Probing charged lepton flavor violation with axion-like particles at Belle II

Kingman Cheung,[a,b,c] **Abner Soffer,**[d] **Zeren Simon Wang**[a,b] **and Yu-Heng Wu**[a]

[a]*Department of Physics, National Tsing Hua University,*
Hsinchu 300, Taiwan

[b]*Center for Theory and Computation, National Tsing Hua University,*
Hsinchu 300, Taiwan

[c]*Division of Quantum Phases and Devices, School of Physics, Konkuk University,*
Seoul 143-701, Republic of Korea

[d]*School of Physics and Astronomy, Tel Aviv University,*
Tel Aviv 69978, Israel

E-mail: cheung@phys.nthu.edu.tw, asoffer@tau.ac.il,
wzs@mx.nthu.edu.tw, yuheng1k996@gapp.nthu.edu.tw

ABSTRACT: We study charged lepton flavor violation associated with a light leptophilic axion-like particle (ALP), X, at the B-factory experiment Belle II. We focus on production of the ALP in the tau decays $\tau \to Xl$ with $l = e, \mu$, followed by its decay via $X \to l^-l^+$. The ALP can be either promptly decaying or long-lived. We perform Monte-Carlo simulations, recasting a prompt search at Belle for lepton-flavor-violating τ decays, and propose a displaced-vertex (DV) search. For both types of searches, we derive the Belle II sensitivity reaches in both the product of branching fractions and the ALP coupling constants, as functions of the ALP mass and lifetime. The results show that the DV search exceeds the sensitivity reach of the prompt search to the relevant branching fractions by up to about a factor of 40 in the long decay length regime.

KEYWORDS: Beyond Standard Model, Effective Field Theories

Contents

1   Introduction

The latest experimental bounds on the neutron electric dipole moment have constrained the strength of Quantum Chromodynamics (QCD) θ-term in the Standard Model (SM) of particle physics to be less than $\sim 10^{-10}$ [1, 2], pointing to a fine-tuning issue, known as the strong CP problem in QCD [3–5]. One of the most compelling solutions to this problem is to introduce a new global symmetry $U(1)_{PQ}$, which is spontaneously broken by a dynamical CP-conserving axion field [3, 4]. Doing so predicts a new pseudoscalar boson, the pseudo-Nambu-Goldstone boson of the broken $U(1)_{PQ}$ symmetry, called the QCD axion, a. The breaking scale of the new symmetry has to be much higher than the electroweak scale: $f_a \gtrsim 10^9$ GeV [6]. Since the mass of the QCD axion, m_a, as well as its couplings to

the SM particles, are both inversely proportional to f_a, the QCD axion should be very light ($10^{-6}\,\mathrm{eV} < m_a < 10^{-2}\,\mathrm{eV}$) and feebly interacting, with a very long lifetime [5].

Similar to the QCD axion, an axion-like particle (ALP), denoted X throughout this work, is a pseudoscalar boson that interacts feebly with SM particles. Unlike the QCD axion, an ALP does not have a linear proportionality relation between its mass m_X and its couplings to the SM particles, and hence does not necessarily solve the strong CP problem. ALPs can be as light as $10^{-22}\,\mathrm{eV}$ [7, 8] or as heavy as $1\,\mathrm{TeV}$ (heavy GUT axions) [9]. Axions and ALPs are considered some of the most plausible candidates for dark matter [10–16].

ALPs can be potentially produced at various experiments, including the Large Hadron Collider and B-factories (see for instance refs. [17–25] for related studies). Depending on their masses and strengths of the couplings to the SM particles, ALPs can be short-lived and decay promptly, or be long-lived and decay after traveling a macroscopic distance, leading to signatures of displaced objects or missing energy. ALPs can couple to gauge bosons as well as quarks and leptons, giving rise to a large class of possible experimental signatures. For instance, using its initial dataset of about $445\,\mathrm{pb}^{-1}$, Belle II has reported the results of a search for a light pseudoscalar boson that couples to photons [26].

In this paper, we restrict ourselves to the case in which the ALPs are leptophilic, i.e., couple at tree level only to the charged leptons, and neglect couplings to neutrinos. In general, the ALP couplings to the charged leptons are not necessarily diagonal, and may have charged-lepton-flavor-violating (LFV) interactions [27]. Since the SM predicts no flavor-changing neutral current (FCNC) at tree level, experimental observations of LFV can be a clear signature of new physics. For muon LFV processes such as $\mu \to e\gamma$ and $\mu \to 3e$, the SM predictions for the branching fractions are about 10^{-55} [28, 29], much lower than the present experimental limits $\mathrm{BR}(\mu \to e\gamma) < 4.2 \times 10^{-13}$ [30] and $\mathrm{BR}(\mu \to 3e) < 1.0 \times 10^{-12}$ [31]. The future experiments Mu3e and MEG II will have an improved sensitivity to these processes by as much as four orders of magnitude [32, 33]. Compared to the muon-related LFV, LFV effects in the τ sector are much more loosely constrained, as we briefly discuss below. Studies of FCNCs with the ALPs can be found for the lepton sector in refs. [34–39] and for the quark sector in refs. [25, 40–42].

In order to study the ALP couplings to the τ lepton, we consider the ongoing B-factory Belle II experiment [43, 44] at the SuperKEKB collider, where energy-asymmetric collisions of electron and positron beams at the center-of-mass energy $\sqrt{s} = 10.58\,\mathrm{GeV}$ take place. Belle II is planned to collect about 4.6×10^{10} $e^+e^- \to \tau^+\tau^-$ events with the design integrated luminosity of $50\,\mathrm{ab}^{-1}$ [45]. The large dataset will enable probing rare τ decays, including LFV decays, with unprecedented precision.

We propose to search for the decays $\tau \to Xl$ with $X \to l^-l^+$, where l can be either an electron or a muon. This restricts the ALP mass to the range $2\,m_l < m_X < m_\tau - m_l$. In this mass range, the strongest published limits on lepton-flavor-conserving ALP couplings stem from the beam-dump experiments conducted at SLAC [46, 47] and a dark-photon search performed at BaBar [21, 48], both of which assumed no tree-level LFV couplings. Bounds on the LFV ALP couplings can also be derived from $\tau \to Xe$, where X is invisible (see studies at ARGUS [49], Belle [50], and Belle II [38, 51–54]). This signature, which is not covered in this work, suffers from large background from $\tau \to l\nu\nu$, requiring more

sophisticated methods to suppress. In section 3.4 we extract stronger constraints from this signature based on test of lepton-flavor universality in leptonic τ decays [55], relevant for the case of interest of non-universal couplings. For past studies on probing LFV associated with an ALP at Belle II, see e.g. refs. [39, 56].

The previous Belle experiment [57] has searched for lepton-flavor-violating τ decays into three charged leptons and obtained limits on the decay branching ratios in the order of magnitude of 10^{-8} [58] (see also ref. [59] for a similar search at the BaBar experiment). In this work we recast that search into the studied theoretical scenarios of a leptophilic ALP, obtaining tighter constraints than previously published ones. Assuming equal selection efficiencies at Belle and Belle II, we calculate the expected Belle II limits on the model-independent as well as model-dependent parameters.

In the main part of this work, we propose a displaced-vertex (DV) search strategy that targets long-lived ALPs. Such long-lived ALPs belong to a large class of long-lived particles (LLPs) widely predicted in various beyond-the-SM models. See refs. [60–62] for reviews of LLP collider searches and models. For past studies of LLPs at Belle II, see, e.g., refs. [63–73]. Experimental LLP searches conducted at BaBar and Belle appear in refs. [74, 75]. We note that ref. [76] proposed for the first time to probe LFV with DVs, considering both muon and τ decays, although no detailed Monte-Carlo (MC) simulation was performed.

The organization of this paper is as follows. In section 2 we introduce the effective model under consideration, give the relevant analytic formulas for τ and X decay rates, and list the two benchmark scenarios we focus on for numerical studies. We summarize the current constraints on the ALP couplings to the charged leptons in section 3. In section 4 we briefly describe the Belle II experiment and detail the simulation procedures that we use to study the sensitivity of prompt and displaced $\tau \to Xl$ searches on the ALP couplings. The numerical results are shown and discussed in section 5, and we conclude in section 6 with a summary and an outlook.

2 Model

We consider a phenomenological model in the framework of effective field theory of dimension-five for a pseudo-Goldstone boson X, corresponding to the ALP considered in this work. The X couples at tree level (only) to the SM charged leptons e, μ, or τ, with the following Lagrangian [6]

$$
\begin{aligned}
\mathcal{L}_X &= \frac{\partial_\mu X}{\Lambda} \bar{l}_\alpha G_{\alpha\beta} \gamma^\mu (1 + \gamma^5) l_\beta \\
&= -i \frac{X}{\Lambda} \bar{l}_\alpha G_{\alpha\beta} ((m_\alpha - m_\beta) + (m_\alpha + m_\beta)\gamma^5) l_\beta,
\end{aligned}
\tag{2.1}
$$

where Λ labels the effective scale below which the model is applicable, G denotes a symmetric matrix of real couplings ($G_{\alpha\beta} = G_{\beta\alpha}$), and $l_{\alpha,\beta} = e, \mu, \tau$ with mass $m_{\alpha,\beta}$. To derive the second line of eq. (2.1), we apply integration by parts and the equations of motion. For simplicity we have assumed that the scalar and pseudoscalar couplings are of equal strength for the same lepton flavor combinations.

In order to simplify the notation, we define the following coupling matrix,

$$g_{\alpha\beta} \equiv G_{\alpha\beta}/\Lambda. \tag{2.2}$$

We can thus express the tree-level decay width of one charged lepton l_α into a lighter lepton l_β plus X with the formula

$$\Gamma(l_\alpha \to l_\beta X) = \frac{m_\alpha^3}{8\pi} \sqrt{\left(1 - \left(\frac{m_\beta + m_X}{m_\alpha}\right)^2\right)\left(1 - \left(\frac{m_\beta - m_X}{m_\alpha}\right)^2\right)}$$
$$\times\, g_{\alpha\beta}^2 \left[\left(1 - \frac{m_\beta^2}{m_\alpha^2}\right)^2 - \frac{m_X^2}{m_\alpha^2}\left(1 + \frac{m_\beta^2}{m_\alpha^2}\right)\right]. \tag{2.3}$$

In the limit $m_{\beta,X} \ll m_\alpha$, eq. (2.3) is reduced to the simple form $m_\alpha^3 g_{\alpha\beta}^2/8\pi$. Our numerical calculations are performed with the full form given in eq. (2.3).

The two-body decay widths of X into a pair of charged leptons are given by

$$\Gamma(X \to l_\beta^- l_\beta^+) = \frac{m_X\, m_\beta^2}{2\pi} g_{\beta\beta}^2 \sqrt{1 - \frac{4m_\beta^2}{m_X^2}}, \tag{2.4}$$

$$\Gamma(X \to l_\alpha^- l_\beta^+ \text{ or } l_\alpha^+ l_\beta^-) \simeq \frac{m_X\, m_\alpha^2}{2\pi} g_{\alpha\beta}^2 \left(1 - \frac{m_\alpha^2}{m_X^2}\right)^2, \tag{2.5}$$

where the two equations are for decays into same-flavor and different-flavor charged leptons, and the second equation is obtained in the limit $m_\beta \ll m_\alpha$.

We note that the flavor-diagonal couplings $g_{\alpha\alpha}$ can induce X decays into a pair of photons, as well as those into a pair of neutrinos, both at one-loop order. Compared to the di-photon channel, the di-neutrino decay is highly suppressed (see the relevant discussions in ref. [21]) and is hence ignored in this work. The corresponding decay width can be calculated with: [21]

$$\Gamma(X \to \gamma\gamma) = 4\pi\alpha^2 m_X^3 |g_{\gamma\gamma}^{\text{eff}}|^2, \tag{2.6}$$

where $g_{\gamma\gamma}^{\text{eff}}$ is an effective coupling that can be derived from the ALP-lepton couplings:

$$g_{\gamma\gamma}^{\text{eff}} = \frac{1}{8\pi^2} \sum_{\alpha=e,\mu,\tau} g_{\alpha\alpha}\, B_1(4m_\alpha^2/m_X^2), \tag{2.7}$$

where the loop function B_1 is defined as follows:

$$B_1(t) = 1 - t f^2(t), \text{ with } f(t) = \begin{cases} \arcsin\left(\frac{1}{\sqrt{t}}\right), & \text{if } t \geq 1 \\ \frac{i}{2}\ln\left(\frac{1+\sqrt{1-t}}{1-\sqrt{1-t}}\right) + \frac{\pi}{2}, & \text{if } t < 1 \end{cases}. \tag{2.8}$$

Note that both the tree-level couplings $g_{\alpha\alpha}$ and the effective coupling $g_{\gamma\gamma}^{\text{eff}}$ are of mass dimension -1, cf. eq. (2.2). With eqs. (2.4), (2.5), and (2.7), we can compute the total decay width and, hence, the proper lifetime of the ALP as a function of its mass and the couplings to the charged leptons.

	Scenario 1	Scenario 2
$g_{\tau\alpha}$ tau decays	$g_{\tau e}$ $\tau \to Xe$	$g_{\tau\mu}$ $\tau \to X\mu$
$g_{\beta\beta}$ X decays	g_{ee} $X \to e^-e^+(\text{sig.})/\gamma\gamma$	$g_{\mu\mu}$ $X \to \mu^-\mu^+(\text{sig.})/\gamma\gamma$

Table 1. Summary of the basic features of the two benchmark scenarios studied in this work, including the couplings switched on and the corresponding induced decays of τ and X. "Sig." indicates the signature studied.

In this paper, we concentrate on the scenarios where we switch on two couplings: one dominating the production and another governing the decay of the ALP. Specifically, the ALP is produced in the decay $\tau \to l_\alpha X$ via the production couplings $g_{\tau\alpha}$ $(\alpha = e, \mu)$, and undergoes the decay $X \to e^-e^+$ or $\mu^+\mu^-$ via the decay couplings $g_{\alpha\alpha}$. For simplicity, we do not consider X decays via the off-diagonal coupling $g_{e\mu}$,[1] but will offer brief comments when we discuss the numerical results in section 5. Moreover, the $g_{\tau\tau}$ coupling is also switched off, so we do not consider ALP production by final-state radiation.

We should point out that in principle, the production coupling $g_{\tau\alpha}$ alone can also induce (lepton/quark-level) four-body X decays via an off-shell τ, as well as, sub-dominantly, the two-body decays $X \to \nu_\tau \bar{\nu}_\alpha$ (or their CP conjugate) via a triangular loop with a W-boson. However, these contributions are negligible compared to those from the tree-level decays considered here, unless $g_{\tau\alpha} \gg g_{\alpha\alpha}$. In the case that the diagonal couplings are indeed negligible, the ALP becomes very long-lived, requiring a missing-energy search strategy [49, 54] that is not the topic of this paper. This particular caveat will not affect our numerical results, and so the highly suppressed X decays via $g_{\tau\alpha}$ couplings are not included in the numerical estimates. For $m_X > m_\tau$, limits on $g_{\tau\alpha}$ can be derived from $e^-e^+ \to l^\pm l^\pm \tau^\mp \tau^\mp$ and $e^-e^+ \to l^\pm \tau^\mp + \text{ missing }$ [56].
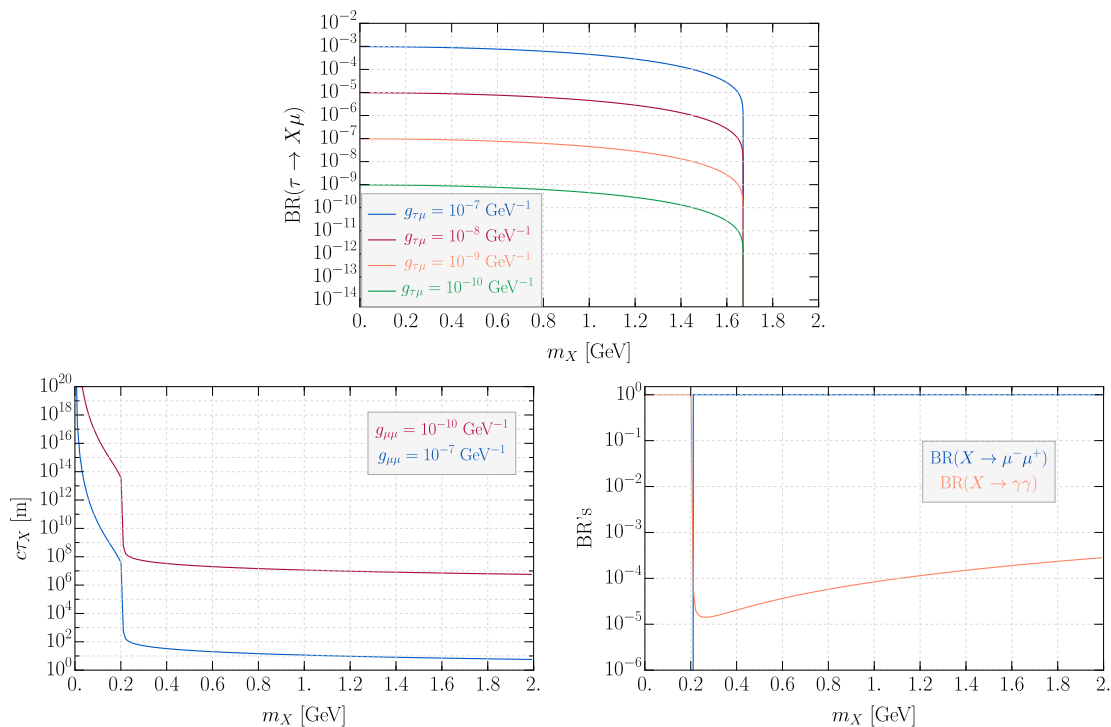
2.1 Benchmark scenarios

In our numerical studies, we focus on two representative benchmark scenarios. In Scenario 1, $g_{\tau e}$ and g_{ee} are switched on, and the other couplings are set to zero. As a result, $g_{\tau e}$ induces the decay $\tau \to Xe$ and g_{ee} leads to $X \to e^-e^+$ and $\gamma\gamma$. In Scenario 2, only $g_{\tau\mu}$ and $g_{\mu\mu}$ are non-zero, so that the ALP is produced via $\tau \to X\mu$ and decays to $\mu^-\mu^+$ and $\gamma\gamma$. We focus on the di-lepton final state and do not consider the di-photon decay mode as a signature of interest in this work. We summarize these two scenarios in table 1.

In figure 1 and figure 2 we present plots of the branching fraction $\text{BR}(\tau \to Xl_\alpha)$, the ALP proper flight distance $c\tau_X$, and its decay branching fractions as functions of m_X for the two benchmark scenarios. The upper plots of figure 1 and figure 2 show the variation of $\text{BR}(\tau \to Xl_\alpha)$ with respect to m_X, for several benchmark values of $g_{\tau\alpha}$. Intriguingly, the X branching fraction plots show that $\text{BR}(X \to \gamma\gamma)$ exceeds $\text{BR}(X \to e^-e^+)$ for $m_X \gtrsim 0.6\,\text{GeV}$ in Scenario 1, despite the loop suppression in $(X \to \gamma\gamma)$. This observation can be understood from the fact that $\Gamma(X \to l_\alpha^- l_\alpha^+) \propto m_X m_\alpha^2$, while $\Gamma(X \to \gamma\gamma) \propto m_X^3$,

---

[1]For a study on this coupling and LFV at Belle II, we refer the reader to ref. [39].
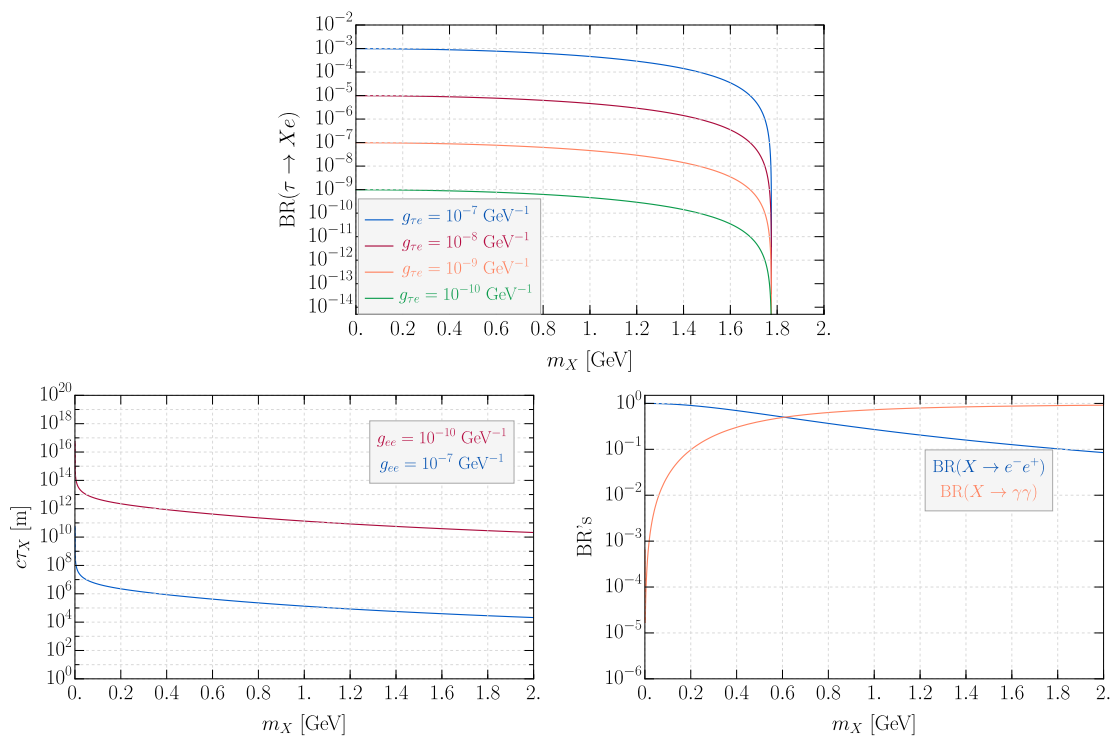
Figure 1. Scenario 1. *Top*: branching ratio of $\tau \to Xe$ vs. m_X for $g_{\tau e} = 10^{-7}$, 10^{-8}, 10^{-9}, $10^{-10}\,\text{GeV}^{-1}$. *Bottom left*: ALP proper flight distance $c\tau_X$ vs. m_X for $g_{ee} = 10^{-10}, 10^{-7}\,\text{GeV}^{-1}$. *Bottom right*: branching ratios of $X \to e^-e^+$ or $\gamma\gamma$, as a function of m_X.

Figure 2. Scenario 2. The same format as in figure 1, replacing e by μ.

cf. eqs. (2.5) and (2.7). Therefore, given the tiny electron mass, $\Gamma(X \to e^- e^+)$ can become subdominant for sufficiently large m_X.

Besides the two benchmark scenarios introduced above, there are, of course, additional scenarios with other possible lepton flavor combinations. In section 5.3, we will qualitatively comment on their numerical results instead of performing detailed studies.

3 Constraints

In this section, we discuss various experimental and astrophysical results from which we derive constraints on the model parameters related to Scenarios 1 and 2 introduced in section 2. The experimental results come from LEP, B-factories, and muon fixed-target experiments, while the astrophysical bounds stem mainly from the supernova 1987A (SN1987A). We also refer the reader to ref. [37] for a comprehensive study on probes of ALPs with LFV.

3.1 LEP

The LEP-2 collider at CERN provided samples of $e^- e^+ \to \tau^- \tau^+$ events at center-of-mass energies $\sqrt{s}$ ranging from 189 to 209 GeV. The inclusive cross section of $e^- e^+ \to \tau^- \tau^+$ at $\sqrt{s} = 207$ GeV was measured to be $\sigma_{\tau\tau}^{\mathrm{LEP}} = 6.05 \pm 0.39\,(\mathrm{stat}) \pm 0.09\,(\mathrm{syst})$ pb with 402 observed events at the ALEPH experiment [77], in good agreement with the SM predictions. If the $g_{\tau e}$ coupling is non-zero, it can induce the same process via a t-channel ALP exchange, modifying the theoretical computation of the inclusive cross section. Thus, we use the experimental measurement to place a bound on the coupling $g_{\tau e}$.

We implement the ALP model (see Lagrangian in eq. (2.1)) using FeynRules [78, 79] to generate an UFO (Universal FeynRules Output) [80] model file, with which we simulate $e^- e^+ \to \tau^- \tau^+$ events and compute the corresponding cross sections for given m_X and $g_{\tau e}$ with MadGraph5 2.7.3 [81]. By scanning over a 2D grid of $(m_X, g_{\tau e})$, we exclude the points where the computed cross section lies outside the two-sided range around the central value of the ALEPH measurement: $[\sigma_{\tau\tau}^{\mathrm{LEP}} - 2\Delta\sigma,\ \sigma_{\tau\tau}^{\mathrm{LEP}} + 2\Delta\sigma]$, where $\Delta\sigma \simeq \sqrt{0.39^2 + 0.09^2}$ pb $= 0.40$ pb is the total experimental uncertainty on $\sigma_{\tau\tau}^{\mathrm{LEP}}$. The theoretical uncertainty is negligible and hence neglected. For $m_X \lesssim 2$ GeV the limit is roughly a constant at $g_{\tau e} < 0.078\,\mathrm{GeV}^{-1}$, because of the relatively large center-of-mass energy. This is a very weak constraint relative to others studied here.

We also note that if one switches on the coupling $g_{\tau\tau}$ in Scenario 1 together with g_{ee}, it can induce τ-pair production via an s-channel X in $e^- e^+$ collisions, thus allowing to probe the product of these two couplings. Such a scenario is beyond the scope of this work. Moreover, the g_{ee} coupling itself affects Møller and Bhabha scattering. A precise measurement of the parity violation asymmetry in Møller scattering has been performed in a fixed target experiment [82]. Nevertheless, the axion model under consideration only has the pseudoscalar coupling to electron after taking the derivative of the field X (see the second line of eq. (2.1)), so it does not contribute to parity violation.

A limit on the ALP interactions using Bhabha scattering excludes values of $\Gamma(X)[\mathrm{BR}(X \to e^- e^+)]^2$ above $1.3 \cdot 10^{-3}$ eV for m_X between 1.75 and 1.88 MeV [83]. This
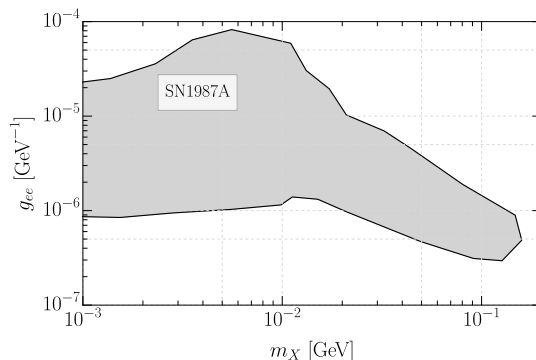
Figure 3. Constraint on g_{ee} with respect to m_X, derived in ref. [89] from the supernova cooling efficiencies.

can be translated to a weak bound of $g_{ee} \gtrsim 0.15 \, \mathrm{GeV}^{-1}$ for $m_X = 1.8 \, \mathrm{MeV}$ in our model, cf. eqs. (2.4) and (2.6).

The coupling $g_{\tau\mu}$ and the coupling product $g_{\mu\mu} g_{\tau\tau}$ could in principle be studied at future high-energy muon colliders [84–87] by measuring the process $\mu^- \mu^+ \to \tau^- \tau^+$.

3.2 Supernova

If the ALP couples to electrons or muons, it can be produced in the supernova environment and contribute to the cooling rate of the supernova core, thus allowing for deriving bounds on these couplings. For instance, ref. [88] obtained an upper bound on $g_{\mu\mu}$ at $10^{-8.1} \, \mathrm{GeV}^{-1}$ for $m_X < 1 \, \mathrm{MeV}$. Also, ref. [89] derived the limits on g_{ee} with SN1987A ranging from $\sim 2.5 \times 10^{-7} \, \mathrm{GeV}^{-1}$ to $\sim 8.8 \times 10^{-5} \, \mathrm{GeV}^{-1}$ for $m_X \lesssim 160 \, \mathrm{MeV}$. In our relevant mass range, we reproduce in figure 3 the SN1987A bounds on g_{ee} here.[2] The gray region is disfavored by SN1987A cooling bounds.

3.3 $\tau \to 3l$

The tightest bounds on the branching fractions for τ LFV decays into three charged leptons were obtained at the Belle experiment [58]. These limits, which are all of order 10^{-8}, are reproduced in table 2. We discuss the resulting bounds on $g_{\tau\alpha} \cdot g_{\alpha\alpha}$ ($\alpha = e, \mu$) in sections 4 and 5.

3.4 Lepton universality in $\tau \to l\nu\bar{\nu}$

When the ALP lifetime is long enough that the ALP rarely decays before traversing the whole detector, the decay $\tau \to l_\alpha X$ yields a signature similar to that of $\tau \to l\nu\bar{\nu}$, namely $\tau \to l + \text{missing}$. It is thus possible to place bounds on the off-diagonal coupling $g_{\tau\alpha}$ for long-lived ALPs. For the coupling $g_{\tau\mu}$, inspired by refs. [90, 91], we focus on the ratio of rates

$$R_{\mu e} = \frac{\Gamma_{\tau \to \mu\nu\bar{\nu}}}{\Gamma_{\tau \to e\nu\bar{\nu}}}, \tag{3.1}$$

---

[2]The coupling 'g_{ae}' in ref. [89] divided by $2m_e$ yields g_{ee} of this paper.

Decay modes	Upper bounds on BR [10^{-8}]
$\tau^- \to e^- e^+ e^-$	2.7
$\tau^- \to \mu^- \mu^+ \mu^-$	2.1
$\tau^- \to e^- \mu^+ \mu^-$	2.7
$\tau^- \to \mu^- e^+ e^-$	1.8
$\tau^- \to e^+ \mu^- \mu^-$	1.7
$\tau^- \to \mu^+ e^- e^-$	1.5

Table 2. Upper bounds on the branching fractions of LFV decays of the τ into three charged leptons, reproduced from ref. [58].

which has been computed accurately in the SM [92]:

$$R_{\mu e}^{\mathrm{SM}} = 0.972559 \pm 0.000005. \tag{3.2}$$

The BaBar collaboration measured this ratio to be [55]

$$R_{\mu e}^{\mathrm{BaBar}} = 0.9796 \pm 0.0039. \tag{3.3}$$

We define the discrepancy between $R_{\mu e}^{\mathrm{SM}}$ and $R_{\mu e}^{\mathrm{BaBar}}$,

$$\Delta R_{\mu e} \equiv R_{\mu e}^{\mathrm{BaBar}}/R_{\mu e}^{\mathrm{SM}} - 1 = 0.0072 \pm 0.0040. \tag{3.4}$$

We also define the theoretical value of this ratio in the presence of the ALP,

$$R^{\mathrm{SM}+X} = R_{\mu e}^{\mathrm{SM}} + \Gamma(\tau \to X\mu)/\Gamma_{\tau \to e\nu\bar\nu}^{\mathrm{SM}}, \tag{3.5}$$

where we use the experimental result $\Gamma_{\tau \to e\nu\bar\nu}^{\mathrm{exp}} = 4.04 \times 10^{-13}$ GeV [5] for $\Gamma_{\tau \to e\nu\bar\nu}^{\mathrm{SM}}$, and $\Gamma(\tau \to X\mu)$ is computed with eq. (2.3). We determine the 95% confidence-level (C.L.) bounds on $g_{\tau\mu}$ vs. m_X from the region that satisfies the following condition:

$$R_{\mu e}^{\mathrm{SM}+X}/R_{\mu e}^{\mathrm{SM}} - 1 < 0.0072 + 2 \times 0.0040, \tag{3.6}$$

where 0.0072 is the central value of $\Delta R_{\mu e}$ given in eq. (3.4) and 0.0040 is its uncertainty.

To derive bounds on $g_{\tau e}$ we follow a similar procedure with the ratios $R_{e\mu} = 1/R_{\mu e}$. In this case, since $R_{e\mu}^{\mathrm{SM}}$ is now larger than $R_{e\mu}^{\mathrm{BaBar}}$, we have

$$\Delta R_{e\mu} \equiv R_{e\mu}^{\mathrm{BaBar}}/R_{e\mu}^{\mathrm{SM}} - 1 = -0.0072 \pm 0.0040. \tag{3.7}$$

The negative value of $\Delta R_{e\mu}$ results in a particularly stringent bound on $\mathrm{BR}(\tau \to Xe)$, after requiring $R_{e\mu}^{\mathrm{SM}+X}/R_{e\mu}^{\mathrm{SM}} - 1 < -0.0072 + 2 \times 0.0040$. The resulting constraints are shown in figure 4. For the kinematically allowed mass range, the constraints on $\mathrm{BR}(\tau \to Xl_\alpha)$ are about 1.4×10^{-4} for Scenario 1 and 2.6×10^{-3} for the second scenario. The corresponding upper limits on $g_{\tau\alpha}$ are as strong as about 4×10^{-8} GeV^{-1} and 2×10^{-7} GeV^{-1} for the two scenarios, respectively, at the low mass regime. After the limits described in section 3.3,
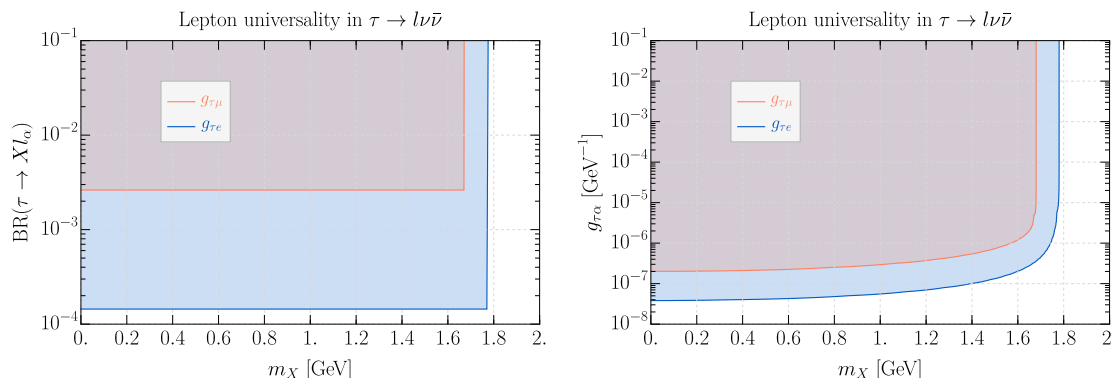
Figure 4. Constraints on $\mathrm{BR}(\tau \to Xl_\alpha)$ (left) and on $g_{\tau\alpha}$ (right) as a function of m_X, derived from the upper bounds on lepton-universality violation in $\tau \to l + \text{missing}$, for Scenarios 1 ($g_{\tau e} \neq 0$) and 2 ($g_{\tau\mu} \neq 0$). These limits are valid for long decay lengths of the X (small diagonal couplings g_{ee} or $g_{\mu\mu}$), so that it rarely decays inside the detector.

these are the tightest existing limits. We note that they are valid for small diagonal couplings, so that the ALP decays inside the detector only rarely.

It should be noted that the lepton-universality constraint will improve after the analysis of ref. [55] is performed with the large Belle II data set. However, since the results of ref. [55] are already limited by the systematics uncertainties, the improvement is expected to be modest.

We notice that ref. [93] has also recast the BaBar search results to the upper bounds on $\mathrm{BR}(\tau \to Xl_\alpha)$. In contrast to our treatment, which takes into account the experimental measurements of the ratios $R_{e\mu}$ and $R_{\mu e}$ of $\Gamma_{\tau \to \mu\nu\bar{\nu}}$ and $\Gamma_{\tau \to e\nu\bar{\nu}}$, the authors of ref. [93] worked on the partial decay widths of the two decay channels separately. As a result, our bound on $\mathrm{BR}(\tau \to Xe)$ is stronger than theirs by about one order of magnitude, while both works obtained similar constraints on $\mathrm{BR}(\tau \to X\mu)$.

3.5 $\tau \to l\gamma$

The ALP couplings to the charged leptons can also lead to the decays $\tau \to l_\alpha\gamma$ via either one or two LFV couplings. In the case of two LFV couplings, the processes must involve the coupling $g_{\mu e}$, which we set to zero throughout the paper. Therefore, we focus on the combinations of one LFV coupling and one diagonal coupling. The relevant coupling combinations are $(g_{\tau\tau}, g_{\tau e})$ and $(g_{\tau e}, g_{ee})$ for Scenario 1, and $(g_{\tau\tau}, g_{\tau\mu})$ and $(g_{\tau\mu}, g_{\mu\mu})$ for Scenario 2. Since we take $g_{\tau\tau} = 0$ in this paper, we derive only bounds on the coupling products $g_{\tau e} \cdot g_{ee}$ and $g_{\tau\mu} \cdot g_{\mu\mu}$ from the experimental upper bounds $\mathrm{BR}(\tau \to e\gamma) < 3.3 \times 10^{-8}$ and $\mathrm{BR}(\tau \to \mu\gamma) < 4.4 \times 10^{-8}$ [94], respectively.

Under these conditions, the ALP contributions to these decays consist of two categories: two-loop diagrams with an effective $g_{\gamma\gamma}^{\mathrm{eff}}$ coupling, and one-loop diagrams without this coupling. The former turns out to be dominant because of the logarithmic dependence on the ultraviolet cutoff. The decay width of $\tau \to l_\alpha\gamma$ is computed with the following analytic expression [36, 37]:

$$\Gamma(\tau \to l_\alpha\gamma) = \frac{m_\tau e^2}{8\pi}\Big(|\mathcal{F}_2^\alpha(0)|^2 + |\mathcal{G}_2^\alpha(0)|^2\Big), \text{ for } m_\tau \gg m_e, m_\mu, \tag{3.8}$$

– 10 –

where $\alpha = e, \mu$, and $\mathcal{F}_2^\alpha(0)$ and $\mathcal{G}_2^\alpha(0)$ are dimensionless form factors consisting of two parts that are naively linear and quadratic in the ALP-lepton couplings:

$$\mathcal{F}_2^\alpha(0) = \mathcal{F}_2^\alpha(0)_{\text{lin.}} + \mathcal{F}_2^\alpha(0)_{\text{quad.}},$$
$$\mathcal{G}_2^\alpha(0) = \mathcal{G}_2^\alpha(0)_{\text{lin.}} + \mathcal{G}_2^\alpha(0)_{\text{quad.}}, \tag{3.9}$$

where

$$\mathcal{F}_2^\alpha(0)_{\text{lin.}} = -\frac{e^2 m_\tau^2}{8\pi^2} g_{\tau\alpha} g_{\gamma\gamma}^{\text{eff}} g_\gamma(x_\tau), \tag{3.10}$$

$$\mathcal{G}_2^\alpha(0)_{\text{lin.}} = -\frac{e^2 m_\tau^2}{8\pi^2} g_{\tau\alpha} g_{\gamma\gamma}^{\text{eff}} g_\gamma(x_\tau), \tag{3.11}$$

$$\mathcal{F}_2^\alpha(0)_{\text{quad.}} = -\frac{m_\tau}{16\pi^2} g_{\tau\alpha} g_{\alpha\alpha} m_\alpha g_2(x_\tau),$$

$$\mathcal{G}_2^\alpha(0)_{\text{quad.}} = +\frac{m_\tau}{16\pi^2} g_{\tau\alpha} g_{\alpha\alpha} m_\alpha g_2(x_\tau), \tag{3.12}$$

where $x_\tau = m_X^2/m_\tau^2 - i\epsilon$ with $\epsilon \to 0^+$ and

$$g_\gamma(x) = 2\log\frac{\Lambda^2}{m_X^2} - \frac{\log x}{x-1} - (x-1)\log\frac{x}{x-1} - 2, \tag{3.13}$$

$$g_2(x) = 1 - 2x + 2(x-1)x\log\frac{x}{x-1}. \tag{3.14}$$

We have assumed $g_{e\mu} = g_{\mu e} = g_{\tau\tau} = 0$, and confined ourselves only to the couplings that are turned on in Scenarios 1 and 2 for simplicity. These couplings, if non-zero, could also lead to $\tau \to l_\alpha \gamma$ decays with two ALP-lepton couplings. For full and complete formulas, see refs. [36, 37].

The resulting bounds on $g_{\tau e} g_{ee}$ and $g_{\tau\mu} g_{\mu\mu}$ are roughly $10^{-6}\,\text{GeV}^{-2}$ for m_X between 0 and $2\,\text{GeV}$. Note that we have fixed $\Lambda = 1\,\text{TeV}$ for the logarithmic dependency (see eq. (3.13)).

3.6 Leptonic decays of the muon

Multiple processes of muon decays into electrons or photons can be induced by the ALP couplings. For instance, the decay $\mu \to e\gamma$ [28, 29], which is extremely rare in the SM, can be mediated either by $g_{\mu e}$ together with g_{ee} or $g_{\mu\mu}$, or by the two LFV couplings $g_{\tau\mu}$ and $g_{\tau e}$. The experimental limits on $\text{BR}(\mu \to e\gamma)$ [30] place very strong bounds on these coupling products. Since in this work we take $g_{\mu e} = 0$ and do not switch on $g_{\tau\mu}$ and $g_{\tau e}$ simultaneously, both of these two contributions vanish (see eqs. (11–14) of ref. [37]).

Similarly, the processes $\mu \to 3e$, $\mu \to e\gamma\gamma$, and $\mu \to e + \text{missing}$ can all be used to derive bounds on the coupling $g_{\mu e}$ (and g_{ee}). However, since we have restricted ourselves to the case of vanishing $g_{\mu e}$, we do not study these processes in detail.

3.7 Leptonic $g - 2$ anomalies

One of the most sensitive avenues for testing the SM is the leptonic anomalous magnetic moments. The long-standing discrepancy in the muon $g - 2$ between the SM predictions and experimental measurements has been an important drive for searching for new physics.

The latest determination of the fine-structure constant $\alpha_{\rm EM}$ makes the electron anomalous magnetic moment consistent with the SM prediction to within 1.6σ [95]:

$$\Delta a_e = a_e^{\rm exp} - a_e^{\rm SM} = (4.8 \pm 3.0) \times 10^{-13}. \tag{3.15}$$

On the other hand, the recently published result of the Fermilab-based Muon $g-2$ experiment [96], when combined with that of the E821 collaboration at BNL [97], has an excess of about 4.2σ over the latest SM theoretical computation [98]:

$$\Delta a_\mu \equiv a_\mu^{\rm exp} - a_\mu^{\rm SM} = (25.1 \pm 5.9) \times 10^{-10}. \tag{3.16}$$

The discrepancy may be explained by an ALP [37, 99, 100]. In principle, both diagonal and off-diagonal couplings could lead to new contributions to $a_\alpha = (g-2)_\alpha/2$ ($\alpha = e, \mu$). However, since in the present work we keep $g_{\mu e}$ switched off and also take equal scalar and pseudoscalar couplings, the LFV couplings contributions vanish [37]. The diagonal coupling contributions to a_l are

$$(\Delta a_{l_\alpha})_{\rm diag.} = -\frac{m_\alpha^2}{4\pi^2} \left[16\pi\, \alpha_{\rm em}\, g_{\gamma\gamma}^{\rm eff}\, g_{\alpha\alpha} \left(\log\frac{\Lambda^2}{m_\alpha^2} - h_2(x_\alpha) \right) + |g_{\alpha\alpha}|^2 h_1(x_\alpha) \right], \tag{3.17}$$

where $x_\alpha = m_X^2/m_\alpha^2$, $g_{\gamma\gamma}^{\rm eff}$ can be computed by eq. (2.7), and

$$h_1(x) = 1 + 2x - (x-1)x\log x + 2x(x-3)\sqrt{\frac{x}{x-4}} \log\left(\frac{\sqrt{x} + \sqrt{x-4}}{2} \right), \tag{3.18}$$

$$h_2(x) = 1 + \frac{x^2}{6}\log x - \frac{x}{3} - \frac{x+2}{3}\sqrt{(x-4)x} \log\left(\frac{\sqrt{x} + \sqrt{x-4}}{2} \right). \tag{3.19}$$

In eq. (3.17), the second term, $-\frac{m_\alpha^2}{4\pi^2}|g_{\alpha\alpha}|^2 h_1(x_\alpha)$, is strictly negative. Therefore, in order to explain the observed discrepancy, the first term in eq. (3.17) must be positive and large enough to overcome the negative second term. We find that for the electron $g-2$ case, the leptophilic axions considered in Scenario 1 can bring the theoretical prediction to reach the 95% lower limit of Δa_e but not up to the central value. On the other hand, for the muon $g-2$ discrepancy the ALPs in Scenario 2 cannot even bring the theoretical prediction to the edge of the 95% lower limit of the experimental value (cf. eq. (3.16)). For these reasons, we do not show their corresponding limits in the sensitivity plots in section 5. Lastly, we note that a very recent lattice calculation of hadronc contributions to the muon $g-2$ pushed the SM prediction to within 1σ of the experimental result [101].

3.8 Beam-dump experiments and dark-photon searches

In our m_X range of interest, results from the E137 and E141 beam-dump experiments at SLAC [102], and similar ones at CHARM [103], Orsay [104], and Fermilab [105], typically constrain the coupling g_{ee} to be outside the range $10^{-5}\,{\rm GeV}^{-1}$ to $1\,{\rm GeV}^{-1}$ when taking the axion to radiate off the electron beam [46, 47]. Moreover, assuming universal diagonal couplings $g_{ee} = g_{\mu\mu} = g_{\tau\tau}$, a dark-photon search performed at BaBar [21, 48] can be recast into bounds on these couplings [21]. The limits obtained exclude g_{ee} above $\sim 6 \times 10^{-3}\,{\rm GeV}^{-1}$
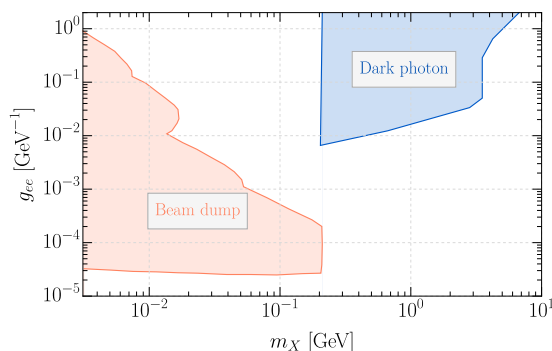
Figure 5. Constraints on g_{ee} as a function of m_X, reproduced from ref. [21]. Note that the BaBar dark photon bound is valid for universal diagonal couplings $g_{ee} = g_{\mu\mu} = g_{\tau\tau}$.

for m_X between about $210\,\mathrm{MeV}$ and $8\,\mathrm{GeV}$.[3] We extract these results from ref. [21] and show them in figure 5.

3.9 Muonium-antimuonium oscillations and $\mu^- \to e^-$ conversion

The bound state of e^- and μ^+ is known as muonium M. In the presence of lepton flavor violation, the muoniums may oscillate into antimuoniums $\bar{M}$ ($e^+\mu^-$). The strongest bound on the probability of spontaneous muonium to antimuonium conversion, $P_{M\bar{M}} < 8.3 \times 10^{-11}$ at 90% C.L., was obtained by the MACS experiment in 1999 [106]. For the leptophilic ALPs, only the coupling $g_{\mu e}$ could result in such an oscillation. Since we keep this coupling off in this work, we do not go into the details of the derivation of the bound.

Similarly, $\mu^- \to e^-$ conversion in nuclei is also a very sensitive probe of LFV [107]. In the context of the leptophilic ALPs, this process also only concerns the coupling $g_{\mu e}$, and is not studied here. Bounds derived from past experiments [108] and those projected for future experiments [109, 110] can be found in ref. [37].

4 Prompt and displaced searches at Belle II

In this section we discuss our recast of the Belle prompt LFV search, projected to the Belle II sensitivity, as well as the displaced search that we propose to conduct. Our results are then shown in section 5.

Belle II is an ongoing B-factory experiment hosted at the KEK laboratory in Tsukuba, Japan. It collides an electron beam of energy $E_{e^-} = 7\,\mathrm{GeV}$ with a positron beam of $E_{e^+} = 4\,\mathrm{GeV}$, reaching a center-of-mass energy of $10.58\,\mathrm{GeV}$. The experiment has so far collected about $200\,\mathrm{fb}^{-1}$, less than the samples collected by the BaBar and Belle experiments, approximately $500\,\mathrm{fb}^{-1}$ and $1\,\mathrm{ab}^{-1}$, respectively. However, Belle II is planned to eventually collect a sample of $50\,\mathrm{ab}^{-1}$. With a cross section $\sigma(e^+e^- \to \tau^+\tau^-) \approx 0.92$ nb, this implies 4.6×10^{10} τ-pair events produced.

---

[3]The parameter $c_{ee}^{\mathrm{eff}}/\Lambda$ used in ref. [21] is related to g_{ee} of the present work by a factor 2: $c_{ee}^{\mathrm{eff}}/\Lambda = 2g_{ee}$.

Decay modes	Baseline efficiency
$\tau \to Xe, X \to e^- e^+$	6.0 %
$\tau \to X\mu, X \to \mu^- \mu^+$	7.6 %

Table 3. Baseline efficiencies reproduced from ref. [58].

4.1 Recast of the prompt search

We first recast the Belle prompt search for lepton-flavor-violating τ decays into three charged leptons, which used a sample of 7.19×10^8 τ-pair events [58].

The background level reported in ref. [58] ranges from 0.01 ± 0.01 events in the $\tau^- \to \mu^+ e^- e^-$ channel to 0.20 ± 0.15 events in the $\tau^- \to e^+ e^- e^-$ channel. We assume that the number of background events will remain small even with the much larger sample of Belle II, sufficiently so that it can be ignored in our Belle II projection.

The event-selection efficiency reported in ref. [58] for each final state is reproduced in table 3. We refer to this as the "baseline efficiency". In addition, since the ALP in the region of sensitivity can be long-lived, its daughter leptons may escape detection in the analysis optimized for prompt decays, reducing the efficiency. We use events simulated with Pythia8.245 [111] and without detector simulation to estimate the efficiencies as functions of the ALP lifetime and mass. While the beams at Belle II collide with a small crossing angle, we simplify this and simulate head-on collisions, with the electron beam momentum in the $+z$ direction.

As required in ref. [58], we require the transverse and longitudinal impact parameters of both charged leptons produced in the ALP decay to satisfy $d_0 < 5\,\text{mm}$ and $z_0 < 30\,\text{mm}$, respectively. The computation of z_0 is performed with

$$z_0 = \left| z - \frac{(x \cdot p_x + y \cdot p_y) \cdot p_z}{p_T^2} \right|, \tag{4.1}$$

where x, y, and z are the production position coordinates of the charged lepton with respect to the IP, p_x, p_y, and p_z are the components of the lepton momentum in the three spatial directions, and $p_T^2 = p_x^2 + p_y^2$ is its transverse momentum squared.

The calculation of the transverse impact parameter is more involved. For straight tracks, it can be approximated as

$$d_0^{\text{naive}} = \frac{|p_x \cdot y - p_y \cdot x|}{p_T}. \tag{4.2}$$

However, given the relatively low momenta and possibly large flight distances, the track curvature resulting from the local magnetic field can be non-negligible. We estimate the radius of the track R as

$$R \simeq \frac{1}{0.3} \frac{p_T}{\text{GeV}} \frac{\text{T}}{B} \text{ meters}, \tag{4.3}$$

where T stands for Tesla and $B = 1.5\,\text{T}$ is the strength of the magnetic field. The true transverse impact parameter d_0 is then computed with

$$d_0 = \sqrt{(R + d_0^{\text{naive}})^2 + r^2 - (d_0^{\text{naive}})^2} - R, \tag{4.4}$$

where r labels the transverse decay position of the ALP from the interaction point (IP).

Furthermore, we require $r < 10\,\mathrm{cm}$. This requirement does not appear explicitly in ref. [58], but we apply it in order to obtain a more realistic simulation of the minimal number of detector hits needed for track reconstruction. This requirement is important only for very light ALPs, which are highly boosted and can pass the d_0 and z_0 requirements even when decaying far from the IP. It is redundant for heavy ALPs, whose daughters are separated by a large angle and would anyway fail the d_0 and z_0 cuts.

4.2 Proposal of a displaced-vertex search

Since the prompt search is not optimized for detection of a long-lived ALP, we propose to complement it with a displaced search, in which the maximal d_0 and z_0 requirements of the prompt analysis are replaced with a minimal requirement on the ALP flight distance, $r > 1\,\mathrm{cm}$. Since this displaced-vertex requirement leads to a large reduction in background (see, e.g., ref. [61]), the background in this search is even smaller than the negligible background in the prompt search.

Given the strong background suppression provided by the displaced vertex and the sub-event background level of the prompt Belle analysis [58], we expect that many of the cuts applied in the prompt analysis can be avoided in the displaced one. For example, based on the efficiencies reported in ref. [58], removing the missing-momentum cut and applying lepton identification criteria on only one of the leptons increases the baseline efficiency by a factor of 1.5 in scenario 1 and 2.3 in scenario 2. These efficiency enhancements relative to the prompt analysis are included into the final sensitivity estimates for the displaced search reported in section 5.

In addition to the baseline efficiency, we determine the detection efficiency variation as a function of the ALP decay position following the procedure used in refs. [63, 72]. We require the ALP to decay inside the fiducial volume of the tracker, defined by the transverse and longitudinal distances from the IP: $1\,\mathrm{cm} < r < 80\,\mathrm{cm}$ and $-40\,\mathrm{cm} < z < 120\,\mathrm{cm}$, respectively. In addition, we account for the tracking efficiency decrease with the transverse distance from the IP. This so-called "displaced-tracking efficiency" $\epsilon^{\mathrm{track}}$ is parameterized as a linear function of r, so that $\epsilon^{\mathrm{track}} = 1$ for $r = 1\,\mathrm{cm}$ and $\epsilon^{\mathrm{track}} = 0$ at $r = 80\,\mathrm{cm}$. In particular, in the large decay length limit, $\epsilon^{\mathrm{track}}$ is effectively reduced to an overall factor of 50% on average, for decays inside the fiducial volume.

4.3 Computation and simulation

The number of signal events at Belle II is computed with:

$$N_S^{\mathrm{Belle\ II}} = 2 \cdot N_{\tau^-\tau^+} \cdot \mathrm{BR}(\tau \to 1\ \mathrm{prong})$$
$$\cdot \mathrm{BR}(\tau \to X l_\alpha) \cdot \epsilon \cdot \mathrm{BR}(X \to l_\alpha^- l_\alpha^+), \qquad (4.5)$$

where $l_\alpha = e$ or μ, $N_{\tau^-\tau^+} = 4.6 \times 10^{10}$ is the total number of τ pair events produced at Belle II, ϵ denotes the final efficiency of event selections as detailed above, and $\mathrm{BR}(\tau \to 1\ \mathrm{prong}) \approx 85\%$ is the SM decay branching ratio of the τ lepton into a single charged particle and neutrinos. Application of this factor assumes that the events are identified as $e^+e^- \to \tau^+\tau^-$ by requiring the topology of one track recoiling against the three leptons

from the signal decay chain. The factor 2 is due to the fact that in each signal event there are two τ leptons. For vanishing background, $N_S^{\text{Belle II}} = 3$ corresponds to the 95% C.L. exclusion limits.

The simulations are performed on a grid in the model parameters. For each grid point, we simulate the process $e^- e^+ \to \tau^- \tau^+$. The τ leptons decay to $X + e$ or $X + \mu$, and the ALP decays to $e^- e^+$ or $\mu^- \mu^+$, respectively. Pythia automatically processes the decay positions of the simulated ALPs according to the exponential decay distributions using the ALP lifetime and boost factor. Then the prompt and displaced sets of event selection criteria are imposed at the generator level to obtain the efficiency separately for each of the two search strategies. Thus, we obtain the efficiencies for the two strategies at each grid point.

5 Numerical results

We now proceed to present numerical results in this section. We concentrate on the two representative benchmark scenarios introduced in section 2 for detailed numerical studies in sections 5.1 and 5.2, while for other further possibilities we will provide a qualitative discussion in section 5.3. Since, as explained in section 4, for both prompt and displaced searches we expect zero background events, we use 3-signal-event isocurves to show exclusion limits at 95% C.L.

5.1 Scenario 1: $g_{\tau e}$ and g_{ee}

In Scenario 1, the couplings $g_{\tau e}$ and g_{ee} are turned on, mediating the production and decay of the ALP, respectively, while the other couplings are all considered as vanishing. Thus, the signal events are $\tau \to Xe$ with $X \to e^- e^+$.

We show in figure 6 the cut-specific event selection efficiencies for the signal events as functions of $c\tau_X$ for the prompt and displaced analysis and for three benchmark masses $m_X = 0.005$ (blue), 0.5 (red), and 1.5 (black) GeV.

Figure 7 presents the estimated 95% C.L. sensitivities at Belle II for the prompt and displaced searches. The upper plot shows the projected limits on the branching fraction product $\text{BR}(\tau \to Xe) \cdot \text{BR}(X \to e^+ e^-)$ as a function of $c\tau_X$ for the three benchmark ALP masses. Since both searches have no background, their maximal sensitivities are approximately equal, with the prompt (displaced) search being more sensitive at shorter (longer) ALP lifetimes. Focusing our attention on the displaced search, which is proposed here for the first time, we observe that it is up to about 40 times more sensitive to the branching-fraction product than the prompt search at large values of $c\tau_X$. This translates to about a factor of 6 in the sensitivity to $g_{\tau e}$ for a given value of g_{ee}, as seen in the lower-left plot. We note that the next-tightest constraint, which comes from lepton universality in $\tau \to l\nu\bar{\nu}$ (see section 3.4), yields $g_{\tau e} \lesssim 10^{-7}\,\text{GeV}^{-1}$ (which is valid only for long-lived ALPs and hence small g_{ee} coupling). Since this is much weaker than the sensitivity reach of the prompt and displaced searches, it is not shown on the same plot.

In the lower-right plot we study the particular case $g_{\tau e} = g_{ee}$, and present the limits on the coupling value vs. m_X for both Belle II and Belle luminosities, together with bounds from existing measurements. Since the $\tau \to l\nu\bar{\nu}$ lepton-universality bound is only valid for
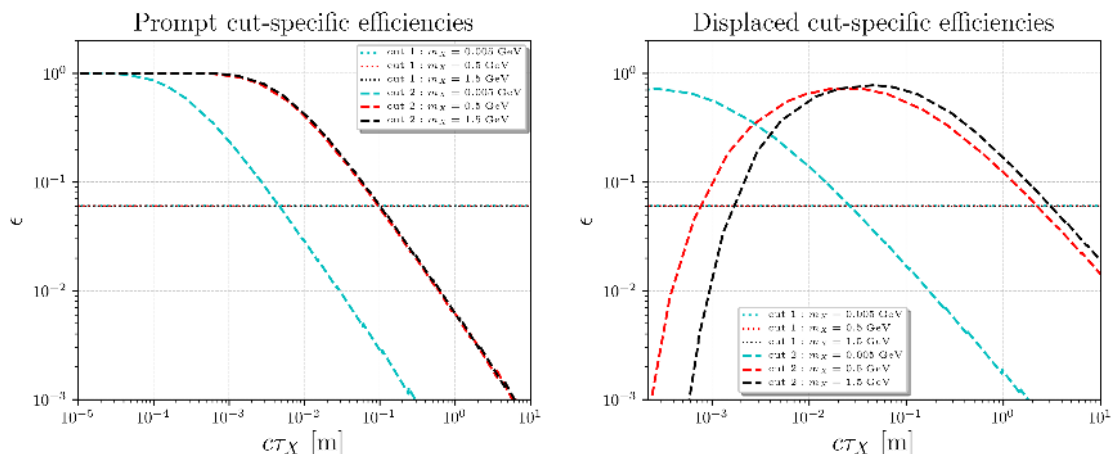
Figure 6. Efficiencies vs. $c\tau_X$ for the signal events of Scenario 1, with three benchmark values of m_X: 0.005, 0.5, 1.5 GeV. *Left*: prompt cut-specific efficiencies. *Right*: displaced cut-specific efficiencies. "Cut 1" denotes the baseline efficiency. For the prompt search, "cut 2" is the efficiency of requiring $d_0 < 5\,\mathrm{mm}$, $z_0 < 30\,\mathrm{mm}$, and $r < 10\,\mathrm{cm}$. For the displaced search, "cut 2" gives the efficiency of applying the fiducial-volume and the displaced-tracking requirements.

long-lived ALPs, we show the corresponding excluded parameter space only for $c\tau_X \gtrsim 1$ m. In this case, the tightest bound in the case of small couplings is obtained from the lepton-universality constraint. However, the prompt $\tau \to 3\ell$ and the displaced search proposed here are uniquely sensitive to a mass-dependent range of coupling values between about 2×10^{-5} and $2 \times 10^{-3}\,\mathrm{GeV}^{-1}$.
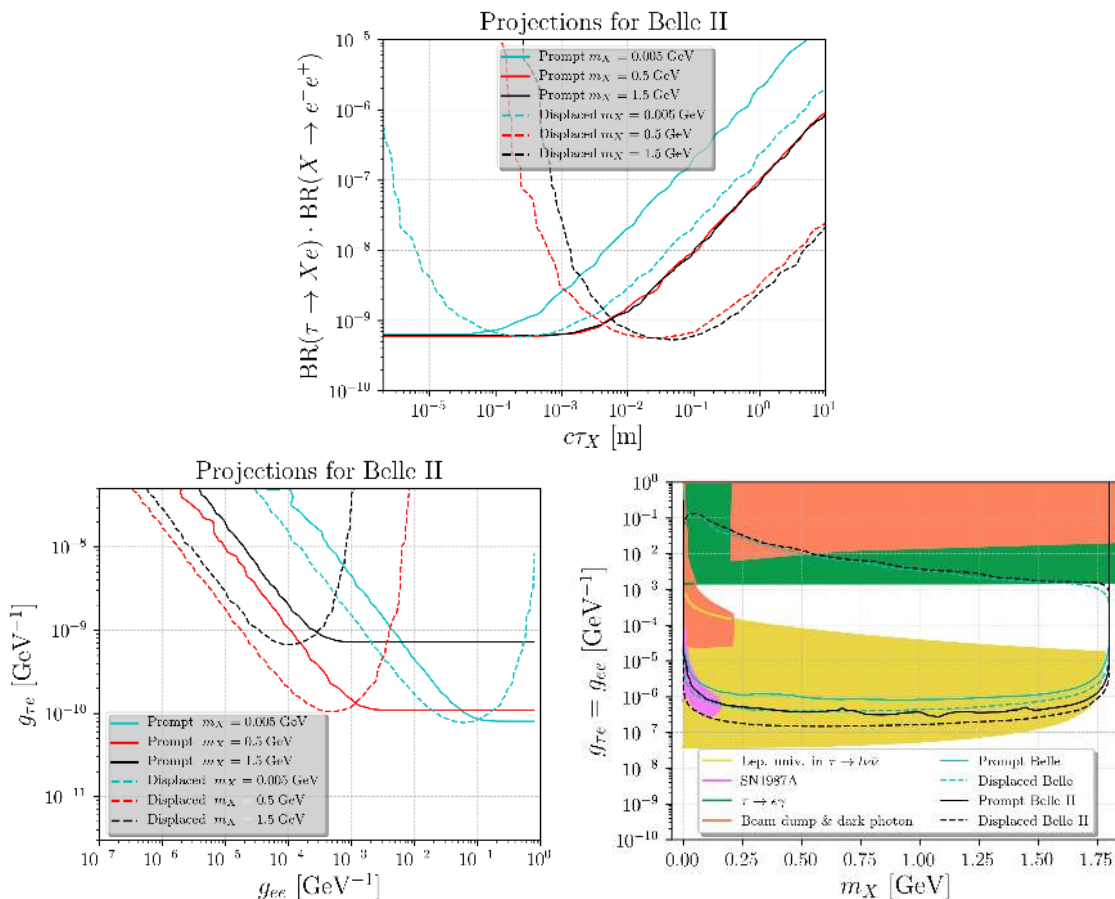
We note that our "model-independent" limits on the branching-fraction product can be used to obtain constraints on other models with similar scattering and decay topologies. Examples include a light new gauge boson Z' [90, 112, 113] or a light CP-even scalar [114–116].

5.2 Scenario 2: $g_{\tau\mu}$ and $g_{\mu\mu}$

In Scenario 2, the non-zero couplings $g_{\tau\mu}$ as well as $g_{\mu\mu}$ are responsible for the production and decay of the ALPs, respectively. The sensitive mass range is now between $2\,m_\mu$ and $m_\tau - m_\mu$.

In figure 8 we show the signal event selection efficiencies as functions of $c\tau_X$, for $m_X = 0.5$, 1.0, and 1.5 GeV. These plots are similar to those of figure 6, except for some differences that mainly stem from the different mass choices, as well as different baseline efficiencies.

We present the final sensitivity estimates for Scenario 2 in figure 9 with the same format as in figure 7. Due to the similar efficiencies of the two scenarios, the branching-fraction bounds are similar. However, the $g_{\mu\mu}$ sensitivity of scenario 2 is significantly enhanced, mainly because the width $\Gamma(X \to ll)$ is proportional to the lepton mass squared. As a result, in the case $g = g_{\tau\mu} = g_{\mu\mu}$ shown in the bottom-right plot, the displaced search can probe g values that are smaller by more than a factor of 10 than those of the lepton-universality constraint. The measurement of $\Gamma(\tau \to \mu\nu\bar{\nu})/\Gamma(\tau \to e\nu\nu)$ is larger than the SM expectation. Therefore, the presence of $g_{\tau\mu}$ and $g_{\mu\mu}$ couplings will bring the prediction
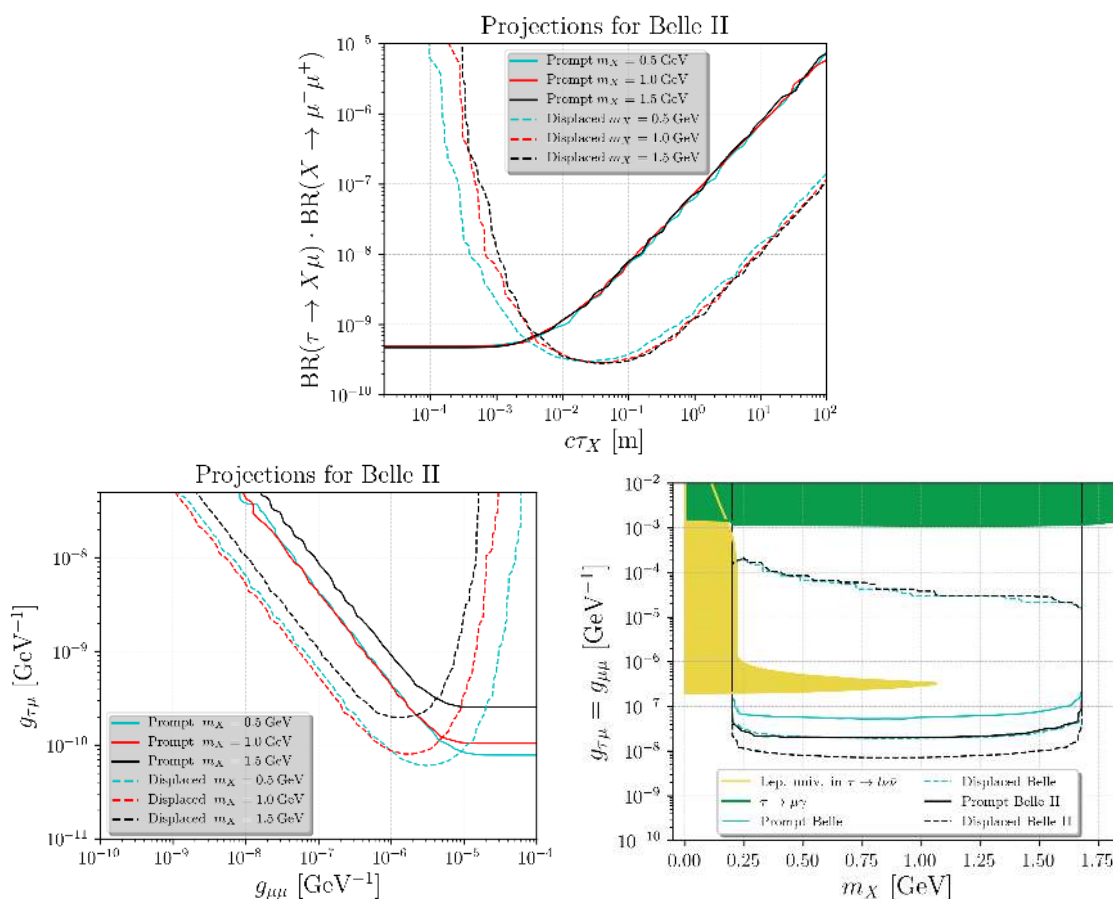
Figure 7. Sensitivity results for Scenario 1 at 95% C.L. *Top*: model-independent limits on the branching-fraction product $\mathrm{BR}(\tau \to Xe) \cdot \mathrm{BR}(X \to e^- e^+)$ vs. $c\tau_X$ for various ALP masses. *Bottom left*: model-dependent bounds from Belle and Belle II on $g_{\tau e}$ vs. g_{ee} for the prompt and displaced searches at Belle II for various ALP masses. *Bottom right*: bounds on the couplings vs. ALP mass in the specific case $g_{\tau e} = g_{ee}$, compared with bounds from lepton universality in $\tau \to l\nu\bar{\nu}$ [55, 92], $\tau \to e\gamma$ [94], SN1987A [89], beam-dump experiments [46, 47], and a dark-photon search [21, 48]. Note that the lepton-universality constraint is only valid for long-lived ALPs, resulting in the upper cutoff of the corresponding region.

closer to the central value of the measurement, so that the valid parameter space satisfies $R_{\mu e}^{SM+X}/R_{\mu e}^{SM} - 1 < 0.0072 + 2 \times 0.004$. In contrast to the case of Scenario 2, the presence of $g_{\tau e}, g_{ee}$ in Scenario 1 will bring the prediction further away from the central value of the measurement, thus resulting in a more stringent constraint. The valid parameter space has to satisfy $R_{e\mu}^{SM+X}/R_{e\mu}^{SM} - 1 < -0.0072 + 2 \times 0.004$. It implies that the region excluded in Scenario 1 is much larger than that of Scenario 2.

We note that, while we focused on $X \to \mu^+\mu^-$ decays that take place inside the tracker fiducial volume, the muon tracks can also be found when the decay takes place in the muon detector, which extends out to a distance of about 3.5 m from the IP. Naively, the somewhat larger decay distance allows for a somewhat stronger reach. However, since the muon detector has no magnetic field, it is impossible to determine the momenta of the
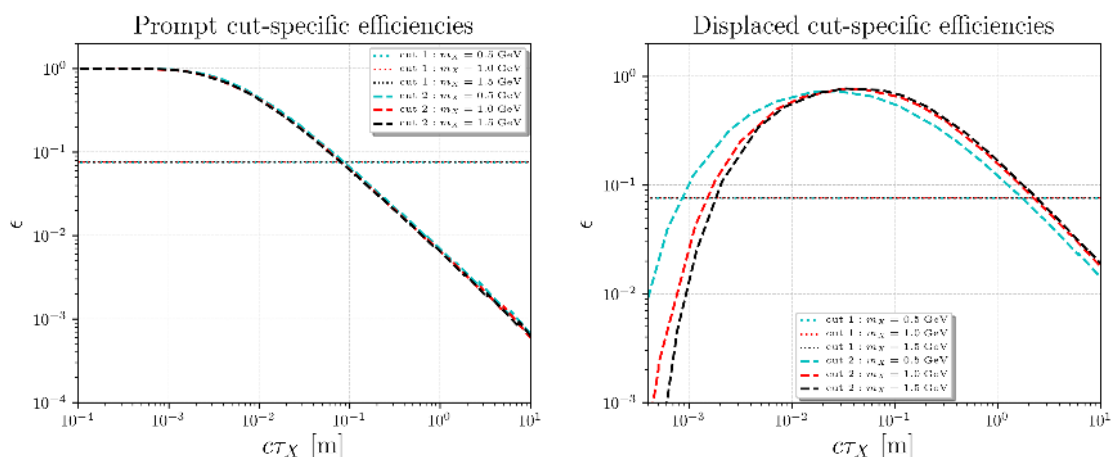
Figure 8. Efficiencies for Scenario 2 ($g_{\tau\mu}$ and $g_{\mu\mu}$ are non-zero). The format is the same as in figure 6, except for the different set of mass choices: $m_X = 0.5$, 1.0, and 1.5 GeV.

Figure 9. Sensitivity results for Scenario 2 ($g_{\tau\mu}$ and $g_{\mu\mu}$ are non-zero). The format is the same as in figure 7, except for the different mass choices.

muons in this case. This would likely result in high background levels that would make this approach less sensitive.

5.3 Other scenarios

We discuss briefly some other coupling scenarios.

One possibility is to take the combinations $(g_{\tau e}, g_{\mu\mu})$ or $(g_{\tau\mu}, g_{ee})$ to be non-zero. However, these results would be rather similar to those of Scenario 2 and Scenario 1, respectively, except at the edges of phase space. This is because the production of the ALP is dominantly proportional to m_τ^3, and is not affected by the light lepton or X masses except at the kinematic threshold (cf. eq. (2.3)). Another possibility is to take both $g_{\tau e} = g_{\tau\mu}$ and $g_{ee} = g_{\mu\mu}$ to be non-vanishing. In this case the ALP decay is dominated by the di-muon channel, because $m_\mu \gg m_e$, and the production rate is approximately doubled compared to that in Scenarios 1 and 2 with the same LFV coupling values. Therefore, the end results are similar to those of Scenario 2 presented here, except for a factor of 2 in the production rate. It is nevertheless worth mentioning that LFV muon decays such as $\mu \to e\gamma$ can place strong constraints on this scenario.

We also consider the possibility of a non-zero $g_{e\mu}$. According to eqs. (2.5) and (2.4), equal values of $g_{e\mu}$ and $g_{\mu\mu}$ lead to equal widths for $X \to e^\pm \mu^\mp$ and $X \to \mu^+\mu^-$ except near the kinematic thresholds. Therefore, we expect the sensitivity for $g_{e\mu}$, together with, e.g., non-vanishing $g_{\tau\mu}$, to be similar to that obtained for Scenario 2. We note that LFV muon-decay processes such as $\mu \to e+$ missing can strongly constrain $g_{e\mu}$ for $m_X < m_\mu$.

Finally, if the coupling $g_{\tau\tau}$ is non-zero, it would lead to ALP-strahlung production of the ALP by radiation off the τ (see ref. [117] for an experimental search at BaBar with this production mechanism) and would loop-induce the decay $X \to \gamma\gamma$ for the m_X range considered in this work.

6 Conclusions

In this paper we investigated the sensitivities of the B-factory experiment Belle II to a leptophilic axion-like particle X with lepton-flavor-violating (LFV) couplings. The decay channels of interest are $\tau \to Xl$, where $l = e, \mu$, with the ALP undergoing the decay $X \to l^+l^-$. The ALP can also decay to a pair of photons via a lepton loop. We do not consider this decay as a search signature, but we do account for it when calculating the signal branching fractions and ALP lifetime. Depending on the ALP lifetime, its decay may be effectively prompt or may take place visibly away from the interaction point. Correspondingly, we studied the sensitivity of both a prompt search and a displaced-vertex search.

In the case of the prompt search, we recast the Belle search for τ decays into three charged leptons [58] to derive constraints on the couplings in the scenarios of interest. The displaced search involves identifying the displaced vertex originating from the long-lived ALP decay. In both cases we used simulated events to determine the efficiency, accounting for the size of the Belle II tracker and for the dependence of the efficiency on the ALP decay position.

We focused on two benchmark scenarios for detailed numerical results: Scenario 1 ($g_{\tau e}$ and g_{ee} are non-zero) and Scenario 2 ($g_{\tau \mu}$ and $g_{\mu\mu}$ are non-zero), taking all other couplings to be vanishing. The signatures for the two scenarios are, respectively, $\tau \to Xe, X \to e^- e^+$ and $\tau \to X\mu, X \to \mu^- \mu^+$. Since the prompt Belle search was background-free [58], we expect this to also be the case at Belle II, particularly with the displaced search. Therefore, we present 3-signal-event isocurves to indicate the Belle II sensitivity at 95% C.L. exclusion limits. As a function of the ALP mass, we show both model-independent limits on the branching-fraction product $\mathrm{BR}(\tau \to Xe) \cdot \mathrm{BR}(X \to e^- e^+)$ vs. $c\tau_X$ and model-dependent limits on the ALP couplings. Generally, the displaced-vertex (prompt) search shows better sensitivity for small (large) values of $g_{\alpha\alpha}$, corresponding to long (short) decay lengths of the ALP. For long decay lengths, the displaced-vertex search can extend the prompt search's sensitivity to the branching-fraction product by a factor of about 40, corresponding to about a factor of 6 in $g_{\tau l}$.

The model-dependent results are also presented, for the case of equal production and decay couplings, in the plane $g = g_{\tau\alpha} = g_{\alpha\alpha}$ vs. m_X. In this case, in Scenario 1, the proposed displaced search is about as sensitive as constraints obtained from lepton-flavor universality tests in $\tau \to l\nu\bar{\nu}$. However, in Scenario 2, the displaced search is sensitive to coupling values that are more than an order of magnitude smaller than those of the lepton-universality constraints.

Generally, we find that Belle II can probe the LFV coupling $g_{\tau\alpha}$ down to below $10^{-10}\,\mathrm{GeV}^{-1}$, with the best sensitivity reached for a different value of $g_{\alpha\alpha}$ in each scenario.

Acknowledgments

References

[1] nEDM collaboration, *Measurement of the permanent electric dipole moment of the neutron*, *Phys. Rev. Lett.* **124** (2020) 081803 [`arXiv:2001.11966`] [INSPIRE].

[2] J. Dragos, T. Luu, A. Shindler, J. de Vries and A. Yousif, *Confirming the Existence of the strong CP Problem in Lattice QCD with the Gradient Flow*, *Phys. Rev. C* **103** (2021) 015202 [`arXiv:1902.03254`] [INSPIRE].

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

[4] R.D. Peccei, *The Strong CP problem and axions*, *Lect. Notes Phys.* **741** (2008) 3 [`hep-ph/0607268`] [INSPIRE].

[5] PARTICLE DATA GROUP collaboration, *Review of Particle Physics*, *PTEP* **2020** (2020) 083C01 [INSPIRE].

[6] J.L. Feng, T. Moroi, H. Murayama and E. Schnapka, *Third generation familons, b factories, and neutrino cosmology*, *Phys. Rev. D* **57** (1998) 5875 [`hep-ph/9709411`] [INSPIRE].

[7] J.E. Kim and D.J.E. Marsh, *An ultralight pseudoscalar boson*, *Phys. Rev. D* **93** (2016) 025027 [`arXiv:1510.01701`] [INSPIRE].

[8] I. De Martino, T. Broadhurst, S.H. Henry Tye, T. Chiueh, H.-Y. Schive and R. Lazkoz, *Recognizing Axionic Dark Matter by Compton and de Broglie Scale Modulation of Pulsar Timing*, *Phys. Rev. Lett.* **119** (2017) 221103 [`arXiv:1705.04367`] [INSPIRE].

[9] V.A. Rubakov, *Grand unification and heavy axion*, *JETP Lett.* **65** (1997) 621 [`hep-ph/9703409`] [INSPIRE].

[10] M. Dine and W. Fischler, *The Not So Harmless Axion*, *Phys. Lett. B* **120** (1983) 137 [INSPIRE].

[11] L.F. Abbott and P. Sikivie, *A Cosmological Bound on the Invisible Axion*, *Phys. Lett. B* **120** (1983) 133 [INSPIRE].

[12] J. Preskill, M.B. Wise and F. Wilczek, *Cosmology of the Invisible Axion*, *Phys. Lett. B* **120** (1983) 127 [INSPIRE].

[13] D.J.E. Marsh, *Axion Cosmology*, *Phys. Rept.* **643** (2016) 1 [`arXiv:1510.07633`] [INSPIRE].

[14] G. Lambiase and S. Mohanty, *Hydrogen spin oscillations in a background of axions and the* 21*n-cm brightness temperature*, *Mon. Not. Roy. Astron. Soc.* **494** (2020) 5961 [`arXiv:1804.05318`] [INSPIRE].

[15] A. Auriol, S. Davidson and G. Raffelt, *Axion absorption and the spin temperature of primordial hydrogen*, *Phys. Rev. D* **99** (2019) 023013 [`arXiv:1808.09456`] [INSPIRE].

[16] N. Houston, C. Li, T. Li, Q. Yang and X. Zhang, *Natural Explanation for* 21 *cm Absorption Signals via Axion-Induced Cooling*, *Phys. Rev. Lett.* **121** (2018) 111301 [`arXiv:1805.04426`] [INSPIRE].

[17] 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].

[18] I. Brivio et al., *ALPs Effective Field Theory and Collider Signatures*, *Eur. Phys. J. C* **77** (2017) 572 [`arXiv:1701.05379`] [INSPIRE].

[19] M.J. Dolan, T. Ferber, C. Hearty, F. Kahlhoefer and K. Schmidt-Hoberg, *Revised constraints and Belle II sensitivity for visible and invisible axion-like particles*, *JHEP* **12** (2017) 094 [*Erratum ibid.* **03** (2021) 190] [`arXiv:1709.00009`] [INSPIRE].

[20] B. Bellazzini, A. Mariotti, D. Redigolo, F. Sala and J. Serra, *R-axion at colliders*, *Phys. Rev. Lett.* **119** (2017) 141804 [`arXiv:1702.02152`] [INSPIRE].

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

[22] S. Knapen, T. Lin, H.K. Lou and T. Melia, *LHC limits on axion-like particles from heavy-ion collisions*, *CERN Proc.* **1** (2018) 65 [`arXiv:1709.07110`] [INSPIRE].

[23] M. Bauer, M. Heiles, M. Neubert and A. Thamm, *Axion-Like Particles at Future Colliders*, *Eur. Phys. J. C* **79** (2019) 74 [`arXiv:1808.10323`] [INSPIRE].

[24] D. Aloni, Y. Soreq and M. Williams, *Coupling QCD-Scale Axionlike Particles to Gluons*, *Phys. Rev. Lett.* **123** (2019) 031803 [arXiv:1811.03474] [InSPIRE].

[25] A. Carmona, C. Scherb and P. Schwaller, *Charming ALPs*, *JHEP* **08** (2021) 121 [arXiv:2101.07803] [InSPIRE].

[26] BELLE-II collaboration, *Search for Axion-Like Particles produced in e^+e^- collisions at Belle II*, *Phys. Rev. Lett.* **125** (2020) 161806 [arXiv:2007.13071] [InSPIRE].

[27] J. Heeck, *Interpretation of Lepton Flavor Violation*, *Phys. Rev. D* **95** (2017) 015022 [arXiv:1610.07623] [InSPIRE].

[28] S.T. Petcov, *The Processes $\mu \to e + \gamma, \mu \to e + \overline{e}, \nu' \to \nu + \gamma$ in the Weinberg-Salam Model with Neutrino Mixing*, *Sov. J. Nucl. Phys.* **25** (1977) 340 [*Erratum ibid.* **25** (1977) 698] [*Erratum ibid.* **25** (1977) 1336] [InSPIRE].

[29] G. Hernández-Tomé, G. López Castro and P. Roig, *Flavor violating leptonic decays of τ and μ leptons in the Standard Model with massive neutrinos*, *Eur. Phys. J. C* **79** (2019) 84 [*Erratum ibid.* **80** (2020) 438] [arXiv:1807.06050] [InSPIRE].

[30] MEG collaboration, *Search for the lepton flavour violating decay $\mu^+ \to e^+\gamma$ with the full dataset of the MEG experiment*, *Eur. Phys. J. C* **76** (2016) 434 [arXiv:1605.05081] [InSPIRE].

[31] SINDRUM collaboration, *Search for the Decay $\mu^+ \to e^+e^+e^-$*, *Nucl. Phys. B* **299** (1988) 1 [InSPIRE].

[32] A. Blondel et al., *Research Proposal for an Experiment to Search for the Decay $\mu \to eee$*, arXiv:1301.6113 [InSPIRE].

[33] A.M. Baldini et al., *MEG Upgrade Proposal*, arXiv:1301.7225 [InSPIRE].

[34] A. Cordero-Cid, G. Tavares-Velasco and J.J. Toscano, *Implications of a very light pseudoscalar boson on lepton flavor violation*, *Phys. Rev. D* **72** (2005) 117701 [hep-ph/0511331] [InSPIRE].

[35] P.S.B. Dev, R.N. Mohapatra and Y. Zhang, *Lepton Flavor Violation Induced by a Neutral Scalar at Future Lepton Colliders*, *Phys. Rev. Lett.* **120** (2018) 221804 [arXiv:1711.08430] [InSPIRE].

[36] M. Bauer, M. Neubert, S. Renner, M. Schnubel and A. Thamm, *Axionlike Particles, Lepton-Flavor Violation, and a New Explanation of a_μ and a_e*, *Phys. Rev. Lett.* **124** (2020) 211803 [arXiv:1908.00008] [InSPIRE].

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

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

[39] M. Endo, S. Iguro and T. Kitahara, *Probing $e\mu$ flavor-violating ALP at Belle II*, *JHEP* **06** (2020) 040 [arXiv:2002.05948] [InSPIRE].

[40] B. Batell, M. Pospelov and A. Ritz, *Multi-lepton Signatures of a Hidden Sector in Rare B Decays*, *Phys. Rev. D* **83** (2011) 054005 [arXiv:0911.4938] [InSPIRE].

[41] J.F. Kamenik and C. Smith, *FCNC portals to the dark sector*, *JHEP* **03** (2012) 090 [arXiv:1111.6402] [InSPIRE].

[42] M.B. Gavela, R. Houtz, P. Quilez, R. Del Rey and O. Sumensari, *Flavor constraints on electroweak ALP couplings*, *Eur. Phys. J. C* **79** (2019) 369 [arXiv:1901.02031] [InSPIRE].

[43] BELLE-II collaboration, *Belle II Technical Design Report*, `arXiv:1011.0352` [INSPIRE].

[44] BELLE-II collaboration, *The Belle II Physics Book*, *PTEP* **2019** (2019) 123C01 [*Erratum ibid.* **2020** (2020) 029201] [`arXiv:1808.10567`] [INSPIRE].

[45] BELLE-II collaboration, *Prospects for τ lepton physics at Belle II*, *SciPost Phys. Proc.* **1** (2019) 003 [`arXiv:1812.04225`] [INSPIRE].

[46] R. Essig, R. Harnik, J. Kaplan and N. Toro, *Discovering New Light States at Neutrino Experiments*, *Phys. Rev. D* **82** (2010) 113008 [`arXiv:1008.0636`] [INSPIRE].

[47] S. Andreas, O. Lebedev, S. Ramos-Sanchez and A. Ringwald, *Constraints on a very light CP-odd Higgs of the NMSSM and other axion-like particles*, *JHEP* **08** (2010) 003 [`arXiv:1005.3978`] [INSPIRE].

[48] BaBar collaboration, *Search for a muonic dark force at BABAR*, *Phys. Rev. D* **94** (2016) 011102 [`arXiv:1606.03501`] [INSPIRE].

[49] ARGUS collaboration, *A Search for lepton flavor violating decays τ → eα, τ → μα*, *Z. Phys. C* **68** (1995) 25 [INSPIRE].

[50] BELLE collaboration, *MC study for the lepton flavor violating tau decay into a lepton and an undetectable particle*, *Nucl. Part. Phys. Proc.* **287-288** (2017) 218 [INSPIRE].

[51] D. Guadagnoli, C.B. Park and F. Tenchini, *τ → ℓ+ invisible through invisible-savvy collider variables*, *Phys. Lett. B* **822** (2021) 136701 [`arXiv:2106.16236`] [INSPIRE].

[52] E. De La Cruz-Burelo, M. Hernandez-Villanueva and A. De Yta-Hernandez, *New method for beyond the Standard Model invisible particle searches in tau lepton decays*, *Phys. Rev. D* **102** (2020) 115001 [`arXiv:2007.08239`] [INSPIRE].

[53] K. Ma, *Polarization Effects in Lepton Flavor Violated Decays Induced by Axion-Like Particles*, `arXiv:2104.11162` [INSPIRE].

[54] F. Tenchini et al., *First results and prospects for tau LFV decay τ → e + α(invisible) at Belle II*, *PoS* **ICHEP2020** (2021) 288 [INSPIRE].

[55] BaBar collaboration, *Measurements of Charged Current Lepton Universality and $|V_{us}|$ using Tau Lepton Decays to $e^- \bar\nu_e \nu_\tau$, $\mu^- \bar\nu_\mu \nu_\tau$, $\pi^- \nu_\tau$, and $K^- \nu_\tau$*, *Phys. Rev. Lett.* **105** (2010) 051602 [`arXiv:0912.0242`] [INSPIRE].

[56] S. Iguro, Y. Omura and M. Takeuchi, *Probing μτ flavor-violating solutions for the muon g − 2 anomaly at Belle II*, *JHEP* **09** (2020) 144 [`arXiv:2002.12728`] [INSPIRE].

[57] BELLE collaboration, *The Belle Detector*, *Nucl. Instrum. Meth. A* **479** (2002) 117 [INSPIRE].

[58] K. Hayasaka et al., *Search for Lepton Flavor Violating Tau Decays into Three Leptons with 719 Million Produced τ⁺τ⁻ Pairs*, *Phys. Lett. B* **687** (2010) 139 [`arXiv:1001.3221`] [INSPIRE].

[59] BaBar collaboration, *Limits on τ Lepton-Flavor Violating Decays in three charged leptons*, *Phys. Rev. D* **81** (2010) 111101 [`arXiv:1002.4550`] [INSPIRE].

[60] D. Curtin et al., *Long-Lived Particles at the Energy Frontier: The MATHUSLA Physics Case*, *Rept. Prog. Phys.* **82** (2019) 116201 [`arXiv:1806.07396`] [INSPIRE].

[61] L. Lee, C. Ohm, A. Soffer and T.-T. Yu, *Collider Searches for Long-Lived Particles Beyond the Standard Model*, *Prog. Part. Nucl. Phys.* **106** (2019) 210 [`arXiv:1810.12602`] [INSPIRE].

[62] J. Alimena et al., *Searching for long-lived particles beyond the Standard Model at the Large Hadron Collider*, *J. Phys. G* **47** (2020) 090501 [`arXiv:1903.04497`] [INSPIRE].

[63] C.O. Dib, J.C. Helo, M. Nayak, N.A. Neill, A. Soffer and J. Zamora-Saa, *Searching for a sterile neutrino that mixes predominantly with ν_τ at B factories*, *Phys. Rev. D* **101** (2020) 093003 [arXiv:1908.09719] [INSPIRE].

[64] C.S. Kim, Y. Kwon, D. Lee, S. Oh and D. Sahoo, *Probing sterile neutrinos in $B(D)$ meson decays at Belle II (BESIII)*, *Eur. Phys. J. C* **80** (2020) 730 [arXiv:1908.00376] [INSPIRE].

[65] D.W. Kang, P. Ko and C.-T. Lu, *Exploring properties of long-lived particles in inelastic dark matter models at Belle II*, *JHEP* **04** (2021) 269 [arXiv:2101.02503] [INSPIRE].

[66] M. Acevedo, A. Blackburn, N. Blinov, B. Shuve and M. Stone, *Multi-track Displaced Vertices at B-Factories*, *JHEP* **09** (2021) 154 [arXiv:2105.12744] [INSPIRE].

[67] S. Dreyer et al., *Physics reach of a long-lived particle detector at Belle II*, arXiv:2105.12962 [INSPIRE].

[68] M. Duerr, T. Ferber, C. Hearty, F. Kahlhoefer, K. Schmidt-Hoberg and P. Tunney, *Invisible and displaced dark matter signatures at Belle II*, *JHEP* **02** (2020) 039 [arXiv:1911.03176] [INSPIRE].

[69] M. Duerr, T. Ferber, C. Garcia-Cely, C. Hearty and K. Schmidt-Hoberg, *Long-lived Dark Higgs and Inelastic Dark Matter at Belle II*, *JHEP* **04** (2021) 146 [arXiv:2012.08595] [INSPIRE].

[70] A. Filimonova, R. Schäfer and S. Westhoff, *Probing dark sectors with long-lived particles at BELLE II*, *Phys. Rev. D* **101** (2020) 095006 [arXiv:1911.03490] [INSPIRE].

[71] X. Chen, Z. Hu, Y. Wu and K. Yi, *Search for dark photon and dark matter signatures around electron-positron colliders*, *Phys. Lett. B* **814** (2021) 136076 [arXiv:2001.04382] [INSPIRE].

[72] S. Dey et al., *Long-lived light neutralinos at Belle II*, *JHEP* **02** (2021) 211 [arXiv:2012.00438] [INSPIRE].

[73] E. Bertholet, S. Chakraborty, V. Loladze, T. Okui, A. Soffer and K. Tobioka, *Heavy QCD Axion at Belle II: Displaced and Prompt Signals*, arXiv:2108.10331 [INSPIRE].

[74] BELLE collaboration, *Search for heavy neutrinos at Belle*, *Phys. Rev. D* **87** (2013) 071102 [*Erratum ibid.* **95** (2017) 099903] [arXiv:1301.1105] [INSPIRE].

[75] BABAR collaboration, *Search for Long-Lived Particles in e^+e^- Collisions*, *Phys. Rev. Lett.* **114** (2015) 171801 [arXiv:1502.02580] [INSPIRE].

[76] J. Heeck and W. Rodejohann, *Lepton flavor violation with displaced vertices*, *Phys. Lett. B* **776** (2018) 385 [arXiv:1710.02062] [INSPIRE].

[77] ALEPH collaboration, *Fermion pair production in e^+e^- collisions at 189–$209\,GeV$ and constraints on physics beyond the standard model*, *Eur. Phys. J. C* **49** (2007) 411 [hep-ex/0609051] [INSPIRE].

[78] N.D. Christensen and C. Duhr, *FeynRules — Feynman rules made easy*, *Comput. Phys. Commun.* **180** (2009) 1614 [arXiv:0806.4194] [INSPIRE].

[79] A. Alloul, N.D. Christensen, C. Degrande, C. Duhr and B. Fuks, *FeynRules 2.0 — A complete toolbox for tree-level phenomenology*, *Comput. Phys. Commun.* **185** (2014) 2250 [arXiv:1310.1921] [INSPIRE].

[80] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer and T. Reiter, *UFO — The Universal FeynRules Output*, *Comput. Phys. Commun.* **183** (2012) 1201 [arXiv:1108.2040] [INSPIRE].

[81] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [arXiv:1405.0301] [INSPIRE].

[82] SLAC E158 collaboration, *Precision measurement of the weak mixing angle in Moller scattering*, *Phys. Rev. Lett.* **95** (2005) 081601 [hep-ex/0504049] [INSPIRE].

[83] A.L. Hallin, F.P. Calaprice, R.A. McPherson and E.R.J. Saettler, *Sensitive search for resonances in low-energy e^+e^- scattering*, *Phys. Rev. D* **45** (1992) 3955 [INSPIRE].

[84] G. Haghighat and M. Mohammadi Najafabadi, *Search for lepton-flavor-violating ALPs at a future muon collider and utilization of polarization-induced effects*, arXiv:2106.00505 [INSPIRE].

[85] J.P. Delahaye et al., *Muon Colliders*, arXiv:1901.06150 [INSPIRE].

[86] M. Boscolo, J.-P. Delahaye and M. Palmer, *The future prospects of muon colliders and neutrino factories*, *Rev. Accel. Sci. Tech.* **10** (2019) 189 [arXiv:1808.01858] [INSPIRE].

[87] H. Al Ali et al., *The Muon Smasher's Guide*, arXiv:2103.14043 [INSPIRE].

[88] R. Bollig, W. DeRocco, P.W. Graham and H.-T. Janka, *Muons in Supernovae: Implications for the Axion-Muon Coupling*, *Phys. Rev. Lett.* **125** (2020) 051104 [*Erratum ibid.* **126** (2021) 189901] [arXiv:2005.07141] [INSPIRE].

[89] G. Lucente and P. Carenza, *Supernova bound on axionlike particles coupled with electrons*, *Phys. Rev. D* **104** (2021) 103007 [arXiv:2107.12393] [INSPIRE].

[90] P. Foldenauer and J. Jaeckel, *Purely flavor-changing Z' bosons and where they might hide*, *JHEP* **05** (2017) 010 [arXiv:1612.07789] [INSPIRE].

[91] W. Altmannshofer, C.-Y. Chen, P.S. Bhupal Dev and A. Soni, *Lepton flavor violating Z' explanation of the muon anomalous magnetic moment*, *Phys. Lett. B* **762** (2016) 389 [arXiv:1607.06832] [INSPIRE].

[92] A. Pich, *Precision Tau Physics*, *Prog. Part. Nucl. Phys.* **75** (2014) 41 [arXiv:1310.7922] [INSPIRE].

[93] D.A. Bryman, S. Ito and R. Shrock, *Upper limits on branching ratios of the lepton-flavor-violating decays $\tau \to \ell\gamma\gamma$ and $\tau \to \ell X$*, *Phys. Rev. D* **104** (2021) 075032 [arXiv:2106.02451] [INSPIRE].

[94] BaBar collaboration, *Searches for Lepton Flavor Violation in the Decays $\tau^\pm \to e^\pm\gamma$ and $\tau^\pm \to \mu^\pm\gamma$*, *Phys. Rev. Lett.* **104** (2010) 021802 [arXiv:0908.2381] [INSPIRE].

[95] 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].

[96] Muon $g-2$ collaboration, *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm*, *Phys. Rev. Lett.* **126** (2021) 141801 [arXiv:2104.03281] [INSPIRE].

[97] Muon $g-2$ collaboration, *Final Report of the Muon E821 Anomalous Magnetic Moment Measurement at BNL*, *Phys. Rev. D* **73** (2006) 072003 [hep-ex/0602035] [INSPIRE].

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

[99] M.A. Buen-Abad, 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].

[100] W.-Y. Keung, D. Marfatia and P.-Y. Tseng, *Axion-Like Particles, Two-Higgs-Doublet Models, Leptoquarks, and the Electron and Muon* $g - 2$, *LHEP* **2021** (2021) 209 [arXiv:2104.03341] [InSPIRE].

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

[102] E.M. Riordan et al., *A Search for Short Lived Axions in an Electron Beam Dump Experiment*, *Phys. Rev. Lett.* **59** (1987) 755 [InSPIRE].

[103] CHARM collaboration, *Search for Axion Like Particle Production in* 400 *GeV Proton-Copper Interactions*, *Phys. Lett. B* **157** (1985) 458 [InSPIRE].

[104] M. Davier and H. Nguyen Ngoc, *An Unambiguous Search for a Light Higgs Boson*, *Phys. Lett. B* **229** (1989) 150 [InSPIRE].

[105] A. Bross, M. Crisler, S.H. Pordes, J. Volk, S. Errede and J. Wrbanek, *A Search for Shortlived Particles Produced in an Electron Beam Dump*, *Phys. Rev. Lett.* **67** (1991) 2942 [InSPIRE].

[106] L. Willmann et al., *New bounds from searching for muonium to anti-muonium conversion*, *Phys. Rev. Lett.* **82** (1999) 49 [hep-ex/9807011] [InSPIRE].

[107] L. Calibbi and G. Signorelli, *Charged Lepton Flavour Violation: An Experimental and Theoretical Introduction*, *Riv. Nuovo Cim.* **41** (2018) 71 [arXiv:1709.00294] [InSPIRE].

[108] SINDRUM II collaboration, *A Search for muon to electron conversion in muonic gold*, *Eur. Phys. J. C* **47** (2006) 337 [InSPIRE].

[109] Mu2e collaboration, *Mu2e Technical Design Report*, arXiv:1501.05241 [InSPIRE].

[110] COMET collaboration, *COMET Phase-I Technical Design Report*, *PTEP* **2020** (2020) 033C01 [arXiv:1812.09018] [InSPIRE].

[111] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [arXiv:1410.3012] [InSPIRE].

[112] J. Heeck, *Lepton flavor violation with light vector bosons*, *Phys. Lett. B* **758** (2016) 101 [arXiv:1602.03810] [InSPIRE].

[113] C.-H. Chen and T. Nomura, $L_\mu - L_\tau$ *gauge-boson production from lepton flavor violating* τ *decays at Belle II*, *Phys. Rev. D* **96** (2017) 095023 [arXiv:1704.04407] [InSPIRE].

[114] T. Flacke, C. Frugiuele, E. Fuchs, R.S. Gupta and G. Perez, *Phenomenology of relaxion-Higgs mixing*, *JHEP* **06** (2017) 050 [arXiv:1610.02025] [InSPIRE].

[115] P.S. Bhupal Dev, R.N. Mohapatra and Y. Zhang, *Displaced photon signal from a possible light scalar in minimal left-right seesaw model*, *Phys. Rev. D* **95** (2017) 115001 [arXiv:1612.09587] [InSPIRE].

[116] P.S.B. Dev, R.N. Mohapatra and Y. Zhang, *Long Lived Light Scalars as Probe of Low Scale Seesaw Models*, *Nucl. Phys. B* **923** (2017) 179 [arXiv:1703.02471] [InSPIRE].

[117] BaBar collaboration, *Search for a Dark Leptophilic Scalar in* e^+e^- *Collisions*, *Phys. Rev. Lett.* **125** (2020) 181801 [arXiv:2005.01885] [InSPIRE].