The Leptonic Higgs Portal

Brian Batell,1 Nicholas Lange,2 David McKeen,3 Maxim Pospelov,2,4 and Adam Ritz2

1Pittsburgh Particle Physics, Astrophysics, and Cosmology Center, Department of Physics and Astronomy, University of Pittsburgh, PA 15260, USA
2Department of Physics and Astronomy, University of Victoria, Victoria, BC V8P 5C2, Canada
3Department of Physics, University of Washington, Seattle, WA 98195, USA
4Perimeter Institute for Theoretical Physics, Waterloo, ON N2J 2W9, Canada

An extended Higgs sector may allow for new scalar particles well below the weak scale. In this work, we present a detailed study of a light scalar S with enhanced coupling to leptons, which could be responsible for the existing discrepancy between experimental and theoretical determinations of the muon anomalous magnetic moment. We present an ultraviolet completion of this model in terms of the lepton-specific two-Higgs doublet model and an additional scalar singlet. We then analyze a plethora of experimental constraints on the universal low energy model, and this UV completion, along with the sensitivity reach at future experiments. The most relevant constraints originate from muon and kaon decays, electron beam dump experiments, electroweak precision observables, rare \(B_d\) and \(B_s\) decays and Higgs branching fractions. The properties of the leptonic Higgs portal imply an enhanced couplings to heavy leptons, and we identify the most promising search mode for the high-luminosity electron-positron colliders as \(e^+ e^- \rightarrow \tau^+ \tau^- + S \rightarrow \tau^+ \tau^- + \ell^+ \ell^-\), where \(\ell = e, \mu\). Future analyses of existing data from BaBar and Belle, and from the upcoming Belle II experiment, will enable tests of this model as a putative solution to the muon \(g-2\) problem for \(m_S < 3.5\) GeV.

1. INTRODUCTION

The LHC discovery of a new particle of mass \(\sim 125\) GeV, with properties consistent with those of the Standard Model Higgs boson [1], provides compelling evidence for the picture of the electroweak symmetry, and its spontaneous breakdown, encapsulated in the Standard Model (SM). It remains an important question to understand whether the entire Higgs sector is minimal, as in the SM, or contains additional states as would be required by supersymmetry, or may be motivated by other scenarios including, for example, models of dark matter.

While the existence of new physics at the TeV scale is still a distinct possibility (see e.g. [2]), in recent years, independent empirical motivations related to dark matter and neutrino masses have pointed to the possibility of a hidden sector, weakly coupled to the SM [3]. The mass scales in the hidden sector can be considered free parameters, and therefore particles much lighter than the electroweak or TeV scales are plausible. On general effective field theory grounds, the leading interactions with a neutral light hidden sector would be through the relevant and marginal interactions involving SM gauge singlets, which have been dubbed “portals” [4] and are the subject of considerable theoretical and experimental study.

In several cases, hypothetical light particles may help to explain certain experimental anomalies and deviations from the SM. It has been appreciated that a rather minimal extension of the SM via an additional vector particle \(V\) (often termed the “dark photon”) that kinetically mixes with the photon through the interaction \((\epsilon/2)V^\mu\nu F_{\mu\nu}\), where \(V^\mu\nu\) and \(F^\mu\nu\) are the V and photon field strengths respectively, can generate an appreciable shift of the muon anomalous magnetic moment [5],

\[
\Delta a_\mu \simeq \frac{\alpha \epsilon^2}{2 \pi} \quad \text{when } m_V \ll m_\mu. \tag{1}
\]

For \(\epsilon \sim 10^{-3}\), such a model offers a correction on the order of the existing discrepancy in \(a_\mu\), with the right sign to alleviate the tension between theory and experimental value of the Higgs doublet, \(\langle H \rangle = 246\) GeV, with properties consistent with those of the SM. The lesson of this observation is that if a new light scalar particle \(S\) couples to leptons with a coupling strength on the order of the SM lepton Yukawa couplings, which in the case of the muon is \(m_\mu/v \simeq 4 \times 10^{-4}\), the muon \(g-2\) problem can be solved. Thus we are motivated to study...
the effective Lagrangian of an elementary scalar $S$,

$$\mathcal{L}_{\text{eff}} = \frac{1}{2} (\partial _\mu S)^2 - \frac{1}{2} m_S^2 S^2 + \sum_{\ell=e,\mu,\tau} g_\ell S \ell \bar{\ell},$$

(3)

with $g_\ell \sim m_\ell / v$ as a promising phenomenological model. Given that $S$ is not the SM Higgs boson, the interaction terms in (3) may appear to contradict SM gauge invariance. Thus, at minimum, Eq. (3) requires an appropriate UV completion, generically in the form of new particles at the electroweak (EW) scale charged under the SM gauge group. On the other hand, if a UV-complete model is found that represents a consistent generalization of (3), the light scalar solution to the muon $g-2$ problem deserves additional attention. Another impetus for studying very light beyond-the-SM (BSM) scalars comes from the existing discrepancy of the muon- and electron-extracted charge radius of the proton [13].

This paper presents a detailed study of light scalars with enhanced coupling to leptons, and provides a viable UV-completion of Eq. (3) through what we dub the ‘leptonic Higgs portal’. We also analyze a variety of phenomenological consequences of the model. The phenomenology of a light scalar coupled to leptons resembles in many ways the phenomenology of the dark photon, but with the distinct feature that the couplings to individual flavors are non-universal and proportional to the mass. As a result, at any given energy the production of such a scalar is most efficient using the heaviest kinematically accessible lepton. We identify the most important search modes for the scalar that could decisively explore its low mass regime. Our main conclusion is that an elementary scalar with coupling to leptons $\ell$ scaling as $m_\ell$ can be very efficiently probed, and in particular the whole mass range consistent with a solution of the muon $g-2$ discrepancy can be accessed through an analysis of existing data and in upcoming experiments.

Our full UV-complete model is based on the lepton-specific two Higgs doublet model with an additional light scalar singlet. The mixing of the singlet with components of the electroweak doublets results in the effective Lagrangian of Eq. (3). The model also induces additional observables, and thus constraints, due to the fact that $S$ receives small but nonvanishing couplings to the SM quarks and gauge bosons. We note that the UV completion presented in this work is not unique. For an alternative UV completion of the same model utilizing vector-like fermions at the weak scale, see Ref. [14]. While many aspects of the low-energy phenomenology based on the effective Lagrangian (3) are similar in both approaches, the UV-dependent effects are markedly different (especially for flavor-changing observables).

This paper is organized as follows. In the next section we discuss light scalars coupled to leptons and a possible UV completion of such models via the leptonic Higgs portal. In Sec. 3 we analyze the constraints and sensitivity levels to light scalars coupled to leptons that are universal, and independent of the UV completion (resulting from muon decays, leptonic kaon decays, electron

![FIG. 1. Left panel: Constraints on the coupling to leptons (in terms of both $\xi_\ell^S = g_\ell (v/m_\ell)$ and $\epsilon_\ell = g_\ell / e$) as a function of the scalar mass, based purely on the effective theory in Eq. (3). The region where $(g - 2)_\mu$ is discrepant at 5$\sigma$ is shaded in red, while the green shaded band shows where the current discrepancy is brought below 2$\sigma$. We show constraints from the beam dumps E137, Orsay, and E141. The projected sensitivities from $\mu \to 3 \ell$, NA48/2, NA62, HPS, analyses of existing data from COMPASS and $B$-factories, as well as a projected sensitivity at BELLE II are also shown. (See Section 3 for details.) Right panel: Constraints on the L2HDM+ϕ UV completion of the effective theory in Eq. (3), as described in Sec. 2. Model independent results are as in the left panel. In addition, for this particular UV completion, there are constraints on the model from searches for $h \to SS \to 2\mu 2\tau / 4\mu$, $B \to K^{\pm} \ell^\pm \ell^\mp$, and $B_s \to \mu^+ \mu^-$. We have set $\tan \beta = 200$, $m_H = m_{H^\pm} = 500$ GeV, and $m_{12} = 100$ GeV. (See Section 4 for details.)]
beam dumps and high-intensity $e^+e^-$ colliders); the results are shown in the left panel of Fig. 1. In Sec. 4 we analyze the constraints and sensitivities that are tied to the specific UV-completion involving the leptonic Higgs portal. These include rare $B$ and Higgs decays; the results are shown in the right panel of Fig. 1. We present some additional discussion and reach our conclusions in Sec. 5.

2. LEPTONIC HIGGS PORTAL

In this section, we discuss a concrete UV-completion of the low-energy Lagrangian in Eq. (3). A simple starting point to couple a singlet field $\varphi$ to the SM is through the Higgs portal,

$$L_{\text{int}} = (A\varphi + \lambda\varphi^2) H^\dagger H,$$

where $H$ is the SM Higgs doublet and $A, \lambda$ are coupling constants. The trilinear term induces mixing between the singlet and the ordinary Higgs boson $h$ after electroweak symmetry breaking, where $H^0 = (v+h)/\sqrt{2}$. The mixing angle is given by

$$\theta = \frac{Av}{m_h^2 - m_\varphi^2},$$

and after field diagonalization the coupling of the light scalar $S$ (mostly comprised of the singlet $\varphi$) to SM fermions is simply given by their SM Yukawa coupling times this mixing angle. Low mass singlets are constrained by $B$ and $K$ meson decays (see, e.g., a collection of theoretical and experimental studies in Refs. [15–22]), and for $m_S < 4$ GeV the mixing angle is limited to $|\theta| < 10^{-3}$. Significant further advances in sensitivity to $\theta$ are possible with the planned SHIP experiment [23]. Therefore, there is no room to accommodate $\theta \sim O(1)$, and consequently no large correction to the muon $g-2$ is allowed within this simple model.

To circumvent this obstacle, we modify the SM by not only adding a singlet but also by introducing a second Higgs doublet that mixes with the singlet. In particular, we are interested in the so-called ‘lepton-specific’ representation of a generic two Higgs doublet model (L2HDM) [24–30]. Calling the two doublets with SM Higgs charge assignments $\Phi_1$ and $\Phi_2$, we assume that $\Phi_1$ couples exclusively to leptons, while $\Phi_2$ couples to quarks. Moreover, we assume that all physical components of $\Phi_{1,2}$ are at the weak scale or above. Taking $(\Phi_2)/(\Phi_1) \equiv \tan \beta$ very large, as well as arranging for the physical bosons of $\Phi_1$ to be heavier than those of $\Phi_2$, we arrive at an “almost SM-like” limit, but with the set of heavier Higgses that couple to leptons possessing couplings enhanced by $\tan \beta$. The mixing term $A_{12} (\Phi_1^\dagger \Phi_2 + \Phi_2^\dagger \Phi_1) \varphi$ will then efficiently mix $\varphi$ with $\Phi_1$, resulting in the light scalar $S$ coupling to leptons with strength

$$g_\ell = \frac{m_\ell}{v} \times \tan \beta \times \theta_\ell,$$

where $\theta_\ell$ is the mixing between $S$ and $\Phi_1$. It is then clear that the desirable outcome of $g_\ell \sim m_\ell/v$ can be achieved in the regime $\tan \beta \gg 1$, $\theta_\ell \ll 1$, and $\tan \beta \times \theta_\ell \sim O(1)$.

We now elaborate on this simple idea and present details of the model. The scalar potential we consider is given by

$$V(\Phi_1, \Phi_2, \varphi) = V_{2\text{HDM}} + V_\varphi + V_{\text{portal}}.$$  (7)

$V_{2\text{HDM}}$ is the main part of the potential that determines the pattern of electroweak symmetry breaking. Its CP-conserving version is given by the familiar expression,

$$V_{2\text{HDM}} = m_1^2 \Phi_1^\dagger \Phi_1 + m_2^2 \Phi_2^\dagger \Phi_2 - m_{12}^2 \left( \Phi_1^\dagger \Phi_2 + \Phi_2^\dagger \Phi_1 \right) + \frac{\lambda_1}{2} \left( \Phi_1^\dagger \Phi_1 \right)^2 + \frac{\lambda_2}{2} \left( \Phi_2^\dagger \Phi_2 \right)^2 + \lambda_3 \left( \Phi_1^\dagger \Phi_1 \right) \left( \Phi_2^\dagger \Phi_2 \right) + \frac{\lambda_4}{2} \left( \Phi_2^\dagger \Phi_2 \right)^2 + \left( \Phi_1^\dagger \Phi_1 \right)^2.$$  (8)

The singlet potential in (7) is a generic polynomial with positive $\varphi^4$ term,

$$V_\varphi = B\varphi + \frac{A_1}{2} \varphi^2 + \frac{A_2}{2} \varphi^3 + \frac{A_3}{4} \varphi^4.$$  (9)

In the portal part of the potential we are most interested in the trilinear terms,

$$V_{\text{portal}} = \left[ A_{11} \Phi_1^\dagger \Phi_1 + A_{22} \Phi_2^\dagger \Phi_2 \right. + \left. A_{12} \left( \Phi_1^\dagger \Phi_2 + \Phi_2^\dagger \Phi_1 \right) \right] \varphi.$$  (10)

Generically, the $A_{11}$ portal term leads to a $1/\tan \beta$ suppressed mixing between $\varphi$ and the electroweak scalars, while, for $\tan \beta \gg 1$, the $A_{22}$ portal coupling is strongly constrained by existing limits on the $A_2 H^\dagger H$ operator. On the other hand, the $A_{12}$ portal is less constrained and leads to efficient mixing with $\varphi$. In what follows we will ignore $A_{11,22}$. Including portal couplings of the form $\Phi_i^\dagger \Phi_j \varphi^2$ would not qualitatively change the phenomenology we are interested in so we ignore them.

The spectrum of the theory at the electroweak scale is dominated by $V_{2\text{HDM}}$, while $V_\varphi$ and $V_{\text{portal}}$ can be regarded as small perturbations. In determining the spectrum at the weak scale, we decompose the doublets assuming each obtains a vacuum expectation value, $(\Phi_a) \equiv v_a$,

$$\Phi_a \supset v_a + \rho_a$$  (11)

for $a = 1, 2$ with $\rho_a$ a real scalar field (we work in unitary gauge and ignore charged components of the doublets for now). The ratio of the VEVs is $v_2/v_1 = \tan \beta$ with $v_1^2 + v_2^2 = v^2 = (246 \text{ GeV})^2$. Furthermore, through a $\varphi$ field redefinition, the coefficient $B$ in Eq. (9) can be chosen so that $\varphi$ does not obtain a VEV.
The elements of the mass matrix of the neutral CP-even scalars in the basis \((\rho_1, \rho_2, \varphi)\) are
\[
M_{11}^2 = m_{12}^2 \tan \beta + \lambda_1 v^2 \cos^2 \beta, \\
M_{22}^2 = m_{12}^2 \cot \beta + \lambda_2 v^2 \sin^2 \beta, \\
M_{12}^2 = -m_{12}^2 + \lambda_3 v^2 \cos \beta \sin \beta, \\
M_{13}^2 = v A_1 \sin \beta, \\
M_{23}^2 = v A_2 \cos \beta, \\
M_{33}^2 = m_0^2.
\]
with \(\lambda_{345} \equiv \lambda_3 + \lambda_4 + \lambda_5\). In the limit that \(A_{12} \ll v, m_{12}\), we can rotate to the mass basis perturbatively,
\[
\begin{pmatrix}
\rho_1 \\
\rho_2 \\
\varphi
\end{pmatrix} \approx \begin{pmatrix}
-\sin \alpha & \cos \alpha & \delta_{11} \\
\cos \alpha & -\sin \alpha & \delta_{12} \\
\delta_{31} & \delta_{32} & 1
\end{pmatrix}
\begin{pmatrix}
h \\
H \\
S
\end{pmatrix},
\]
with small mixing angles \(\delta_{ij}\), and \(\alpha\) satisfying
\[
\tan 2\alpha = \frac{2M_{12}^2}{M_{11}^2 - M_{22}^2}.
\]
The masses of the physical states \(h\) and \(H\) are
\[
m_{h,H}^2 = \frac{1}{2} \left[ M_{11}^2 + M_{22}^2 \mp \sqrt{\left(M_{11}^2 - M_{22}^2\right)^2 + 4 \left(M_{12}^2\right)^2} \right],
\]
while the mass of \(S\) is
\[
m_S^2 \approx m_0^2 + \delta_{11} M_{13}^2 + \delta_{23} M_{23}^2.
\]
We will see that \(S\) can be rendered light while coupling dominantly to leptons (when \(\tan \beta\) is large) below, putting off questions of fine-tuning for the time being.

In the L2HDM, the Yukawa interactions of \(\Phi_1\) and \(\Phi_2\) with fermions are given by
\[
-\mathcal{L}_Y = \overline{\psi} \lambda_1 \Phi_1 \psi_R + \overline{\psi} \lambda_2 \Phi_2 \psi_R + \overline{\psi} \lambda_3 \Phi_3 \psi_R + \text{h.c.,}
\]
suppressing generation indices and using first generation notation. The Yukawa content of this model is exactly the same as in the SM, ensuring a pattern of minimal flavor violation (MFV). In particular, there are no flavor-changing neutral currents (FCNCs) mediated by either of the Higgs fields at tree level. The only difference with the SM is through the appearance of the vacuum angle \(\beta\) in the mass-Yukawa coupling relation,
\[
m_e = \cos \beta \times \frac{Y_e v}{\sqrt{2}}, \\
m_{u(d)} = \sin \beta \times \frac{Y_{u(d)} v}{\sqrt{2}}.
\]
In the large \(\tan \beta\) regime, the size of the Yukawa couplings in the quark sector is almost the same as in the SM, but in the lepton sector all Yukawa couplings are enhanced by \(\tan \beta\).

Upon diagonalization of the Higgs mass matrix, the Yukawa interactions of the physical states are
\[
-\mathcal{L}_Y = \sum_{\phi = S, h, H} \frac{\phi^* m_{\phi} \phi}{v} \psi^c \psi.
\]

### Table I

<table>
<thead>
<tr>
<th>(\phi)</th>
<th>(S)</th>
<th>(h)</th>
<th>(H)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\ell)</td>
<td>(\delta_{13} / \epsilon_3)</td>
<td>(-\delta_{12} / \epsilon_2)</td>
<td>(\epsilon_3 / \epsilon_2)</td>
</tr>
<tr>
<td>(q)</td>
<td>(\delta_{32} / s_3)</td>
<td>(c_2 / s_2)</td>
<td>(s_2 / c_2)</td>
</tr>
<tr>
<td>(W, Z)</td>
<td>(\delta_{13} / \epsilon_3)</td>
<td>(-\delta_{32} s_3 \sin (\beta - \alpha))</td>
<td>(-\delta_{32} c_2 \cos (\beta - \alpha))</td>
</tr>
</tbody>
</table>

### Table II

<table>
<thead>
<tr>
<th>(\phi)</th>
<th>(S)</th>
<th>(h)</th>
<th>(H)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\ell)</td>
<td>(- (v A_{12} / m_H^2) \tan \beta)</td>
<td>(\epsilon_3 \mp \tan \beta)</td>
<td>(\pm \tan \beta)</td>
</tr>
<tr>
<td>(q)</td>
<td>(- (v A_{12} / m_H^2) x \cot \beta)</td>
<td>(1)</td>
<td>(\mp \epsilon_3 \cot \beta)</td>
</tr>
<tr>
<td>(W, Z)</td>
<td>(- (v A_{12} / m_H^2) (r + x) \cot \beta)</td>
<td>(1)</td>
<td>(\pm (1 - \epsilon_3) \cot \beta)</td>
</tr>
</tbody>
</table>

Defined this way, \(\xi_{\psi, \lambda}^\phi = 1\) is a coupling of SM Higgs strength. In Table I, we show these couplings in terms of the angles \(\alpha\) and \(\beta\), and in Table II provide approximate values in the regime of interest.

We assume that \(h\) has SM-like couplings to the gauge bosons and quarks, which means that \(\cos (\beta - \alpha) \simeq 0\) and \(\cos \alpha \simeq \sin \beta\). Furthermore, if \(\tan \beta \gg 1\), then \(H\) and \(S\) will couple much more strongly to leptons than to quarks. This can be accomplished by choosing \(\alpha \simeq 0\) (and negative) and \(\beta \simeq \pi / 2\). In this case, we can make \(h\) arbitrarily SM-like, consistent with the observations of the ATLAS and CMS experiments, while allowing \(m_H\) and \(\tan \beta\) to vary (again ignoring questions of fine tuning for now).

Given this pattern of masses and couplings, we can find the singlet mixing angles,
\[
\delta_{13} \simeq -\frac{v A_{12}}{m_H^2}, \\
\delta_{23} \simeq -\frac{v A_{12}}{m_H^2} \left[ 1 + \frac{\epsilon_3}{1 - m_H^2 / m_H^2} \right] \cot \beta,
\]
or
\[
\xi_{\ell}^S \simeq -\frac{v A_{12}}{m_H^2} \tan \beta, \\
\xi_{q}^S \simeq -\frac{v A_{12}}{m_H^2} \left[ 1 + \frac{\epsilon_3}{1 - m_H^2 / m_H^2} \right] \cot \beta.
\]
Recall that the Yukawa couplings of $S$ are $g_{\ell,q} = \frac{\xi_{\ell,q} m_{\ell,q}}{v}$.

We can re-express the mass shift of the lightest scalar from Eq. (19) due to electroweak symmetry breaking in terms of more physical parameters,

$$m_S^2 \simeq m_0^2 - \left( \frac{m_H \xi_S}{\tan \beta} \right)^2.$$  \hspace{1cm} (27)

The cancellation between $\delta m_S^2$ and $m_0^2$ to obtain a GeV-scale value of $m_S$ represents a (mild) fine-tuning in this theory. We have checked that the hierarchy of the mass scales, $m_S \ll m_{h,H}$ is indeed possible without inducing an instability of the corresponding minimum in the scalar potential.

3. Universal constraints on the (leptonic) light scalar

We subdivide all the possible constraints on the light scalar $S$ into two groups. The first, model independent, group relies exclusively on the coupling to leptons in Eq. (3), and comes mostly from low and medium energy processes, and does not use any of the additional particles brought in by the UV completion. We present the second, model dependent, group of constraints in the next Section.

Although we introduced the notation $g_\ell = \frac{\xi_\ell}{m_\ell}$ in describing a particular UV completion in Sec. 2, we will make use of this parameterization and present results in this Section in terms of $\xi_\ell$, i.e. normalizing $g_\ell$ on the SM Higgs Yukawa coupling.

A. Lifetimes and decay modes of $S$

We will concentrate on the masses in the range 1 MeV to a few GeV for $m_S$. (A region from $\sim$ 200 keV to $2m_\ell \simeq$ 1 MeV may represent an interesting blind spot [31, 32], but is not treated in this paper.) In this mass range, the dominant decay modes of $S$ are to leptons, with partial width given by

$$\Gamma_{S \rightarrow \ell\bar{\ell}} = \frac{g_\ell^2}{8\pi} \frac{m_S}{2} \left( 1 - \frac{4m_\ell^2}{m_S^2} \right)^{3/2}.$$  \hspace{1cm} (28)

Depending on the coupling strength and the boost of the $S$ particle produced, the decay length of $S$ can be macroscopic, or rather prompt. For example, for $m_S = 1$ GeV, the proper decay length is

$$\tau(m_S = 1\text{ GeV}) \simeq 3 \times 10^{-6} \text{ cm} \times \left( \frac{1}{\xi_\ell} \right)^2,$$  \hspace{1cm} (29)

and the decay is prompt.

The $\gamma\gamma$ decay fraction may become noticeable (up to $\sim 20\%$ just below $m_S = 2m_\mu$) due to the loop-induced coupling to photons. In our model, the scaling $g_\ell \propto m_\ell$ allows for unambiguous determinations of the corresponding branching ratios. We plot the branching ratios of $S$ as a function of its mass in Fig. 2 noting that the decay is always dominated by the heaviest kinematically allowed lepton pair.

B. Muon anomalous magnetic moment

A loop of light scalars contributes to the anomalous magnetic moments of fermions. A straightforward calculation gives

$$a_\ell = \frac{g_\ell^2}{8\pi^2} \int_0^1 \frac{(1-z)^2(1+z)}{(1-z)^2 + z(m_S/m_\ell)^2},$$  \hspace{1cm} (30)

which, in the limits of a very light and a very heavy scalar, reduces to $3g_\ell^2/(16\pi^2)$ and $g_\ell^2/(4\pi^2) \times (m_\ell^2/m_S^2) \log(m_S/m_\ell)$ respectively. Equation (30) and the $g_\ell \propto m_\ell$ dependence lead to $a_\ell$ scaling as the second (fourth) power of lepton mass in the limit of a light (heavy) scalar. The tau lepton $g - 2$ receives the largest contribution from scalar exchange, but is not measured to the required precision (and, in fact, the $a_\tau$ sign is not experimentally determined). The strongest constraints come from $g - 2$ of the muon, and if the the current discrepancy, which we take to be $(26.1 \pm 8.0) \times 10^{-10}$ [33], is interpreted as new physics, it suggests a non-zero range for $\xi_\mu$ shown in Fig. 1. Notice that, in contrast to the dark photon case, the highly precise measurements of electron $g - 2$ do not provide competitive sensitivity. For the rest of the paper, we will treat the suggested muon $g - 2$ band as a target of opportunity, and investigate other observables that could provide complementary sensitivity to $g_\mu$ in this range.

To facilitate comparison with the dark photon case, we show results in Fig. 1 (left panel) in terms of both $\xi_\ell = g_\ell (v/m_\ell)$ and $\epsilon_{\ell,\mu}$, where $-e$ is the charge of the electron, which is the coupling strength to the electron of
a dark photon with kinetic mixing angle $\epsilon_{\text{eff}}$. Expressed in terms of $\epsilon_{\text{eff}}$, regions determined by the coupling to the electron are in roughly the same place as those in the dark photon case (modulo small differences due to scalar vs. vector properties), while those determined by couplings to $\mu$ and $\tau$ move to smaller values of $\epsilon_{\text{eff}}$ by factors of $\sim m_{\mu,\tau}/m_e$.

Note that in our UV-completion via the leptonic Higgs portal, there are additional contributions to $a_\mu$ from the heavy neutral and charged Higgs states. These contributions are subdominant to that of $S$, unless some of the neutral scalars are light, below the mass of the weak bosons. In this work, we will assume the heavy Higgs bosons are much heavier than this, so that the dominant contribution to $a_\mu$ comes from $S$, but see e.g. Refs. [29, 30, 34] for a recent study exploring this region of parameter space in the lepton-specific 2HDM.

C. Beam-dump and fixed target constraints

The coupling of the scalar $S$ to electrons is considerably smaller than to muons, $g_e/g_\mu = (m_e/m_\mu) \simeq 0.005$. Consequently, low mass scalars with $m_S < 2m_\mu$ can have displaced decays, or even travel a macroscopic distance before decaying. Fig. 1 shows constraints from older beam dump experiments, such as E137 and E141. In both cases, the scalars $S$ are produced in an underlying bremsstrahlung-like process, $e + \text{Nucleus} \rightarrow e + S + \text{Nucleus}$. Notice that these experiments firmly rule out scalars with masses below 30 MeV as candidates for the solution of muon $g-2$ discrepancy. Consequently, for the rest of the constraints, we will concentrate on $m_S > 10$ MeV. It is also important to note the modification of the shape of the excluded region compared to the case of dark photons, universally coupled to all leptons. In the scalar model above $m_S = 210$ MeV, there is no sensitivity in the beam dump experiments due to abrupt shortening of the lifetime of $S$ by the muon pair decay channel.

The JLab experiment HPS [35] utilizes a fixed target, scattering electrons on tungsten, producing scalars through their couplings to electrons. It has the capability to detect displaced decays within a few cm from the target, and will be sensitive to the scalar $S$ in the relevant mass range. Translating the projected sensitivity to the dark photon parameter space to the case of the leptonic scalar, we arrive at the sensitivity reach of HPS shown in Fig. 1. Above the muon threshold, the scalar decays are too prompt to be detected in this fashion. At the same time, muon fixed target experiments have a chance of probing this parameter space for the model. This possibility was discussed in Ref. [36], in connection with a possible search for an axion-like particle in $\mu + \text{Nucleus} \rightarrow \mu + \text{Nucleus} + a(\rightarrow \mu^+\mu^-)$ at the COMPASS facility at CERN [37]. Recasting the projected sensitivity in the case of the scalar particles, we obtain an $O(1)$ sensitivity to $\xi_S^\mu$, shown in Fig. 1.

It is also possible that proton beam dump and fixed target experiments could be sensitive to $S$. Indeed, primary mesons produced subsequently lead to muons, which in turn can radiate the scalar using a larger coupling, $g_\mu$. The challenge in such a set-up would be to identify a clean way of detecting electron-positron pairs (or for the case of the fixed target experiments, possibly muon pairs) that result from scalar decays. A planned high-energy proton beam dump experiment, SHIP [23], as well as the existing Fermilab experiment SeaQuest [38], may present advantageous venues, as the high-energy and relatively short distance to the detector will increase chances for detecting displaced decays.

As a separate note, it is worth mentioning that recent studies of the LHCb sensitivity to dark photons [39] may open a new pathway to probe dark scalars as well. The search suggested in [39] will not directly apply to a leptophilic scalar $S$. Nonetheless, LHCb provides an attractive opportunity to search for $S$ via its production in association with muons. The large boosts available at LHCb may facilitate such searches via displaced decays of $S$.

D. Future sensitivity from muon decay

Flavor-violating muon decays will be scrutinized in a series of upcoming experiments. Of particular interest for the model discussed in this paper is the $\mu^+ \rightarrow e^+e^+\nu\bar{\nu}$ search, planned at the Paul Scherrer Institute [40], which will have exquisite energy resolution for the final state leptons.

In the present model, the flavor-violating decays of muons are absent, but the exotic scalars $S$ can be radiated on-shell in the process $\mu^+ \rightarrow \nu_S e^+ S \rightarrow \nu_S e^+ e^+ e^-$. The momenta for the electron and one of the two positrons in the final state must reconstruct the mass of the scalar, $(p_{e+} + p_{e^-})^2 = m_S^2$. Therefore, a scalar signal would be a bump in the invariant mass of the electron-positron pairs, superimposed on the SM background $\mu^+ \rightarrow \nu\nu e^+ e^+ e^-$. Making use of the recent study of a future dark photon search in this set-up [41], we recast the projected sensitivity for the case of the leptonic scalar $S$. The signals for $S$ and $V$ were simulated using MadGraph. For the scalar, emission from the initial muon line dominates, since $g_\mu \gg g_e$. The resulting sensitivity reach is shown in Fig. 1.

Note that the projections of Ref. [41] assume a prompt decay of the intermediate $e^+e^-$ resonance. However, for a small portion of the low mass, small $\xi_S^\mu$ parameter space where the experiment has sensitivity, the decay length of the $S$ particle can be longer than $O(\text{cm})$, which is approximately the radius of the innermost silicon detector. Thus, a more careful study must be carried out to assess the sensitivity in this region. The displaced decays may in fact help to reduce the level of background if, of course, the vertex can be cleanly reconstructed. See also Ref. [41] for further discussion of a potential search involving displaced decays.
E. Kaon decays

Another well-studied source of muons is via kaon decays. A new particle coupled to muons can be emitted in the decay $K^+ \rightarrow \mu^+ \nu S$. Note that charge conjugated processes are understood to be implicitly included throughout this section. (For recent discussions of scalar and vector emission in similar processes, see Refs. [42, 43].) For this study, we will concentrate on the past experiment NA48/2 [44] and the on-going experiment NA62 [45].

Depending on the mass of the scalar, it will decay to either $\mu^+\mu^-$ or $e^+e^-$. The first case is relatively straightforward. The SM rate for a similar process, $K^+ \rightarrow \mu^+\nu\mu^+\mu^-$, was beyond the reach of previous experiments, and only upper limits on the corresponding branching fraction exist. On the other hand, for the electron-positron decays of $S$ there are significant sources of known background. The first source is due to a rare SM decay $K^+ \rightarrow \mu^+\nu e^+e^-$. This process has been measured for the invariant mass of a pair in excess of 150 MeV [46] with a branching ratio of $7 \times 10^{-8}$. Below 150 MeV, there is a significant background due to the SM process $K^+ \rightarrow \mu^+\nu\pi^0$, with subsequent Dalitz decay of the neutral pion $\pi^0 \rightarrow e^+e^-\gamma$ that would mimic the signal if the photon is not detected. Finally, there is also some background from pion/muon mis-identification in the underlying $K^+ \rightarrow \pi^+\pi^0$ decay and the Dalitz decay of $\pi^0$.

Even though NA48/2 data has been collected, the corresponding analysis has not yet been done, and therefore both experiments need to be viewed in terms of potential future sensitivity levels. We derive them using the calculated signal rate in our model, and the published detector resolution for electron-positron pairs. To estimate the backgrounds, we use known kaon branching ratios and assume that the probability of missing a photon is $\sim 10^{-3}$. We also extend $K^+ \rightarrow \mu^+\nu e^+e^-$ to the entire range of $m_{ee}$ using simulations. Above muon threshold we set the rate of the signal to 5 events to derive the corresponding sensitivity limits. The projected sensitivity is shown in Fig. 1.

F. Associated production of scalars with $\tau\bar{\tau}$ at lepton colliders

High-luminosity $B$-factories, such as BaBar and Belle, have collected an integrated luminosity of $\sim 1\text{ ab}^{-1}$, and among other things have produced a significant sample of $\tau^+\tau^-$ pairs. The upcoming experiment Belle II is aiming to expand this dataset by a factor of $O(100)$. Given lepton couplings proportional to mass, the associated production of scalars $S$ from the taus,

$$e^+e^- \rightarrow \tau^+\tau^- + (S \rightarrow e^+e^- \text{ or } \mu^+\mu^-),$$

may represent the best chance for discovering or limiting the parameter space for such particles. The search for exotic particles in association with taus is a relatively unexplored subject, with only one specific case analyzed to date [47, 48].

The production cross section for (31) can be calculated analytically. We present the corresponding result as a function of the scalar mass in Fig. 3. To set the scale of the expected event rate for a 1 GeV mass scalar, we take parameters within the muon $g - 2$ band, and translate to the scale of the coupling to $\tau$-leptons, $g_\tau^2 \sim 1.3 \times 10^{-3}$. This leads to a very large number of produced scalars in the combined BaBar and Belle dataset, on the order of $5 \times 10^4$. Simulating the QED backgrounds using MadGraph, and requiring that at least one of the taus decay leptonically, we arrive at the sensitivity curves shown in Fig. 1. These sensitivity projections rely on a “bump hunt” in $\mu^+\mu^-$ (or $e^+e^-$) over the smoothly distributed QED background. Notice that for $m_S > 2m_\tau$, the dominant decay mode of the scalar is the tau pair, and the sensitivity is reduced due to the lack of stable leptons reconstructing to the invariant mass $m_S$. The decay to muons in this mass range is suppressed by $(m_\mu/m_\tau)^2$. Also, for scalar masses below $2m_\mu$ the decay length of scalars become comparable to the size of the detector, leading to reduced sensitivity. We account for this by introducing a requirement that the $S$ decays occur within 25 cm of the beam pipe.

It is worth emphasizing that an analysis of process (31) represents perhaps the most effective way of probing the parameter space of the leptonic scalar model in a wide mass range, from a few MeV to $\sim 3.5$ GeV.
4. CONSTRAINTS ON LIGHT SCALARS DUE TO THEIR ELECTROWEAK PROPERTIES

In this Section we analyze constraints that depend on the embedding of the simple framework of Eq. (3) into the SM. We focus on those that are a consequence of our choice of the L2HDM+φ scenario outlined in Sec. 2; in other models, constraints could differ.

A. Higgs decays

The SM-like Higgs $h$ can decay to pairs of light scalars through both $V_{2HDM}$ and $V_{portal}$ after electroweak symmetry breaking via the operator $C_{hSS}hSS$. In the SM-like limit,

$$C_{hSS} \simeq \left( \frac{m_h^2}{2 \tan \beta} + 2m_{t2}^2 \right) \frac{(\xi^S)^2}{\tan \beta}. \quad (32)$$

The decays $h \to SS \to 4\tau$ [49] and $h \to SS \to 2\mu2\tau$ [50] have been probed at the LHC, but not observed. These null results can be interpreted as an upper limit on $\xi^S$. As suggested in [51], the $2\mu2\tau$ final state offers better reach than the $4\tau$ search for $m_S > 2m_\tau$. There is also a very strong bound from a CMS search for Higgs decays into two highly collimated $\mu^+\mu^-$ pairs, $h \to SS \to (2\mu)(2\mu)$, which limits the branching for this process to less than $10^{-5}$ [52]. This bound is particularly relevant for $2m_\mu < m_S < 2m_\tau$. In Fig. 1 (right panel), we show the limit from this search for $\tan \beta = 200$, $m_{12} = 1$ TeV. The constraints become important for the muon $g-2$-motivated parameter space once $m_S$ is in the multi-GeV regime.

We note here that a light scalar could appear in $Z$ decays, through $Z \to \tau^+\tau^-S$. For $m_S$ between 100 MeV and 1 GeV, this branching for this mode is roughly $10^{-7}(\xi^S)^2$. Future facilities that produce $\mathcal{O}(10^9)$ $Z$ bosons could potentially observe this decay.

B. $B$-meson decays

Although its coupling to quarks and $W$ bosons is suppressed, the scalar mediates quark flavor-changing transitions at one-loop, leading to, for instance, rare $B$ decays like $B \to K\mu^+\mu^-$ (or more generically, $B \to X_s\mu^+\mu^-$) and $B_s \to \mu^+\mu^-$. At large $\tan \beta$ and $\xi^S = 1$, the leading term in the effective Lagrangian mediating $b \to s$ transitions relevant for these decays is

$$\mathcal{L}_{b\to s} \simeq -\frac{3V_{tb}^*V_{ts}m_m^2}{16\pi^2} \frac{m_{B_s}^2 \xi^S}{m_b^2 \tan^2 \beta} S_S b_L + \text{h.c.} \quad (33)$$

This operator can mediate the decay $B_s \to S^* \to \mu^+\mu^-$ through an off-shell $S$ and can lead to the decay $B \to KS^{(*)} \to K\mu^+\mu^-, Ke^+e^-$. If $2m_e < m_S < 2m_\tau$, the decays $B \to KS$ and $B \to K^+S$ can proceed with $S$ decaying to $\mu^+\mu^-$ subsequently; this is subject to strong constraints from the lack of a bump in the $\mu^+\mu^-$ invariant mass in $B \to K^+\mu^+\mu^-$ at LHCb [21]. We show limits on $\xi^S$ that result from these decay modes in Fig. 1, taking $\tan \beta = 200$, $m_H = m_{H\pm} = 500$ GeV. (The degeneracy of the heavy Higgs masses weakens electroweak precision constraints.) Notice that for the mass range $2m_\tau < m_S < m_B - m_{12}^S$, the sensitivity is degraded as $S$ would primarily decay to a tau pair.

We note in passing that the constraint on $\xi^S$ could be weakened by a factor $\sim m_H^2/m_b^2$ if $\xi^S \sim -1$ [cf. Eq. (24)] which is consistent with the data on Higgs properties.

For $m_S < 2m_\mu$, the important search channels are $B \to X_s e^+e^-$. These modes are better suited for searches at Belle II, and sensitivities below branchings of $10^{-8}$ will also cover the remaining ‘triangular’ parameter space in Fig. 1 (right panel).

C. Electroweak precision constraints

Enhanced couplings of the lepton-specific Higgses will also induce one-loop corrections to leptonic branching ratios of the $Z$-boson. Here we analyze $R_7$, defined as $R_7 \equiv (\Gamma(Z \to Z_h\phi)/\Gamma(Z \to \gamma\gamma))$, where $\Gamma(Z \to Z_h\phi) \propto \sum_{q=u,d,s,c,b}(g_{q\ell}^2 + g_{b\ell}^2)$ and $\Gamma(Z \to \gamma\gamma) \propto ((|g_{\ell\gamma}|^2 + |g_{b\gamma}|^2)$, with $g_{u\ell} = t_3 - Q_S^2$, $g_{b\ell} = -Q_S^2$, $SW$ stands for the sine of the weak mixing angle. Perturbations to $R_7$ can be expressed in terms of corrections to $S_W^0$ and modifications of the $Z\tau\tau$ vertices by the scalar loops,

$$\frac{\Delta R_7}{R_7} = 4.3\delta g_{\ell\gamma} - 3.7\delta g_{b\gamma} - 0.8\delta S_W^0 \quad \quad (34)$$

$$\simeq 46g_{W_3}^2 + 1.9 \times 10^{-3} T \quad \quad (35)$$

with $g_A = g_L - g_R$.

Interpreting the PDG fit, $R_7 = 20.764 \pm 0.045$ as the constraint, $-2 \times 10^{-3} \leq \Delta R_7/R_7 \leq 2 \times 10^{-3}$, we compare it to the result of the one-loop calculation in our model. The corrections to $\delta g_{\ell\gamma}$ and $\delta g_{b\gamma}$ can be obtained in the L2HDM model following [53, 54], and we present the ensuing constraint in the right panel of Fig. 1. The contributions due to loops of scalars that are (mostly) components of electroweak doublets are negligible for $m_{H,H\pm,A} \gtrsim 300$ GeV, even for $\tan \beta$ as large as 200, as taken in Fig. 1.

Additionally, we mention that as long as there is some degeneracy in the masses of at least two heavy scalars (at the order of ~ 50 GeV), corrections to the oblique electroweak parameters $S, T$, and $U$ are not constraining.

5. DISCUSSION AND CONCLUSIONS

We have analyzed a simplified model of a light ‘dark scalar’ that couples predominantly to leptons. This hid-
den sector model has a very distinct phenomenology, differing in several ways from the phenomenology of the canonical dark photon model. It is interesting that the coupling of a light scalar $S$ to leptons can still be of order $m_{\text{e}}/v$, and thus capable of inducing a large shift in the anomalous magnetic moment of the muon, without being excluded by direct searches. This is because the coupling to electrons relative to muons is suppressed by $m_{\text{e}}/m_{\text{\mu}}$, and many constraints that have ruled out the minimal version of the dark photon model as an explanation of the muon $g - 2$ discrepancy do not have any constraining power.

The simplified model (3) does not, however, respect the SU(2)$_L$ × U(1) gauge symmetry of the SM and needs a UV completion. This implies that either the field $S$ or the fermion fields in (3) cannot have well defined charge assignments. One possible UV completion, investigated in this paper, defines $S$ predominantly as a singlet scalar with a small admixture of an SU(2) doublet. On the other hand, one can consider the possibility of lepton fields in (3) arising from a mixing between the ‘normal’ SM fields and heavy vector-like leptons [14], so that mixing with a pure singlet $S$ becomes possible.

The UV completion of the model proposed here is based on the lepton-specific two Higgs doublet model, augmented by an additional light singlet. In the large $\tan \beta$ regime, the Yukawa couplings of the lepton-specific Higgs bosons $h_1 \supset (H, A, H^\pm)$ to leptons are enhanced relative to their SM values. If an additional singlet field $\phi$ mixes with $h_1$, the end result can be a new light boson $S$ with couplings to leptons that scale as $m_1$ and are of order the SM Yukawa couplings, proportional to the product of a small mixing angle $\theta$ and large $\tan \beta$. At the same time, the couplings of $S$ to quarks and weak gauge bosons are suppressed, which softens all constraints from the FCNC processes derived from $K, B$ physics. Moreover, there are no charged lepton flavor violating processes, since flavour conservation is built into the Yukawa structure of the model. (For the alternate UV completion with vector-like fermions [14], flavor symmetry in the charged lepton sector is likely to be broken. At the same time, the pure singlet nature of $S$ in this type of UV completion may allow flavor changing processes to be kept separate for the quark and lepton sectors, thus avoiding strong constraints from hadronic FCNC.)

We have analyzed a wide selection of constraints and sensitivity limits from the existing experiments, and from upcoming searches. The production of scalars is enhanced in processes that involve muons and tau leptons. We have studied muon and kaon decays, and shown that future experiments and analyses of the existing data (e.g. by NA48/2, BaBar and Belle experiments) are capable of reaching the levels of sensitivity to the parameter space suggested by the muon $g - 2$ discrepancy. The mass range $m_S < 2m_{\text{\mu}}$ naturally leads to longer lived bosons, and may be probed through experiments that have sensitivity to displaced decays, such as the HPS experiment at JLab.

Perhaps the most sensitive current search for a leptonic dark scalar can be performed by the BaBar and Belle collaborations, using existing data. The process of interest involves tau pair production with an associated emission of the scalar. The large datasets generated by the two experiments will allow a sensitive analysis of $\tau^+\tau^-\mu^+\mu^-$ and $\tau^+\tau^-e^+e^-$ production, looking for a peak in the invariant mass of electrons and muons. Even without extra data that should be collected at Belle II, the two $B$-factories should comprehensively test the dark scalar model in the wide mass range spanning almost three orders of magnitude.

The constraints and projected sensitivity reach for many experiments are summarized in the two panels of Fig. 1. The results in the left panel are based only on the simplified model (3) and use only the $g_1 \propto m_1$ scaling and absence of invisible decay channels for $S$. Much stronger constraints are derived for $m_S > 2m_{\text{\mu}}$ using quark flavor physics, within the lepton-specific 2HDM UV completion. One should still keep in mind that the strong constraints shown in the right panel of Fig. 1 are indeed very sensitive to the type of UV completion, and can in principle be avoided with a different microscopic model of (3).

Note added: Following the completion of this work, the BaBar Collaboration released a preprint [55] with an analysis that constrains any light vector particle ($V$) in the $e^+e^- \rightarrow \mu^+\mu^- V \rightarrow \mu^\pm\mu^-\mu^\pm\mu^-$ channel. This limit can be appropriately recast for the scalar model, and we show the resulting constraint as the solid black line in Fig. 1. This is now the strongest model-independent constraint over a large region of the $2m_{\text{\mu}} < m_S < 2m_\tau$ mass range. However, unlike the case of a vector coupled to $L_\mu - L_\tau$, the limit from [55] does not rule out the $g - 2$ band in that region. The reason is that the scalar contribution to $g - 2$ is somewhat larger than that of the vector and its production cross section is smaller at the same mass and coupling to muons. The constraint can be improved even further if BaBar performs the corresponding $e^+e^- \rightarrow \tau^+\tau^- S$ analysis.

ACKNOWLEDGMENTS

We would like to thank B. Echenard, E. Goudzovski, I. Nugent, M. Roney and B. Shuve for helpful discussions. The work of M.P. and A.R. is supported in part by NSERC, Canada, and research at the Perimeter Institute is supported in part by the Government of Canada through NSERC and by the Province of Ontario through MEDT. The work of B.B. is supported in part by the U.S. Department of Energy under grant No. DE-SC0015634. The work of D.M. is supported by the U.S. Department of Energy under grant No. DE-FG02-96ER40956.