposed as possible dark matter candidates [21–26] and, more generally, have been thought of as having an impact on cosmological observables [27–38]. Furthermore, they have been considered in flavor model building [39–46] and are known to emerge in mechanisms of neutrino mass generation, where they go by the name of Majorons [47–52]. A recent and complete review that includes models with ALPs is Ref. [53].

In addition to the model building efforts, part of the community has extensively worked to improve and complete the ALP effective description: the literature ranges from new theoretical perspectives [54–65] to more experimentally oriented analyses of potential ALP signals at both colliders [56,66–88] and low-energy facilities [89–122]. A very recent review on this topic can be found in Ref. [123].

Contrary to the common understanding that a generic ALP cannot solve the strong *CP* problem, recent results demonstrate [97,124–133] that the inverse proportionality between the mass m_a and the scale f_a traditionally touted for the QCD axion can be relaxed. It follows that the parameter space accommodating a consistent solution to the strong *CP* problem exceeds that of the very constrained QCD axion band. Specifically, this allows one to consider ALPs as effective descriptions of axions that indeed solve the strong *CP* problem yet whose mass is not described by Eq. (1.1).

In accordance with the traditional QCD axion perspective, most of the ALP literature *a priori* assumes the hierarchy of scales

$$m_a \ll f_a. \tag{1.2}$$

This identification of the characteristic scale f_a as an apparent cutoff scale appears naively reasonable on two fronts. First, ALP EFTs constitute a textbook instance of a bottom-up EFT approach wherein the connection to the actual EFT cutoff scale Λ is not necessarily made manifest. This occurs because the approach embodied by ALP EFTs has f_a, and not the cutoff Λ, setting an apparent scale of validity for the EFT (see Sec. II). Although this may be viewed as a notational rather than a fundamentally physical issue, it is relevant to recall that the relation between f_a and Λ cannot, in general, be established if the EFT approach is truly ultraviolet (UV) agnostic. Since any sensible ALP EFT will require $f_a \leq \Lambda$, the assumption in Eq. (1.2) ensures a physical mass well below the true scale of validity and thus the consistency of the EFT. Furthermore, Eq. (1.2) also happens to be an experimentally justified working hypothesis in contexts where the ALP is assumed light with respect to the Fermi scale. Indeed, the experimental bounds on the actual EFT cutoff scale Λ, resulting from both collider and flavor facilities, turn out to be in the TeV region for $\mathcal{O}(1)$ Wilson coefficients. As we shall later see, f_a is not expected to veer off far below that scale unless an extremely strongly coupled UV theory is assumed.

However, caution is needed when blindly assuming the general relation Eq. (1.2) for two reasons. On the one hand, the scale Λ can be much smaller than traditionally assumed. Indeed, if the Wilson coefficients happened to be substantially suppressed, Λ would approach the scale of a traditionally light ALP mass. On the other hand, and all prejudice aside, an ALP may well be very massive when compared to other scales in the standard ALP EFT description. There is no fundamental reason why m_a should be strictly constrained by the value of f_a before the latter is set in relation to an EFT cutoff Λ. Poignantly, particle physics already showcases a striking example of an EFT featuring pNGBs with masses m_π larger than their associated characteristic decay scale f_π, namely, pions in χPT. Another example is the model proposed in Ref. [134], recently introduced as a solution to the strong *CP* problem featuring a GeV-scale axion. The model is currently under scrutiny by the community [135] and appears to be excluded. Interestingly, the exclusion arises from its incompatibility with experimental data on pion physics, rather than from a theoretical inconsistency.

This observation alone calls for prudence, and, all things considered, a fundamental reappraisal of the working hypothesis $m_a \ll f_a$ prevalent in the literature appears necessary.

In the next sections, *we claim that the direct comparison between m_a and f_a should be taken with a grain of salt and we exemplify the consequences of relaxing the $m_a \ll f_a$ condition in a simplified scenario.* The study focuses on a leptophilic ALP whose dominant phenomenology resides in charged lepton decays and in the magnetic dipole moment of the electron and the muon. While most of the studies of ALPs assume the parametric hierarchy $m_a \ll f_a$, with m_a being either very light or even in the GeV range [56,66–122,136], *we will focus on the $m_a > f_a$ region, that has hitherto received little attention in phenomenological analyses, thus providing results complementary to the existing investigations. As a part of this analysis, we point out the possibility to completely relax the present -3.8σ tension between the cesium determination of the anomalous magnetic dipole moment of the electron and its current experimental measurement within the context of a leptophilic ALP.*

The structure of the paper includes in Sec. II a review of the relations between m_a, f_a, and Λ within the naive dimensional analysis description, followed in Sec. III by a description of the EFT Lagrangian that represents our starting point for the analysis. In Sec. IV, we list the physical processes that will be used in the numerical study, present the explicit dominant contributions to the expression of their observables, and identify specific benchmark textures for

the ALP-lepton couplings. Section V illustrates the results of our numerical analysis, first focusing on the different flavor textures and then on the possible explanation of the $(g-2)_e$ anomaly. The ALP-aca code [120] is the main tool adopted for this analysis, in order to correctly account for the renormalization group equations (RGE) contributions. We then conclude in Sec. VI. Additional material on the RGE analysis can be found in the Appendix, and a *Mathematica* file in Supplemental Material [137] provides the complete expressions for the observables used, including a comparison with the available literature.

II. DIMENSIONAL ANALYSIS AND COUPLING REGIME

In this section, we discuss the $m_a \ll f_a$ relation that is commonly assumed in the ALP effective description. At first sight, comparing m_a with f_a seems legitimate as both quantities have the same mass dimension. However, they fundamentally embody two different and independent concepts: m_a is a mass parameter, while f_a is a decay constant that governs the ALP couplings. This comparison of m_a vs f_a is actually specific of ALP EFTs and not of UV constructions: considering, for example, the traditional QCD axion model, the relation in Eq. (1.1) is well-defined as the same types of quantities appear on both sides of the equation, i.e., a mass and a decay constant. To clarify the fundamental relation between these conceptually different quantities, we need to dive deeper into the details of the EFT construction.

Independently of the specific high-energy dynamics that gives rise to the ALP, the low-energy description only deals with particles lighter than the EFT cutoff Λ which represents the scale at which heavier degrees of freedom have been integrated out. In EFTs, Λ often coincides with the mass of the heavier degrees of freedom,[1] meaning that any good EFT description cannot describe particles with masses larger than Λ. The same general principle also applies to ALP EFTs: a sensible EFT description requires an ALP mass below the cutoff. Thus, $m_a < \Lambda$ is a general consistency condition for the low-energy description to be meaningful. Since both quantities refer to mass parameters, the comparison of m_a with Λ is not only legitimate but also conceptually valid.

We now take a step forward considering the pNGB nature of the ALP, that is, as already stated in the Introduction, treating it as a real, CP-odd scalar field, whose interactions present a shift-symmetry invariance, except for the anomalous couplings and its mass term. This aspect allows us to harness the well-known framework of naive dimensional analysis (NDA) [138–140] that encodes the usual Λ counting rule in EFT combined with the 4π counting rule,[2] altogether providing a systematic way to normalize EFT operators. The 4π counting rule is very useful in the ALP description due to its close relation to the nature of NGBs, as is the case of χPT. We can write the NDA master formula as follows:

$$\frac{\Lambda^4}{g_*^2}\left[\frac{\partial}{\Lambda}\right]^{n_\partial}\left[\frac{g_*\phi}{\Lambda}\right]^{n_\phi}\left[\frac{g_* A}{\Lambda}\right]^{n_A}\left[\frac{g_*\psi}{\Lambda^{3/2}}\right]^{n_\psi} \tag{2.1}$$

where n_i is the number of i-element insertions in any effective operator, with ϕ being any scalar field, A any gauge boson, and ψ any fermion. In particular, g_* is a measure of the strength of the underlying dynamics that gives rise to the low-energy operators: $g_* = 4\pi$ represents the strongly interacting regime, while $g_* = 1$ stands for the weakly interacting one, with the intermediate values interpolating between these two limits and characterizing couplings of moderate strength.

In the case of χPT and focusing on the simpler case of pions, the Lagrangian is written in terms of $U \equiv \exp 2i\boldsymbol{\pi}/f_\pi$, where $\boldsymbol{\pi}$ is the pion matrix and f_π is the pion decay constant. Expanding U in terms of f_π implies that any meson insertion comes with a $1/f_\pi$ factor, i.e., an insertion rule given by $[\boldsymbol{\pi}/f_\pi]^{n_\pi}$. We can thus compare with Eq. (2.1) and identify a relation between Λ and f_π,

$$\Lambda = g_* f_\pi. \tag{2.2}$$

This has already been pointed out in the pioneering article by Georgi and Manohar [138] and beautifully makes the point that the only meaningful scale in the game is the cutoff Λ and not f_π. Focusing, for instance, on pion scattering, the derivative expansion characteristic of χPT is an expansion in powers of ∂/Λ that is valid for small pion momenta up to $\Lambda = 4\pi f_\pi \sim 1$ GeV. If instead of Λ we used f_π as the cutoff of the theory and thus as the momentum expansion parameter, the fact that $m_\pi > f_\pi$ would result in the failure of χPT for pion scattering already at the threshold $p \sim m_\pi$. Evidently, this not the case.

We can easily generalize the previous discussion to the ALP scenario: although the characteristic scale suppressing the ALP couplings can be taken to be f_a, the scale that defines the validity of the EFT description is Λ. As in the case of the pions, there is no contradiction in the possibility of m_a larger than f_a as long as m_a remains below Λ.

Despite the above discussion, we note that the comparison between m_a and f_a can still provide useful information. If an ALP signal consistent with a mass larger than f_a is detected, this suggests that the underlying dynamics may be strongly interacting. Likewise, if the signal points to

[1] If the heavy fields contribute to the effective operators only at loop level, then the cutoff Λ should not be directly identified with their mass. In such cases, Λ rather parametrizes the suppression scale of the induced operators, which can differ from the heavy mass by loop factors (e.g., $\Lambda \sim 16\pi^2 M$).

[2] The 4π counting is equivalent to $\hbar$ counting in the EFT, as also discussed previously in Refs. [140–142].

$m_a \lesssim f_a$, then a weakly interacting UV shall be favored as the origin of the ALP. We can exemplify this concept adopting the NDA master formula Eq. (2.1) as applied to a simple model containing a complex scalar field $\phi(x)$ transforming under the PQ symmetry. The associated Lagrangian reads

$$\mathcal{L} \supset \partial_\mu \phi^* \partial^\mu \phi - [-\Lambda^2 \phi^* \phi + g_*^2 (\phi^* \phi)^2], \tag{2.3}$$

and when $g_* = 4\pi$ the 1-loop corrections to the quartic couplings are as large as the tree-level parameter, thus describing a strongly interacting regime.

At the minimum,

$$\phi(x) = \frac{\langle \phi \rangle + \rho(x)}{\sqrt{2}} e^{ia(x)/f_a}, \qquad \langle \phi \rangle = \frac{\Lambda}{g_*}, \tag{2.4}$$

with $\rho(x)$ being the radial degree of freedom with mass Λ and a being the ALP. Once we focus on the interactions describing the ALP, canonical kinetic terms imply a relation between f_a and $\langle \phi \rangle$, that is, $f_a = \langle \phi \rangle$. This relation is equivalent to the one in Eq. (2.2) valid in χPT,

$$\Lambda = g_* f_a. \tag{2.5}$$

Let us now add a derivative coupling with a generic fermion ψ: still according to the NDA master formula,

$$\frac{g_*}{\Lambda} \partial_\mu \phi \bar{\psi} \psi = \frac{1}{f_a} \partial_\mu \phi \bar{\psi} \psi \Rightarrow \frac{1}{f_a} \partial_\mu a \bar{\psi} \psi, \tag{2.6}$$

where in the last step we retained only the ALP degree of freedom.

This minimal and simplified example leads to the idea that *the ALP may arise from either a weakly or a strongly interacting regime, but the form of the ALP interactions is insensitive to this aspect. As a consequence, analyses of ALP couplings are not able to tell anything about the g_* factor of Eq. (2.2) that instead can be investigated with direct or indirect searches of the radial mode ρ. In contrast, if the ALP mass is larger than f_a, this alone signals that the particle must arise from strong dynamics. It is important to stress, however, that the reverse implication fails: $m_a < f_a$ does not guarantee a weakly interacting ALP.*

III. THE EFFECTIVE DESCRIPTION

Our starting point is the effective ALP Lagrangian in the so-called chirality-preserving or derivative basis, which at the cutoff scale Λ can be written as follows:

$$\begin{aligned}\mathcal{L}_{\partial a}(\Lambda) = {} & \frac{1}{2}(\partial_\mu a)^2 - \frac{1}{2} m_a^2 a^2 + \frac{a}{f_a}\left(\frac{g'^2 c_{\tilde{B}}}{16\pi^2} B_{\mu\nu}\tilde{B}^{\mu\nu} + \frac{g^2 c_{\tilde{W}}}{16\pi^2} W^i_{\mu\nu}\tilde{W}^{i,\mu\nu} + \frac{g_s^2 c_{\tilde{G}}}{16\pi^2} G^a_{\mu\nu}\tilde{G}^{a,\mu\nu}\right) \\ & + \frac{\partial^\mu a}{f_a}(\overline{q'_L}\gamma_\mu c'_q q'_L + \overline{u'_R}\gamma_\mu c'_u u'_R + \overline{d'_R}\gamma_\mu c'_d d'_R + \overline{\ell'_L}\gamma_\mu c'_\ell \ell'_L + \overline{e'_R}\gamma_\mu c'_e e'_R),\end{aligned} \tag{3.1}$$

to first order of the expansion in $1/f_a$. Here g_s, g, and g' denote the strong, weak, and hypercharge couplings, respectively; $G^a_{\mu\nu}$, $W^i_{\mu\nu}$, and $B_{\mu\nu}$ are the $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ field strengths and $\tilde{X}^{\mu\nu} \equiv \frac{1}{2}\epsilon^{\mu\nu\rho\sigma} X_{\rho\sigma}$, with $\epsilon^{0123} = 1$, their associated dual fields. The left-handed (LH) and right-handed (RH) fermions $f' = \{q'_L, u'_R, d'_R, \ell'_L, e'_R\}$ are triplets in flavor space, while the corresponding ALP couplings, $c'_f = \{c'_q, c'_u, c'_d, c'_\ell, c'_e\}$, are general 3×3 Hermitian matrices defined in the fermion flavor basis. No ALP–Higgs operator has been included in the Lagrangian of Eq. (3.1), being redundant when ALP–fermion operators are included [10].

Performing unitary transformations on the fermion fields, beyond affecting the SM Lagrangian, induces a change in the second line of Eq. (3.1). Denoting the transformations as $\psi'_L = V_\psi \psi_L$ and $\psi'_R = U_\psi \psi_R$, with V_ψ and U_ψ being unitary 3×3 matrices in the flavor space, the ALP-fermion couplings conserve the same form but with new coefficients c_ψ:

$$\begin{aligned} & c_u \equiv U_u^\dagger c'_u U_u, \qquad c_d \equiv U_d^\dagger c'_d U_d, \qquad c_e \equiv U_e^\dagger c'_e U_e \\ & c_q \equiv V_q^\dagger c'_q V_q, \qquad c_\ell \equiv V_\ell^\dagger c'_\ell V_\ell. \end{aligned} \tag{3.2}$$

This will be useful when discussing the change to any other flavor basis or the fermion mass basis.

It is possible to write the Lagrangian in Eq. (3.1) in the so-called chirality-flipping or Yukawa-like basis, that has some advantages in the phenomenological study. By performing fermion chiral transformations, we can obtain

$$\begin{aligned}\mathcal{L}_a(\Lambda) = {} & \frac{1}{2}(\partial_\mu a)^2 - \frac{1}{2} m_a^2 a^2 + \frac{a}{f_a}\left(\frac{g'^2 \hat{c}_{\tilde{B}}}{16\pi^2} B_{\mu\nu}\tilde{B}^{\mu\nu} + \frac{g^2 \hat{c}_{\tilde{W}}}{16\pi^2} W^i_{\mu\nu}\tilde{W}^{i,\mu\nu} + \frac{g_s^2 \hat{c}_{\tilde{G}}}{16\pi^2} G^a_{\mu\nu}\tilde{G}^{a,\mu\nu}\right) \\ & - \frac{a}{f_a}(\overline{q'_L} H \hat{Y}'_d d'_R + \overline{q'_L}\,\tilde{H}\,\hat{Y}'_u u'_R + \overline{\ell'_L} H \hat{Y}'_e e'_R + \text{H.c.}).\end{aligned} \tag{3.3}$$

In the previous expression, the ALP-fermion couplings are proportional to

$$\hat{Y}'_d = i(Y'_d c'_d - c'_q Y'_d), \qquad \hat{Y}'_u = i(Y'_u c'_u - c'_q Y'_u),$$
$$\hat{Y}'_e = i(Y'_e c'_e - c'_\ell Y'_e), \tag{3.4}$$

where Y'_ψ are the Yukawa coupling matrices in the original flavor basis. On the other hand, the anomalous terms receive contributions from the fermion chiral transformations according to the Fujikawa method [143],

$$\hat{c}_{\tilde{B}} = c_{\tilde{B}} - K_B, \qquad \hat{c}_{\tilde{W}} = c_{\tilde{W}} - K_W,$$
$$\hat{c}_{\tilde{G}} = c_{\tilde{G}} - K_G, \tag{3.5}$$

where the K_X are defined as

$$K_B \equiv \mathrm{Tr}[N_c(N_L \mathcal{Y}_q^2 c'_q - \mathcal{Y}_u^2 c'_u - \mathcal{Y}_d^2 c'_d) + N_L \mathcal{Y}_\ell^2 c'_\ell - \mathcal{Y}_e^2 c'_e], \tag{3.6}$$

$$K_W \equiv T_F \mathrm{Tr}[N_c c'_q + c'_\ell], \tag{3.7}$$

$$K_G \equiv T_F \mathrm{Tr}[N_L c'_q - c'_u - c'_d], \tag{3.8}$$

with $T_F = \frac{1}{2}$, $N_L = 2$, $N_c = 3$, and the hypercharges

$$\mathcal{Y}_q = \frac{1}{6}, \qquad \mathcal{Y}_u = \frac{2}{3}, \qquad \mathcal{Y}_d = -\frac{1}{3}$$
$$\mathcal{Y}_\ell = -\frac{1}{2}, \qquad \mathcal{Y}_e = -1. \tag{3.9}$$

The Lagrangians in Eqs. (3.1) and (3.3) are defined at the cutoff scale Λ. At lower scales, quantum corrections generically modify the ALP-fermion couplings. The RGE [60,95,144–146] state that the Wilson coefficients $c_{\tilde{B}}$, $c_{\tilde{W}}$, and $c_{\tilde{G}}$ of the anomalous terms do not receive RGE corrections at one loop, while the Wilson coefficients c'_ψ do receive corrections at this order. Our main focus in this paper is to explore regions of the parameter space where $m_a > f_a$, keeping a conservative approach to the hierarchy between the PQ and the electroweak (EW) symmetry breaking, $\Lambda \gg v_{\mathrm{EW}}$. For consistency, the condition $f_a > v_{\mathrm{EW}}/4\pi$ must be respected, where $v_{\mathrm{EW}} = 246$ GeV is the vacuum expectation value of the Higgs model.

At the end, the RGE running from the cutoff scale Λ down to m_a is taken into consideration, using the ALPaca [120] code to implement this part of the analysis. The relevant Lagrangian to match with depends on whether m_a is larger or smaller than v_{EW}. In the former case, $m_a > v_{\mathrm{EW}}$, the leading contributions to the fermionic Lagrangian read

$$\mathcal{L}^{>v}_{\partial a}(m_a) \supset \frac{\partial^\mu a}{f_a}(\overline{q_L}\gamma_\mu c_q^L q_L + \overline{u_R}\gamma_\mu c_u^R u_R + \overline{d_R}\gamma_\mu c_d^R d_R + \overline{\ell_L}\gamma_\mu c_\ell^L \ell_L + \overline{e_R}\gamma_\mu c_e^R e_R), \tag{3.10}$$

where $c^{L,R}$ are the running Wilson coefficients defined at the scale m_a. If, instead, $m_a < v_{\mathrm{EW}}$, then the members of the $SU(2)_L$-doublets should be treated differently and the coefficients undergo a different running under the RGE. For example, focusing for simplicity on the lepton sector, the Lagrangian would contain three different terms,

$$\mathcal{L}^{<v}_{\partial a}(m_a) \supset \frac{\partial^\mu a}{f_a}(\overline{e_L}\gamma_\mu c^L e_L + \overline{e_R}\gamma_\mu c^R e_R + \overline{\nu_L}\gamma_\mu c^\nu \nu_L), \tag{3.11}$$

where both c^L and c^ν derive from c_ℓ^L and differ only for the RGE contributions.

For the purposes of this work, we restrict ourselves to leptonic observables, as they offer a cleaner and more precise ground compared to hadronic ones. *We will focus on leptophilic ALPs, i.e., ALPs that couple predominantly to leptons, with negligible couplings to quarks.* Nevertheless, running the leptonic couplings from a high scale Λ down to the scale of the ALP mass m_a will generate couplings to quarks, especially to the top (t) and bottom (b) flavors. Therefore, it is important to gauge the impact this could have on hadronic observables. For this purpose, we first perform the RGE running from the scale Λ down to the ALP mass m_a, where the ALP is integrated out, and we match to the relevant effective Lagrangian at that scale, constructed only with SM fields. In the case of $m_a > v_{\mathrm{EW}}$, the sensible description is the Standard Model effective field theory (SMEFT) [147,148] Lagrangian. We integrate out the ALP in the chirality-flipping basis of Eq. (3.3) to simplify the procedure and, matching onto the SMEFT Warsaw basis [148] and applying Fierz redefinitions, we find that the following $d = 6$ operators receive contributions:

$$\mathcal{L}_a^{4\psi} = \mathcal{C}_{le}\mathcal{Q}_{le} + \mathcal{C}_{qu}^{(1)}\mathcal{Q}_{qu}^{(1)} + \mathcal{C}_{qd}^{(1)}\mathcal{Q}_{qd}^{(1)} + \mathcal{C}_{ledq}\mathcal{Q}_{ledq} + \mathcal{C}_{quqd}^{(1)}\mathcal{Q}_{quqd}^{(1)} + \mathcal{C}_{quqd}^{(8)}\mathcal{Q}_{quqd}^{(8)} + \mathcal{C}_{lequ}^{(1)}\mathcal{Q}_{lequ}^{(1)}, \tag{3.12}$$

where

$$
\begin{aligned}
\mathcal{C}_{le} &= -\frac{v^2}{4m_a^2 f_a^2}(\hat{Y}'_e)^{pt}(\hat{Y}'^{\dagger}_e)^{sr} & \mathcal{Q}_{le} &= \left(\overline{e'^p_L}\gamma_\mu e'^r_L\right)\left(\overline{e'^s_R}\gamma^\mu e'^t_R\right)\\
\mathcal{C}^{(1)}_{qu} &= -\frac{v^2}{12m_a^2 f_a^2}(\hat{Y}'_u)^{pt}(\hat{Y}'^{\dagger}_u)^{sr} & \mathcal{Q}^{(1)}_{qu} &= \left(\overline{q'^p_L}\gamma_\mu q'^r_L\right)(\overline{u'^s_R}\gamma^\mu u'^t_R)\\
\mathcal{C}^{(1)}_{qd} &= -\frac{v^2}{12m_a^2 f_a^2}(\hat{Y}'_d)^{pt}(\hat{Y}'^{\dagger}_d)^{sr} & \mathcal{Q}^{(1)}_{qd} &= \left(\overline{q'^p_L}\gamma_\mu q'^r_L\right)\left(\overline{d'^s_R}\gamma^\mu d'^t_R\right)\\
\mathcal{C}_{ledq} &= -\frac{v^2}{2m_a^2 f_a^2}(\hat{Y}'_e)^{pr}(\hat{Y}'^{\dagger}_d)^{st} & \mathcal{Q}_{ledq} &= \left(\overline{\ell'^p_L}e'^r_R\right)\left(\overline{d'^s_R}q'^t_L\right)\\
\mathcal{C}^{(1)}_{quqd} &= \frac{v^2}{12m_a^2 f_a^2}(\hat{Y}'_u)^{sr}(\hat{Y}'_d)^{pt} & \mathcal{Q}^{(1)}_{quqd} &= \left(\overline{q'^p_L}u'^r_R\right)\varepsilon_{jk}\left(\overline{q'^s_L}d'^t_R\right)\\
\mathcal{C}^{(8)}_{quqd} &= \frac{v^2}{2m_a^2 f_a^2}(\hat{Y}'_u)^{sr}(\hat{Y}'_d)^{pt} & \mathcal{Q}^{(8)}_{quqd} &= \left(\overline{q'^p_L}T^A u'^r_R\right)\varepsilon_{jk}\left(\overline{q'^s_L}T^A d'^t_R\right)\\
\mathcal{C}^{(1)}_{lequ} &= \frac{v^2}{2m_a^2 f_a^2}(\hat{Y}'_e)^{pr}(\hat{Y}'_u)^{st} & \mathcal{Q}^{(1)}_{lequ} &= \left(\bar{\ell}'^p_L e'^r_R\right)\varepsilon_{jk}\left(\bar{q}'^s_L u'^t_R\right),
\end{aligned}
\tag{3.13}
$$

where T^A are the $SU(3)_c$ generators and ε_{jk} is the 2-dimensional Levi-Civita tensor. We made explicit the flavor contractions in the expressions of the Wilson coefficients and of the operators. To the best of our knowledge, this has not been presented elsewhere in the literature, where the general treatment involves $m_a \ll v_{\rm EW}$.

Observing the Wilson coefficients listed above, we can notice that most of them are strongly suppressed. On the one hand, there are suppressing factors of v^2/f_a^2 or v^2/m_a^2, depending on whether f_a or m_a constitutes the larger scale. On the other hand, the $\hat{Y}'$ terms contain, by definition, the fermion Yukawas and the Lagrangian coefficients c'_ψ of the ALP-fermion couplings. In particular, as already stated, we will focus on leptophilic ALPs and therefore c'_q, c'_u, and c'_d are vanishing at the scale Λ. Moreover, an explicit computation including the RGE effects reveals that the largest entries of the quark $\hat{Y}'$ turn out to be of $\mathcal{O}(10^{-4})$ and associated to the t and b flavors. We will enter more into details in the next section, after having defined the ALP-lepton coupling textures.

On the whole, all Wilson coefficients $\mathcal{C}$ except for $\mathcal{C}_{le}$ are extremely suppressed and the corresponding operators turn out to have a negligible effect in our study. On the contrary, the operator $\mathcal{Q}_{le}$ may have a relevant role in the numerical analysis, depending on the specific ALP-lepton coupling texture considered. In the next section, we will indeed provide the explicit expressions for various charged lepton decays that can be described by this operator.

Regarding the ALP couplings to gauge bosons, and especially the coupling to photons, we remark that they play a fundamental role due to their effect on a large variety of observables, and will therefore be discussed in more detail in the next section. As mentioned before, the coefficients $c_{\tilde{B}}$ and $c_{\tilde{W}}$ do not receive any correction through the RGE running at one-loop (1L) level. Consistently with our hypothesis of leptophilic ALPs, where only the coefficients c'_ℓ and c'_e in Eq. (3.1) are nonvanishing, one may then conclude that no interactions with photons are then generated. However, this is not the case even at scales larger than $\Lambda_{\rm QCD}$, the QCD confinement scale, where contributions are indeed expected due to the nonperturbative effects described within the chiral perturbation approach. The explicit computation of the $a\gamma\gamma$ coupling at 1L, described by the classic triangle Feynman diagram with internal charged leptons, is nonvanishing [see Ref. [119] for a complete and updated discussion on this aspect]. In the end, the effective description of the ALP-photon couplings at any scale $m_a > m_\tau$ can be approximately written as

$$
\mathcal{L}^{<v}_{\partial a}(m_a) \supset \frac{a}{f_a}\frac{\alpha_{\rm em}}{4\pi}c^{\rm eff}_{a\gamma\gamma}F_{\mu\nu}\tilde{F}^{\mu\nu} \tag{3.14}
$$

with

$$
c^{\rm eff}_{a\gamma\gamma} = -\sum_i (c^L_{ii} - c^R_{ii}), \tag{3.15}
$$

for the case under consideration of a leptophilic ALP. It will also be useful to specify the effective coupling of the ALP with one photon and one Z, as it enters in several penguin diagrams:

$$
\mathcal{L}^{<v}_{\partial a}(m_a) \supset \frac{a}{f_a}\frac{\alpha_{\rm em}}{2\pi s_w c_w}c^{\rm eff}_{a\gamma Z}F_{\mu\nu}\tilde{Z}^{\mu\nu} \tag{3.16}
$$

with

$$
c^{\rm eff}_{a\gamma Z} = -1.88\times 10^{-2}\sum_i (c^L_{ii} - c^R_{ii}), \tag{3.17}
$$

TABLE I. Summary of leptonic observables used in our analysis, as well as their predictions within the SM context with massive active neutrinos and current experimental limits and the future prospects (in parentheses).

Observable	SM prediction + m_ν	Experimental limits
$(g-2)_e/2$	Cs: 0.0011596521816l(23) [150] Rb: 0.001159652180252(95) [151]	0.00115965218062(12) [149]
$(g-2)_\mu/2$	0.00116592033(62) [152]	0.00116592059(22) [153]
$\mathcal{BR}(\mu \to e\gamma)$	$\lesssim \mathcal{O}(10^{-55})$ [154–156]	$<1.5 \times 10^{-13}$ [157] ($<6 \times 10^{-14}$ [157,158])
$\mathcal{BR}(\mu \to 3e)$	$\lesssim \mathcal{O}(10^{-55})$ [154,159]	$<1.0 \times 10^{-12}$ [160] (Mu3e: $< 5 \times 10^{-16}$ [161])
$\mathcal{BR}(\mu + N \to e + N)$	$\lesssim \mathcal{O}(10^{-54})$ [162]	$\mathcal{BR}_{\mathrm{Au}} < 7 \times 10^{-13}$ [163] (Mu2e: $\mathcal{BR}_{\mathrm{Al}} < 2.87 \times 10^{-17}$ [164]) (COMET: $\mathcal{BR}_{\mathrm{Al}} < 2 \times 10^{-17}$ [165])
$\mathcal{BR}(\tau \to e\gamma)$	$\lesssim \mathcal{O}(10^{-49})$ [155]	$<3.3 \times 10^{-8}$ [149]
$\mathcal{BR}(\tau \to \mu\gamma)$	$\lesssim \mathcal{O}(10^{-49})$ [155]	$<4.2 \times 10^{-8}$ [149]
$\mathcal{BR}(\tau \to 3e)$	$\lesssim \mathcal{O}(10^{-56})$ [154,159]	$<2.7 \times 10^{-8}$ [149]
$\mathcal{BR}(\tau \to 3\mu)$	$\lesssim \mathcal{O}(10^{-55})$ [154,159]	$<1.9 \times 10^{-8}$ [149]
$\mathcal{BR}(\tau^- \to e^-\mu^+\mu^-)$	$\lesssim \mathcal{O}(10^{-56})$ [154,159]	$<2.7 \times 10^{-8}$ [149]
$\mathcal{BR}(\tau^- \to \mu^- e^+ e^-)$	$\lesssim \mathcal{O}(10^{-55})$ [154,159]	$<1.8 \times 10^{-8}$ [149]

for all the ALP mass values considered in our analysis, where s_w and c_w stand for the sine and cosine of the Weinberg angle, $\sin^2\theta_w = 0.2312$ [149].

Before concluding this section, we comment on shift-symmetry-breaking terms and their possible effects. Let us again focus on the minimal scenario illustrated in the previous section, whose Lagrangian is given by Eq. (2.3). By definition, any shift-symmetry-breaking dynamics can only operate below the scale at which the PQ symmetry is spontaneously broken, namely, Λ. Naively, this is traditionally treated by introducing a "small" explicit symmetry breaking parameter, to which any such contribution is proportional. The advantage of this approach is that, once this parameter is set to zero, no explicit symmetry breaking effect survives. However, this does not imply that the resulting contributions to shift-symmetry-breaking observables must be of the same order as the small parameter itself.

The pion mass again provides a useful example. Even neglecting electromagnetic interactions, the pion mass is much larger than the quark masses that explicitly break chiral symmetry. In this case, the dominant contribution to the pion mass arises from chromomagnetic interactions, which would ultimately vanish in the limit of massless quarks.

Generalizing this discussion, the size of explicit shift-symmetry-breaking effects is in fact strongly model-dependent, and the ALP mass term does not necessarily play any special role beyond constituting the lowest dimensional operator that explicitly breaks the shift symmetry. Therefore, in our high-energy Lagrangian, we shall not consider any other shift-symmetry-breaking terms beyond the anomalous couplings and the ALP mass term; without any specific underlying rationale, the opposite ansatz would have us dealing with a generic CP-odd scalar and not with an ALP.

IV. OBSERVABLES

Having established the theoretical discussion, we proceed with a data-driven approach. We will focus on leptonic processes to study the viability of the parameter space within the range $m_a > f_a$, although we will also integrate the complementary window $m_a < f_a$. As already stated, these observables offer a cleaner and more precise ground compared to hadronic ones. Although the analysis focuses on specific leptophilic ALP-coupling textures, it allows us to firmly establish that the $m_a > f_a$ parameter space is viable and interesting. Last but not least, this approach is self-consistent without the need of any *ad hoc* hypothesis, besides requiring a leptophilic ALP. In this case, the hadron phenomenology is negligible because the ALP-quark and ALP-gluon couplings are extremely suppressed.

The selected processes are as follows:

(i) The anomalous magnetic dipole moments of the electron and muon, $(g-2)_e$ and $(g-2)_\mu$.

(ii) Charged lepton flavor violating (cLFV) processes: radiative charged lepton decays, μ–e conversion in nuclei, and τ decays into lighter charged leptons.

They have a high experimental precision, especially when the electron and the muon are involved. A summary with the SM predictions, current experimental values, and future prospects is given in Table I. A comment is in order: when dealing with the μ–e conversion in nuclei, we consider the best present bound from SINDRUM II with gold, and the future prospects with aluminum from COMET.

For the anomalous magnetic dipole moment of the electron, we need to distinguish between the two determinations, involving the measurement of $\alpha_{\rm em}$ with cesium [150] and with rubidium [151], that lead to two different tensions with the current experimental measure:

$$\delta a_e^{\rm Cs} = (-9.9 \pm 2.6) \times 10^{-13}\,(-3.8\sigma),$$
$$\delta a_e^{\rm Rb} = (3.7 \pm 1.5) \times 10^{-13}\,(+2.4\sigma), \tag{4.1}$$

where $\delta a_e^X \equiv a_e^{\rm Exp} - a_e^X$.

Remarkably, the cesium determination shows an anomaly that, if consolidated, would call for a new physics contribution to be explained. We will discuss this possibility in the next section.

For the $(g-2)_\mu$, the situation changed considerably in recent years. Indeed, the lattice-QCD calculation precision has increased and allows for a new precise determination of the hadron vacuum polarization contribution that results in a major upward shift of the total SM prediction. Compared against the current experimental determination, this implies that there is no tension between the SM and experiment at the current level of precision. Therefore, we will treat it at the same level as the other constraints, and not as an anomaly.

Besides the above list of lepton flavor conserving and violating observables, one may further consider cosmological and astrophysical bounds, as well as collider constraints. We do not perform a dedicated analysis of these observables in this work. Rather, we provide qualitative arguments to assess their relevance, while noting that a dedicated collider study may be required in selected regions of the parameter space.

Focusing first on astrophysical searches [for a review, see, for instance, Part III of Ref. [123]] that include, in particular, studies on the ALP production mechanism in the Sun, red giants, neutron stars, and supernovae, very strong constraints can be derived from energy-loss arguments for ALP masses smaller than $m_a \sim 0.1$ GeV. Our study applies only to masses larger than this quantity, and therefore, we can avoid discussing these observables.

On the other hand, for ALP masses in the GeV range or above, cosmological bounds are typically relevant only for extremely small ALP-photon couplings [for a review, see, for instance, Part II of Ref. [123]]. For larger couplings, the ALP lifetime becomes very short, so that the particle decays well before epochs such as big bang nucleosynthesis. In addition, these cosmological limits are intrinsically model-dependent, as they rely on assumptions about the thermal history of the Universe, in particular the reheating temperature and the efficiency of thermal production mechanisms. Based on the ALP-photon coupling considerations in the previous section, the ALP-photon coupling in our scenarios is at least four orders of magnitude larger than the region where cosmological bounds become relevant. Consequently, the corresponding lifetime is extremely short on cosmological timescales, and the particle decays long before it could affect standard cosmological observables, rendering those constraints irrelevant for the parameter space we will explore.

Finally, turning our attention to collider searches, the most relevant constraints are on the ALP–photon coupling and arise from channels with final states containing two or more photons [56,73,76]. In particular, the dominant bounds are driven by large electron positron searches for associated ALP production of the type $e^+e^- \to \gamma a \to 3\gamma$, by limits from exotic decays of on-shell Z bosons, such as $Z \to \gamma a \to 3\gamma$, and by LHC analyses targeting diphoton resonances or photon-fusion production of ALPs subsequently decaying into photon pairs. The present data constrain the region of the parameter space for ALP masses in 1 GeV $\lesssim m_a \lesssim$ 1 TeV and ALP-photon couplings that can be translated in our notation to be $f_a/|c_{a\gamma\gamma}| \gtrsim 30$ GeV. While this estimate suggests that the corresponding condition is fulfilled over much of the parameter space considered below, it identifies selected low-f_a regions as potentially testable by dedicated collider searches. We return to this possibility in the phenomenological discussion.

Furthermore, when focusing on the future experimental facilities [76], the sensitivity to ALP-photon couplings increases through dedicated searches. Promising channels include associated production with photons, Z or Higgs bosons at future e^+e^- colliders such as FCC-ee, CEPC, ILC and CLIC, as well as diphoton resonance searches, photon-fusion production, and exotic Higgs or Z decays at the HL-LHC and possible future hadron colliders. Dedicated analyses of these signatures could therefore provide sensitivity to regions of our parameter space.

Moreover, additional constraints may arise from collider probes of ALP couplings to top quarks. Searches for heavy pseudoscalars decaying into $t\bar{t}$ pairs, as well as production in association with top quarks, provide bounds on the effective top coupling. For $m_a \gtrsim 400$ GeV, current analyses translate into limits of the order $|c_{tt}|/f_a \lesssim \mathcal{O}(10^{-2}\text{–}10^{-3})$ GeV^{-1}, under specific assumptions on the ALP-gluon coupling and the ALP total decay rate [166–169]. In the scenario considered here, the induced top coupling obtained after matching to the Yukawa basis satisfies $(\hat{Y}'_u)^{33} \lesssim \mathcal{O}(10^{-4})$, as discussed in the Appendix, which lies well below the current experimental sensitivity. Consequently, collider observables involving top quarks do not impose relevant constraints on the parameter space explored in this work.

We will consider two different scenarios: (i) an ALP that couples only to the first two generations and (ii) an ALP that couples to the three families. Indeed, since the ALP–fermion couplings are proportional to the fermion masses, an ALP that has a tree-level coupling to τ leads to a noticeably different phenomenology compared to a τ-phobic ALP. Besides this, we adopt specific benchmark

flavor structures to analyze how the results depend on the underlying scenarios.

In what follows, in order to simplify the notation, we will redefine the coefficients c_ℓ^L and c_e^R in Eq. (3.10) as c^L and c^R, respectively. We recall that they account for the running RGE effects from the cutoff scale Λ down to the ALP mass m_a.[3] We will better clarify the relation between these coefficients and the Lagrangian ones in the next section, after specifying the ALP-lepton coupling textures.

A. $e-\mu$ framework

We first consider the simplified scenario of an ALP coupling only to the first two families, e, μ. The leading-order contributions to the relevant observables, computed for an off-shell ALP in the limit $m_a \gg m_\mu$, are given by

$$a_e^{\rm ALP} \approx \frac{1}{8\pi^2}\frac{m_e m_\mu}{f_a^2}\left\{\frac{m_\mu^2}{m_a^2}\mathrm{Re}[c_{\mu e}^L c_{\mu e}^{R\,*}]\left(-\frac{3}{2}+\log\left(\frac{m_a^2}{m_\mu^2}\right)\right)+\frac{\alpha_{\rm em}}{\pi}\frac{m_e}{m_\mu}(c_{ee}^L-c_{ee}^R)c_{a\gamma\gamma}^{\rm eff}\left(-\frac{3}{2}+\log\left(\frac{\Lambda^2}{m_a^2}\right)\right)\right\}, \tag{4.2}$$

$$\begin{aligned} a_\mu^{\rm ALP} \approx \frac{1}{8\pi^2}\frac{m_\mu^2}{f_a^2}\Bigg\{&\frac{m_\mu^2}{m_a^2}\left[-(c_{\mu\mu}^L-c_{\mu\mu}^R)^2\left(-\frac{11}{6}+\log\left(\frac{m_a^2}{m_\mu^2}\right)\right)+\frac{1}{6}(|c_{\mu e}^L|^2+|c_{\mu e}^R|^2)\right.\\ &\left.+\frac{m_e}{m_\mu}\mathrm{Re}[c_{\mu e}^L c_{\mu e}^{R\,*}]\left(-\frac{13}{6}+\log\left(\frac{m_a^2}{m_e^2}\right)\right)\right]+\frac{\alpha_{\rm em}}{\pi}(c_{\mu\mu}^L-c_{\mu\mu}^R)c_{a\gamma\gamma}^{\rm eff}\left(-\frac{3}{2}+\log\left(\frac{\Lambda^2}{m_a^2}\right)\right)\Bigg\}, \end{aligned} \tag{4.3}$$

$$\begin{aligned} \mathcal{BR}(\mu\to e\gamma) \approx \frac{\alpha_{\rm em}}{1024\pi^4}\frac{m_\mu^5}{f_a^4\Gamma_\mu}(|c_{\mu e}^L|^2+|c_{\mu e}^R|^2)\Bigg[&\frac{m_\mu^2}{m_a^2}(c_{\mu\mu}^L-c_{\mu\mu}^R)\left(-\frac{5}{3}+\log\left(\frac{m_a^2}{m_\mu^2}\right)\right)\\ &-\frac{\alpha_{\rm em}}{\pi}c_{a\gamma\gamma}^{\rm eff}\left(-\frac{3}{2}+\log\left(\frac{\Lambda^2}{m_a^2}\right)\right)\Bigg]^2, \end{aligned} \tag{4.4}$$

$$\mathcal{BR}(\mu\to 3e) \approx \frac{\alpha_{\rm em}}{3\pi}\left(-\frac{11}{4}+\log\left(\frac{m_\mu^2}{m_e^2}\right)\right)\mathcal{BR}(\mu\to e\gamma), \tag{4.5}$$

$$\begin{aligned} \mathcal{BR}(\mu^- N\to e^- N) \approx &\frac{\xi}{512\pi^4}\frac{m_\mu^4}{f_a^4}(|c_{\mu e}^L|^2+|c_{\mu e}^R|^2)\\ &\times\left[\frac{2m_\mu^2}{3m_a^2}(c_{\mu\mu}^L-c_{\mu\mu}^R)\left(-\frac{11}{6}+\log\left(\frac{m_a^2}{m_\mu^2}\right)\right)-\frac{\alpha_{\rm em}}{\pi}c_{a\gamma\gamma}^{\rm eff}\left(-\frac{3}{2}+\log\left(\frac{\Lambda^2}{m_a^2}\right)\right)\right]^2, \end{aligned} \tag{4.6}$$

where Γ_μ is the total muon decay rate and the prefactor ξ is defined as

$$\xi = \frac{8\alpha_{\rm em}^5 m_\mu Z_{\rm eff}^4 Z F_p^2}{\Gamma_{\rm capt}}, \tag{4.7}$$

with $\Gamma_{\rm capt}$ being the total muon capture rate, Z ($Z_{\rm eff}$) the (effective) atomic charge, and F_p the nuclear form factor. Explicit values for Au and Al are [94]

$$\xi_{\rm Au} \approx 5.2\times 10^{12}, \qquad \xi_{\rm Al} \approx 3.5\times 10^{12}. \tag{4.8}$$

Notice that, in all these expressions, log stands for the natural logarithm.

A careful reader may notice that these expressions only contain ALP couplings with the charged leptons and with photons, but not with a photon and a Z. The reason resides again in the fact that only the dominant contributions are shown in the expressions (and also in those for the three-family scenario reported below). In particular, when focusing on masses $m_a \gg m_\tau$, the form factors associated with the contributions involving $c_{a\gamma Z}^{\rm eff}$ are extremely suppressed (see *Mathematica* file

[3]The additional evolution down to the mass scale of the relevant charged lepton, the scale at which the various observables are computed, is simply the purely SM RGE running and thus negligible.

in Supplemental Material), and therefore, we have not shown them explicitly in the above formulas, but nevertheless we will consider them in our numerical analysis.

B. Three-family framework

In the case of an ALP coupling to three generations, the ALP couplings to τ play a dominant role. Not only should the τ phenomenology be considered, but also the expressions for the μ and electron processes introduced above become substantially modified. All in all, ignoring the specific hierarchies among the ALP-lepton couplings, the expressions of the relevant observables simplify as follows: for an off-shell ALP in the limit $m_a \gg m_\tau$,

$$a_e^{\rm ALP} \approx \frac{1}{8\pi^2}\frac{m_e m_\tau}{f_a^2}\left\{\frac{m_\tau^2}{m_a^2}\mathrm{Re}[c_{\tau e}^L c_{\tau e}^{R\,*}]\left(-\frac{3}{2}+\log\left(\frac{m_a^2}{m_\tau^2}\right)\right)+\frac{m_e}{m_\tau}\left[\frac{\alpha_{\rm em}}{\pi}(c_{ee}^L-c_{ee}^R)c_{a\gamma\gamma}^{\rm eff}\left(-\frac{3}{2}+\log\left(\frac{\Lambda^2}{m_a^2}\right)\right)\right]\right\}, \tag{4.9}$$

$$a_\mu^{\rm ALP} \approx \frac{1}{8\pi^2}\frac{m_\mu m_\tau}{f_a^2}\left\{\frac{m_\tau^2}{m_a^2}\mathrm{Re}[c_{\tau\mu}^L c_{\tau\mu}^{R\,*}]\left(-\frac{3}{2}+\log\left(\frac{m_a^2}{m_\tau^2}\right)\right)+\frac{m_\mu}{m_\tau}\left[\frac{\alpha_{\rm em}}{\pi}(c_{\mu\mu}^L-c_{\mu\mu}^R)c_{a\gamma\gamma}^{\rm eff}\left(-\frac{3}{2}+\log\left(\frac{\Lambda^2}{m_a^2}\right)\right)\right]\right\}, \tag{4.10}$$

$$\mathcal{BR}(\mu\to e\gamma) \approx \frac{\alpha_{\rm em}}{1024\pi^4}\frac{m_\mu^3 m_\tau^2}{f_a^4\Gamma_\mu}\left\{\frac{m_\tau^4}{m_a^4}(|c_{\tau e}^L|^2|c_{\tau\mu}^R|^2+|c_{\tau e}^R|^2|c_{\tau\mu}^L|^2)\left(-\frac{3}{2}+\log\left(\frac{m_a^2}{m_\tau^2}\right)\right)^2+\frac{\alpha_{\rm em}^2}{\pi^2}\frac{m_\mu^2}{m_\tau^2}(c_{a\gamma\gamma}^{\rm eff})^2(|c_{\mu e}^L|^2+|c_{\mu e}^R|^2)\left(-\frac{3}{2}+\log\left(\frac{\Lambda^2}{m_a^2}\right)\right)^2\right\}, \tag{4.11}$$

$$\mathcal{BR}(\tau\to e\gamma) \approx \frac{\alpha_{\rm em}}{1024\pi^4}\frac{m_\tau^5}{f_a^4\Gamma_\tau}(|c_{\tau e}^L|^2+|c_{\tau e}^R|^2)\left[\frac{m_\tau^2}{m_a^2}(c_{\tau\tau}^L-c_{\tau\tau}^R)\left(-\frac{5}{3}+\log\left(\frac{m_a^2}{m_\tau^2}\right)\right)-\frac{\alpha_{\rm em}}{\pi}c_{a\gamma\gamma}^{\rm eff}\left(-\frac{3}{2}+\log\left(\frac{\Lambda^2}{m_a^2}\right)\right)\right]^2, \tag{4.12}$$

$$\mathcal{BR}(\tau\to\mu\gamma) \approx \frac{(|c_{\tau\mu}^L|^2+|c_{\tau\mu}^R|^2)}{(|c_{\tau e}^L|^2+|c_{\tau e}^R|^2)}\mathcal{BR}(\tau\to e\gamma), \tag{4.13}$$

$$\mathcal{BR}(\mu^- N\to e^- N) \approx \frac{2\xi}{\alpha_{\rm em}}\frac{\Gamma_\mu}{m_\mu}\mathcal{BR}(\mu\to e\gamma), \tag{4.14}$$

$$\mathcal{BR}(\mu\to 3e) \approx \frac{\alpha_{\rm em}}{3\pi}\left(-\frac{11}{4}+\log\left(\frac{m_\mu^2}{m_e^2}\right)\right)\mathcal{BR}(\mu\to e\gamma), \tag{4.15}$$

$$\mathcal{BR}(\tau\to 3e) \approx \frac{\alpha_{\rm em}}{3\pi}\left(-\frac{11}{4}+\log\left(\frac{m_\tau^2}{m_e^2}\right)\right)\mathcal{BR}(\tau\to e\gamma), \tag{4.16}$$

$$\mathcal{BR}(\tau\to 3\mu) \approx \frac{\alpha_{\rm em}}{3\pi}\left(-\frac{11}{4}+\log\left(\frac{m_\tau^2}{m_\mu^2}\right)\right)\mathcal{BR}(\tau\to\mu\gamma), \tag{4.17}$$

$$\mathcal{BR}(\tau^- \to \mu^- e^+ e^-) \approx \frac{\alpha_{\rm em}^2}{3072\pi^5}\frac{m_\tau^5}{f_a^4\Gamma_\tau}(|c_{\tau\mu}^L|^2+|c_{\tau\mu}^R|^2) \times\left\{\frac{m_\tau^4}{m_a^4}(c_{\tau\tau}^L-c_{\tau\tau}^R)^2\left[-\frac{11}{36}\left(-\frac{56}{33}+\log\left(\frac{m_a^2}{m_\tau^2}\right)\right)\left(-\frac{4}{3}+\log\left(\frac{m_a^2}{m_\tau^2}\right)\right)\right.\right. \left.\left.+\left(-\frac{5}{3}+\log\left(\frac{m_a^2}{m_\tau^2}\right)\right)^2\left(-3+\log\left(\frac{m_\tau^2}{m_e^2}\right)\right)\right]+\frac{\alpha_{\rm em}^2}{\pi^2}(c_{a\gamma\gamma}^{\rm eff})^2\left(-\frac{3}{2}+\log\left(\frac{\Lambda^2}{m_a^2}\right)\right)^2\left(-3+\log\left(\frac{m_\tau^2}{m_e^2}\right)\right)\right\}, \tag{4.18}$$

$$\mathcal{BR}(\tau^- \to e^- \mu^+ \mu^-) \approx \frac{\alpha_{\rm em}^2}{3072\pi^5}\frac{m_\tau^5}{f_a^4\Gamma_\tau}(|c_{\tau e}^L|^2+|c_{\tau e}^R|^2) \times\left\{\frac{m_\tau^4}{m_a^4}(c_{\tau\tau}^L-c_{\tau\tau}^R)^2\left[-\frac{11}{36}\left(-\frac{56}{33}+\log\left(\frac{m_a^2}{m_\tau^2}\right)\right)\left(-\frac{4}{3}+\log\left(\frac{m_a^2}{m_\tau^2}\right)\right)\right.\right. \left.\left.+\left(-\frac{5}{3}+\log\left(\frac{m_a^2}{m_\tau^2}\right)\right)^2\left(-3+\log\left(\frac{m_\tau^2}{m_\mu^2}\right)\right)\right]+\frac{\alpha_{\rm em}^2}{\pi^2}(c_{a\gamma\gamma}^{\rm eff})^2\left(-\frac{3}{2}+\log\left(\frac{\Lambda^2}{m_a^2}\right)\right)^2\left(-3+\log\left(\frac{m_\tau^2}{m_\mu^2}\right)\right)\right\}. \tag{4.19}$$

In the above expressions, Γ_τ stands for the total τ decay rate. The attributes of the ALP couplings mentioned in the two-family case also apply here.

A few more comments are in order. The expressions in Eqs. (4.18) and (4.19) only show the one-loop contributions proportional to $\alpha_{\rm em}^2$. These are indeed the dominant contributions especially for large m_a values. Nevertheless, the tree-level contributions and the interference ones are in general not negligible and are taken into consideration in our numerical analysis. On the other hand, our analysis does not include the processes $\tau^- \to \mu^+ e^- e^-$ and $\tau^- \to e^+ \mu^- \mu^-$ since a dedicated analytical study beyond the present literature would be necessary, a task which we leave for a separate work. Nevertheless, their impact in our numerical results is expected to be negligible given the present and future experimental sensitivities.

Last but not least, we point out that, for certain flavor textures for which the ALP-τ couplings are not negligible yet still small relative to other ALP-lepton couplings, the complete expressions for the above observables should be used. Indeed, due to the enhancement of the τ-mass, it may occur that contributions proportional to "small" ALP-τ couplings are as important as those involving "large" ALP-μ couplings. These expressions can be found in the *Mathematica* file in Supplemental Material.

C. Benchmark flavor textures

In order to explore the richness of the parameter space, we focus on a set of flavor textures for the ALP-lepton couplings that allow us to extract relevant information for each case. We will first focus on the two-family case and then move to the full three-family one. There are two competing ingredients that determine the expected results in each case: the dominant couplings present in the different observables and the present (and/or future) experimental sensitivities of these observables. The discussion that follows will be driven by both in order to identify flavor structures that lead to different phenomenological regions in the parameter space.

Before going into details of the different textures, it is necessary to specify the scale at which they are defined and the connection with the Lagrangian coefficients c_ψ. The working hypothesis is that the high-energy Lagrangian in Eq. (3.1) at the scale Λ is defined in the minimal lepton flavor violation (MLFV) [170–178] flavor basis with the charged lepton Yukawa diagonal[4] or any other flavor model where the three lepton generations are ruled by a single symmetry group [e.g., [179–182]]. The neutrino sector discussion depends on the specific mechanism assumed to generate masses for the active neutrinos. For example, within the minimal extension of the SM involving the Weinberg operator [183], a consistent MLFV treatment would be to consider the PMNS matrix fully contained in the Wilson coefficient of the Weinberg operator. It follows that the relation between the coefficients c_ψ and c'_ψ is instead highly simplified and $c_\ell = c'_\ell$ and $c_e = c'_e$.

The RGE running introduces corrections to these couplings moving from Λ to the scale of the ALP mass m_a. These effects have been studied at length in the literature [59,60,95], and we avoid reporting here the explicit expressions of the relevant β-functions. Instead, we use the ALPaca tool [120] in order to correctly take into account the RGE effects and conclude that the following textures

[4]Without attempting to identify and construct an explicit UV completion that yields this condition, one may envisage a scenario consistent with the MLFV framework in which the flavor symmetry is broken at a scale $\Lambda' > \Lambda$. In such a setup, a nontrivial ALP-lepton flavor structure can emerge from the action of heavy mediators that remain dynamical down to the scale Λ, where the PQ symmetry is subsequently broken.

are stable under the running; i.e., the quantum corrections turn out smaller or at most of the same order of the present entries. This is due to a combination of two main factors: (i) tree-level couplings with gluons or quarks are vanishing by construction at the scale Λ; (ii) the suppression of the electron and muon masses with respect to the tau mass. It follows that the tau-sector receives the largest corrections, but they have no impact in our analysis. We will come back to this aspect with a more numerical estimate at the end of each section.

1. e − μ framework

When the ALP only couples to the first two generations, the formulas in Eqs. (4.2)–(4.6) may provide some insights into simplifying the multidimensional parameter space. A warning message is in order: these expressions only show the leading contributions, while subleading ones may contain other dependencies on the ALP–lepton couplings. The discussion that follows mainly refers to the dominant contributions, and we will comment on the subleading ones when necessary.

First of all, the RH and LH couplings, c^R and c^L, typically enter expressions for the observables through combinations that are (anti)symmetric under the exchange of c^R and c^L. Consequently, switching off one of the two chiralities or introducing a strong hierarchy between them leads to the same qualitative behavior as assuming identical flavor structures for both. For this reason, we restrict our analysis to the latter scenario.

There are, however, notable exceptions. The magnetic dipole moments of the electron and the muon, as well as the $\mu^- N \to e^- N$ decay, contain terms proportional to the product $c^R c^L$ and, therefore, are sensitive to the relative size of the two chiralities. Nevertheless, assuming a similar flavor structure for c^R and c^L does not limit the generality of our conclusions: a large hierarchy between the two can in general be reabsorbed into a redefinition of the scale f_a, yielding a phenomenologically equivalent setup to the case with c^R and c^L of comparable magnitude.

Focusing now on the internal flavor structure of the couplings, we can discuss the impact of the diagonal and off-diagonal entries. Most expressions present the product of the two kinds of entries in some combination, that is, $c_{ii}^{L/R} \times c_{ij}^{L/R}$. The magnetic dipole moments of both the electron and muon show other types of dependencies: they depend on either a product of only off-diagonal entries or alternatively a combination of only diagonal entries, explicitly or through the effective ALP coupling with photons that is tantamount to saying the difference between the flavor conserving chiral couplings. The simultaneous presence of different combinations may translate into a more difficult understanding of the numerical results with respect to the analytical dependence of the observables from the ALP couplings.

All in all, we can identify three flavor structures that present different phenomenologies. The first is a fully diagonal structure, which would imply that any flavor changing process is extremely suppressed, as indeed the expressions in Eqs. (4.4)–(4.6) identically vanish. The only relevant observables are the magnetic dipole moments of the electron and muon that receive contributions from the photon penguin diagram.

If, instead, we consider a texture with only off-diagonal entries, or with off-diagonal entries very much dominant with respect to the diagonal ones, one may expect a completely different phenomenological scenario. However, this is not the case as the leading contributions of the flavor changing processes are suppressed with respect to those of the magnetic dipole moments. All in all, these two extreme scenarios give rise to very similar constraints on the parameter space. We thus consider the following texture as a benchmark case:

$$\text{Texture 2(a)}: \; c^{L/R} \sim \begin{pmatrix} 0 & \mathcal{O}(1) \\ \mathcal{O}(1) & 0 \end{pmatrix}. \tag{4.20}$$

A more democratic texture with all the entries of the same order of magnitude leads, instead, to a much richer phenomenology, where the flavor changing processes play a relevant role considering the present constraints and the future expected sensitivities:

$$\text{Texture 2(b)}: \; c^{L/R} \sim \begin{pmatrix} \mathcal{O}(1) & \mathcal{O}(1) \\ \mathcal{O}(1) & \mathcal{O}(1) \end{pmatrix}. \tag{4.21}$$

These two textures are stable under the running between the scales we consider in our analysis. To illustrate this point, we show in the Appendix the explicit result of assuming these textures at the high-scale $\Lambda = 10$ TeV and applying the RGE running down to any scale of the ALP mass down to $m_a = 50$ GeV, using ALPaca. In the $e - \mu$ sector, there is no relevant correction that modifies these initial structures: the running contributes to the flavor conserving τ-entry with a value of $\sim 10^{-6}$ in the best scenario. The latter may represent a dangerous outcome as the τ-contributions are generically enhanced by m_τ/m_μ or m_τ/m_e to some powers, but a closer look at the expressions in Eqs. (4.9)–(4.19) for the three-family framework reveals that the $c_{\tau\tau}^{L,R}$ couplings always appear proportional to the flavor violating τ-entries, $c_{e\tau}^{L,R}$ or $c_{\mu\tau}^{L,R}$, which receive contributions from the running that are even more suppressed.[5] As a consequence, these contributions are safely subleading with respect to those in Eqs. (4.2)–(4.6) involving exclusively the ALP-lepton couplings with the first two generations.

[5]The one-loop contributions to these flavor violating entries are suppressed by the neutrino masses and are thus negligible. The subsequent contributions appear at two-loop level and thus also very much suppressed.

TABLE II. $e-\mu$ framework. Dominant contribution to $\hat{Y}'(m_a)$, performing the RGE running from $\Lambda = 10$ TeV down to $m_a \in [100, 400]$ GeV, using ALPaca. The initial values of the couplings assume the LH and RH couplings to have opposite (equal) signs (positive for LH, negative for RH) for the diagonal (off-diagonal) terms.

	Dominant $\hat{Y}'(m_a)$
Texture 2(a)	$(\hat{Y}'_d) = 0, (\hat{Y}'_u) = 0, (\hat{Y}'_e)^{12} \sim (\hat{Y}'_e)^{21} \sim 10^{-4}$
Texture 2(b)	$(\hat{Y}'_d)^{33} \sim 10^{-6}, (\hat{Y}'_u)^{33} \sim 10^{-4}, (\hat{Y}'_e)^{22} \sim 10^{-3}$

In the Appendix, we also report the dominant contributions to the low-energy hadronic couplings. These correspond to the flavor-conserving interactions of the ALP with the top and bottom quarks and are sufficiently suppressed to produce no appreciable impact on hadronic observables. This can be verified by combining the effective analysis of Eq. (3.12) with the numerical results presented in Table II and implementing them in FLAVIO [184]. While this discussion does not provide an exhaustive treatment of hadronic phenomenology, as, for instance, it does not include observables induced at one loop, it nonetheless offers a sound justification for neglecting hadronic constraints when considering these textures.

2. Three-family framework

Once the ALP couples to the three lepton generations, more observables can be considered and the relevant expressions are Eqs. (4.9)–(4.19). As for the previous case, the couplings c^R and c^L are both present, and the case of hierarchies among them provides the same qualitative phenomenology as the case with similar magnitudes for many observables. Indeed, although considering the extreme case where either $c^R = 0$ or $c^L = 0$, leading to several contributions to vanish, the presence of the photon penguin terms implies just a $\sim 10^{-2}$ suppression with respect to the case with both sizable couplings. This suppression is generically easily reabsorbed into a redefinition of the scale f_a.

If we now focus on the flavor structure of these textures, the flavor diagonal couplings enter either explicitly or through the effective ALP-photon coupling. Although the latter are suppressed by powers of $\alpha_{\rm em}$, they turn out to be dominant for large m_a values. The reason can be easily appreciated, for example, looking closer at Eq. (4.9). The purely axion-mediated contribution in the first line is proportional to m_τ^2/m_a^2, while the photon penguin in the second line is proportional to $\alpha_{\rm em} m_e/(\pi m_\tau)$ with the m_a-dependence entering only in the logarithm: the latter contribution mildly varies with m_a, while the first scales quadratically with m_a.

It is then interesting to analyze to which extent the off-diagonal entries provide specific phenomenological signatures. Therefore, in the following, we will consider the next textures.

$$\text{Texture 3(a)}: \; c^{L/R} \sim \begin{pmatrix} \mathcal{O}(1) & \mathcal{O}(1) & \mathcal{O}(1) \\ \mathcal{O}(1) & \mathcal{O}(1) & \mathcal{O}(1) \\ \mathcal{O}(1) & \mathcal{O}(1) & \mathcal{O}(1) \end{pmatrix} \tag{4.22}$$

$$\text{Texture 3(b)}: \; c^{L/R} \sim \begin{pmatrix} \mathcal{O}(1) & \mathcal{O}(1) & \lesssim \mathcal{O}(10^{-3}) \\ \mathcal{O}(1) & \mathcal{O}(1) & \lesssim \mathcal{O}(10^{-3}) \\ \lesssim \mathcal{O}(10^{-3}) & \lesssim \mathcal{O}(10^{-3}) & \lesssim \mathcal{O}(10^{-3}) \end{pmatrix} \tag{4.23}$$

$$\text{Texture 3(c)}: \; c^{L/R} \sim \begin{pmatrix} \lesssim \mathcal{O}(10^{-4}) & \lesssim \mathcal{O}(10^{-4}) & \lesssim \mathcal{O}(10^{-4}) \\ \lesssim \mathcal{O}(10^{-4}) & \mathcal{O}(1) & \mathcal{O}(1) \\ \lesssim \mathcal{O}(10^{-4}) & \mathcal{O}(1) & \mathcal{O}(1) \end{pmatrix} \tag{4.24}$$

$$\text{Texture 3(d)}: \; c^{L/R} \sim \begin{pmatrix} \mathcal{O}(1) & \lesssim \mathcal{O}(10^{-4}) & \mathcal{O}(1) \\ \lesssim \mathcal{O}(10^{-4}) & \lesssim \mathcal{O}(10^{-4}) & \lesssim \mathcal{O}(10^{-4}) \\ \mathcal{O}(1) & \lesssim \mathcal{O}(10^{-4}) & \mathcal{O}(1) \end{pmatrix} \tag{4.25}$$

Texture 3(a) [Eq. (4.22)] represents a completely democratic scenario. The elements in the diagonal are taken as $\mathcal{O}(1)$, but smaller values would lead to the same phenomenological results, according to the previous discussion.

Modifying the previous texture to suppress the couplings with τ would converge toward texture 2(b), as expected. Textures of the type of 3(b) [Eq. (4.23)] will then be discussed together with texture 2(b) and referred to as τ-phobic cases.

A third type of flavor benchmark texture that may be considered is the e-phobic case or texture 3(c) [Eq. (4.24)], where the $\mu-\tau$ sector has the largest couplings with the ALP. In this case, all contributions to the observables that involve an electron are then extremely suppressed, and the only observables relevant to the analysis are expected to be $a_\mu^{\rm ALP}$ and the $\tau \to \mu\gamma$, $\tau \to 3\mu$, and $\tau^- \to \mu^- e^+ e^-$ decays. These observables have however lower sensitivities, and therefore, the analysis is less constraining.

The last trivial modification of texture 3(a) is when the muon sector is suppressed, that is, a μ-phobic scenario, corresponding to texture 3(d) [Eq. (4.25)]. This is complementary to the previous case, where contributions to observables involving a muon are extremely suppressed, and the relevant processes are expected to be $a_e^{\rm ALP}$ and the $\tau \to e\gamma$, $\tau \to 3e$, and $\tau^- \to e^-\mu^+\mu^-$ decays. Besides the muon magnetic dipole moment, the other observables are not well known and their impact on the numerical analysis should be weak.

Anticipating the results from the numerical analysis for textures 3(c) and 3(d), the expected subleading contributions to $\mu \to e\gamma$ and $\mu^- N \to e^- N$ are not so negligible and have an impact in the plots, although in many cases not so evident.

As for the two-family case, these coupling matrices are assumed at the high-scale Λ and one should carefully check their stability under the RGE running from that scale down to the ALP mass. By adopting the ALPaca code, we have done this exercise and found out that the largest contributions are smaller or at most equal to the present entries. Moreover, once again, we see that the induced ALP couplings to third-generation quarks are tiny, thus leading to a negligible hadron phenomenology. Appendix contains the corresponding discussion and shows the explicit results with $\Lambda = 10$ TeV and $m_a \in [50, 400]$ GeV.

V. PHENOMENOLOGICAL ANALYSIS

We have performed a χ^2 analysis of all relevant observables at the 3σ level for the different coupling textures, taking into account the observables listed in the previous section. Given that the two determinations of $(g-2)_e$ are not mutually consistent, we begin with a conservative approach based on the rubidium determination, and subsequently address the anomaly implied by the cesium result.

Considering the number of parameters involved, m_a, Λ, $c_{ij}^{L,R}$, we assume real Wilson coefficients to simplify the analysis, and moreover, we do not present our results in terms of a global fit. Instead, we marginalize with respect to two parameters at a time, fixing the others to specific values. The aim is to illustrate specific features of the textures with two-dimensional plots where the allowed parameter space accounts for the present experimental constraints and the future prospects.

In doing that, we avoid *ad hoc* cancellations among different contributions, by assigning specific values and signs to the fixed parameters in each plot. For example, the flavor conserving couplings enter most of the time as a difference between the LH and RH ones, $(c_{ii}^L - c_{ii}^R)$: to prevent these terms from vanishing, we generically assume these coefficients to be equal in modulus but with opposite sign, $c_{ii}^L = -c_{ii}^R$.

Moreover, as is common in an EFT analysis, bounds do not apply individually on the Lagrangian coefficients or on the scale, that is, $c^{L,R}$ and f_a or m_a in our case, but on their ratios, namely, $c^{L,R}/f_a$ or $c^{L,R}/m_a$. This is why most analyses are done by fixing the value of the scale and investigating the coefficient parameter space, or, on the contrary, by fixing the coefficients to specific values and studying the bounds on the scale. We will follow both approaches.

Before proceeding with the discussion of the textures, we present a study of the $m_a \times f_a$ parameter space, fixing all the coefficients as discussed above. The result can be seen in Fig. 1. The plot on the left is obtained in the specific case of texture 3(a), by fixing the coefficients as $c^L = -c^R = 1$ for both the flavor conserving and violating entries. The darker blue region is excluded at 3σ considering the present experimental bounds, while the lighter one also includes future prospects. The gray region instead is excluded due to a theoretical consistency condition: $m_a < 4\pi f_a$, assuming in this case the cutoff to be $\Lambda = 4\pi f_a$, which would reflect a strongly interacting theory. The red line refers to the condition $m_a = f_a$ and only serves to divide the two regions of the parameter space, where m_a is larger or smaller than f_a.

The processes that dominate the fit are $\mu \to e\gamma$ for the present available data, and $\mu N \to eN$ for future prospects. As expected from the previous expressions of these observables, we can appreciate the effect of two competing tendencies. On the one hand, the contributions to the photon penguin are dominant over almost the entire parameter space except for the spike, where they (partially) vanish. Their dependence on the ALP mass appears exclusively in the logarithmic terms and is suppressed by $(\alpha_{\rm em}/f_a^2)^n$. On the other hand, the direct charged lepton contributions, which are proportional to $1/(m_a f_a)^{2n}$, are relevant only in the spike region and for low values of m_a. All in all, the unconstrained parameter space is very large and *there is no experimental preference for the allowed white regions with* $m_a < f_a$ *or with* $m_a > f_a$. The latter is only further constrained by the theoretical consistency limit that starts to dominate for $m_a \gtrsim 130$ TeV ($m_a \gtrsim 280$ TeV) for current (future) prospects. The viability of this region of the parameter space is typically overlooked by the prejudice

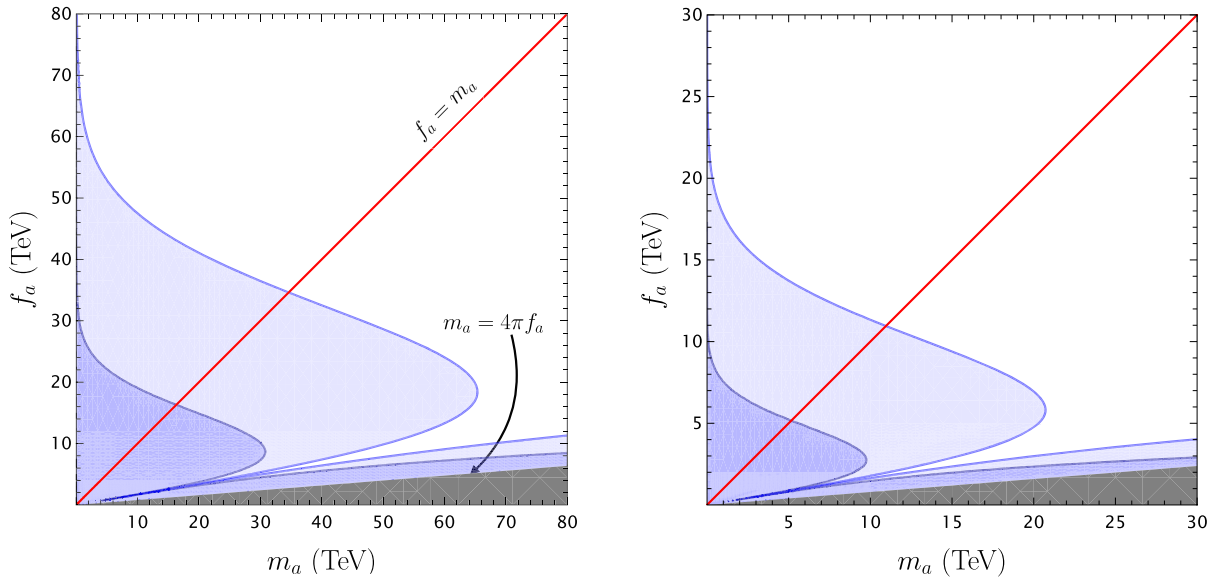


FIG. 1. $m_a \times f_a$ parameter space for texture 3(a), having fixed all the Lagrangian coefficients to $c^L = -c^R = 1$ (left) and to $c^L_{ii} = -c^R_{ii} = 1$ and $c^L_{ij} = -c^R_{ij} = 0.1$ (right). The darker (lighter) blue region denotes the parameter space excluded by present experimental constraints (future prospects), at 3σ. The red line represents the condition $m_a = f_a$. The gray region is excluded by the theoretical consistency condition $m_a < 4\pi f_a$.

of assuming an ALP with the same hierarchical relation between its mass and the scale f_a as for the traditional QCD axion, i.e., $m_a \ll f_a$.

The plot on the right closely resembles the one on the left, with the only difference being the choice of flavor violating couplings, which are now fixed to $c^L_{ij} = -c^R_{ij} = 0.1$. Although this setup no longer corresponds to texture 3(a), it is nevertheless useful to illustrate how the exclusion region depends on the overall magnitude of the flavor violating entries. Since the dominant observables scale as $|c^L_{ij}|^2|c^R_{ij}|^2$, lowering $c^{L,R}_{ij}$ by one order of magnitude suppresses the exclusion bounds by a factor of 3. More importantly, this modification does not alter the shape of the excluded region; it simply produces a uniform rescaling in the $m_a \times f_a$ plane. However, for textures with even smaller off-diagonal entries, the relative relevance among the observables may completely change. Indeed, observables with a stronger dependence on the flavor conserving charged lepton couplings may provide the strongest constraints. For example, certain tau-lepton decay modes may also become relevant through their dependence on $c^{L,R}_{\tau\tau}$, although their constraints remain systematically weaker than those from $\mu \to e\gamma$, which continues to provide the dominant bound. At the same time, such a strong suppression would push the exclusion limit of both f_a and m_a to very low scales, entering the region where our description breaks down and thus no conclusion can be deduced.

All in all, in view of the above considerations, we refrain from performing a systematic scan over the coupling values. Instead, we concentrate on textures with genuinely distinct internal structures that cannot be mapped onto each other through a mere rescaling of their entries.

Figure 2 illustrates the same parameter space as in Fig. 1, but for textures 2(a) (left) and 2(b) (right). First of all, comparing the four plots, we can notice that the experimental bounds are much weaker in the case of texture 2(a), presenting even a different profile shape. The difference in scales arises from the contributions of $c^{\text{eff}}_{a\gamma\gamma}$, which are absent in texture 2(a). Textures 2(b) and 3(a) present a similar scale, with the relevant difference being only the contribution from $c_{\tau\tau}$ inside $c^{\text{eff}}_{a\gamma\gamma}$. In the regime of the spike where the explicit charged lepton contribution is relevant, we can appreciate the effect of the τ mass enhancement: when τ couplings are active, the corresponding contributions are enhanced by m_τ/m_μ terms and, correspondingly, the values of m_a and f_a have to be larger in order to saturate the $\mu \to e\gamma$ and $\mu^- N \to e^- N$ constraints that dominates the experimental bounds.

More into details, in the plot on the left for texture 2(a), the present bounds are mainly given by the anomalous $g-2$ of the muon and the future prospects from charged lepton decays are not visible as they are weaker. For this texture, the induced ALP-photon coupling is extremely suppressed and thus the photon penguin does not provide interesting contributions. We remain with the pure charged lepton contributions and their $1/(m_a f_a)^4$ dependence. Instead, in the plot on the right for texture 2(b), the present constraints arise from $\mu \to e\gamma$ decay and the future prospects from $\mu^- N \to e^- N$, and in particular the effective

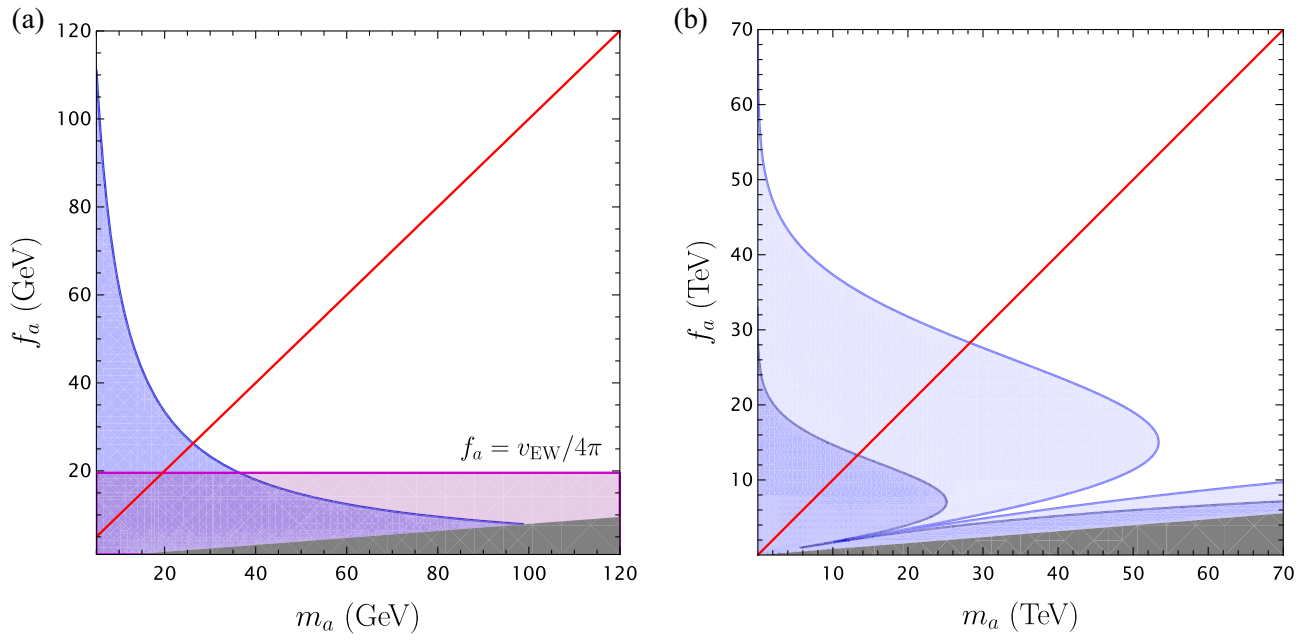


FIG. 2. $m_a \times f_a$ parameter space for textures 2(a) on the left and 2(b) on the right, having fixed all the Lagrangian coefficients to $c^L = -c^R = 1$. The darker (lighter) blue region denotes the parameter space excluded by present experimental constraints (future prospects) at 3σ. The red line represents the condition $m_a = f_a$. The gray and purple regions are excluded by the theoretical consistency conditions $m_a < 4\pi f_a$ and $f_a < v_{\rm EW}/4\pi$, respectively.

photon coupling is sizable and consequently the photon penguin contribution is dominant.

For texture 2(a), there are values of the ALP mass for which the values of f_a already fall below the scale of validity of our description, $f_a \sim 30$ GeV. Although this plot strictly holds for $c^L_{\mu e} = 1 = -c^R_{\mu e}$, slight changes to these values do not lead to different results, and therefore, we expect it to be unbounded from the lepton observables considered. Texture 2(b), instead, is perfectly consistent with our EFT treatment. As for the previously discussed texture 3(a), there is no experimental preference for any of the two regions separated by the red line.

In what follows, for each texture, we will include plots that illustrate the possible existing correlations between different ALP-lepton couplings: in most of the cases, the behavior that we can naively infer from the formulas associated with the most relevant observables can be easily identified in the plots. However, in some cases, the expected leading contributions are not so dominant and/or observables with weak sensitivities also play a relevant role in the χ^2 analysis, and consequently the understanding of the parameter space is more involved. Moreover, we will always focus on comparing $m_a < f_a$ and $m_a > f_a$ to underline that both regions deserve to be equally studied.

A. Analysis of the textures

1. Texture 2(a)

Texture 2(a) only presents flavor violating entries, and the relevant observables are $g-2$ of the electron and muon. Indeed, the muon decays and the $\mu^- N \to e^- N$ conversion are suppressed, as they depend on the product of flavor conserving and violating couplings. As the plot on the left of Fig. 2 shows, the scales of m_a and f_a corresponding to the present experimental bounds are extremely low and in part violate the consistency limit of our description, $f_a > v_{\rm EW}/4\pi$.

To better examine the low scales allowed by this texture, Fig. 3 showcases three plots for the parameter space $c^{L,R}_{\mu e}$, assuming an extreme value for $f_a = 30$ GeV and progressively increasing values of m_a. The relevant aspect to be underlined refers to the values of the coefficients: the region for $|c^{L,R}_{\mu e}| \leq 1$ is always unconstrained, for both present data and future prospects. Larger coefficient values, $c^{L,R} > 1$, would simply correspond to a redefinition of f_a to even smaller values, and thus, it would not represent a qualitatively different setup. We can therefore conclude that the suppression of the ALP-lepton couplings in this case, due to their proportionality to m_e and m_μ, leads to an undetectable ALP in any of the leptonic observables considered, even in the near future.

2. Texture 2(b)

This texture presents a democratic structure in the $e-\mu$ sectors, while the τ-couplings are suppressed. The only relevant observables are $g-2$ of the electron and muon, and the decays of the muon. The most present experimental constraint is due to $\mu \to e\gamma$, while the dominant future prospect is from $\mu^- N \to e^- N$. Indeed, this is reflected in the parameter space of the plots in Fig. 4. In order to

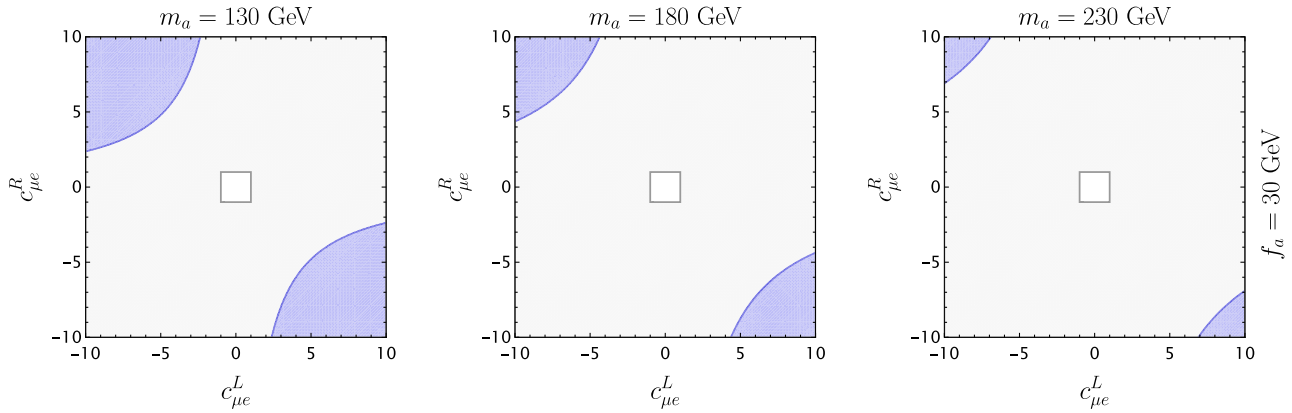


FIG. 3. Texture 2(a). $c^L_{\mu e} \times c^R_{\mu e}$ parameter space for fixed $c^{L,R}_{ii} = 0$. $f_a = 30$ GeV in all the plots. $m_a = 130(180)[230]$ GeV in the plot on the left (center) [right]. The colored regions are excluded at 3σ by the current experimental bounds. The gray region corresponds to $|c^{L,R}_{\mu e}| > 1$.

understand the contours of the darker (lighter) blue regions, we can focus on Eq. (4.4) [Eq. (4.6)]. In particular, in the plots in the top panel, we can identify the circular profile of the blue regions corresponding to the $(|c^L_{\mu e}|^2 + |c^R_{\mu e}|^2)$ combination within the $\mathcal{BR}(\mu \to e\gamma)$ and $\mathcal{BR}(\mu^- N \to e^- N)$ expressions.

The profiles in the plots in the middle and bottom panels can be explained by focusing on the flavor conserving couplings in Eqs. (4.4) and (4.6), namely, the two relevant observables for the present constraints and future prospects, respectively. Recalling the definition of the effective photon coupling, the focus is on the combination $(c^L_{\mu\mu} - c^R_{\mu\mu})^2$: fixing $c^L_{\mu\mu} = 1$ implies the possibility of a partial cancellation of this combination upon variation of $c^R_{\mu\mu}$. The value of the observables Eqs. (4.4) and (4.6) becomes suppressed, thus reducing the constraints on the corresponding parameter space. This behavior is obtained when $c^R_{\mu\mu} \to 1$; however, when it turns negative, no cancellation is in action and the parameter space becomes excluded by both the present bounds and the future prospects.

The plots are presented for a specific value of f_a chosen based on the plot in Fig. 2 in order to be close to the region where experimental bounds are relevant. Three values of m_a are showcased in order to illustrate the variation of the parameter space for the given couplings as a function of m_a. While in the plots in the top row this dependence is trivial, in the others it is more involved: we observe how progressively increasing the masses (from left to right) opens up the available parameter space. The scaling of the bounds agrees with the expected behavior dominated by the photonic penguin diagram contribution.

3. Texture 3(a)

This texture is a natural generalization of the previous texture to the case of three generations and presents a completely democratic structure among its entries. Since the τ contribution is switched on, τ-decays can have a nonvanishing contribution, which is enhanced by m_τ/m_μ with respect to the muon observables, although the former has lower precision than the latter. Performing our numerical analysis, the processes that yield the strongest present and future bounds are, respectively, $\mu \to e\gamma$ and $\mu N \to eN$.

Therefore, as in texture 2(b), the dominant contributions arise from the couplings $c^{L,R}_{\mu e}$ as well as the diagonal ones, due to the enhancement of the photon penguin terms. In Fig. 5, we provide the same plots as in Fig. 4, noticing the difference in the scale as a result of the contribution of $c_{\tau\tau}$ through $c^{\text{eff}}_{a\gamma\gamma}$ as already stated before.

For this texture, we also consider the case where all the couplings of one chirality vanish. In this case, the only surviving contributions are those proportional to products of couplings with the same chirality, such as those in the second line of Eq. (4.11) for $\mathcal{BR}(\mu \to e\gamma)$. The plots in Fig. 6 explicitly show how the parameter space varies switching off the c^R coupling: in the higher panel, $c^R = -1$, while in the lower one $c^R = 0$. The complementary case with $c^L = 0$ is equivalent. The top panel is shown at the same scale as Fig. 5, $f_a = 14$ TeV, whereas the chiral bottom panel uses $f_a = 7$ TeV. The latter reflects the weaker bounds after the opposite-chirality contributions are removed.

4. Texture 3(b)

Adopting the same coupling structure as in the previous texture 3(a), a suppression of $\mathcal{O}(10^{-3})$ in the tau sector effectively reduces the theory to a two-lepton scenario where only ALP couplings with electrons and muons are relevant. The processes that provide the strongest constraints on the parameter space remain the same as in the previous textures 2(b) and 3(a), and the plots in Fig. 4 also apply to the texture 3(b) case.

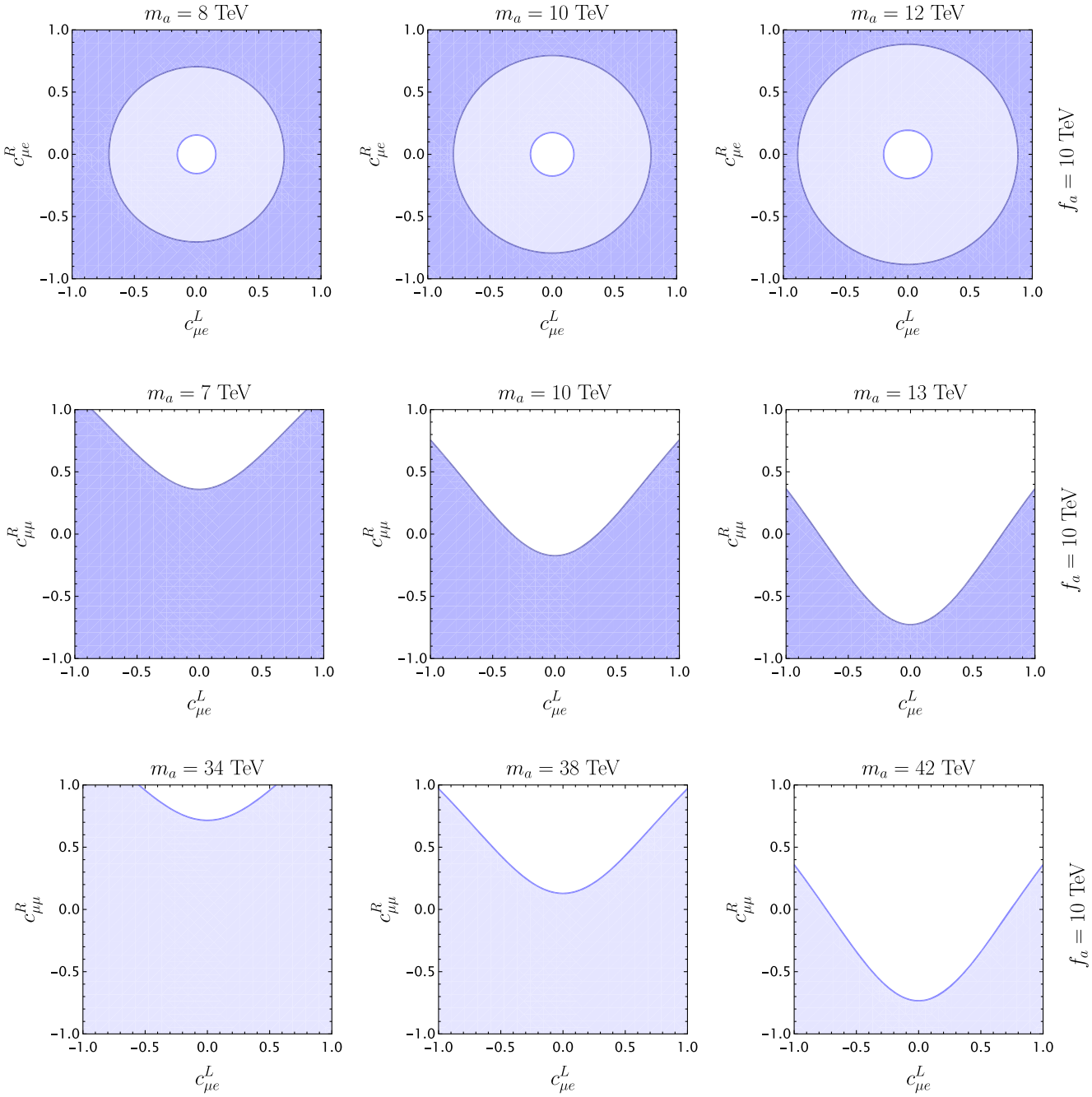


FIG. 4. Texture 2(b). Top row: $c^L_{\mu e} \times c^R_{\mu e}$ parameter space for fixed $c^L_{ii} = 1 = -c^R_{ii}$. Middle and bottom rows: $c^L_{\mu e} \times c^R_{\mu\mu}$ parameter space for fixed $c^L_{ii} = 1 = -c^R_{ij}$. In all three cases, the chosen values of m_a (on top of the plots) increase from left to right. $f_a = 10$ TeV in all the plots. The plots showcase the 3σ exclusion regions for the current (future) experimental bounds in darker (lighter) blue.

The contribution from diagonal ALP-τ couplings represents no major impact on $c^{\text{eff}}_{a\gamma\gamma}$, making this texture essentially equivalent to texture 2(b). On that line, it is useful to comment that if the τ-entries were $\sim\mathcal{O}(10^{-1})$, the contribution to $c^{\text{eff}}_{a\gamma\gamma}$ would be able to represent an increment in the scale with respect to texture 2(b). On the contrary, if these entries were an order of magnitude smaller ($\sim\mathcal{O}(10^{-4})$), the corresponding texture would receive contributions from the RGE running, effectively leading at low-energy to the original texture 3(b).

5. Texture 3(c)

This analysis also focuses on a two-family framework, suppressing all terms related to the electron by four orders

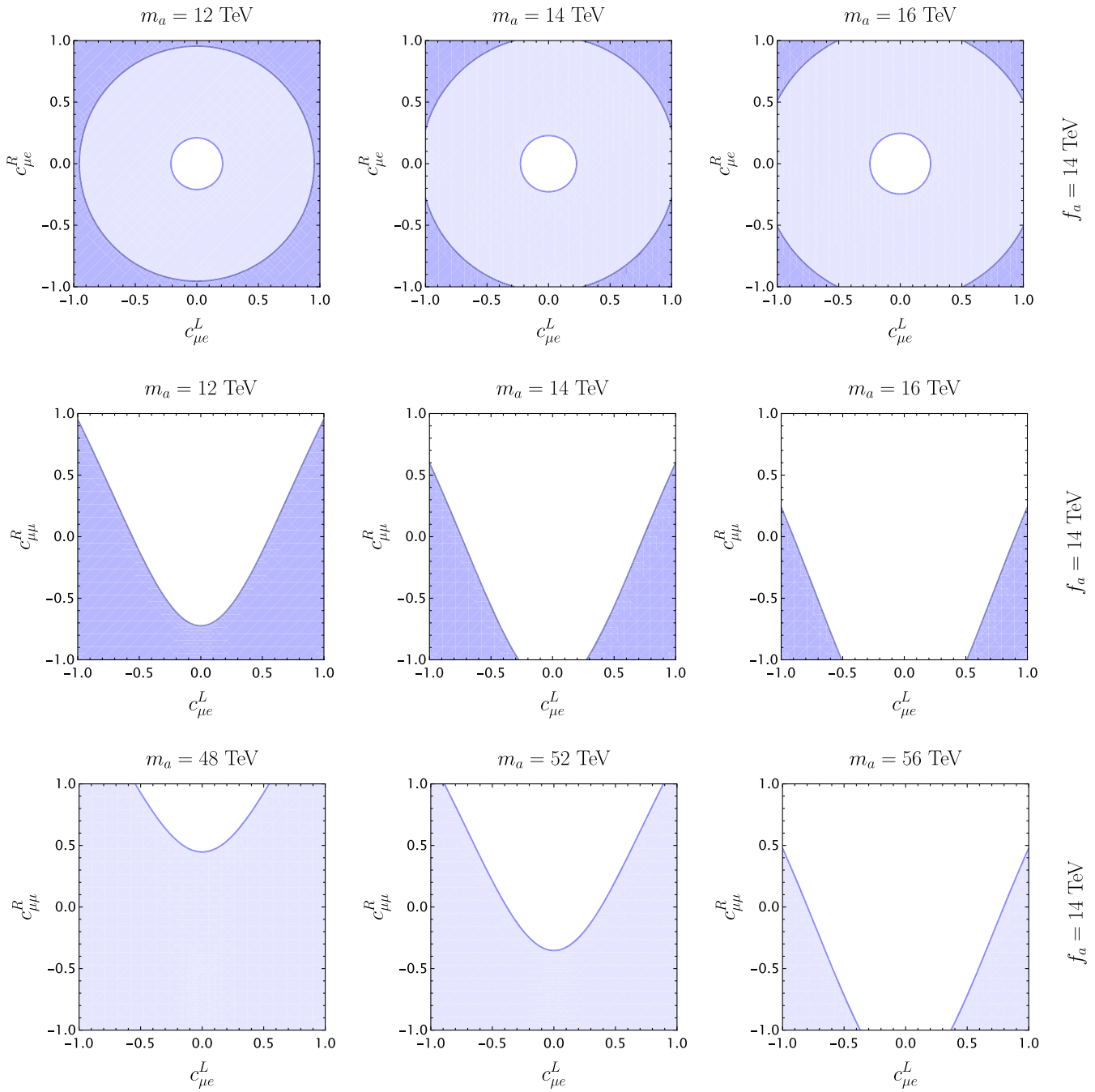


FIG. 5. Texture 3(a). Top row: $c^L_{\mu e} \times c^R_{\mu e}$ parameter space for fixed $c^L_{ii} = 1 = -c^R_{ii}$. Middle and bottom rows: $c^L_{\mu e} \times c^R_{\mu\mu}$ parameter space for fixed $c^L_{ii} = 1 = -c^R_{ij}$. In all three cases, the chosen values of m_a (on top of the plots) increase from left to right. $f_a = 14$ TeV in all the plots. The plots showcase the 3σ exclusion regions for the current (future) experimental bounds in darker (lighter) blue.

of magnitude compared to the other terms. This choice is made intentionally, as the process $\mu \to e\gamma$ is no longer the most constraining, having been surpassed by $\tau \to \mu\gamma$ for present bounds. For future prospects, the process $\mu^- N \to e^- N$ improves significantly with respect to the current situation, but still cannot provide the strongest bound, although it competes with $\tau \to \mu\gamma$.

The dominant contributions arise from the couplings $c^{L,R}_{\tau\mu}$ and $c_{\tau\tau}$ due to the enhancement of the photon penguin terms. However, due to the suppression of the muon observables and the lesser precision in the τ ones, the scale is expected to be lower than the analog texture 3(b).

In the plots in Fig. 7, we illustrate the parameter space $c^L_{\tau\mu} \times c^R_{\tau\mu}$ on the higher panel, and $c^L_{\tau\mu} \times c^R_{\tau\tau}$ on the lower

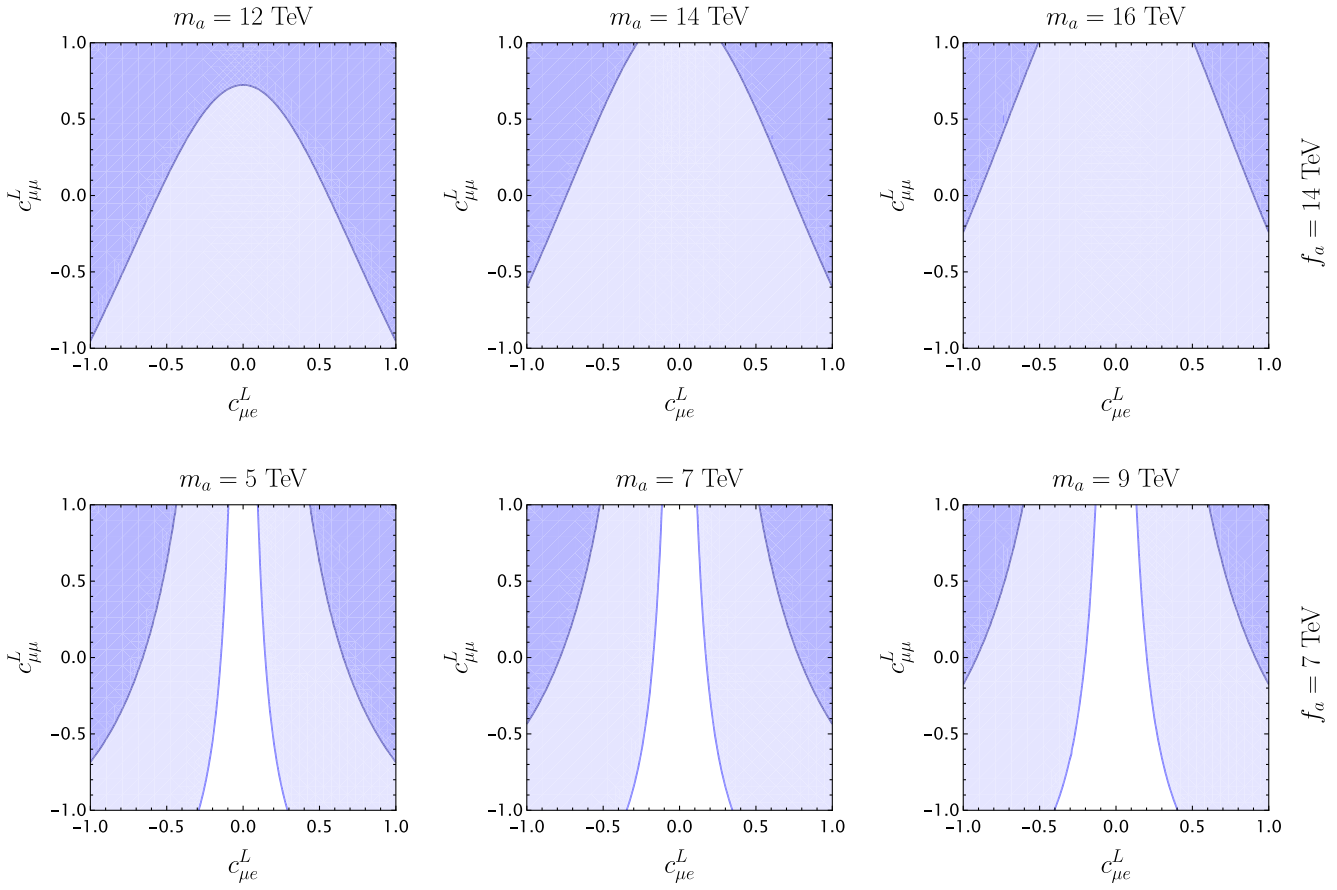


FIG. 6. Texture 3(a). $c^L_{\mu e} \times c^L_{\mu\mu}$ parameter space (for fixed $c^L_{ii} = 1 = -c^R_{ii}$). Top row: $c^R = -1$ and $f_a = 14$ TeV. Bottom row: $c^R = 0$ and $f_a = 7$ TeV. In all three cases, the chosen values of m_a (on top of the plots) increase from left to right. The plots showcase the 3σ exclusion regions for the current (future) experimental bounds in darker (lighter) blue.

one. The shape of the different plots is equivalent to the ones provided in Fig. 4, only changing the scale, as stated previously. Since the future prospects of $\mu^- N \to e^- N$ can at best compete with $\tau \to \mu\gamma$, the change in shape and scale in both plots is minimal.

The ALP-e couplings have negligible effects in the χ^2, and as a result no meaningful conclusion can be deduced from plots focusing on the $c^L_{ie} \times c^R_{ie}$ parameter space.

As commented for the previous texture, smaller values of the flavor-violating entries would lead to quantum corrections running from the high-energy scale down to m_a, while texture 3(c) is stable under RGE running. Increasing the suppressed entries by an order of magnitude would recover the processes $\mu \to e\gamma$ and $\mu^- N \to e^- N$ as the more constraining ones for present bounds and future sensitivities, respectively.

6. Texture 3(d)

In this texture, all the couplings associated with the muon are $\mathcal{O}(10^{-4})$, representing the complementary framework with respect to textures 3(b) and 3(c).

The discussion of this texture becomes equivalent to the previous one, but replacing $\tau \to \mu\gamma$ by $\tau \to e\gamma$ as the process that yields the strongest constraints. For future prospects, the process $\mu^- N \to e^- N$ can compete with $\tau \to e\gamma$ but becomes less stringent than for $\tau \to \mu\gamma$ in the previous texture.

The dominant contributions arise from the couplings $c^{L,R}_{\tau e}$ and $c_{\tau\tau}$, but are not necessary from the photon penguin terms. In fact, a remarkable contribution for this texture arises from the $(g-2)_e$ observable, as the scales involved make the photonic contribution no longer relevant compared with the nonphotonic ones. That contribution, being $\mathrm{Re}[c^L_{\tau e} c^R_{\tau e}]$ in Eq. (4.9), is responsible for altering the shape of the plots in Fig. 8 compared to the ones in Fig. 7. For future prospects, the lesser constraints of $\mu^- N \to e^- N$ can slightly change the shape and scale in both plots.

The ALP-μ couplings have negligible effects in the χ^2, and therefore, plots illustrating the parameter space $c^L_{\mu i} \times c^R_{\mu i}$ would not provide insightful results.

This texture is also stable under RGE running effects, while entries smaller than $\sim 10^{-4}$ would be affected by

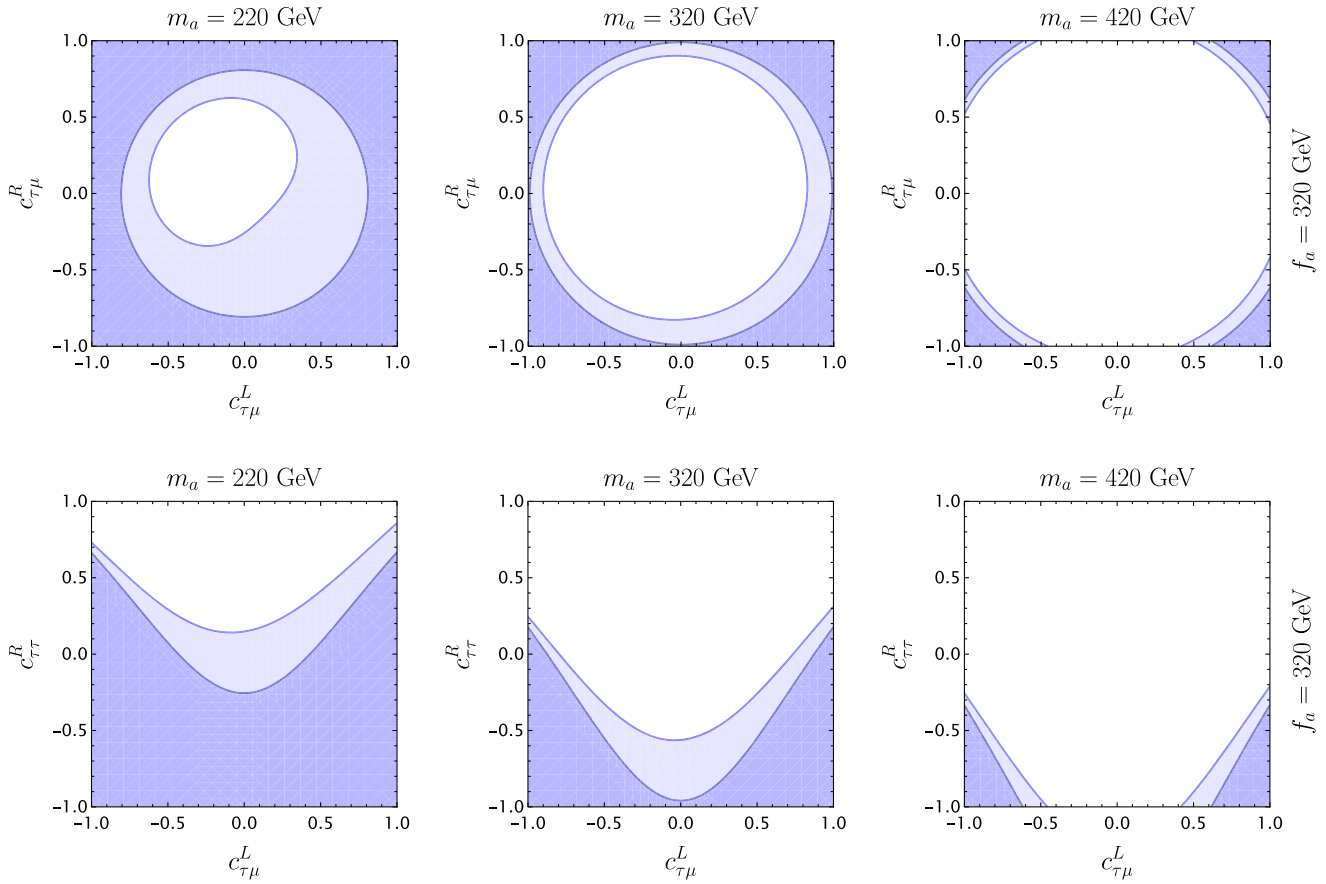


FIG. 7. Texture 3(c). Top row: $c^L_{\tau\mu} \times c^R_{\tau\mu}$ parameter space. Bottom row: $c^L_{\tau\mu} \times c^R_{\tau\tau}$ parameter space. While the couplings that appear in the axes are kept free, those ones that refer to the electron are fixed to $c^L = 10^{-4} = -c^R$ and the rest are fixed to $c^L = 1 = -c^R$. $f_a = 320$ GeV in all the plots and $m_a = 220(320)[420]$ GeV in the left (middle) [right] column. The plots showcase the 3σ exclusion regions for the current (future) experimental bounds in darker (lighter) blue.

quantum corrections. As for texture 3(c), increasing the suppressed entries by an order of magnitude would recover the processes $\mu \to e\gamma$ and $\mu^- N \to e^- N$ as the more constraining ones for present bounds and future sensitivities, respectively.

B. The a_e^{Cs} anomaly

In the previous analysis, we considered the rubidium determination for $(g-2)_e$ that presents a tension with the present experimental measurement of $+2.4\sigma$. This is a more conservative approach, while if we adopt the cesium determination the tension increases up to -3.8σ and thus represents an anomaly. In this section, we discuss whether and under which conditions a leptophilic ALP can explain this tension. As already mentioned in our analysis, which rests on the consistency requirement that $f_a > v_{\mathrm{EW}}/4\pi$, we particularly focus on the $m_a > f_a$ parameter space. This is complementary to other studies that instead focus on the opposite regime; see, for example, Refs. [93,94].

The expression for the dominant ALP contribution to $(g-2)_e$ is Eq. (4.2) in the $e-\mu$ family case and Eq. (4.9) for the three-family framework. Both flavor violating and conserving ALP-lepton couplings enter these expressions. The first ones have a chiral enhancement that is m_μ/m_e and m_τ/m_e, respectively, due to the pure lepton-mediated diagram. This does not occur for the flavor conserving ones, associated to the photon penguin contributions. On the other hand, the latter have a milder m_a dependence and turn out to be dominant at large ALP masses.

The plots in Fig. 9 show our results for texture 2(a) within the e–μ framework (left) and for texture 3(d) within the three-family setup (right). In all panels, the orange region denotes the values of the parameter space $m_a \times f_a$ that resolve the cesium tension at 3σ, assuming $c^L_{\mu e} = 1 = -c^R_{\mu e}$ in the e–μ plot and $c^L_{\tau e} = 1 = -c^R_{\tau e}$ in the three-family one. Present experimental bounds and future sensitivities are shown in blue—notice the barely visible blue region on the lower left corner of the plot on the left.

Starting with the e–μ scenario, an ALP obeying texture 2(a) can indeed account for the cesium anomaly while remaining fully compatible with current and projected constraints. The only additional requirement to keep in

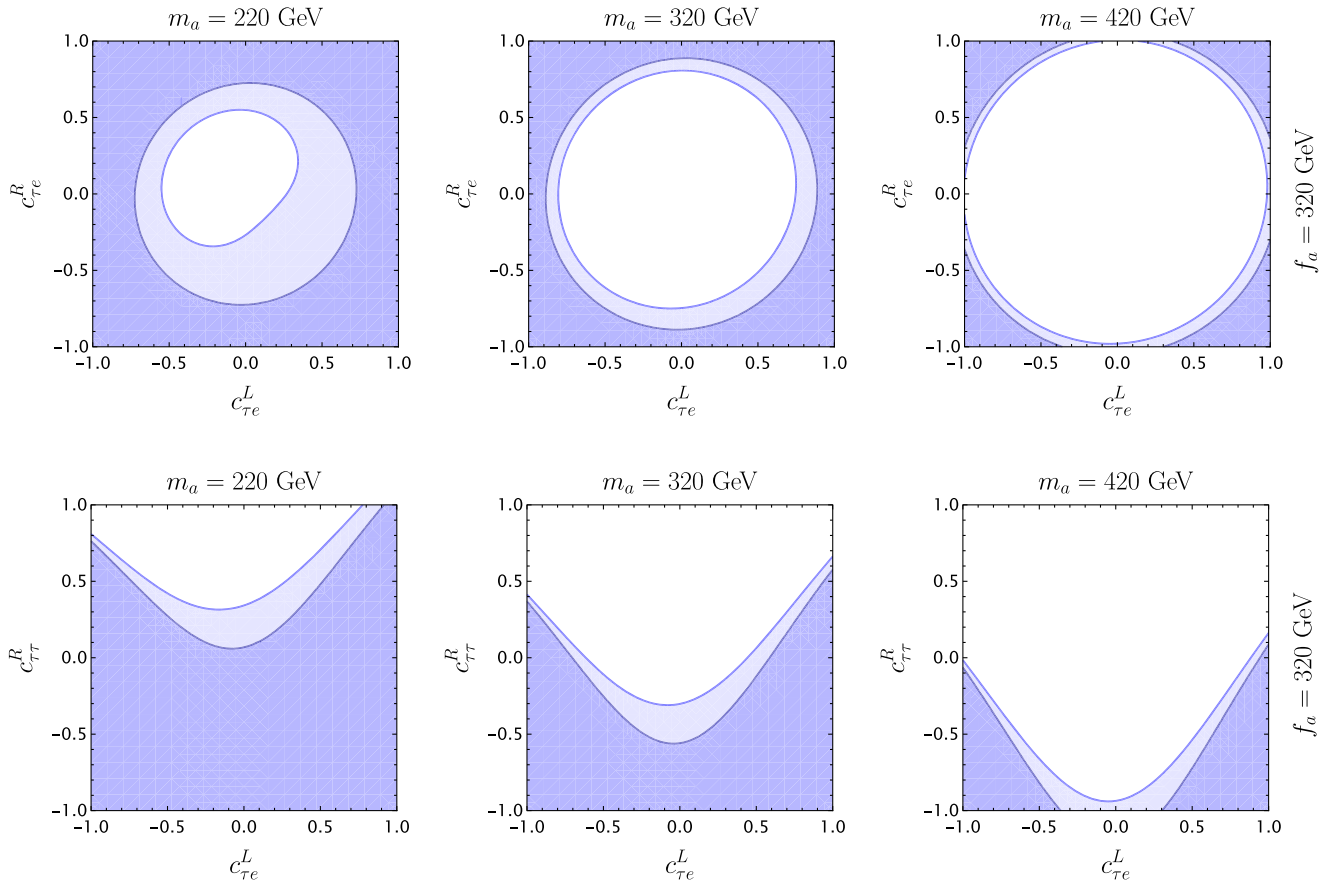


FIG. 8. Texture 3(d). Top row: $c^L_{\tau e} \times c^R_{\tau e}$ parameter space. Bottom row: $c^L_{\tau e} \times c^R_{\tau\tau}$ parameter space. While the couplings that appear in the axes that are kept free, those ones that refer to the muon are fixed to $c^L = 10^{-4} = -c^R$ and the rest are fixed to $c^L = 1 = -c^R$. $f_a = 320$ GeV in all the plots and $m_a = 220(320)[420]$ GeV in the left (middle) [right] column. The plots showcase the 3σ exclusion regions for the current (future) experimental bounds in darker (lighter) blue.

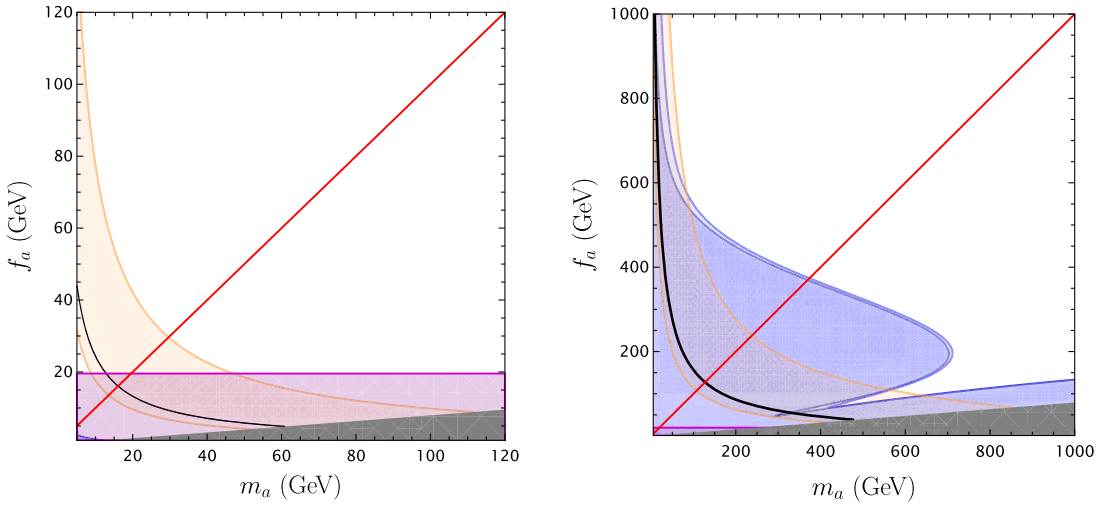


FIG. 9. $m_a \times f_a$ parameter space. The orange region identifies the parameter space where the cesium anomaly can be solved at 3σ, with the black line representing the best fit values. The left plot refers to texture 2(a), while the right plot to texture 3(d). The darker (lighter) blue regions are excluded by the present experimental bound (by the future prospects) at 3σ. The red line represents the condition $m_a = f_a$. The gray and purple regions are excluded by the theoretical consistency conditions $m_a < 4\pi f_a$ and $f_a < v_{\rm EW}/4\pi$, respectively.

mind is $f_a > v_{\rm EW}/4\pi$, which implies that only the region with $m_a < f_a$ is consistent with our effective description. The analogous analysis for texture 2(b) excludes an ALP explanation of the anomaly, as the $\mu \to e\gamma$ limits impose a much stronger lower bound on both m_a and f_a.

Turning now to the three-family case with texture 3(d), we find that present data, dominated by $\tau \to e\gamma$, exclude this solution over a large portion of the parameter space, though not entirely. The point with $m_a \approx 500$ GeV and $f_a \approx 100$ GeV evades the leptonic constraints and lies close to the sensitivity suggested by the current collider estimate; it is therefore a promising target for a dedicated collider search. The alternative with $m_a \approx 50$ GeV and $f_a \gtrsim 700$ GeV reduces the $(g-2)_e$ tension to the 1σ level. In the region $m_a > f_a$, the allowed parameter space arises from the cancellation between the pure lepton mediated contribution and the photon penguin one.

Considering now the other three-generation textures, this ALP solution becomes strongly excluded. Indeed, once the e–μ couplings are switched on, then the $\mu \to e\gamma$ and $\mu^- N \to e^- N$ bounds are much more relevant, thus excluding the region of the parameter space $m_a \times f_a$ interesting for this anomaly.

VI. FINAL REMARKS

In this work, we have revisited one of the most widely used assumptions in the ALP literature, namely, the parametric hierarchy $m_a \ll f_a$. While this relation is often taken as a defining feature of ALP effective descriptions, it is not a fundamental consistency requirement of the EFT framework. The meaningful theoretical constraint is instead $m_a < \Lambda$, where Λ denotes the true EFT cutoff. The comparison between m_a and f_a is therefore model-dependent and, in general, does not carry intrinsic physical significance. Moreover, using NDA, the region $m_a > f_a$ is naturally interpreted as pointing toward a strongly coupled ultraviolet completion.

Motivated by this theoretical reassessment, we have performed a dedicated phenomenological analysis of leptophilic ALPs in the regime $m_a > f_a$, a region that has received little attention in previous studies. Additionally, we have implemented the full RGE from the cutoff scale down to m_a. Integrating out the considered leptophilic ALP and matching onto the SMEFT Warsaw basis, we could conclude that all induced quark operators are strongly suppressed, both by Yukawa insertions and by loop effects. This establishes that the hadronic sector does not play a significant role in the phenomenology under consideration, thereby justifying our focus on purely leptonic observables. On the other hand, we have rederived the one-loop expressions for the anomalous magnetic dipole moments of the electron and muon, as well as for radiative and three-body charged lepton flavor violating decays and μ–e conversion in nuclei, for both the two-family and three-family frameworks. In the latter case, we have shown explicitly how τ-induced contributions can dominate due to the m_τ enhancement, even when the corresponding flavor couplings are comparatively small. The interplay between tree-level, one-loop, and interference terms has been carefully assessed for large ALP mass, resulting in a study complementary to existing literature.

From the phenomenological perspective, our analysis demonstrates that sizable regions of parameter space with $m_a > f_a$ remain viable under current experimental constraints. Moreover, this region can lead to distinctive patterns in charged lepton flavor violation and in the anomalous magnetic moments. In particular, if no specific hierarchy is considered in the ALP-lepton couplings and they are taken as $\mathcal{O}(1)$, then the present experimental constraints, dominated in general by the $\mu \to e\gamma$ process, imply the condition $m_a f_a \gtrsim (17\ \text{TeV})^2$. Otherwise, if the flavor violating couplings in any of the flavor sectors are suppressed by one order of magnitude, this bound is relaxed to $m_a f_a \gtrsim (6\ \text{TeV})^2$. Considering future prospects, dominated by the $\mu^- N \to e^- N$ conversion, these conditions become $m_a f_a \gtrsim (35\ \text{TeV})^2$ and $m_a f_a \gtrsim (11\ \text{TeV})^2$. We have also considered specific benchmark scenarios for the ALP-lepton couplings to investigate the dependence of the parameter space on the flavor violating ALP couplings. Specifically, we focused on electron-phobic, muon-phobic, and tau-phobic frameworks, assessing the stability of the assumed textures under RGE running and clearly concluding that the region of parameter space with $m_a > f_a$ is very much unconstrained under both present experimental data and future prospects.

Furthermore, we have identified the conditions under which the existence of an ALP can accommodate the cesium determination of $(g-2)_e$, which currently presents a -3.8σ tension with the corresponding experimental measure. In the two-family case of a tau-phobic ALP, we found that only by suppressing the flavor conserving ALP coupling to the muon can the ALP hypothesis explain the present tension, as in this way it is possible to evade the $\mu \to e\gamma$ and $\mu^- N \to e^- N$ limits. The values of the ALP mass and decay constant f_a have to satisfy the relation $(10\ \text{GeV})^2 \lesssim m_a f_a \lesssim (30\ \text{GeV})^2$, together with the consistency condition $f_a > v_{\rm EW}/4\pi$. In the three-family scenario, we highlighted the fundamental role played by the flavor violating couplings $c_{\tau e}^{L,R}$, and by the relative chiral structure of the ALP-lepton interactions with their proportionality to the lepton masses. We concluded that a leptophilic ALP can indeed solve the anomaly in the cesium determination of $(g-2)_e$ if its couplings to muons are suppressed with respect to those with taus and electrons. The low-f_a point at $m_a \approx 500$ GeV and $f_a \approx 100$ GeV is a promising target for a dedicated collider search, while the solution with $m_a \approx 50$ GeV and $f_a \gtrsim 700$ GeV provides a complementary high-f_a possibility.

Overall, our results show that the region $m_a > f_a$ is not only theoretically consistent but also phenomenologically

rich and experimentally testable. The widespread identification of f_a with the EFT cutoff has led to an overly restrictive exploration of ALP parameter space. Relaxing this prejudice opens new avenues both for model building and for experimental searches, particularly in precision leptonic observables.

Abandoning the leptophilic scenario considered here, one has to take into consideration collider bounds associated with the ALP-gauge field couplings and ALP-top interactions, and the various constraints arising from the hadronic sector. These studies are complementary with respect to our analysis and are deferred to a future investigation.

ACKNOWLEDGMENTS

We are grateful to Emmanuel Stamou, Enrique Fernández Martínez, and Pablo Quílez for valuable discussions. The work of M. F. Z. is supported by the Spanish MIU through the National Program FPU (Grant No. FPU22/03625). The work of A. L.-O. is supported by the FPI Grant No. PRE2022-105383 funded by MCIN/AEI/10.13039/501100011033 and ESF+. The work of S. R. H. is supported by the FPI Grant No. PREP2022-000805 funded by MCIN/AEI/10.13039/501100011033. We acknowledge partial financial support by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 101086085-ASYMMETRY and by the Spanish Research Agency (Agencia Estatal de Investigación) through the Grant IFT Centro de Excelencia Severo Ochoa No. CEX2020-001007-S and by the Grant No. PID2022-137127NB-I00 funded by MCIN/AEI/10.13039/501100011033. This article is based upon work from COST Action COSMIC WISPers CA21106, supported by COST (European Cooperation in Science and Technology).

DATA AVAILABILITY

The Wolfram *Mathematica* notebook containing the analytical expressions for the observables used in this work, together with the assumptions employed in their derivation, is available in Supplemental Material [137].

APPENDIX: RUNNING AND MATCHING OF THE ALP COUPLINGS

In Sec. IV C, we discuss the different coupling textures [see Eqs. (4.20)–(4.25)] used in our phenomenological analysis. A question one might have is whether the running would noticeably impact the couplings, thus completely altering the coupling texture at low energies, at which the processes of our analysis are computed and compared with experiment. Here we will show that this is not the case. For this, we run the couplings from a set scale $\Lambda = 10$ TeV, corresponding to the highest scale present in our analysis, down to the mass of the ALP, $m_a \in [50, 400]$ GeV, using ALPaca [120]. For energy scales below the ALP mass, where the ALP is already integrated out, we do not account for the EW running, since its contributions are negligible. In the following, we report the order of magnitude of the couplings associated with the different textures. As initial values, we assume that the diagonal elements of c^L and c^R have opposite signs, while for the off-diagonal elements we assume them to be equal.

$$\text{Texture 2(a)}:\ c^{L/R} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{\text{RGE}} |c^{L/R}| \sim \begin{pmatrix} 0 & \mathcal{O}(1) & 0 \\ \mathcal{O}(1) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{A1}$$

$$\text{Texture 2(b)}:\ c^{L/R} = \begin{pmatrix} \pm 1 & 1 & 0 \\ 1 & \pm 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{\text{RGE}} |c^{L/R}| \sim \begin{pmatrix} \mathcal{O}(1) & \mathcal{O}(1) & 0 \\ \mathcal{O}(1) & \mathcal{O}(1) & 0 \\ 0 & 0 & \mathcal{O}(10^{-4}) \end{pmatrix} \tag{A2}$$

$$\text{Texture 3(a)}:\ c^{L/R} = \begin{pmatrix} \pm 1 & 1 & 1 \\ 1 & \pm 1 & 1 \\ 1 & 1 & \pm 1 \end{pmatrix} \xrightarrow{\text{RGE}} |c^{L/R}| \sim \begin{pmatrix} \mathcal{O}(1) & \mathcal{O}(1) & \mathcal{O}(1) \\ \mathcal{O}(1) & \mathcal{O}(1) & \mathcal{O}(1) \\ \mathcal{O}(1) & \mathcal{O}(1) & \mathcal{O}(1) \end{pmatrix} \tag{A3}$$

$$\text{Texture 3(b)}:\ c^{L/R} = \begin{pmatrix} \pm 1 & 1 & 10^{-3} \\ 1 & \pm 1 & 10^{-3} \\ 10^{-3} & 10^{-3} & \pm 10^{-3} \end{pmatrix} \xrightarrow{\text{RGE}} |c^{L/R}| \sim \begin{pmatrix} \mathcal{O}(1) & \mathcal{O}(1) & \mathcal{O}(10^{-3}) \\ \mathcal{O}(1) & \mathcal{O}(1) & \mathcal{O}(10^{-3}) \\ \mathcal{O}(10^{-3}) & \mathcal{O}(10^{-3}) & \mathcal{O}(10^{-3}) \end{pmatrix} \tag{A4}$$

$$\text{Texture 3(c)}:\ c^{L/R} = \begin{pmatrix} \pm 10^{-4} & 10^{-4} & 10^{-4} \\ 10^{-4} & \pm 1 & 1 \\ 10^{-4} & 1 & \pm 1 \end{pmatrix} \xrightarrow{\text{RGE}} |c^{L/R}| \sim \begin{pmatrix} \mathcal{O}(10^{-4}) & \mathcal{O}(10^{-4}) & \mathcal{O}(10^{-4}) \\ \mathcal{O}(10^{-4}) & \mathcal{O}(1) & \mathcal{O}(1) \\ \mathcal{O}(10^{-4}) & \mathcal{O}(1) & \mathcal{O}(1) \end{pmatrix} \tag{A5}$$

$$\text{Texture 3(d)}:\ c^{L/R} \sim \begin{pmatrix} \pm 1 & 10^{-4} & 1 \\ 10^{-4} & \pm 10^{-4} & 10^{-4} \\ 1 & 10^{-4} & \pm 1 \end{pmatrix} \xrightarrow{\text{RGE}} |c^{L/R}| \sim \begin{pmatrix} \mathcal{O}(1) & \mathcal{O}(10^{-4}) & \mathcal{O}(1) \\ \mathcal{O}(10^{-4}) & \mathcal{O}(10^{-4}) & \mathcal{O}(10^{-4}) \\ \mathcal{O}(1) & \mathcal{O}(10^{-4}) & \mathcal{O}(1) \end{pmatrix} \tag{A6}$$

TABLE III. Three-family framework. Dominant contributions to $\hat{Y}'(m_a)$, performing the RGE running from $\Lambda = 10$ TeV down to $m_a \in [100, 400]$ GeV, using ALPaca. The initial values of the couplings assume the LH and RH couplings to have opposite (equal) signs (positive for LH, negative for RH) for the diagonal (off-diagonal) terms.

	Dominant $\hat{Y}'(m_a)$
Texture 3(a)	$(\hat{Y}'_d)^{33} \sim 10^{-5}, (\hat{Y}'_u)^{33} \sim 10^{-4}, (\hat{Y}'_e)^{33} \sim 10^{-2}$
Texture 3(b)	$(\hat{Y}'_d)^{33} \sim 10^{-6}, (\hat{Y}'_u)^{33} \sim 10^{-4}, (\hat{Y}'_e)^{22} \sim 10^{-3}$
Texture 3(c)	$(\hat{Y}'_d)^{33} \sim 10^{-6}, (\hat{Y}'_u)^{33} \sim 10^{-4}, (\hat{Y}'_e)^{33} \sim 10^{-2}$
Texture 3(d)	$(\hat{Y}'_d)^{33} \sim 10^{-6}, (\hat{Y}'_u)^{33} \sim 10^{-4}, (\hat{Y}'_e)^{33} \sim 10^{-2}$

Notice that after RGE the couplings alter their exact values but not the overall order of magnitude. Additionally, in the derivative basis, the ALP coupling to gauge bosons is scale-invariant. Therefore, they will not be generated via running.

Another relevant point we briefly discussed was the fact that we do not need to consider observables with quarks since the Wilson coefficients associated with the integration out of the ALP, Eq. (3.13), are all suppressed with the exception of $\mathcal{C}_{le}$. Such smallness is related to the fact that the quark $\hat{Y}'$ is suppressed by the Yukawa and the ALP–quark couplings generated via running. Indeed, an explicit calculation shows that the biggest contributions to the quark $\hat{Y}'$ are related to the t and b quarks, as could be expected. The dominant contribution to this quantity for the different textures can be seen in Table II in the two-family case, and Table III for the three-family case. For these calculations, we fix the input parameters $\Lambda = 10$ TeV and $m_a \in [100, 400]$ GeV.

In Eq. (3.13), we demonstrate that these quantities enter the Wilson coefficients in pairs and are additionally suppressed by $v^2/(m_a^2 f_a^2)$. Therefore, all operators involving quarks will be suppressed at least by a factor $10^{-6} v^2/(m_a^2 f_a^2)$.

[1] R. D. Peccei and H. R. Quinn, *CP* conservation in the presence of instantons, Phys. Rev. Lett. **38**, 1440 (1977).
[2] S. Weinberg, A new light boson?, Phys. Rev. Lett. **40**, 223 (1978).
[3] F. Wilczek, Problem of strong P and T invariance in the presence of instantons, Phys. Rev. Lett. **40**, 279 (1978).
[4] A. R. Zhitnitsky, On possible suppression of the axion hadron interactions (in Russian), Sov. J. Nucl. Phys. **31**, 260 (1980).
[5] M. Dine, W. Fischler, and M. Srednicki, A simple solution to the strong *CP* problem with a harmless axion, Phys. Lett. B **104**, 199 (1981).
[6] J. E. Kim, Weak interaction singlet and strong *CP* invariance, Phys. Rev. Lett. **43**, 103 (1979).
[7] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Can confinement ensure natural *CP* invariance of strong interactions?, Nucl. Phys. **B166**, 493 (1980).
[8] G. Grilli di Cortona, E. Hardy, J. Pardo Vega, and G. Villadoro, The QCD axion, precisely, J. High Energy Phys. 01 (2016) 034.
[9] B. B. Brandt, G. Endrődi, J. J. Hernández Hernández, G. Markó, and L. Pannullo, The axion-photon coupling from lattice quantum chromodynamics, arXiv:2603.29153.
[10] H. Georgi, D. B. Kaplan, and L. Randall, Manifesting the invisible axion at low-energies, Phys. Lett. **169B**, 73 (1986).
[11] L. Merlo, F. Pobbe, and S. Rigolin, The minimal axion minimal linear σ model, Eur. Phys. J. C **78**, 415 (2018); **79**, 963(E) (2019).
[12] I. Brivio, M. B. Gavela, S. Pascoli, R. del Rey, and S. Saa, The axion and the Goldstone Higgs, Chin. J. Phys. (Taipei) **61**, 55 (2019).
[13] J. Alonso-González, L. Merlo, F. Pobbe, S. Rigolin, and O. Sumensari, Testable axion-like particles in the minimal linear σ model, Nucl. Phys. **B950**, 114839 (2020).

[14] J. Alonso-Gonzalez, J. M. Lizana, V. Martinez-Fernandez, L. Merlo, and S. Pokorski, Probing effective field theory approach in the *CP* violating minimal linear σ model, Eur. Phys. J. C **81**, 538 (2021).

[15] B. Bellazzini, A. Mariotti, D. Redigolo, F. Sala, and J. Serra, *R*-axion at colliders, Phys. Rev. Lett. **119**, 141804 (2017).

[16] E. Witten, Some properties of O(32) superstrings, Phys. Lett. **149B**, 351 (1984).

[17] K.-S. Choi, I.-W. Kim, and J. E. Kim, String compactification, QCD axion and axion-photon-photon coupling, J. High Energy Phys. 03 (2007) 116.

[18] P. Svrcek and E. Witten, Axions in string theory, J. High Energy Phys. 06 (2006) 051.

[19] A. Arvanitaki, S. Dimopoulos, S. Dubovsky, N. Kaloper, and J. March-Russell, String axiverse, Phys. Rev. D **81**, 123530 (2010).

[20] M. Cicoli, M. Goodsell, and A. Ringwald, The type IIB string axiverse and its low-energy phenomenology, J. High Energy Phys. 10 (2012) 146.

[21] G. Gelmini, D. N. Schramm, and J. W. F. Valle, Majorons: A simultaneous solution to the large and small scale dark matter problems, Phys. Lett. **146B**, 311 (1984).

[22] V. Berezinsky and J. W. F. Valle, The KeV majoron as a dark matter particle, Phys. Lett. B **318**, 360 (1993).

[23] M. Lattanzi and J. W. F. Valle, Decaying warm dark matter and neutrino masses, Phys. Rev. Lett. **99**, 121301 (2007).

[24] F. Bazzocchi, M. Lattanzi, S. Riemer-Sørensen, and J. W. F. Valle, X-ray photons from late-decaying majoron dark matter, J. Cosmol. Astropart. Phys. 08 (2008) 013.

[25] M. Lattanzi, S. Riemer-Sorensen, M. Tortola, and J. W. F. Valle, Updated CMB and x- and γ-ray constraints on Majoron dark matter, Phys. Rev. D **88**, 063528 (2013).

[26] F. S. Queiroz and K. Sinha, The Poker face of the Majoron dark matter model: LUX to keV line, Phys. Lett. B **735**, 69 (2014).

[27] R. Z. Ferreira and A. Notari, Observable windows for the QCD axion through the number of relativistic species, Phys. Rev. Lett. **120**, 191301 (2018).

[28] F. D'Eramo, R. Z. Ferreira, A. Notari, and J. L. Bernal, Hot axions and the H_0 tension, J. Cosmol. Astropart. Phys. 11 (2018) 014.

[29] M. Escudero and S. J. Witte, A CMB search for the neutrino mass mechanism and its relation to the Hubble tension, Eur. Phys. J. C **80**, 294 (2020).

[30] F. Arias-Aragón, F. D'eramo, R. Z. Ferreira, L. Merlo, and A. Notari, Cosmic imprints of XENON1T axions, J. Cosmol. Astropart. Phys. 11 (2020) 025.

[31] F. Arias-Aragon, E. Fernandez-Martinez, M. Gonzalez-Lopez, and L. Merlo, Neutrino masses and Hubble tension via a Majoron in MFV, Eur. Phys. J. C **81**, 28 (2021).

[32] F. Arias-Aragón, F. D'Eramo, R. Z. Ferreira, L. Merlo, and A. Notari, Production of thermal axions across the ElectroWeak phase transition, J. Cosmol. Astropart. Phys. 03 (2021) 090.

[33] R. Z. Ferreira, A. Notari, and F. Rompineve, Dine-Fischler-Srednicki-Zhitnitsky axion in the CMB, Phys. Rev. D **103**, 063524 (2021).

[34] M. Escudero and S. J. Witte, The hubble tension as a hint of leptogenesis and neutrino mass generation, Eur. Phys. J. C **81**, 515 (2021).

[35] T. Araki, K. Asai, K. Honda, R. Kasuya, J. Sato, T. Shimomura, and M. J. S. Yang, Resolving the Hubble tension in a $U(1)_{L_\mu - L_\tau}$ model with the Majoron, Prog. Theor. Exp. Phys. **2021**, 103B05 (2021).

[36] F. D'Eramo, F. Hajkarim, and S. Yun, Thermal axion production at low temperatures: A smooth treatment of the QCD phase transition, Phys. Rev. Lett. **128**, 152001 (2022).

[37] F. D'Eramo, F. Hajkarim, and S. Yun, Thermal QCD axions across thresholds, J. High Energy Phys. 10 (2021) 224.

[38] F. D'Eramo, E. Di Valentino, W. Giarè, F. Hajkarim, A. Melchiorri, O. Mena, F. Renzi, and S. Yun, Cosmological bound on the QCD axion mass, redux, J. Cosmol. Astropart. Phys. 09 (2022) 022.

[39] A. Davidson and K. C. Wali, Minimal flavor unification via multigenerational Peccei-Quinn symmetry, Phys. Rev. Lett. **48**, 11 (1982).

[40] F. Wilczek, Axions and family symmetry breaking, Phys. Rev. Lett. **49**, 1549 (1982).

[41] Y. Ema, K. Hamaguchi, T. Moroi, and K. Nakayama, Flaxion: A minimal extension to solve puzzles in the standard model, J. High Energy Phys. 01 (2017) 096.

[42] L. Calibbi, F. Goertz, D. Redigolo, R. Ziegler, and J. Zupan, Minimal axion model from flavor, Phys. Rev. D **95**, 095009 (2017).

[43] F. Arias-Aragon and L. Merlo, The minimal flavour violating axion, J. High Energy Phys. 10 (2017) 168; 11 (2019) 152(E).

[44] F. Arias-Aragón, E. Fernández-Martínez, M. González-López, and L. Merlo, Dynamical minimal flavour violating inverse seesaw, J. High Energy Phys. 09 (2022) 210.

[45] L. Di Luzio, A. W. M. Guerrera, X. P. Díaz, and S. Rigolin, On the IR/UV flavour connection in non-universal axion models, J. High Energy Phys. 06 (2023) 046.

[46] A. Greljo, A. Smolkovič, and A. Valenti, Froggatt-Nielsen ALP, J. High Energy Phys. 09 (2024) 174.

[47] Y. Chikashige, R. N. Mohapatra, and R. D. Peccei, Spontaneously broken lepton number and cosmological constraints on the neutrino mass spectrum, Phys. Rev. Lett. **45**, 1926 (1980).

[48] Y. Chikashige, R. N. Mohapatra, and R. D. Peccei, Are there real Goldstone bosons associated with broken lepton number?, Phys. Lett. **98B**, 265 (1981).

[49] G. B. Gelmini and M. Roncadelli, Left-handed neutrino mass scale and spontaneously broken lepton number, Phys. Lett. B **99**, 411 (1981).

[50] A. de Giorgi, L. Merlo, X. Ponce Díaz, and S. Rigolin, The minimal massive Majoron Seesaw Model, J. High Energy Phys. 03 (2024) 094.

[51] Q. Liang, X. Ponce Díaz, and T. T. Yanagida, Axion detection experiments can probe Majoron models, Phys. Rev. Lett. **134**, 151803 (2025).

[52] A. Greljo, X. Ponce Díaz, and A. E. Thomsen, Insights on the cosmic origin of matter from proton stability, J. Cosmol. Astropart. Phys. 11 (2025) 043.

[53] C. Albertus *et al.*, WISPedia—the WISPs Encyclopedia: Cosmic WISPers 2026—V1.0, arXiv:2602.09089.
[54] K. Choi, K. Kang, and J. E. Kim, Effects of η' in low-energy axion physics, Phys. Lett. B **181**, 145 (1986).
[55] A. Salvio, A. Strumia, and W. Xue, Thermal axion production, J. Cosmol. Astropart. Phys. 01 (2014) 011.
[56] I. Brivio, M. B. Gavela, L. Merlo, K. Mimasu, J. M. No, R. del Rey, and V. Sanz, ALPs effective field theory and collider signatures, Eur. Phys. J. C **77**, 572 (2017).
[57] G. Alonso-Álvarez, M. B. Gavela, and P. Quilez, Axion couplings to electroweak gauge bosons, Eur. Phys. J. C **79**, 223 (2019).
[58] M. B. Gavela, R. Houtz, P. Quilez, R. Del Rey, and O. Sumensari, Flavor constraints on electroweak ALP couplings, Eur. Phys. J. C **79**, 369 (2019).
[59] M. Chala, G. Guedes, M. Ramos, and J. Santiago, Running in the ALPs, Eur. Phys. J. C **81**, 181 (2021).
[60] J. Bonilla, I. Brivio, M. B. Gavela, and V. Sanz, One-loop corrections to ALP couplings, J. High Energy Phys. 11 (2021) 168.
[61] F. Arias-Aragón and C. Smith, Leptoquarks, axions and the unification of B, L, and Peccei-Quinn symmetries, Phys. Rev. D **106**, 055034 (2022).
[62] F. Arias-Aragón, J. Quevillon, and C. Smith, Axion-like ALPs, J. High Energy Phys. 03 (2023) 134.
[63] L. Di Luzio, H. Gisbert, G. Levati, P. Paradisi, and P. Sørensen, *CP*-violating axions: A theory review, arXiv:2312.17310.
[64] L. Di Luzio, G. Levati, and P. Paradisi, The chiral Lagrangian of *CP*-violating axion-like particles, J. High Energy Phys. 02 (2024) 020.
[65] C. H. de Lima, D. McKeen, and M. Swiatlowski, Low-multiplicity jets as probes of GeV-scale light-quark-coupled particles, arXiv:2604.19864.
[66] J. Jaeckel, M. Jankowiak, and M. Spannowsky, LHC probes the hidden sector, Phys. Dark Universe **2**, 111 (2013).
[67] K. Mimasu and V. Sanz, ALPs at Colliders, J. High Energy Phys. 06 (2015) 173.
[68] J. Jaeckel and M. Spannowsky, Probing MeV to 90 GeV axion-like particles with LEP and LHC, Phys. Lett. B **753**, 482 (2016).
[69] A. Alves, A. G. Dias, and K. Sinha, Diphotons at the *Z*-pole in models of the 750 GeV resonance decaying to axion-like particles, J. High Energy Phys. 08 (2016) 060.
[70] S. Knapen, T. Lin, H. K. Lou, and T. Melia, Searching for axionlike particles with ultraperipheral heavy-ion collisions, Phys. Rev. Lett. **118**, 171801 (2017).
[71] M. Bauer, M. Neubert, and A. Thamm, LHC as an axion factory: Probing an axion explanation for $(g-2)_\mu$ with exotic Higgs decays, Phys. Rev. Lett. **119**, 031802 (2017).
[72] A. Mariotti, D. Redigolo, F. Sala, and K. Tobioka, New LHC bound on low-mass diphoton resonances, Phys. Lett. B **783**, 13 (2018).
[73] M. Bauer, M. Neubert, and A. Thamm, Collider probes of axion-like particles, J. High Energy Phys. 12 (2017) 044.
[74] C. Baldenegro, S. Fichet, G. von Gersdorff, and C. Royon, Searching for axion-like particles with proton tagging at the LHC, J. High Energy Phys. 06 (2018) 131.
[75] N. Craig, A. Hook, and S. Kasko, The photophobic ALP, J. High Energy Phys. 09 (2018) 028.
[76] M. Bauer, M. Heiles, M. Neubert, and A. Thamm, Axion-like particles at future colliders, Eur. Phys. J. C **79**, 74 (2019).
[77] M. B. Gavela, J. M. No, V. Sanz, and J. F. de Trocóniz, Nonresonant searches for axionlike particles at the LHC, Phys. Rev. Lett. **124**, 051802 (2020).
[78] G. Haghighat, D. Haji Raissi, and M. Mohammadi Najafabadi, New collider searches for axionlike particles coupling to gluons, Phys. Rev. D **102**, 115010 (2020).
[79] D. Wang, L. Wu, J. M. Yang, and M. Zhang, Photon-jet events as a probe of axionlike particles at the LHC, Phys. Rev. D **104**, 095016 (2021).
[80] A. de Giorgi, L. Merlo, and J.-L. Tastet, Probing HNL-ALP couplings at colliders, Fortschr. Phys. **71**, 2300027 (2023).
[81] J. Bonilla, I. Brivio, J. Machado-Rodríguez, and J. F. de Trocóniz, Nonresonant searches for axion-like particles in vector boson scattering processes at the LHC, J. High Energy Phys. 06 (2022) 113.
[82] F. A. Ghebretinsaea, Z. S. Wang, and K. Wang, Probing axion-like particles coupling to gluons at the LHC, J. High Energy Phys. 07 (2022) 070.
[83] V. E. Vileta, B. Gavela, R. Houtz, and P. Quilez, Discrete Goldstone bosons, Phys. Rev. D **107**, 035009 (2023).
[84] L. Calibbi, Z. Huang, S. Qin, Y. Yang, and X. Yin, Testing axion couplings to leptons in Z decays at future e^+e^- colliders, Phys. Rev. D **108**, 015002 (2023).
[85] M. B. Marcos, A. de Giorgi, L. Merlo, and J.-L. Tastet, ALPs and HNLs at LHC and Muon Colliders: Uncovering new couplings and signals, SciPost Phys. **18**, 084 (2025).
[86] F. Arias-Aragón, L. Darmé, G. G. di Cortona, and E. Nardi, Atoms as electron accelerators for measuring the cross section of $e^+e^- \to$ hadrons, Phys. Rev. Lett. **134**, 061802 (2025).
[87] A. Biekötter and K. Mimasu, Axions and axion-like particles: Collider searches, arXiv:2508.19358.
[88] Y. Ema, P. J. Fox, M. Hostert, T. Menzo, M. Pospelov, A. Ray, and J. Zupan, Long-lived axionlike particles from tau decays, Phys. Rev. D **112**, 115028 (2025).
[89] E. Izaguirre, T. Lin, and B. Shuve, Searching for axionlike particles in flavor-changing neutral current processes, Phys. Rev. Lett. **118**, 111802 (2017).
[90] W. J. Marciano, A. Masiero, P. Paradisi, and M. Passera, Contributions of axionlike particles to lepton dipole moments, Phys. Rev. D **94**, 115033 (2016).
[91] L. Merlo, F. Pobbe, S. Rigolin, and O. Sumensari, Revisiting the production of ALPs at B-factories, J. High Energy Phys. 06 (2019) 091.
[92] D. Aloni, C. Fanelli, Y. Soreq, and M. Williams, Photoproduction of axionlike particles, Phys. Rev. Lett. **123**, 071801 (2019).
[93] M. Bauer, M. Neubert, S. Renner, M. Schnubel, and A. Thamm, Axionlike particles, lepton-flavor violation, and a new explanation of a_μ and a_e, Phys. Rev. Lett. **124**, 211803 (2020).

[94] C. Cornella, P. Paradisi, and O. Sumensari, Hunting for ALPs with lepton flavor violation, J. High Energy Phys. 01 (2020) 158.

[95] M. Bauer, M. Neubert, S. Renner, M. Schnubel, and A. Thamm, The low-energy effective theory of axions and ALPs, J. High Energy Phys. 04 (2021) 063.

[96] L. Calibbi, D. Redigolo, R. Ziegler, and J. Zupan, Looking forward to lepton-flavor-violating ALPs, J. High Energy Phys. 09 (2021) 173.

[97] L. Di Luzio, R. Gröber, and P. Paradisi, Hunting for *CP*-violating axionlike particle interactions, Phys. Rev. D **104**, 095027 (2021).

[98] M. Bauer, M. Neubert, S. Renner, M. Schnubel, and A. Thamm, Flavor probes of axion-like particles, J. High Energy Phys. 09 (2022) 056.

[99] A. Carmona, C. Scherb, and P. Schwaller, Charming ALPs, J. High Energy Phys. 08 (2021) 121.

[100] A. W. M. Guerrera and S. Rigolin, Revisiting $K \to \pi a$ decays, Eur. Phys. J. C **82**, 192 (2022).

[101] J. Alda, A. W. M. Guerrera, S. Peñaranda, and S. Rigolin, Leptonic meson decays into invisible ALP, Nucl. Phys. **B979**, 115791 (2022).

[102] E. Bertholet, S. Chakraborty, V. Loladze, T. Okui, A. Soffer, and K. Tobioka, Heavy QCD axion at Belle II: Displaced and prompt signals, Phys. Rev. D **105**, L071701 (2022).

[103] H.-C. Cheng, L. Li, and E. Salvioni, A theory of dark pions, J. High Energy Phys. 01 (2022) 122.

[104] J. Bonilla, A. de Giorgi, B. Gavela, L. Merlo, and M. Ramos, The cost of an ALP solution to the neutral B-anomalies, J. High Energy Phys. 02 (2023) 138.

[105] J. Bonilla, A. de Giorgi, and M. Ramos, Neutral *B*-anomalies from an *on-shell* scalar exchange, arXiv: 2211.05135.

[106] A. de Giorgi and G. Piazza, A lesson from $R_{\tau\tau}^{K^{(*)}}$ and $R_{\nu\nu}^{K^{(*)}}$ at Belle II, Fortschr. Phys. **72**, 2300200 (2024).

[107] A. W. M. Guerrera and S. Rigolin, ALP production in weak mesonic decays, Fortschr. Phys. **71**, 2200192 (2023).

[108] J. Bonilla, B. Gavela, and J. Machado-Rodríguez, Limits on ALP-neutrino couplings from loop-level processes, Phys. Rev. D **109**, 055023 (2024).

[109] F. Arias-Aragón, V. Brdar, and J. Quevillon, New directions for axionlike particle searches combining nuclear reactors and haloscopes, Phys. Rev. Lett. **132**, 211802 (2024).

[110] L. Di Luzio, A. W. M. Guerrera, X. Ponce Díaz, and S. Rigolin, Axion-like particles in radiative quarkonia decays, J. High Energy Phys. 06 (2024) 217.

[111] A. de Giorgi, M. Fuentes Zamoro, and L. Merlo, Visible GeV ALP from TeV vector-like leptons, Fortschr. Phys. **73**, 2400165 (2025).

[112] J. Alda, G. Levati, P. Paradisi, S. Rigolin, and N. Selimovic, Collider and astrophysical signatures of light scalars with enhanced τ couplings, J. High Energy Phys. 06 (2025) 008.

[113] J. Alda, C. Broggini, G. Di Carlo, L. Di Luzio, D. Piatti, S. Rigolin, and C. Toni, Weak nuclear decays deep-underground as a probe of axion dark matter, Phys. Rev. D **111**, 035022 (2025).

[114] F. Arias-Aragón, L. Darmé, G. G. di Cortona, and E. Nardi, Production of dark sector particles via resonant positron annihilation on atomic electrons, Phys. Rev. Lett. **132**, 261801 (2024); **133**, 219901(E) (2024).

[115] F. Arias-Aragón, M. Giannotti, G. G. di Cortona, and F. Mescia, Axion-induced pair production: A new strategy for axion detection, Phys. Rev. D **111**, 043021 (2025).

[116] D. Bisht, S. Chakraborty, and A. Samanta, A comprehensive study of ALPs from B-decays, J. High Energy Phys. 07 (2025) 092.

[117] L. Calibbi, T. Li, L. Mukherjee, and Y. Yang, Probing ALP lepton flavor violation at μTRISTAN, Phys. Rev. D **110**, 115009 (2024).

[118] J. Martin Camalich and R. Ziegler, Flavor phenomenology of light dark sectors, Annu. Rev. Nucl. Part. Sci. **75**, 223 (2025).

[119] J. Alda, M. Fuentes Zamoro, L. Merlo, X. Ponce Díaz, and S. Rigolin, Comprehensive ALP searches in meson decays, arXiv:2507.19578.

[120] J. Alda, M. Fuentes Zamoro, L. Merlo, X. Ponce Díaz, and S. Rigolin, ALPaca: The ALP automatic computing algorithm, arXiv:2508.08354.

[121] M. Ardu, L. Calibbi, M. Fedele, and F. Mescia, ALP production in Lepton flavour violating meson, tau and gauge boson decays, arXiv:2604.11889.

[122] A. Eberhart, M. Fedele, F. Kahlhoefer, E. Ravensburg, and R. Ziegler, Leptophilic ALPs in laboratory experiments, J. High Energy Phys. 12 (2025) 055.

[123] A. Arza *et al.*, The COSMIC WISPers white paper: The physics case for Weakly Interacting Slim Particles, arXiv:2603.03433.

[124] L. Di Luzio, F. Mescia, and E. Nardi, Redefining the axion window, Phys. Rev. Lett. **118**, 031801 (2017).

[125] L. Di Luzio, F. Mescia, and E. Nardi, Window for preferred axion models, Phys. Rev. D **96**, 075003 (2017).

[126] M. K. Gaillard, M. B. Gavela, R. Houtz, P. Quilez, and R. Del Rey, Color unified dynamical axion, Eur. Phys. J. C **78**, 972 (2018).

[127] A. Hook, S. Kumar, Z. Liu, and R. Sundrum, High quality QCD axion and the LHC, Phys. Rev. Lett. **124**, 221801 (2020).

[128] L. Di Luzio, M. Giannotti, E. Nardi, and L. Visinelli, The landscape of QCD axion models, Phys. Rep. **870**, 1 (2020).

[129] L. Di Luzio, B. Gavela, P. Quilez, and A. Ringwald, An even lighter QCD axion, J. High Energy Phys. 05 (2021) 184.

[130] L. Di Luzio, B. Gavela, P. Quilez, and A. Ringwald, Dark matter from an even lighter QCD axion: Trapped misalignment, J. Cosmol. Astropart. Phys. 10 (2021) 001.

[131] B. Gavela, P. Quílez, and M. Ramos, The QCD axion sum rule, J. High Energy Phys. 04 (2024) 056.

[132] P. Cox, T. Gherghetta, and A. Paul, A common origin for the QCD axion and sterile neutrinos from $SU(5)$ strong dynamics, J. High Energy Phys. 12 (2023) 180.

[133] A. de Giorgi and M. Ramos, Extra-dimensional axion patterns, Phys. Rev. D **111**, 075006 (2025).

[134] H. Murayama, GeV-scale QCD axion, arXiv:2601.02465.

[135] L. Di Luzio, M. Redi, A. Strumia, A. Tesi, and A. V. Titov, The too visible QCD axion, J. High Energy Phys. 06 (2026) 151.

[136] P. Escribano and A. Vicente, Ultralight scalars in leptonic observables, J. High Energy Phys. 03 (2021) 240.
[137] See Supplemental Material at http://link.aps.org/supplemental/10.1103/xf8c-823x for a *Mathematica* notebook with the complete analytical expressions of the observables used in this work, the assumptions employed in their derivation, and a comparison with the available literature.
[138] A. Manohar and H. Georgi, Chiral quarks and the nonrelativistic quark model, Nucl. Phys. **B234**, 189 (1984).
[139] A. G. Cohen, D. B. Kaplan, and A. E. Nelson, Counting 4 pis in strongly coupled supersymmetry, Phys. Lett. B **412**, 301 (1997).
[140] B. M. Gavela, E. E. Jenkins, A. V. Manohar, and L. Merlo, Analysis of general power counting rules in effective field theory, Eur. Phys. J. C **76**, 485 (2016).
[141] J. Elias-Miro, J. R. Espinosa, E. Masso, and A. Pomarol, Higgs windows to new physics through $d = 6$ operators: constraints and one-loop anomalous dimensions, J. High Energy Phys. 11 (2013) 066.
[142] J. R. Espinosa, Effective QFT and BSM, lectures at the Taller de Altas Energías 2014, Centro de Ciencias de Benasque Pedro Pascual, Benasque, Spain (2014).
[143] K. Fujikawa, Path integral measure for gauge invariant Fermion theories, Phys. Rev. Lett. **42**, 1195 (1979).
[144] J. Martin Camalich, M. Pospelov, P. N. H. Vuong, R. Ziegler, and J. Zupan, Quark flavor phenomenology of the QCD axion, Phys. Rev. D **102**, 015023 (2020).
[145] S. Das Bakshi, J. Machado-Rodríguez, and M. Ramos, Running beyond ALPs: Shift-breaking and *CP*-violating effects, J. High Energy Phys. 11 (2023) 133.
[146] L. C. Bresciani, G. Brunello, G. Levati, P. Mastrolia, and P. Paradisi, Renormalization of effective field theories via on-shell methods: The case of axion-like particles, J. High Energy Phys. 10 (2025) 190.
[147] W. Buchmuller and D. Wyler, Effective Lagrangian analysis of new interactions and flavor conservation, Nucl. Phys. **B268**, 621 (1986).
[148] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek, Dimension-six terms in the standard model Lagrangian, J. High Energy Phys. 10 (2010) 085.
[149] S. Navas *et al.* (Particle Data Group), Review of particle physics, Phys. Rev. D **110**, 030001 (2024).
[150] R. H. Parker, C. Yu, W. Zhong, B. Estey, and H. Müller, Measurement of the fine-structure constant as a test of the Standard Model, Science **360**, 191 (2018).
[151] L. Morel, Z. Yao, P. Cladé, and S. Guellati-Khélifa, Determination of the fine-structure constant with an accuracy of 81 parts per trillion, Nature (London) **588**, 61 (2020).
[152] R. Aliberti *et al.*, The anomalous magnetic moment of the muon in the Standard Model: An update, Phys. Rep. **1143**, 1 (2025).
[153] D. P. Aguillard *et al.* (Muon g-2 Collaboration), Detailed report on the measurement of the positive muon anomalous magnetic moment to 0.20 ppm, Phys. Rev. D **110**, 032009 (2024).
[154] S. T. Petcov, The processes $\mu \to e\gamma$, $\mu \to e\bar{e}e$, $\nu' \to \nu\gamma$ in the weinberg–salam model with neutrino mixing, Yad. Fiz. **25**, 641 (1977) [Sov. J. Nucl. Phys. **25**, 340 (1977)]; Yad. Fiz. **25**, 1336(E) (1977) [Sov. J. Nucl. Phys. **25**, 698(E) (1977)].
[155] T.-P. Cheng and L.-F. Li, *Gauge Theory of Elementary Particle Physics* (Oxford University Press, New York, 2000).
[156] L. Calibbi and G. Signorelli, Charged lepton flavour violation: An experimental and theoretical introduction, Riv. Nuovo Cimento Soc. Ital. Fis. **41**, 71 (2018).
[157] K. Afanaciev *et al.* (MEG II Collaboration), New limit on the $\mu^+ \to e^+\gamma$ decay with the MEG II experiment, Eur. Phys. J. C **85**, 1177 (2025).
[158] K. Afanaciev *et al.* (MEG II Collaboration), A search for $\mu^+ \to e^+\gamma$ with the first dataset of the MEG II experiment, Eur. Phys. J. C **84**, 216 (2024); **84**, 1042(E) (2024).
[159] G. Hernández-Tomé, G. López Castro, and P. Roig, Flavor violating leptonic decays of τ and μ leptons in the Standard Model with massive neutrinos, Eur. Phys. J. C **79**, 84 (2019); **80**, 438(E) (2020).
[160] U. Bellgardt *et al.* (SINDRUM Collaboration), Search for the decay $\mu^+ \to e^+e^+e^-$, Nucl. Phys. **B299**, 1 (1988).
[161] D. Redigolo, M. Tammaro, and A. Tesi, Large *CP* violation in flavor violating muon decays, Eur. Phys. J. C **85**, 103 (2025).
[162] A. Ilakovac and A. Pilaftsis, Flavor violating charged lepton decays in seesaw-type models, Nucl. Phys. **B437**, 491 (1995).
[163] W. H. Bertl *et al.* (SINDRUM II Collaboration), A search for muon to electron conversion in muonic gold, Eur. Phys. J. C **47**, 337 (2006).
[164] L. Bartoszek *et al.* (Mu2e Collaboration), Mu2e technical design report, arXiv:1501.05241.
[165] M. Moritsu (COMET Collaboration), Search for muon-to-electron conversion with the COMET experiment, Universe **8**, 196 (2022).
[166] A. M. Sirunyan *et al.* (CMS Collaboration), Search for heavy Higgs bosons decaying to a top quark pair in proton-proton collisions at $\sqrt{s} = 13$ TeV, J. High Energy Phys. 04 (2020) 171; 03 (2022) 187(E).
[167] F. Esser, M. Madigan, V. Sanz, and M. Ubiali, On the coupling of axion-like particles to the top quark, J. High Energy Phys. 09 (2023) 063.
[168] A. Anuar, A. Biekötter, T. Biekötter, A. Grohsjean, S. Heinemeyer, L. Jeppe, C. Schwanenberger, and G. Weiglein, ALP-ine quests at the LHC: hunting axion-like particles via peaks and dips in $t\bar{t}$ production, J. High Energy Phys. 12 (2024) 197.
[169] F. Esser, M. Madigan, A. Salas-Bernardez, V. Sanz, and M. Ubiali, Di-Higgs production via axion-like particles, J. High Energy Phys. 10 (2024) 164.
[170] G. D'Ambrosio, G. F. Giudice, G. Isidori, and A. Strumia, Minimal flavor violation: An effective field theory approach, Nucl. Phys. **B645**, 155 (2002).
[171] V. Cirigliano, B. Grinstein, G. Isidori, and M. B. Wise, Minimal flavor violation in the lepton sector, Nucl. Phys. **B728**, 121 (2005).
[172] S. Davidson and F. Palorini, Various definitions of minimal flavour violation for leptons, Phys. Lett. B **642**, 72 (2006).

[173] M. B. Gavela, T. Hambye, D. Hernandez, and P. Hernandez, Minimal flavour Seesaw models, J. High Energy Phys. 09 (2009) 038.
[174] R. Alonso, G. Isidori, L. Merlo, L. A. Munoz, and E. Nardi, Minimal flavour violation extensions of the seesaw, J. High Energy Phys. 06 (2011) 037.
[175] D. N. Dinh, L. Merlo, S. T. Petcov, and R. Vega-Álvarez, Revisiting minimal lepton flavour violation in the light of leptonic *CP* violation, J. High Energy Phys. 07 (2017) 089.
[176] R. Alonso, M. B. Gavela, D. Hernandez, and L. Merlo, On the potential of leptonic minimal flavour violation, Phys. Lett. B **715**, 194 (2012).
[177] R. Alonso, M. B. Gavela, G. Isidori, and L. Maiani, Neutrino mixing and masses from a minimum principle, J. High Energy Phys. 11 (2013) 187.
[178] R. Alonso, M. B. Gavela, D. Hernández, L. Merlo, and S. Rigolin, Leptonic dynamical Yukawa couplings, J. High Energy Phys. 08 (2013) 069.
[179] G. Blankenburg, G. Isidori, and J. Jones-Perez, Neutrino masses and LFV from minimal breaking of $U(3)^5$ and $U(2)^5$ flavor Symmetries, Eur. Phys. J. C **72**, 2126 (2012).
[180] T. Feldmann, C. Luhn, and P. Moch, Lepton-flavour violation in a Pati-Salam model with gauged flavour symmetry, J. High Energy Phys. 11 (2016) 078.
[181] R. Alonso, E. Fernandez Martinez, M. B. Gavela, B. Grinstein, L. Merlo, and P. Quilez, Gauged lepton flavour, J. High Energy Phys. 12 (2016) 119.
[182] F. Arias-Aragón, C. Bouthelier-Madre, J. M. Cano, and L. Merlo, Data driven flavour model, Eur. Phys. J. C **80**, 854 (2020).
[183] S. Weinberg, Baryon and lepton nonconserving processes, Phys. Rev. Lett. **43**, 1566 (1979).
[184] D. M. Straub, FLAVIO: A Python package for flavour and precision phenomenology in the standard model and beyond, arXiv:1810.08132.