Angling for insights: illuminating light new physics at Mu3e through angular correlations

Simon Knapen ⬡,[a,b] Kevin Langhoff ⬡,[b,a] Toby Opferkuch ⬡[b,a] and Diego Redigolo ⬡[c]

[a]*Theory Group, Lawrence Berkeley National Laboratory,*
Berkeley, CA 94720, U.S.A.

[b]*Berkeley Center for Theoretical Physics, University of California,*
Berkeley, CA 94720, U.S.A.

[c]*INFN, Sezione di Firenze,*
Via G. Sansone 1, 50019 Sesto Fiorentino, Italy

E-mail: smknapen@lbl.gov, klanghoff@berkeley.edu,
tobyopferkuch@berkeley.edu, d.redigolo@gmail.com

Abstract: We examine the capability of Mu3e to probe light new physics scenarios that produce a prompt electron-positron resonance and demonstrate how angular observables are instrumental in enhancing the experimental sensitivity. We systematically investigate the effect of Mu3e's expected sensitivity on the parameter space of the dark photon, as well as on axion-like particles and light scalars with couplings to muons and electrons.

https://doi.org/10.1007/JHEP07(2024)194

Contents

1 Introduction

The coming decade promises impressive progress in muon physics, with a host of new experiments that will be coming online or are already taking data. The muon $g-2$ experiment has reached an unprecedented level of precision on the muon magnetic dipole moment [1] and it will soon be complemented by a suite of new probes of charged lepton flavor violation (CLFV). Experiments such as Mu2e [2], COMET [3] and DeeMe [4] focus on $\mu \to e$ conversions, while MEG II [5] and Mu3e [6, 7] will be hunting for exotic muon decay channels as the hallmarks of beyond Standard Model (SM) physics. While this program has originally been envisioned as a probe of heavy new physics at mass scales as high as 10^4–10^5 TeV, see e.g. [8–12], it also provides a unique opportunity to test light new particles with extremely feeble couplings to the SM, see e.g. [13–18].

In this work we zoom in on searches for light, weakly coupled particles at Mu3e that decay promptly to an e^+e^- pair [13, 19]. The Mu3e experiment [6, 7] at the Paul Scherrer Institute (PSI) in Switzerland will study around 2.5×10^{15} muon decays during phase-I of
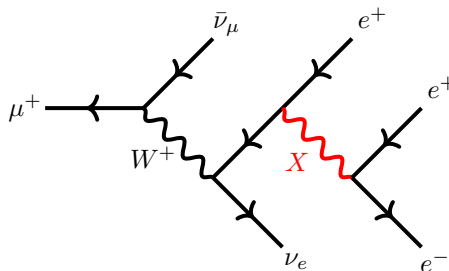
Figure 1. Example diagram of a light new particle X produced in muon decays. Depending on the model, emissions from the muon and/or the neutrinos are also considered.

operations and 5.5×10^{16} muon decays during phase-II with the primary goal of increasing the sensitivity to the CLFV decay $\mu^+ \to e^+ e^+ e^-$. We focus our attention on the lepton flavor conserving decay $\mu^+ \to e^+ \bar{\nu}_\mu \nu_e X$, where the new particle X can be radiated from any charged SM particle before decaying back to a dilepton pair ($X \to e^+ e^-$). This results in three tracks plus missing energy, as exemplified in figure 1. This channel was already considered in [13, 19] with the main motivation of extending the reach on the parameter space of the minimal dark photon in the mass range between approximately 10–100 MeV. In this paper we expand on these previous analyses in two ways: *i)* We broaden the physics case for Mu3e by considering different new physics models that can be probed in $\mu^+ \to e^+ \bar{\nu}_\mu \nu_e X$. *ii)* We further study the impact of angular correlations in distinguishing signal from background, and show that a few simple variables can already deliver an $\mathcal{O}(1)$ improvement to the sensitivity without the need for modifications to the detector or data taking strategy. In the event of a discovery, angular correlations, including those obtained when accounting for muon polarization, can also be used to pin down the properties of the new particle.

The main SM background originates from radiative muon decay ($\mu^+ \to e^+ \bar{\nu}_\mu \nu_e \gamma^*$) where the off-shell SM photon splits into an electron-positron pair ($\gamma^* \to e^+ e^-$). This process was computed at leading order in refs. [20, 21] and more recently at next-to-leading order in refs. [22–24] and constitutes an irreducible background for all the signals considered here, besides also being the main background for the flagship CLFV Mu3e search: $\mu^+ \to e^+ e^+ e^-$. For the CLFV decay, the absence of missing energy allows one to impose very effective kinematic cuts, e.g. by demanding that the three tracks reconstruct the invariant mass of the muon. Such a cut is much too strong for the models we consider here, and differentiating $\mu^+ \to e^+ \bar{\nu}_\mu \nu_e X$ from the background therefore requires a different strategy. In particular, a search for a resonance in the invariant mass distributions of the $e^+ e^-$ pairs is already highly effective [13, 19], but nevertheless a large irreducible SM background remains. In this work we show that the sensitivity of a vanilla 'bump hunt' analysis can be enhanced by an $\mathcal{O}(1)$ factor, by considering additional correlations in the angles and energies of the final-state charged leptons.

We apply our analysis to light new physics scenarios where the new particle X is either spin-0 or spin-1. For the purpose of our analysis, we can parametrize the interactions of X with muons and electrons as

$$\begin{aligned}
\mathcal{L}_0 &\supset X \left[\bar{\mu}(g_S^\mu + g_P^\mu \gamma_5)\mu + \bar{e}(g_S^e + g_P^e \gamma_5)e \right] , \\
\mathcal{L}_1 &\supset X^\alpha \left[\bar{\mu}\gamma_\alpha(g_V^\mu + g_A^\mu \gamma_5)\mu + \bar{e}\gamma_\alpha(g_V^e + g_A^e \gamma_5)e \right] ,
\end{aligned} \tag{1.1}$$

where for the scalar model, the couplings $g_S^\ell(g_P^\ell)$ parametrize the strength to parity even (odd) operator. Similarly, the $g_V^\ell(g_A^\ell)$ parametrize the strength of a spin-1 X^α to the vector (axial) current. We neglected extra possible interactions of X with neutrinos, whose phenomenological relevance for Mu3e will be discussed in appendix A. We further assumed that interactions of the scalar X with photons are suppressed relative to its interaction with the charged leptons. After obtaining results and physical insights with the general setup in eqs. (1.1), we map those results onto motivated models, and compare with existing bounds.

The remainder of this paper is organized as follows: in section 2 we briefly describe our simulation setup and analysis strategy. We then discuss in section 3 the physical origin of the kinematical variables which allow us to improve the discrimination between signal and background. In section 4 we will discuss explicit models that fit in the generic parametrization of eqs. (1.1) and compare our forecasted reach at Mu3e with current constraints. We complement the discussion in the appendices: additional results and models are presented in appendix A, followed by a discussion on the effects of the muon polarization in appendix B. In appendix C we give further details on our simulation framework and validate our approximate detector modeling with public experimental results. In appendix D we give explicit formulas for the differential widths of every model discussed in the paper.

2 Analysis strategy

In the relevant regime of couplings for Mu3e, the decay width of the new resonance is much smaller than the detector resolution. The main observable is therefore a narrow resonance in the invariant mass distributions of the e^+e^--pairs, while the dominant SM background is a smooth falling distribution from the internal conversion process $\mu^+ \to \bar{\nu}_\mu e^+ \nu_e(\gamma^* \to e^+e^-)$. In this section we describe our variables and general strategy to enhance the sensitivity of a plain resonance search; all details on our event generation and our modeling of the detector response are deferred to appendix C.

Since the final state contains two positrons, there is a combinatoric ambiguity in identifying the e^+e^- pair originating from the new resonance. We label the two possible invariant mass combinations of these pairs with $m_{ee,1}$ and $m_{ee,2}$. There are two main approaches to mitigating the combinatoric background. The first is to treat both combinations as individual events, effectively doubling the number of events in the m_{ee} distribution. Once smearing and detector effects are included this can lead to a small number of spurious signal events where, depending on the value of m_X, the invariant mass of the electron with the second positron may also be close to m_X. This is shown in figure 2, where the thin dashed lines show signal and background distributions utilizing this method.

The second approach, which we will adopt in what follows, involves choosing the e^+e^- pair for which the invariant mass is closest to an arbitrarily chosen hypothesized signal mass m_X. In other words, we define the variable

$$\hat{m}_X \equiv \begin{cases} m_{ee,1} \text{ for } |m_{ee,1} - m_X| < |m_{ee,2} - m_X|, \\ m_{ee,2} \text{ for } |m_{ee,1} - m_X| > |m_{ee,2} - m_X|. \end{cases} \tag{2.1}$$

This choice automatically maximizes the signal efficiency, but artificially skews the background distribution by creating a feature centered around $m_{ee} = m_X$. This feature (solid black
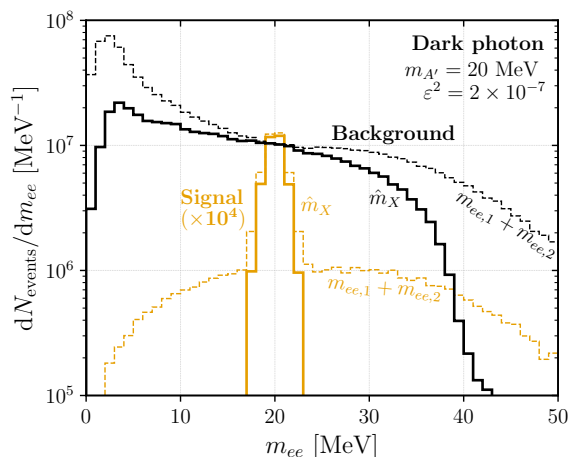
Figure 2. Illustrative example of the signal (**orange lines**) and background (*black lines*) distributions in the invariant mass plane of the electron-positron pair m_{ee} after smearing and detector acceptance are applied, see appendix C for details. In the notation of eq. (1.1), we show the signal for the case of a vector spin-1 resonance with $g_V^\mu = g_V^e = e\epsilon$ (dark photon model) with mass $m_X = 20\,\mathrm{MeV}$ and coupling $\varepsilon^2 = 2 \times 10^{-7}$ is multiplied by a factor 10^4. Both the background and signal are shown for two cases: (i) the naive approach where one considers both the invariant mass combinations of the electron-positron pair (*thin dashed* line labeled $m_{ee,1} + m_{ee,2}$) and (ii) our approach through the use of the observable $\hat{m}_X$, see eq. (2.1), which picks the combination closest to the mass hypothesis of the new resonance (*thick solid* line labeled $\hat{m}_X$).

line in figure 2) is however extremely broad compared to both the resonance width and the mass resolution of the detector. We therefore do not expect it to be a significant source of systematic uncertainties in a data-driven background estimation.

We define p_X as the four-vector of the e^+e^- pair with invariant mass $\hat{m}_X$ and label the positron assigned to the hypothesized resonance [through eq. (2.1)] with e_X^+. The remaining, 'solitary' positron is denoted by e_s^+. This means that whether a positron is assigned as 'e_X^+' or 'e_s^+' depends on the hypothesized value for m_X. With this convention in mind, we label the four-momentum of 'solitary' positron $p_{e_s^+}$ and the four-momentum of the positron assigned to the resonance with $p_{e_X^+}$. The electron four-momentum is denoted by p_{e^-}. We further define the track energies normalized to the muon mass m_μ through

$$x_{e^-} \equiv \frac{2E_{e^-}}{m_\mu}, \quad x_{e_X^+} \equiv \frac{2E_{e_X^+}}{m_\mu}, \text{ and } x_{e_s^+} \equiv \frac{2E_{e_s^+}}{m_\mu}, \tag{2.2}$$

which take values between 0 and 1.

In the next section, we will describe how angular correlations between these three tracks can help discriminating the signal from the background. Concretely, the variables we will need are

$$\cos\theta_{e_s^+ e_X^+} \equiv \vec{p}_{e_s^+} \cdot \vec{p}_{e_X^+} / (|\vec{p}_{e_s^+}||\vec{p}_{e_X^+}|), \tag{2.3}$$

$$\cos\theta_{e^- e_s^+} \equiv \vec{p}_{e_s^+} \cdot \vec{p}_{e^-} / (|\vec{p}_{e_s^+}||\vec{p}_{e^-}|), \tag{2.4}$$

where the vector arrows label three-momenta. There is additional angular information if the polarization of the muon beam is taken into account; however, it turns out that these

additional angles yield less distinguishing power. They are however more sensitive to the differences between the models presented in the previous section, and are therefore well suited to distinguish models from one another in the event of a positive discovery. We defer this discussion to appendix B.

A rigorous experimental analysis would presumably construct a global likelihood for the full kinematics of both signal and background. Here we opt for a simplified approach, which is computationally faster and has the advantage that the effect of the angular correlations can be isolated easily. Concretely, we first place a cut on the $\hat{m}_X$ distribution in a symmetric window around our hypothesis for m_X. The optimal size of this window is determined by maximizing $S/\sqrt{B}$, with S and B the number of signal and background events for a given window size. We subsequently bin the events in this window in two distinguishing variables from the set of $\{\cos\theta_{e^-e_s^+}, \cos\theta_{e^-e_X^+}, x_{e_s^+}\}$. We must limit ourselves to binning in two dimensions because of size limitations of our background Monte Carlo sample, while in section 3 we discuss why these variables are the most effective in separating the signal from the background.

To estimate the additional sensitivity provided by any chosen pair of distinguishing variables we construct a binned Poisson likelihood over both variables

$$\mathcal{L}(S_i, B_i, n_i) = \prod_i \frac{(S_i + B_i)^{n_i}}{n_i!} e^{-(S_i+B_i)}. \tag{2.5}$$

Here S_i (B_i) are the bin-wise signal (background) expectation values, n_i are the number of observed events in the i-th bin. The n_i are drawn from the background-only distribution ($S_i = 0$). As our test statistic, we then define the log-likelihood ratio using eq. (2.5)

$$\lambda \equiv -2\log\left[\frac{\mathcal{L}(S_i, B_i, n_i)}{\mathcal{L}(S_i = 0, B_i, n_i)}\right] \approx \sum_i S_i^2/B_i. \tag{2.6}$$

In the high statistics limit, the Poisson likelihood in eq. (2.5) asymptotes to a Gaussian and the likelihood ratio can be approximated by the second equality in eq. (2.6), where we took $n_i \approx B_i$. To extract the 95% expected exclusion limit (one-sided), we solve for $\lambda = 2.71$. We performed this procedure both with the Poisson and Gaussian likelihoods, and established that the latter is always an excellent approximation. Additional details and validation figures can be found in appendix C.

3 Angular correlations

Because of the difference in matrix elements, we expect that the full angular distribution may provide additional signal versus background discrimination. Making use of all this information is however non-trivial, as the statistics of any realistic Monte Carlo sample will tend to be smaller than the number of actual background events Mu3e will record. This means that it is difficult to build a reliable, high-dimensional likelihood function without a data-driven background estimation.

We address this difficulty in two ways: firstly, we implement a custom set of cuts in Madgraph, which allows us to generate a separate, high-statistics background sample tailored to each signal model point. This method is described and validated in appendix C. Secondly,
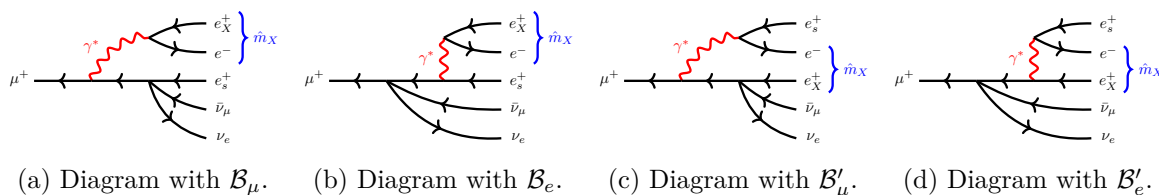
(a) Diagram with $\mathcal{B}_\mu$.   (b) Diagram with $\mathcal{B}_e$.   (c) Diagram with $\mathcal{B}'_\mu$.   (d) Diagram with $\mathcal{B}'_e$.

Figure 3. Feynman diagrams for the SM background in eq. (3.2). The left two diagrams identify the electron-positron pair with invariant mass closest to $\hat{m}_X$ as those that came from the virtual photon. The right two diagrams identify the other combination. It is these second two diagrams which are exclusive to the background leading to an altered distribution in $\cos\theta_{e^-e_s^+}$ compared to the signal.

we analyze the amplitudes of both signal and background to identify a small set of the most promising variables with which to perform a likelihood analysis. In this section we will describe the physics behind these variables.

3.1 Model-independent variable

Before we discuss any dynamics related to a specific model, it is valuable to consider first the general features of both signal and background amplitudes. Note that for the sake of clarity we will set the electron mass to zero in what follows, though m_e was set to its physical value for all our numerical results. The general form of the amplitude for the signal process, irrespective of the specific model considered, will be

$$\mathcal{M}_{\text{sig}} = \frac{\mathcal{S}_\mu}{(p_\mu - p_X)^2 - m_\mu^2} + \frac{\mathcal{S}_e}{(p_X + p_{e_s^+})^2} \, . \tag{3.1}$$

Here the first and second terms correspond to the case where X is radiated off the final state e_s^+ and initial state μ^+, respectively, while $\mathcal{S}_{e/\mu}$ are the numerators of the amplitude which are model dependent but do not have additional poles.

This must be compared to the background which also has a propagator from the virtual photon. Concretely, there are four diagrams, shown in figure 3, which correspond to an amplitude of the form

$$\mathcal{M}_{\text{bkg}} = \frac{1}{(p_\mu - p_X)^2 - m_\mu^2} \left[\frac{\mathcal{B}_\mu}{p_{e^-} \cdot p_{e_X^+}} + \frac{\mathcal{B}'_\mu}{p_{e^-} \cdot p_{e_s^+}} \right] + \frac{1}{(p_X + p_{e_s^+})^2} \left[\frac{\mathcal{B}_e}{p_{e^-} \cdot p_{e_X^+}} + \frac{\mathcal{B}'_e}{p_{e^-} \cdot p_{e_s^+}} \right] \, . \tag{3.2}$$

The $\mathcal{B}_\mu$, $\mathcal{B}'_\mu$, $\mathcal{B}_e$ and $\mathcal{B}'_e$ terms correspond respectively to the diagrams 3(a), 3(c), 3(b) and 3(d) in figure 3. If we investigate the pole structure in eq. (3.2) we find

$$p_{e^-} \cdot p_{e_X^+} = m_X^2/2 \, , \tag{3.3}$$

$$p_{e^-} \cdot p_{e_s^+} = E_{e^-} E_{e_s^+} (1 - \cos\theta_{e^-e_s^+}) \, , \tag{3.4}$$

which implies that the terms which scale as $(p_{e^-} \cdot p_{e_s^+})^{-1}$ will dominate the amplitude when e^- and e_s^+ become collinear. This feature is most pronounced for large m_X where the other terms are suppressed by virtue of eq. (3.3).
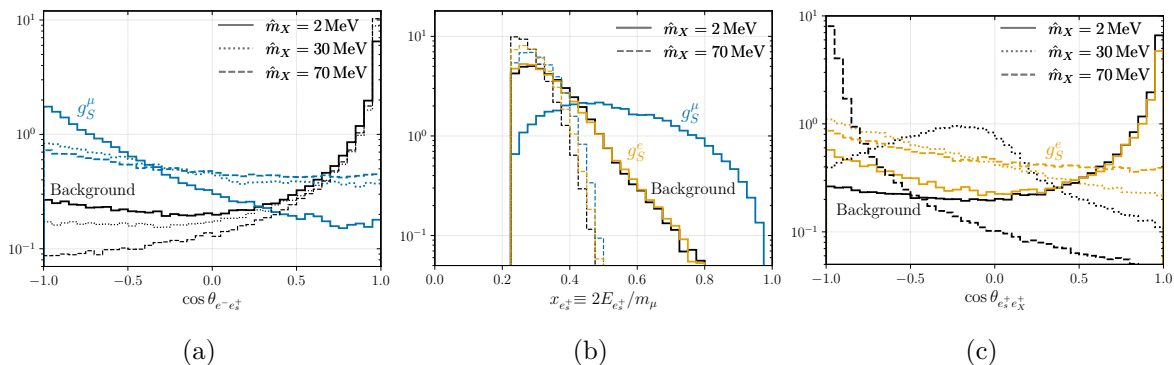
Figure 4. Normalized distributions for the background (*black lines*) and signal for a scalar coupled predominantly to muons $g_S^\mu \gg g_S^e \neq 0$ (**blue lines**) and for a scalar coupled to electrons $g_S^e \neq 0$) (**orange lines**) for the kinematic variables included in our analysis. For the background distributions we require $\hat{m}_X$ to be within a window of $\pm 2.5\,\mathrm{MeV}$ around the respective value quoted in the legend. All distributions are normalized to one.

The pole at $\cos\theta_{e^-e_s^+} = 1$ is due to the off-shell SM photon in figures 3(c) and 3(d) which is absent for all signal models. As a result $\cos\theta_{e^-e_s^+}$ is a powerful model-independent variable to help separate signal and background. This is shown in figure 4(a), which shows a clear peak for the background (black lines) across all values of $\hat{m}_X$. For illustration we also show a particular signal model (blue lines corresponding to a scalar coupled predominantly to muons $g_S^\mu \gg g_S^e \neq 0$) which is anti-correlated with the background in part due to the absence of the same pole.

3.2 Model-dependent variables

To discuss the angular variables that are model-dependent we consider a CP-even scalar coupled predominantly to either muons or electrons (through g_S^μ or g_S^e in eq. (1.1)). However, the observations discussed here depend only on whether the muon or the electron coupling dominates and will therefore also apply to the case of the CP-odd scalar.

The squared amplitudes for a scalar coupled dominantly to electrons or muons are respectively

$$\frac{1}{2}\sum_{\mathrm{spins}} |\mathcal{M}_e|^2 = 16 G_F^2 (g_S^e)^2 \frac{|\mathcal{M}_{e,1}|^2 |\mathcal{M}_{e,2}|^2}{(p_X + p_{e_s^+})^4} \,, \tag{3.5}$$

$$\frac{1}{2}\sum_{\mathrm{spins}} |\mathcal{M}_\mu|^2 = 16 G_F^2 (g_S^\mu)^2 \frac{|\mathcal{M}_{\mu,1}|^2 |\mathcal{M}_{\mu,2}|^2}{\left[(p_\mu - p_X)^2 - m_\mu^2\right]^2} \,, \tag{3.6}$$

with

$$
\begin{aligned}
|\mathcal{M}_{e,1}|^2 &= 2 E_{\nu_e} m_\mu \,, \\
|\mathcal{M}_{e,2}|^2 &= 4(p_X \cdot p_{e_s^+})(p_X \cdot p_{\nu_\mu}) - 2 m_X^2 (p_{e_s^+} \cdot p_{\nu_\mu}) \,, \\
|\mathcal{M}_{\mu,1}|^2 &= 4 m_\mu (E_X - 2 m_\mu)(p_X \cdot p_{\nu_e}) + 8 m_\mu E_{\nu_e} \left(m_\mu^2 - \frac{1}{4} m_X^2 \right) \,, \\
|\mathcal{M}_{\mu,2}|^2 &= 2 p_{e_s^+} \cdot p_{\bar{\nu}_\mu} \,.
\end{aligned} \tag{3.7}
$$

In eq. (3.5) and eq. (3.6) we recognize the pole structure we anticipated in eq. (3.1). Concretely, we can expand the denominators as

$$(p_X + p_{e_s^+})^2 = 2 E_X E_{e_s^+} (1 - \beta_X \cos\theta_{e_s^+ X}) + m_X^2 \,, \tag{3.8}$$

$$(p_\mu - p_X)^2 - m_\mu^2 = m_X^2 - 2 m_\mu E_X \,, \tag{3.9}$$

where β_X is the velocity of X. Inserting eq. (3.8) into eq. (3.5) shows that at low m_X, the amplitude for the scalar coupled to electrons prefers $E_{e_s^+}$ to be as small as possible. The same pattern can be found in the background amplitude, looking at the second line in eq. (3.2). This behavior is not present if the scalar X couples predominantly to muons, as can be seen by plugging eq. (3.9) in eq. (3.6). We therefore expect $E_{e_s^+}$ to be larger in the muon coupled case. This is shown in figure 4(b), where we used the dimensionless energy variable $x_{e_s^+}$ defined in the previous section. In the low m_X regime we therefore expect $x_{e_s^+}$ to be a useful variable for models which predominantly couple to muons.

With the variable above we have not yet fully exhausted all the angular information in the event. Using the angle $\cos\theta_{e_s^+ e_X^+}$ additional discrimination between background and signal can be achieved. The $\cos\theta_{e_s^+ e_X^+}$ distributions for signal and background are shown in figure 4(c), for different choices of m_X. For the background, we see that both positrons shift from being colinear for low m_X, to back-to-back configurations at higher masses. The former is easily understood in terms of the soft-colinear emission preferred by eq. (3.2). For high m_X, the pieces scaling as $(m_X)^{-2}$ are suppressed but the remaining terms still prefer the e^- and e_s^+ to be colinear. Momentum conservation then forces e_X^+ to be back-to-back with e_s^+ and e^-, as both neutrinos are very soft in this regime. A priori, one would therefore expect from figure 4(c) that $\cos\theta_{e_s^+ e_X^+}$ is the most powerful either at low or at high m_X. This is not the case however, as in both those limits this angle is strongly correlated with $\cos\theta_{e^- e_s^+}$ and it therefore adds little to the significance. It however does contribute at intermediate m_X.

3.3 Summary

To assess the enhancement in reach achievable by incorporating our new kinematic variables, we supplement the bump hunt strategy with a 2D likelihood in two pairs of variables: $(\cos\theta_{e^- e_s^+}, \cos\theta_{e_X^+ e_s^+})$ and $(\cos\theta_{e^- e_s^+}, x_{e_s^+})$. Subsequently, we normalize this outcome against the sensitivity obtained using a conventional bump-hunt method, resulting in the relative improvement in reach, as depicted in figure 5. Here colors refer to the various signal models while the dashed versus solid lines refer to the pair of kinematical variables chosen.

We observe an approximate twofold or better increase in reach across all models and mass selections, except in the case of the dark photon model at low m_X. In this low mass range, the dark photon signal closely resembles an off-shell SM photon exchange, leading to diminished sensitivity. Conversely, when increasing the signal resonance's mass, distinctions in angular correlations between the signal and background become more pronounced. In such scenarios, the variable $\cos\theta_{e^- e_s^+}$ predominantly contributes to sensitivity. Additionally, for (pseudo)scalar couplings to electrons, we find that the inclusion of $\cos\theta_{e_s^+ e_X^+}$ yields the most favorable results, whereas incorporating $x_{e_s^+}$ proves advantageous for (pseudo)scalar couplings to muons, as anticipated in the preceding section.
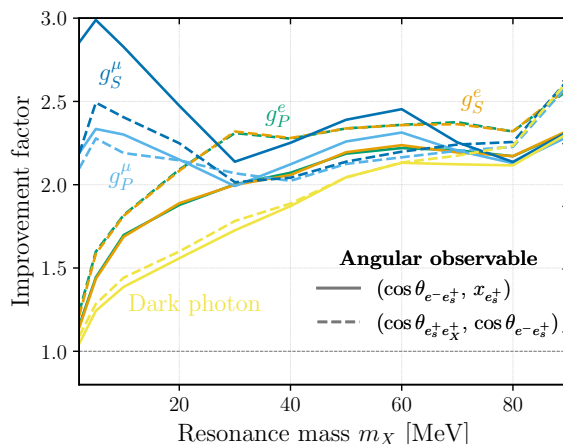
Figure 5. Improvement of the reach shown as a ratio of the rates (couplings squared) from including the two combinations of kinematic observables. The dark photon model (*yellow lines*) used here as a benchmark refers to $g_V^\mu = g_V^e = e\epsilon$. The scalar (*blue lines*) or pseudoscalar (*light-blue lines*) coupled predominantly to muons refers to $g_{S,P}^\mu \gg g_{S,P}^e \neq 0$. While lastly the scalar (*orange lines*) or pseudoscalar (*green lines*) coupled only to electrons refers to $g_{S,P}^e \neq 0$, all of the above in the notation of eq. (1.1).

4 Results

In this section we describe a few benchmark models whose unconstrained parameter space can be explored through the decay $\mu^+ \to e^+ \bar{\nu}_\mu \nu_e (X \to e^+ e^-)$. We discuss the dark photon in section 4.1, light scalars and axion-like particles in section 4.2. We comment on other models in appendix A where the Mu3e reach is likely not competitive with existing experiments.

4.1 Dark photon

Dark photons are present whenever the standard model is embedded into a theory which contains an extra U(1) gauge symmetry [25–31]. They feature in a variety of phenomenological models, such as those where it acts as a mediator particle to a dark sector [32] as well as attempts to explain the anomalous magnetic moment of the muon [33].

At low energies, the interaction between the dark photon and the SM comes through the kinetic mixing with the SM photon

$$\mathcal{L} \supset -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{4} F'_{\mu\nu} F'^{\mu\nu} - \frac{\varepsilon}{2} F_{\mu\nu} F'^{\mu\nu} \,, \tag{4.1}$$

where the mixing parameter ε can be naturally taken to be arbitrarily small. Upon diagonalization of these kinetic terms the dark photon interaction with SM fermions (f) are generated and the relevant Lagrangian for this paper is the following:

$$\mathcal{L} = -\sum_f \varepsilon q_f \bar{f} \gamma^\mu f A'_\mu + m_{A'}^2 A'_\mu A'^\mu \,, \tag{4.2}$$

with q_f the electric charge of f. In the notation of eq. (1.1) the dark photon has $g_V^\mu = g_V^e = e\varepsilon$ and $g_A^\mu = g_A^e = 0$ and its mass $m_{A'}$ can be considered as independent from its coupling to the SM.
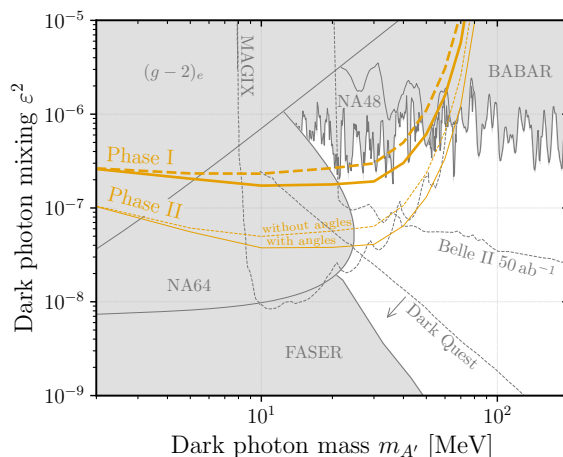
Figure 6. Projections for the 95% upper limit on ε^2 (**orange lines**) for the dark photon model at Mu3e phase I ($2.5 \times 10^{15}\mu^+$) and phase II ($5.5 \times 10^{16}\mu^+$). The *solid* and *dashed* curves indicate our projected sensitivity with and without the inclusion of angular variables described in section 3. **Gray shaded** areas are the existing exclusion limits from FASER [36], NA64 [34], the electron magnetic dipole moment [37], NA48 [38] and BABAR [39]. **Thin dashed gray** lines indicate the expected sensitivity of Belle II with the full luminosity of 50 ab^{-1} [40], the MAGIX [41] proposal at MESA and the DarkQuest proposal at Fermilab [42]. The existing exclusions in this figure were generated with the DarkCast package [43].

If the dark photon is lighter than the other potential states in the dark sector, it will decay back to SM particles through the interaction in eq. (4.2). For $2m_e < m_{A'} < m_\mu$, its proper decay length is given by

$$c\tau_{A'} \approx 0.5\,\mathrm{mm}\left(\frac{10^{-4}}{\epsilon}\right)^2\left(\frac{10\,\mathrm{MeV}}{m_{A'}}\right). \tag{4.3}$$

As we will see, this means that the dark photon decay always occurs promptly compared the vertex resolution ($\sim 3\,\mathrm{mm}$) for the range of $m_{A'}$ and ϵ which is accessible to Mu3e but not constrained by other experiments such as NA64 [34].

Our results for the dark photon model are shown in figure 6, with and without the inclusion of the angular variables from the previous section. Without including the angular correlations, our results are in excellent agreement with prior work [19, 35] for $m_{A'} \gtrsim 30\,\mathrm{MeV}$. For $m_{A'} \lesssim 30\,\mathrm{MeV}$ we find slightly weaker sensitivity, which is due to our Monte Carlo overestimating the background in this region (See figure 14 in appendix C). Such low masses are however already disfavored by limits from NA64.

By including the angular information, we find the sensitivity to the rate, or equivalently to ε^2, can be improved by roughly a factor of two over the remaining mass range of interest. The enhanced reach means that phase I of Mu3e will be well positioned to cover the gap at low mass between NA64 and NA48. This gap will likely be tested before Mu3e phase I by dedicated experiments such as PADME [44] and the special configuration of MEG II [45]. However, both these experimental configurations are expected to quickly lose sensitivity when the dark photon mass is changed, therefore, leaving a lot of unexplored parameter space to be probed by Mu3e already in phase I.

Phase II of Mu3e will carve out significantly more parameter space competing with other complementary probes such as Belle II [40], the proposed MAGIX experiment at the MESA accelerator in Mainz [41] and the DarkQuest proposal at Fermilab [42], as well as eventually LHCb [46].

4.2 Light scalars and axion-like particles

Light scalars and pseudoscalars in the 1–$100\,\mathrm{MeV}$ mass range are ubiquitous in many extensions of the SM, ranging from dark sector models [32, 47, 48], heavy QCD axion models [49–53] to generic models containing pseudo-Nambu-Goldstone bosons associated to spontaneously broken approximate global symmetries [54–59]. Independently of the UV origin of the light scalar and its CP nature, we can ask if it is natural for it to be lighter than the muon mass and at the same time interacting with muons and/or electrons with a coupling strength large enough to be testable at Mu3e. Assuming that the couplings of the light scalar are generated by some UV physics at a scale Λ_{UV}, generically the expected one-loop correction to the light scalar mass induced by these couplings can be estimated as

$$\delta m_X \sim \frac{g_{S,A}^{\mu,e}}{4\pi}\Lambda_{\mathrm{UV}}\,.\tag{4.4}$$

Requiring the above correction to be smaller or of the same order of the light scalar mass (i.e. $\delta m_X \lesssim m_X$) we get a rough prediction for the scale at which new particles are expected to be present

$$\Lambda_{\mathrm{UV}} \lesssim 625\,\mathrm{GeV}\left(\frac{4\times 10^{-4}}{g_{S,A}^{\mu,e}}\right)\left(\frac{m_X}{20\,\mathrm{MeV}}\right)\,,\tag{4.5}$$

which shows that seeing light scalars within the reach of Mu3e is compatible with not having found heavy electroweak charged states at around or above electroweak scale.

For $2m_e < m_X < m_\mu - m_e$, the available visible decay channels are into di-lepton and di-photon pairs. The former decay width is directly controlled by the couplings introduced in eq. (1.1). For $m_X \gg m_e$ the decay width of X is given by

$$\Gamma_{X\to e^+e^-} \approx \frac{g_e^2}{8\pi}m_X\,,\tag{4.6}$$

where $g_e \equiv \sqrt{(g_S^e)^2 + (g_A^e)^2}$. The scalar decay width into diphotons is unavoidably loop-generated anytime the scalar couples to muon, resulting in the ratio of partial widths

$$\frac{\Gamma_{X\to\gamma\gamma}}{\Gamma_{X\to e^+e^-}} \approx 2\times 10^{-3}\left(\frac{m_X}{10\,\mathrm{MeV}}\right)^2\left(\frac{g_\mu/g_e}{10^3}\right)^2\,.\tag{4.7}$$

This shows that the proper decay length of X is dominated by the di-electron channel, even for a large hierarchy between g_μ and g_e. (This assumes that the diphoton width does not get significant contributions from UV physics.) Under these circumstances we can estimate the decay width of the scalar as

$$c\tau_X \approx 5\,\mathrm{mm}\left(\frac{10^{-5}}{g_e}\right)^2\left(\frac{10\,\mathrm{MeV}}{m_X}\right)\,.\tag{4.8}$$
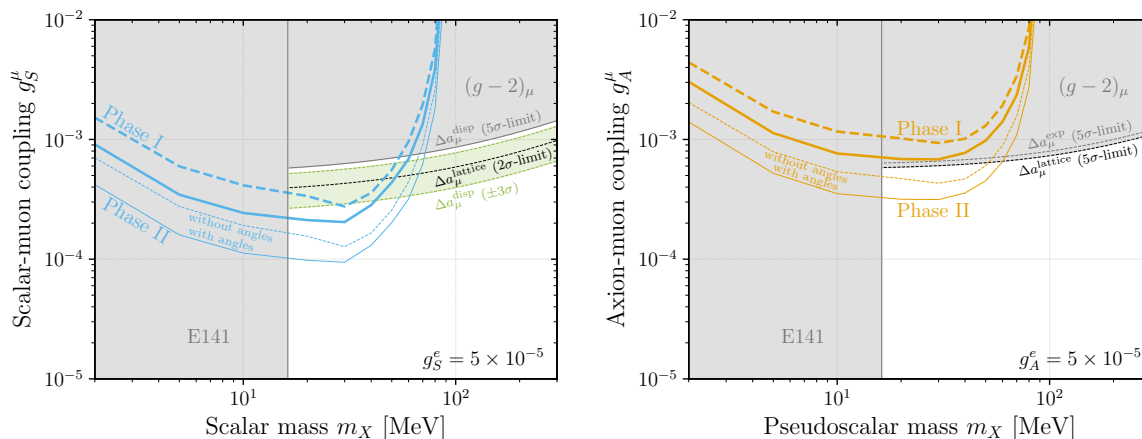
Figure 7. Parameter space of a CP-even scalar (*left*) and a CP-odd scalar (*right*) with hierarchy between the muon and electron couplings smaller than the SM mass hierarchy: $g^e_{S,A}/g^\mu_{S,A} > m_e/m_\mu$. Projections for the 95% upper limit on g^μ_S and g^μ_P are shown in **blue** and **orange**, respectively. The *thick* lines correspond to phase I of Mu3e with $2.5 \times 10^{15} \mu^+$ and the *thin* lines to the phase II with $5.5 \times 10^{16} \mu^+$. The *solid* and *dashed* lines show the reach with and without the use of angular correlations. The *gray shaded* areas show constraints from Orsay [60] and the $(g-2)_\mu$ measurement. In the scalar case the best-fit region for $\Delta a^{\mathrm{disp}}_\mu$ defined in eq. (4.11) is shown as a *green* shaded region while the limit from $\Delta a^{\mathrm{lattice}}_\mu$ defined in eq. (4.12) is indicated as a *thin-dashed black* line.

A "prompt" search at Mu3e requires the X decay to occur within 3 mm of the corresponding μ decay, which means that this strategy is limited to $g_e \gtrsim 10^{-5}$ before events begin to be lost due to vertexing. Within this range of couplings it makes sense to ask where a prompt search at Mu3e can have the largest impact. In general, the physics capabilities of Mu3e are most unique when the electron couplings of X are substantially smaller than the muon couplings. This is because stringent bounds from electron beam dumps [61, 62] probing the same region of parameter space are relaxed.

In figure 7 we show a benchmark scenario where the electron couplings are suppressed with respect to the muon, but not to the extent that they scale as the ratio of the electron mass over the muon mass ($g^e_{S,A}/g^\mu_{S,A} > m_e/m_\mu$). In this scenario the muon coupling dominates the signal branching ratio while the electron coupling is sufficiently small to not be constrained (see figure 8) but sufficiently large to produce prompt decays at Mu3e. One should keep in mind that generating the scalar interactions in eq. (1.1) with arbitrary hierarchies needs some amount of alignment in the ultraviolet theory, to avoid introducing dangerous sources of flavor violation. However, allowing for such an alignment opens up a richer coupling structure compared to the vector case.

Comparing the right and the left plot of figure 7 one can see that the reach for the pseudo-scalar is substantially worse than for the scalar, to the extent that only phase II of Mu3e will be able to probe unconstrained parameter space. This is due to an unfortunate cancellation in the muon decay amplitude for CP-odd scalars emitted from the muon line. This cancellation can be easily understood by looking at the process at the amplitude level: in the $m_X \ll m_\mu$ limit, we can factorize the light particle emission from the muon line from the rest of the process by expanding in the virtuality of the off-shell muon. In this limit the CP-odd

Scalar $(g_P^{\mu,e} = 0)$	Pseudo-scalar $(g_S^{\mu,e} = 0)$	Figure
$g_S^e = 5 \times 10^{-5}$	$g_P^e = 5 \times 10^{-5}$	Figure 7
$g_S^e = m_e/m_\mu g_S^\mu$	$g_P^e = m_e/m_\mu g_P^\mu$	Figure 8
$g_S^\mu = 0$	$g_P^\mu = 0$	Figure 9

Table 1. List of benchmarks considered for the scalar and pseudoscalar model. Figure 8 and figure 9 can be found in appendix A.

scalar emission goes to zero with respect to the scalar one because $\bar{u}(p-k)\gamma_5 u(p) \sim \mathcal{O}(k)$ whereas $\bar{u}(p-k)u(p) \sim \mathcal{O}(m_\mu)$.

Currently, the most relevant existing bounds come from the E141 beam dump [63] and the anomalous magnetic dipole moment of the muon $(g-2)_\mu$. The contribution to $(g-2)_\mu$ at one-loop are given by (see for example refs. [64, 65])

$$a_\mu^S = \frac{g_S^{\mu\,2}}{8\pi^2} \int_0^1 \mathrm{d}x \frac{(1-x)^2(1+x)}{(1-x)^2 + x\frac{m_X^2}{m_\mu^2}}, \tag{4.9}$$

$$a_\mu^P = -\frac{g_P^{\mu\,2}}{8\pi^2} \int_0^1 \mathrm{d}x \frac{x^3}{x^2 + (1-x)\frac{m_X^2}{m_\mu^2}}, \tag{4.10}$$

for the cases where X is either a CP-even scalar or a CP-odd scalar. As these contributions come with opposite signs we will treat them separately. To date, there is still a substantial tension between the SM predictions and the measurement, as well as between different theoretical predictions. To characterize this discrepancy we show for the CP even scalar case two different results: *i)* the best-fit region (at $\pm 3\sigma$) in green assuming the dispersive methods of extracting the hadronic vacuum polarization (HVP) contributions are correct [66], *ii)* the 2σ upper limit in black using the recent BMW result for HVP [67]. In formulas, we define

$$\Delta a_\mu^{\mathrm{disp}} = a_\mu^{\mathrm{exp}} - a_\mu^{\mathrm{disp}} = 249(48) \times 10^{-11}, \tag{4.11}$$
$$\Delta a_\mu^{\mathrm{lattice}} = a_\mu^{\mathrm{exp}} - a_\mu^{\mathrm{lattice}} = 105(62) \times 10^{-11}, \tag{4.12}$$

with $a_\mu^{\mathrm{disp}} = 116591810(43) \times 10^{-11}$ [66], $a_\mu^{\mathrm{lattice}} = 116592160(57) \times 10^{-11}$ and the recent FNAL update $a_\mu^{\mathrm{exp}} = 116592055(24) \times 10^{-11}$ [1] where the theory and experimental errors are added in quadrature. Due to the leading sign of the pseudoscalar case the contribution a_μ^P always increases the tension between experiment and dispersive predictions. Consequently, we show two conservative limits: *i)* the 5σ limit in black where the maximal contribution from the axion equal to the size of the BMW lattice uncertainty and *ii)* the future case in gray where the experimental uncertainty determines the allowed size of the pseudoscalar couplings again at 5σ.

Alternative scenarios for the scalar couplings are summarized in table 1 and discussed in more detail in appendix A. In a nutshell, figure 8 shows that a prompt search at Mu3e is not the best way to constrain the theoretically appealing scenarios where $g_{S,A}^e/g_{S,A}^\mu = m_e/m_\mu$. This is because the scalar X is already too long-lived in the regime that is still allowed by existing bounds. We leave an investigation of a displaced search and a comparison with the

reach of other experiments [15, 46, 68, 69] for future work. Figure 9 shows that when the muon coupling is negligible with respect to the electron coupling, the Mu3e expected reach is always surpassed by the present bounds on $\pi^+ \to e^+ \nu X$ [70] with $X \to e^+ e^-$, which benefits from the chirality suppression of the SM background $\pi^+ \to e^+ \nu \gamma^*$ with $\gamma^* \to e^+ e^-$.

5   Conclusions and outlook

In summary, Mu3e has a well-established capability to detect or constrain light $e^+ e^-$ resonances, particularly dark photons. In this study, we systematically explored alternative models that produce this signature and found that the most compelling cases are associated with the dark photon and (pseudo)scalars primarily coupling to muons. We also investigated the impact of kinematic variables, such as track angles and energy spectra, on enhancing sensitivity in a standard bump-hunt search. The inclusion of two such variables demonstrated an approximately twofold improvement in sensitivity to the $\mu^+ \to X e^+ \bar{\nu}_\mu \nu_e$ branching ratio, without requiring detector upgrades or changes in data collection strategies.

Our proposed analysis may extend to the study of exotic decays involving other Standard Model particles, where a light particle is emitted from SM muons and decays into a dilepton pair. An example of this is the $K^+ \to \mu^+ \nu X$ decays mentioned in ref. [71]. The potential reach of NA62 and future kaon facilities for these final states warrants further investigation.

However, our study has two significant limitations. Firstly, generating a high-statistics Monte Carlo background sample for a study that simultaneously bins in more than two kinematic variables, in addition to the $e^+ e^-$ invariant mass, proved challenging. Nonetheless, in practical experimental scenarios, this limitation can be addressed by extracting background estimates from sidebands in the $e^+ e^-$ invariant mass spectrum. Machine learning techniques, such as the CWoLa prescription [72], may offer effective solutions to this challenge, potentially surpassing our current estimates.

Finally, our analysis focused solely on prompt decays of X. For scalar models, especially, this limitation is significant (see appendix A). The primary sources of background in a displaced analysis are expected to arise from photon conversions within the detector material and tracks from unrelated Michel decays that cross randomly. The former can be mitigated with a veto on regions in which there is detector material, while the latter is suppressed to an extent by the timing capabilities of the detector. The primary limitation of such a search is likely to be signal acceptance, a topic we defer to future work.
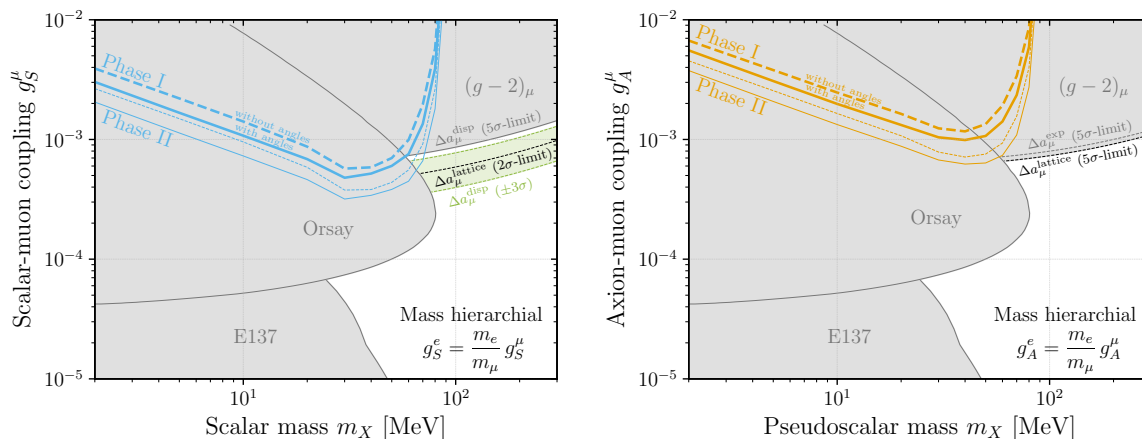
Figure 8. Parameter space of a CP-even scalar (*left*) and a CP-odd scalar (*right*) with hierarchy between the muon and electron couplings equal to the SM mass hierarchy: $g^e_{S,A}/g^\mu_{S,A} = m_e/m_\mu$. The **gray shaded** areas show constraints from beam dump experiments [61, 62] and the $(g-2)_\mu$ measurement. In the scalar case the best-fit region for $\Delta a^{\rm disp}_\mu$ defined in eq. (4.11) is shown in *green* while the limit from $\Delta a^{\rm lattice}_\mu$ defined in eq. (4.12) is indicated as a *dashed black* line. Projections for the 95% upper limit on g^μ_S **light-blue** and g^μ_A **orange** are weaker than current limits.

A Additional models

In this appendix, we provide additional information about potential models that could have been of interest for $\mu^+ \to X e^+ \bar\nu_\mu \nu_e$ but, as it turned out, were subjected to stronger constraints from other experiments.

First we further explore the parameter space of CP-even and CP-odd scalars considering the case $g^e_{S,P}/g^\mu_{S,P} = m_e/m_\mu$ in figure 8 and the one of pure electron coupling in figure 9. In figure 8 the scalar particle tends to decay displaced leading to a reduction of the signal strength for $\mu^+ \to X e^+ \bar\nu_\mu \nu_e$ with a prompt $X \to e^+ e^-$. The future Mu3e sensitivity is weaker than current beam dumps constraints from Orsay [61] and E137 [62]. The suppression is exacerbated for CP-odd scalar because of their reduced rate compared to the CP-even scalar as explained in section 4.2. In figure 9 the constraints from pion decays derived from SINDRUM data [73] are always stronger than the Mu3e sensitivity because of the chirality suppression of the corresponding SM background [70].

Second, we explore models of light vectors gauging $\mathrm{U}(1)_{L_\mu - L_e}$ whose current can be

$$J^\alpha_{L_\mu - L_e} = (L^\dagger_\mu \bar\sigma^\alpha L_\mu - \bar\mu^\dagger \bar\sigma^\alpha \bar\mu) - (L^\dagger_e \bar\sigma^\alpha L_e - \bar e^\dagger \bar\sigma^\alpha \bar e), \tag{A.1}$$

or

$$\tilde J^\alpha_{L_\mu - L_e} = (L^\dagger_\mu \bar\sigma^\alpha L_\mu + \bar\mu^\dagger \bar\sigma^\alpha \bar\mu) - (L^\dagger_e \bar\sigma^\alpha L_e + \bar e^\dagger \bar\sigma^\alpha \bar e), \tag{A.2}$$

the later of which we dub "twisted" $\mathrm{U}(1)_{L_\mu - L_e}$ in what follows. These models provide an example of a new resonance which also couples to neutrinos hence generalizing the simplified lagrangian of eq. (1.1). Figure 10 shows how the expected reach at Mu3e is weaker than the present constraints from neutrino scattering and oscillation experiments [74–76] and NA64 [77]. Note that the neutrino scattering experiments are slightly weaker than the oscillation constraints and are therefore omitted from the figure.
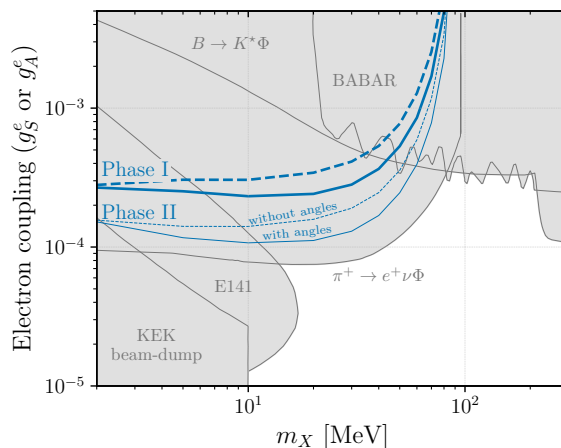
Figure 9. Parameter space of CP-even or a CP-odd scalar which couples exclusively to electrons/positrons. In this case the rates are exactly identical in the limit where the electron is massless. Projections for the 95% upper limit at Mu3e shown in **blue** are weaker than current constraints from pion decays derived in ref. [70] from SINDRUM data [73] (**gray shaded**).
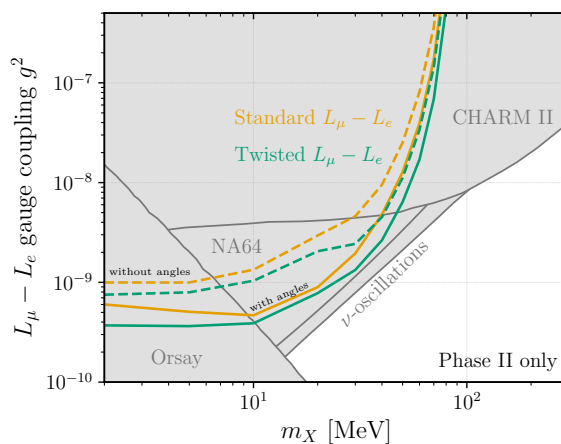

Figure 10. Parameter space of the standard and twisted $L_\mu - L_e$ gauge bosons defined in eq. (A.1) and eq. (A.2). Projections for the 95% upper limit at Mu3e shown in **orange** and **green** respectively are weaker than current constraints from NA64 [77] and neutrino scattering and oscillation experiments [76, 79, 80] (**gray shaded**).

In general the neutrino constraints (see for example ref. [78] for a summary) together with the existing constraints on the electron coupling itself tend to make the of neutrino coupling small enough to be neglected in the signal rate for $\mu^+ \to X e^+ \bar{\nu}_\mu \nu_e$ also for general scalar mediator. This observation justifies a posteriori the parametrization of eq. (1.1).

B Impact of muon polarization

Accounting for the muon polarization is potentially useful when using a muon beam obtained from pion decaying at rest, as is the case in the PSI beam line. The theoretical expectation is that these "surface" muons are 100% polarized in the opposite direction of the momentum
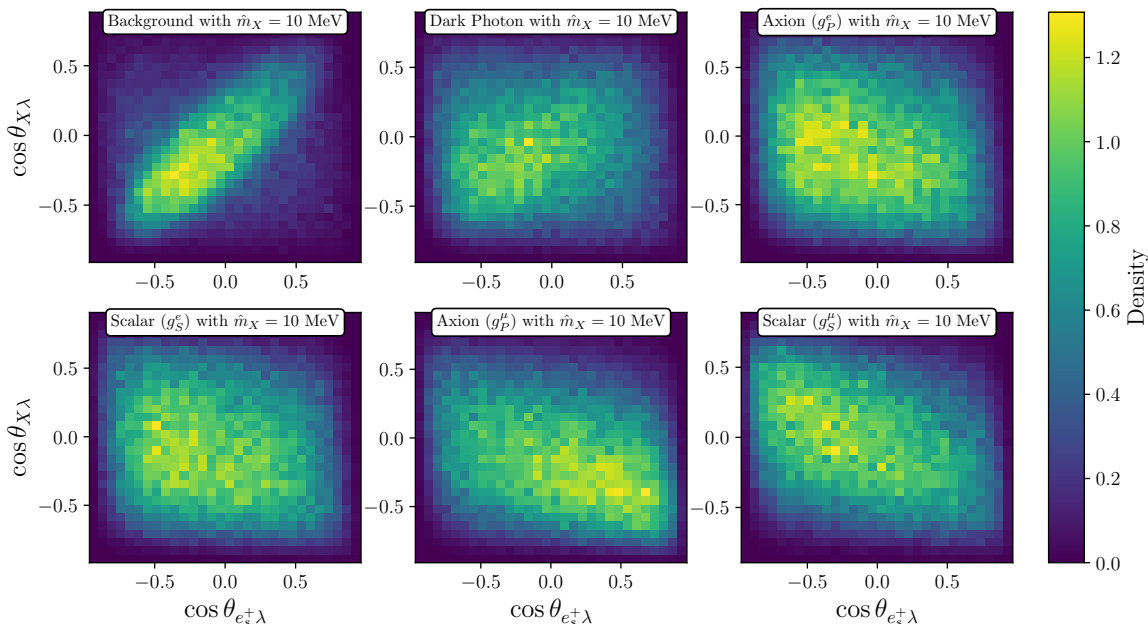
Figure 11. The distributions over the angles $\cos\theta_{e_s^+\lambda}$ and $\cos\theta_{X\lambda}$ for $\hat{m}_X = 10\,\text{MeV}$ with 10^5 simulated events for the five different signal models plus the SM background.

vector. In practice, the polarization of the stopped muons depends on depolarization effects and should be determined experimentally. As a reference for this study we use the measured averaged polarization at the MEG experiment [81] at PSI which is 86% in the opposite direction of the momentum vector. The only detail which is new when accounting for polarization is the inclusion of two additional angles; the angle between the momentum of e_s^+ and the polarization vector λ, which we call $\cos\theta_{e_s^+\lambda}$, and the angle between the momentum of X and the polarization vector, which we call $\cos\theta_{X\lambda}$. The distributions with respect to these two variables show mild differences depending on the signal model as can be seen in figure 11.

Although the dependence on the polarization is modest, it does allow one to improve the constraints, as shown in figure 12. Moreover, we expect that the correlations between those variables and the polarization dependent $\cos\theta_{e_s^+\lambda}$ and $\cos\theta_{X\lambda}$ to be fairly mild. In a full, data-driven experimental analysis we therefore recommend the inclusion of variables which are sensitive to the muon polarization, in addition to the kinematical variables discussed in section 3.

C Simulation framework and validation

The reconstructed final state consists of two positrons and an electron with four-momenta p_{p_1}, p_{p_2} and p_e, respectively. There are therefore two possible combinations comprising of an electron-positron pair, whose invariant masses are denoted with m_{ee_1} and m_{ee_2}. In section 2 our search strategy comprised of first performing a bump hunt in the m_{ee} distribution before examining the angular distributions of the positron and the reconstructed new physics state X to further extract signal from background.
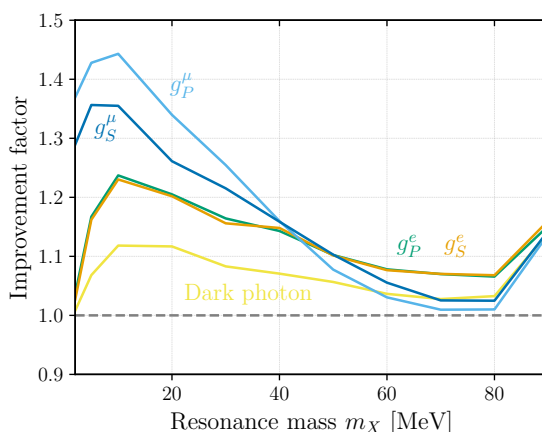
Figure 12. Improvement of the reach shown as a ratio of the rates (couplings squared) using the polarization dependent angles $\cos\theta_{e_s^+\lambda}$ and $\cos\theta_{X\lambda}$. The same models as figure 5 are shown.

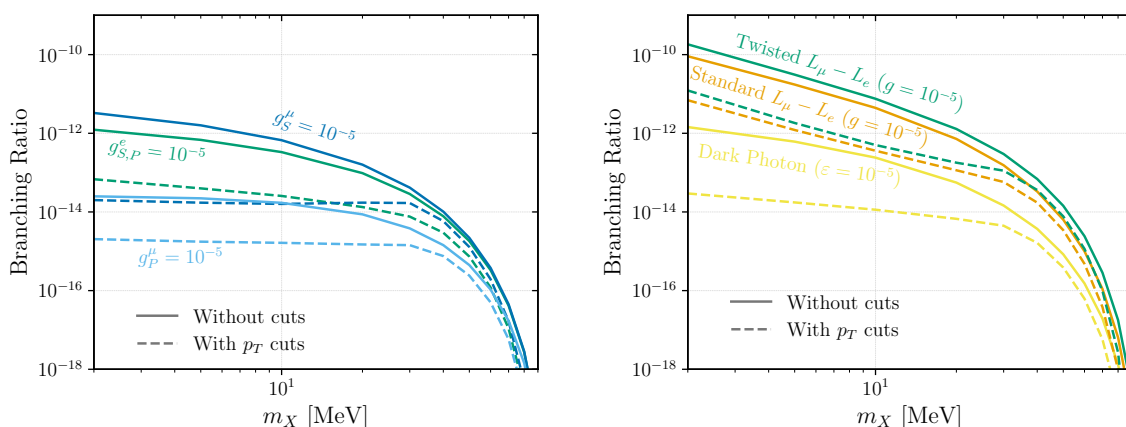**Figure 13.** The branching ratios for $\mu^+ \to Xe^+\bar{\nu}_\mu\nu_e$, $X \to e^+e^-$, with and without $p_T > 10\,\text{MeV}$ cut on all tracks, for the relevant coupling set to $g = 10^{-5}$. *Left:* Spin 0 particles with a coupling predominantly to either electrons and muons. For the electron coupling, the scalar and pseudo-scalar cases have identical branching ratios, up to m_e/m_μ suppressed corrections. For muon coupling, the branching ratio to a pseudo-scalar is suppressed relative to the branching ratio to the scalar, as explained in section 4. *Right:* Spin 1 particles, specifically the dark photon and both $L_\mu - L_e$ models. The different models are labeled the same as in figure 5 and 10.

C.1 Truth-level simulation

To simulate signal events, we used the `FeynRules 2.3` [82] to generate a `Universal FeynRules Output` [83] model file for the models defined in the previous section. We then generated events using `MadGraph5_aMC@NLO 3.4.1` [84], taking care to specify the polarization of the muon. In addition we have also analytically computed the $\mu^+ \to e^+\bar{\nu}_\mu\nu_e X$ decay rates for the various signal models (see appendix D).

With these amplitudes and our own Monte-Carlo tool, we have validated the differential decay distributions obtained with MadGraph. The branching ratios for the various models are shown in figure 13, with and without p_T cuts on the tracks.
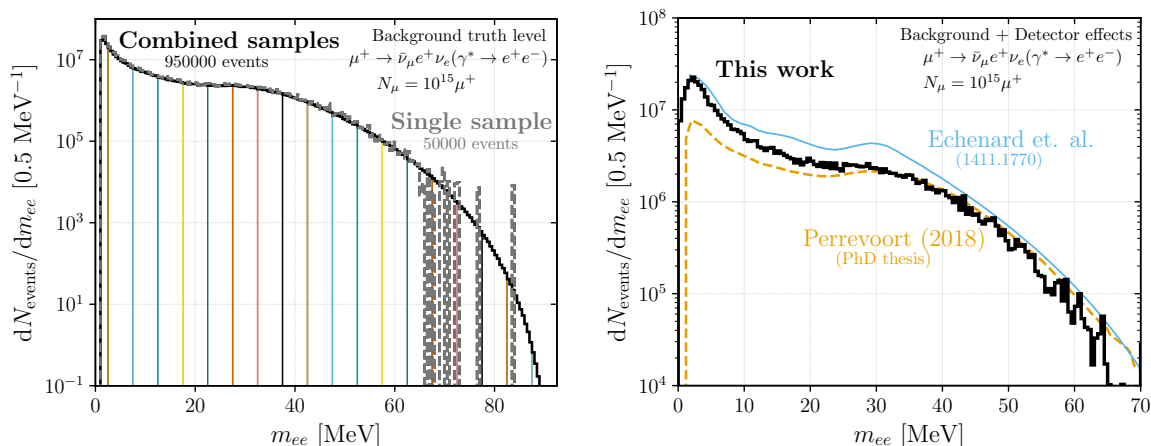
Figure 14. Invariant mass distribution $m_{ee} \equiv m_{ee,1} + m_{ee,2}$ for the dominant Standard Model background, internal conversion $\mu^+ \to \bar{\nu}_\mu e^+ \nu_e (\gamma^* \to e^+ e^-)$. (*Left:*) Validation of our phase-space slicing method to generate sufficient statistics across the full kinematic range of the invariant mass. Colored histograms are the individual samples centered around a mass hypothesis $|m_{ee,i} - m_X| \leq 2.5$ MeV, with $i = 1, 2$, while the black histogram is the sum over these samples. The gray histogram is from a single, inclusive background sample. (*Right:*) The full background m_{ee}-distribution (solid-black) including detector effects (see appendix C.2). We also show the background distributions from refs. [35] (thin-orange) and [13] (thin-blue). For both panels we take the total number of muon decays $N_\mu = 10^{15}$.

The dominant background is the internal conversion process $\mu^+ \to e^+ e^- e^+ \bar{\nu}_\mu \nu_e$ [13, 35], which we also simulate with MadGraph. (See ref. [21] for the relevant amplitude and differential decay rates of this process.) In order to perform the search over the full range of kinematically accessible masses m_X, and study the subsequent kinematic variables in section 3, a sufficiently large number of events must be generated. This is largely through the difficulty in sampling the high-mass endpoint of the m_{ee} distribution as this corresponds to an extreme configuration in five-body phase space where the neutrinos are very soft. It is therefore not feasible to generate a single, statistically satisfactory background sample. Instead we generate a separate background sample for each mass hypothesis of the dark sector particle (m_X). We do so by demanding that either $|m_{ee,1} - m_X| \leq 2.5$ MeV or $|m_{ee,2} - m_X| \leq 2.5$ MeV at the generator level, and weight the sample with its corresponding fiducial width. This cut is not available in the standard MadGraph cards and thus requires a modification to the MadGraph code responsible for imposing kinematic cuts. Our implementation can be found on GitHub ⟲, while in the left-hand side panel of figure 14 we validate this approach. Here we show in grey-dashed the m_{ee} distribution (including both possible $e^+ e^-$-pairs) for a single run of MadGraph generating 50000 events. At large invariant mass there is a clear degradation in the statistics due to the extreme configuration of phase-space that is required. This is overcome through the above phase-space slicing, shown as the colored histograms, producing a smooth distribution even towards the kinematic endpoint at large m_{ee}. Lastly we note that for all the background and signal samples we place a generator cut of $p_T \geq 10$ MeV on all the electrons/positrons.

Muons produced at the πE5 beam line at PSI are roughly 86% polarized as measured by the MEG experiment [81]. Since the muons at Mu3e are expected to be similarly polarized, we
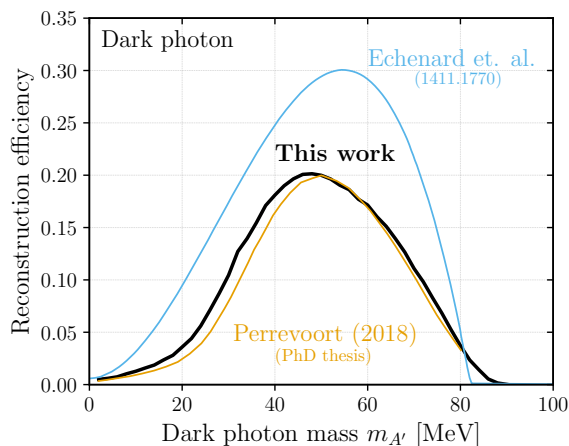
Figure 15. Reconstruction efficiency of the signal events as a function of the dark photon mass. For comparison we also show the efficiency curves of refs. [35] (**thin-orange**) and [13] (**thin-light-blue**).

Cuts	Phase space only			Internal conv. bkg.		Dark photon $(m_{A'} = 40\,\text{MeV})$	
	Ref. [35]	Ref. [6]	This work	Ref. [35]	This work	Ref. [35]	This work
$p_T \geq 12\,\text{MeV}$	–	–	73%	–	2.7%	–	64%
			(73%)		(2.8%)		(65%)
Geometric acceptance (short tracks)	43%	38%	62%	–	1.9%	–	47%
			(85%)		(67%)		(73%)
Rescaling factor	–	–	38%	–	1.1%	–	29%
			(61%)		(61%)		(61%)
Track reconstruction (short tracks)	39%	34%	32%	0.51%	0.68%	18%	19%
	(89%)	(90%)	(83%)	(0.51%)	(59%)	(18%)	(68%)
Vertex quality and timing	26%	$\sim 28\%^a$	29%	0.46%	0.63%	17%	18%
	(93%)	(82%)	(93%)	(90%)	(93%)	(94%)	(93%)
Final efficiency	36%	28%	29%	0.46%	0.63%	17%	18%

^{a}Extracted assuming long-track efficiency for vertexing

Table 2. Cut-flow table for the various factors affecting the final detector efficiency. The percentages given are the running total efficiency while the numbers in parentheses are the percentage of events lost relative to the previous step. Note that here we choose to apply an additional rescaling factor to match the geometric acceptance, which is tuned to the flat phase space events and validated against both signal and background efficiencies. We use a hyphen ($-$) to denote steps that either are not applicable (the p_T cut or rescaling factor for refs. [6, 35]) or where information is absent. See text for additional details and figure 15 for signal efficiency at alternative dark photon masses. Lastly, due to the ad hoc rescaling factor applied here, statistical uncertainties are not quoted as we expect systematic uncertainties to dominate.

randomly mirror 7% of the events in our samples along the polarization axis of the muon. We also assume the muons decay uniformly at rest within the target material of the experiment.

C.2 Detector modeling

The Mu3e experiment consists of four layers of pixel trackers around a large conical target in which the muons are stopped, along with timing layers to assist in background rejection. It

has sizeable angular coverage for electron/positron tracks with transverse momenta larger than approximately 10 MeV, while the neutrinos escape as missing energy.

Mu3e separates the observed charged leptons into *short tracks* and *long tracks*. Roughly, the difference is that short tracks pass through four layers of pixel trackers, be that traversing the detector once, or for low p_T tracks passing only through the two inner pixel layers before being bent back where they traverse back through the same layers again. For long tracks at least two more hits in the pixel layers are required. Effectively, these additional hits allow for a more precise measurement of their momentum vector as compared to that of the short tracks. The track reconstruction efficiencies for short track and long tracks have been simulated by the Mu3e collaboration, and are parametrized as a function of the total track momentum (**p**) and its angle with the beam axis (θ) (figures (19.1) and (19.4) in ref. [6]). These parametrizations describe the probability that a track is reconstructed, *if* it passed through at a tracking layer at least four times. It therefore does not include the geometric acceptance, which is defined as the probability of passing through the pixel layers at least four times. We must model this separately, as we describe below.

To model the geometric acceptance we first begin by placing a p_T cut of 12 MeV on the electron/positron tracks, which rejects all charged leptons which would not make it through the first two pixel layers given the uniform 1 T magnetic field in the detector. The remaining loss of tracks occurs as the length of the target in the beam-direction is similar to that of the inner pixel tracker layers. Consequently if a muon is stopped at the start/end of the target there is a sizeable probability that the lower p_T tracks are sufficiently bent to miss the start/end of the inner pixel layer. For muons that decay more centrally this probability is effectively zero as all that matters is the gyro-radius of the track, which is accounted for by the p_T cut. To account for these effects we firstly use the aforementioned track reconstruction map in **p** versus θ, however, rather than use the z-axis which gives the probability of reconstructing a track, we use the non-zero region to build an approximate acceptance map of the detector. The resulting efficiency, quoted in the second line in table 2, is effectively the "best case" efficiency, for a muon decaying at the center of the detector. This is therefore an overestimate as it does not yet account for the extended size of the target. To address this remaining discrepancy we apply an additional efficiency factor to match the result of the Mu3e collaboration. Fortunately Mu3e provides this factor for Monte Carlo events uniformly distributed in phase space, see table (22.1) of ref. [6]. This ends up being an additional factor of 61.1% in our reconstruction efficiency.[1] Once we tuned our efficiency with a uniform phase space distribution, we validate it on both our signal and background events, see the additional rows of table 2 which show good agreement with those stated in ref. [35]. After this step we can then apply the track reconstructions efficiencies mentioned above, through the digitization of figures (19.1) and (19.4) from ref. [6]. The latter of which governs the probability that a short track event can be 'upgraded' to a long track, which plays an important role in smearing below. The final step involves a number of cuts selecting for quality of the vertex and timing in a given event. We cannot simulate these effects, but instead rely on statements in ref. [35] where its is claimed that timing is $\sim 98\%$ efficient, while vertex cuts are $\sim 95\%$ efficient. A complete breakdown of these steps and

---

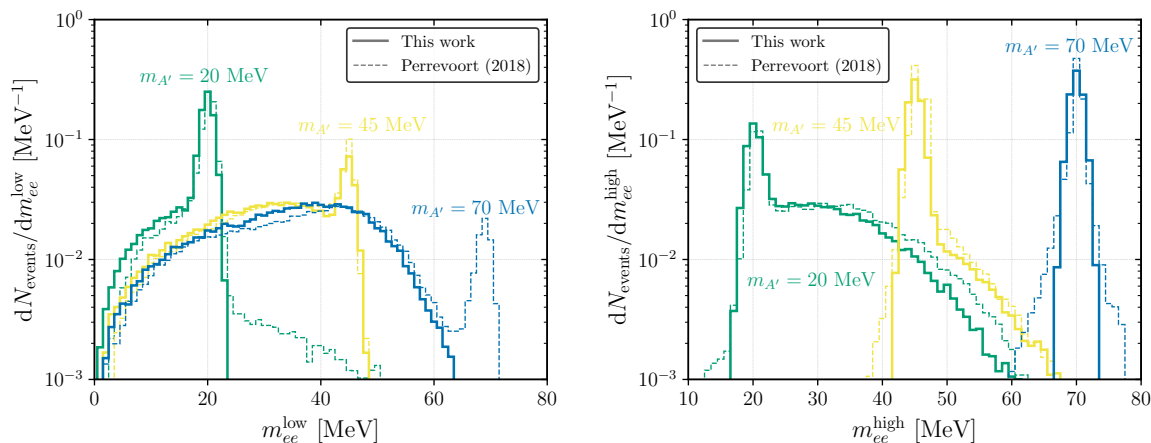[1]See also table (4.1) of ref. [35] for a similar value.

Figure 16. Signal invariant mass distributions for the dark photon signal model. To compare to ref. [35] (dashed lines) we show the low and high mass combinations of electron-positron pairs for $m_{A'} = 20, 45, 70 \, \mathrm{MeV}$ (green, yellow, blue).

comparisons to both refs. [6, 35] is given in table 2, where a hyphen is used to indicate steps where specific information is either not given by the experimental collaboration or is not relevant for them as it included automatically in their full detector simulation. Here we see that tuning our analysis pipeline on the flat phase-space events yields good agreement when applied to both signal and background events. This is especially apparent in figure 15 which shows the efficiency for the dark photon signal model over a range of dark photon masses.

Importantly, the electron-positron pair invariant mass (m_{ee}) resolution depends not only on the momentum resolution but also on the angular resolution. In fact, it is this angular resolution which dominates the invariant mass resolution. Unfortunately, the Mu3e Technical Design Report (TDR) [6] does not give the angular resolution of the tracks, but the invariant mass resolution for the positron-electron pair is reported in ref. [35] as a function of the number of long tracks and the invariant mass value. We therefore use this result in our analysis, but since we are interested in angular correlations, our analysis would benefit from information about the angular resolution which we are unfortunately unable to incorporate. To determine the smearing in the m_{ee} distribution we use figure (6.12) in ref. [35], drawing from a Gaussian distribution with a width given by the reported RMS value which depends on the number of reconstructed long tracks. Note that we have implemented both m_{ee} dependent and independent smearing (6.12a) versus (6.12b) of ref. [35], however, there is no appreciable difference between the two for the application at hand. To validate the smearing we show the distributions for both signal (figure 16) and background (right-hand panel of figure 14). For the background we see that there are sizeable differences at low-invariant masses. This discrepancy is larger than the mismatch in the efficiencies given in table 2, and likely points to additional effects not included in the rescaling factor to match the geometric acceptance. However, given that we over-predict background events in this region, our limits will simply be conservative at the bottom end of the mass range. For the signal smearing (figure 16) we see good agreement around the resonance peaks between the solid (our results) and dashed lines (ref. [35]) for the different combinations of the electron-positron pairs invariant mass.

However, away from the peaks we see that our distributions are narrower compared to ref. [35]. This is because we assumed Gaussian smearing, when in reality a Crystal ball function with fatter tails would better describe the smearing, but it is not possible to implement given the information available. This also leads to the absence of a peak in the low-m_{ee} combination for $m_{A'} = 70\,\mathrm{MeV}$. Essentially a low-probability fluctuation is responsible for causing a signal event to be sufficiently smeared to appear in the low-invariant mass m_{ee} distribution, which given our Gaussian modeling never occurs for our signal samples of 10^5 events.

C.3 Additional analysis details

We are interested in performing a hypothesis test on each signal, parametrized by (m_X, g), with g representing the coupling governing the signal strength for the model of interest $(\varepsilon, g_{S,P}^{\mu,e}, g_{V,A}^{\mu,e})$ where the two relevant hypotheses are:

- $H_0 = \{$**Background only hypothesis:** Standard Model only$\}$

- $H_1 = \{$**Signal hypothesis:** Signal model described by parameters (m_X, g) exists$\}$

To test these hypotheses we define $\hat{m}_X$ as the invariant mass of the electron-positron pair closest to the hypothesis mass m_X and take events in a window $m_X \pm \delta$ around this resonance. The width of this window, δ, is chosen through maximization of the test statistic $S/\sqrt{B}$, which for our samples resulted in $\delta = 1.5\,\mathrm{MeV}$. The events that satisfy this cut are then binned in variables introduced in section 3.

As we are performing a binned likelihood analysis, the maximum number of angular bins we can consider faces a Monte Carlo limitation arising from our ability to sample these distributions across the entire domain of the angles while maintaining a satisfactory number of event per bin. Practically speaking the number of bins is therefore limited by the size of our Monte Carlo samples. This is primarily an issue for the background samples and is the key motivation for the splicing in phase-space. To demonstrate that our samples are sufficiently large, we show in figure 17 the improvement in reach from our angular analysis as a function of the number of bins per angle for several signal models. To ensure that bins with insufficient statistics are not driving the gain in sensitivity we discard bins from the likelihood that contain less than 10 (5) background Monte Carlo events with non-zero weights and at least one signal event; the corresponding increase in sensitivity is shown by the thick solid (thick dashed) curves. On the opposite y-axis we also show with thin lines the percentage of bins that are thrown away given this criteria. From the behavior we see that the gain in sensitivity quickly increases with a modest number of bins per angle before plateauing. Further improvement requires extremely fine binning with a relaxation of the criteria for the number of events with non-zero weights per bin. We therefore conclude our likelihood captures most of the relevant physics provided that we use ~ 20 or more bins per variable. For the results presented in the main body of the paper a filled circle indicates our chosen binning (30 bins per angle), which we see is conservative compared to the finer binning results with fewer events per bin. For the (pseudo)scalar case coupled to electrons relaxing the event requirement per bin and increasing the number of bins from 30 to 60 results only in a 1.5% improvement in the reach.
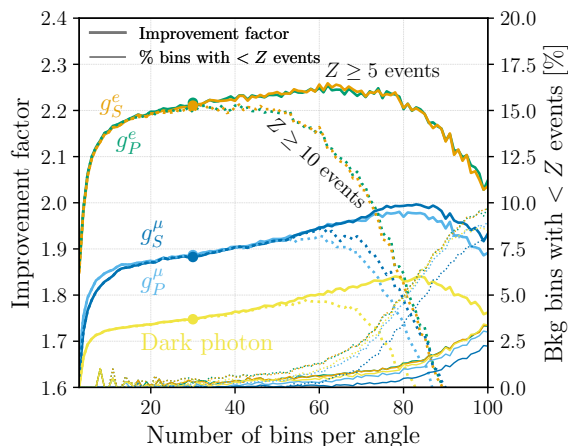
Figure 17. Improvement factor including angular information in the reach projection as a function of the total number of bins in each angular direction for $m_X = 40\,\text{MeV}$ and the angular observables $\cos\theta_{e_s^+ e_X^+}$ and $\cos\theta_{e^- e_s^+}$. The thick solid (dotted) lines indicate the improvement factor requiring at least 10 (5) Monte Carlo events with non-zero weights per bin for the background given non-zero signal events. The thin lines (with associated y-axis on the right-hand side) indicate the percentage of bins rejected (and therefore not contributing to the likelihood) given this requirement. The colors indicate the coupling structure of the signal model.

D Signal differential decay rate formulas

In this appendix we will give some analytic expressions for the differential decay rate of $\mu^+ \to e^+ \nu_e \bar{\nu}_\mu X$. We first discuss the scenario where the muon polarization is ignored, and later consider the effect of including polarization.

D.1 Unpolarized muon decay

The differential decay rate is:

$$\mathrm{d}\Gamma = \frac{\langle|\mathcal{M}|^2\rangle}{2m_\mu}(2\pi)^4\delta^{(4)}(p_\mu - \sum_f p_f)\prod_f \frac{\mathrm{d}^3 p_f}{(2\pi)^3 2E_f} \tag{D.1}$$

where the index f iterates over final state particles and $\langle|\mathcal{M}|^2\rangle$ is the helicity averaged square amplitude. To lighten the notation, we will also use the symbol of the particle to denote the four momentum in this section. E.g. "p_ν^α" will be abreviated as "$\bar{\nu}^\alpha$" etc. The neutrinos are not observable, and so it is beneficial to integrate over their phase space. This process is aided by the fact that the squared amplitude for this process can always be written as

$$\left\langle|\mathcal{M}|^2\right\rangle = \nu_\alpha \bar{\nu}_\beta \mathcal{A}^{\alpha\beta}\,, \tag{D.2}$$

where $\mathcal{A}^{\alpha\beta}$ does not depend on either ν or $\bar{\nu}$. We can then factorize the final state phase space measure as

$$\delta^{(4)}(\mu - \sum_f f) = \frac{1}{2}\int \mathrm{d}^4 Z \delta^{(4)}\left(Z - \nu - \bar{\nu}\right)\delta^{(4)}\left(\mu - Z - e - X\right)\,. \tag{D.3}$$

Using the first delta function, we can integrate over the neutrinos using the identity

$$\int \frac{\mathrm{d}^3\nu}{2E_\nu} \frac{\mathrm{d}^3\bar{\nu}}{2E_{\bar{\nu}}} \delta^{(4)}\left(Z - \nu - \bar{\nu}\right) \nu^\alpha \bar{\nu}^\beta = \frac{\pi}{24}\left[\eta^{\alpha\beta} Z^2 + 2Z^\alpha Z^\beta\right].$$ (D.4)

The second delta function can be removed by integrating over $\mathrm{d}^4 Z$ which results in fixing $Z \to \bar{Z} = \mu - e - X$. The result of this process is the following:

$$\mathrm{d}\Gamma = \frac{\pi}{24}\left[\eta_{\alpha\beta} \bar{Z}^2 + 2\bar{Z}_\alpha \bar{Z}_\beta\right] \frac{\mathcal{A}^{\alpha\beta}}{2m_\mu(2\pi)^8} \frac{\mathrm{d}^3 e}{2E_e} \frac{\mathrm{d}^3 X}{2E_X}.$$ (D.5)

Then defining the Lorentz invariant dimensionless quantities

$$x \equiv 2(\mu \cdot e)/m_\mu^2,$$ (D.6)

$$y \equiv 2(\mu \cdot X)/m_\mu^2,$$ (D.7)

$$z \equiv 2(e \cdot X)/m_\mu^2,$$ (D.8)

$$r \equiv (m_X/m_\mu)^2.$$ (D.9)

This decay rate can therefore be written as

$$\mathrm{dBR}(\mu \to \nu\bar{\nu}eX) = \left\{\frac{F(x,y,z)}{(r+z)^2} + \frac{G(x,y,z)}{(r-y)^2} + \frac{H(x,y,z)}{(r+z)(r-y)}\right\} \frac{\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z}{8\pi^2 r},$$ (D.10)

where we divided by the muon width $\Gamma_\mu = G_F^2 m_\mu^5/192\pi^3$ and used the result (valid when polarization is not taken into account)

$$\frac{\mathrm{d}^3 e}{2E_e} \frac{\mathrm{d}^3 X}{2E_X} = \frac{\pi^2}{4} m_\mu^4 \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z.$$ (D.11)

Here the function F corresponds to the square of the diagram with an electron propagator, G corresponds to the diagram with a muon propagator, and H corresponds to the interference term.

However, it is useful to now consider the rest frame of the muon where we can write $z = \frac{1}{2}xy(1 - vc)$, where $v = |\boldsymbol{X}|/E_X = \sqrt{1 - 4r/y^2}$ is the speed of particle X in the muon rest frame and $c = \boldsymbol{X} \cdot \boldsymbol{e}/(|\boldsymbol{e}||\boldsymbol{X}|)$ is the cosine of the angle between the momentum of e and X in the muon rest frame. This angle is physically intuitive and simplifies the limits of integration. In particular, if we consider $v = v(y)$ and $z = z(x,y,c)$, then the branching ratio becomes

$$\mathrm{dBR}(\mu \to \nu\bar{\nu}eX) = \frac{vxy}{16\pi^2 r} \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}c \left\{\frac{F(x,y,z)}{(r+z)^2} + \frac{G(x,y,z)}{(r-y)^2} + \frac{H(x,y,z)}{(r+z)(r-y)}\right\},$$ (D.12)

where the limits of integration are given by

$$c \in (-1, 1),$$ (D.13)

$$y \in \left(2\sqrt{r}, 1 + r\right),$$ (D.14)

$$x \in \left(0, \frac{4 - 4y + y^2\left(1 - v^2\right)}{4 - 2y + 2cyv}\right).$$ (D.15)

For the spin-1 model described by eq. (1.1) we get the following results:

$$F_1 = (g_V^e - g_A^e)^2 \Big[z^2(z(2x+y-2)+x(-2x-2y+3))$$
$$+ rz\left(-4x^2+x(-8y+7z+6)-4y^2+6y(z+1)-8z\right)$$
$$+ r^2(z(4x+5y-6)+x(2x+2y-3))-r^3x \Big],$$
(D.16)

$$G_1 = g_V^{\mu\,2}\Big[4y^2(z(2x+y-2)+x(-2x-2y+3))$$
$$+ r\left(-8(y+2)z(4x+3y-3)+4x(4x(y+2)+y(7y+6)-12)+16(y+2)z^2\right)$$
$$+ r^2\left(8x^2-4x(4y+2z+15)+4(5y+14)z\right)-4r^3x\Big]$$
$$+ g_A^{\mu\,2}\Big[4y^2(z(2x+y-2)+x(-2x-2y+3))$$
$$+ r\left(4xy^2+8(y-4)(x-z)(x+y-z-1)+8(y-4)(x-z)(x+2y-z-2)\right)$$
$$+ r^2(4x(2x-4y+21)-4z(2x-5y+22))-4r^3x\Big]$$
$$+ g_V^\mu g_A^\mu\Big[4y\left(-(y+2)z(2x+y)+xy(2x+2y-1)+4z^2+2z\right)$$
$$+ r\left(8z(x(4y-2)+y(3y-2)+1)-4xy(4x+7y-4)-16(y-1)z^2\right)$$
$$+ r^2(4z(2x-5y+4)-4x(2x-4y+1))+4r^3x\Big],$$
(D.17)

$$H_1 = (g_V^e - g_A^e)g_V^\mu\Big[2yz(z(2x+y-2)+x(-2x-2y+3))$$
$$+ 2r(z(-6x^2-(x+6)y+x+y^2+6)+z^2(2x-y+2)$$
$$+ x(2x+y)(2x+2y-3))+2r^2(z(7x-y+4)+x(-10x-7y+7))+10r^3x]$$
$$+ (g_V^e - g_A^e)g_A^\mu\Big[-2(z((y+2)z(2x+y)+xy(-2x-2y+1)-4z^2-2z))$$
$$- 2r(z(-6x^2-x(y-5)+y(y+2)-4)+z^2(2x-y-8)$$
$$+ x(2x+y)(2x+2y-1))+2r^2(10x^2+x(7y-7z-3)+(y+2)z)-10r^3x\Big].$$
(D.18)

While for the spin-0 model described by eq. (1.1) we get the following results:

$$F_0 = (g_S^e - g_P^e)^2 r\Big[\left(z^2(x+2y-2)-2yz(x+y)+3yz\right)+r^2(2(y-1)z-x(-2x-2y+3))-r^3x\Big],$$
(D.19)

$$G_0 = g_S^{\mu\,2}r\Big[2(8(-2x^2+x(-2y+2z+3)+(y-2)z)+2(y-4)(-2z(x+y)+xy+2z^2+z))$$
$$+ 4r(2x^2-2xz+9x+2yz-10z)-4r^2x\Big]$$
$$+ g_P^{\mu\,2}r\Big[4y(-2z(x+y)+xy+2z^2+z)+4r(2x^2-2xz-3x+2yz+2z)-4r^2x\Big]$$
$$+ g_S^\mu g_P^\mu r\Big[((y-4)(-2z(x+y)+xy+2z^2+z)+y(-2z(x+y)+xy+2z^2+z))$$
$$+ 2r(2x^2-2xz+x+2yz-4z)-2r^2x\Big],$$
(D.20)

$$H_0 = 0,$$
(D.21)

D.2 Polarized muon decay

To describe the decay of a polarized muon we define the covariant spin vector

$$S = \left(\frac{|\boldsymbol{\mu}|}{m_\mu}, \frac{E_\mu \boldsymbol{\mu}}{m_\mu |\boldsymbol{\mu}|} \right) , \tag{D.22}$$

satisfying $S^2 = -1$ and $S \cdot \mu = 0$ and the helicity operator

$$H = \frac{1}{2} \gamma_5 \slashed{S} . \tag{D.23}$$

In the rest frame of the muon, the covariant spin vector can be taken to be $S \to (0, \hat{\boldsymbol{z}})$ where $H = \frac{1}{2} \gamma_5 \gamma_3$.

We define the squared amplitude coming from the decay of a muon with helicty $h = \pm 1/2$ as $|\mathcal{M}_h|^2$. One simple way of calculating this squared amplitude is by inserting a helicty projection operator in from of the muon spinor

$$v(\boldsymbol{\mu}) \to \left(\frac{1}{2} \pm H \right) v(\boldsymbol{\mu}) . \tag{D.24}$$

We say a muon has polarization $P_\mu \in (-1, 1)$ when the muon's helicity has a expectation value $\langle H \rangle = 2 P_\mu$. The polarization averaged amplitude with polarization P_μ will then be given by

$$\langle |\mathcal{M}| \rangle_{P_\mu} \equiv \frac{1 + P_\mu}{2} |\mathcal{M}_{+1/2}|^2 + \frac{1 - P_\mu}{2} |\mathcal{M}_{-1/2}|^2 . \tag{D.25}$$

It is this value which replaces $\langle |\mathcal{M}|^2 \rangle$ in eq. (D.1).

Introducing the additional four vector S also introduces two additional angles. We can define these as

$$c_e = \frac{2}{m_\mu x} (S \cdot e) , \tag{D.26}$$

$$c_X = \frac{2}{m_\mu v y} (S \cdot X) , \tag{D.27}$$

which are just the cosines of the angle of the $\boldsymbol{e}$ and $\boldsymbol{X}$ with the polarization vector in the muon rest frame.

We would then like to write down the generalization of the expression eq. (D.12). This requires two modifications; first the F, G, and H functions must be modified to include polarization, second we must ensure not to integrate over the two additional angles related to polarization.

First we discuss the generalization of F, G, and H. This is done simply by modifying the functions to be

$$\begin{aligned} F(x, y, z) \to F^{(P_\mu)}(x, y, z, c_e, c_X) &= F(x, y, z) \\ &+ P_\mu c_e F^{[c_e]}(x, y, z) + P_\mu v c_X F^{[c_X]}(x, y, z) , \end{aligned} \tag{D.28}$$

and similarly for G and H.

The measure, $\mathrm{d}\Omega_e\,\mathrm{d}\Omega_X$ can only be reduced to three variables. We will therefore use the angles defined by

$$\hat{\boldsymbol{e}} = (s_e, 0, c_e)\,, \tag{D.29}$$

$$\hat{\boldsymbol{X}} = (s_X \cos\varphi, s_X \sin\varphi, c_X)\,, \tag{D.30}$$

where $s_i = \sqrt{1 - c_i^2}$. In this coordinate system we have

$$\frac{\mathrm{d}^3 e}{2E_e}\frac{\mathrm{d}^3 X}{2E_X} = \frac{\pi m_\mu^4 vxy}{32}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}c_e\,\mathrm{d}c_X\,\mathrm{d}\varphi\,, \tag{D.31}$$

such that the differential branching ratio formula becomes

$$\mathrm{dBR}(\mu \to \nu\bar{\nu}eX) =$$

$$\left\{\frac{F^{(P_\mu)}(x,y,z,c_e,c_X)}{(r+z)^2} + \frac{G^{(P_\mu)}(x,y,z,c_e,c_X)}{(r-y)^2} + \frac{H^{(P_\mu)}(x,y,z,c_e,c_X)}{(r+z)(r-y)}\right\}\frac{vxy}{64\pi^3 r}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}c_e\,\mathrm{d}c_X\,\mathrm{d}\varphi. \tag{D.32}$$

With these variables we have

$$z(x,y,c(c_X,c_e,\varphi)) = \frac{1}{2}xy(1 - vc(c_X,c_e,\varphi))\,, \tag{D.33}$$

$$c(c_X,c_e,\varphi) = s_e s_X \cos\varphi + c_e c_X\,, \tag{D.34}$$

and the limits of integration will be the extension of the previous case, namely

$$\begin{aligned}
\varphi &\in (0, 2\pi)\,, & y &\in \left(2\sqrt{r}, 1+r\right)\,, \\
c_e &\in (-1, 1)\,, & x &\in \left(0, \frac{4 - 4y + y^2\left(1 - v^2\right)}{4 - 2y + 2c(c_X,c_e,\varphi)yv}\right)\,, \\
c_X &\in (-1, 1)\,,
\end{aligned} \tag{D.35}$$

The additional components of F, G, and H for a spin-1 particle are

$$F_1^{[c_e]} = (g_V^e - g_A^e)^2\,x\left[z^2(1 - 2x - y + 2z) + rz(-4x - 4y + 7z + 2) + r^2(2x + y + 4z - 1) - r^3\right]\,, \tag{D.36}$$

$$F_1^{[c_X]} = (g_V^e - g_A^e)^2\,y\left[z^2(z - x) + 2rz(-2x - 2y + 3z + 1) + r^2(x + 5z)\right]\,, \tag{D.37}$$

$$G_1^{[c_e]} = x(-r + 2x + y - 2z - 1)\left[g_V^{\mu\,2}\left(r(r+4) - y^2\right) + g_A^{\mu\,2}(r - y)^2 + 2g_V^\mu g_A^\mu r(y - r)\right]\,, \tag{D.38}$$

$$\begin{aligned}
G_1^{[c_X]} &= g_V^{\mu\,2}4y\left[(z - x)\left(5r^2 + r(-4x - 6y + 4z + 2) + y^2\right)\right]\,, \\
&\quad + g_A^{\mu\,2}4y\Big[z\left(5r^2 - 2r(6x + 3y + 1) + 4x(y + 2) + y(y + 4) - 4\right) + \\
&\qquad x\left(-7r^2 + 2r(4x + 5y - 2) + y(-4x - 3y + 2)\right) + 4(r - 2)z^2\Big] \\
&\quad + g_V^\mu g_A^\mu 2y\left[-z\left(2y(-3r + x - 1) + r(5r - 10x + 8) + y^2\right) + x(3r - y)(2r - 2x - 2y + 3) - 4rz^2\right]\,, \tag{D.39}
\end{aligned}$$

$$H_1^{[c_e]} = 2(g_V^e - g_A^e)x \times$$
$$\left[g_V^\mu \left\{ z\left(-7r^2 + 6rx + r + y(2x + y - 1)\right) + r\left(-5r^2 + 5r(2x + y - 1) - 4x(x + y) + 2x\right) - 2z^2(r + y) \right\} \right.$$
$$\left. + g_A^\mu r\left\{ 5r^2 + r(-10x - 10y + 7z + 5) + 4x^2 + 2x(5y - 3z - 3) - 7yz + 5(y - 1)y + 2z^2 + 3z \right\} \right],$$
$$(D.40)$$

$$H_1^{[cx]} = 2(g_V^e - g_A^e)y \times$$
$$\left[g_V^\mu \left\{ r^2(2x + z) + r\left(-2x^2 + x(-2y + z + 1) + z(-y + z + 2)\right) + yz(x - z) \right\} \right.$$
$$\left. g_A^\mu \left\{ z\left(-r^2 + r(6x + y) - 2x(x + y) + 3x\right) + z^2(-r + 2x + y - 2) + rx(3r - 4x - 3y + 2) \right\} \right].$$
$$(D.41)$$

For a spin-0 particle we have

$$F_0^{[c_e]} = (g_V^e - g_A^e)^2 rx \left(z(z - y) + r(2x + y - 1) - r^2 \right), \tag{D.42}$$

$$F_0^{[cx]} = (g_V^e - g_A^e)^2 ry(r(x + 2z) + z(-x - 2y + 2z + 1)), \tag{D.43}$$

$$G_0^{[c_e]} = 2rx(1 + r - 2x - y + 2z) \left[2g_V^{\mu\,2}(r - 2y + 4) + 2g_A^{\mu\,2}r + g_V^\mu g_A^\mu(r - y) \right], \tag{D.44}$$

$$G_0^{[cx]} = 2ry\left[g_V^{\mu\,2} \left\{ 2x(3r - 3y + 6z + 2) - 4z(r - y + z) - 8x^2 - 2z \right\} \right.$$
$$g_A^{\mu\,2} \left\{ 2r(x - 2z) - 2\left(-2z(x + y) + xy + 2z^2 + z\right) \right\}$$
$$\left. g_V^\mu g_A^\mu \left\{ (x - z)(2r - 2x - 2y + 2z + 3) \right\} \right], \tag{D.45}$$

$$H_0^{[c_e]} = 0, \quad H_0^{[cx]} = 0. \tag{D.46}$$

References

[1] 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].

[2] MU2E collaboration, *Mu2e Technical Design Report*, `arXiv:1501.05241` [`DOI:10.2172/1172555`] [INSPIRE].

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

[4] N. Teshima, *Status of the DeeMe Experiment, an Experimental Search for μ-e Conversion at J-PARC MLF*, *PoS* **NuFact2019** (2020) 082 [`arXiv:1911.07143`] [INSPIRE].

[5] MEG II collaboration, *The design of the MEG II experiment*, *Eur. Phys. J. C* **78** (2018) 380 [`arXiv:1801.04688`] [INSPIRE].

[6] MU3E collaboration, *Technical design of the phase I Mu3e experiment*, *Nucl. Instrum. Meth. A* **1014** (2021) 165679 [`arXiv:2009.11690`] [INSPIRE].

[7] Mu3e collaboration, *The Mu3e Experiment*, in the proceedings of the *Snowmass 2021*, Seattle, U.S.A., July 17–26 (2022) [arXiv:2204.00001] [inSPIRE].

[8] Y. Kuno and Y. Okada, *Muon decay and physics beyond the standard model*, *Rev. Mod. Phys.* **73** (2001) 151 [hep-ph/9909265] [inSPIRE].

[9] Y. Okada, K.-I. Okumura and Y. Shimizu, $\vec{\mu} \to e\gamma$ *and* $\vec{\mu} \to 3e$ *processes with polarized muons and supersymmetric grand unified theories*, *Phys. Rev. D* **61** (2000) 094001 [hep-ph/9906446] [inSPIRE].

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

[11] S. Davidson, *Completeness and complementarity for* $\mu \to e\gamma$, $\mu \to e\bar{e}e$ *and* $\mu A \to eA$, *JHEP* **02** (2021) 172 [arXiv:2010.00317] [inSPIRE].

[12] P.D. Bolton and S.T. Petcov, *Measurements of* $\mu \to 3e$ *Decay with Polarised Muons as a Probe of New Physics*, *Phys. Lett. B* **833** (2022) 137296 [arXiv:2204.03468] [inSPIRE].

[13] B. Echenard, R. Essig and Y.-M. Zhong, *Projections for Dark Photon Searches at Mu3e*, *JHEP* **01** (2015) 113 [arXiv:1411.1770] [inSPIRE].

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

[15] I. Galon, D. Shih and I.R. Wang, *Dark photons and displaced vertices at the MUonE experiment*, *Phys. Rev. D* **107** (2023) 095003 [arXiv:2202.08843] [inSPIRE].

[16] Y. Jho, S. Knapen and D. Redigolo, *Lepton-flavor violating axions at MEG II*, *JHEP* **10** (2022) 029 [arXiv:2203.11222] [inSPIRE].

[17] M. Hostert, T. Menzo, M. Pospelov and J. Zupan, *New physics in multi-electron muon decays*, *JHEP* **10** (2023) 006 [arXiv:2306.15631] [inSPIRE].

[18] R.J. Hill, R. Plestid and J. Zupan, *Searching for new physics at* $\mu \to e$ *facilities with* μ^+ *and* π^+ *decays at rest*, *Phys. Rev. D* **109** (2024) 035025 [arXiv:2310.00043] [inSPIRE].

[19] Mu3e collaboration, *The Rare and Forbidden: Testing Physics Beyond the Standard Model with Mu3e*, *SciPost Phys. Proc.* **1** (2019) 052 [arXiv:1812.00741] [inSPIRE].

[20] D.Y. Bardin, T.G. Istatkov and G. Mitselmakher, *On* $\mu^+ \to e^+ \nu_e \bar{\nu}_\mu e^+ e^-$ *decay*, *Yad. Fiz.* **15** (1972) 284 [inSPIRE].

[21] A. Flores-Tlalpa, G. López Castro and P. Roig, *Five-body leptonic decays of muon and tau leptons*, *JHEP* **04** (2016) 185 [arXiv:1508.01822] [inSPIRE].

[22] M. Fael and C. Greub, *Next-to-leading order prediction for the decay* $\mu \to e\,(e^+e^-)\,\nu\,\bar{\nu}$, *JHEP* **01** (2017) 084 [arXiv:1611.03726] [inSPIRE].

[23] G.M. Pruna, A. Signer and Y. Ulrich, *Fully differential NLO predictions for the rare muon decay*, *Phys. Lett. B* **765** (2017) 280 [arXiv:1611.03617] [inSPIRE].

[24] P. Banerjee, T. Engel, A. Signer and Y. Ulrich, *QED at NNLO with McMule*, *SciPost Phys.* **9** (2020) 027 [arXiv:2007.01654] [inSPIRE].

[25] P. Fayet, *Effects of the Spin 1 Partner of the Goldstino (Gravitino) on Neutral Current Phenomenology*, *Phys. Lett. B* **95** (1980) 285 [inSPIRE].

[26] L.B. Okun, *Limits of electrodynamics: paraphotons?*, *Sov. Phys. JETP* **56** (1982) 502 [inSPIRE].

[27] H. Georgi, P.H. Ginsparg and S.L. Glashow, *Photon Oscillations and the Cosmic Background Radiation*, *Nature* **306** (1983) 765 [inSPIRE].

[28] B. Holdom, *Two U(1)'s and Epsilon Charge Shifts*, *Phys. Lett. B* **166** (1986) 196 [INSPIRE].

[29] R. Essig et al., *Working Group Report: New Light Weakly Coupled Particles*, in the proceedings of the *Snowmass 2013: Snowmass on the Mississippi*, Minneapolis, U.S.A., July 29 – August 06 (2013) [arXiv:1311.0029] [INSPIRE].

[30] J.L. Hewett et al., *Fundamental Physics at the Intensity Frontier*, arXiv:1205.2671 [DOI:10.2172/1042577] [INSPIRE].

[31] J. Jaeckel and A. Ringwald, *The Low-Energy Frontier of Particle Physics*, *Ann. Rev. Nucl. Part. Sci.* **60** (2010) 405 [arXiv:1002.0329] [INSPIRE].

[32] M. Pospelov, *Secluded U(1) below the weak scale*, *Phys. Rev. D* **80** (2009) 095002 [arXiv:0811.1030] [INSPIRE].

[33] P. Fayet, *U-boson production in e^+e^- annihilations, psi and Upsilon decays, and Light Dark Matter*, *Phys. Rev. D* **75** (2007) 115017 [hep-ph/0702176] [INSPIRE].

[34] 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].

[35] A.-K. Perrevoort, *Sensitivity Studies on New Physics in the Mu3e Experiment and Development of Firmware for the Front-End of the Mu3e Pixel Detector*, Ph.D. thesis, Universität Heidelberg, 69117 Heidelberg, Germany (2018) [INSPIRE].

[36] FASER collaboration, *First Physics Results from the FASER Experiment*, in the proceedings of the *57th Rencontres de Moriond on Electroweak Interactions and Unified Theories*, La Thuile, Italy, March 18–25 (2023) [arXiv:2305.08665] [INSPIRE].

[37] A. Bodas, R. Coy and S.J.D. King, *Solving the electron and muon $g - 2$ anomalies in Z' models*, *Eur. Phys. J. C* **81** (2021) 1065 [arXiv:2102.07781] [INSPIRE].

[38] NA48/2 collaboration, *Search for the dark photon in π^0 decays*, *Phys. Lett. B* **746** (2015) 178 [arXiv:1504.00607] [INSPIRE].

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

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

[41] L. Doria et al., *Search for light dark matter with the MESA accelerator*, in the proceedings of the *13th Conference on the Intersections of Particle and Nuclear Physics*, Palm Springs, U.S.A., May 29 – June 03 (2018) [arXiv:1809.07168] [INSPIRE].

[42] A. Berlin, S. Gori, P. Schuster and N. Toro, *Dark Sectors at the Fermilab SeaQuest Experiment*, *Phys. Rev. D* **98** (2018) 035011 [arXiv:1804.00661] [INSPIRE].

[43] P. Ilten, Y. Soreq, M. Williams and W. Xue, *Serendipity in dark photon searches*, *JHEP* **06** (2018) 004 [arXiv:1801.04847] [INSPIRE].

[44] E. Nardi et al., *Resonant production of dark photons in positron beam dump experiments*, *Phys. Rev. D* **97** (2018) 095004 [arXiv:1802.04756] [INSPIRE].

[45] H. Benmansour, *Search for the X17 particle with the MEG-II apparatus*, (2022) [https://agenda.infn.it/event/28874/contributions/169428/attachments/93998/128533/ICHEP2022_Benmansour_0707.pdf].

[46] P. Ilten, J. Thaler, M. Williams and W. Xue, *Dark photons from charm mesons at LHCb*, *Phys. Rev. D* **92** (2015) 115017 [arXiv:1509.06765] [INSPIRE].

[47] B. Batell, A. Freitas, A. Ismail and D. Mckeen, *Flavor-specific scalar mediators*, *Phys. Rev. D* **98** (2018) 055026 [arXiv:1712.10022] [INSPIRE].

[48] P. Ballett, S. Pascoli and M. Ross-Lonergan, *U(1)' mediated decays of heavy sterile neutrinos in MiniBooNE*, *Phys. Rev. D* **99** (2019) 071701 [arXiv:1808.02915] [INSPIRE].

[49] S. Dimopoulos, *A Solution of the Strong CP Problem in Models With Scalars*, *Phys. Lett. B* **84** (1979) 435 [INSPIRE].

[50] B. Holdom and M.E. Peskin, *Raising the Axion Mass*, *Nucl. Phys. B* **208** (1982) 397 [INSPIRE].

[51] J.M. Flynn and L. Randall, *A Computation of the Small Instanton Contribution to the Axion Potential*, *Nucl. Phys. B* **293** (1987) 731 [INSPIRE].

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

[53] K. Choi and H.D. Kim, *Small instanton contribution to the axion potential in supersymmetric models*, *Phys. Rev. D* **59** (1999) 072001 [hep-ph/9809286] [INSPIRE].

[54] C. Kilic, T. Okui and R. Sundrum, *Vectorlike Confinement at the LHC*, *JHEP* **02** (2010) 018 [arXiv:0906.0577] [INSPIRE].

[55] G. Ferretti and D. Karateev, *Fermionic UV completions of Composite Higgs models*, *JHEP* **03** (2014) 077 [arXiv:1312.5330] [INSPIRE].

[56] A. Belyaev et al., *Singlets in composite Higgs models in light of the LHC 750 GeV diphoton excess*, *Phys. Rev. D* **94** (2016) 015004 [arXiv:1512.07242] [INSPIRE].

[57] B. Bellazzini et al., *R-axion at colliders*, *Phys. Rev. Lett.* **119** (2017) 141804 [arXiv:1702.02152] [INSPIRE].

[58] G. Ferretti, *Gauge theories of Partial Compositeness: Scenarios for Run-II of the LHC*, *JHEP* **06** (2016) 107 [arXiv:1604.06467] [INSPIRE].

[59] K.S. Jeong, T.H. Jung and C.S. Shin, *Axionic Electroweak Baryogenesis*, *Phys. Lett. B* **790** (2019) 326 [arXiv:1806.02591] [INSPIRE].

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

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

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

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

[64] B. Batell et al., *Muon anomalous magnetic moment through the leptonic Higgs portal*, *Phys. Rev. D* **95** (2017) 075003 [arXiv:1606.04943] [INSPIRE].

[65] A. Freitas, J. Lykken, S. Kell and S. Westhoff, *Testing the Muon g − 2 Anomaly at the LHC*, *JHEP* **09** (2014) 155 [*Erratum ibid.* **09** (2014) 155] [arXiv:1402.7065] [INSPIRE].

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

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

[68] S. Gninenko, *Addendum to the NA64 Proposal: Search for the $A' \to invisible$ and $X \to e^+e^-$ decays in 2021*, CERN-SPSC-2018-004, CERN, Geneva (2018).

[69] N. Baltzell et al., *The Heavy Photon Search Experiment*, arXiv:2203.08324 [INSPIRE].

[70] 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].

[71] E. Goudzovski et al., *New physics searches at kaon and hyperon factories*, *Rept. Prog. Phys.* **86** (2023) 016201 [arXiv:2201.07805] [INSPIRE].

[72] J.H. Collins, K. Howe and B. Nachman, *Extending the search for new resonances with machine learning*, *Phys. Rev. D* **99** (2019) 014038 [arXiv:1902.02634] [INSPIRE].

[73] SINDRUM collaboration, *Limits for short-lived neutral particles emitted in μ^+ or π^+ decay*, *Phys. Lett. B* **175** (1986) 101 [INSPIRE].

[74] S. Bilmis et al., *Constraints on Dark Photon from Neutrino-Electron Scattering Experiments*, *Phys. Rev. D* **92** (2015) 033009 [arXiv:1502.07763] [INSPIRE].

[75] M. Lindner, F.S. Queiroz, W. Rodejohann and X.-J. Xu, *Neutrino-electron scattering: general constraints on Z' and dark photon models*, *JHEP* **05** (2018) 098 [arXiv:1803.00060] [INSPIRE].

[76] P. Coloma, M.C. Gonzalez-Garcia and M. Maltoni, *Neutrino oscillation constraints on $U(1)'$ models: from non-standard interactions to long-range forces*, *JHEP* **01** (2021) 114 [*Erratum ibid.* **11** (2022) 115] [arXiv:2009.14220] [INSPIRE].

[77] NA64 collaboration, *Probing light dark matter with positron beams at NA64*, *Phys. Rev. D* **109** (2024) L031103 [arXiv:2308.15612] [INSPIRE].

[78] P. Coloma et al., *Constraining new physics with Borexino Phase-II spectral data*, *JHEP* **07** (2022) 138 [*Erratum ibid.* **11** (2022) 138] [arXiv:2204.03011] [INSPIRE].

[79] TEXONO collaboration, *A Search of Neutrino Magnetic Moments with a High-Purity Germanium Detector at the Kuo-Sheng Nuclear Power Station*, *Phys. Rev. D* **75** (2007) 012001 [hep-ex/0605006] [INSPIRE].

[80] BOREXINO collaboration, *Limiting neutrino magnetic moments with Borexino Phase-II solar neutrino data*, *Phys. Rev. D* **96** (2017) 091103 [arXiv:1707.09355] [INSPIRE].

[81] MEG collaboration, *Muon polarization in the MEG experiment: predictions and measurements*, *Eur. Phys. J. C* **76** (2016) 223 [arXiv:1510.04743] [INSPIRE].

[82] A. Alloul et al., *FeynRules 2.0 — A complete toolbox for tree-level phenomenology*, *Comput. Phys. Commun.* **185** (2014) 2250 [arXiv:1310.1921] [INSPIRE].

[83] C. Degrande et al., *UFO — The Universal FeynRules Output*, *Comput. Phys. Commun.* **183** (2012) 1201 [arXiv:1108.2040] [INSPIRE].

[84] 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].