Leptophilic ALPs in laboratory experiments

Alexander Eberhart [ID],[a] Marco Fedele [ID],[b,c] Felix Kahlhoefer [ID],[a,d] Eike Ravensburg [ID][e,f] and Robert Ziegler [ID][a]

[a] *Institute for Theoretical Particle Physics (TTP), Karlsruhe Institute of Technology (KIT),*
  Wolfgang-Gaede-Str. 1, D-76131 Karlsruhe, Germany

[b] *Departament de Física Teòrica, IFIC, Universitat de València — CSIC,*
  Parque Científico, Catedrático José Beltrán 2, E-46980 Paterna, Spain

[c] *PRISMA+ Cluster of Excellence & Mainz Institute for Theoretical Physics,*
  Johannes Gutenberg University,
  D-55099 Mainz, Germany

[d] *Institute for Astroparticle Physics (IAP), Karlsruhe Institute of Technology (KIT),*
  Wolfgang-Gaede-Str. 1, D-76131 Karlsruhe, Germany

[e] *Oskar Klein Centre, Department of Physics, Stockholm University,*
  SE-10691 Stockholm, Sweden

[f] *CP3-Origins, University of Southern Denmark,*
  Campusvej 55, DK-5230 Odense M, Denmark

  E-mail: alexander.m.eberhart@durham.ac.uk, mfedele@uni-mainz.de,
  kahlhoefer@kit.edu, ravensburg@cp3.sdu.dk, robert.ziegler@kit.edu

Abstract: We study the collider phenomenology of leptophilic axion-like particles (ALPs), i.e. pseudoscalar particles that couple only to charged leptons. Loops of charged leptons induce effective interactions of the ALPs with photons, which depend on the momenta of the interacting particles and differ between pseudoscalar and derivative lepton couplings. We systematically discuss the form of the interaction with photons for general external momenta and identify the regimes when it can be safely approximated by an effective coupling constant. We use these results to derive novel constraints from LEP and calculate state-of-the-art limits from E137 and NA64 for four different scenarios, in which the ALPs couple either to a single lepton generation or universally to all, for both pseudoscalar and derivative lepton couplings. We collect complementary bounds from astrophysics, flavour, and other laboratory experiments to chart the allowed parameter space of leptophilic ALPs in the MeV-GeV mass range.

Contents

1 Introduction

Some of the simplest modifications of the Standard Model (SM) involve a single axion-like particle (ALP), which is a pseudo-Goldstone boson of a suitable, spontaneously broken global symmetry. In contrast to the QCD axion [1–4], this symmetry is subject to a large explicit breaking, which induces an ALP mass that is simply taken as a free parameter. Of particular interest are masses in the MeV–GeV range, which lead to a very rich phenomenology at particle colliders [5], depending on the precise form of ALP interactions with SM particles. Though the explanation of the dark matter abundance is one of the main motivations of these scenarios (see, e.g. refs. [6–34]), it is important to map current limits from laboratory experiments, astrophysics and cosmology into the general parameter space of the ALP Lagrangian.

While this task has been performed in much detail for ALP couplings to photons, couplings to leptons have received far less attention (notable exceptions include refs. [34–40]). In the present work, we aim to systematically study the model-independent laboratory constraints on "leptophilic ALPs", which, by definition, couple dominantly to charged leptons.[1] As a result, we provide an exhaustive overview of the relevant constraints from beam-dump, flavour and collider experiments on leptophilic ALPs, which complement limits from astrophysics [37, 42–44].

We restrict ourselves to flavour-diagonal couplings and consider four simple scenarios, where either the coupling to a single lepton family dominates (e, μ, τ) or the ALP couples in a flavour-universal way to leptons. In contrast to previous analyses, we explicitly consider two cases of possible leptonic ALP interactions, which take the form of either pseudoscalar or derivative couplings. It is well-known that these interactions are largely equivalent for tree-level processes involving a single ALP but differ for multi-ALP interactions and, in particular, by an effective ALP-photon coupling induced by the chiral anomaly. Therefore, the two leptophilic interactions lead to different predictions for observables that involve an effective ALP-photon-photon interaction, which is induced by a leptonic loop. The associated vertex function depends on the external momenta and, thus, on the typical energy scales of the relevant process (apart from a constant part depending on whether pseudoscalar or derivative couplings are considered). This energy dependence implies that the phenomenology of the leptophilic ALP cannot be simply obtained from existing limits on the ALP-photon coupling but has to be re-calculated using the full momentum dependence for all processes involving ALP-photon interactions [29, 37]. The complete analysis of the resulting limits from E137 [45], NA64 [46, 47] and LEP [48] represents the main result of this work.

Although such loop-induced processes are suppressed with respect to tree-level diagrams involving only leptons, we will show that beam dump experiments like E137 can give highly relevant constraints on leptophilic ALP, even when both the production and decay of the ALPs proceed exclusively via photon couplings. When the relevant energy scales are much larger than the ALP and lepton masses, the loop function becomes suppressed. The resulting vertex inherits this suppression for pseudoscalar interactions, while the derivative vertex is essentially given by the constant part induced by the chiral anomaly. This implies that, e.g. limits from high-energy colliders like LEP that have been derived for effective photon couplings [48] are

---

[1]The case of ALPs that couple dominantly to neutrinos was recently discussed in ref. [41].

relevant only for leptophilic ALPs with derivative couplings, and similar conclusions hold for highly energetic leptonic beam dumps. More generally, a large separation of scales often permits the approximation of the fully off-shell vertex function by the on-shell coupling, which largely simplifies calculations, as we discuss in detail. We also point out the great potential of muon beam-dump experiments, such as NA64μ [49], to constrain models of leptophilic ALPs [50]. Indeed, NA64μ already sets relevant constraints on several of the scenarios that we consider, which will substantially improve in coming years with larger data sets.

The rest of this work is organised as follows. In section 2, we define the basic setup of the leptophilic ALP models and discuss the difference between the pseudoscalar and derivative basis and the resulting ALP-photon-photon vertex function. In section 3, we use these results in order to calculate ALP production at leptonic beam dumps in all relevant channels, focusing on E137 and NA64. Constraints on leptophilic ALPs from other experiments are discussed in section 4, and in particular, we derive novel limits by recasting LEP1 data on multi-photon events (largely following ref. [48]). Here, we also collect existing bounds in the literature from B factories, W^+ boson and rare meson decays, SN 1987A, leptonic anomalous magnetic moments and cosmology. Our results are presented in section 5, which mainly consist of novel limits on leptophilic ALPs from E137 and NA64 that have been carefully derived including the track-length distribution of the beam electrons as well as secondary electrons and positrons, which we combine with other existing limits in order to provide a concise overview of all relevant constraints on leptophilic ALPs. We summarise our conclusions in section 6 and include various appendices with details on the experimental setup in appendix A and the cross-section calculations in appendix C. We also compare our results for the exclusion limit from E137 with previous analyses in appendix D.

2 Setup

In this section, we introduce the effective Lagrangians for leptophilic ALP models, discuss in detail the phenomenological and theoretical differences between pseudoscalar and derivative couplings and study the general induced one-loop ALP couplings to photons.

2.1 ALP couplings to fermions

The ALP is assumed to be a pseudoscalar that is not charged under the SM gauge group, arising as a pseudo-Goldstone boson from the spontaneous breaking of a global Peccei-Quinn (PQ) symmetry at some UV scale Λ. At scales much smaller than Λ, the effective Lagrangian is governed by the non-linearly realised PQ symmetry under which the ALP shifts, $a \to a + \text{const}$. While for the QCD axion, this symmetry is only broken by QCD instantons, for the ALP that we consider here, there is also a soft breaking by an explicit ALP mass term. Apart from this mass term, the most general Lagrangian then involves shift-invariant couplings to fermions and anomalous couplings to gauge bosons, just as the QCD axion [51].

Here, we focus on leptophilic ALPs, which are described by an effective Lagrangian only involving charged leptons. Up to higher-dimensional operators and restricting to flavour-diagonal couplings, the general ALP Lagrangian reads

$$\mathcal{L}_D = \frac{1}{2}\partial_\mu a\,\partial^\mu a - \frac{1}{2}m_a^2 a^2 + \frac{\partial_\mu a}{2\Lambda}\sum_{\ell=e,\mu,\tau} C_\ell\,\overline{\psi}_\ell\gamma^\mu\gamma_5\psi_\ell\,, \tag{2.1}$$

with real couplings C_ℓ. Equivalently, one can change the field basis and write this Lagrangian in terms of pseudoscalar couplings to fermions. Under the chiral ALP-dependent rotation of the lepton fields

$$\psi_\ell \to e^{i\frac{C_\ell}{2\Lambda}a\gamma_5}\psi_\ell\,, \qquad\qquad \overline{\psi}_\ell \to \overline{\psi}_\ell\, e^{i\frac{C_\ell}{2\Lambda}a\gamma_5}\,, \qquad\qquad (2.2)$$

the derivative couplings above are cancelled when plugging the transformed fields into the lepton kinetic terms, and the ALP appears in the lepton mass term. Since this transformation is anomalous, i.e. it does not leave the fermion path integral measure invariant, it also gives rise to a contribution to the ALP couplings to photons, so that in the new field basis, the ALP Lagrangian reads (see e.g. refs. [5, 34] for details)

$$\mathcal{L}_{\rm D} = \frac{1}{2}\partial_\mu a\,\partial^\mu a - \frac{1}{2}m_a^2 a^2 - ia \sum_{\ell=e,\mu,\tau}\frac{m_\ell}{\Lambda}C_\ell\overline{\psi}_\ell\gamma_5\psi_\ell + \frac{\alpha}{4\pi}\frac{a}{\Lambda}F^{\mu\nu}\tilde{F}_{\mu\nu}\sum_{\ell=e,\mu,\tau}C_\ell\,, \qquad (2.3)$$

up to higher-order terms in a/Λ, and $\tilde{F}^{\mu\nu} = \epsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}/2$ with $\epsilon^{0123} = 1$. Note that the truncation explicitly breaks the shift symmetry of lepton couplings, which is restored by higher-order terms where the shift symmetry just corresponds to a re-phasing of the leptons.

It is well-known [52, 53] that all physical observables remain unchanged under non-linear field redefinitions of the form in eq. (2.2), which is just a change of variables in the path integral formulation of QFT. In the context of effective field theories, it has been shown in ref. [54] that such field redefinitions are largely equivalent to using the equations of motion for the involved fields (even when off-shell), although care has to be taken beyond the leading order [55]. Working consistently at leading order in $1/\Lambda$, i.e. only using vertices involving a single ALP, the Lagrangians in eq. (2.1) and eq. (2.3) therefore make identical predictions, even though the latter features ALP couplings to *both* leptons and photons.

This equivalence implies that a leptophilic Lagrangian in the pseudoscalar basis defined by

$$\mathcal{L}_{\rm P} = \frac{1}{2}\partial_\mu a\,\partial^\mu a - \frac{1}{2}m_a^2 a^2 - ia\sum_{\ell=e,\mu,\tau}\frac{m_\ell}{\Lambda}C_\ell\overline{\psi}_\ell\gamma_5\psi_\ell\,, \qquad\qquad (2.4)$$

does not give the same results as (2.1) when the corresponding process involves effective vertices to photons. For this reason, there are two "leptophilic" ALP models with distinct phenomenology, which in the following we refer to as the "derivative" (eq. (2.1)) and "pseudoscalar" (eq. (2.4)) model. To follow the conventional notation in the literature, we also introduce the dimensionless coupling $g_{a\ell}$ via the relation

$$\frac{C_\ell}{\Lambda} = \frac{g_{a\ell}}{m_\ell}\,, \qquad\qquad (2.5)$$

such that the pseudoscalar interactions can be written as $\mathcal{L}_{\rm P} \supset -ig_{a\ell}a\overline{\psi}_\ell\gamma_5\psi_\ell$. In order to clearly distinguish between derivative and pseudoscalar leptophilic models, we use the notation C_ℓ/Λ exclusively for the derivative interaction and the coupling $g_{a\ell}/m_\ell$ exclusively for the pseudoscalar interaction. Moreover, we will discuss below four different scenarios, depending on the values of C_ℓ for the three families. Either the ALP couples to a single lepton species, e.g. $C_e \neq 0$, while $C_\mu = C_\tau = 0$, or it has flavour-universal couplings in the derivative basis, i.e. $C_e = C_\mu = C_\tau$, corresponding to universal PQ charges.

Finally, we briefly discuss the question of which type of leptophilic interaction (eq. (2.1) or eq. (2.4)) is best motivated from a UV perspective. Simple UV completions seem to suggest the pseudoscalar form, which can arise from a renormalisable Yukawa coupling with a non-linearly realised Goldstone boson in a straightforward way. For example, one can imagine a standard QCD axion scenario like the DFSZ model [56, 57], in addition to some small explicitly PQ breaking mass term, for example induced by M_{Planck}-suppressed operators that generally spoil the solution to the Strong CP Problem in QCD axion models (the so-called PQ Quality Problem [58]). In DFSZ scenarios, SM fermions are charged under the PQ symmetry, which is spontaneously broken by the Higgs fields (at least two) and a SM singlet scalar. The axion is then a linear combination of the pseudoscalar components in the Higgs fields and the singlet and couples to fermions through the Yukawa interactions (see e.g. refs. [59, 60] for the explicit construction of generalised DFSZ models)

$$\mathcal{L}_{\mathrm{DFSZ}} \supset -\bar{u}_{\mathrm{L}} m_U u_{\mathrm{R}} e^{i\chi_{Hu}\frac{a}{v_{\mathrm{PQ}}}} - \bar{d}_{\mathrm{L}} m_D d_{\mathrm{R}} e^{i\chi_{Hd}\frac{a}{v_{\mathrm{PQ}}}} - \bar{e}_{\mathrm{L}} m_E e_{\mathrm{R}} e^{i\chi_{Hd}\frac{a}{v_{\mathrm{PQ}}}} + \mathrm{h.c.}\,, \qquad (2.6)$$

where $\chi_{H_{u,d}}$ are the PQ charges of the Higgs doublets $H_{u,d}$ and $v_{\mathrm{PQ}} \propto \Lambda$ is the PQ breaking scale.

Expanding the exponential functions, one obtains mass terms for the fermions as well as pseudoscalar ALP-fermion interaction terms, while a tree-level coupling between ALPs and photons is absent in this basis. A chiral redefinition of the fermion fields would trade the pseudoscalar couplings of eq. (2.6) for derivative couplings and generally non-vanishing couplings to gauge bosons, unless the PQ charges are such that the electromagnetic (EM) anomaly coefficient vanishes. In the four leptophilic scenarios discussed above, this anomaly is indeed non-vanishing since, by construction, only leptons contribute, and the leptonic PQ charge matrix is not traceless. Thus, in order to obtain one of these scenarios with derivative lepton couplings and a vanishing photon coupling, one would need to extend the simple DFSZ UV model with new heavy fermions that have suitable PQ charges to cancel the EM anomaly, along the lines of common KSVZ models [61, 62].

On the other hand, since the derivative coupling is the only dimension-5 operator that preserves the shift-symmetry of the ALP, other shift-symmetric couplings (such as those to gauge bosons, e.g. in a KSVZ model) will generally induce this form of coupling through RGE-evolution [63]. An example is the photophobic ALP discussed in ref. [64], which, however, certainly requires more model-building effort than the above-mentioned DFSZ-like constructions. Moreover, as discussed in ref. [64], it comes with additional interactions with the electroweak gauge bosons (which we do not consider here).

In the remainder of this article, we choose to remain agnostic about the nature of the UV completions of both bases and simply derive constraints for both scenarios in eq. (2.1) and eq. (2.4), which will deviate when probing ALP couplings to photons. These couplings are effectively generated from lepton loops and differ for the two interactions by a constant, as we discuss in more detail now.

2.2 Effective ALP coupling to photons

Even though leptophilic ALPs do not couple to photons at tree level, an effective ALP-photon interaction is induced by the triangle diagram shown in figure 1, which corresponds to an
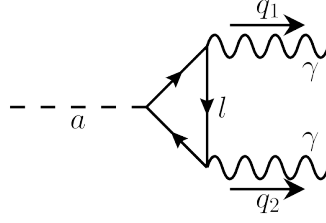
Figure 1. Triangle diagram inducing the effective ALP-photon interaction.

off-shell ALP-photon-photon vertex function $\Gamma^{\mu\nu}_{a^*\gamma^*\gamma^*}(q_1, q_2)$ that depends on the outgoing photon momenta $q_{1,2}$, with incoming ALP momentum $q_a = q_1 + q_2$. It is convenient to introduce an effective coupling as a scalar form factor by

$$\Gamma^{\mu\nu}_{a^*\gamma^*\gamma^*}(q_1, q_2) = i g^{\text{eff}}_{a^*\gamma^*\gamma^*}(q_a^2, q_1^2, q_2^2) q_{1\alpha} q_{2\beta} \epsilon^{\mu\nu\alpha\beta} , \tag{2.7}$$

such that $g^{\text{eff}}_{a^*\gamma^*\gamma^*}(q_a^2, q_1^2, q_2^2)$ is given by a scalar one-loop integral with three external momenta q_a, q_1, q_2, which can be expressed in terms of Passarino-Veltman functions. For ALPs with a derivative coupling to a single lepton species of the form in eq. (2.1), the effective off-shell coupling reads [37] (the superscript "D" denotes the derivative basis)

$$g^{\text{eff,D}}_{a^*\gamma^*\gamma^*}(q_a^2, q_1^2, q_2^2) = \frac{\alpha}{\pi} \frac{C_\ell}{\Lambda} \Big[1 + 2m_\ell^2 C_0 \Big(q_1^2, q_2^2, q_a^2, m_\ell^2, m_\ell^2, m_\ell^2 \Big) \Big] , \tag{2.8}$$

where C_0 is the Passarino-Veldman three-point function defined as

$$C_0 \Big(q_1^2, q_2^2, (q_1 + q_2)^2, m_1^2, m_2^2, m_3^2 \Big) \equiv \int \frac{\mathrm{d}^4 k}{i\pi^2} \frac{1}{(k^2 - m_1^2)(k_1^2 - m_2^2)(k_{12}^2 - m_3^2)} , \tag{2.9}$$

with $k_1 \equiv k - q_1$ and $k_{12} = k - q_1 - q_2$, following the conventions in ref. [65]. If the ALP couples to several leptons, the effective coupling is given by the sum over all contributions.

Since the Lagrangian in eq. (2.1) is equivalent to the one in eq. (2.3) for single ALP vertices, the scalar vertex function for the pseudoscalar coupling in eq. (2.4) differs from eq. (2.8) by a constant, which is given by the coefficient of the $\frac{1}{4} a F \tilde{F}$ operator in eq. (2.3), $\alpha C_\ell / (\pi \Lambda)$. Indeed, for the effective coupling in the pseudoscalar basis (denoted by the superscript "P"), one has

$$g^{\text{eff,P}}_{a^*\gamma^*\gamma^*}(q_a^2, q_1^2, q_2^2) = \frac{\alpha}{\pi} \frac{g_{a\ell}}{m_\ell} \Big[2m_\ell^2 C_0 \Big(q_1^2, q_2^2, q_a^2, m_\ell^2, m_\ell^2, m_\ell^2 \Big) \Big] , \tag{2.10}$$

where we used $g_{a\ell}/m_\ell = C_\ell/\Lambda$.

It is instructive to evaluate the fully off-shell couplings for the simplifying cases where one or more external particles are on-shell. If a single photon is on-shell, we find (in agreement with ref. [36])

$$g^{\text{eff,D}}_{a^*\gamma^*\gamma}(q_a^2, q_1^2) = \frac{\alpha}{\pi} \frac{C_\ell}{\Lambda} \left(1 - \int_0^1 dx \int_0^1 dy \frac{m_\ell^2}{m_\ell^2 - x(1-x)(q_a^2 - 2y q_a q_1)} \right) , \tag{2.11}$$

which is valid in the derivative basis, i.e. using the Lagrangian in eq. (2.1). For the analogous result in the pseudoscalar basis, one simply needs to drop the constant part.

Taking the ALP on-shell footnote The alternative case that both photons are on-shell but the ALP is off-shell is discussed in ref. [66]. $q_a^2 = m_a^2$, the vertex function depends only on a single off-shell photon squared momentum $q_1^2 \equiv t$, and reads

$$g_{a\gamma^*\gamma}^{\text{eff,D}}(t) = \frac{\alpha}{\pi} \frac{C_\ell}{\Lambda} \left(1 - \int_0^1 dx \int_0^1 dy \frac{m_\ell^2}{m_\ell^2 - x(1-x)(m_a^2 - y(m_a^2 + t))} \right) , \qquad (2.12)$$

where we have taken eq. (2.11) and set $2q_a q_1 = m_a^2 + t$. This equality becomes evident by rewriting the momentum q_2 of the on-shell photon as $q_2^2 = (q_a - q_1)^2 = m_a^2 - 2q_a q_1 + t = 0$.

Finally, if all particles are on-shell, i.e. if $t = 0$, the vertex function becomes a constant

$$g_{a\gamma\gamma}^{\text{eff,D}} = \frac{\alpha}{\pi} \frac{C_\ell}{\Lambda} \left(1 - \int_0^1 dx \int_0^1 dy \frac{m_\ell^2}{m_\ell^2 - x(1-x)(1-y)m_a^2} \right) = \frac{\alpha}{\pi} \frac{C_\ell}{\Lambda} \left(1 - \tau_\ell f(\tau_\ell)^2 \right) , \quad (2.13)$$

with $\tau_\ell = 4m_\ell^2/m_a^2$ and the function

$$f(\tau) = \begin{cases} \arcsin\left(\frac{1}{\sqrt{\tau}}\right), & \text{if} \quad \tau \geq 1 \\ \frac{\pi}{2} + \frac{i}{2} \log\left(\frac{1+\sqrt{1-\tau}}{1-\sqrt{1-\tau}}\right), & \text{if} \quad 0 < \tau < 1 \end{cases} , \qquad (2.14)$$

consistent with the expression derived in refs. [5, 37]. This constant can be identified with the effective Lagrangian coupling

$$\mathcal{L}_\gamma = \frac{g_{a\gamma\gamma}^{\text{eff}}}{4} a F^{\mu\nu} \tilde{F}_{\mu\nu} . \qquad (2.15)$$

and controls the ALP decay into two photons. Since, asymptotically,

$$\tau f^2(\tau) = \begin{cases} 1 + \frac{1}{3\tau} & \tau \gg 1 \\ \frac{\tau}{4} \left(\pi + i \ln \frac{4}{\tau} \right)^2 & \tau \ll 1 \end{cases} , \qquad (2.16)$$

the effective photon coupling defined in eq. (2.15) for a derivatively coupled ALP in eq. (2.1) has the limits

$$g_{a\gamma\gamma}^{\text{eff,D}} \propto 1 - \tau_\ell f(\tau_\ell)^2 \propto \begin{cases} -\frac{1}{12} (m_a/m_\ell)^2 & \text{if} \quad m_a \ll m_\ell \\ 1 & \text{if} \quad m_a \gg m_\ell \end{cases} . \qquad (2.17)$$

Instead, the effective photon coupling for an ALP coupled to the pseudoscalar current in eq. (2.4) has the asymptotic behaviour

$$g_{a\gamma\gamma}^{\text{eff,P}} \propto -\tau_\ell f(\tau_\ell)^2 \propto \begin{cases} -1 - \frac{1}{12} (m_a/m_\ell)^2 & m_a \ll m_\ell \\ \mathcal{O}\left(\log^2(m_a/m_\ell) (m_\ell/m_a)^2 \right) & m_a \gg m_\ell \end{cases} . \qquad (2.18)$$

in agreement with ref. [37].

Therefore, for couplings to leptons significantly lighter than the ALP ($\tau_\ell \ll 1$), the contribution to the effective photon coupling coming from C_0 itself is negligible. In the case of derivative couplings, this leaves only a constant that results in an effective coupling of $g_{a\gamma\gamma}^{\text{eff,D}} \approx \alpha C_\ell/(\pi\Lambda)$. Instead, for pseudoscalar couplings, the ALP-photon coupling is strongly

suppressed in this limit. Conversely, in the case of couplings to leptons significantly heavier than the ALP ($\tau_\ell \gg 1$), the limits are effectively reversed. The derivative coupling leads to a strong suppression in $g_{a\gamma\gamma}^{\text{eff,D}}$ due to cancellation between the constant and C_0, while the pseudoscalar coupling leads to a similar result as the tree-level coupling $g_{a\gamma\gamma}^{\text{eff,P}} \approx -\alpha g_{a\ell}/(\pi m_\ell)$, albeit with the opposite sign. The ALP decay rate to two photons will feature the same limiting behaviour, see section 2.3. To summarise, the effective on-shell photon couplings are given by

$$g_{a\gamma\gamma}^{\text{eff,D}} = \frac{\alpha}{\pi}\frac{C_\ell}{\Lambda}\left(1 - \tau_\ell f(\tau_\ell)^2\right), \qquad\qquad g_{a\gamma\gamma}^{\text{eff,P}} = -\frac{\alpha}{\pi}\frac{g_{a\ell}}{m_\ell}\tau_\ell f(\tau_\ell)^2, \qquad (2.19)$$

for derivative and pseudoscalar couplings, respectively.

The loop function can also be calculated analytically for the case in which one of the photons is off-shell, i.e. eq. (2.12). In this case, one finds

$$2m_\ell^2 C_0(0, t, m_a^2, m_\ell^2, m_\ell^2, m_\ell^2) = B\left(\frac{4m_\ell^2}{m_a^2}, \frac{4m_\ell^2}{t}\right), \qquad (2.20)$$

where the function $B(\tau_1, \tau_2)$ is defined as

$$B(\tau_1, \tau_2) = \frac{\tau_1\tau_2}{\tau_1 - \tau_2}\left(f(\tau_1)^2 - f(\tau_2)^2\right). \qquad (2.21)$$

Since t can also be negative, we extend the definition of f to

$$f(\tau) = \begin{cases} \arcsin\left(\frac{1}{\sqrt{\tau}}\right), & \text{if } \tau \geq 1 \\ \frac{\pi}{2} + \frac{i}{2}\log\left(\frac{1+\sqrt{1-\tau}}{1-\sqrt{1-\tau}}\right), & \text{if } 0 < \tau < 1 \\ \frac{i}{2}\log\left(\frac{\sqrt{1-\tau}+1}{\sqrt{1-\tau}-1}\right), & \text{if } \tau < 0 \end{cases} . \qquad (2.22)$$

Note that eq. (2.20) is consistent with eq. (2.13) since in the limit of $t \to 0$

$$B(\tau_1, \tau_2) \xrightarrow{\tau_2 \to \infty} -\tau_1 f(\tau_1)^2, \qquad (2.23)$$

as $\tau_2 f(\tau_2)^2$ becomes constant. Instead, in the limit of $t \gg m_\ell^2$, one has

$$B(\tau_1, \tau_2) \xrightarrow{\tau_2 \to 0} -\tau_2 f(\tau_2)^2 \propto \tau_2 \ln^2 \tau_2, \qquad (2.24)$$

which is particularly relevant for ALP searches with large momentum transfer.

The limit for C_0 in eq. (2.20) is sufficient for our phenomenological discussion since at least one of the photons will be on-shell, or at least approximately so, see section 3.2. To summarise the resulting effective coupling, one has

$$g_{a\gamma^*\gamma}^{\text{eff,D}}(t) = \frac{\alpha}{\pi}\frac{C_\ell}{\Lambda}\left[1 + B\left(\frac{4m_\ell^2}{m_a^2}, \frac{4m_\ell^2}{t}\right)\right], \qquad g_{a\gamma^*\gamma}^{\text{eff,P}}(t) = \frac{\alpha}{\pi}\frac{g_{a\ell}}{m_\ell}B\left(\frac{4m_\ell^2}{m_a^2}, \frac{4m_\ell^2}{t}\right), \qquad (2.25)$$

for derivative and pseudoscalar coupling, respectively. For $t \gg m_\ell^2$, the vertex function for the pseudoscalar couplings is suppressed, while it approaches a constant for the derivative coupling. In the following, we will often suppress the argument, as it is clear from the subscript whether we refer to the on- or off-shell coupling.

2.3 ALP decay rates

The decay rates for the decay of an ALP into a pair of photons or leptons are [5]

$$\Gamma_{a\to\gamma\gamma} = \frac{|g_{a\gamma\gamma}^{\mathrm{eff}}|^2 m_a^3}{64\pi}\,, \tag{2.26}$$

$$\Gamma_{a\to\ell\ell} = \frac{g_{a\ell}^2 m_a}{8\pi}\sqrt{1 - \frac{4m_\ell^2}{m_a^2}} = \frac{C_\ell^2 m_\ell^2 m_a}{8\pi\Lambda^2}\sqrt{1 - \frac{4m_\ell^2}{m_a^2}}\,, \tag{2.27}$$

where the effective on-shell photon coupling for derivative and pseudoscalar couplings is given in eq. (2.19) with the limiting values in eqs. (2.17) and (2.18), respectively. For leptophilic ALPs with arbitrary couplings to the three lepton species, the total decay rate is given by

$$\Gamma_a = \Gamma_{a\to\gamma\gamma} + \sum_{l=e,\mu,\tau} \Gamma_{a\to\ell\ell}\,. \tag{2.28}$$

3   ALP production at lepton beam dumps

Beam-dump experiments are a valuable asset in the search for physics beyond the Standard Model (BSM). The general setup of a beam-dump experiment is depicted schematically in figure 2. A high-intensity beam, usually consisting of protons, electrons or muons, impinges upon a target material. Inside this target, countless Standard Model processes take place, creating particles such as pions or kaons [67]. However, a small fraction of the interactions might produce a weakly interacting BSM particle X, for example, an ALP. In order to shield the detectors from the SM background, the beam dump is positioned downstream of the target. The dump absorbs the shower of particles, and only weakly interacting particles are able to pass. Once past the dump, the particle X may decay back into Standard Model particles, which leads to a signature in the detectors placed downstream of the dump [68].

Beam-dump experiments are well-suited to searches for particles in the mass range $\mathcal{O}(\mathrm{MeV}) - \mathcal{O}(\mathrm{GeV})$ with tiny couplings to Standard Model particles. These feebly interacting particles have long lifetimes and decay lengths and are consequently difficult to detect in collider experiments. In contrast, the long decay volumes of beam-dump experiments allow for greater sensitivity to these types of particles since they are able to decay inside the experiment and be detected [68, 69].

The most sensitive beam-dump experiments for leptophilic ALPs are the ones employing lepton beams. Concretely, we consider the E137 beam-dump experiment performed at SLAC [45] to search for neutral metastable particles, as well as the NA64 beam-dump experiment running at the CERN Super Proton Synchrotron. The latter performs three different types of searches:

- The standard *invisible mode* of the NA64 experiment differs from E137 and figure 2 in that the experiment employs an active beam dump. Instead of the detector being downstream of the target and detecting potential decay products, the target itself is the detector, and the experiment searches for missing energy [46].

- In contrast to the invisible setup, NA64 run in *visible mode* uses a similar search strategy to the one employed by E137 [47].
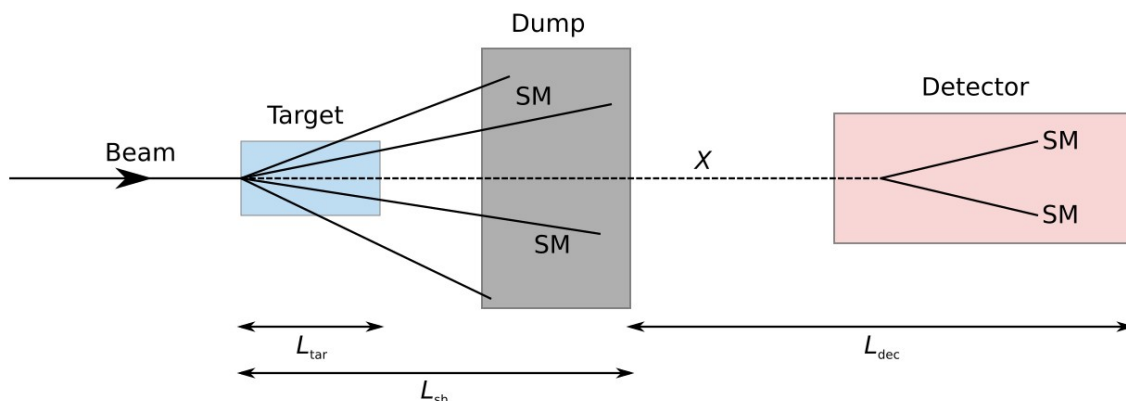
Figure 2. Schematic illustration of a beam dump. A BSM particle X is produced at the target and passes through the dump. It can be detected if it decays into SM particles in the decay volume.

- The NA64μ experiment [49] employs a muon beam to search for new particles that couple primarily to muons, such as the gauge boson of a new $\mathrm{U}(1)_{L_\mu - L_\tau}$ gauge symmetry. Clearly, this *muon mode* also possesses great sensitivity to muonphilic ALP models.

The details of the different experiments are summarised in appendix A.

In the remainder of this section, we explain how the constraints for these experiments are calculated. Complementary constraints from other types of experiments will be discussed in section 4.

3.1 ALP yield in E137

For electron beam-dump experiments like E137 or NA64 run in visible mode, the general formula for the expected number of ALPs that can be detected by the experiment is given by the expression [70]

$$
N_a = \frac{N_e X}{M_{\text{target}}/N_{\text{avo}}} \int_{E_{\min}}^{E_{\max}} \mathrm{d}E \int_{x_{\min}}^{x_{\max}} \mathrm{d}x \int_0^T \mathrm{d}t \, I_e(E_0, E, t) \frac{\mathrm{d}\sigma}{\mathrm{d}x} e^{-L_{\text{sh}}\left(\frac{1}{l_a} + \frac{1}{l_\lambda}\right)} \left(1 - e^{-\frac{L_{\text{dec}}}{l_a}}\right),
\tag{3.1}
$$

with the number of electrons on target (EOT) N_e, the unit length of the radiation target X, the molar mass of the target atom M_{target}, and the Avogadro constant N_{avo}. Furthermore, L_{sh} denotes the length of the beam dump, while L_{dec} is the length of the decay volume, i.e. the distance from the end of the beam dump to the detector. A summary of the experimental parameters of E137 and NA64 is given in appendix A.3.

The variable $x = E_a/E$ defines the ratio of the energy transferred from an electron (or muon for muon beam dumps) of energy E to an ALP with energy E_a. The lower integration bound for x can be determined from the requirement that the ALP energy must be greater than the energy threshold of the detector E_{cut}, as well as the ALP rest mass. Assuming the ALP and electron/muon initial and final states are collinear [71], the upper bound can be approximated by requiring energy conservation in the case that the outgoing electron/muon

is at rest, with the entire energy (excluding the lepton's rest mass) being transferred to the ALP. This means the bounds for x are

$$x_{\min} = \frac{\max(m_a, E_{\text{cut}})}{E}, \quad x_{\max} = \frac{E - m_\ell}{E} = 1 - \frac{m_\ell}{E}. \tag{3.2}$$

In this work, we only consider ALPs with mass below $1\,\text{GeV}$, i.e. below the energy threshold, such that $x_{\min} = E_{\text{cut}}/E$.

The upper integration boundary for the electron energy is given by the incident beam energy $E_{\max} = E_0$. The energy of the electrons decreases after impacting the beam dump due to shower creation and scattering processes [72]. This is not the case for muon beams (see section 3.4 below). The lower energy bound is given for the electron energy $E_{\min}$ at which the bounds $x_{\min,\max}$ are equal. This integration range for the energy E is, therefore, in the interval between

$$E_{\min} = E_{\text{cut}} + m_\ell, \quad E_{\max} = E_0. \tag{3.3}$$

The function $I_e(E_0, E, t)$ denotes the energy distribution function for electrons with an energy E at a depth t inside the beam dump. The maximum depth for t is given by $T = \rho L_{\text{tar}}/X$, where ρ is the density of the target and L_{tar} is the thickness of the target.

In this work, we consider the primary shower of electrons $I_e^{(1)}$, as well as the secondary shower of electrons and positrons $I_{e,p}^{(2)}$, which are created through pair production from photons in the shower [72]. We use the expressions for $I_e^{(1,2)}$ derived in ref. [72]. However, we do not use the simplified distributions, which assume that $E_0 - E \ll E_0$, but instead use the full expressions for the distributions, see appendix B. Since the electron energy distribution is the only function of t in the integral (3.1), we define the integrated energy distribution

$$I_{\text{int}}(E_0, E, T) \equiv \int_0^T \text{d}t\, I_e(E_0, E, t). \tag{3.4}$$

The integrated distributions for first and second-generation electrons/positrons in E137 are shown in figure 3. In addition, we indicate the distribution of first-generation electrons obtained when using the approximation $E_0 - E \ll E_0$. The approximate expression differs substantially from the exact first-generation distribution. However, it coincidentally matches the more complete result, which includes second-generation effects, quite well.

The absorption of ALPs by electrons in the target is described by l_λ, which leads to exponential attenuation. This is relevant for thick targets but can be neglected here [70]. The ALP decay length l_a depends on the couplings as well as the energy and mass of the ALP. It can be written as

$$l_a = \frac{p_a}{m_a}\frac{1}{\Gamma_a} = \frac{\sqrt{E_a^2 - m_a^2}}{m_a}\frac{1}{\Gamma_a} = \frac{\sqrt{x^2 E^2 - m_a^2}}{m_a}\frac{1}{\Gamma_a}. \tag{3.5}$$

The first exponential function in eq. (3.1) can be understood as the probability that the ALP only decays after leaving the beam dump, such that the decay products do not get absorbed. Conversely, the last term describes the probability that the ALP does not decay behind the detector, i.e. that it decays inside the decay volume and can be detected.
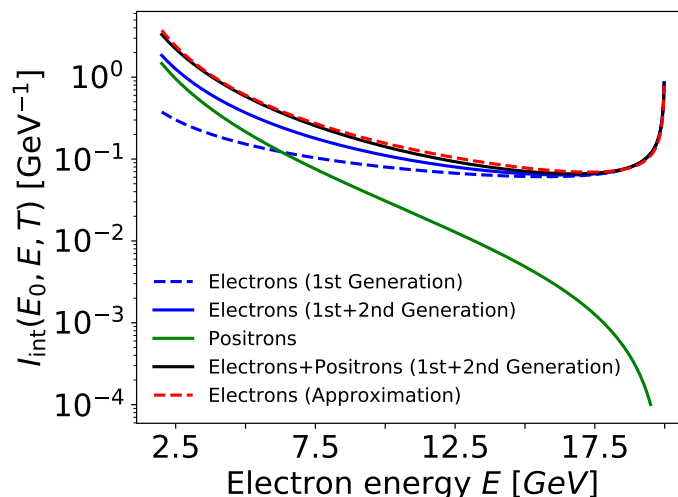
Figure 3. Integrated electron/positron energy distributions for E137.

With the simplifications discussed above, the general formula we use for the expected number of detected ALPs is

$$N_a = \frac{N_e X}{M_{\text{target}}/N_{\text{avo}}} \int_{E_{\min}}^{E_0} \mathrm{d}E \int_{x_{\min}}^{x_{\max}} \mathrm{d}x \, I_{\text{int}}(E_0, E, T) \frac{\mathrm{d}\sigma}{\mathrm{d}x} e^{-\frac{L_{\text{sh}}}{l_a}} \left(1 - e^{-\frac{L_{\text{dec}}}{l_a}}\right). \quad (3.6)$$

The only missing ingredient is the differential cross section $\mathrm{d}\sigma/\mathrm{d}x$ for the production of ALPs, which will be discussed next.

3.2 Dark Bremsstrahlung and Primakoff production

The main ALP production mechanisms in beam-dump experiments are Dark Bremsstrahlung (DB) and Primakoff production, shown in the top row of figure 4. In the DB diagrams, the leptophilic ALP couples to the beam electron/muon at tree level, while for the Primakoff process, it only couples to the photons via the effective loop coupling.

Due to the high energy of the beam compared to the ALP and lepton masses, the photon that is exchanged between the nucleus and the electron/muon is approximately on-shell [73]. As a result, the nucleus can be replaced by an effective flux of photons, which is described by the Improved Weizsäcker-Williams (IWW) approximation. The IWW approximation allows the cross section to be calculated using the reduced $2 \to 2$ scattering amplitude $\overline{|\mathcal{M}|}^2$. Thus, for a lepton scattering with a nucleus

$$\ell^-(p) + N(P_i) \to \ell^-(p') + N(P_f) + a(k), \quad (3.7)$$

the simplified photoproduction processes are

$$\ell^-(p) + \gamma(q) \to \ell^-(p') + a(k). \quad (3.8)$$

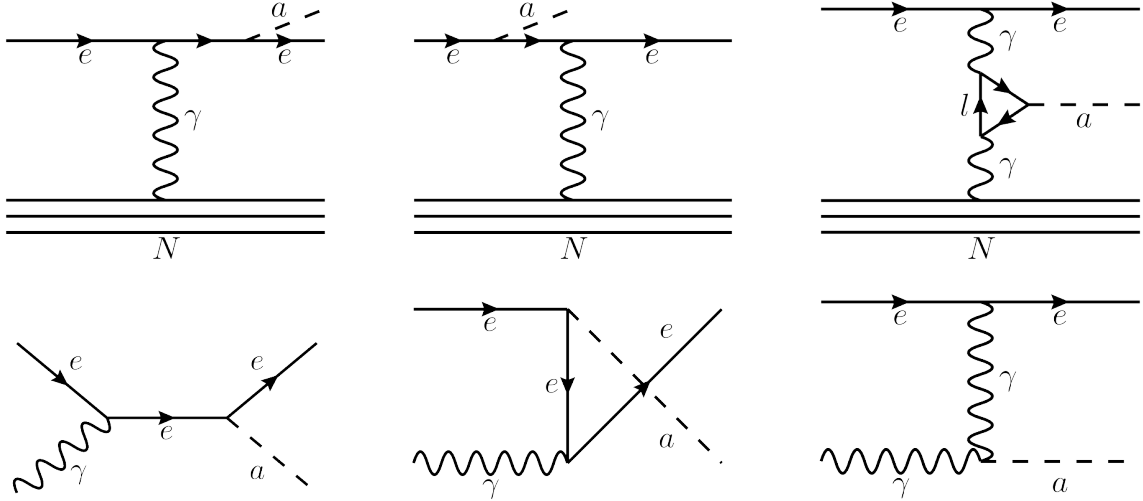The simplified DB and Primakoff diagrams are shown in the bottom row of figure 4.

Figure 4. Top row: Feynman diagrams contributing to the process $eN \to eNa$. The diagrams correspond from left to right to the DB s-channel process, the DB u-channel process and the Primakoff process. Bottom row: corresponding Feynman diagrams contributing to the process $e\gamma \to ea$. Note that the lepton loop has been replaced with the effective Primakoff coupling.

The differential cross section for the tree-level DB process can be calculated analytically (see appendix C.1) and is found to be

$$
\frac{\mathrm{d}\sigma_{\mathrm{DB}}}{\mathrm{d}x} = \frac{\alpha}{(4\pi)^2} \sqrt{x^2 - \frac{m_a^2}{E^2}} \chi \frac{1-x}{x}
$$

$$
\times \frac{e^2 E^2 g_{a\ell}^2 \theta_{a,\mathrm{max}}^2 x^5}{3(1-x)\left(m_a^2(1-x) + m_\ell^2 x^2\right)^2 \left(m_a^2(1-x) + x^2\left(E^2\theta_{a,\mathrm{max}}^2 + m_\ell^2\right)\right)^3}
$$

$$
\times \left[3m_a^4(1-x)^2 x^2 \left(E^2\theta_{a,\mathrm{max}}^2 + 3m_\ell^2\right) + 3x^6 \left(E^2\theta_{a,\mathrm{max}}^2 m_\ell + m_\ell^3\right)^2 \right.
$$

$$
\left. + m_a^2(1-x)x^4 \left(2E^4\theta_{a,\mathrm{max}}^4 + 9E^2\theta_{a,\mathrm{max}}^2 m_\ell^2 + 9m_\ell^4\right) + 3m_a^6(1-x)^3 \right],
$$

where $\theta_{a,\mathrm{max}}$ denotes the largest angle relative to the beam axis that the ALP can have in order to be detected. The Primakoff process and the interference term involve the momentum-dependent effective coupling to photons, so we do not obtain a simple analytical expression for the corresponding differential cross sections. Instead, we evaluate them numerically.

In order to analyse the mass and energy dependence of the cross sections, we plot the contributions to σ (normalised by the coupling) coming from DB, Primakoff, and interference processes in E137 in figure 5. We assume that the ALP couples exclusively to electrons and compare the normalised cross sections for two masses: $m_a = 10\,\mathrm{MeV}$ and $m_a = 450\,\mathrm{MeV}$.

For $m_a = 10\,\mathrm{MeV}$, the DB cross section is clearly the main contributor to the total cross section. This is to be expected, considering these interactions are mediated by tree-level interactions between the ALPs and electrons. However, it would be wrong to conclude that the Primakoff process does not give important contributions elsewhere, as can be seen for $m_a = 450\,\mathrm{MeV}$. For this mass, Primakoff production outweighs DB production, assuming the ALP exhibits a derivative coupling to electrons (while for pseudoscalar coupling, the
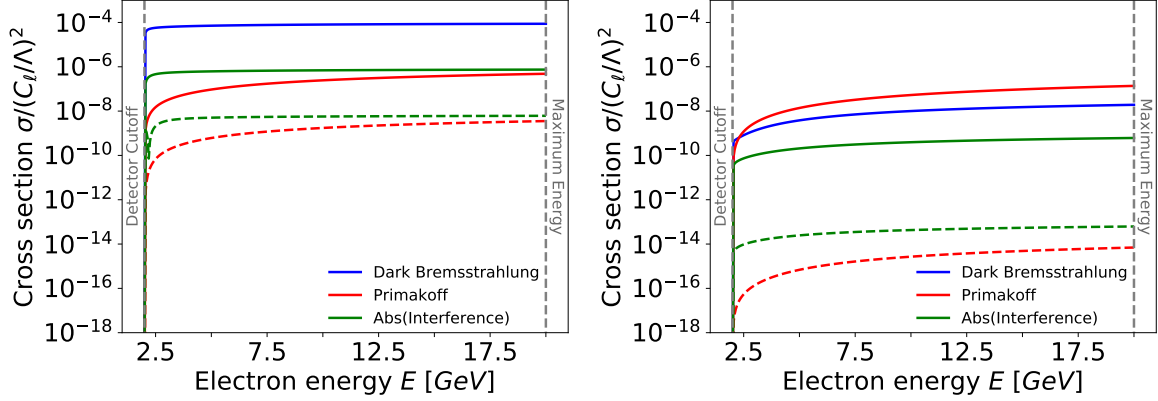
Figure 5. Normalised cross sections for non-vanishing ALP coupling to electrons as a function of the electron energy E. The ALP mass is $m_a = 10\,\text{MeV}$ on the left, $m_a = 450\,\text{MeV}$ on the right. The solid lines represent derivative couplings, while the dashed lines represent pseudoscalar couplings.

Primakoff contribution is again negligible). This mainly happens because, while it is true that all cross sections decrease for an increasing ALP mass, the Primakoff one for derivative couplings decreases only very slowly with respect to the other ones, hence eventually becoming dominant. This difference arises due to the fact that scale-invariant contribution to the effective photon coupling remains constant, while the loop factor is suppressed.

It should be noted that comparing σ is only illustrative, as the actual integration over x weighs $\mathrm{d}\sigma/\mathrm{d}x$ with the exponential functions given at the end of eq. (3.6). However, it is clear that while the Primakoff production of leptophilic ALPs occurs only via an effective loop coupling, it can still have an impact on constraints that can be set on the ALP mass and couplings.

Finally, there are still the contributions arising from secondary positrons created in the shower. Both the DB and Primakoff processes give exactly the same contributions when the electron line in the Feynman diagrams is replaced with a positron line. The only difference between the contributions are the electron and positron energy distribution functions $I_{e,p}$. This means that instead of considering the electron and positron production processes separately, we summarise them together by using the total electron-positron distribution function

$$I_{\text{int}}(E_0, E, T) = \int_0^T \mathrm{d}t\Big(I_e^{(1)}(E_0, E, t) + I_e^{(2)}(E_0, E, t)\Big) + \int_0^T \mathrm{d}t I_p^{(2)}(E_0, E, t) \qquad (3.10)$$

$$= \int_0^T \mathrm{d}t\Big(I_e^{(1)}(E_0, E, t) + 2I_e^{(2)}(E_0, E, t)\Big). \qquad (3.11)$$

For muonphilic ALPs, there can also be ALP bremsstrahlung from secondary muons produced in the electromagnetic shower [74]. Comparing our results to the ones in ref. [74], we find that the loop-induced Primakoff process dominates ALP production for muonphilic ALPs. Therefore, we do not include the contributions of secondary muons in the present work.

3.3 Annihilation processes

In addition to the contributions from DB and Primakoff production, we also consider ALP production from processes in which a positron created in a secondary shower annihilates
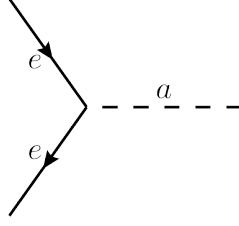
Figure 6. Resonant ALP production through the process $e^- e^+ \to a$.

with an electron inside the beam dump. Since only positrons from the shower contribute to these processes, the first important change compared to DB and Primakoff production is that the integrated energy distribution appearing in eq. (3.6) only contains the contributions from positrons, i.e.

$$I_{\mathrm{int}}(E_0, E, T) = \int_0^T \mathrm{d}t \, I_p^{(2)}(E_0, E, t) \,. \tag{3.12}$$

If the annihilation of the positron and electron has a centre-of-mass energy that is exactly $\sqrt{s} = m_a$, an ALP can be produced resonantly:

$$e^-(p_1) + e^+(p_2) \to a(k) \,. \tag{3.13}$$

The Feynman diagram for this process is shown in figure 6. The matrix element for this process is simply

$$\overline{|\mathcal{M}_{\mathrm{R}}|}^2 = \frac{g_{ae}^2 s}{2} \,, \tag{3.14}$$

and the differential cross section is given by

$$\frac{\mathrm{d}\sigma_{\mathrm{R}}}{\mathrm{d}x} = \frac{\pi g_{ae}^2}{4 m_e \sqrt{1 - \frac{4 m_e^2}{m_a^2}}} \delta\left(E - \left(\frac{m_a^2}{2 m_e} - m_e \right) \right) \delta\left(x - \left(1 + \frac{m_e}{E} \right) \right) \,. \tag{3.15}$$

The first delta distribution in eq. (3.15) is equivalent to the condition $s = m_a^2$. After inserting this cross section into eq. (3.6), the two delta distributions cancel out the integrations, leaving behind

$$N_{a,\mathrm{res}} = \frac{N_e X}{M_{\mathrm{target}}/N_{\mathrm{avo}}} \frac{\pi g_{ae}^2}{4 m_e \sqrt{1 - \frac{4 m_e^2}{m_a^2}}} I_{\mathrm{int}}(E_0, E_*, T) e^{-\frac{L_{\mathrm{sh}}}{l_{a*}}} \left(1 - e^{-\frac{L_{\mathrm{dec}}}{l_{a*}}} \right) \,. \tag{3.16}$$

Here, E_* and l_{a*} are the positron energy and decay length, respectively, after applying the conditions given by the two delta distributions in eq. (3.15). The mass range in which resonant production can be detected is bounded from below by the fact that the ALP must have an energy larger than the cutoff energy of the detector and from above by the maximum possible positron energy, i.e.

$$E_a = E + m_e = \frac{m_a^2}{2 m_e} > E_{\mathrm{cut}} \,, \tag{3.17}$$

$$E = \frac{m_a^2}{2 m_e} - m_e < E_0 \,. \tag{3.18}$$

This means resonantly produced ALPs can only be detected when the ALP mass is in the interval $\sqrt{2m_e E_{\text{cut}}} < m_a < \sqrt{2m_e(E_0 + m_e)}$.

In addition to the resonant production discussed above, electron-positron annihilation can also proceed non-resonantly. The corresponding cross sections are discussed in appendix C.2.

3.4 ALP yields in NA64

While eq. (3.6) holds for E137 and the visible mode setup of NA64, there are some relevant changes when calculating the expected yield of ALPs for the invisible setup as well as the muon setup of NA64.

NA64 in invisible mode. Since the invisible setup of NA64 is an active beam dump that measures the missing energy of an event instead of detecting particles downstream of the dump, one can consider the decay volume to be infinitely large, i.e. one can set $L_{\text{dec}} \to \infty$. This simplifies the equation for N_a to

$$N_a = \frac{N_e X}{M_{\text{target}}/N_{\text{avo}}} \int_{E_{\text{min}}}^{E_0} \mathrm{d}E \int_{x_{\text{min}}}^{x_{\text{max}}} \mathrm{d}x I_{\text{int}}(E_0, E, T) \frac{\mathrm{d}\sigma}{\mathrm{d}x} e^{-\frac{L_{\text{sh}}}{l_a}} \,, \tag{3.19}$$

where the various quantities that enter this equation are given in appendix A.3. In principle, it would also be possible to simplify the expression for $\mathrm{d}\sigma/\mathrm{d}x$, since there is no longer a restriction on the emission angle θ_a of the ALP. Since the main contribution to the cross section comes from small angles [75], one can safely extend the integration to large angles of order unity. Concretely, in analogy to E137, we introduce an arbitrary cutoff for the angle based on the geometry of the detector given by

$$\theta_{a,\text{max}} = \frac{R_{\text{sh}}}{L_{\text{sh}}} \,, \tag{3.20}$$

where $R_{\text{sh}} \approx 60 \,\text{cm}$ is the approximate radius of the detector, in this case of the hadronic calorimeters (HCALs).

NA64μ. The general setup of the NA64μ experiment is similar to that of NA64 run in invisible mode, with it being an active beam dump. This means that the formula (3.19) applies to NA64μ as well, with the caveat that the electron lines in the DB and Primakoff diagrams in figure 4 need to be replaced with muon lines, and the EOT N_e should be replaced with the muons on target (MOT) N_μ. Since there are no muons for the anti-muons in the beam to annihilate with, the contribution discussed in section 3.3 is absent. Finally, an important difference is the energy distribution function of the muons compared to the electrons. Since muons are much more massive than electrons, they lose significantly less energy due to bremsstrahlung compared to electrons, allowing them to easily penetrate the beam dump [49]. This means the muon's energy distribution function is independent of the depth t. Therefore, we can use the thin-target approximation to calculate the distribution function for muons, which is given by [72]

$$I_\mu(E_0, E, t) = \delta(E_0 - E) \quad \Rightarrow \quad I_{\text{int}}(E_0, E, T) = T\delta(E_0 - E) \,. \tag{3.21}$$

This means the expected yield of ALPs simplifies to

$$N_a = \frac{N_\mu X}{M_{\text{target}}/N_{\text{avo}}} T \int_{x_{\min}}^{x_{\max}} \mathrm{d}x \frac{\mathrm{d}\sigma}{\mathrm{d}x} e^{-\frac{L_{\text{sh}}}{l_a}} \,, \tag{3.22}$$

where it is implied that $E = E_0$ in the expressions of $x_{\min,\max}$, l_a, and the cross section. It is important to note that this makes N_a very sensitive to the length T of the beam dump. This is in contrast with the other electron beam dumps, in which the electron shower attenuates after several radiation lengths.

4   Other constraints

In this section we review other experimental and theoretical constraints of relevance for leptophilic ALPs.

4.1   Perturbative unitarity

A simple but important constraint on the ALP coupling to leptons can be obtained by considering the scattering process $\ell^+\ell^- \to \ell^+\ell^-$ via s-channel ALP exchange. Even though the ALP interaction is non-renormalisable, the matrix element for this process becomes independent of the centre-of-mass energy for $\sqrt{s} \gg m_a$. The requirement of unitarity of the $J = 0$ partial wave a_0, i.e. $|\text{Re } a_0| < 1/2$, implies [38]

$$\frac{C_\ell}{\Lambda} = \frac{g_{a\ell}}{m_\ell} < \frac{1}{m_\ell}\sqrt{\frac{8\pi}{3}} \,. \tag{4.1}$$

4.2   LEP bounds on exotic Z boson decays

Ref. [48] studied constraints on ALP couplings to photons from LEP measurements on the Z pole. Here, we apply this analysis to the present setup, where photon couplings are induced via lepton loops. As we will see, these loops are strongly suppressed by m_ℓ^2/M_Z^2 for the case of pseudoscalar ALP couplings, meaning relevant constraints from LEP only arise for derivative couplings.

There are two processes of interest: $e^+e^- \to Z \to a\gamma$ via on-shell production of a Z boson, and $e^+e^- \to \gamma^* \to a\gamma$ via an off-shell photon at the Z peak. For light ALPs with mass below m_π, the photons from the radiative ALP decay are very collimated, so that both processes look like the forbidden SM decay $Z \to 2\gamma$ [76]. For Z-boson decays we follow ref. [48] and use the bound

$$\text{BR}(Z \to a\gamma) \times \text{BR}(a \to \gamma\gamma) \times P(l_a) \le 5.2 \times 10^{-5} \,, \tag{4.2}$$

which is inferred from the 95% confidence level LEP limit $\text{BR}(Z \to \pi^0\gamma) \le 5.2 \times 10^{-5}$ [77]. Here, we conservatively include only the fraction $\text{BR}(a \to \gamma\gamma)$ of ALPs decaying into two photons since it is unclear whether a leptonically decaying ALP could mimic a π^0 decay when strongly boosted. Moreover, the factor

$$P(l_a) = 1 - e^{-\frac{L}{l_a}} \,, \tag{4.3}$$

gives the probability that the ALP decays inside the decay volume and can be detected. Here, l_a is the ALP decay length in the laboratory frame (see section 3.1), and we take $L = 10\,\mathrm{cm}$. The above limits are used for $m_a \leq m_{\pi^0} = 135\,\mathrm{MeV}$. For larger ALP masses up to $1\,\mathrm{GeV}$ we use the limits obtained in ref. [48] from simulating angular distributions for the three-photon decay ($e^+e^- \to Z \to a + \gamma \to 3\gamma$), which were compared bin by bin to the distributions given in ref. [77]. The resulting bounds are weaker than those for $m_a \leq m_{\pi^0}$ by about 30%.

These constraints can be directly interpreted as a bound on lepton couplings, using the decay rates for derivative (D) and pseudoscalar (P) couplings [36]

$$\Gamma(Z \to a\gamma)_{\mathrm{D}} = \frac{m_Z^3}{96\pi^3} \frac{\alpha^2}{s_{\mathrm{w}}^2 c_{\mathrm{w}}^2} \frac{|C_\ell|^2}{\Lambda^2} \left(s_{\mathrm{w}}^2 - 1/4\right)^2 |1 + B(\tau_\ell, \tau_Z)|^2 \left(1 - \frac{m_a^2}{m_Z^2}\right)^3,$$

$$\Gamma(Z \to a\gamma)_{\mathrm{P}} = \frac{m_Z^3}{96\pi^3} \frac{\alpha^2}{s_{\mathrm{w}}^2 c_{\mathrm{w}}^2} \frac{|g_{a\ell}|^2}{m_\ell^2} \left(s_{\mathrm{w}}^2 - 1/4\right)^2 |B(\tau_\ell, \tau_Z)|^2 \left(1 - \frac{m_a^2}{m_Z^2}\right)^3, \tag{4.4}$$

with s_{w} denoting the sine of the Weinberg angle, $\tau_Z = 4m_\ell^2/M_Z^2$, τ_ℓ is defined below eq. (2.13), and $B(\tau_1, \tau_2)$ given in eq. (2.21). Since here we consider $m_a \ll m_Z$, one has $\tau_Z \ll \tau_\ell$, so that to good approximation

$$B(\tau_\ell, \tau_Z) = \tau_Z \left(f^2(\tau_\ell) - f^2(\tau_Z)\right) \ll 1. \tag{4.5}$$

Constraints on the pseudoscalar scenario are thus much weaker than those in the derivative scenario. In fact, they are weaker than the perturbative unitarity bounds for all leptons and, therefore, irrelevant.

For the derivative scenario, the constraints are more relevant but still rather weak due to the small Z boson couplings to SM leptons resulting from the accidental cancellation $s_w^2 \approx 1/4$. As we will see below, the LEP limits from Z decays are coincidentally of the same magnitude as the limits from ALP production via off-shell photons.

Indeed, the same signature as $Z \to a\gamma$ is generated from ALP production via off-shell photons $e^+e^- \to \gamma^* \to a\gamma$ with $s = M_Z^2$. Therefore, the cross-section is limited by

$$\begin{aligned}
\sigma(e^+e^- \to \gamma^* \to a\gamma)|_{s=M_Z^2} &\leq \sigma_{\mathrm{limit}}(e^+e^- \to Z \to a\gamma)|_{s=M_Z^2} \\
&= \frac{12\pi}{m_Z^2} \mathrm{BR}(Z \to e^+e^-)\mathrm{BR}_{\mathrm{limit}}(Z \to a\gamma) \\
&= \frac{7.9 \times 10^{-9}}{\mathrm{GeV}^2} \times \frac{1}{\mathrm{BR}(a \to \gamma\gamma)} \times \frac{1}{P(l_a)}, \tag{4.6}
\end{aligned}$$

where we have used the narrow-width approximation in the second line and eq. (4.2) in the third line. The cross section at the Z pole can be calculated using the general vertex function with a single off-shell photon in the derivative and pseudoscalar scenario, respectively (cf. eq. (2.25))

$$\begin{aligned}
\Gamma_{a\gamma^*\gamma}^{\mu\nu,\mathrm{D}} &= i\frac{\alpha C_\ell}{\pi\Lambda} \epsilon^{\mu\nu\alpha\beta} q_\alpha k_\beta \left[1 + B(\tau_\ell, \tau_Z)\right], \\
\Gamma_{a\gamma^*\gamma}^{\mu\nu,\mathrm{P}} &= i\frac{\alpha g_{a\ell}}{\pi m_\ell} \epsilon^{\mu\nu\alpha\beta} q_\alpha k_\beta B(\tau_\ell, \tau_Z), \tag{4.7}
\end{aligned}$$

where the external on-shell photon has outgoing momentum q and polarisation index μ, k is the outgoing momentum of the axion, and $p = k + q$ is the inflowing momentum of the off-shell photon with polarisation index ν, with $p^2 = s = M_Z^2$. The total cross-section is given by

$$\sigma(e^+ e^- \to \gamma^* \to a\gamma)_{\mathrm{D}} = \frac{\alpha^3}{24\pi^2} \frac{|C_\ell|^2}{\Lambda^2} \frac{\left(1 - \frac{m_a^2}{m_Z^2}\right)^3 \left(1 + \frac{2m_\ell^2}{m_Z^2}\right)}{\sqrt{1 - 4m_\ell^2/M_Z^2}} |1 + B(\tau_\ell, \tau_Z)|^2 \, ,$$

$$\sigma(e^+ e^- \to \gamma^* \to a\gamma)_{\mathrm{P}} = \frac{\alpha^3}{24\pi^2} \frac{|g_{a\ell}|^2}{m_\ell^2} \frac{\left(1 - \frac{m_a^2}{m_Z^2}\right)^3 \left(1 + \frac{2m_\ell^2}{m_Z^2}\right)}{\sqrt{1 - 4m_\ell^2/m_Z^2}} |B(\tau_\ell, \tau_Z)|^2 \, , \qquad (4.8)$$

Again, constraints on the lepton couplings in the pseudoscalar scenario are weaker than those from perturbative unitarity, such that constraints from ALP production at LEP are relevant only in the derivative scenario. They are accidentally of the same size as the constraints from Z boson decays, since

$$\frac{\Gamma_Z}{12\pi M_Z \mathrm{BR}(Z \to e^+ e^-)} \frac{e^2 s_{\mathrm{w}}^2 c_{\mathrm{w}}^2}{(s_{\mathrm{w}}^2 - 1/4)^2} \simeq 0.94 \, . \qquad (4.9)$$

The resulting constraints on the lepton couplings in the derivative scenario are shown in the $(m_a - C_\ell/\Lambda$ plane) in figure 7 for all three individual lepton couplings as well as the universal case, with the limits from perturbative unitarity shown in grey. Note that, as explained above, for $m_a \leq m_{\pi^0}$ the bounds are slightly stronger than above, due to highly collimated photons looking like the forbidden $Z \to \gamma\gamma$ decay [48]; moreover, the spikes in correspondence of $2m_e$ and $2m_\mu$ originate from the leptonic ALP decay becoming allowed, hence lowering the corresponding $\mathrm{BR}(a \to \gamma\gamma)$; finally, the plateau observed for the muon case between $\sim 100\,\mathrm{MeV}$ and $2m_\mu$, and for the tau case above $\sim 400\,\mathrm{MeV}$, correspond to the ALP mass values where $P(l_a) \simeq 1$.

4.3 B factories

In ref. [78], the BaBar collaboration reported results from a search for dark photons A' produced in the reaction $e^+ e^- \to \gamma A'$, $A' \to \ell^+ \ell^-$ with $\ell = e, \mu$ in terms of a 90% CL upper bound on the kinetic mixing parameter ϵ as a function of the dark photon mass $m_{A'}$. We reinterpret these results in terms of the couplings of leptophilic ALPs by comparing the cross section for dark photon production with the one for non-resonant production of ALPs, see appendix C.2.

Since the BaBar search focuses on dark photon decays to light charged leptons, the bound does not apply to models of ALPs coupled exclusively to tau leptons. For ALPs coupled exclusively to electrons, we only reinterpret the bounds up to the muon production threshold, $m_a < 2m_\mu$. Similarly, for ALPs coupled exclusively to muons, we only reinterpret the BaBar bound for $m_a > 2m_\mu$. Note that even though there is no tree-level coupling to electrons in this case, the ALP can still be produced through the s-channel process $e^- e^+ \to \gamma^* \to a\gamma$ via the effective photon coupling induced through the muon loop.

In ref. [79], the Belle II collaboration reported results from their search for a muonphilic scalar S via the process $e^- e^+ \to \mu^- \mu^+ S$ with $S \to \mu^- \mu^+$. The Belle II collaboration assumed

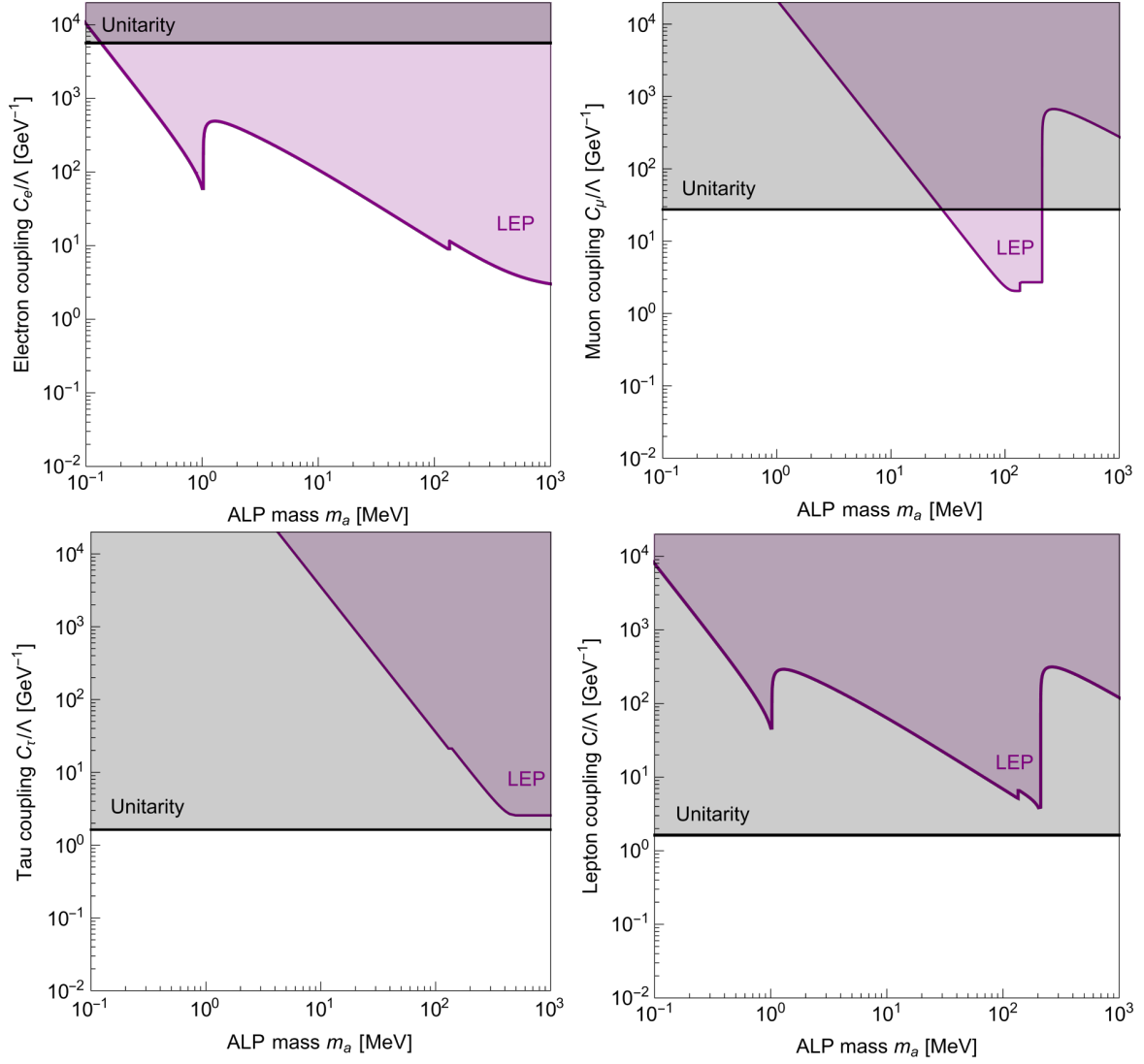**Figure 7.** Constraints on ALP derivative couplings from LEP searches for $Z \to 3\gamma$ from direct ALP production either via on-shell or off-shell photons (purple), see text for details. Also shown are the limits from perturbative unitarity (gray), which always dominate over analogous constraints on pseudoscalar couplings.

an interaction Lagrangian of the form $\mathcal{L}_{\text{int}} \supset -i g_{a\mu} a \overline{\psi}_\mu \psi_\mu$, which allows us to directly use their results, neglecting the small differences of scalars compared to pseudoscalars. Moreover, this constraint applies to both the derivative and pseudoscalar coupling cases since it involves a tree-level coupling between the muon and the ALP.

4.4 Meson and W^+ boson decays

Ref. [35] pointed out the possibility of searching for leptophilic ALPs using rare meson decays such as $\pi^+ \to e^+ \nu a$ [80] or $K^+ \to e^+ \nu a$ [81], or W^+ boson decays such as $W^+ \to e^+ \nu a$. The resulting constraints can be directly applied to our setup, noting that the derivative scenario corresponds to the "weak-violating" ALP in ref. [35] and the pseudoscalar scenario

to the "weak-conserving" ALP. These terms are defined above their eq. (7), from which, along with their eq. (1), one can directly see that the "weak-violating" model is identical with our derivative coupling case in eq. (2.1). On the other hand, the "weak-conserving" model is better mapped to our scenarios using the basis in their eq. (3), which for the "weak-conserving" model contains the pseudoscalar coupling, but also the anomalous coupling between ALP and the photon, so it is not identical with the pseudoscalar scenario in eq. (2.4). Nevertheless, as can also be seen from their figure 1, the limits from rare meson decays in ref. [35] only constrain the first and last terms in their eq. (3), and the last term vanishes for the "weak-conserving" ALP. Thus, in practice, this case constrains our pseudoscalar scenario.

The decay $\pi^+ \to e^+ \nu a$ can also occur in air showers, in which case the subsequent ALP decays may be observed with Super-Kamiokande [82]. For the case of ALPs coupling only to tau leptons, ref. [83] derived constraints from BESIII searches for $J/\psi \to \gamma + \mathrm{inv}$ [84] and $J/\psi \to 3\gamma$ [85].

4.5 Supernova bounds

A burst of ALPs can be produced by the electrons, muons or photons that are present in large numbers in the hot and dense plasmas of core-collapse supernovae. In 1987, one of these explosions, named SN 1987A, was observed in the Large Magellanic Cloud, around 50 kpc away from Earth. This is the only supernova to date for which the accompanying neutrino emission was measured by experiments here on Earth. The duration of this neutrino burst is in overall agreement with simulations [86], which can be used to constrain ALP models. If any exotic particle would be produced in the central neutrinosphere of the nascent proto-neutron star in the supernova core and was coupled so weakly that it escaped without re-depositing its energy, it would provide an additional cooling channel that is not present in simulations. Following approximate, order-of-magnitude physical arguments supported by early simulations of the explosion, the additional cooling power cannot be larger than the neutrino luminosity — otherwise the duration of the neutrino burst would be significantly shortened, contrary to observations [87, 88].

Furthermore, one can derive a bound from the ALPs transferring energy into the mantle of the progenitor star, which would contribute to the explosion energy of the circumstellar material [89, 90]. For typical supernovae such as SN 1987A, this explosion energy is observed to be around 10^{51} erg, about one percent of the total neutron star binding energy released in the explosion [91]. There are SNe with explosion energies lower by about an order of magnitude, which might also be used to constrain ALP energy deposition [92]. Finally, we point out that recently refs. [44, 93] have argued for modifications of both the cooling and energy deposition constraint in the strong coupling regime, which were not included in some of the older references we cite below.

For ALPs coupled to electrons, ref. [44] computed a collection of SN constraints, including the above mentioned cooling and explosion energy bounds, which are the relevant ones for the regions of parameter space studied in this work. The loop-induced ALP-photon interactions in the SN plasma are claimed to be negligible, in contrast to the finding of ref. [37],[2] and

---

[2]We point our here that even following the prescription that ref. [44] argues for, namely to evaluate the effective loop-induced ALP-photon coupling as in vacuum but with the thermal electron mass instead

thus are not included in either ALP production or reabsorption rates. In this way, the bound could both be underestimated or overestimated at its upper end, depending on which production and reabsorption effects would dominate. However, as long as loop-effects are ignored, pseudoscalar and derivative couplings lead to the same phenomenology and we can show the constraints for both cases in figure 10. We show the most conservative constraint shown in ref. [44]. For ALP masses $m_a < 1\,\text{MeV}$, ref. [44] does not show any constraints. Since scattering processes in the plasma should be independent of the ALP mass for such light ALPs, we extrapolate the cooling constraint to lower masses but choose to cut off the energy deposition bound since in principle it depends on the decay length of the ALP, which of course changes significantly with m_a especially at the threshold $m_a = 2m_e$.

For interactions with muons, the cooling constraint on the derivative coupling was presented in ref. [94]. Note, however, that these results have been criticized in refs. [91] and [95], mainly pointing out that since absorption is only considered in the radial direction, i.e., is underestimated, and hence the upper end of the bound overestimated, and furthermore that the muons are assumed to be non-relativistic even though typical thermal energies are $3T \simeq m_\mu$. The latter effect also tends to overestimate the constraint in the strong coupling regime. Hence, we show the bound as a dashed line in the left panel of figure 11 to indicate where SN phenomenology should become relevant, but caution that there are potentially large numerical uncertainties on the exact value of the bound. The constraint is different for pseudoscalar couplings since the differing loop-induced photon coupling significantly affects the upper end of the bound for $5\,\text{MeV} \lesssim m_a \lesssim 2m_\mu \simeq 211\,\text{MeV}$, and has been calculated in ref. [91]. More importantly, for $m_a \lesssim 10\,\text{MeV}$, ref. [91] furthermore estimates that the explosion energy constraint in this case is $g_{a\mu}/m_\mu \gtrsim 2.26 \cdot 10^{-2}\,\text{GeV}^{-1}$. The bound is approximated by demanding that the total energy emitted in ALPs, in the "Stefan-Boltzmann approximation" discussed and shown to be a good estimate in simple models of SN-ALP emission [95], does not exceed the SN explosion energy. We show this constraint, which is stronger than the cooling bound in the relevant parameter region, in the right panel of figure 11. Also at larger ALP masses there should be a constraint on the muon coupling. However, since $m_a \gtrsim T$ for heavier ALPs — especially near the upper end of the bound that constrains ALPs produced in the colder outer regions of the proto-neutron star — the estimates of production, absorption, and decay rates in ref. [91] do not hold anymore, and the bound is not simply extrapolated to larger masses.

For ALPs coupled only to tau leptons, one can safely integrate them out, given that the tau mass is much larger than typical temperatures reached in the proto-neutron star. This leaves only a constant on-shell coupling to photons for the pseudoscalar case, which could be constrained by the usual analyses, while for the derivative case, the coupling is suppressed by m_a^2/m_τ^2 or E_a^2/m_τ^2 with the ALP energy E_a (for ALP production or reabsorption via

---

of m_e, would leave a nearly maximal $g_{a\gamma^*\gamma}^{\text{eff},\,\text{P},\text{D}}\left(|t| \sim T^2 \gtrsim m_{e,\text{thermal}}^2\right) \lesssim \alpha C_e/\pi\Lambda$ for both the pseudoscalar as well as the derivative-coupling case. This is because the thermal electron mass is set by temperature and chemical potential, both of which also determine the typical energy in the Primakoff process — see section 2.2 and ref. [37]. And since photon coalescence becomes important for $m_a \gtrsim \mathcal{O}(10\,\text{MeV}) > m_{e,\text{thermal}}$, at least in the derivative-coupling case, also that process (or its inverse) is not significantly suppressed as $g_{a\gamma\gamma}^{\text{eff},\,\text{D}}\left(m_a > m_{e,\text{thermal}}\right) \simeq \alpha C_e/\pi\Lambda$. Nevertheless, ref. [44] improves over the work in ref. [37] in many ways and hence we include their results here.

the Primakoff process). In ref. [83], SN constraints on $g_{a\gamma}$ are translated into constraints on the derivative coupling C_τ/Λ. However, the fact that the effective photon coupling is energy dependent if one of the photons is off-shell, as in the Primakoff process, is not taken into account (not many details are given, but the loop-function seems only to be evaluated using the ALP mass). We also note that due to the energy-dependence of the effective ALP-photon coupling, the constraints on $g_{a\gamma}$ cannot simply be rescaled but the energy integrals have to be reevaluated using the respective loop-functions. Thus, we choose not the show these constraints, but remark that they should be present in some form in this region of parameter space.

In the universal coupling case, production from electrons, muons, and the loop-induced photon interaction (also receiving a contribution from $g_{a\tau}$) have to be taken into account, and the same holds true for reabsorption and decay events. Thus, neither of the bounds on electron or muon couplings are accurate estimates for this case, and we thus choose not to show any SN constraints for this scenario either, but again remark that they should be present somewhere in the parameter space.

4.6 Proton beam-dump experiments and forward-physics detectors

Proton beam dump experiments provide sensitive probes of the effective ALP-photon coupling [96, 97]. A detailed reinterpretation of these constraints in the context of leptophilic ALPs is not straightforward since the typical momentum transfer t from the target nucleus to the ALP can be comparable to m_ℓ^2, such that the full momentum dependence of the lepton loop needs to be included for an accurate estimate.

To obtain a first estimate of the expected constraints, we implemented the constant piece of the loop-induced photon coupling $g_{a\gamma\gamma}^{\text{eff}} = \alpha C_\ell/(\pi\Lambda)$ in the public code ALPINIST [98], which allows for a reinterpretation of proton-beam dump data for ALPs with different coupling structures. The resulting constraints are found to be weaker than the ones from E137 and NA64. Given the results from ref. [99] for muonphilic scalars, we anticipate the same conclusion to hold also for muonphilic ALPs, which can be produced via Bremsstrahlung from secondary muons. Since a more accurate calculation is not expected to give significantly stronger bounds, we refrain from a detailed study and omit these constraints from our analysis. For the same reason, we do not attempt to reinterpret the recent search for ALPs coupled to photons at FASER [100].

4.7 Lepton $g-2$

Leptophilic ALPs contribute to the anomalous magnetic moment of the charged leptons, $a_\ell \equiv (g-2)_\ell/2$ with $\ell = e, \mu, \tau$. The main contributions for the muonic case are shown in figure 8.

At the one-loop level, there is no difference between derivative and pseudoscalar couplings due to the direct coupling of the ALP to the external charged lepton lines. Conversely, the two-loop contribution, which features an effective ALP-photon coupling, distinguishes the two coupling types. The expression for the one-loop contribution is given in ref. [5], while the expressions for the two-loop contributions can be found in refs. [21, 34] and [101] for
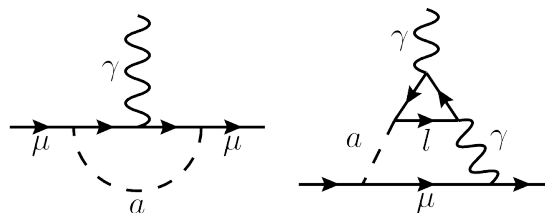
Figure 8. One-loop contributions (left) and two-loop contributions (right) of muonphilic ALPs to $(g-2)_\mu$.

pseudoscalar and derivative couplings, respectively. Note that, in the latter case, the result depends on the UV cutoff scale Λ, which we will set here to $\Lambda \sim 1\,\mathrm{TeV}$.

The theoretical situation regarding the muon anomalous magnetic moment is currently under scrutiny, see e.g. ref. [102] and references therein. Indeed, up to a few years ago, the theoretical consensus was a value [103] that currently shows a negative 5.2σ discrepancy with the current experimental world average [104], a discrepancy that would only be exacerbated by ALPs contributions. However, the latest lattice determinations of the so-called *leading order hadronic vacuum polarisation* (LOHVP) contribution increase the theoretical prediction of a_μ, effectively resulting in a compatibility at the 0.9σ level [105, 106]. Yet these determinations are incompatible with the previous LOHVP determinations at the 4σ level. While waiting for the resolution of this theoretical controversy, we decided to adopt a conservative approach here and include in our results only a dashed line corresponding to the ALP parameter values such that its presence would decrease a_μ^{theory} by $\delta a_\mu = -10^{-9}$, i.e. corresponding to a negative 2σ shift from the current experimental value, when current theoretical and experimental errors are combined. Nevertheless, this result must not be considered as a binding constraint, as nothing prevents the presence of additional new processes contributing to a_μ and counteracting the ALP effect.

A similar conservative approach is required as well for the electron anomalous magnetic moment, albeit for different reasons. In this case, the discrepant determinations of the LOHVP do not play a relevant role because this contribution is negligible due to its electron mass suppression. However, at the current level of accuracy of 10^{-13}, the theoretical prediction is sensitive to different measurements of the fine-structure constant employed in its determination, where a 5.4σ difference has been observed in its two latest experimental determinations [107, 108]. The latest experimental determination of a_e, measured with a precision similar to the theoretical uncertainty, sits between the two theoretical determinations [109]. Therefore, we adopt a similar approach to the muon case, showing in our results a dashed line corresponding to ALP parameter values such that they would decrease a_e^{theory} by 10^{-13}. Once again, this constraint does not have to be considered binding due to possible additional contributions to this channel.

Finally, we also consider the effect a tauphilic ALP would have on the anomalous magnetic moment of the tau lepton, a_τ. The recent determination of a_τ from photon fusion at CMS [110] provides $a_\tau = \left(0.9^{+3.2}_{-3.1}\right) \times 10^{-3}$. Given this experimental uncertainty, we choose once again to mark the parameter space where the presence of a tauphilic ALP decreases a_τ^{theory} by 10^{-3} just by a dashed line.

4.8 ALP lifetime

Axion-like particles decaying into leptons and photons in the early universe may change its expansion rate, the temperature of the electron-photon bath relative to the one of neutrinos and the primordial element abundances [111]. While a detailed calculation of these constraints is challenging, in particular for the muonic decay mode, it is expected that ALP decays are harmless for cosmology if the ALP lifetime is shorter than $0.1\,\mathrm{s}$. To give the reader a feeling for how relevant this consideration is, we indicate a dashed line labelled "$\tau_a = 0.1\,\mathrm{s}$" in our plots. This line should, however, not be interpreted as a hard constraint since additional effects can change the cosmological evolution, and the resulting constraints may be weaker.

5 Results

In this section, we present and discuss the bounds on the leptophilic ALP models obtained from the beam-dump experiments E137 and NA64 discussed in section 3 as well as the other constraints discussed in section 4. In order to remain agnostic about the UV completion of this scenario, we analyse different coupling structures independently:

- Scenario E: non-vanishing electron coupling $\hat{g}_{ae} \neq 0$,

- Scenario M: non-vanishing muon coupling $\hat{g}_{a\mu} \neq 0$,

- Scenario T: non-vanishing tau coupling $\hat{g}_{a\tau} \neq 0$,

- Scenario U: universal lepton coupling $\hat{g}_{ae} = \hat{g}_{a\mu} = \hat{g}_{a\tau}$,

where we have introduced the definitions $\hat{g}_{a\ell} \equiv g_{a\ell}/m_\ell$ or $\hat{g}_{a\ell} \equiv C_\ell/\Lambda$ for the pseudoscalar or the derivative case, respectively. For each scenario, we consider both derivative and pseudoscalar couplings and set all other ALP couplings to zero. All bounds are calculated at 90% confidence level, except the ones from LEP and $W^+ \to e^+\nu a$, which are at 95% confidence level.

5.1 Effect of the coupling type on the E137 bound

Before presenting our full results, we compare the effect of the choice of coupling type on the bounds derived from the E137 experiment. The comparisons of the bounds for all scenarios are shown in figure 9.

In this figure, aside from comparing the derivative coupling given in eq. (2.1) and the pseudoscalar coupling in eq. (2.4), we also show a comparison with a simplifying approximation, in which the effective off-shell coupling of the ALP to photons is set to the constant on-shell value, in order to judge the relevance of the momentum dependence of the effective photon coupling. We approximate the couplings according to the limits of the derivative and pseudoscalar couplings given in eqs. (2.17) and (2.18), respectively. Since the ALP mass is larger than that of the electron in the most relevant regions of parameter space, we assume $\tau_e = 4m_e^2/m_a^2 \ll 1$. For the muon and tau contribution, on the other hand, we can approximate $\tau_{\mu,\tau} \gg 1$ in the parameter regions probed by E137. In the on-shell

approximation, we thus approximate the derivative coupling by the constant values

$$g_{a\gamma^*\gamma}^{\text{eff,D}} \to g_{a\gamma\gamma}^{\text{eff,D}} \approx g_{a\gamma\gamma,\text{const}}^{\text{eff,D}} = \begin{cases} \frac{\alpha C_e}{\pi\Lambda}, & \text{for} \quad \text{Scenario E}, \\ -\frac{\alpha C_\mu}{\pi\Lambda}\frac{m_a^2}{12m_\mu^2}, & \text{for} \quad \text{Scenario M}, \\ -\frac{\alpha C_\tau}{\pi\Lambda}\frac{m_a^2}{12m_\tau^2}, & \text{for} \quad \text{Scenario T}, \\ \frac{\alpha C_\ell}{\pi\Lambda}\left(1 - \frac{m_a^2}{12m_\mu^2} - \frac{m_a^2}{12m_\tau^2}\right), & \text{for} \quad \text{Scenario U}. \end{cases} \tag{5.1}$$

Similarly, the approximate pseudoscalar couplings read

$$g_{a\gamma^*\gamma}^{\text{eff,P}} \to g_{a\gamma\gamma}^{\text{eff,P}} \approx g_{a\gamma\gamma,\text{const}}^{\text{eff,P}} = \begin{cases} 0, & \text{for} \quad \text{Scenario E}, \\ \frac{\alpha g_{a\mu}}{\pi m_\mu}\left(-1 - \frac{m_a^2}{12m_\mu^2}\right), & \text{for} \quad \text{Scenario M}, \\ \frac{\alpha g_{a\tau}}{\pi m_\tau}\left(-1 - \frac{m_a^2}{12m_\tau^2}\right), & \text{for} \quad \text{Scenario T}, \\ \frac{\alpha g_{a\ell}}{\pi m_\ell}\left(-1 - \frac{m_a^2}{12m_\mu^2} - 1 - \frac{m_a^2}{12m_\tau^2}\right), & \text{for} \quad \text{Scenario U}. \end{cases} \tag{5.2}$$

Specifically, in the approximation for the case of pseudoscalar couplings to electrons only, we set the loop-induced coupling to photons to zero and consider only the tree-level contribution from electrons. For Scenario U, we also consider a simplified tree-level model of ALPs, where we set all effective loop-induced photon couplings to zero in order to estimate their relevance for the final exclusion bounds.

Beginning with Scenario E in the top-left panel of figure 9, one can see the bounds on the derivative coupling, the pseudoscalar one and their respective constant approximations. For clarity, only the contributions coming from DB and Primakoff production are shown, but not the positron annihilation processes. For ALP masses $m_a \lesssim 100\,\text{MeV}$, the agreement of the bounds between all couplings is close to perfect. However, for larger masses, the differences between the couplings become evident. While the lower bounds on the derivative coupling remain relatively constant until reaching the tip, the pseudoscalar bounds are weaker for the higher ALP mass range.

The derivative and pseudoscalar couplings perfectly match their respective constant approximations. This is unsurprising, especially at the tip of the plot, as the ALP mass is much larger than the electron mass, making eq. (5.1) an excellent approximation for the effective photon coupling in the derivative basis. Similarly, the fact that the bound for the pseudoscalar coupling can be reproduced by setting the loop-induced photon coupling to zero agrees with the expectation from eq. (2.18), as the effective photon coupling in the pseudoscalar basis is strongly suppressed by the electron mass.

We note that this excellent agreement between the approximated couplings and the exact couplings is still present at the very left side of the plot, where $m_a \approx 10\,\text{MeV}$ and the approximation $\tau \ll 1$ starts to become invalid. This agreement can be explained by the fact that the Primakoff production mechanism, which is the contribution sensitive to the effective coupling, is dwarfed by the tree-level DB production channel, as can be seen in figure 5. No matter which coupling type is used, their differing effects are irrelevant compared to DB production, hence the agreement is good for low masses as well. At the tip of the exclusion, on the other hand, the Primakoff production actually outweighs the DB production (see figure 5), such that the two coupling scenarios give slightly different results.
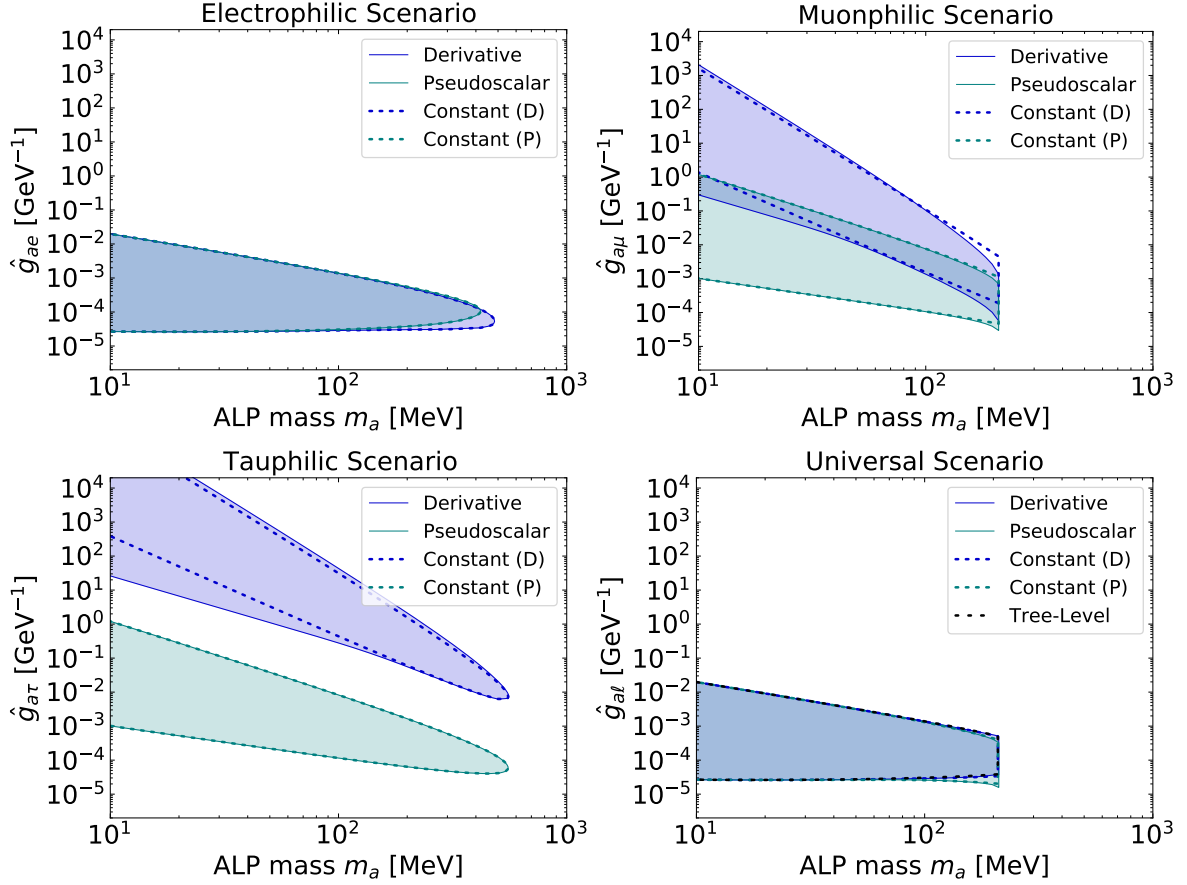
Figure 9. Bounds derived from E137 using different types of couplings to electrons (top-left), muons (top-right), tau leptons (bottom-left) and universal couplings to all leptons (bottom-right).

This difference becomes much more pronounced in the top-right panel of figure 9, which shows the comparison of the bounds derived using Scenario M. Since in this scenario only the muon coupling is present at tree level, the only production mechanism in E137 is Primakoff production. Below the threshold $m_a = 2m_\mu \approx 211\,\mathrm{MeV}$, the ALP can only decay into photons, whereas above the muon threshold, the decay into muons proceeds so fast that the ALPs can never leave the dump and reach the detector. As a result, the bounds on this model are very sensitive to the type of coupling used. Indeed, since the effective coupling in the derivative basis is smaller than in the pseudoscalar basis for large lepton masses, the bounds for the derivative coupling are shifted to higher $\hat{g}_{a\mu}$ to compensate.

In addition to the derivative and pseudoscalar couplings, we also show the bounds derived using constant couplings. For the constant approximation of the pseudoscalar coupling, the bounds are generally in very good agreement with the bounds derived from the full pseudoscalar coupling, except close to the threshold $m_a = 2m_\mu$ where the approximation $\tau \gg 1$ becomes invalid. For the constant approximation of the derivative coupling, the bounds of the full and approximate couplings start to deviate significantly for low masses. In other words, it is essential to consider the full momentum dependence of the effective photon coupling to accurately describe the bounds on the ALP-muon coupling for the derivative coupling case.

One can observe the same effect in Scenario T, which is shown in the bottom-left panel of figure 9. Since the ALP masses are far below the tau threshold, the ALPs can only decay into photons via the effective coupling. As in Scenario M, the ALPs can only be produced via Primakoff production. As a result, the qualitative discussion for Scenarios M and T are almost identical. The only major difference is that, since the relevant ALP mass range remains well below $2m_\tau$, the constant coupling approximation works very well for $m_a \gtrsim 100\,\text{MeV}$. For lower masses, however, the momentum dependence of the couplings in the derivative case again becomes apparent.

Finally, the bottom-right panel of figure 9 considers Scenario U, i.e. universal couplings to all leptons. The general shape of the bounds is similar to Scenario E for $m_a < 2m_\mu$, but features the cutoff at $m_a = 2m_\mu$ seen in Scenario M. Since there is now once again a tree-level coupling to the beam electrons, the DB production of ALPs is the dominant process for lower ALP masses, such that pseudoscalar and derivative couplings agree in this mass range. The difference between the two coupling types only becomes apparent at masses around $m_a \gtrsim 100\,\text{MeV}$, when the loop-induced decays into photons start to compete with the tree-level decays into electrons [8].

Similarly to Scenario E, both the derivative coupling and the pseudoscalar one are described quite accurately by their relative constant coupling approximation in most of the mass range. However, the additional muon and tau couplings in Scenario U lead to a non-trivial momentum dependence in the effective Primakoff coupling near the threshold $m_a = 2m_\mu$, particularly in the pseudoscalar case. By coincidence, the derivative case is well approximated by considering only tree-level couplings, i.e. setting the effective ALP-photon coupling to zero and neglecting both Primakoff production and decays into photons.

5.2 Combined constraints

Having discussed the pertinent differences between pseudoscalar and derivative couplings for the bounds from E137, as well as the relevance of the momentum dependence in the effective couplings, we now turn to the various other constraints on leptophilic ALPs. Our main results are shown in figures 10 to 13, corresponding to the four scenarios introduced above. In each case, we show in the left (right) panel the constraints for derivative (pseudoscalar) couplings. The bounds that have been (re)computed by us, namely the ones from sections 3 and 4.2, are shown in the foreground, each highlighted with a differently coloured area. Conversely, the laboratory constraints from sections 4.1, 4.3,and 4.4, which have been taken directly from the literature, are shown in the background as a unified, gray-shaded area. Supernova constraints, where applicable and known, are indicated as dashed lines since they are somewhat approximate (and even contain some errors) in this strong coupling regime, cf. section 4.5 — they nevertheless show that SN phenomenology should become relevant in the indicated region of parameter space. Non-binding constraints, namely the ones coming from anomalous magnetic moments described in section 4.7 and the ALP lifetime discussed in section 4.8, are indicated for reference as dashed lines. Finally, for the NA64μ experiment, we also indicate projected bounds for a total of 10^{14} MOT, compared to the current $1.98 \cdot 10^{10}$ MOT.

Let us start our discussion from Scenario E, for which bounds are given in figure 10. Contrary to the previous section, where only contributions from DB and Primakoff production
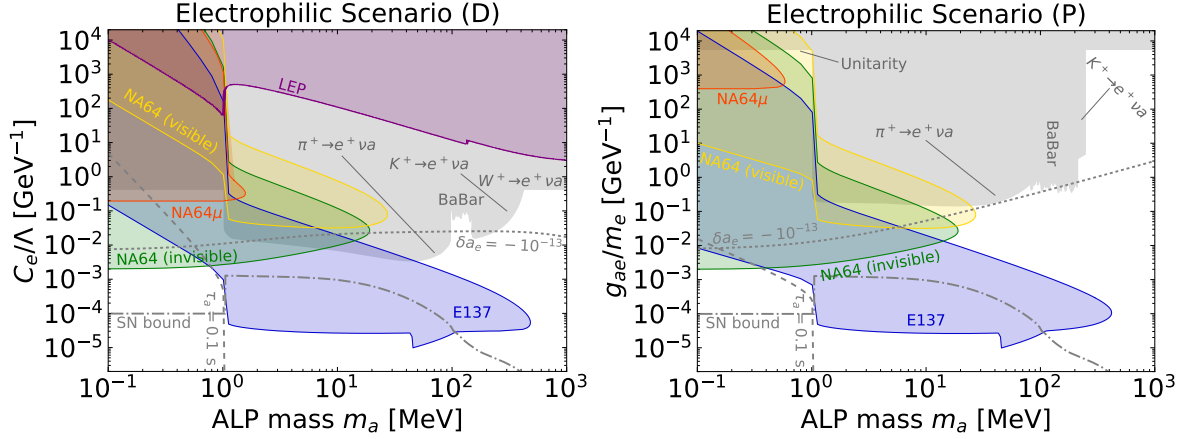
Figure 10. Laboratory bounds on an ALP with derivative (left) and pseudoscalar (right) coupling to electrons.

were considered for estimating E137 constraints, an additional bump is now present because we include the resonant ALP production yields as well.[3] We find that E137, which ran over 40 years ago, is still able to set some of the most stringent constraints on the electrophilic ALP models, especially for high ALP masses. However, the bounds set by NA64, both in its visible and invisible modes, complement those by E137, particularly for higher values of the coupling. This is particularly relevant in the pseudoscalar case, where these additional constraints cover all the otherwise unconstrained regions for ALP masses below 10 MeV and couplings above $10^{-2}\,\mathrm{GeV}^{-1}$. While most of the parameter space under consideration is now covered for the derivative case, this is not the case in the pseudoscalar one. In particular, the absence of constraints coming from LEP and W^+ decays leave the large coupling region for an ALP mass above $\sim 200\,\mathrm{MeV}$ unconstrained. To determine the viable parameter regions of smaller couplings, it would be desirable to perform a detailed calculation of the constraints from SNe for pseudoscalar couplings, including the effect of loop-induced couplings to photons.

Moving onto Scenario M, for which bounds are shown in figure 11, we observe, as expected, that the strongest constraints come from NA64μ, which is exceptionally well suited to probing muonphilic ALPs. Contrary to the previous scenario, E137 is only relevant for couplings between about $10^{-3}\,\mathrm{GeV}^{-1}$ and $10^{-4}\,\mathrm{GeV}^{-1}$, and masses above 100 (10) MeV, for the derivative (pseudoscalar) case. However, if NA64μ extends its programme to 10^{14} MOT, it is expected to push its constraints down to $10^{-5}\,\mathrm{GeV}^{-1}$, hence totally encompassing E137 ones. Complementary constraints also come from Belle II and from LEP (only in the derivative case) and SNe. We therefore conclude that, as for Scenario E, the parameter space is more constrained in the derivative coupling case than in the pseudoscalar coupling one.

Concerning Scenario T, from figure 12, it becomes clear that currently not many constraints are stronger than the unitarity bounds. We also observe that, in this scenario, the main additional bound comes from E137, which is stronger in the pseudoscalar coupling case than in the derivative one. As a result, in this scenario, the pseudoscalar coupling

---

[3]While resonant ALP production is not possible in NA64μ, it could in principle play a role in NA64 run in invisible and visible modes. However, due to the high cutoff energy $E_{\mathrm{cut}} = 20\,\mathrm{GeV}$ of NA64, the condition $m_a > \sqrt{2 m_e E_{\mathrm{cut}}}$ derived in section 3.3 cannot be satisfied in the parameter regions probed by NA64.
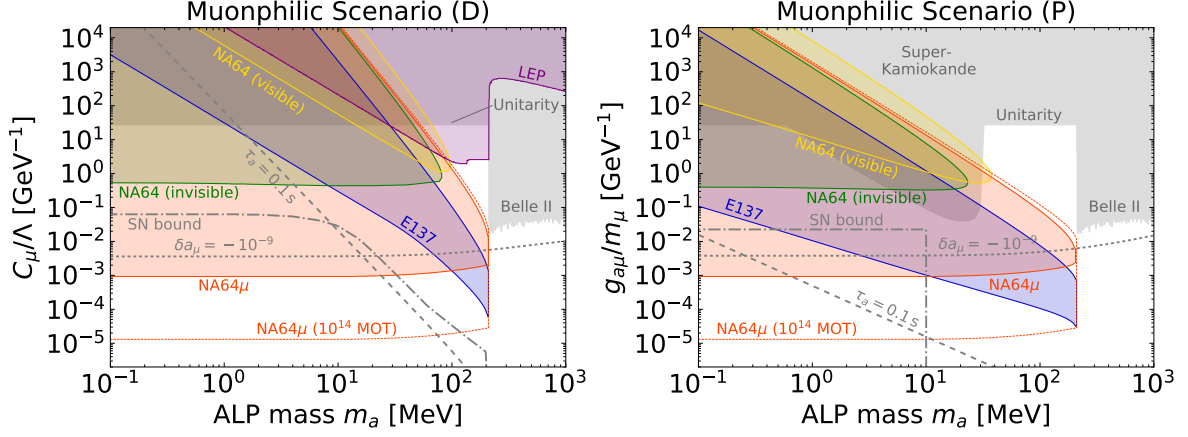
Figure 11. Laboratory bounds on an ALP with derivative (left) and pseudoscalar (right) coupling to muons.
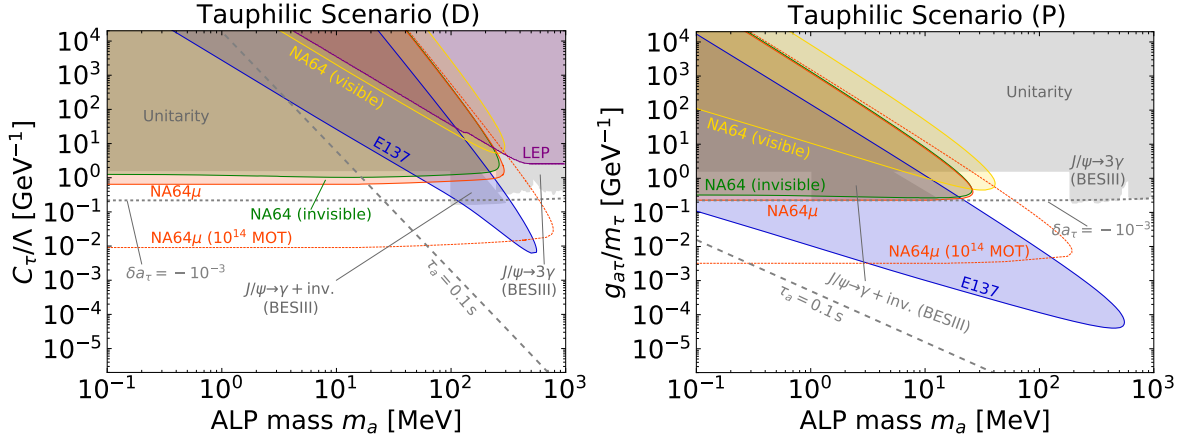

Figure 12. Laboratory bounds on an ALP with derivative (left) and pseudoscalar (right) coupling to tau leptons.

case is more tightly constrained than the derivative one. Current bounds from NA64 have marginal relevance in the former case, namely only in the ALP mass range 10–30 MeV, while in the latter one, they improve the unitarity bound by a factor of 2. The final NA64μ results will be able to provide much stronger constraints, improving the current bound by about two orders of magnitudes in the coupling and encompassing almost completely the ones from E137 in the derivative case.

We conclude with Scenario U, for which constraints are reported in figure 13. As expected, if the ALP couples to all charged leptons, the different constraints from the three previous scenarios become relevant simultaneously. We therefore find that the combination of E137, NA64μ, BaBar and Belle II measurements exclude ALP couplings above about $10^{-2}\,\text{GeV}^{-1}$. The strongest constraint comes again from E137, which pushes the bound down to almost $10^{-5}\,\text{GeV}^{-1}$ in the 1–200 MeV ALP mass range, while for smaller ALP masses NA64 excludes couplings larger than $10^{-3}\,\text{GeV}^{-1}$. With 10^{14} MOT, NA64μ is expected to push the coupling bounds down to about $10^{-5}\,\text{GeV}^{-1}$ for ALP masses up to the muon threshold, going beyond
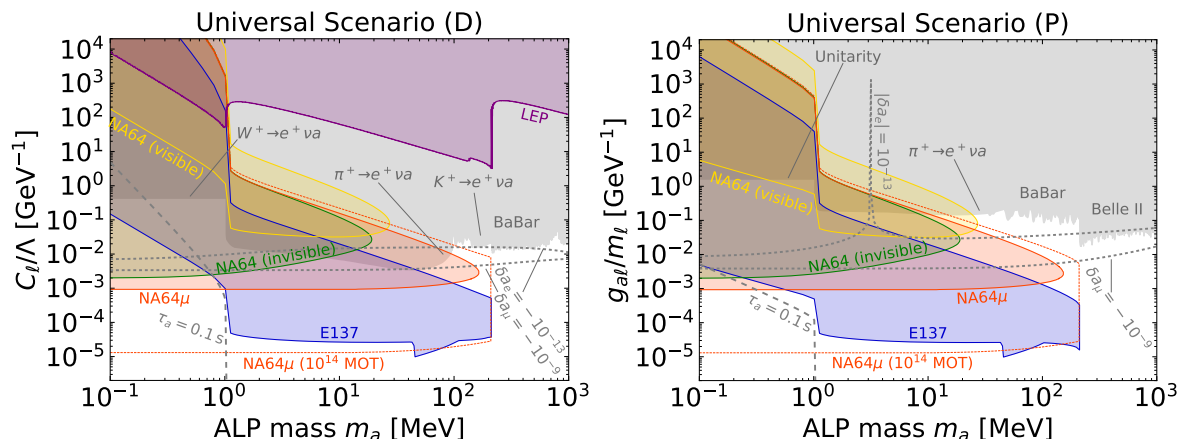
Figure 13. Laboratory bounds on an ALP with universal derivative (left) and pseudoscalar (right) coupling to all charged leptons.

the E137 reach and covering the vast majority of the currently unconstrained parameter space studied here. For the case of universal couplings to charged leptons, there is no relevant difference between the derivative and the pseudoscalar coupling cases, except a small unconstrained window for ALP masses in the 30–200 MeV mass range for the latter case.

6 Conclusion

Axion-like particles (ALPs) arising as pseudo-Goldstone bosons from a spontaneously broken approximate global symmetry are one of the most well-motivated ways to extend the Standard Model with a new state below the GeV scale coupled feebly to known particles. While couplings of ALPs to SM gauge bosons and SM quarks have been studied in great detail, the case of leptophilic ALPs, which couple dominantly to charged leptons, has received much less attention. In the present work, we have filled this gap by providing an exhaustive overview of the relevant constraints from beam-dump, flavour and collider experiments, as well as from theoretical considerations and astrophysics.

First of all, we have pointed out that the precise definition of leptophilic ALPs is rather subtle since the assumed absence of tree-level couplings to photons leads to a distinction between ALPs with pseudoscalar couplings and ALPs with derivative couplings. We take a complete model-independent perspective and discuss both interactions in parallel, highlighting their differences.

In both cases, effective interactions of ALPs to photons are generated through lepton loops, which depend on the momenta of the ALP and the two photons. If all particles are on-shell, one reproduces the standard photon coupling, but for off-shell particles, the loop function depends non-trivially on the momenta, which are specific to the relevant experimental setup. We have shown that this effect plays an important role in the analysis of electron beam-dump experiments such as E137 and collider constraints from, e.g. LEP and have derived suitable approximations that can be used to simplify the analysis.

In our analysis, we have then considered separately the three cases where ALPs couple only to one generation of charged leptons, as well as the case of universal couplings to all

charged leptons. In all cases, we find that E137 gives highly relevant constraints, even if both the production and decay of the ALPs proceed exclusively via loop-induced photon couplings. Further relevant constraints come in particular from searches for missing energy and displaced decays at NA64, in particular when using a muon beam. Indeed, future runs of NA64μ constitute a highly promising way to further probe the parameter space of leptophilic ALPs.

For cases in which the ALPs couple to muons, beam dump experiments are not sensitive to ALP masses above the threshold $m_a > 2m_\mu$, where ALPs rapidly decay back into muons. In this parameter region, future searches for di-muon resonances at Belle II [79] are particularly promising. Below the threshold, relevant sensitivity may come from the planned proton beam-dump experiment SHiP [112] as well as from the proposed forward-physics experiment FASER2 [113], both of which can exploit the presence of secondary muons in hadron showers to search for leptophilic ALPs. A detailed study of this production mode, including Monte Carlo simulations of the secondary muon distributions, is left for future work. Finally, an improved understanding of the bounds from SN 1987A for all the different coupling scenarios is needed in order to identify the most interesting regions of parameter space at very small couplings. Combining all of these different strategies will enable us to further explore the parameter space of leptophilic ALPs.

Acknowledgments

A Experimental details

A.1 E137

An electron beam with an energy of 20 GeV was directed at an aluminium target. Over the course of the experiment, a total of 30 C (Coulombs) of electrons were deposited in the dump, corresponding to $N_e = 1.87 \cdot 10^{20}$ electrons on target (EOT). The target was placed in front of a hill 179 m in length, which absorbed all SM particles except neutrinos. The detector was placed behind the hill, providing a decay length of 204 m. The detector consisted of an 8-radiation-length shower calorimeter, which could detect photons and electrons. For

the first phase of the experiment, the detector was arranged in a 2×3 mosaic of $1\,\mathrm{m} \times 1\,\mathrm{m}$ chambers, which was later changed to $3\,\mathrm{m} \times 3\,\mathrm{m}$ chambers. The cutoff energy for the detector, i.e. the minimum ALP energy to be considered a candidate, was set to $2\,\mathrm{GeV}$. No candidate events were observed.

A.2 NA64

Invisible mode. The detector consists of one electromagnetic calorimeter (ECAL) and three HCAL. The ECAL consists of lead and scintillator plates, having a total length of $45\,\mathrm{cm}$ and a width of $23\,\mathrm{cm} \times 23\,\mathrm{cm}$. The ECAL is the active target and detects the energy deposition of electrons. Each HCAL is made of iron and scintillator plates intended to detect charged and neutral hadrons. The three HCALs have a combined length of approximately $5\,\mathrm{m}$, and a width of $60\,\mathrm{cm} \times 60\,\mathrm{cm}$. Using this setup, the calorimeters are able to determine the energy deposited inside the beam dump. If an ALP or other weakly interacting particle were to be produced inside the dump and was long-lived enough, it would escape the detector and carry away a fraction of the original energy of the beam. This missing energy can be measured and registered as an event [46].

The main beam has an energy of $100\,\mathrm{GeV}$. The energy cutoff for the detector is set at $50\,\mathrm{GeV}$. In the runs from 2016–2022, NA64 had a total of $N_e = 9.37 \cdot 10^{11}$ EOT. In the 2016 run, a fourth HCAL was placed in line with the other three HCALs, which was moved downstream to be a zero-degree HCAL to help with background detection in later runs. A detailed description of the layout of the experiment can be found in ref. [114]. No candidate events were observed [115].

Visible mode. The setup used for the visible mode is described in ref. [47]. The beam is directed toward an electromagnetic calorimeter composed of tungsten and scintillator plates (WCAL). The WCAL serves as an active beam dump, with the first layers being the main regions in which BSM particles would be produced. The rest of the WCAL serves as a dump for the SM particles created in other processes, while weakly interacting particles would pass through the dump. After a decay length of approximately $3.5\,\mathrm{m}$, the WCAL is followed by the ECAL and HCALs discussed above. If an ALP or any other particle was produced and subsequently decayed inside the decay volume, the ECAL would register the energy of the decay products. If the energy that is deposited in the ECAL is the same as the missing energy in the WCAL, i.e. $E_{\mathrm{ECAL}} = E_{\mathrm{beam}} - E_{\mathrm{WCAL}}$, then this is a candidate event [116].

In 2017, the beam was run at an energy of $100\,\mathrm{GeV}$. The length of the WCAL was varied from $290\,\mathrm{mm}$, with $2.4 \cdot 10^{10}$ EOT, to $220\,\mathrm{mm}$, with $3 \cdot 10^{10}$ EOT. The cutoff energy for the detector was set to $20\,\mathrm{GeV}$. In 2018, the beam energy was increased to $150\,\mathrm{GeV}$, with $3 \cdot 10^{10}$ EOT, with the cutoff energy being set to $30\,\mathrm{GeV}$. No candidate events were observed [47, 116].

Muon mode. The experimental setup for NA64μ is illustrated in ref. [49]. As with the original NA64 run with an electron beam, the experiment is built to detect the invisible decay of BSM particles by measuring the missing energy of an event. A muon beam with an energy of $160\,\mathrm{GeV}$ is directed at an ECAL, similar to the one in previous runs. The ECAL has a total length of $49\,\mathrm{cm}$. Further downstream, a magnetic spectrometer is placed, which uses a bending magnet (MS2) to deflect the muon beam, as well as two HCALs to remove

Experiment	Target	Z	A	X $[\mathrm{g\,cm^{-2}}]$	ρ $[\mathrm{g\,cm^{-3}}]$	M_{target} $[\mathrm{g\,mol^{-1}}]$
E137	Al	13	27	24.01	2.7	26.98
NA64 (invisible,μ)	Pb	82	207	6.37	11.35	207.2
NA64 (visible)	W	74	184	6.76	19.3	183.84

Table 1. Material parameters for E137 and NA64 [118].

Experiment	EOT/MOT	E_0 [GeV]	E_{cut} [GeV]	L_{sh} [m]	L_{dec} [m]	T	$\theta_{a,\mathrm{max}}$
E137	$1.87 \cdot 10^{20}$	20	2	179	204	2012.9	$4.4 \cdot 10^{-3}$
NA64 (invisible)	$9.37 \cdot 10^{11}$	100	50	5.5	∞	979.8	0.11
NA64 (visible):							
Run 1 (2017)	$2.4 \cdot 10^{10}$	100	20	0.29	3.5	82.8	$6.9 \cdot 10^{-2}$
Run 2 (2017)	$3 \cdot 10^{10}$	100	20	0.22	3.5	62.8	$6.9 \cdot 10^{-2}$
Run 3 (2018)	$3 \cdot 10^{10}$	150	30	0.22	3.5	62.8	$6.9 \cdot 10^{-2}$
NA64μ	$1.98 \cdot 10^{10}$	160	80	11	∞	87.3	$5.5 \cdot 10^{-2}$

Table 2. Experimental parameters for E137 and NA64.

residuals from the upstream detectors [49]. The muon momentum can then be reconstructed. Candidate events are defined as events in which the reconstructed muon momentum is smaller than 80 GeV and the energy deposition in the calorimeters is compatible with a minimum ionising particle, such that the missing energy is larger than 80 GeV. This setup results in a shielding of approximately 11 m, which the ALP must traverse before decaying. After the 2021 test run and the 2022–2023 runs, the experiment accumulated a total of $1.98 \cdot 10^{10}$ muons on target (MOT). No candidate events were observed [117].

A.3 Summary of experimental parameters

For ease of reference, the values of all experimental parameters we use to calculate N_a are listed in tables 1 and 2. Since the experimental parameters were varied for the different runs of NA64 operated in visible mode, the sum over all runs is taken when determining the total number N_a

$$N_a = \sum_{\mathrm{runs}} N_{a,i}. \tag{A.1}$$

The invisible mode of the NA64 experiment has three HCALs downstream of the ECAL, giving a total length of approximately 5.5 m. In the first run in 2016, the experiment had a fourth HCAL, which was later moved to a zero-degree HCAL. This means the length of the shielding of the 2016 run was greater compared to the later runs. Nevertheless, we approximate $L_{\mathrm{sh}} = 5.5\,\mathrm{m}$ for all runs, as the error would be minimal on account of the relatively low contribution of the $4.5 \cdot 10^{10}$ EOT from this run to the total $9.37 \cdot 10^{11}$ EOT.

B Details on electron energy distribution functions

For the calculation of the electron and positron energy distributions, we use the expressions that were derived in ref. [72], namely, eq. (22) for the first generation electrons, eq. (24) for the first generation photons, and eq. (12) for the second generation electrons and positrons:

$$I_e^{(1)}(t, E) = \frac{1}{E_0} \frac{\ln(E_0/E)^{bt-1}}{\Gamma(bt)}, \tag{B.1}$$

$$\begin{aligned}
I_\gamma^{(1)}(t, k) = \frac{1}{k} \int_0^t dt' \frac{e^{-c(t-t')}}{\Gamma(bt'+1)} \left(\ln \frac{1}{u}\right)^{bt'} \\
\times \left[u + \sum_{n=0}^\infty \frac{b(-1)^n - u^2}{n!(n+bt'+1)} \left(\ln \frac{1}{u}\right)^{n+1} \right],
\end{aligned} \tag{B.2}$$

$$\begin{aligned}
I_{e,p}^{(2)}(t, E) = \int_0^t dt' \int_E^{E_0} dE' \frac{1}{E'} \frac{\ln(E'/E)^{b(t-t')-1}}{\Gamma(b(t-t'))} \\
\times \int_{E'}^{E_0} dk\, 2I_\gamma^{(1)}(t', k) \left[b\left(1 - \frac{k}{E'}\right) + \left(\frac{k}{E'}\right)^2 \right] \frac{E'^2}{k^3},
\end{aligned} \tag{B.3}$$

where E_0 is the initial beam energy, $u = k/E_0$, $b = 4/3$, and $c = 7/9$.

C Details on cross section calculations

C.1 Dark Bremsstrahlung and Primakoff production

It is convenient to define modified Mandelstam variables $\tilde{s}$, $\tilde{u}$, and t_2 in the same way as [71], noting the use of the opposite sign of the metric. Assuming the incoming lepton has an energy that is much greater than the masses of the ALP and lepton, one can make the approximations [71]

$$\tilde{s} = -\frac{\tilde{u}}{1-x}, \quad \tilde{u} = -xE^2\theta_a^2 - m_a^2\frac{1-x}{x} - m_\ell^2 x, \quad t_2 = \frac{x}{1-x}\tilde{u} + m_a^2. \tag{C.1}$$

In the IWW approximation, the cross section is given by [71]

$$\left(\frac{d\sigma}{dx}\right)_{\text{IWW}} = \frac{\alpha}{(4\pi)^2} \sqrt{x^2 - \frac{m_a^2}{E^2}} \chi \frac{1-x}{x} \int_{\tilde{u}_{\min}}^{\tilde{u}_{\max}} d\tilde{u} \frac{\overline{|\mathcal{M}|}^2}{\tilde{u}^2}, \tag{C.2}$$

with the boundaries

$$\tilde{u}_{\min} = -xE^2\theta_{a,\max}^2 - m_a^2\frac{1-x}{x} - m_\ell^2 x, \quad \tilde{u}_{\max} = -m_a^2\frac{1-x}{x} - m_\ell^2 x, \tag{C.3}$$

where the maximum scattering angle $\theta_{a,\max}$ depends on the geometry of the detector. The effective flux of photons χ is given by

$$\chi = \int_{t_{\min}}^{t_{\max}} dt \frac{t - t_{\min}}{t^2} F^2(t). \tag{C.4}$$

Here, F is the form factor. The inelastic form factor can be neglected here as its contribution is negligible compared to the elastic form factor [71], given by

$$F(t) = \left(\frac{a^2 t}{1 + a^2 t}\right)\left(\frac{1}{1 + t/d}\right) Z. \tag{C.5}$$

The first part accounts for the atomic form factor, while the second part accounts for the nuclear form factor [70, 119]. Here, $a = 111Z^{1/3}/m_e$ and $d = 0.164\,\mathrm{GeV}^2 A^{-2/3}$. The integers Z and A are the atomic number and mass number of the target atom, respectively.

Using FeynCalc [65], we calculate that the spin-averaged, squared amplitude for the simplified DB processes is

$$\overline{|\mathcal{M}_{\mathrm{DB}}|}^2 = \frac{e^2 g_{a\ell}^2 \Big[2m_a^2(\tilde{s} + \tilde{u})\left(m_\ell^2(\tilde{s} + \tilde{u}) + \tilde{s}\tilde{u}\right) - \tilde{s}\tilde{u}(\tilde{s} + \tilde{u})^2 - 2m_a^4\tilde{s}\tilde{u} \Big]}{\tilde{s}^2 \tilde{u}^2} . \tag{C.6}$$

Using eq. (C.2) and the substitutions in eq. (C.1), we then obtain eq. (3.9). Since the DB diagrams have tree-level couplings to the ALP, the choice between a pseudoscalar or derivative ALP-lepton coupling has no effect on the amplitude.

For the Primakoff process, the amplitude is

$$\overline{|\mathcal{M}_{\mathrm{P}}|}^2 = \frac{e^2 \left| g_{a\gamma^*\gamma}^{\mathrm{eff}}(t_2) \right|^2 \Big[-t_2\left(\tilde{s}^2 + \tilde{u}^2\right) - 2m_\ell^2(\tilde{s} + \tilde{u})^2 \Big]}{4t_2^2} , \tag{C.7}$$

where the effective off-shell coupling $g_{a\gamma^*\gamma}^{\mathrm{eff}}(t_2)$ is defined in eq. (2.25) and should be understood as the sum over the three effective couplings coming from the couplings to the leptons, i.e.

$$g_{a\gamma^*\gamma}^{\mathrm{eff}}(t_2) = \sum_{\ell=e,\mu,\tau} g_{a\gamma^*\gamma}^{\mathrm{eff}}(t_2, m_\ell) . \tag{C.8}$$

Additionally, the interference term between the two processes gives the contribution

$$\overline{|\mathcal{M}_{\mathrm{I}}|}^2 = \frac{e^2 \,\mathrm{Re}\left(g_{a\gamma^*\gamma}^{\mathrm{eff}}(t_2) \right) g_{a\ell} m_\ell (\tilde{s} + \tilde{u})^3}{\tilde{s} t_2 \tilde{u}} . \tag{C.9}$$

C.2 Non-resonant production

The Feynman diagrams for the non-resonant production of an ALP

$$e^-(p_1) + e^+(p_2) \to a(k_1) + \gamma(k_2) \tag{C.10}$$

are shown in figure 14. The positron can annihilate with the electron through either an s-channel, t-channel or u-channel diagram, of which the latter two are tree-level diagrams. The squared, spin-averaged matrix elements for the tree-level, loop-level, and interference terms are

$$
\begin{aligned}
\overline{|\mathcal{M}_{\mathrm{NR,tree}}|}^2 =& \frac{e^2 g_{ae}^2}{(t - m_e^2)^2 (u - m_e^2)^2} \\
& \times \Big[\left(t - m_e^2\right)\left(u - m_e^2\right)\left(2m_a^4 + (s - m_a^2)^2\right) + 2m_a^2\left(s - m_a^2\right)\left(tu - m_e^4\right) \Big] , \\
\overline{|\mathcal{M}_{\mathrm{NR,loop}}|}^2 =& \frac{e^2 \left| g_{a\gamma^*\gamma}^{\mathrm{eff}}(s) \right|^2}{4s^2} \\
& \times \Big[m_a^4\left(2m_e^2 + s\right) - 2m_a^2 s\left(m_e^2 + s + t\right) + 2s\left(m_e^4 - 2m_e^2 t + t(s+t)\right) + s^3 \Big] , \\
\overline{|\mathcal{M}_{\mathrm{NR,inter}}|}^2 =& \frac{e^2 \,\mathrm{Re}\left(g_{a\gamma^*\gamma}^{\mathrm{eff}}(s) \right) g_{ae} m_e \left(s - m_a^2\right)^3}{s\left(t - m_e^2\right)\left(u - m_e^2\right)} ,
\end{aligned}
\tag{C.11}
$$

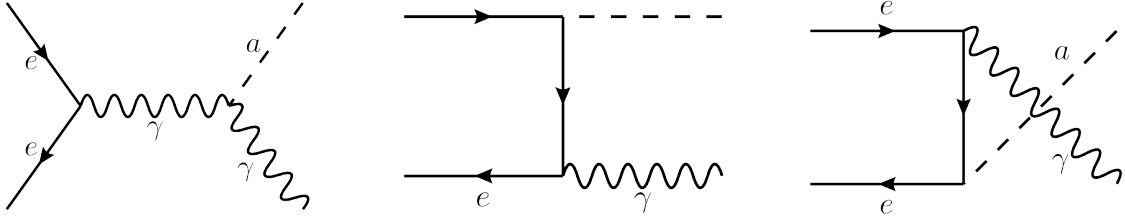**Figure 14.** Feynman diagrams contributing to the process $e^- e^+ \to a\gamma$.

respectively. The Mandelstam variables are defined in the standard way

$$s = (p_1 + p_2)^2, \quad t = (p_1 - k_1)^2, \quad u = (p_1 - k_2)^2 \,. \tag{C.12}$$

The cross section for non-resonant production of ALPs is

$$\frac{d\sigma_{\mathrm{NR}}}{dx} = \frac{1}{16\pi s} \frac{1}{1 - \frac{m_e}{E}} \overline{|\mathcal{M}_{\mathrm{NR}}|}^2 \theta(\theta_{a,\mathrm{max}} - \theta_a) \,, \tag{C.13}$$

with the angle θ_a defined via

$$\cos(\theta_a) \equiv \frac{2(E + m_e)E_a - s - m_a^2}{2|\mathbf{p}||\mathbf{k_a}|} = \frac{2(E + m_e)(E_a - m_e) - m_a^2}{2|\mathbf{p}||\mathbf{k_a}|} \,. \tag{C.14}$$

The final Heaviside step function ensures that the trajectory of the ALP is within the angular coverage of the detector. For non-resonant annihilation, the integration boundaries in eqs. (3.2) and (3.3) are not applicable. Instead, using the Mandelstam variable $s = 2m_e(E + m_e)$, the integration boundaries are

$$x_{\mathrm{min}} = \frac{1}{E} \cdot \max\left[E_{\mathrm{cut}} \,, \frac{1}{4m_e}\left(s + m_a^2 - \sqrt{1 - \frac{4m_e^2}{s}}(s - m_a^2) \right) \right] \,, \tag{C.15}$$

$$x_{\mathrm{max}} = \frac{1}{E} \cdot \frac{1}{4m_e}\left(s + m_a^2 + \sqrt{1 - \frac{4m_e^2}{s}}(s - m_a^2) \right) \,, \tag{C.16}$$

$$E_{\mathrm{min}} = \frac{s_{\mathrm{min}}}{2m_e} - m_e \,, \quad E_{\mathrm{max}} = E_0 \,, \tag{C.17}$$

$$s_{\mathrm{min}} = \max\left[m_a^2 \,, m_e \frac{(m_a^2 - 2m_e E_{\mathrm{cut}})\left(\sqrt{E_{\mathrm{cut}}^2 - m_a^2} + E_{\mathrm{cut}} \right) + m_a^2 m_e}{m_a^2 - 2m_e E_{\mathrm{cut}} + m_e^2} \right] \,. \tag{C.18}$$

It is important to note that for $m_a > \sqrt{2m_e E_{\mathrm{cut}}}$, the diagrams contributing to the non-resonant production of ALPs exhibit an infrared divergence at the lower integration bound $s_{\mathrm{min}} = m_a^2$ due to the low-energy photon. This could be remedied by including one-loop corrections to the ALP-lepton vertex, which are, however, beyond the scope of the present work. We therefore neglect the sub-leading non-resonant production in this region of parameter space and focus solely on the dominant contribution from resonant production.

Non-resonant production of ALPs can also be compared to the production of Dark Photons in BaBar to convert bounds on Dark Photons into bounds on ALPs. The differential production cross section for Dark Photons for high lepton beam energies is approximately [120]

$$\frac{\mathrm{d}\sigma_{A'}}{\mathrm{d}\cos(\theta)} = \frac{2\pi\epsilon^2\alpha^2}{s}\left(1 - \frac{m_{A'}^2}{s}\right)\frac{1 + \cos^2(\theta) + \frac{m_{A'}^2/s}{(1-m_{A'}^2/s)^2}}{1 - \cos^2(\theta)}\,, \qquad (C.19)$$

with the Dark Photon mass $m_{A'}$ and the kinetic mixing parameter ϵ. In order for the Dark Photon to be detected, we require that the angles satisfy $\cos(\theta_{\min}) < \cos(\theta) < \cos(\theta_{\max})$ in the centre-of-mass system, where $\cos(\theta_{\min}) \approx -0.92$ and $\cos(\theta_{\max}) \approx 0.87$ [120]. By integrating and matching the cross sections of Dark Photons and ALPs

$$\int_{\cos(\theta_{\min})}^{\cos(\theta_{\max})} \mathrm{d}\cos(\theta)\,\frac{\mathrm{d}\sigma_{\mathrm{NR}}}{\mathrm{d}\cos(\theta)} = \int_{\cos(\theta_{\min})}^{\cos(\theta_{\max})} \mathrm{d}\cos(\theta)\,\frac{\mathrm{d}\sigma_{A'}}{\mathrm{d}\cos(\theta)}\,, \qquad (C.20)$$

we can derive constraints on our leptophilic ALP model using the constraints on $\sigma_{A'}$ in ref. [78].

D Comparison of E137 with other implementations

In this appendix, we compare our implementation of E137 and the resulting constraints on the ALP-lepton couplings with previous studies in the literature. In ref. [70], the concurrent effects of an ALP-photon and ALP-electron coupling were analysed. The authors give the constraints on a derivative coupling $g_{a\bar{e}e}^{\mathrm{eff}}$ for various values of $g_{a\gamma\gamma}^{\mathrm{eff}}$, which are however taken to be constant. Our results can be compared by taking the constraints in ref. [70] for $g_{a\gamma\gamma}^{\mathrm{eff}} = 0$ and comparing them to our tree-level model constraints from figure 9 ("Constant (P)" for Scenario E). This comparison is shown in figure 15. For low ALP masses, our constraints are in good agreement with those in ref. [70], but we obtain substantially stronger bounds for high masses.

It should be noted that ref. [70] does not consider the contributions to ALP production arising from electron-positron annihilation. Therefore, we also compare our constraints with those given in ref. [35], based on ref. [45], where this contribution is included. The original analysis from ref. [45] did not discuss the bounds for large couplings, which were estimated in ref. [35] assuming the ALPs to be monoenergetic. Clearly, this approximation is not very accurate and leads to a very different upper bound. A more accurate derivation of the upper bound was performed in ref. [29] using a formalism similar to ours. Indeed, the upper bound obtained in this case[4] agrees well with our results.

The shape of the lower bounds from refs. [29, 45], which include the contribution from positron annihilation, matches ours relatively well, but the bound is shifted to larger couplings. This shift may be attributed to the different implementations of the track length. Indeed, the track length distribution that we use (see figure 3) is larger for low energies than the one given in figure 14 of ref. [45], which is based on the (unpublished) EGS shower code. We note that our implementation agrees very well with the results from ref. [121], which has run GEANT4 simulations for E137. Another potential difference between our analysis and the one from ref. [45] is that the E137 analysis uses a much more sophisticated implementation

---

[4]Note that ref. [29] only considers ALPs with universal couplings to all leptons. The constraint on ALPs coupled only to electrons shown in figure 15 was kindly provided by Sophie Mutzel.
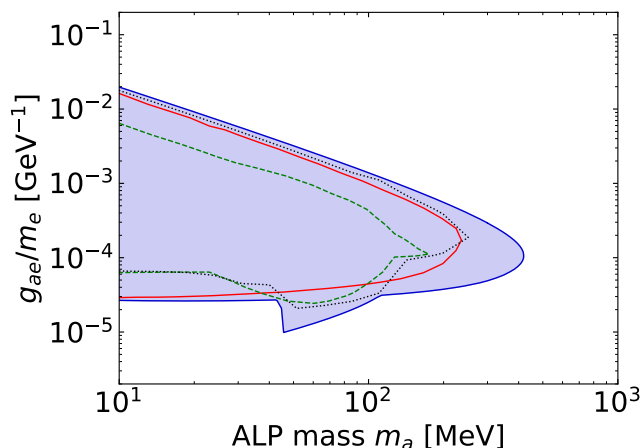
Figure 15. Constraints on the tree-level electron coupling from E137. The result from this work is shown as the blue-shaded region, while the results from refs. [35, 70], and [29] are shown as solid red, dashed green, and dotted black lines, respectively.

of the detector acceptance, tracking the ALP decay products and their interactions with the detector, which is beyond the scope of our analysis.

As discussed in appendix C.2, our bounds have a discontinuity at $m_a = \sqrt{2m_e E_{\rm cut}} \approx$ 45.2 MeV, because we conservatively neglect the non-resonant production channels. Including these channels is expected to further improve the agreement with ref. [45], which finds a continuous bound.

Data Availability Statement. This article has no associated data or the data will not be deposited.

Code Availability Statement. This article has no associated code or the code will not be deposited.

References

[1] F. Wilczek, *Problem of Strong P and T Invariance in the Presence of Instantons*, *Phys. Rev. Lett.* **40** (1978) 279 [INSPIRE].

[2] S. Weinberg, *A New Light Boson?*, *Phys. Rev. Lett.* **40** (1978) 223 [INSPIRE].

[3] R.D. Peccei and H.R. Quinn, *CP Conservation in the Presence of Instantons*, *Phys. Rev. Lett.* **38** (1977) 1440 [INSPIRE].

[4] R.D. Peccei and H.R. Quinn, *Constraints Imposed by CP Conservation in the Presence of Instantons*, *Phys. Rev. D* **16** (1977) 1791 [INSPIRE].

[5] M. Bauer, M. Neubert and A. Thamm, *Collider Probes of Axion-Like Particles*, *JHEP* **12** (2017) 044 [arXiv:1708.00443] [INSPIRE].

[6] Y. Nomura and J. Thaler, *Dark Matter through the Axion Portal*, *Phys. Rev. D* **79** (2009) 075008 [arXiv:0810.5397] [INSPIRE].

[7] M. Freytsis and Z. Ligeti, *On dark matter models with uniquely spin-dependent detection possibilities*, *Phys. Rev. D* **83** (2011) 115009 [arXiv:1012.5317] [INSPIRE].

[8] M.J. Dolan, F. Kahlhoefer, C. McCabe and K. Schmidt-Hoberg, *A taste of dark matter: Flavour constraints on pseudoscalar mediators*, *JHEP* **03** (2015) 171 [*Erratum ibid.* **07** (2015) 103] [arXiv:1412.5174] [INSPIRE].

[9] A. Berlin, S. Gori, T. Lin and L.-T. Wang, *Pseudoscalar Portal Dark Matter*, *Phys. Rev. D* **92** (2015) 015005 [arXiv:1502.06000] [INSPIRE].

[10] J.J. Fan, S.M. Koushiappas and G. Landsberg, *Pseudoscalar Portal Dark Matter and New Signatures of Vector-like Fermions*, *JHEP* **01** (2016) 111 [arXiv:1507.06993] [INSPIRE].

[11] J.M. No, *Looking through the pseudoscalar portal into dark matter: Novel mono-Higgs and mono-Z signatures at the LHC*, *Phys. Rev. D* **93** (2016) 031701 [arXiv:1509.01110] [INSPIRE].

[12] M. Bauer, U. Haisch and F. Kahlhoefer, *Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators*, *JHEP* **05** (2017) 138 [arXiv:1701.07427] [INSPIRE].

[13] S. Baek, P. Ko and J. Li, *Minimal renormalizable simplified dark matter model with a pseudoscalar mediator*, *Phys. Rev. D* **95** (2017) 075011 [arXiv:1701.04131] [INSPIRE].

[14] A. Kamada, H. Kim and T. Sekiguchi, *Axionlike particle assisted strongly interacting massive particle*, *Phys. Rev. D* **96** (2017) 016007 [arXiv:1704.04505] [INSPIRE].

[15] K. Kaneta, H.-S. Lee and S. Yun, *Dark photon relic dark matter production through the dark axion portal*, *Phys. Rev. D* **95** (2017) 115032 [arXiv:1704.07542] [INSPIRE].

[16] S. Banerjee et al., *Cornering pseudoscalar-mediated dark matter with the LHC and cosmology*, *JHEP* **07** (2017) 080 [arXiv:1705.02327] [INSPIRE].

[17] G. Arcadi et al., *Pseudoscalar Mediators: A WIMP model at the Neutrino Floor*, *JCAP* **03** (2018) 042 [arXiv:1711.02110] [INSPIRE].

[18] Y. Hochberg et al., *Strongly interacting massive particles through the axion portal*, *Phys. Rev. D* **98** (2018) 115031 [arXiv:1806.10139] [INSPIRE].

[19] A. Berlin et al., *Dark Matter, Millicharges, Axion and Scalar Particles, Gauge Bosons, and Other New Physics with LDMX*, *Phys. Rev. D* **99** (2019) 075001 [arXiv:1807.01730] [INSPIRE].

[20] P. deNiverville and H.-S. Lee, *Implications of the dark axion portal for SHiP and FASER and the advantages of monophoton signals*, *Phys. Rev. D* **100** (2019) 055017 [arXiv:1904.13061] [INSPIRE].

[21] D. Buttazzo, P. Panci, D. Teresi and R. Ziegler, *Xenon1T excess from electron recoils of non-relativistic Dark Matter*, *Phys. Lett. B* **817** (2021) 136310 [arXiv:2011.08919] [INSPIRE].

[22] L. Darmé, F. Giacchino, E. Nardi and M. Raggi, *Invisible decays of axion-like particles: constraints and prospects*, *JHEP* **06** (2021) 009 [arXiv:2012.07894] [INSPIRE].

[23] S.-F. Ge, X.-D. Ma and P. Pasquini, *Probing the dark axion portal with muon anomalous magnetic moment*, *Eur. Phys. J. C* **81** (2021) 787 [arXiv:2104.03276] [INSPIRE].

[24] S. Gola, S. Mandal and N. Sinha, *ALP-portal majorana dark matter*, *Int. J. Mod. Phys. A* **37** (2022) 2250131 [arXiv:2106.00547] [INSPIRE].

[25] V. Domcke, K. Schmitz and T. You, *Cosmological relaxation through the dark axion portal*, *JHEP* **07** (2022) 126 [arXiv:2108.11295] [INSPIRE].

[26] A.S. Zhevlakov, D.V. Kirpichnikov and V.E. Lyubovitskij, *Implication of the dark axion portal for the EDM of fermions and dark matter probing with NA64e, NA64μ, LDMX, M3, and BaBar*, *Phys. Rev. D* **106** (2022) 035018 [arXiv:2204.09978] [InSPIRE].

[27] M. Bauer, G. Rostagni and J. Spinner, *Axion-Higgs portal*, *Phys. Rev. D* **107** (2023) 015007 [arXiv:2207.05762] [InSPIRE].

[28] P. Panci, D. Redigolo, T. Schwetz and R. Ziegler, *Axion dark matter from lepton flavor-violating decays*, *Phys. Lett. B* **841** (2023) 137919 [arXiv:2209.03371] [InSPIRE].

[29] A. Bharucha, F. Brümmer, N. Desai and S. Mutzel, *Axion-like particles as mediators for dark matter: beyond freeze-out*, *JHEP* **02** (2023) 141 [arXiv:2209.03932] [InSPIRE].

[30] P.J. Fitzpatrick et al., *Dark matter through the axion-gluon portal*, *Phys. Rev. D* **108** (2023) 075003 [arXiv:2306.03128] [InSPIRE].

[31] D.K. Ghosh, A. Ghoshal and S. Jeesun, *Axion-like particle (ALP) portal freeze-in dark matter confronting ALP search experiments*, *JHEP* **01** (2024) 026 [arXiv:2305.09188] [InSPIRE].

[32] J.A. Dror, S. Gori and P. Munbodh, *QCD axion-mediated dark matter*, *JHEP* **09** (2023) 128 [arXiv:2306.03145] [InSPIRE].

[33] F. Capozzi et al., *New constraints on ALP couplings to electrons and photons from ArgoNeuT and the MiniBooNE beam dump*, *Phys. Rev. D* **108** (2023) 075019 [arXiv:2307.03878] [InSPIRE].

[34] G. Armando, P. Panci, J. Weiss and R. Ziegler, *Leptonic ALP portal to the dark sector*, *Phys. Rev. D* **109** (2024) 055029 [arXiv:2310.05827] [InSPIRE].

[35] W. Altmannshofer, J.A. Dror and S. Gori, *New Opportunities for Detecting Axion-Lepton Interactions*, *Phys. Rev. Lett.* **130** (2023) 241801 [arXiv:2209.00665] [InSPIRE].

[36] M. Bauer et al., *Flavor probes of axion-like particles*, *JHEP* **09** (2022) 056 [arXiv:2110.10698] [InSPIRE].

[37] R.Z. Ferreira, M.C.D. Marsh and E. Müller, *Strong supernovae bounds on ALPs from quantum loops*, *JCAP* **11** (2022) 057 [arXiv:2205.07896] [InSPIRE].

[38] C. Cornella, P. Paradisi and O. Sumensari, *Hunting for ALPs with Lepton Flavor Violation*, *JHEP* **01** (2020) 158 [arXiv:1911.06279] [InSPIRE].

[39] L. Calibbi, D. Redigolo, R. Ziegler and J. Zupan, *Looking forward to lepton-flavor-violating ALPs*, *JHEP* **09** (2021) 173 [arXiv:2006.04795] [InSPIRE].

[40] E. Bertuzzo, A.L. Foguel, G.M. Salla and R.Z. Funchal, *New Limits on Leptophilic Axionlike Particles and Majorons from ArgoNeuT*, *Phys. Rev. Lett.* **130** (2023) 171801 [arXiv:2202.12317] [InSPIRE].

[41] J. Bonilla, B. Gavela and J. Machado-Rodríguez, *Limits on ALP-neutrino couplings from loop-level processes*, *Phys. Rev. D* **109** (2024) 055023 [arXiv:2309.15910] [InSPIRE].

[42] A. Caputo and G. Raffelt, *Astrophysical Axion Bounds: The 2024 Edition*, *PoS* **COSMICWISPers** (2024) 041 [arXiv:2401.13728] [InSPIRE].

[43] P. Carenza and G. Lucente, *Supernova bound on axionlike particles coupled with electrons*, *Phys. Rev. D* **104** (2021) 103007 [*Erratum ibid.* **110** (2024) 049901] [arXiv:2107.12393] [InSPIRE].

[44] D.F.G. Fiorillo, T. Pitik and E. Vitagliano, *Supernova production of axionlike particles coupling to electrons, reloaded*, *Phys. Rev. D* **112** (2025) 083008 [arXiv:2503.15630] [InSPIRE].

[45] J.D. Bjorken et al., *Search for Neutral Metastable Penetrating Particles Produced in the SLAC Beam Dump*, *Phys. Rev. D* **38** (1988) 3375 [INSPIRE].

[46] NA64 collaboration, *NA64 Status Report 2020*, CERN-SPSC-2020-017 (2020).

[47] NA64 collaboration, *Improved limits on a hypothetical X(16.7) boson and a dark photon decaying into e^+e^- pairs*, *Phys. Rev. D* **101** (2020) 071101 [arXiv:1912.11389] [INSPIRE].

[48] J. Jaeckel and M. Spannowsky, *Probing MeV to 90 GeV axion-like particles with LEP and LHC*, *Phys. Lett. B* **753** (2016) 482 [arXiv:1509.00476] [INSPIRE].

[49] NA64 collaboration, *First Results in the Search for Dark Sectors at NA64 with the CERN SPS High Energy Muon Beam*, *Phys. Rev. Lett.* **132** (2024) 211803 [arXiv:2401.01708] [INSPIRE].

[50] H. Li, Z. Liu and N. Song, *Probing axion and flavored new physics with the NA64μ experiment*, *Phys. Rev. D* **112** (2025) 115013 [arXiv:2501.06294] [INSPIRE].

[51] H. Georgi, D.B. Kaplan and L. Randall, *Manifesting the Invisible Axion at Low-energies*, *Phys. Lett. B* **169** (1986) 73 [INSPIRE].

[52] J.S.R. Chisholm, *Change of variables in quantum field theories*, *Nucl. Phys.* **26** (1961) 469 [INSPIRE].

[53] S. Kamefuchi, L. O'Raifeartaigh and A. Salam, *Change of variables and equivalence theorems in quantum field theories*, *Nucl. Phys.* **28** (1961) 529 [INSPIRE].

[54] C. Arzt, *Reduced effective Lagrangians*, *Phys. Lett. B* **342** (1995) 189 [hep-ph/9304230] [INSPIRE].

[55] J.C. Criado and M. Pérez-Victoria, *Field redefinitions in effective theories at higher orders*, *JHEP* **03** (2019) 038 [arXiv:1811.09413] [INSPIRE].

[56] A.R. Zhitnitsky, *On Possible Suppression of the Axion Hadron Interactions* (in Russian), *Sov. J. Nucl. Phys.* **31** (1980) 260 [INSPIRE].

[57] M. Dine, W. Fischler and M. Srednicki, *A Simple Solution to the Strong CP Problem with a Harmless Axion*, *Phys. Lett. B* **104** (1981) 199 [INSPIRE].

[58] M. Kamionkowski and J. March-Russell, *Planck scale physics and the Peccei-Quinn mechanism*, *Phys. Lett. B* **282** (1992) 137 [hep-th/9202003] [INSPIRE].

[59] L. Di Luzio, M. Giannotti, E. Nardi and L. Visinelli, *The landscape of QCD axion models*, *Phys. Rept.* **870** (2020) 1 [arXiv:2003.01100] [INSPIRE].

[60] M. Badziak, G. Grilli di Cortona, M. Tabet and R. Ziegler, *Flavor-violating Higgs decays and stellar cooling anomalies in axion models*, *JHEP* **10** (2021) 181 [arXiv:2107.09708] [INSPIRE].

[61] J.E. Kim, *Weak Interaction Singlet and Strong CP Invariance*, *Phys. Rev. Lett.* **43** (1979) 103 [INSPIRE].

[62] M.A. Shifman, A.I. Vainshtein and V.I. Zakharov, *Can Confinement Ensure Natural CP Invariance of Strong Interactions?*, *Nucl. Phys. B* **166** (1980) 493 [INSPIRE].

[63] M. Chala, G. Guedes, M. Ramos and J. Santiago, *Running in the ALPs*, *Eur. Phys. J. C* **81** (2021) 181 [arXiv:2012.09017] [INSPIRE].

[64] N. Craig, A. Hook and S. Kasko, *The Photophobic ALP*, *JHEP* **09** (2018) 028 [arXiv:1805.06538] [INSPIRE].

[65] V. Shtabovenko, R. Mertig and F. Orellana, *FeynCalc 10: Do multiloop integrals dream of computer codes?*, *Comput. Phys. Commun.* **306** (2025) 109357 [arXiv:2312.14089] [INSPIRE].

[66] J. Bonilla, I. Brivio, M.B. Gavela and V. Sanz, *One-loop corrections to ALP couplings*, *JHEP* **11** (2021) 168 [arXiv:2107.11392] [INSPIRE].

[67] F. Ertas, *Phenomenology of light pseudoscalar particles*, Ph.D. thesis, RWTH Aachen University, Aachen, Germany (2021) [INSPIRE].

[68] C. Cesarotti and R. Gambhir, *The new physics case for beam-dump experiments with accelerated muon beams*, *JHEP* **05** (2024) 283 [arXiv:2310.16110] [INSPIRE].

[69] D. Kim, J. Yu, J.-C. Park and H. Kim, *The Beam-Dump Ceiling and Its Experimental Implication: The Case of a Portable Experiment*, arXiv:2401.09529 [INSPIRE].

[70] J. Liu, Y. Luo and M. Song, *Investigation of the concurrent effects of ALP-photon and ALP-electron couplings in Collider and Beam Dump Searches*, *JHEP* **09** (2023) 104 [arXiv:2304.05435] [INSPIRE].

[71] Y.-S. Liu and G.A. Miller, *Validity of the Weizsäcker-Williams approximation and the analysis of beam dump experiments: Production of an axion, a dark photon, or a new axial-vector boson*, *Phys. Rev. D* **96** (2017) 016004 [arXiv:1705.01633] [INSPIRE].

[72] Y.-S. Tsai and V. Whitis, *Thick target bremsstrahlung and target consideration for secondary particle production by electrons*, *Phys. Rev.* **149** (1966) 1248 [INSPIRE].

[73] K.J. Kim and Y.-S. Tsai, *Improved weizsacker-williams method and its application to lepton and w boson pair production*, *Phys. Rev. D* **8** (1973) 3109 [INSPIRE].

[74] L. Marsicano et al., *Probing Leptophilic Dark Sectors at Electron Beam-Dump Facilities*, *Phys. Rev. D* **98** (2018) 115022 [arXiv:1812.03829] [INSPIRE].

[75] R.R. Dusaev, D.V. Kirpichnikov and M.M. Kirsanov, *Photoproduction of axionlike particles in the NA64 experiment*, *Phys. Rev. D* **102** (2020) 055018 [arXiv:2004.04469] [INSPIRE].

[76] M. Chala, M. Duerr, F. Kahlhoefer and K. Schmidt-Hoberg, *Tricking Landau-Yang: How to obtain the diphoton excess from a vector resonance*, *Phys. Lett. B* **755** (2016) 145 [arXiv:1512.06833] [INSPIRE].

[77] L3 collaboration, *Tests of QED at LEP energies using $e^+e^- \to \gamma\gamma(\gamma)$ and $e^+e^- \to \ell^+\ell^-\gamma\gamma$*, *Phys. Lett. B* **353** (1995) 136 [INSPIRE].

[78] BaBar collaboration, *Search for a Dark Photon in e^+e^- Collisions at BaBar*, *Phys. Rev. Lett.* **113** (2014) 201801 [arXiv:1406.2980] [INSPIRE].

[79] Belle-II collaboration, *Search for a $\mu^+\mu^-$ resonance in four-muon final states at Belle II*, *Phys. Rev. D* **109** (2024) 112015 [arXiv:2403.02841] [INSPIRE].

[80] SINDRUM collaboration, *Limits for Shortlived Neutral Particles Emitted μ^+ or π^+ Decay*, *Phys. Lett. B* **175** (1986) 101 [INSPIRE].

[81] A.A. Poblaguev et al., *Experimental study of the radiative decays $K^+ \to \mu^+\nu e^+e^-$ and $K^+ \to e^+\nu e^+e^-$*, *Phys. Rev. Lett.* **89** (2002) 061803 [hep-ex/0204006] [INSPIRE].

[82] K. Cheung, J.-L. Kuo, P.-Y. Tseng and Z.S. Wang, *Atmospheric axionlike particles at Super-Kamiokande*, *Phys. Rev. D* **106** (2022) 095029 [arXiv:2208.05111] [INSPIRE].

[83] J. Alda et al., *Collider and astrophysical signatures of light scalars with enhanced τ couplings*, *JHEP* **06** (2025) 008 [arXiv:2407.18296] [INSPIRE].

[84] BESIII collaboration, *Search for the decay $J/\psi \to \gamma +$ invisible*, *Phys. Rev. D* **101** (2020) 112005 [arXiv:2003.05594] [INSPIRE].

[85] BESIII collaboration, *Search for diphoton decays of an axionlike particle in radiative J/ψ decays*, *Phys. Rev. D* **110** (2024) L031101 [`arXiv:2404.04640`] [INSPIRE].

[86] D.F.G. Fiorillo et al., *Supernova simulations confront SN 1987A neutrinos*, *Phys. Rev. D* **108** (2023) 083040 [`arXiv:2308.01403`] [INSPIRE].

[87] G.G. Raffelt, *Stars as laboratories for fundamental physics: The astrophysics of neutrinos, axions, and other weakly interacting particles*, Chicago University Press, Chicago, U.S.A. (1996).

[88] G.G. Raffelt, *Astrophysical axion bounds*, *Lect. Notes Phys.* **741** (2008) 51 [`hep-ph/0611350`] [INSPIRE].

[89] S.W. Falk and D.N. Schramm, *Limits From Supernovae on Neutrino Radiative Lifetimes*, *Phys. Lett. B* **79** (1978) 511 [INSPIRE].

[90] A. Sung, H. Tu and M.-R. Wu, *New constraint from supernova explosions on light particles beyond the Standard Model*, *Phys. Rev. D* **99** (2019) 121305 [`arXiv:1903.07923`] [INSPIRE].

[91] A. Caputo, G. Raffelt and E. Vitagliano, *Muonic boson limits: Supernova redux*, *Phys. Rev. D* **105** (2022) 035022 [`arXiv:2109.03244`] [INSPIRE].

[92] A. Caputo, H.-T. Janka, G. Raffelt and E. Vitagliano, *Low-Energy Supernovae Severely Constrain Radiative Particle Decays*, *Phys. Rev. Lett.* **128** (2022) 221103 [`arXiv:2201.09890`] [INSPIRE].

[93] D.F.G. Fiorillo, T. Pitik and E. Vitagliano, *Energy Transfer by Feebly Interacting Particles in Supernovae: The Trapping Regime*, *Phys. Rev. Lett.* **135** (2025) 071005 [`arXiv:2503.13653`] [INSPIRE].

[94] D. Croon, G. Elor, R.K. Leane and S.D. McDermott, *Supernova Muons: New Constraints on Z' Bosons, Axions and ALPs*, *JHEP* **01** (2021) 107 [`arXiv:2006.13942`] [INSPIRE].

[95] A. Caputo, G. Raffelt and E. Vitagliano, *Radiative transfer in stars by feebly interacting bosons*, *JCAP* **08** (2022) 045 [`arXiv:2204.11862`] [INSPIRE].

[96] B. Döbrich et al., *ALPtraum: ALP production in proton beam dump experiments*, *JHEP* **02** (2016) 018 [`arXiv:1512.03069`] [INSPIRE].

[97] B. Döbrich, J. Jaeckel and T. Spadaro, *Light in the beam dump - ALP production from decay photons in proton beam-dumps*, *JHEP* **05** (2019) 213 [*Erratum ibid.* **10** (2020) 046] [`arXiv:1904.02091`] [INSPIRE].

[98] J. Jerhot et al., *ALPINIST: Axion-Like Particles In Numerous Interactions Simulated and Tabulated*, *JHEP* **07** (2022) 094 [`arXiv:2201.05170`] [INSPIRE].

[99] C. Rella, B. Döbrich and T.-T. Yu, *Searching for muonphilic dark sectors with proton beams*, *Phys. Rev. D* **106** (2022) 035023 [`arXiv:2205.09870`] [INSPIRE].

[100] FASER collaboration, *Shining light on the dark sector: search for axion-like particles and other new physics in photonic final states with FASER*, *JHEP* **01** (2025) 199 [`arXiv:2410.10363`] [INSPIRE].

[101] M.A. Buen-Abad, J.J. Fan, M. Reece and C. Sun, *Challenges for an axion explanation of the muon g − 2 measurement*, *JHEP* **09** (2021) 101 [`arXiv:2104.03267`] [INSPIRE].

[102] C. Davies, *Muon g-2*, *PoS* **LATTICE2024** (2025) 019 [`arXiv:2503.03364`] [INSPIRE].

[103] T. Aoyama et al., *The anomalous magnetic moment of the muon in the Standard Model*, *Phys. Rept.* **887** (2020) 1 [`arXiv:2006.04822`] [INSPIRE].

[104] MUON G-2 collaboration, *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.20 ppm*, *Phys. Rev. Lett.* **131** (2023) 161802 [arXiv:2308.06230] [INSPIRE].

[105] S. Borsanyi et al., *Leading hadronic contribution to the muon magnetic moment from lattice QCD*, *Nature* **593** (2021) 51 [arXiv:2002.12347] [INSPIRE].

[106] A. Boccaletti et al., *High precision calculation of the hadronic vacuum polarisation contribution to the muon anomaly*, arXiv:2407.10913 [INSPIRE].

[107] R.H. Parker et al., *Measurement of the fine-structure constant as a test of the Standard Model*, *Science* **360** (2018) 191 [arXiv:1812.04130] [INSPIRE].

[108] 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* **588** (2020) 61 [INSPIRE].

[109] X. Fan, T.G. Myers, B.A.D. Sukra and G. Gabrielse, *Measurement of the Electron Magnetic Moment*, *Phys. Rev. Lett.* **130** (2023) 071801 [arXiv:2209.13084] [INSPIRE].

[110] CMS collaboration, *Observation of $\gamma\gamma \to \tau\tau$ in proton-proton collisions and limits on the anomalous electromagnetic moments of the τ lepton*, *Rept. Prog. Phys.* **87** (2024) 107801 [arXiv:2406.03975] [INSPIRE].

[111] P.F. Depta, M. Hufnagel and K. Schmidt-Hoberg, *Updated BBN constraints on electromagnetic decays of MeV-scale particles*, *JCAP* **04** (2021) 011 [arXiv:2011.06519] [INSPIRE].

[112] SHiP collaboration, *BDF/SHiP at the ECN3 high-intensity beam facility*, CERN-SPSC-2023-033 (2023).

[113] J. Adhikary et al., *Scientific program for the Forward Physics Facility*, *Eur. Phys. J. C* **85** (2025) 430 [arXiv:2411.04175] [INSPIRE].

[114] NA64 collaboration, *NA64 status report 2023*, CERN-SPSC-2023-023 (2023).

[115] NA64 collaboration, *Search for Light Dark Matter with NA64 at CERN*, *Phys. Rev. Lett.* **131** (2023) 161801 [arXiv:2307.02404] [INSPIRE].

[116] NA64 collaboration, *NA64 Status Report 2018*, CERN-SPSC-2018-015 (2018).

[117] NA64 collaboration, *NA64 Status Report 2024*, CERN-SPSC-2024-024 (2024).

[118] PARTICLE DATA GROUP collaboration, *Review of particle physics*, *Phys. Rev. D* **110** (2024) 030001 [INSPIRE].

[119] P. Fayet and M.O. Olea-Romacho, *Searching for a new light gauge boson with axial couplings in muon beam dump experiments*, *JHEP* **07** (2024) 223 [arXiv:2405.02104] [INSPIRE].

[120] R. Essig, P. Schuster and N. Toro, *Probing Dark Forces and Light Hidden Sectors at Low-Energy e+e- Colliders*, *Phys. Rev. D* **80** (2009) 015003 [arXiv:0903.3941] [INSPIRE].

[121] L. Marsicano et al., *Dark photon production through positron annihilation in beam-dump experiments*, *Phys. Rev. D* **98** (2018) 015031 [arXiv:1802.03794] [INSPIRE].