PHYSICAL REVIEW D **92,** 055021 (2015)

Dark matter relic from muon anomalies

Geneviève Bélanger,[1] Cédric Delaunay,[1] and Susanne Westhoff[2]
[1]*LAPTh, Université Savoie Mont Blanc, CNRS B.P. 110, F-74941 Annecy-le-Vieux, France*
[2]*PITTsburgh Particle-physics Astrophysics and Cosmology Center (PITT-PACC), Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA*
(Received 31 July 2015; published 17 September 2015)

We show that the recently reported anomalies in $b \to s\mu^+\mu^-$ transitions, as well as the long-standing $g_\mu - 2$ discrepancy, can be addressed simultaneously by a new massive Abelian gauge boson with loop-induced coupling to muons. Such a scenario typically leads to a stable dark matter candidate with a thermal relic density close to the observed value. Dark matter in our model couples dominantly to leptons, hence signals in direct detection experiments lie well below the current sensitivity. The LHC, in combination with indirect detection searches, can test this scenario through distinctive signatures with muon pairs and missing energy.

DOI: 10.1103/PhysRevD.92.055021 PACS numbers: 12.60.Cn, 13.20.He, 14.60.Ef

I. INTRODUCTION

The lack of an acceptable dark matter (DM) candidate within the Standard Model (SM) is a pressing phenomenological motivation for the existence of new physics (NP) beyond the SM. Dark matter may very well be close to the weak scale, emerging from theories addressing the electroweak (EW) hierarchy problem of the SM. However, these theories generally predict a variety of new phenomena and particles around the TeV scale, whose existence still remains to be established experimentally. Despite the lack of a NP discovery, a few measurements are in mild tensions with SM predictions. In particular, there is a growing array of anomalies involving muons [1–5].[1] Moreover, DM searches at direct detection experiments strongly limit its interactions with quarks (see e.g. Ref. [9]), suggesting that DM around the EW scale might preferentially couple to leptons. In this paper, we thus take a data-driven approach, entertaining the possibility that the observed muon-related anomalies are the first signals of leptophilic DM at the EW scale.

The list of anomalies involving muons starts with the long-standing puzzle of the anomalous magnetic moment of the muon, $a_\mu \equiv (g_\mu - 2)/2$. The Brookhaven National Laboratory measurement [1] exceeds the SM prediction by about three standard deviations [10],

$$\Delta a_\mu \equiv a_\mu^{\rm exp} - a_\mu^{\rm SM} = (287 \pm 80) \times 10^{-11}. \qquad (1)$$

More recently, the LHCb experiment reported on a series of anomalies in (semi-) leptonic B meson decays, which together point to a possible new source of $b \to s\mu^+\mu^-$ transitions at short distances. The perhaps most tantalizing deviation occurs in the ratio $R_K = \mathcal{B}(B^+ \to K^+\mu^+\mu^-)/\mathcal{B}(B^+ \to K^+e^+e^-)$, which was observed about 2.6σ below the theoretically clean SM prediction, $R_K^{\rm SM} - 1 \sim 10^{-4}$ [4],

$$R_K^{\rm exp} = 0.745^{+0.090}_{-0.074}({\rm stat}) \pm 0.036({\rm syst}). \qquad (2)$$

Furthermore, the measured decay rate for $B^0 \to K^{0*}\mu^+\mu^-$ was found to exceed the SM prediction in a particular region of phase space [2], a result supported by the latest LHCb data [3]. Also the recent measurement of the differential distribution in $B_s^0 \to \phi\mu^+\mu^-$ deviates from the SM prediction by about 3.5σ [5]. The observed decay rate for $B_s^0 \to \mu^+\mu^-$ is slightly below, but compatible with the SM prediction [11]. While none of these LHCb anomalies by themselves are significant enough to claim a discovery, it is intriguing that they point to a common Lorentz structure when interpreted as a signal of NP [12–17]. A simultaneous explanation of Δa_μ, however, requires more sophisticated model building. In this paper, we propose a simple toy model, which addresses both the $g_\mu - 2$ and the various LHCb anomalies.

Most NP interpretations of the $b \to s\mu^+\mu^-$ anomalies postulate the existence of a new state in the range of $\Lambda \sim 1$–10 TeV with *tree-level* couplings to muons and quarks, for instance a Z' gauge boson [12,18–20] or a scalar leptoquark [21–24]. The same interactions typically contribute to $g_\mu - 2$ at the one-loop level, but yield too small a contribution to explain the discrepancy in Eq. (1), due to the suppression by the high scale, $a_\mu \propto m_\mu^2/\Lambda^2$ [13]. A way out is to generate the coupling of Z' to muon pairs only *radiatively*, so that contributions to $b \to s\mu^+\mu^-$ transitions and $g_\mu - 2$ are both induced at the one-loop level by NP around the EW scale. As we argue in this work, this requires a richer NP sector with an electrically neutral state,

[1]We do not consider the proton size anomaly observed in muonic hydrogen atoms [6]. NP interpretations of this measurement are very challenging [7,8].

which is stable if the tree-level $Z'\mu^+\mu^-$ coupling is forbidden by a (spontaneously broken) symmetry. Hence, addressing the aforementioned muon-related anomalies generally *yields* a DM candidate, which, by construction, is mostly leptophilic. The same NP interactions dominate DM annihilation in the early Universe. We will demonstrate that a minimal model with the above properties typically leads to a stable DM candidate with a thermal relic density of the order of the observed value.

The remainder of the paper is organized as follows. We introduce our phenomenological model in Sec. II and identify the parameter space to simultaneously accommodate the $g_\mu - 2$ and $b \to s\mu^+\mu^-$ anomalies in Sec. III. We then discuss the implications of collider constraints in Sec. IV. The resulting DM phenomenology is analyzed in Sec. V and Sec. VI. We conclude and give an outlook on future experimental tests of our model in Sec. VII.

II. A MINIMAL MODEL

We consider the extension of the SM by a new (dark) sector, consisting of heavy leptons L, L^c and quarks Q, Q^c with vectorlike gauge couplings, as well as two complex scalars, ϕ and χ. Interactions with the SM are mediated by a new gauge boson Z' associated with an Abelian $\mathrm{U}(1)_X$ symmetry, under which only the particles of the dark sector are charged. The quantum numbers of all new particles are listed in Table I.[2] Since no chiral fermion carries $\mathrm{U}(1)_X$ charge, the model is anomaly free [25]. Notice that our choice of $\mathrm{U}(1)_X$ charges forbids tree-level Z' couplings with SM leptons. This is crucial in order to simultaneously address the $g_\mu - 2$ and $b \to s\mu^+\mu^-$ anomalies with NP around the weak scale.

Besides canonical kinetic terms, the relevant new interactions in the Lagrangian are

$$\mathcal{L}_{\mathrm{NP}} \supset \epsilon B_{\mu\nu}X^{\mu\nu} - \lambda_{\chi H}|\chi|^2|H|^2 - \lambda_{\phi H}|\phi|^2|H|^2 - V(\phi,\chi) - [y(\bar{l}L)\chi + w(\bar{q}Q)\phi + \mathrm{H.c.}], \quad (3)$$

where

$$V(\phi,\chi) \equiv (r\phi\chi^2 + \mathrm{H.c.}) + \lambda_\phi|\phi|^4 + \lambda_\chi|\chi|^4, \quad (4)$$

and $l^T = (\nu_{\ell L}, \ell_L)^T$, $q^T = (u_L, d_L)^T$ and H are the SM lepton, quark and Higgs doublets, while $B_{\mu\nu}$ and $X_{\mu\nu}$ are the hypercharge and $\mathrm{U}(1)_X$ field strength tensors, respectively.

[2]A Z' model that explains the $b \to s\mu^+\mu^-$ anomalies with tree-level couplings to quarks and leptons and provides an inherent DM candidate has recently been proposed in Ref. [20]. Unlike in our model, the dark leptons mix with SM leptons, which results in a different phenomenology of dark matter and lepton observables. In particular, the $g_\mu - 2$ contribution in the model in Ref. [20] is typically much smaller than the needed shift in Eq. (1).

TABLE I. New fields and their quantum numbers. All SM fields are neutral under $\mathrm{U}(1)_X$.

	Spin	$\mathrm{SU}(3)_c$	$\mathrm{SU}(2)_L$	$\mathrm{U}(1)_Y$	$\mathrm{U}(1)_X$
L, L^c	1/2	**1**	**2**	−1/2	1
Q, Q^c	1/2	**3**	**2**	1/6	−2
ϕ	0	**1**	**1**	0	2
χ	0	**1**	**1**	0	−1

The scalar couplings y and w are *a priori* generic complex 3×3 matrices in flavor space. The $\mathrm{U}(1)_X$ symmetry is spontaneously broken by the vacuum expectation value (VEV) of ϕ, leading to a Z' mass of $m_{Z'} = 2g'\langle\phi\rangle$. It furthermore lifts the mass degeneracy between the components of $\chi = (\chi_0 + i\chi')/\sqrt{2}$ by

$$\delta \equiv \frac{m_{\chi'}^2}{m_{\chi_0}^2} - 1 = -2\frac{r\langle\phi\rangle}{m_{\chi_0}^2}. \quad (5)$$

We further assume that the singlet χ is *inert*, i.e., does not develop a VEV. There thus remains an exact $\mathbb{Z}_2$ symmetry, under which only χ and L are odd, which ensures that the lightest state of the spectrum is stable. For definiteness, we choose $r < 0$ and $m_L > m_{\chi_0}$, so that the scalar component of χ, χ_0 is a DM candidate.[3] As we will argue in the following sections, the above framework accommodates the $g_\mu - 2$ and $b \to s\mu^+\mu^-$ anomalies simultaneously, without conflicting with current collider constraints, only in particular regions of parameter space, where the DM candidate χ_0 is around the weak scale and is relatively lighter than the other dark-sector states,

$$m_{\chi_0} \lesssim m_{\chi'} \sim m_Q \sim m_{Z'}, \quad (6)$$

while the dark-lepton mass m_L may be either close to the DM state or well above it. We do not address the dynamical origin of such a spectrum, but simply achieve it by tuning the bare mass squared of χ and $r\langle\phi\rangle$. We use $1/\delta$ as a measure of this tuning, which remains mild in the parameter region of interest. Besides addressing the anomalies, this small hierarchy of the spectrum also ensures that the lightest dark-sector state is a spinless SM-singlet DM candidate.

We now turn to the new interactions in Eq. (3). The first term describes kinetic mixing between the hypercharge $\mathrm{U}(1)_Y$ and the dark $\mathrm{U}(1)_X$ field tensors, which is strongly constrained by electroweak precision measurements [26] and collider searches [27]. The second term, the scalar Higgs portal, is bounded by direct detection and collider

[3]Choosing the neutral dark lepton L^0 as the lightest state of the spectrum would *a priori* also yield a DM candidate. However, it would already have been observed in direct DM detection experiments through Z-boson mediated interactions with nuclei.

searches [28]. Since both interactions are not relevant to our discussion, we set $\epsilon = \lambda_{\chi H} = 0$.[4]

The last two terms mediate interactions between the dark sector and the SM. The Z' couples with strength g' to the current $j^\rho \equiv j_0^\rho + \delta j^\rho$, where j_0^ρ is the tree-level $\mathrm{U}(1)_X$ current and

$$\delta j^\rho \equiv \sum_{f,f'} \frac{\Gamma_{ff'}(q^2)}{1+\delta_{ff'}} (\bar{f}_L \gamma^\rho f'_L) + \mathrm{H.c.} \tag{7}$$

is the part of the current induced after $\mathrm{U}(1)_X$ breaking. Here q^2 is the squared Z' momentum, $f, f' = \ell, \nu_\ell, u, d$, and $\delta_{ff'} = 1$ for $f = f'$ and zero otherwise. SM quarks mix with dark quarks Q after $\mathrm{U}(1)_X$ breaking, while SM leptons and dark leptons L do not mix due to the different charge assignments under $\mathrm{U}(1)_X$ (see Table I). The Z' coupling to SM quarks thus arises at tree level through the mixing $w(\bar{q}Q)\langle\phi\rangle + \mathrm{H.c.}$. It is q^2 independent and isospin universal at leading order. In contrast, the Z' coupling to leptons is radiatively induced at the one-loop level through the exchange of dark-sector fields via the interaction $y(\bar{l}L)\chi + \mathrm{H.c.}$, see Fig. 1. In the limit $m_{\ell,\nu_\ell} \ll m_{\chi,L}$, which we envisage here, the lepton form factors satisfy $\Gamma_{\ell\ell}(q^2) \simeq \Gamma_{\nu_\ell\nu_\ell}(q^2)$, since the isospin components of $L = (L^0, L^-)$ are mass degenerate at leading order. Calculating the one-loop diagrams of Fig. 1 in the limit $q^2 = 0$ yields

$$\Gamma_{\ell\ell}(0) = \frac{|y|^2}{32\pi^2} F_{Z'}(\tau, \delta), \tag{8}$$

where $\tau \equiv m_L^2/m_{\chi_0}^2$ and the loop function $F_{Z'}$ is given in Eq. (A2).

Notice that $F_{Z'}(\tau, \delta)$ vanishes in the limit of unbroken $\mathrm{U}(1)_X$, $\delta \to 0$, where scalar and pseudoscalar components are degenerate. This stems from the fact that the form factor $\Gamma_{\ell\ell}(0)$ formally arises from the local operator $(\phi^* D_\rho \phi)(\bar{\ell}_L \gamma^\rho \ell_L) + \mathrm{H.c.}$ after $\mathrm{U}(1)_X$ breaking, induced by the one-loop diagram in Fig. 1 with two insertions of the VEV $\langle\phi\rangle$.

As of the flavor structure of dark-sector interactions, we follow a phenomenological approach and introduce only scalar couplings y and w between flavors that are required to explain the observed anomalies. We thus assume that the dark sector only couples to the left-handed fermions $\mu_L, \nu_{\mu L}$ and $\bar{b}_L s_L + \bar{s}_L b_L$. The magnitude of the $Z' \bar{b}_L s_L$ coupling is experimentally constrained from B_s meson mixing.

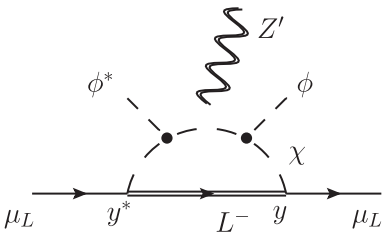

FIG. 1. Dominant one-loop contribution to the operator $(\phi^* D_\rho \phi)(\bar{\mu}_L \gamma^\rho \mu_L)$ yielding $\Gamma_{\mu\mu}(0) \neq 0$. The Z' line is understood to be attached wherever possible. A similar diagram with $\mu \to \nu_\mu$ and $L^- \to L^0$ induces the Z' coupling to neutrinos.

Allowing for a NP contribution of $\mathcal{O}(10\%)$ to the mass difference ΔM_{B_s} leads to the bound [13,30][5]

$$|\Gamma_{bs}| \lesssim 2.4 \times 10^{-3} \left(\frac{m_{Z'}/g'}{300\ \mathrm{GeV}} \right). \tag{9}$$

We do not address here the origin of these peculiar flavor structures in the lepton and quark sectors.[6] We acknowledge, however, that explicit flavor completions of this model could lead to correlated effects in other meson-physics and leptonic observables [19,31–36].

III. EXPLAINING MUON-RELATED ANOMALIES

New contributions to the muon anomalous magnetic moment are induced at the one-loop level through the diagram in Fig. 2, yielding

$$\Delta a_\mu^{\mathrm{NP}} = \frac{|y|^2}{32\pi^2} \frac{m_\mu^2}{m_{\chi_0}^2} F_g(\tau)[1 + A(\tau, \delta)], \tag{10}$$

with m_μ the muon mass and $A(\tau, \delta) \equiv \frac{F_g(\tau/(1+\delta))}{(1+\delta)F_g(\tau)}$. The loop function F_g is given in Eq. (A1). For a degenerate dark spectrum, we have $F_g(1) = 1/12$ and $A(1, 0) = 1$, so that the discrepancy Δa_μ in Eq. (1) is accommodated for $m_{\chi_0} \simeq 45|y|\mathrm{GeV}$. As we will see, the LHCb anomalies typically require a significant scalar versus pseudoscalar mass splitting $\delta \gg 1$, with a dark-lepton versus DM mass splitting roughly in the interval $\tau \in [1, \delta]$. Equation (1) is then accommodated for $m_{\chi_0} \simeq 32|y|(1 + 2/\delta)\mathrm{GeV}$ ($\tau \simeq 1$) or $m_{\chi_0} \simeq 55|y|/\sqrt{\delta}\mathrm{GeV}$ ($\tau \simeq \delta \gg 1$). (See Appendix for details.)

[4]Even if absent at the tree-level, kinetic mixing is loop induced through $B_\mu - X_\mu$ vacuum polarization effects with exchanged dark fermions L and Q. While for the charges assumed in Table I, these loop contributions are logarithmically UV sensitive [29], they can be made finite in complete models by introducing additional vectorlike fermions with appropriate $\mathrm{U}(1)_X$ and SM charges [25].

[5]This bound could be significantly weakened, if the Z' also coupled to flavor-changing right-handed down-quark currents with a strength of $\mathcal{O}(10\%)$ of their left-handed counterparts [30]. We do not consider such a scenario here.

[6]Constructing an UV completion of our model with such flavor structures is by itself an interesting endeavor, which we leave for future work.

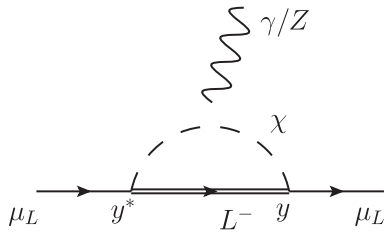

FIG. 2. One-loop NP contribution to $g_\mu - 2$ and the Z coupling to muons. The photon γ is to be attached to L^-, and the Z boson to either of the fermion lines.

Consider now the $b \to s\mu^+\mu^-$ anomalies. At the b-quark mass scale, the NP amplitude is described by the effective Hamiltonian

$$\mathcal{H}_{\rm eff}^{\rm NP} = -\frac{\alpha G_F}{2\sqrt{2}\pi} V_{tb}V_{ts}^* \sum_i C_i \mathcal{O}_i + \text{H.c.}, \tag{11}$$

where α and G_F are, respectively, the fine-structure and Fermi constants, V_{ij} are CKM matrix elements, and the sum runs over the operators ($\ell = e, \mu$)

$$\mathcal{O}_9^\ell \equiv \bar{b}\gamma_\rho(1-\gamma_5)s\bar{\ell}\gamma^\rho\ell, \tag{12}$$

$$\mathcal{O}_{10}^\ell \equiv \bar{b}\gamma_\rho(1-\gamma_5)s\bar{\ell}\gamma^\rho\gamma_5\ell. \tag{13}$$

A global fit to leptonic, semileptonic and radiative B decays favors effective couplings to muons [21,37]

$$C_9^\mu = -C_{10}^\mu \simeq -0.5 \tag{14}$$

and negligible electron coefficients $C_{9,10}^e \simeq 0$.[7] In our model, $C_{9,10}^\mu$ are induced at the one-loop level through the diagram in Fig. 3, which gives

$$C_9^\mu = -C_{10}^\mu = -\frac{g'^2}{4m_{Z'}^2}\Lambda_{\rm SM}^2 \frac{|V_{tb}V_{ts}^*|}{V_{tb}V_{ts}^*}\Gamma_{bs}\Gamma_{\mu\mu}, \tag{15}$$

where $\Lambda_{\rm SM} = [2\sqrt{2}\pi/(\alpha G_F|V_{tb}V_{ts}^*|)]^{1/2} \simeq 50$ TeV is the scale of the SM contributions. Accommodating the $b \to s\mu^+\mu^-$ anomalies as in Eq. (14), while respecting the ΔM_{B_s} bound from Eq. (9), yields a lower bound on the muon form factor,

[7]A slightly better fit is obtained for a pure vector NP coupling to muon pairs, i.e., for $C_9^\mu \simeq -1$ and $C_{10}^\mu \simeq 0$ [12,37]. In scenarios where C_9^μ is induced at the loop level (as in our model), it is difficult to achieve such a large effect. However, from a model-building perspective the solution $C_9^\mu = -C_{10}^\mu \approx -0.5$ with chiral interactions is more inviting, as it allows a natural implementation of weak isospin gauge invariance.

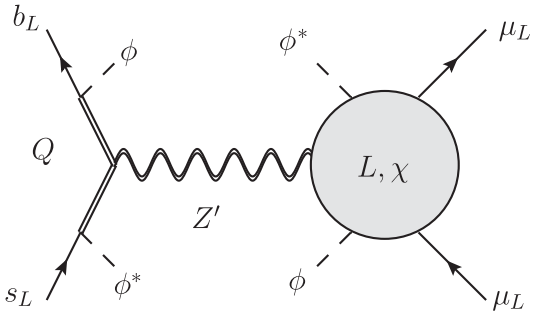

FIG. 3. Leading NP contribution to the Wilson coefficients $C_{9,10}^\mu$. The shaded disk denotes the radiatively induced Z' coupling to muon pairs at zero momentum exchange, as shown in Fig. 1.

$$\Gamma_{\mu\mu}(m_b^2) \gtrsim 0.029\left(\frac{m_{Z'}/g'}{300\ \text{GeV}}\right). \tag{16}$$

Since the Z' coupling to muons is small, cf. Eq. (8), this bound is fulfilled only for a light Z' boson, as well as a large mass splitting δ. In a spectrum with moderate tuning, $\delta \lesssim 10$, this leads to

$$m_{Z'} \lesssim 110\text{–}270\ \text{GeV}\left(\frac{g'}{3}\right)\left(\frac{|y|}{3}\right)^2, \tag{17}$$

where the two mass values correspond to $\tau \simeq \delta$ and $\tau \sim 1$, respectively. For $m_{Z'}$ around the weak scale, LHCb anomalies thus require rather large couplings y and g', while a perturbative upper bound of $m_{Z'} \lesssim 8.5$ TeV ($\tau \simeq \delta$) or $m_{Z'} \lesssim 20$ TeV ($\tau \simeq 1$) applies for $g' = y = 4\pi$. We discuss collider constraints on such a Z' in Sec. IV. To summarize, assuming perturbative couplings $g' \simeq 3$ and a mild tuning of the dark spectrum, $\delta \simeq 10$, both the $g_\mu - 2$ and LHCb anomalies are accommodated for parameters values interpolating between the two limiting cases

$$\begin{aligned}(1)\ &\tau \simeq \delta, \quad |y| \simeq 6, \quad g' \simeq 3, \quad \delta \simeq 10,\\ &m_{\chi_0} \simeq 100\ \text{GeV}, \quad m_{Z'} \simeq 300\ \text{GeV},\end{aligned} \tag{18}$$

$$\begin{aligned}(2)\ &\tau \simeq 1, \quad |y| \simeq 2, \quad g' \simeq 3, \quad \delta \simeq 10,\\ &m_{\chi_0} \simeq 70\ \text{GeV}, \quad m_{Z'} \simeq 150\ \text{GeV}.\end{aligned} \tag{19}$$

IV. COLLIDER CONSTRAINTS

We now analyze the relevant constraints on our model from EW precision measurements at the LEP experiments, as well as from the first LHC run. First of all, our model implies sizable radiative corrections to the $Z\mu\bar{\mu}$, $Z\nu_\mu\bar{\nu}_\mu$ and $W^+\mu\bar{\nu}_\mu$ couplings from the one-loop diagram in Fig. 2. The EW gauge couplings are shifted by ($V = W, Z$)

$$\frac{\delta g}{g_{\rm SM}} = \frac{|y|^2}{32\pi^2} F_V(\tau, r_q), \tag{20}$$

where the one-loop function F_V is found in Eq. (A4) and $r_q \equiv q^2/m_L^2$. The vertex correction is m_V^2/m_L^2 suppressed at the V pole, $q^2 = m_V^2$. While the QED part of the couplings at zero momentum is protected by gauge invariance, the weak isospin part is not corrected at one-loop level, since the SM Higgs doublet does not directly couple to the dark sector.[8] Despite this parametric suppression, the lightness of the dark-sector states, $m_{\chi_0} \simeq 100$ GeV and $m_L \simeq 100$–400 GeV, together with a relatively large Yukawa coupling, $y \gtrsim 2$, typically shifts the SM gauge couplings by one to a few per mill, which is in mild tension ($\sim 1\sigma$–3σ) with LEP data [10]. This tension may be relieved in a more sophisticated version of the minimal model considered here.

Our model also predicts a series of signatures at hadron colliders, most notably muon pair production through a resonant Z', as well as signals of large missing energy with muon pairs and/or jets. Current LHC limits on a Z' resonance with SM-like couplings to fermions are around $m_{Z'} \gtrsim 3$ TeV [39,40]. However, Z' production in our model only occurs through sea-quark ($b\bar{s} + s\bar{b}$) annihilations and is thus strongly suppressed. On the other hand, the Z' dominantly decays into muon pairs and neutrinos. Given the conditions Eqs. (9) and (16) on the Z' couplings to SM fermions and for Z' masses that accommodate the LHCb anomalies as in Eq. (17), the branching ratios to leptons are both ~40%–50%, depending on the value of the Yukawa coupling.[9] We thus find the cross section for Z' production with decay into $\mu^+\mu^-$ to be of $\mathcal{O}$(fb), which is an order of magnitude below current limits at the 8 TeV LHC. The next-to-leading Z' branching ratio is $\mu^+\mu^-\chi_0\chi_0$, ranging from 2% to 10% for large Yukawa couplings. Along the same lines, monojet signatures from the direct production of a DM pair in association with a hard jet from initial state radiation (ISR) lie at least one order of magnitude below the current LHC sensitivity [42,43].

The EW production of L^+L^- pairs leads to a signature with dimuons and missing energy, which resembles the one used in searches for smuons, the supersymmetric partners of the muon. The only difference with our signal lies in the spin of the produced particles. However, it was shown in Ref. [44] that results for slepton searches in simplified models could safely be applied to the production of fermion pairs decaying into a fermion and a scalar. We use SmodelS [45], a tool designed to decompose the signal of any NP model into simplified topologies, and compare the predictions to the exclusion limits set by the ATLAS and CMS slepton searches [46,47]. We find that the 8 TeV LHC sets strong constraints on the mass of L^-, even stronger than for smuons, because L^+L^- pair production cross sections are significantly larger. Dark-lepton masses $m_L \lesssim 450$ GeV are excluded, except if the mass splitting with the DM is sufficiently small, $m_L - m_{\chi_0} \lesssim 60$ GeV. In this region, the dimuon signal is overwhelmed with SM background. Similar searches at LEP2 lead to the lower bound of $m_L \gtrsim 100$ GeV [48].

V. DARK MATTER RELIC ABUNDANCE

The DM candidate in our model is the lightest component of the scalar χ, which we assume to be χ_0. It is largely leptophilic, as follows from the charge assignments in Table I. For a spectrum as in Eq. (6), DM annihilation proceeds dominantly into $\mu^+\mu^-$ and $\nu_\mu\bar{\nu}_\mu$ through t-channel exchange of L^- and L^0, respectively, as shown in Fig. 4. Coannihilation processes with the other dark states are negligible for $\tau \simeq \delta \gg 1$, while annihilation into Z' pairs is negligible as long as $m_{Z'} \gtrsim m_{\chi_0}$. The resulting annihilation cross section is d-wave suppressed in the chiral limit, $\sigma_{\bar{l}l} v \sim v^4$ [49]. Adding up final-state muons and neutrinos, the thermal average is[10]

$$\langle \sigma_{\bar{l}l} v \rangle = \frac{a_d}{x^2 m_{\chi_0}^2} + \mathcal{O}(x^{-3}), \qquad a_d \equiv \frac{|y|^4}{2\pi(1+\tau)^4}, \tag{21}$$

where v is the relative DM-DM velocity, and $x \equiv m_{\chi_0}/T$, the DM mass-to-temperature ratio. For the parameter set in Eq. (18), $a_d \simeq 0.02$ is a typical value. In the freeze-out approximation [51], the relic density is

$$\Omega_\chi^{\bar{l}l} h^2 \simeq 8.5 \times 10^{-11} \left(\frac{3 x_f^3 m_{\chi_0}^2}{a_d \sqrt{g_*} \mathrm{GeV}^2} \right) \times [1 + R_f]^{-1}, \tag{22}$$

where $g_* \simeq 80$ is the effective number of relativistic degrees of freedom at $T_f \simeq 1$–10 GeV and

$$R_f \simeq 2.6 \times 10^{-18} \left(\frac{\sqrt{g_*} m_{\chi_0}}{a_d \mathrm{GeV}} \right) x_f^{3/2} \exp(x_f) \tag{23}$$

accounts for the reduced DM abundance at freeze-out. Since DM annihilation $\langle \sigma_{\bar{l}l} v \rangle \propto x^{-2}$ is d-wave suppressed

[8] This implies that the operator $(H^\dagger D_\rho H)\bar{l}\gamma^\rho l + \mathrm{H.c.}$ (relevant for the $Z\mu^+\mu^-$ coupling), as well as $H^\dagger \sigma \cdot W_{\mu\nu} H B^{\mu\nu}$ and $|H^\dagger D_\mu H|^2$ (which respectively shift the S and T parameters [38]), are only induced at the two-loop level. Effects on S and T are thus sufficiently small to evade the LEP bounds.

[9] We use CalcHEP 3.4 [41] to compute the Z' production cross sections and partial decay widths assuming the spectrum in Eq. (18).

[10] In our analytic discussion, we work at leading order in $v \simeq 0.1$–0.3 at freeze-out in a nonrelativistic expansion, while in our numerical analysis the relic density is computed with micrOMEGAs [50].

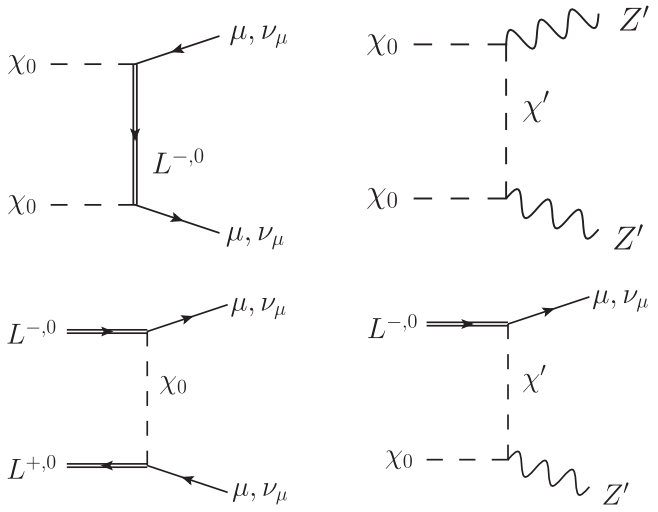

FIG. 4. Dominant amplitudes for DM-DM annihilation (top panels) and DM-L or L-L coannihilation (bottom panels) in the early Universe.

during freeze-out, DM decoupling is slightly delayed, which results in a larger (power of) $x_f \simeq 25$ in Eq. (23) compared to s- or p-wave annihilation. In a scenario as in Eq. (18), $R_f \simeq \mathcal{O}(1)$ is a significant correction.

Without this correction and for fixed mass splittings τ and δ, the relic density scales as $\Omega_\chi^{\bar{l}l} h^2 \propto m_{\chi_0}^2/|y|^4$, i.e., like $(\Delta a_\mu \times C_9^\mu)^{-1}$, in our model. Using Eqs. (10) and (15), Eq. (22) reads ($R_f = 0$)

$$\Omega_\chi^{\bar{l}l} h^2 \simeq 0.01 \times I(\tau,\delta)\left(\frac{x_f}{25}\right)^3\left(\frac{100\ \text{GeV}}{m_{Z'}/g'}\right) \times \left(\frac{287\times 10^{-11}}{\Delta a_\mu}\right)\left(\frac{0.5}{|C_9|}\right), \tag{24}$$

where $I(\delta,\tau) \equiv F_g(\tau)[1 + A(\tau,\delta)]F_{Z'}(\tau,\delta)(1+\tau)^4$. Assuming parameter values of Eq. (18) (with $\tau \gg 1$) in order to accommodate the muon anomalies yields $I(\tau,\delta) \simeq 20$–40. Recalling that $R_f \sim \mathcal{O}(1)$, this gives a relic abundance close to the observed value from Planck data [52]

$$\Omega_\chi^{\text{obs}} h^2 = 0.1199 \pm 0.0022. \tag{25}$$

However, for parameter values of Eq. (19) (with $\tau \simeq 1$), $I(\tau,\delta) \simeq \mathcal{O}(1)$ and the resulting relic density is typically a factor of 5–10 smaller than observed. We stress that the d-wave suppression of scalar DM annihilation into lepton pairs through a heavy fermion mediator plays an important role in the above prediction. Had DM annihilation been p-wave (s-wave) dominated, the same parameter space would have predicted a relic density smaller by a factor of about 10 (hundred).

Since the two-body process in Eq. (21) is strongly suppressed,[11] parametrically subdominant annihilation channels can become relevant. The leading such contribution is three-body annihilation $\chi_0\chi_0 \to \mu^+\mu^-\gamma$ with an extra photon in the final state. Virtual internal Bremsstrahlung (VIB) from the dark lepton L^+ lifts the kinematic suppression, so that annihilation proceeds through an s-wave [49,53]. The thermally averaged cross section for this process is [54]

$$\langle \sigma_{\mu\bar{\mu}\gamma} v\rangle = \frac{\alpha}{32\pi^2}\frac{|y|^4}{m_{\chi_0}^2} F_\gamma(\tau), \tag{26}$$

with the function F_γ defined in Eq. (A6). Following Ref. [54], we find that VIB suppresses the predicted relic density in Eq. (22) by a factor $R_\gamma \equiv \Omega_\chi^{\bar{l}l+\mu\bar{\mu}\gamma}/\Omega_\chi^{\bar{l}l} < 1$, which in the chiral limit and at zero relative DM velocity $v = 0$ approximates[12]

$$R_\gamma^{-1} \simeq 1 + \frac{3\alpha}{16\pi} x_f^2 (1+\tau)^4 F_\gamma(\tau). \tag{27}$$

The function $(1+\tau)^4 F_\gamma(\tau)$ strongly decreases with τ, yielding to a suppression factor of $R_\gamma \simeq 0.5(0.9)$ for $\tau \simeq 1$ ($\tau \gg 1$). Therefore, VIB is a significant contribution only in particular regions of parameter space, where coannihilation in addition strongly depletes the relic density.

In order to assess the robustness of the prediction for the relic density in Eq. (24) more quantitatively, we broadly scan around the parameter sets in Eqs. (18) and (19) that accommodate the muon anomalies, varying the parameters in the following ranges:

$$\begin{aligned} &50 \le m_{\chi_0} \le 450\ \text{GeV}, \qquad 200 \le m_{Z'} \le 1000\ \text{GeV},\\ &1 \lesssim \tau \le 20, \qquad 0 \lesssim \delta \le 20, \qquad 1 \le |y| \le 4\pi,\\ &1 \le g' \le 5. \end{aligned} \tag{28}$$

We restrict the tuning among the scalar components to maximally 10%. Furthermore, we impose the condition that the $g_\mu - 2$ and the LHCb anomalies quoted in Sec. I are explained within one standard deviation, and that the collider limits on dimuon and missing energy production from Sec. IV are satisfied. We then compute the resulting relic density with micrOMEGAs [50]. As shown in Fig. 5, for many parameter points where the dominant DM

[11]Contributions from the finite muon mass, which induce s-wave annihilation into dileptons, are of $\mathcal{O}(10^{-3})$ and thus are negligible.

[12]This expression is only an estimate of the total $\mathcal{O}(\alpha)$ correction to DM annihilation into the final state $\mu^+\mu^-$. A complete calculation that takes account of the full DM-velocity dependence requires us to also include virtual corrections. We leave a thorough treatment of the extra-photon contribution for future implementation in micrOMEGAs.

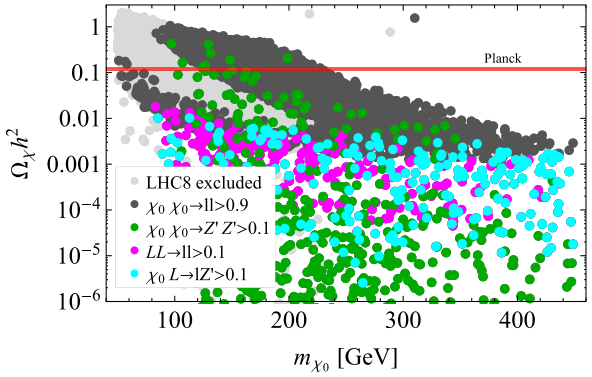

FIG. 5 (color online). DM relic density $\Omega_\chi^{\bar{l}l}h^2$ as a function of the DM mass, m_{χ_0}, as predicted by the muon-related collider anomalies. Light gray points: excluded by 8 TeV LHC and LEP2 constraints on di-lepton plus missing energy signals from L^+L^- production Dark gray points: DM annihilation dominated by $\chi_0\chi_0 \to \mu^+\mu^-, \nu_\mu\bar{\nu}_\mu$. Green points: additional contributions (∼10% and more) from $\chi_0\chi_0 \to Z'Z'$. Magenta and cyan points: coannihilation through $LL \to ll$ and $\chi_0 L \to Z'l$, respectively, accounts for more than ∼10% of the total $\langle\sigma v\rangle$. Red band: 3σ range of $\Omega_\chi^{\rm obs}$ from Planck [52]. (See text for details.)

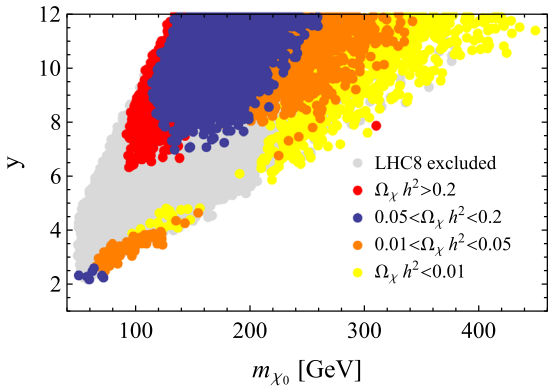

FIG. 6 (color online). Dark Yukawa coupling y as a function of m_{χ_0} required to accommodate the muon anomalies within one standard deviation. Gray points are excluded by 8 TeV LHC and LEP2 constraints on dilepton plus missing energy signals from L^+L^- production. All other points correspond to dominant DM annihilation $\chi_0\chi_0 \to \mu^+\mu^-, \nu_\mu\bar{\nu}_\mu$, i.e., to the dark gray points in Fig. 5. Colors depict different amounts of the associated DM relic abundance; $\Omega_\chi^{\rm obs}h^2$ is obtained in the blue regions.

annihilation process is $\chi_0\chi_0 \to \bar{l}l$ (dark gray points) the relic density falls within an order of magnitude of $\Omega_\chi^{\rm obs}$, as expected from the above discussion. However, we also find that the predicted relic density can be much lower than the observed value, if the spectrum significantly deviates from Eq. (6). First of all, Ω_χ rapidly decreases when the DM mass increases above $m_{\chi_0} \gtrsim 200$ GeV. Satisfying Δa_μ with heavier DM masses requires a lighter dark-lepton L or, equivalently, smaller values of τ, leading to a significantly stronger DM annihilation into muon(-neutrino) pairs, see Eq. (21). Secondly, when $m_{Z'} \lesssim m_{\chi_0}$, the process $\chi_0\chi_0 \to Z'Z'$ efficiently depletes the DM relic abundance (the green points in Fig. 5). Moreover, when $m_L \simeq m_{\chi_0}$, corresponding to $\tau \simeq 1$, coannihilation processes with dark leptons (see Fig. 4) can also contribute to significantly reducing the DM density (magenta points in Fig. 5). This is particularly pronounced when Z' can be produced in the final state via the coannihilation processes $\chi_0 L^\pm(L^0) \to Z'\mu^\pm(\nu_\mu)$ (the cyan points in Fig. 5).

The two regions where the relic density is in agreement with the observed value closely resemble the parameter limits favored by the muon anomalies from Eqs. (18) and (19), respectively. These are

(1) The *heavy lepton scenario*. In this region, $m_{\chi_0} \approx$ 120–250 GeV and $m_L \gtrsim 450$ GeV, corresponding to $\tau \gg 1$, because of collider constraints on dilepton plus missing energy signals. As a result, the Yukawa coupling is strong, $|y| \gtrsim 6$, to ensure a large enough DM annihilation rate into lepton pairs.

(2) The *compressed scenario*. This second region corresponds to $m_{\chi_0} \approx 50$–80 GeV and $m_L \approx 100$–120 GeV, i.e., to $\tau \simeq 1$. It evades collider constraints because of the relatively small gap between DM and dark-lepton masses. Since DM annihilation is larger for $\tau \simeq 1$, see Eq. (21), a smaller Yukawa coupling is needed to ensure enough DM annihilation. However, as the muon anomalies request $|y| \gtrsim 2$, it is more challenging to obtain the observed relic density $\Omega_\chi^{\rm obs}h^2 \sim 0.1$ in the compressed scenario.

In Fig. 6, we illustrate the range of y needed to obtain various amounts of DM relic abundance Ω_χ, as a function of the DM mass, for the case of dominant DM annihilation into lepton pairs, $\chi_0\chi_0 \to \bar{l}l$.

VI. DARK MATTER DETECTION

The DM candidate in our model is dominantly leptophilic, which makes direct detection very challenging. Furthermore, since Z and Higgs boson couplings to χ_0 pairs are absent at tree level and loop-induced vector interactions are momentum suppressed, any signal of spin-independent DM-nucleus scattering lies below the sensitivity of current direct detection experiments.[13]

On the other hand, VIB in DM annihilation, discussed in Sec. V, plays an important role for indirect detection today. The suppression of tree-level DM annihilation into dilepton states, $\sigma_{\bar{l}l}v \sim v^4$, is even stronger than during freeze-out, as velocities today are around $v_{\rm halo} \sim 10^{-3}$ in our galactic

[13]The leading contribution to spin-independent (SI) scattering arises at the two-loop level from the effective interaction $\chi_0\chi_0\bar{q}q$. For the spectrum in Eq. (18), we estimate its size to be $\sigma_{\rm SI}^p \sim \mathcal{O}(10^{-50})$ cm^2, which lies below the coherent neutrino background [55] for DM masses around the weak scale. Spin-dependent interactions are strongly velocity suppressed in the nonrelativistic limit.

neighborhood. Photons from VIB are thus the dominant signal in gamma-ray searches of indirect detection experiments. The annihilation cross sections due to VIB in our two scenarios from Eqs. (18) and (19), integrated over the photon energy spectrum, are

$$(1)\ \langle\sigma_{\mu\bar{\mu}\gamma}v\rangle \simeq 3\times10^{-27}\ \mathrm{cm}^3/\mathrm{s}\quad \text{and}$$
$$(2)\ \langle\sigma_{\mu\bar{\mu}\gamma}v\rangle \simeq 7\times10^{-25}\ \mathrm{cm}^3/\mathrm{s}, \tag{29}$$

respectively. The smaller cross section in the heavy lepton scenario (1) is mainly due to the suppression $\langle\sigma_{\mu\bar{\mu}\gamma}v\rangle \sim 1/\tau^4$ in the limit of large mass splitting, $\tau \gg 1$.

For small mass splitting $\tau \simeq 1$, the energy spectrum of VIB photons is peaked close to the end point $E_\gamma = m_{\chi_0}$ [53], which leaves a characteristic signature in indirect detection experiments. The Fermi satellite has performed a dedicated search for spectral features in VIB using data from space regions near the Galactic center collected during pass 7 [56]. Assuming a spectral shape characteristic for $1 \lesssim \tau \lesssim 2$, this search sets an experimental upper bound on VIB,

$$\langle\sigma_{\mu\bar{\mu}\gamma}v\rangle \lesssim 10^{-27}\ \mathrm{cm}^3/\mathrm{s}, \tag{30}$$

for $m_{\chi_0} \lesssim 100$ GeV.

More inclusive analyses of Fermi data from dwarf galaxies constrain the overall gamma-ray flux emitted in DM annihilations [57,58]. These searches are sensitive to the sum of secondary photons (from final-state radiation off muons and muon decays) and VIB. Though the cross section bounds are generally weaker than from the dedicated VIB search, they constrain $\langle\sigma_{\mu\bar{\mu}\gamma}v\rangle$ for scenarios with large τ, which the VIB spectral search is not sensitive to.

In Fig. 7, we compare the predictions of DM annihilation through VIB in our model with the bounds from Fermi. We focus only on points in the scan range of Eq. (28), which satisfy the muon-related anomalies and yield a DM relic density of $0.01 < \Omega_\chi h^2 < 1$. The upper bound on $\langle\sigma_{\mu\bar{\mu}\gamma}v\rangle$ from Fermi's VIB search was derived in Ref. [56] (the dashed red line). It excludes part of the compressed scenario (2) with $1 \le \tau \le 2$ (the blue points), where VIB is particularly large [see Eq. (29)]. The bound cannot directly be applied to scenarios with $\tau > 2$ (the dark gray points), for which the photon energy peak is much less pronounced and the VIB search thus significantly loses sensitivity. Inclusive gamma-ray searches in nearby dwarf galaxies lead to a less strigent upper bound on $\langle\sigma_{\mu\bar{\mu}\gamma}v\rangle$ (the black line), which we adopt from [59] under the assumption that VIB dominates photon emission in DM annihilation. Parameter space regions with $\tau > 2$ and a relic density comparable to the observed value lie at least an order of magnitude below the bound from dwarf galaxies. Notice that the Fermi bounds on the cross section are based on data from pass 7. With the recently released pass 8 data set [58],

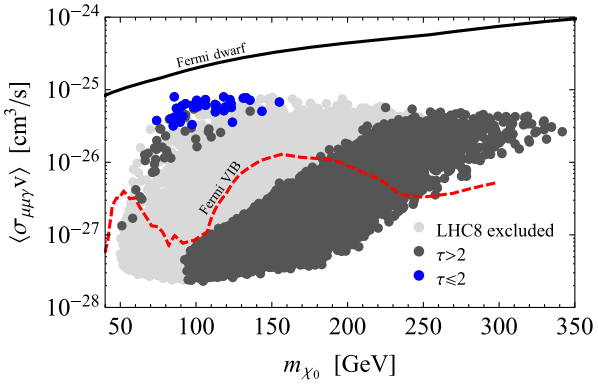

FIG. 7 (color online). Thermally averaged cross section for DM annihilation into $\mu^+\mu^-\gamma$ in our galactic neighborhood today as a function of the DM mass, as predicted by muon-related collider anomalies. All points satisfy $0.01 \lesssim \Omega_\chi h^2 \lesssim 1$. Light gray points are excluded by 8 TeV LHC and LEP2 constraints on dilepton plus missing energy signals from L^+L^- production. Points with $\tau \le 2$ ($\tau > 2$) are shown in blue (dark gray). The black line denotes the Fermi LAT upper bound on photon emission in nearby dwarf galaxies, while the dashed red line represents the Fermi LAT limit on spectral features in VIB signals [56], which applies only for $1 \lesssim \tau \lesssim 2$, i.e., to the blue points.

we expect these bounds to improve by about an order of magnitude. The inclusive gamma-ray search might thus start probing the heavy lepton scenario, as well as the part of the compressed scenario, which does not display a significant spectral feature.

VII. CONCLUSIONS AND OUTLOOK

We have shown that a minimal phenomenological model featuring a new Abelian force with radiative couplings to left-handed muons can simultaneously explain the observed muon anomalies in $g_\mu - 2$ and $b \to s\mu^+\mu^-$ transitions through one-loop effects of new particles charged under the associated $\mathrm{U}(1)_X$ symmetry. The spontaneous breaking of $\mathrm{U}(1)_X$ down to a $\mathbb{Z}_2$ symmetry induces a sizable mass splitting among the scalar states of the dark sector, which controls the radiative Z' coupling to muon pairs and ensures that the lightest new state is a cosmologically stable neutral scalar. Hence, addressing both the $g_\mu - 2$ and $b \to s\mu^+\mu^-$ anomalies in this way typically yields a leptophilic DM candidate, which is hardly visible in direct DM detection searches.

All muon-related anomalies can simultaneously be accommodated with Z' and DM masses around the EW scale and large dark couplings $y, g' \gtrsim 2$. While DM annihilation is governed by the same interactions that explain the muon anomalies, the large Yukawa coupling y typically allows for a thermal DM relic density surprisingly close to the observed value, unless significant coannihilation is present. This result partially relies on the fact that the dominant DM annihilation into lepton pairs

is d-wave velocity suppressed, thus favoring larger couplings compared to scenarios with s- or p-wave annihilation. The velocity suppression furthermore depletes final-state and secondary photon emission from DM annihilation into muon pairs in our galactic neighborhood today. Signals in indirect detection experiments are therefore dominantly due to hard photons from VIB.

Another key ingredient of the model is a new vectorlike lepton doublet, which in particular mediates tree-level DM annihilation to leptons. Being $\mathbb{Z}_2$ odd, these dark leptons do not mix with SM leptons, nor do they couple to the SM Higgs field at tree level. This prevents the otherwise severe constraints from Higgs and EW precision observables, most notably oblique corrections. The only significant effect on low-energy observables lies in the Z-boson coupling to muon pairs or invisible states, which is in mild tension with precision LEP data at the Z pole.

At high-energy hadron colliders, dark leptons around the EW scale are produced in pairs through Drell-Yan-like processes. We recast available LHC searches for muon pairs with missing energy in the context of supersymmetry and find a general lower bound on the dark-lepton mass, $m_L \gtrsim 450$ GeV. A good fit to all muon anomalies with a relic density around the observed value is still possible in two characteristic parameter regions. These are (1) the heavy lepton scenario with dark leptons above 450 GeV and a DM candidate in the range $125 \lesssim m_{\chi_0} \lesssim 250$ GeV, and (2) the compressed scenario with a light dark sector around 70 GeV and a small splitting between the dark-lepton and DM masses, $m_L - m_{\chi_0} \lesssim 60$ GeV, to which current dimuon searches are insensitive.

Accommodating the LHCb anomalies in the heavy lepton scenario typically requires a very large Yukawa coupling not far from the nonperturbative regime, which suggests an UV completion of our model not far above the TeV scale. Weaker couplings are nevertheless possible in the compressed scenario thanks to the larger DM annihilation cross section at freeze-out in the limit of a small mass gap with the dark lepton. However, this small mass splitting induces a sharply peaked photon energy spectrum from VIB in DM annhihilation today. Dedicated searches for such a signature by the Fermi experiment partially exclude the compressed scenario as a simultaneous explanation of muon anomalies and thermal DM. Interestingly, indirect detection experiments are thus complementary to collider searches, which are not sensitive to the compressed scenario. Probing the heavy lepton scenario through indirect detection, however, is difficult, since the photon energy peak from VIB is much less pronounced for large mass gaps and inclusive gamma-ray searches are less sensitive.

We point out that B_s meson mixing significantly limits the $Z'\bar{b}s$ coupling to at most a few per mill. Given the loop suppression of the Z' coupling to muon pairs in our model, the $b \to s\mu^+\mu^-$ anomalies imply new contributions to B_s meson mixing right around the corner, while they are only a possibility in most Z' models explaining the LHCb anomalies with tree-level couplings to leptons. This is an interesting implication of simultaneously addressing the $g_\mu - 2$ discrepancy. Moreover, the absence of a Z' coupling to valence quarks and its constrained interaction with bottom and strange quarks strongly limits direct Z' production at hadron colliders. In particular, resonant Z' production followed by decays in muon pairs is well below the sensitivity of the 8 TeV LHC. Similar conclusions apply to monojet searches.

Prospects are good that our model can be tested in the near future. Since dedicated DM searches through direct detection are not sensitive to our scenario, the 14 TeV LHC will be the most important experiment to probe it. We expect that the reach for dark muons should be improved by close to a factor of 2 during run II, similar to what was found for smuons [60,61]. Thus, the parameter space of the heavy lepton scenario will be fully covered, since the $g_\mu - 2$ anomaly sets an upper bound $m_L \lesssim 800$ GeV for couplings $y < 4\pi$. The spectrum in the compressed scenario (which is already under pressure from indirect DM detection searches) is more difficult to access through direct searches. One possibility would be to exploit the monojet signature from the direct production of dark muons in association with an ISR jet in the highly compressed region, where the dark muon decay leads to an invisible signature. Precision measurements of Z-boson couplings at a future e^+e^- collider would give complementary indirect information about the dark sector.

ACKNOWLEDGMENTS

We wish to thank Diego Guadagnoli, Kenneth Lane, Gilad Perez and Alexander Pukhov for the discussions. The work of C. D. is supported by the "Investissements d'avenir, Labex ENIGMASS." S. W. is supported in part by the U.S. National Science Foundation under Grant No. PHY-1212635.

APPENDIX: AUXILIARY FUNCTIONS

The $g_\mu - 2$ loop function from the diagrams in Fig. 2 is

$$F_g(\tau) \equiv \frac{1}{6(\tau-1)^4}[\tau^3 - 6\tau(\tau - \log\tau) + 3\tau + 2]. \quad \text{(A1)}$$

For degenerate dark (pseudo-)scalar and dark-lepton masses, one has $F_g(\tau = 1) = 1/12$, while in the decoupling limit of the dark lepton the function decreases as $F_g(\tau \gg 1) \simeq 1/(6\tau)$.

For the dominant Z' coupling to SM lepton pairs at zero momentum, corresponding to the diagrams in Fig. 1, the loop function is

$$F_{Z'}(\tau,\delta) \equiv \int_0^1 du \int_0^{1-u} dv \log\left[1+\frac{v\delta}{\tau+(u+v)(1-\tau)}\right] + \log\left[1-\frac{v\delta}{\tau+(u+v)(1+\delta-\tau)}\right]. \quad \text{(A2)}$$

For degenerate scalar and pseudoscalar masses, $F_{Z'}(\tau,0)=0$, while in the limit of large mass splitting Eq. (A2) approximately gives

$$F_{Z'}(\tau,\delta\gg 1) \simeq \frac{1}{2}\log\delta - \frac{\tau^2(1+\log\tau)-3\tau+2}{2(\tau-1)^2}. \quad \text{(A3)}$$

Moreover, Eq. (A2) approaches a constant when both mass splittings are comparably large, $F_{Z'}(\tau\sim\delta\gg 1)\simeq 1/4$.

The one-loop corrections to the $V=W,Z$ couplings to SM leptons from Fig. 2 are proportional to the loop function

$$F_V(\tau,r_q) = 2\int_0^1 du\int_0^{1-u} dv\left[\frac{1-u-v}{u+v}\log f(\tau,u+v)+\frac{1-f(\tau,u+v)+2uvr_q}{f(\tau,u+v)-uvr_q}-\log(f(\tau,u+v)-uvr_q)\right], \quad \text{(A4)}$$

where $f(\tau,x)\equiv x+(1-x)/\tau$ and $r_q\equiv q^2/m_L^2$, with q^2 the squared momentum of the external gauge boson. The loop function falls off at low q^2 as

$$F_V(\tau,r_q\ll 1) \simeq \frac{r_q\tau}{36(\tau-1)^4}[(\tau-1)(7\tau^2-29\tau+16) + 6(3\tau-2)\log\tau]. \quad \text{(A5)}$$

Finally, the function controlling the approximate three-body correction from VIB to the relic density is

$$F_\gamma(\tau) = (\tau+1)\left[\frac{\pi^2}{6}-\log^2\left(\frac{\tau+1}{2\tau}\right)-2\mathrm{Li}_2\left(\frac{\tau+1}{2\tau}\right)\right] + \frac{4\tau+3}{\tau+1}+\frac{4\tau^2-3\tau-1}{2\tau}\log\left(\frac{\tau-1}{\tau+1}\right), \quad \text{(A6)}$$

where $\mathrm{Li}_2(x)=-\int_0^1 dy\log(1-xy)/y$ is the polylogarithm function of second order.

[1] G. Bennett *et al.* (Muon $g-2$ Collaboration), Phys. Rev. D **73**, 072003 (2006).
[2] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. Lett. **111**, 191801 (2013).
[3] R. Aaij *et al.* (LHCb Collaboration), Reports No. LHCb-CONF-2015-002 and CERN-LHCb-CONF-2015-002, 2015.
[4] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. Lett. **113**, 151601 (2014).
[5] R. Aaij *et al.* (LHCb Collaboration), arXiv:1506.08777.
[6] R. Pohl *et al.*, Nature (London) **466**, 213 (2010).
[7] V. Barger, C.-W. Chiang, W.-Y. Keung, and D. Marfatia, Phys. Rev. Lett. **106**, 153001 (2011).
[8] D. Tucker-Smith and I. Yavin, Phys. Rev. D **83**, 101702 (2011).
[9] D. Akerib *et al.* (LUX Collaboration), Phys. Rev. Lett. **112**, 091303 (2014).
[10] K. Olive *et al.* (Particle Data Group), Chin. Phys. C **38**, 090001 (2014).
[11] V. Khachatryan *et al.* (CMS and LHCb Collaborations), Nature (London) **522**, 68 (2015).
[12] S. Descotes-Genon, J. Matias, and J. Virto, Phys. Rev. D **88**, 074002 (2013).
[13] W. Altmannshofer and D. M. Straub, Eur. Phys. J. C **73**, 2646 (2013).
[14] F. Beaujean, C. Bobeth, and D. van Dyk, Eur. Phys. J. C **74**, 2897 (2014); **74**, 3179(E) (2014).
[15] D. Ghosh, M. Nardecchia, and S. A. Renner, J. High Energy Phys. 12 (2014) 131.
[16] W. Altmannshofer and D. M. Straub, Eur. Phys. J. C **75**, 382 (2015).
[17] T. Hurth, F. Mahmoudi, and S. Neshatpour, J. High Energy Phys. 12 (2014) 053.
[18] R. Gauld, F. Goertz, and U. Haisch, Phys. Rev. D **89**, 015005 (2014).
[19] W. Altmannshofer, S. Gori, M. Pospelov, and I. Yavin, Phys. Rev. D **89**, 095033 (2014).
[20] D. Aristizabal Sierra, F. Staub, and A. Vicente, Phys. Rev. D **92**, 015001 (2015).
[21] G. Hiller and M. Schmaltz, Phys. Rev. D **90**, 054014 (2014).
[22] S. Biswas, D. Chowdhury, S. Han, and S. J. Lee, J. High Energy Phys. 02 (2015) 142.
[23] B. Gripaios, M. Nardecchia, and S. A. Renner, J. High Energy Phys. 05 (2015) 006.
[24] I. de Medeiros Varzielas and G. Hiller, J. High Energy Phys. 06 (2015) 072.
[25] P. J. Fox, J. Liu, D. Tucker-Smith, and N. Weiner, Phys. Rev. D **84**, 115006 (2011).
[26] J. Kumar and J. D. Wells, Phys. Rev. D **74**, 115017 (2006).
[27] E. J. Chun, J.-C. Park, and S. Scopel, J. High Energy Phys. 02 (2011) 100.
[28] A. Djouadi, O. Lebedev, Y. Mambrini, and J. Quevillon, Phys. Lett. B **709**, 65 (2012).

[29] B. Holdom, Phys. Lett. **166B**, 196 (1986).
[30] A. J. Buras, F. De Fazio, and J. Girrbach, J. High Energy Phys. 02 (2013) 116.
[31] S. L. Glashow, D. Guadagnoli, and K. Lane, Phys. Rev. Lett. **114**, 091801 (2015).
[32] A. Crivellin, L. Hofer, J. Matias, U. Nierste, S. Pokorski, and J. Rosiek, arXiv:1504.07928 [Phys. Rev. D (to be published)].
[33] A. Crivellin, G. D'Ambrosio, and J. Heeck, Phys. Rev. Lett. **114**, 151801 (2015).
[34] A. Crivellin, G. D'Ambrosio, and J. Heeck, Phys. Rev. D **91**, 075006 (2015).
[35] C.-J. Lee and J. Tandean, J. High Energy Phys. 08 (2015) 123.
[36] A. Celis, J. Fuentes-Martin, M. Jung, and H. Serodio, Phys. Rev. D **92**, 015007 (2015).
[37] W. Altmannshofer and D. M. Straub, arXiv:1503.06199.
[38] M. E. Peskin and T. Takeuchi, Phys. Rev. Lett. **65**, 964 (1990).
[39] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. D **90**, 052005 (2014).
[40] V. Khachatryan *et al.* (CMS Collaboration), J. High Energy Phys. 04 (2015) 025.
[41] A. Belyaev, N. D. Christensen, and A. Pukhov, Comput. Phys. Commun. **184**, 1729 (2013).
[42] G. Aad *et al.* (ATLAS Collaboration), Eur. Phys. J. C **75**, 299 (2015).
[43] V. Khachatryan *et al.* (CMS Collaboration), Eur. Phys. J. C **75**, 235 (2015).
[44] C. Arina, M. E. C. Catalan, S. Kraml, S. Kulkarni, and U. Laa, J. High Energy Phys. 05 (2015) 142.
[45] S. Kraml, S. Kulkarni, U. Laa, A. Lessa, V. Magerl *et al.*, arXiv:1412.1745.
[46] G. Aad *et al.* (ATLAS Collaboration), J. High Energy Phys. 05 (2014) 071.
[47] V. Khachatryan *et al.* (CMS Collaboration), Eur. Phys. J. C **74**, 3036 (2014).
[48] J. Abdallah *et al.* (DELPHI Collaboration), Eur. Phys. J. C **31**, 421 (2003).
[49] F. Giacchino, L. Lopez-Honorez, and M. H. G. Tytgat, J. Cosmol. Astropart. Phys. 10 (2013) 025.
[50] G. Bélanger, F. Boudjema, A. Pukhov, and A. Semenov, Comput. Phys. Commun. **192**, 322 (2015).
[51] P. Gondolo and G. Gelmini, Nucl. Phys. **B360**, 145 (1991).
[52] P. Ade *et al.* (Planck Collaboration), arXiv:1502.01589.
[53] T. Bringmann, L. Bergstrom, and J. Edsjo, J. High Energy Phys. 01 (2008) 049.
[54] T. Toma, Phys. Rev. Lett. **111**, 091301 (2013).
[55] J. Billard, L. Strigari, and E. Figueroa-Feliciano, Phys. Rev. D **89**, 023524 (2014).
[56] T. Bringmann, X. Huang, A. Ibarra, S. Vogl, and C. Weniger, J. Cosmol. Astropart. Phys. 07 (2012) 054.
[57] M. Ackermann *et al.* (Fermi-LAT Collaboration), Phys. Rev. Lett. **107**, 241302 (2011).
[58] M. Ackermann *et al.* (Fermi-LAT Collaboration), arXiv:1503.02641.
[59] A. Geringer-Sameth and S. M. Koushiappas, Phys. Rev. Lett. **107**, 241303 (2011).
[60] I. Melzer-Pellmann and P. Pralavorio, Eur. Phys. J. C **74**, 2801 (2014).
[61] J. Eckel, M. J. Ramsey-Musolf, W. Shepherd, and S. Su, J. High Energy Phys. 11 (2014) 117.