Inelastic Dark Matter at the LHC Lifetime Frontier: 
ATLAS, CMS, LHCb, CODEX-b, FASER, and MATHUSLA

Asher Berlin

Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA
(Dated: October 5, 2018)

Felix Kling

Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA

Recent years, an extensive program has emerged in the search for GeV-scale hidden sectors, involving beam dump, fixed-target, and collider experiments [1, 2]. If such hidden sectors couple to the Standard Model through mediators heavier than \( \sim 10 \) GeV, their production at low-energy accelerators is kinematically suppressed, leaving open significant pockets of viable parameter space. We investigate this scenario in models of inelastic dark matter, which give rise to visible signals at various existing and proposed LHC experiments, such as ATLAS, CMS, LHCb, CODEX-b, FASER, and MATHUSLA. These experiments can leverage the large center of mass energy of the LHC to produce GeV-scale dark matter from the decays of dark photons in the cosmologically motivated mass range of \( \sim 1 \) – 100 GeV. We also provide a detailed calculation of the radiative dark matter-nucleon/electron elastic scattering cross section, which is relevant for estimating rates at direct detection experiments.

I. INTRODUCTION

In recent years, an extensive program has emerged in the search for GeV-scale hidden sectors, involving beam dump, fixed-target, and collider experiments [1, 2]. If the mediator coupling hidden sectors to the Standard Model (SM) are heavier than roughly \( \sim 10 \) GeV (corresponding to the center of mass energy of \( B \)-factories), their production at low-energy accelerators is kinematically suppressed. As a result, there are large regions of viable parameter space for GeV-scale particles that couple to the SM at the \( \mathcal{O}(10^{-2}) \)-level through \( \mathcal{O}(10) \) GeV mediators.

These scenarios are directly motivated from considerations of light dark matter (DM). It is well-known that models of thermal DM at or below the GeV scale often require the presence of new light mediators in order to facilitate the depletion of the DM energy density in the early universe [3]. Although stated less often, SM measurements (e.g., of the invisible widths of the \( Z \) and Higgs bosons) extend these claims to masses as large as \( m_{\text{DM}} \sim \text{few} \times 10 \) GeV (see, e.g., Refs. [4, 5]).

For \( m_{\text{DM}} \lesssim 50 \) GeV, measurements of the cosmic microwave background (CMB) severely restrict how these new mediators can couple DM to the SM [6]. For instance, considerations of the CMB necessitate that the DM annihilation rate to electromagnetically charged particles is suppressed at late times. This requirement typically also leads to small annihilation rates today, worsening the discovery potential for indirect detection experiments searching for the products of local DM annihilations. In this work, we focus on a class of models that naturally leads to negligible annihilation rates at the time of recombination.

Although originally motivated as an explanation for the longstanding DAMA anomaly [7], models of inelastic DM (iDM) constitute a viable and compelling paradigm for light thermal DM. In these models, DM couples to the SM only by interacting with a nearly degenerate excited state. For relative mass-splittings larger than \( \sim \mathcal{O}(10^{-6}) \), inelastic scattering at direct detection experiments is kinematically suppressed by the small DM virial velocity, \( v \sim \mathcal{O}(10^{-3}) \). However, this kinematic suppression is overcome in relativistic settings, such as accelerators, where the DM and excited state can be directly produced. For mass-splittings above a few MeV, the excited state can decay back to DM and a pair of SM fermions, often on collider timescales. In this case, searches for visible displaced vertices at colliders and fixed-target experiments are sensitive to cosmologically motivated masses and couplings.

To date, the majority of existing studies on visible signals of iDM have focused on mediator masses at or below the GeV scale (see, however, Ref. [8]). In this work, we investigate mediators heavier than \( \sim 10 \) GeV, where most high-intensity accelerators have little sensitivity due to their limited center of mass energy. This large open region of viable parameter space is well-suited for high-energy machines, such as the LHC. In recent years, different search strategies have been proposed for detecting long-lived particles (LLPs) at the existing LHC experiments, ATLAS, CMS, and LHCb. Furthermore, various dedicated experiments have been suggested, such as CODEX-b, FASER, and MATHUSLA. In this paper, we focus on the potential sensitivity of these existing and proposed experiments to visible signals of iDM near the GeV scale.

The remainder of this paper is structured as follows. In Sec. II, we motivate the existence of additional me-
diators for DM below the weak scale. In Sec. III, we introduce the particular iDM model that we investigate throughout this work and discuss various aspects of its cosmology in Sec. IV. In Sec. V, we summarize existing constraints on these models. In Sec. VI, we present strategies for the LHC to search for displaced lepton-jets and time-delayed tracks at ATLAS and CMS and displaced vertices at LHCb, CODEX-b, FASER, and MATHUSLA. In Sec. VII, we discuss details related to our modeling of iDM production at the LHC and present our results in Sec. VIII. We summarize our conclusions in Sec. IX. Detailed discussions of experimental energy thresholds, kinetically-mixed dark photons, loop-induced DM-nucleon/electron elastic scattering cross sections, and dark photon production via vector meson mixing are provided in Appendices A, B, C, and D, respectively.

II. THERMAL DARK MATTER BELOW THE WEAK SCALE

Thermal DM that couples to the SM solely through the electroweak force must be heavier than the GeV scale. This is most often stated in terms of the Lee-Weinberg bound [3], which in its modern form relates the DM mass to the weak scale \( m_Z \sim 100 \text{ GeV} \), the temperature at matter-radiation equality \( T_{\text{eq}} \sim \text{eV} \), and the Planck mass \( m_{\text{Pl}} \sim 10^{18} \text{ GeV} \),

\[
m_{\text{DM}} \gtrsim \frac{m_Z^2}{(T_{\text{eq}} m_{\text{Pl}})^{1/2}} \sim \text{GeV}.
\]  

Eq. (1) implies that sub-GeV thermal DM motivates the presence of new light mediators [9]. This philosophy has guided various experimental developments in the search for well-defined “thermal targets” [1, 2, 10].

This picture can be sharpened with the inclusion of additional experimental inputs. It is well-known that thermal DM that is lighter than \( \sim 50 \text{ GeV} \) and annihilates through the exchange of an electroweak boson is strongly constrained by null results of recent direct detection experiments as well as by the invisible Z and Higgs widths [4, 5]. This can be stated parametrically, similar to Eq. (1). If \( m_{\text{DM}} \lesssim m_Z \), then the DM annihilation rate \( \langle \sigma v \rangle \) is smaller than the value required by thermal freeze-out \( \sigma v_{\text{th}} \sim 1/T_{\text{eq}} m_{\text{Pl}} \),

\[
\frac{\langle \sigma v \rangle}{\sigma v_{\text{th}}} \lesssim \mathcal{O}(10^{-2}) \times \frac{\Gamma_{\text{inv}} T_{\text{eq}} m_{\text{Pl}}}{m_Z^2} \lesssim \mathcal{O}(10^{-2}) ,
\]  

where \( \Gamma_{\text{inv}} \) is the additional contribution to the invisible Z width from decays to DM pairs, and in the second inequality we have enforced \( \Gamma_{\text{inv}} \lesssim \mathcal{O}(\text{MeV}) \) [11]. A similar statement holds for couplings to the SM Higgs. Such claims can be avoided, for instance, if \( 2 m_{\text{DM}} \approx m_{Z,h} \), in which case invisible Z or h decays to DM particles are kinematically suppressed and DM annihilations are resonantly enhanced. Therefore, aside from this tuning, DM lighter than \( \sim 50 \text{ GeV} \) motivates the presence of additional light mediators that are nearby in mass to the DM itself. A well-motivated example is the dark photon, as discussed in Sec. III. In the next section, we also introduce a class of models in which the DM-dark photon coupling is directly responsible for thermal freeze-out in the early universe.

III. INELASTIC DARK MATTER

Models of inelastic DM (iDM) were first proposed to resolve the longstanding DAMA anomaly [7]. Although such explanations are no longer viable in light of searches conducted by more recent direct detection experiments (see, e.g., Ref. [12]), the overall framework still constitutes a compelling paradigm for light thermal DM.

In this section, we outline a concrete model that we will investigate throughout this work. We focus on a particular implementation of iDM, involving a Dirac pair of two-component Weyl fermions (\( \eta \) and \( \xi \)) that are oppositely charged under a broken \( U(1)_D \) symmetry (\( Q_D \xi = \pm 1 \)). The corresponding massive gauge boson is the dark photon, denoted as \( A' \). The dark photon couples to SM hypercharge through the kinetic mixing term

\[
\mathcal{L} \supset \frac{\epsilon}{2 \cos \theta_w} A'_{\mu \nu} B^{\mu \nu} ,
\]  

where \( \epsilon \ll 1 \) is a dimensionless parameter controlling the size of the mixing and \( \theta_w \) is the Weinberg angle [13, 14]. For example, this kinetic mixing is generated radiatively if there exist any particles charged under both \( U(1)_D \) and hypercharge. In this case, the natural expected size is \( \epsilon \sim \epsilon_D/16 \pi^2 \sim 10^{-3} \), where \( \epsilon_0 = 10 \text{ MeV} \). The corresponding massive gauge boson is the dark photon, denoted as \( A' \). A detailed discussion of kinetic mixing for general \( A' \) masses is provided in Appendix B.

A Dirac mass, \( m_D \), involving the DM spinors, \( \eta \) and \( \xi \), is allowed by all symmetries of the theory. In the case that a dark Higgs is responsible for the mass of the \( A' \), it is natural to imagine that the spontaneous breaking of \( U(1)_D \) also gives rise to Majorana mass terms \( (\delta_\eta \xi) \) for these hidden sector fermions. At energies below the scale of \( U(1)_D \)-breaking, we parametrize the effective Lagrangian as

\[
-\mathcal{L} \supset m_D \eta \xi + \frac{1}{2} \delta_\eta \eta^2 + \frac{1}{2} \delta_\xi \xi^2 + \text{h.c.} ,
\]  

where \( m_D, \delta_\eta, \delta_\xi > 0 \). Since \( \delta_\eta \xi \neq 0 \) explicitly breaks \( U(1)_D \), it is technically natural to take \( \delta_\eta \xi \ll m_D \). In this limit, the mass eigenstates correspond to a pseudo-
Dirac pair given by
\[
\chi_1 \simeq \frac{i}{\sqrt{2}} (\eta - \xi) \\
\chi_2 \simeq \frac{1}{\sqrt{2}} (\eta + \xi),
\]
where the phase of \( \chi_1 \) is fixed to guarantee a positive mass term. \( \chi_1 \) and \( \chi_2 \) are nearly degenerate, with masses \( m_1 \simeq m_2 \) given by
\[
m_{1,2} \simeq m_D + \frac{1}{2} (\delta_\eta + \delta_\xi).\tag{6}
\]

For later convenience, we define the fractional mass-splitting between \( \chi_1 \) and \( \chi_2 \),
\[
\Delta \equiv \frac{m_2 - m_1}{m_1} \simeq \frac{\delta_\eta + \delta_\xi}{m_D} \ll 1.\tag{7}
\]

In four-component notation, \( \chi_{1,2} \) are described by Majorana spinors. In the pseudo-Dirac limit, they couple off-diagonally to the dark photon, i.e.,
\[
\mathcal{L} \supset ie_D A'_\mu \bar{\chi}_1 \gamma^\mu \chi_2 + \mathcal{O}(\delta_\eta, \delta_\xi/m_D),\tag{8}
\]
where \( e_D \) is the \( U(1)_D \) gauge coupling. In general, there also exist terms that couple \( A' \) diagonally to \( \chi_1 \) and \( \chi_2 \). However, these are either suppressed by the mass hierarchy, \( \delta_\eta, \delta_\xi/m_D \ll 1 \), or vanish exactly when \( \delta_\eta = \delta_\xi \).

Throughout this work, we will ignore such contributions, and assume that \( A' \) couples purely off-diagonally (inelastically) to \( \chi_1 \chi_2 \) pairs.

## IV. COSMOLOGY

Kinetic mixing between \( U(1)_D \) and hypercharge couples \( \chi_{1,2} \) to the SM. In particular, if the degree of kinetic mixing satisfies
\[
\epsilon \gtrsim \mathcal{O}(10^{-8}) \times \left( \frac{m_{A'}}{\text{GeV}} \right)^{1/2},\tag{9}
\]
then the \( A' \) equilibrates with the visible sector in the early universe. In this case, \( \chi_{1,2} - A' \) interactions often lead to a large thermal \( \chi_{1,2} \) number density that must be depleted at later times in order to be cosmologically viable. Since \( m_2 \gtrsim m_1 \), the depletion of \( \chi_2 \) in the early universe can efficiently proceed through dynamics within the hidden sector, e.g., through scattering \( \chi_2 \chi_2 \rightarrow \chi_1 \chi_1 \) or decays \( \chi_2 \rightarrow \chi_1 + \cdots \). \( \chi_1 \) and \( \chi_2 \) are odd under a global \( Z_2 \) subgroup of \( U(1)_D \), which is preserved by the Lagrangian terms of Eqs. \( (4) \) and \( (8) \). Since \( \chi_1 \) is the lightest state charged under this symmetry, it is stable on cosmological timescales and plays the role of DM.

If \( m_{A'} \gtrsim m_A \), then DM freeze-out is dictated by annihilations to pairs of on-shell dark photons \( (\chi_1 \chi_1 \rightarrow A' A') \), followed by \( A' \rightarrow e^+ e^- , \mu^+ \mu^- , \cdots \). Such processes are unsuppressed at late times and for \( m_{A'} \lesssim \text{few} \times 10 \text{ GeV} \) are severely constrained from considerations of the CMB and various indirect detection searches (see Ref. [15] for a recent review). If instead, \( m_{A'} \gtrsim m_1 \), then freeze-out predominantly proceeds through coannihilations directly into SM particles, \( \chi_1 \chi_2 \rightarrow A'^* + f \bar{f} \), where \( f \) is a SM fermion. For \( m_f \ll m_1 \simeq m_2 \ll m_{A'} \), the thermally-averaged annihilation rate for this process is parametrically of the form
\[
\langle \sigma v \rangle \sim \mathcal{O}(10^2) \frac{\alpha_D \alpha_{em} \epsilon^2 m_1^2}{m_{A'}^4} e^{-\Delta x},\tag{10}
\]
where \( \alpha_D \equiv e_D^2/4\pi, x \equiv m_1/T, \) and \( T \) is the temperature of the photon bath. The number density of \( \chi_2 \) is Boltzmann suppressed at late times. As a result, the coannihilation rate in Eq. \( (10) \) is exponentially small at temperatures much below the \( \chi_1 - \chi_2 \) mass-splitting (\( \Delta x \gg 1 \)).

As we will see below, for \( m_{A'} \gtrsim 10 \text{ GeV} \), the degree of kinetic mixing is restricted by terrestrial experiments to \( \epsilon \lesssim \text{few} \times 10^{-2} \). Imposing this limit on \( \epsilon \) and demanding that Eq. \( (10) \) is of the required size for thermal freeze-out, i.e., \( \langle \sigma v \rangle \sim \langle \sigma v_{bh} \rangle \sim 1/T m_{P1} \), then implies an upper bound on the mass-splitting,
\[
\Delta \lesssim \mathcal{O}(0.1) \left( 1 - \log \left( \frac{m_{1}}{10 \text{ GeV}} \right) \left( \frac{m_{A'}}{3 m_1} \right)^2 \right),\tag{11}
\]

where we have fixed \( x \simeq 20 \) to the typical value at freeze-out. For mass-splittings of this size, the exponential Boltzmann factor in Eq. \( (10) \) corresponds to only a mild suppression at temperatures relevant for the freeze-out of \( \chi_1 \). However, at much later times, such as during recombination, the annihilation rate is completely negligible. This significantly alleviates strong bounds derived from distortions of the CMB [6, 16].

We numerically solve the Boltzmann equations for the relic abundance of \( \chi_1 \), incorporating hadronic and leptonic final states, as detailed in Refs. [8, 21–23]. Our results are shown in Fig. 1. We do not incorporate intermediate GeV-scale hadronic resonances in this figure since our main focus throughout this work is on \( m_1 \gtrsim \text{few} \times \text{GeV} \). Although the dark photon predominantly mixes with the SM photon for \( m_{A'} \lesssim \mathcal{O}(10) \text{ GeV} \), it has additional non-negligible mixing with the \( Z \) for larger masses. A detailed discussion of \( A' - \gamma, Z \) mixing is given in Appendix B. In addition to \( A' \) exchange, we include processes where the DM annihilates through an intermediate \( Z \left( \chi_1 \chi_2 \rightarrow A'^* \right) \), which is also described by Eq. (6). These contributions have a sizable effect on the DM relic abundance near the \( Z \)-resonance at \( m_1 \simeq m_Z/2 \). We have also checked

\[1 \) This is due to an enhanced charge-conjugation symmetry under which \( A' \rightarrow -A' \) and \( \chi_{1,2} \rightarrow \mp \chi_{1,2} \).

\[2 \) Compared to the relic abundance calculation performed in
FIG. 1. Viable parameter space in the $\epsilon - m_{A'}$ plane for a model of inelastic dark matter with $\alpha_D = 0.1$. Along the colored contours, $\chi_1$ acquires an abundance (through $\chi_1\chi_2 \to A'^*Z \to f\bar{f}$) that is in agreement with the observed dark matter energy density. Solid, dashed, and dotted contours correspond to relative $\chi_1 - \chi_2$ mass-splittings of $\Delta = 0.02$, 0.05, and 0.1, respectively, while the orange and blue sets of contours correspond to the $A' - \chi_1$ mass ratios of $m_{A'}/m_1 = 3$ and 5, respectively. Existing constraints from a monophoton search at BaBar [17, 18] and precision electroweak measurements at LEP [19, 20] are shown as solid gray regions. The large area of viable and cosmologically motivated parameter space corresponding to $10 \, \text{GeV} \lesssim m_{A'} \lesssim 100 \, \text{GeV}$ and $10^{-3} \lesssim \epsilon \lesssim \text{few} \times 10^{-2}$ is the primary region of interest in this work.

that our analytic results agree well with the numerical output of the publicly available codes FeynRules [25] and MicrOMEGAS [26].

In Fig. 1, we show regions of parameter space in the $\epsilon - m_{A'}$ plane where $\chi_1$ freezes out with an abundance that is in agreement with the observed DM energy density for various choices of model parameters. The solid, dashed, and dotted colored contours correspond to the fractional mass-splittings of $\Delta = 0.02$, 0.05, and 0.1, respectively, whereas the orange and blue sets of contours correspond to the mass ratios $m_{A'}/m_1 = 3$ and $m_{A'}/m_1 = 5$. We have fixed $\alpha_D = 0.1$ throughout. As shown in Fig. 1, for fixed $\alpha_D$ and $m_1$, smaller mass-splittings ($\Delta$) and mass ratios ($m_{A'}/m_1$) enhance the rate for $\chi_1 - \chi_2$ coannihilations in the early universe and facilitate smaller values of $\epsilon$ in the cosmologically favored parts of parameter space.\(^3\) Also shown in Fig. 1 are existing constraints from measurements of electroweak precision observables at LEP [19, 20] and a monophoton search at BaBar [18]. These bounds, along with several others, will be discussed in detail in Sec. V.

If $\alpha_D$ is comparable to SM gauge couplings, then $\alpha_D \gg \alpha_{em}\epsilon^2$ for the cosmologically motivated parameter space of Fig. 1. As a result, decays of $A'$ to $\chi_1\chi_2$ pairs dominate over those directly to SM particles if kinematically accessible. In the limit that $m_{A'} \gg m_1 \simeq m_2$, the corresponding partial width is

$$\Gamma(A' \to \chi_1\chi_2) \simeq \frac{\alpha_D m_{A'}}{3}.$$  \hspace{1cm} (12)

For dark photons produced in terrestrial experiments, such decays are the dominant production mechanism for DM and its excited state. Once a $\chi_1\chi_2$ pair is produced in this manner, $\chi_1$ leaves the detector as missing energy/momentum. The heavier counterpart, $\chi_2$, is unstable and eventually undergoes a three-body decay to DM and a pair of SM particles through an off-shell $A'$, i.e., $\chi_2 \to \chi_1A'^* \to \chi_1f\bar{f}$, giving rise to hidden valley-like signatures [28]. The rate for $\chi_2$ to decay to $\chi_1$ and a pair of SM leptons is approximately given by

$$\Gamma(\chi_2 \to \chi_1\ell^+\ell^-) \simeq \frac{4\epsilon^2\alpha_{em}\alpha_D\Delta^5 m_1^5}{15\pi m_{A'}^4},$$  \hspace{1cm} (13)

where we have taken $m_{A'} \gg m_1 \simeq m_2 \gg m_\ell$. For $\epsilon \ll 1$ and $\Delta \lesssim 0.1$, the proper lifetime of $\chi_2$ is easily macroscopic,

$$c\tau_{\chi_2} \sim \mathcal{O}(m) \times \left(\frac{\alpha_D}{0.1}\right)^{-1} \left(\frac{\epsilon}{10^{-2}}\right)^{-2} \left(\frac{m_{A'}}{3m_1}\right)^4 \left(\frac{\Delta}{0.05}\right)^{-5} \left(\frac{m_1}{10 \, \text{GeV}}\right)^{-1}.$$  \hspace{1cm} (14)

In experiments with baselines significantly smaller than the boosted lifetime of $\chi_2$ (such as the $B$-factories BaBar and Belle II), the process $A' \to \chi_1\chi_2$ often registers as an invisible decay. On the other hand, longer baseline experiments can efficiently search for the visible decay products from the long-lived $\chi_2$. While previous studies have shown that existing and future $B$-factories and fixed-target experiments will definitively test large regions of the cosmologically motivated parameter space for $m_{A'} \lesssim 10 \, \text{GeV}$ [24, 29], heavier masses remain relatively unconstrained (see, however, Ref. [8]). Amongst the proposed experiments in the search for long-lived and weakly-coupled hidden sectors, those that are able to leverage the high center of mass energy of the LHC.

Ref.s [8, 24], we do not find a significant enhancement of the annihilation rate at $m_{A'} \simeq m_Z$, due to interference between $A'$- and $Z$-exchange diagrams that contribute to the annihilation amplitude with opposite signs. This is shown explicitly in Eq. (B22) of Appendix B.
(such as the ones discussed in this work) pose a strong advantage in this regard. In the remainder of this work, we focus on the details of \( A' \) production and the detection of \( \chi_2 \) in the large open region of parameter space corresponding to \( 10 \text{ GeV} \lesssim m_{A'} \lesssim 100 \text{ GeV} \).

V. EXISTING CONSTRAINTS

For the sake of completeness, we briefly review in this section the multitude of existing constraints on models of light iDM. These constraints can be classified into model-independent bounds and those arising from invisible or visible signals. Below, we will only discuss the most stringent of these that are directly relevant to the models under investigation in this work. For a more complete discussion, we refer the reader to the relevant sections of, e.g., Refs. [24, 29].

Experimental measurements that are independent of unknown dynamics in the hidden sector constitute a highly powerful probe of these models. A prime example of this arises from electroweak precision measurements performed at LEP. The kinetic mixing of Eq. (3) generically leads to various shifts in electroweak precision observables, and recent fits give a model-independent upper bound of \( \epsilon \lesssim 3 \times 10^{-2} \) for \( m_{A'} \lesssim 100 \text{ GeV} \) [19, 20]. For dark photon masses of \( 10 \lesssim m_{A'} \lesssim 100 \text{ GeV} \), these are the most stringent existing constraints. Similar tests of electroweak precision observables at the LHC are expected to improved upon these limits by a factor of \( \lesssim 2 \) in \( \epsilon \) [20].

For \( m_{A'} \lesssim 10 \text{ GeV} \), a monophoton search at BaBar places an upper bound on the kinetic mixing of \( \epsilon \lesssim 10^{-3} \). In models of iDM, this signal arises from \( e^+ e^- \rightarrow \gamma + A' \rightarrow \gamma + \chi_1 \chi_2 \). The \( \chi_2 \) (along with the \( \chi_1 \)) is registered as missing energy provided that it escapes the detector or if the leptons from its decay \( (\chi_2 \rightarrow \chi_1 \ell^+ \ell^-) \) fall below the required energy thresholds [17, 18]. For \( \Delta \lesssim 0.1 \), we find that \( \chi_2 \) is regarded as an invisible final state in this search for most relevant regions of parameter space. A similar search with much higher luminosity is ultimately expected to be performed at Belle II in the near future, with an improved sensitivity to \( \epsilon \) by as much as an order of magnitude [30–33].

Although not the primary focus of this paper, we also briefly discuss existing constraints for sub-GeV hidden sector masses. In particular, for \( m_{\chi_1,2} \lesssim m_{A'} \lesssim \text{GeV} \), strong bounds can be derived from searches for energy deposition from \( \chi_1,2 \) scatterings or \( \chi_2 \) decays in detectors placed downstream of high-intensity beam dump experiments, such as LSND [34, 35], E137 [36, 37], and MiniBooNE [38, 39]. For sizable \( \alpha_D \) and \( \chi_1 - \chi_2 \) mass-splittings of \( \Delta \sim \mathcal{O}(0.1) \), these searches lead to upper bounds of roughly \( \epsilon \lesssim \mathcal{O}(10^{-5}) - \mathcal{O}(10^{-3}) \) for \( 10 \text{ MeV} \lesssim m_1 \lesssim 1 \text{ GeV} \), respectively [24, 29].

VI. INELASTIC DARK MATTER AT THE LHC

In this section, we discuss dedicated searches at existing and proposed LHC experiments for visible signals of iDM. We consider on-shell production of dark photons for \( m_{A'} \lesssim \text{few} \times 100 \text{ GeV} \). A detailed discussion of the various \( A' \) production mechanisms will be provided later in Sec. VII. If \( m_{A'} \gtrsim m_1 + m_2 \), once produced, the dark photon promptly decays to a \( \chi_1 \chi_2 \) pair. \( \chi_1 \) is cosmologically stable and easily escapes the instrumental geometry without interacting. However, since \( \chi_2 \) is unstable and naturally long-lived, we will focus on dedicated searches for visible displaced vertices arising from its decay, \( \chi_2 \rightarrow \chi_1 f \bar{f} \), where \( f \) is a SM fermion. We now proceed by outlining proposed experimental searches at ATLAS, CMS, LHCb, MATHUSLA, CODEX-b, and FASER.

A. ATLAS/CMS

As is common in many DM models, the production of iDM at the LHC leads to missing (transverse) energy. In the parameter space of interest (\( m_{A'} \lesssim 100 \text{ GeV} \), as in Fig. 1), the missing energy is typically small (\( E_T \sim m_{A'} \)). As a result, the detection of such a signal is challenging given the trigger requirements and large background rates. However, if the decay length of the excited iDM state is not too large (\( c\tau_{\chi_2} \lesssim \mathcal{O}(m) \)), it often decays inside the ATLAS and CMS detectors leaving an additional visible component to the signal. In the following, we will consider two types of proposed search strategies for these LLPs, looking for either a displaced lepton-jet or a time-delayed signal.

1. Displaced Muon-Jet

A displaced decay of the excited iDM state, \( \chi_2 \), inside the LHC detectors is a striking signature and provides a powerful handle for background rejection. Roughly a quarter of the time, \( \chi_2 \) decays into muons when kinematically accessible, leading to a particularly clean signal: a so-called *displaced muon-jet* (DMJ). Such a search has been proposed by the authors of Ref. [8], and in the following we adapt their analysis strategy.

As suggested in Ref. [8], we utilize the monojet + \( E_T \) trigger, which requires missing energy balanced by a recoil jet with transverse momentum larger than \( p_{T,j} > 120 \text{ GeV} \). Such a trigger has been used in previous DM searches at both 8 TeV [40] and 13 TeV [41] runs of the LHC and we assume that a similar performance can be achieved at a high-luminosity run.

Due to the small mass-splitting between the iDM states (\( \Delta \ll 1 \)), the \( \chi_2 \) decay products are typically soft, such that \( p_{T,\mu} \sim \Delta \times p_{T,\chi_2} \ll p_{T,j} \). We require each muon to have a transverse momentum \( p_{T,\mu} > 5 \text{ GeV} \) which
FIG. 2. Schematic drawings of a timing layer at CMS (top-left), MATHUSLA (top-right), CODEX-b (bottom-left), and FASER (bottom-right), along with their locations with respect to the LHC ring. The red shaded region indicates the decay volume for each experiment.

roughly corresponds to the threshold for soft muon identification [42]. In order to be able to reconstruct the tracks with sufficient precision, they must hit the outer layers of the tracking system. We therefore require a radial displacement of the $\chi_2$ decay vertex of $r_{\chi_2} < 30$ cm. Finally, we demand that the muon tracks are sufficiently displaced and require a transverse impact parameter of $d_\mu > 1$ mm. In summary, this search strategy for displaced muon-jets at ATLAS and CMS requires

$$\text{DMJ : } p_{T,j} > 120 \text{ GeV}$$
$$p_{T,\mu} > 5 \text{ GeV}$$
$$r_{\chi_2} < 30 \text{ cm}$$
$$d_\mu > 1 \text{ mm}.$$  \hspace{1cm} (15)

When estimating the reach, we assume the expected integrated luminosity of the high-luminosity (HL) LHC, $\mathcal{L} = 3 \text{ ab}^{-1}$, and that backgrounds can be reduced to a negligible level as argued by the authors of Ref. [8].

2. Time-Delayed Tracks

An alternative search strategy using precision timing has been proposed in Ref. [43]. If $\chi_2$ decays after traversing a macroscopic distance, its decay products will arrive at the calorimeter delayed in time compared to SM particles that are promptly produced at the interaction point (IP). This time delay is due both to the reduced speed of the massive iDM state, $v_{\chi_2}$, and the increased path length of the displaced decay, $l_{\chi_2} + l_\ell$ ($\ell$ is a lepton from the decay of $\chi_2$), compared to a SM track with path length $l_{SM}$. We can estimate the time delay as $\Delta t = l_{\chi_2}/v_{\chi_2} + l_\ell/c - l_{SM}/c$, where for simplicity we have assumed that the decay products move along straight lines at the speed of light.

The CMS collaboration recently proposed the installation of a precision timing detector with a resolution of ~ 30 picoseconds. This timing layer would be located in front of the electromagnetic calorimeter (ECAL) with a radial size of $R = 1.17$ m and extending up to $z = 3.04$ m along the beam axis. While this upgrade was originally intended for pile-up reduction, its potential implementation in searches for LLPs has been investigated in Ref. [43]. A schematic drawing of this setup is shown in the top-left panel of Fig. 2.

As in Sec. VI A 1, we require the excited iDM state to decay leptonically. Following Ref. [43], we demand that at least one of the $\chi_2$ decay products has a time delay of $\Delta t > 0.3$ ns and require a recoil jet of $p_{T,j} > 30$ GeV to timestamp the primary vertex. Since no vertex reconstruction is required for signal identification, this search can make use of the entire decay volume inside the ECAL and access radial and longitudinal displacements of the decay vertex of $r_{\chi_2} < 1.17$ m and $z_{\chi_2} < 3.04$ m, re-
spectively. If there is to be a significant time delay, $\chi_2$ should decay sufficiently far away from the primary vertex. We therefore impose a minimal radial decay distance of $r_{\chi_2} > 5$ cm. Although the timing layer is expected to have a very low energy threshold, we require the leptons to have a transverse momentum of $p_{T,\ell} > 3$ GeV. This ensures that they travel with light speed along approximately straight trajectories within the magnetic field of the detector.

We consider two trigger options: a conservative scenario using the conventional trigger and an optimistic one using a timing-based trigger. For the conservative estimate, we use the conventional monojet + $E_T$ trigger and require a recoil jet with $p_{T,j} > 120$ GeV balancing the missing energy. The cuts in this scenario are

$$\text{Timing} (E_T) : \quad p_{T,j} > 120 \text{ GeV}$$
$$p_{T,\ell} > 3 \text{ GeV}$$
$$\Delta t > 0.5 \text{ ns}$$
$$5 \text{ cm} < r_{\chi_2} < 1.17 \text{ m}$$
$$z_{\chi_2} < 3.04 \text{ m}.$$ \hfill (16)

The presence of a leptonic signature in association with a sizable amount of missing transverse momentum eliminates most potential background sources. While this search does not rely on tracking information for signal reconstruction, it can provide a powerful handle for background rejection. For instance, tracking could be used to veto delayed tracks that are consistent with being produced at or close to the beam line. Alternatively, timing information alone might be sufficient to suppress backgrounds, as discussed in detail by the authors of Ref. [43]. In the following, we will assume that SM backgrounds can be reduced to negligible levels.

The search strategy outlined above is primarily limited by triggering requirements. To explore the full capability of precision timing in LLP searches, we also consider the optimistic case in which the time delay alone can be used for triggering. This allows us to lower the requirement on the transverse momentum of the recoil jet to $p_{T,j} > 30$ GeV, which is sufficient to timestamp the primary vertex. The relevant cuts are then given by

$$\text{Timing} (\Delta t) : \quad p_{T,j} > 30 \text{ GeV}$$
$$p_{T,\ell} > 3 \text{ GeV}$$
$$\Delta t > 0.3 \text{ ns}$$
$$5 \text{ cm} < r_{\chi_2} < 1.17 \text{ m}$$
$$z_{\chi_2} < 3.04 \text{ m}.$$ \hfill (17)

In both cases, we assume an integrated luminosity of $\mathcal{L} = 3 \text{ ab}^{-1}$ when estimating the reach.

B. Displaced Muons at LHCb

LHCb is a dedicated $B$-physics experiment with coverage in the forward direction and is expected to run with a triggerless readout in the near future, overcoming one of the main challenges of ATLAS and CMS. Furthermore, it has been shown that the detector’s vertexing and invariant mass resolution as well as its ability for particle identification makes it an excellent instrument to search for displaced muon pairs from, e.g., displaced visible decays of dark photons [44, 45]. Similar searches are also sensitive to soft energy deposition from the showering and hadronization in confining hidden sectors [46, 47].

We estimate LHCb’s sensitivity to displaced muons originating from $\chi_2$ decays in models of iDM by applying the baseline post-module search criteria of Ref. [45],

$$\text{LHCb} : \quad p_{T,\mu} > 0.5 \text{ GeV}$$
$$p_\mu > 10 \text{ GeV}$$
$$6 \text{ mm} < r_{\chi_2} < 22 \text{ mm}$$
$$2 < \eta_{\chi_2} < 5$$
$$2 < \eta_\mu < 5.$$ \hfill (18)

The first two requirements on the muon momenta suppress contributions from fake muons, whereas the remaining cuts on the transverse displacement and rapidity ensure that the displaced vertex is sufficiently separated from the beamline and efficiently registered in the Vertex Locator (VELO) with a dimuon identification efficiency of $\sim 50\%$. We note that slightly relaxing the muon momentum thresholds in this analysis could enlarge the discovery potential of LHCb to the signals discussed here. However, lower thresholds also significantly complicate estimates of background contributions to the signal region, and hence, we do not consider this possibility further.

Ref. [45] adopted an estimated background of 25 events per mass bin (originating from interactions with the detector material) in a search for resonant dimuons from the visible decays of long-lived dark photons. Although the muons from the three-body decay of $\chi_2$ do not reconstruct a resonance, such processes populate a significant fraction of events near the kinematic limit of $m_{\mu\mu} \lesssim \Delta \times m_1$. Strategies to additionally suppress backgrounds by leveraging kinematic handles of such processes at long-baseline experiments has been recently discussed in Ref. [48]. A dedicated analysis of this signal and the relevant backgrounds at LHCb is beyond the scope of this work. In estimating the projected sensitivity of LHCb to visible signals of iDM, we will conservatively demand 100 signal events, assuming the integrated luminosity expected by the end of the HL-LHC era, $\mathcal{L} = 300 \text{ fb}^{-1}$.

C. MATHUSLA

The sensitivity of LHC experiments that surround the various IPs (such as ATLAS, CMS, and LHCb) is suppressed for small mass-splittings, $\Delta \lesssim \mathcal{O}(0.1)$. In this case, the lifetime of $\chi_2$ is sufficiently large that only a small number of decays occur within the inner tracking
systems and its decay products are often too soft to overcome limitations from triggering/backgrounds.

Additional experiments have been proposed to perform dedicated searches for LLPs. The largest of these proposed detectors is MATHUSLA (the MAssive Timing Hodoscope for Ultra-Stable neutral pArticles) [49, 50]. Its planned location is on the surface near ATLAS or CMS. A schematic drawing, showing its position and size, is shown in the top-right panel of Fig. 2. The detector consists of a 200 m × 200 m × 20 m decay volume, positioned ∼100 m downstream from the IP and ∼100 m above the LHC beam. This large detector volume, covering ∼10% of the full solid angle, ensures a sizable geometric acceptance. The rock that is present between the LHC ring and the proposed detector location is expected to reduce QCD backgrounds originating from the IP to a negligible level.

The decay volume is covered with a scintillating layer to veto incoming charged particles such as high-energy muons from the IP. Placed on top of the decay volume is a ∼5 m thick tracking system, which is envisioned to consist of five tracking layers with a timing resolution of ∼1 ns and a spatial resolution of ∼1 cm. This would allow for displaced vertex reconstruction and reliable separation between the upward going LLP signal and the downward going cosmic ray background. Searches for LLPs are assumed to be background free.

In estimating MATHUSLA’s sensitivity, we require $\chi_2$ to decay in the decay volume,

\[
\text{MATHUSLA} : \begin{align*}
100 \text{ m} &< x_{\chi_2} < 120 \text{ m} \\
-100 \text{ m} &< y_{\chi_2} < 100 \text{ m} \\
100 \text{ m} &< z_{\chi_2} < 300 \text{ m}.
\end{align*}
\]  

Final state energy thresholds ranging from 200−600 MeV have been discussed in Ref. [49]. For our baseline analysis, we assume a luminosity of $\mathcal{L} = 3 \text{ ab}^{-1}$ and require an energy deposition of 600 MeV per track. In Appendix A, we discuss the effect of modifying this threshold on the projected sensitivity.

D. CODEX-b

While a detector located at the surface necessarily requires a large decay volume to collect a sufficient number of LLP decays, similar physics could be probed with a smaller decay volume if it is located closer to the IP. A possible location for such a detector has been identified in the LHCb cavern. As part of the the upcoming Run 3 upgrade of LHCb, a relocation of the data acquisition system is planned, and a large unoccupied space, shielded by a 3 m thick concrete radiation shield, will become available. CODEX-b (the COmpact Detector for EXotics at LHCb) has been proposed to be constructed in this space [51].

A schematic drawing of the detector location and size is shown in the bottom-left panel of Fig. 2. The decay volume is expected to have dimensions of 10 m × 10 m × 10 m and to be positioned ∼5 m downstream from the IP at a transverse distance of ∼26 m. This detector design therefore covers approximately 1% of the solid angle. All six sides of the detector would be equipped with a sextet of RPC tracking layers with an effective granularity of ∼1 cm and a timing resolution of ∼1 ns. This would allow for the reconstruction of a displaced vertex inside the decay volume with an $\mathcal{O}(1)$ signal efficiency. We follow Ref. [51] and additionally require the track energies to be above a threshold of ∼600 MeV. In Appendix A, we discuss the effect of modifying this threshold on the projected sensitivity.

To reduce background rates to a manageable level, the detector needs to be sufficiently shielded from the large flux of SM particles produced at the IP. In addition to the existing concrete shielding, an additional lead shield located close to the IP and covering the detector geometric acceptance is needed. A ∼4.5 m thick lead shield, corresponding to 25 nuclear interaction lengths, would reduce the background rates from most SM particles (e.g., neutrons, kaons, and pions) to negligible levels. The remaining background would predominantly arise from penetrating muons, leading to secondary neutrons or kaons at the far end of the concrete shield which enter the decay volume undetected and produce tracks by scattering on air. An additional scintillating veto to detect charged particles could remove such muon induced events if installed in the lead shield. In the following, we assume that CODEX-b can operate with negligible background. In order to estimate the corresponding sensitivity at the HL-LHC, we take LHCb’s ultimate luminosity to be $\mathcal{L} = 300 \text{ fb}^{-1}$. In our analysis, we require that $\chi_2$ decays inside the CODEX-b decay volume and that its decay products are sufficiently energetic,

\[
\text{CODEX-b} : \begin{align*}
26 \text{ m} &< x_{\chi_2} < 36 \text{ m} \\
-3 \text{ m} &< y_{\chi_2} < 7 \text{ m} \\
5 \text{ m} &< z_{\chi_2} < 15 \text{ m} \\
E_{\text{track}} &> 600 \text{ MeV}.
\end{align*}
\]  

E. FASER

New physics at high-energy colliders, such as the LHC, has traditionally been expected in the high transverse momentum region. This is also the case for the experiments discussed above, which are sensitive to LLPs that are produced with large transverse momentum. However, if new particles are light and weakly-coupled, this focus may be misguided. In this case, large event rates are only available at low transverse momentum that is comparable to the mass of the LLP, $p_T \sim M_{\text{LLP}}$. Such particles are predominantly produced in the very forward direction, collimated around the beam collision axis, and hence escape detection in typical LHC detectors. Moreover, due to their weak coupling to the SM and the large boost expected in the forward direction, these particles are natu-
rally long-lived and travel a macroscopic distance before decaying. A detector placed in the very forward region along the beam collision axis may therefore be optimal to detect these decays. FASER (the ForwArD Search ExpeRiment) is an experiment designed to take advantage of this opportunity and search for light LLPs in the very forward region. Previous studies have established FASER’s potential to discover light new particles [52–55].

As shown in the bottom-right panel of Fig. 2, FASER would be placed along the beam collision axis, several hundred meters downstream of the ATLAS or CMS IP (and after the LHC tunnel begins to curve). A particularly promising location has been identified a few meters outside of the main LHC tunnel, 480 m downstream from the ATLAS IP, in the side tunnel TI18. This space was formerly used to connect the SPS and LEP tunnels but is currently empty and unused. At this location, the beam collision axis intersects with TI18 close to where it merges with the main LHC tunnel at the construction hall UJ18. For concreteness, we assume that the decay volume of FASER has a cylindrical shape with a depth of $D = 10\,\text{m}$ and a radius of $R = 1\,\text{m}$.

Long-lived particles that are produced in the very forward direction and decay in the detector typically have very large energies on the order of $\sim \text{TeV}$. The energetic decay products lead to a striking signature at FASER consisting of charged tracks with very high energy, originating from a vertex inside the detector and with a combined momentum pointing back to the IP. A detector that aims to make use of kinematic features to distinguish signal from background therefore needs to be able to measure the individual tracks with sufficient resolution and identify their charges. A tracking-based technology, supplemented with a magnet and calorimeter to allow for energy measurements, would make up the key components of the FASER detector.

The shielding provided by the rock that surrounds the detector’s location as well as the forward LHC infrastructure consisting of magnets and absorbers would eliminate most potential background processes. The only known particles that can transport TeV energies through $\sim 100\,\text{m}$ of rock between the IP and FASER are muons and neutrinos. A detailed analysis using FLUKA, taking into account the exact layout of the LHC tunnels and models of radiation-matter interactions, has shown that the dominant source of background is radiative processes associated with muons from the $pp$ collision debris [56]. Such backgrounds can be identified by the presence of a high-energy muon traversing the full detector and can be suppressed by using a scintillating charged particle veto layer at the front of the detector. Additional backgrounds from neutrino interactions with the detector are small and generally have different kinematics. The study has also shown that no high-energy particles are expected to enter FASER from infrastructure-induced backgrounds, such as proton showers in the dispersion suppressor or from beam-gas interactions. In the following, we assume that backgrounds can be reduced to negligible levels.

In order to reduce the trigger rate at low energies, FASER requires a large visible energy deposition from LLP decay products, i.e., $E_{\text{vis}} \gtrsim 100\,\text{GeV}$. For many models, this choice is dictated by the kinematics of the signal, while for iDM, lower energy thresholds are optimal for small DM mass-splittings. In Appendix A, we discuss the effect of modifying this threshold on the projected sensitivity.

In estimating the projected sensitivity of FASER, we assume a luminosity of $L = 3\,\text{ab}^{-1}$. We also require that the excited iDM state, $\chi_2$, decays within the detector volume and that its decay products are sufficiently energetic,

\[
\text{FASER: } \begin{cases} 
  r_{\chi_2} < 1\,\text{m} \\
  470\,\text{m} < z_{\chi_2} < 480\,\text{m} \\
  E_{\text{vis}} > 100\,\text{GeV} 
\end{cases}
\]

\[
(21)
\]

VII. PRODUCTION AT THE LHC

At the LHC, the iDM states, $\chi_1$ and $\chi_2$, are pair-produced from the prompt decays of dark photons. In the
top panel of Fig. 3, we present the $\chi_1\chi_2$ production cross section per $\epsilon^2$, $\sigma(pp \to \chi_1\chi_2)/\epsilon^2$, as a function of the dark photon mass, $m_{A'}$, for the dominant production channels: meson decays, dark Bremsstrahlung, and Drell-Yan. The solid lines correspond to the total production cross section within the forward hemisphere ($p_{T,A'} > 0$). The lighter dashed lines instead show the production cross section in the very forward region, requiring $\chi_2$ (from the decay $A' \to \chi_1\chi_2$) to be within 2 mrad of the beam collision axis, which corresponds to the angular acceptance of FASER. The systematic uncertainties of the various production rates are shown as shaded regions around the central prediction. The central and bottom panels show the corresponding relative uncertainties for total and forward production, respectively.

If the dark photon is lighter than a few hundred MeV, it can be produced from the on-shell decays of pseudoscalar mesons, such as neutral pions and $\eta$-mesons. Following Ref. [52], we generate the meson spectra using the Monte Carlo code EPOS-LHC [57], as implemented in the simulation package CRMC [58], and subsequently decay the mesons into dark photons, i.e., $\pi^0, \eta \to \gamma A'$. The systematic uncertainties are estimated by comparing to the generators QGSJET-II-04 [59] and SIBYLL 2.3 [60] and are at the 20% level for both the total and forward cross section.

Heavier dark photons above the meson mass threshold are mainly produced via dark Bremsstrahlung and Drell-Yan. We model forward dark Bremsstrahlung using the Fermi-Weizsäcker-Williams approximation [61], as outlined in Ref. [52]. In particular, we follow Refs. [62, 63] by using a proton form-factor which incorporates off-shell mixing with vector mesons, such as $\rho$ and $\omega$ mesons, leading to an enhanced production rate around $m_{A'} \approx 775$ MeV. The validity requirements of the Fermi-Weizsäcker-Williams approximation ($E_{A'}, E_{\text{beam}} - E_{A'} \gg m_p, m_{A'}, p_{T,A'}$) are naturally fulfilled for highly energetic and forward dark photons within the FASER acceptance. However, the approximate formalism is no longer valid for larger $p_{T,A'} (\gg A_{QCD})$ and we therefore do not include Bremsstrahlung production for CODEX-b or MATHUSLA. Theoretical uncertainties in the Bremsstrahlung calculation are estimated by varying the cut on the dark photon transverse momentum, $p_{T,A'} < 1$ GeV, by a factor of two. Note that roughly 10% of all dark photons from meson decays and Bremsstrahlung are produced in the very forward direction (corresponding to the geometric acceptance of FASER), as shown by the dashed lines in Fig. 3.

We follow a different approach to estimate the dark photon production rate at higher transverse momentum, which is relevant for MATHUSLA and CODEX-b. According to the model of vector meson dominance, the photon couples to hadronic states through mixing with intermediate vector mesons. Such interactions also induce an effective mixing between the dark photon and vector mesons, such as the $\rho$. As a result, any physical process producing vector mesons should also lead to the production of dark photons. We perform a rough estimate of the resulting dark photon spectrum by rescaling the inclusive $\rho$ meson spectrum provided by the Monte Carlo generator EPOS-LHC by the appropriate mixing factors. We discuss this in more detail in Appendix D.

At larger momentum transfers, Drell-Yan processes, such as $q\bar{q} \to A'Z \to \chi_1\chi_2$, also contribute to iDM production. The results were obtained using MadGraph 5 [64] and PYTHIA 8 [65] using the parton distribution function (PDF) NNPDF 3.1 NNLO [66] and the FeynRules model file of Ref. [20]. This model file includes the general $A' - \gamma, Z$ mixing angles and masses, as discussed in Appendix B, as well as interference effects between the $A'$ and $Z$ contributions. As seen in Fig. 3, $A' - Z$ mixing leads to an enhancement of the production cross section for $m_{A'} \approx m_Z$. Interference effects are typically small except around the $Z$-peak when the two resonances overlap, $m_{A'} \pm \Gamma_{A'} \approx m_Z \pm \Gamma_Z$. In this case, the cross section also depends on the value of $\alpha_D$ which determines the width of the dark photon and hence the size of the interference. The overall cross section in Fig. 3 approximately scales with the kinetic mixing parameter as $\sigma \propto \epsilon^2$. Small deviations from this simple scaling occur if $m_{A'} \approx m_Z$ and $\epsilon$ is sufficiently large such that $\Gamma(Z \to \chi_1\chi_2) \propto \epsilon^2 \alpha_D$ significantly alters the total $Z$ decay width. However, $\epsilon$ values of this size are already excluded by LEP [19, 20]. In modeling Drell-Yan production, the dominant systematic uncertainty arises from the choice of scale at which the PDFs are evaluated. We estimate the corresponding scale uncertainty by varying the renormalization and factorization scales ($\mu$) by a factor of two around their central values, $\mu^2 = m_{A'}^2$. As shown by the blue shaded region in the central and bottom panels of Fig. 3, this theoretical uncertainty is larger for small dark photon masses and becomes $O(1)$ for $m_{A'} \approx 2$ GeV.

While GeV-scale dark photons (and the $\chi_1\chi_2$ pairs from their decays) that are produced from exotic meson decays or Bremsstrahlung are typically very collimated around the beam collision axis, heavier states have a broader spectrum. This is illustrated in the left panel of Fig. 4 where we show the $\chi_2$ kinematic distribution from Drell-Yan production in the $p_{T,\chi_2} - \theta_{\chi_2}$ plane for $m_{A'} = 10$ GeV, $m_{A'}/m_1 = 3$, $\Delta = 0.1$, $\epsilon = 10^{-3}$, and $\alpha_D = 0.1$. This representative set of parameters can be probed by CODEX-b, FASER, and MATHUSLA, as shown by the circled gray star in the top-left panel of Fig. 5. The transverse momentum of $\chi_2$ originates mainly from the decay of the dark photon and therefore peaks around $p_{T,\chi_2} \approx (m_{A'}^2 - 4m_1^2)^{1/2}/2$. Larger masses lead to a more broad and energetic $p_T$ spectrum. The right-vertical axis of Fig. 4 indicates the boosted decay length of $\chi_2$.

The right panel of Fig. 4 shows similar $\chi_2$ kinematic distributions, but now only including events that enter the detector geometry and weighted by the probability to decay within FASER, MATHUSLA, and CODEX-b. Regions of phase space contribute in which $\chi_2$ is sufficiently boosted such that the decay length, $\beta \gamma c \ell_{\chi_2}$, is not much
smaller than the distance between the detector and IP. For the chosen benchmark parameters, FASER is sensitive to the entire available energy spectrum within the angular acceptance, while MATHUSLA and CODEX-b are only able to probe the high-energy tail.

VIII. RESULTS

We now discuss the projected sensitivities of the various searches described in Sec. VI. Our main results are shown in Figs. 5-7. In Figs. 5 and 6, we focus on the cosmologically motivated region of $m_{A'} \sim \text{few } \times m_1 \sim \mathcal{O}(10)$ GeV.

Fig. 5 illustrates existing constraints (shaded gray) and the projected sensitivity of proposed LHC searches (solid color) in the $\epsilon - m_1$ plane for $m_{A'}/m_1 = 3$, $\alpha_D = 0.1$, and various choices of the fractional DM mass-splitting ($\Delta = 0.1, 0.05, 0.03$, and $0.01$). Along the black contour of each panel, the abundance of $\chi_1$ agrees with the measured DM energy density. Also shown are the projected reach of the existing Belle II [33] and SeaQuest [29] experiments (dot-dashed). The sensitivity of a monophoton search at Belle II is estimated by rescaling the 20 fb$^{-1}$ background study up to 50 ab$^{-1}$ [30–33], and the SeaQuest projection assumes a Phase II luminosity of 10$^{20}$ protons on target.

For hidden sector masses slightly above the kinematic threshold of $B$-factors ($m_{A'} = 3 m_1 \gtrsim 10$ GeV), dedicated LLP searches at CODEX-b (orange), FASER (dark blue), and MATHUSLA (green) are projected to be sensitive to much of the cosmologically motivated parameter space. The right panel of Fig. 4 aids in understanding the qualitative features of Fig. 5 and the comparative reach of these experiments. The forward layout of FASER requires the production of highly energetic $\chi_1\chi_2$ pairs, enabling the ability to probe shorter proper $\chi_2$ lifetimes and hence larger couplings. In contrast, the geometries of CODEX-b and MATHUSLA favor off-axis events that are correspondingly softer. Due to the decreased boost of such events, CODEX-b and MATHUSLA have optimal sensitivity to smaller couplings and longer proper lifetimes. Note that the projected reach of CODEX-b is smaller but qualitatively similar to that of MATHUSLA, due to the former’s smaller angular coverage and luminosity.

Displaced muon searches at LHCb (yellow) and ATLAS/CMS (light blue) are sensitive to smaller displacements and therefore are expected to cover regions of parameter space with larger couplings, provided that the mass-splitting is above the muon threshold ($m_1 > 2m_\mu/\Delta$). As shown in Fig. 5, LHCb is more sensitive to smaller masses (compared to ATLAS/CMS), since in this case the $\chi_2$ particles are produced soft and forward and LHCb benefits from its low muon thresholds. In contrast, ATLAS and CMS have enhanced sensitivity to heavier scenarios ($m_{A'} \gtrsim m_2$), since $\chi_2$ is produced more isotropically and its decay products are more energetic. The proposed LHC timing search utilizing the monojet trigger (solid cyan) extends this reach towards smaller couplings due to the increased size of the accessible decay volume. This search is limited at large couplings by the small $\chi_2$ decay length. For smaller masses, $\chi_2$ is highly boosted, leading to more collinear decay products and time delays below the required threshold to suppress backgrounds, $\Delta t \lesssim \mathcal{O}(\text{ns})$. Also shown in Fig. 5 is the projected reach of a timing analysis using only a timing-
FIG. 5. Existing constraints (shaded gray) and projected sensitivities (color) to models of fermionic inelastic dark matter in the ε − m_1 plane for m_A/m_1 = 3, α_D = 0.1, and various choices of the χ_1 − χ_2 fractional mass-splitting (Δ = 0.1, 0.05, 0.03, and 0.01). Along the black contour, the abundance of χ_1 matches the observed dark matter energy density. The solid gray regions are excluded from LEP [19, 20] and BaBar [18]. The solid colored contours show the projected reach of various proposed searches for displaced vertices at the LHC, such as at ATLAS and CMS (light blue) [8], LHCb (yellow) [44–46], CODEX-b (orange) [51], FASER (dark blue) [52], and MATHUSLA (green) [67]. We also show the sensitivity of a precision timing search at CMS [43], utilizing a conventional monojet (solid cyan) or an optimistic timing-based (dashed cyan) trigger. See the corresponding subsections of Sec. VI for further details. Also shown in dot-dashed are the projected sensitivities of Belle II (red) [33] and SeaQuest (purple) [29]. In the top-left panel, the circled gray star corresponds to the choice of model parameters in Fig. 4.

Based trigger (dashed cyan), which has enhanced sensitivity compared to one utilizing the conventional monojet trigger.

Fig. 6 examines the restricted thermal DM parameter space in the α_D − m_1 plane for m_A/m_1 = 3 and various choices of the mass-splitting, Δ. For each panel of this figure, we instead fix ε to the specific value required for χ_1 to freeze out with an abundance that matches the observed DM energy density. Hence, every point in this parameter space is cosmologically viable. In this case, Eq. (10) implies that for a fixed value of the DM coannihilation cross section at freeze-out, larger values of α_D favor smaller ε, suppressing the A' production rate at accelerators. For m_1 ∼ 10(10) GeV, the remaining viable parameter space favors sub-10% mass-splittings (Δ < 0.1), and sizable α_D, i.e., values larger than or comparable to the strength of SM gauge couplings.

We have ignored the RG-evolution of α_D (e.g., from m_1 ∼ 10 GeV up to LHC energies), which is needed for a more careful comparison between cosmological and ac-
accelerator processes. We have bounded $\alpha_D$ from above in Fig. 6 by demanding that if $\alpha_D$ is defined at $\mu \sim m_1$, then it remains perturbative up to a scale of $\mu \sim 10$ TeV. This is shown for large $\alpha_D$ as the dashed gray region in Fig. 6.

As is evident in Figs. 5 and 6, the combined strength of various searches, such as FASER and LHCb, will test most remaining regions of cosmologically viable parameter space for mass-splittings greater than a few percent, while CODEX-b and MATHUSLA will probe thermal relics for even smaller mass-splittings, $\Delta \sim 1\%$. Due to their shorter effective baselines, proposed searches at LHCb and ATLAS/CMS provide the dominant reach for $\Delta \gtrsim 0.1$, while CODEX-b, FASER, and MATHUSLA are increasingly important for DM mass-splittings smaller than 10\%.

In Figs. 5 and 6, we have shown results for $\Delta = 0.1, 0.05, 0.03,$ and $0.01$. Since the lifetime of $\chi_2$ scales as $\tau_{\chi_2} \propto \epsilon^{-2} \Delta^{-3}$ (see Eq. (14)), the projected sensitivities of the various LLP searches shift towards larger values of $\epsilon$ for smaller mass-splittings. Throughout, we have focused on $\Delta \lesssim 0.1$, since cosmologically motivated values of $\epsilon$ are often in tension with existing constraints for $m_1 \gtrsim 10$ GeV and mass-splittings larger than $\sim 10\%$, as discussed in Sec. IV. Although DM freeze-out in the early universe is largely insensitive to the particular value of $\Delta$ for $\Delta \lesssim 0.1$, the majority of accelerator searches discussed in this work rely on detecting the visible de-
FIG. 7. As in the top-left panel of Fig. 5, existing constraints (shaded gray) and projected sensitivities (color) to models of fermionic inelastic dark matter in the $\epsilon - m_1$ plane for $m_{A'}/m_1 = 3$, $\Delta = 0.1$, and $\alpha_D = 0.1$, and now extended to smaller DM masses. Constraints from the SLAC E137 beam dump are shown as a gray dot-dashed contour since they suffer from uncertainties pertaining to the energy threshold of the analysis [29]. In addition to the various LHC searches discussed in Sec. VI, we also show the projected reach from low-energy accelerators as dot-dashed colored contours, such as Belle II (red) [33], SeaQuest (purple) [29], dedicated searches for displaced vertices at BaBar (magenta) [8], MiniBooNE (green) [24], JSNS (red) [68], BDX (brown) [24], and LDMX (pink) [10, 29, 69]. Also shown is the future sensitivity from electroweak precision tests (EWPT) at the LHC (dark teal) [20].

cay products of $\chi_2$, whose total energy is directly controlled by the size of $\Delta$. In particular, for $\Delta \lesssim 0.01$, the energy of these decay products often falls under the required experimental thresholds and/or $\chi_2$ is so long-lived that it rarely decays inside any of the instruments discussed. Hence, for sub-percent mass-splittings, these accelerator searches do not have significant reach. In this case, future direct detection experiments will have complementary sensitivity for hidden sector gauge couplings near the perturbative limit, i.e., $\alpha_D \gtrsim O(1)$. For $\Delta \gtrsim 10^{-6}$, inelastic scattering with SM particles is kinematically suppressed, and loop-induced elastic processes dominate. This point is discussed in more detail in Appendix C.

Although it is not the primary focus of this work, we note that similar signals of IDM have been considered for sub-GeV masses. Fig. 7 compares the LHC searches considered here (solid color) to the multitude of previous studies focusing on low-energy accelerators and fixed-target experiments (dot-dashed color) for the standard benchmark of $m_{A'}/m_1 = 3$, $\Delta = 0.1$, and $\alpha_D = 0.1$. The parameter space shown in Fig. 7 is similar to that in the top-left panel of Fig. 5, but is now extended down to DM masses of $m_1 = 10$ MeV. It is interesting to note that in addition to providing leading sensitivity to such models when $m_1 \sim O(10)$ GeV, proposed LHC experiments such as FASER and MATHUSLA would be competitive with other searches at low-energy accelerators for GeV-scale masses, such as SeaQuest and MiniBooNE.

In this section, we have fixed the dark photon-to-DM mass ratio to $m_{A'}/m_1 = 3$. For slightly larger values of $m_{A'}/m_1$, both the thermal relic line and the projected sensitivities of various LLP searches shift to larger values of $\epsilon$ compared to Fig. 5. We have not considered $m_{A'} \gg 3m_1$, since the amount of viable parameter space that is cosmologically motivated shrinks significantly (see Fig. 1). For $m_{A'} \lesssim m_1 + m_2$, the signals discussed above may still occur through the production of off-shell dark photons, although at a suppressed rate. In this case, future LHC searches for visibly decaying dark photons ($A' \rightarrow \ell^+\ell^-$) are an efficient means of detection [20]. Furthermore, for $m_{A'} \lesssim m_1$, there are no longer specific $e$-targets in parameter space that are cosmologically motivated, since new processes such as $\chi_1\chi_1 \rightarrow A'A'$ facilitate freeze-out in the early universe for much smaller values of $\epsilon$ [70]. While such scenarios are still interesting and worthy of study, a dedicated analysis is beyond the scope of this work.

IX. CONCLUSION

GeV-scale hidden sectors have been extensively searched for at low-energy beam dump, fixed-target, and
collider experiments. However, many of these searches lack sensitivity if light hidden sectors couple to the Standard Model through mediators heavier than $\sim 10$ GeV, as is often appreciated within the context of hidden valleys [28].

In this work, we have focused on models in which a GeV-scale pseudo-Dirac dark matter multiplet couples to the Standard Model through dark photons of mass $O(10)$ GeV. In this case, cosmological considerations motivate dark matter fractional mass-splittings smaller than $\sim 10\%$. Furthermore, for mass-splittings larger than $O(\text{MeV})$, the excited dark matter state is often unstable on collider timescales, and future searches for its displaced decay products at existing and proposed LHC experiments can leverage the large center of mass energy to test the majority of remaining motivated parameter space below $\sim 100$ GeV.

In this study, we have focused on dedicated searches at ATLAS, CMS, and LHCb, as well as the proposed CODEX-b, FASER, and MATHUSLA experiments. For dark photons heavier than $\sim 10$ GeV, such experiments constitute the best avenue towards detection. It is enticing to note that even for sub-GeV hidden sector masses, the projected reach of FASER and MATHUSLA are often competitive with higher-intensity and lower-energy beam dump and fixed-target experiments. This warrants more detailed studies of other qualitatively similar scenarios, such as models in which dark matter is the lightest state of a strongly interacting hidden sector [71, 72].

The visible energy deposition in models of inelastic dark matter depends crucially on the dark matter mass-splitting, and is typically suppressed for splittings below $O(100)$ MeV. Hence, unlike, e.g., minimal models of visibly decaying dark photons or dark scalars, enhancing the sensitivity of future LHC searches to models of inelastic dark matter requires optimizing the energy threshold of the analysis to guarantee sufficient signal efficiency while still maintaining adequate background discrimination.

ACKNOWLEDGEMENTS

We thank Nikita Blinov, Stefania Gori, Yoni Kahn, Simon Knapen, Gordan Krnjaic, Zhen Liu, Natalia Toro, and Mike Williams for useful discussions. We would also like to personally blame Gordan Krnjaic for his strong work ethic, which is largely responsible for the sheer multitude of contours in Fig. 7. AB is supported by the U.S. Department of Energy under Contract No. DE-AC02-76SF00515. FK is supported by the U.S. National Science Foundation under the grant PHY-1620638. AB is grateful to the hospitality of the Kavli Institute for Theoretical Physics in Santa Barbara, CA, supported in part by the National Science Foundation under Grant No. NSF PHY11-25915 where some of the research reported in this work was carried out.
Appendix A: Energy Thresholds

Fully exploring the cosmologically motivated iDM parameter space with the LHC depends crucially on the energy thresholds implemented in the various analyses of Sec. VI. In this appendix, we briefly discuss how slightly varying these energy/momentum thresholds alters the projected sensitivity of future searches.

In Fig. 8, we show the projections of ATLAS, CMS, CODEX-b, FASER, and MATHUSLA for different energy thresholds, compared to the default baseline cuts of Sec. VI. Darker/lighter shades correspond to larger/smaller minimum energy thresholds for the decay products of $\chi_2$. For CODEX-b and MATHUSLA, we demand a minimum energy of 1200, 600, or 300 MeV per track. For FASER, the total visible energy deposition is restricted to be greater than 200, 100, or 50 GeV. For a displaced muon-jet search at ATLAS/CMS we require a minimal muon $p_T$ of 10, 5, or 2.5 GeV. For the timing analysis at CMS, we implement cuts on the minimum transverse lepton momentum of 6, 3, or 1.5 GeV. As mentioned in Sec. VI, it is beyond the scope of this work to consider such variations for LHCb, although it could significantly enlarge the discovery potential.

Fig. 8 implies that the projected sensitivities are often significantly enhanced for slightly lower energy thresholds. Hence, for sub-10% DM mass-splittings, the ability to robustly explore regions of cosmology motivated parameter space for a given experimental search depends on the instrument’s capability to implement adequately low energy thresholds while still maintaining sufficient background rejection.

Appendix B: Kinetic Mixing

In this appendix, we briefly review the model of a kinetically-mixed dark photon, closely following the discussion in Ref. [20]. We imagine that at low-energies, the massive gauge boson of a broken $U(1)_B$ symmetry ($A'$) kinetically mixes with that of SM hypercharge ($B$). The relevant starting point is

$$\mathcal{L} \supset -\frac{1}{4} B_{\mu \nu}^2 - \frac{1}{4} A'^2_{\mu \nu} + \frac{1}{2} m_Z^2 A'_{\mu} A'_{\nu} + \frac{\epsilon}{2 \cos \theta_w} A'_{\mu \nu} B_{\mu \nu},$$

where $\theta_w$ is the Weinberg angle and the $Z$ boson is defined as in the SM. Above, the first two lines contain the relevant kinetic and mass terms for the SM hypercharge, $Z$, and $U(1)_B$ gauge bosons, while the third line introduces kinetic mixing between the two sectors, controlled by the dimensionless parameter, $\epsilon$. The $Z$ and $A'$ mass parameters, $m_Z$ and $m_{A'}$, are written in this way to indicate that these are the mass terms in the unmixed basis, and hence correspond to the physical masses only in the limit that $\epsilon \to 0$. We remain agnostic in regards to the source of $m_{A'}$, which may arise, for instance, from an extended $U(1)_B$ Higgs sector or the Stueckelberg mechanism.

Let us first focus on the kinetic terms of Eq. (B1), which can be diagonalized by performing the following field redefinition on the hypercharge sector,

$$B_{\mu} \rightarrow B_{\mu} + \frac{\epsilon}{\cos \theta_w} A'_{\mu}.$$

In particular, applying Eq. (B2) to the kinetic terms of Eq. (B1) gives

$$\mathcal{L}_{\text{kin}} = -\frac{1}{4} B_{\mu \nu}^2 - \frac{1}{4} \left( 1 - \frac{\epsilon^2}{\cos^2 \theta_w} \right) A'^2_{\mu \nu}.$$

In order to canonically normalize the kinetic term of $A'$, we perform another redefinition,

$$A'_{\mu} \rightarrow \frac{A'_{\mu}}{(1 - \frac{\epsilon^2}{\cos^2 \theta_w})^{1/2}}.$$

After this rescaling, we recover the canonical form, i.e.,

$$\mathcal{L}_{\text{kin}} = -\frac{1}{4} B_{\mu \nu}^2 - \frac{1}{4} A'^2_{\mu \nu}.$$

Now, we must examine the mass terms. To recap, Eqs. (B2) and (B4) can be written as a single transformation,

$$B_{\mu} \rightarrow B_{\mu} + \frac{\epsilon}{\cos \theta_w} \frac{1}{(1 - \frac{\epsilon^2}{\cos^2 \theta_w})^{1/2}} A'_{\mu}$$

$$A'_{\mu} \rightarrow \frac{A'_{\mu}}{(1 - \frac{\epsilon^2}{\cos^2 \theta_w})^{1/2}},$$

which allowed us to write the kinetic terms in a canonical form. The first line above corresponds to a shift of the SM $Z$ and photon fields,

$$Z_{\mu} \rightarrow Z_{\mu} - \frac{\epsilon \tan \theta_w}{(1 - \frac{\epsilon^2}{\cos^2 \theta_w})^{1/2}} A'_{\mu}$$

$$A_{\mu} \rightarrow A_{\mu} + \frac{\epsilon}{(1 - \frac{\epsilon^2}{\cos^2 \theta_w})^{1/2}} A'_{\mu}.$$
Note that the $A' - A'$ (bottom-right) entry in Eq. (B8) differs from that of Ref. [20]. To diagonalize this matrix, we need to perform one final field redefinition (which leaves the kinetic terms unchanged),

$$\begin{pmatrix} Z'_{\mu} \\ A'_\mu \end{pmatrix} \rightarrow \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} Z_{\mu} \\ A'_\mu \end{pmatrix} , \quad (B10)$$

where the angle $\alpha$ is defined as

$$\tan \alpha = \frac{1}{2} [1 - \eta^2 - \delta^2 - \text{sign}(1 - \delta^2) \sqrt{4 \eta^2 + (1 - \eta^2 - \delta^2)^2}] . \quad (B11)$$

The masses of the $Z$-like and $A'$-like states are then given by

$$m^2_{Z,A'} = \frac{\tilde{m}^2_Z}{2} \left(1 + \eta^2 + \delta^2 \pm \text{sign}(1 - \delta^2) \sqrt{(1 + \eta^2 + \delta^2)^2 - 4\delta^2} \right) , \quad (B12)$$

where we have adopted the convention of Ref. [20] ($\alpha \rightarrow 0$ as $\epsilon \rightarrow 0$ for $m_Z \neq m_{A'}$, regardless of the sign of $m_Z - m_{A'}$). It is useful to write down the single overall transformation that combines Eqs. (B6) and (B10),

$$\begin{pmatrix} A' \\ Z \\ A \end{pmatrix} \rightarrow C \times \begin{pmatrix} A' \\ Z \\ A \end{pmatrix} , \quad (B13)$$

where the matrix, $C$, is given by

$$C \equiv \begin{pmatrix} (\eta/\epsilon) \cos \alpha \cot \theta_w & (\eta/\epsilon) \sin \alpha \cot \theta_w & 0 \\ -\sin \alpha + \eta \cos \alpha & \cos \alpha - \eta \sin \alpha & 0 \\ \eta \cos \alpha \cot \theta_w & \eta \sin \alpha \cot \theta_w & 1 \end{pmatrix} . \quad (B14)$$

This single transformation diagonalizes the entire Lagrangian of Eq. (B1).

Let us now apply this transformation to the interaction terms. If the original interaction Lagrangian (in the $\epsilon \rightarrow 0$ limit) is parametrized as

$$\mathcal{L}_\text{int} = A'_\mu \bar{J}_D^\mu + \sum_f \left( Z_{\mu} \bar{f} \gamma^\mu (g_v + g_0 \gamma^5) f \right) + \epsilon Q_f A'_\mu \bar{f} \gamma^\mu f , \quad (B15)$$

then Eq. (B14) implies that, in the mass basis, we have,

$$\mathcal{L}_\text{int} \rightarrow (C_{A'Z} \bar{A}'_\mu + C_{A'Z} \bar{Z}_\mu) J_D^\mu$$

$$+ \sum_f \left( A'_\mu \bar{f} \gamma^\mu \left( [g_v C_{Z,A'} + \epsilon Q_f C_{A'A'}] + g_0 C_{Z,A'} \gamma^5 \right) f \right) + Z_\mu \bar{f} \gamma^\mu \left( [g_v C_{Z,Z} + \epsilon Q_f C_{A,Z}] + g_0 C_{Z,Z} \gamma^5 \right) f , \quad (B16)$$

where $J_D^\mu$ corresponds to the $U(1)_D$ current, the sum is performed over the SM fermions ($f$), and we have ignored couplings involving the massless photon (since it does not couple to $J_D^\mu$).

It is instructive to take various limits of Eq. (B14). If $\epsilon \ll 1$ and $m_{A'} \ll m_Z$, then Eq. (B14) becomes

$$C(\epsilon \ll 1, m_{A'} \ll m_Z) \simeq \begin{pmatrix} 1 - \epsilon \tan \theta_w & 0 \\ 0 & 1 \end{pmatrix} . \quad (B17)$$

In this case, Eq. (B16) is approximately

$$\mathcal{L}_\text{int}(\epsilon \ll 1, m_{A'} \ll m_Z) \simeq \left( A'_\mu - \epsilon \tan \theta_w Z_\mu \right) J_D^\mu$$

$$+ \sum_f \epsilon Q_f A'_\mu \bar{f} \gamma^\mu f + \cdots , \quad (B18)$$

which is the standard form of the $A' - \text{SM}$ interaction that is most often studied in the literature. Another interesting limit arises when $\epsilon \ll 1$ and $m_{A'} \simeq m_Z$. In this case, Eq. (B14) is given by

$$C(\epsilon \ll 1, m_{A'} \simeq m_Z) \simeq \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ \epsilon & \epsilon & \sqrt{2} \end{pmatrix} \quad (B19)$$

It might seem unsettling that the $A' - Z$ mixing does not vanish as $\epsilon \rightarrow 0$. However, this is simply a result of the convention for $\alpha$ in Eq. (B11) when $m_{A'} \simeq m_Z$. For two degenerate states of identical spin, one can always choose separate bases related by arbitrary orthogonal transformations. Regardless, physical processes must remain unchanged. We will provide a concrete example below.

If we instead choose to parametrize the original interaction Lagrangian of Eq. (B15) as

$$\mathcal{L}_\text{int} = A'_\mu \bar{J}_D^\mu + Z_\mu \bar{J}_Z^\mu + A'_\mu \bar{J}_{\text{em}}^\mu , \quad (B20)$$

where $J_{D,Z,\text{em}}$ are the currents corresponding to the associated gauge bosons, then applying the redefinition of Eq. (B19) gives

$$\mathcal{L}_\text{int} \rightarrow \frac{1}{\sqrt{2}} \left( A'_\mu (J_D^\mu - J_Z^\mu + \epsilon J_{\text{em}}^\mu) \right)$$

$$+ \frac{1}{\sqrt{2}} Z_\mu (J_D^Z + J_Z^Z + \epsilon J_{\text{em}}^Z) + A'_\mu J_{\text{em}}^\mu . \quad (B21)$$

At energies well below the $A'$ and $Z$ masses, we then obtain the corresponding four-Fermi theory coupling the $U(1)_D$ and SM sectors,

$$-\mathcal{L}_\text{int} \simeq \frac{1}{2} \left( \frac{1}{m_{A'}} - \frac{1}{m_Z^2} \right) J_D^\mu J_Z^\mu$$

$$+ \frac{\epsilon}{2} \left( \frac{1}{m_{A'}^2} + \frac{1}{m_Z^2} \right) J_D^\mu J_{\text{em}}^\mu$$

$$\simeq \frac{\epsilon}{m_Z^2} J_D^\mu J_{\text{em}}^\mu , \quad (B22)$$

which shows that the two sectors indeed decouple for $\epsilon \rightarrow 0$. Note also the lack of a resonant enhancement for $m_{A'} \simeq m_Z$. 
Appendix C: Direct Detection

In this appendix, we provide a calculation for DM elastic scattering in underground direct detection experiments. Let us first briefly summarize the basic formalism. For more comprehensive discussions and reviews, see, e.g., Refs. [73–75]. Let $\chi_1$ and $f$ denote a Majorana DM and SM fermion field, respectively. We define the SM currents,

$$O_f^{(0)} = m_f \bar{f} f$$
$$O_f^{(2)\mu\nu} = \frac{1}{2} \bar{f} \left[ i \gamma^\mu \gamma^\nu + i \gamma^\nu \gamma^\mu - \frac{1}{2} g^{\mu\nu} \gamma_5 \right] f . \quad \text{(C1)}$$

The full theory of Sec. III is matched to the effective low-energy Lagrangian

$$\mathcal{L}_\text{EFT} = c_f^{(0)} \bar{\chi}_1 \chi_1 O_f^{(0)} + \frac{c_f^{(2)}}{m_1^2} \bar{\chi}_1 i \gamma_5 \partial^\mu \chi_1 O_f^{(2)\mu\nu} , \quad \text{(C2)}$$

where $c_f^{(0)}$ and $c_f^{(2)}$ are the spin-0 and spin-2 Wilson coefficients, respectively, and an implicit summation over the SM fermions ($f$) is assumed.

In the case of DM-nucleon scattering ($f = q$, where $q$ is a SM quark), we define the SM form-factors,

$$\langle N | O_q^{(0)} | N \rangle = m_N f_q^{(0)} ,$$
$$\langle N | O_q^{(2)\mu\nu} | N \rangle = \frac{1}{m_N} \left( k^\mu k^\nu - \frac{1}{4} m_N^2 g^{\mu\nu} \right) f_q^{(2)} , \quad \text{(C3)}$$

where $N = p, n$ denotes a proton or neutron, respectively, and the form-factors, $f_q^{(0)}$, are given by the lattice and nucleon PDFs, respectively [73–75]. This can be adapted in the case of electron-scattering ($f = e$), in which case we have

$$\langle e | O_e^{(0)} | e \rangle = m_e ,$$
$$\langle e | O_e^{(2)\mu\nu} | e \rangle = \frac{1}{m_e} \left( k^\mu k^\nu - \frac{1}{4} m_e^2 g^{\mu\nu} \right) . \quad \text{(C4)}$$

These matrix elements, together with the Wilson coefficients of Eq. (C2), define the spin-independent amplitudes for nucleon and electron scattering

$$\mathcal{M}_N = m_N \left( \sum_{q = u, d, s} f_q^{(0)} c_q^{(0)} + \frac{3}{4} \sum_{q = N} f_q^{(2)} c_q^{(2)} \right) ,$$
$$\mathcal{M}_e = m_e \left( c_e^{(0)} + \frac{3}{4} c_e^{(2)} \right) . \quad \text{(C5)}$$

We then have for the spin-independent scattering cross section,

$$\sigma_{\text{SI}} = \frac{4}{\pi} \mu_{1i}^2 |\mathcal{M}_i|^2 , \quad \text{(C6)}$$

where $i = N, e$ and $\mu_{1i}$ is the $\chi_1 - i$ reduced mass.

For the model of fermionic iDM in Sec. III, inelastic scattering ($\chi_1 f \rightarrow \chi_2 f$) is kinematically suppressed for $\Delta \gtrsim 10^{-6}$. In this case, the leading order contribution to elastic scattering is given by one-loop diagrams, as shown in Fig. 9. Similar diagrams exist for $Z$-exchange and, and $\chi_1$ also couples to the gluon field-strength at the two-loop level. A detailed calculation is beyond the scope of this work, but we do not expect these contributions to significantly alter our final results.

In the limit that $m_f \ll m_1 < m_{A'}$, we calculate the Wilson coefficients from the diagram of Fig. 9 to be

$$c_f^{(0)} \approx - \frac{3 Q_f^2 \alpha_\text{em} \alpha_D e^2}{2 m_{A'}} m_1 \left( 1 - 4 \log \frac{m_{A'}}{m_1} \right) ,$$
$$c_f^{(2)} \approx - \frac{4 Q_f^2 \alpha_\text{em} \alpha_D e^2}{9 m_{A'}} m_1 \left( 1 - 12 \log \frac{m_{A'}}{m_1} \right) . \quad \text{(C7)}$$

In the case of $\chi_1$-electron scattering, this gives a simple analytic result for the cross section,

$$\sigma_{\text{SI}}^{(\text{electron})} \approx \frac{\alpha_\text{em}^2 \alpha_D^2 e^4 m_e^2}{9\pi m_{A'}^2} \left( 11 - 60 \log \frac{m_{A'}}{m_1} \right)^2 . \quad \text{(C8)}$$

Although a detailed calculation for the DM-nucleon cross section should include the appropriate form-factors, as in the top line of Eq. (C5), the form of Eq. (C8) is still useful as a parametric estimate,

$$\sigma_{\text{SI}}^{(\text{nucleon})} \approx \frac{\alpha_\text{em}^2 \alpha_D^2 e^4}{9\pi} \frac{m_1^2}{m_{A'}^3} . \quad \text{(C9)}$$

We demand that the DM-nucleon elastic scattering cross section is larger than the coherent neutrino background, i.e., $\sigma_{\text{SI}}^{(\text{nucleon})} \gtrsim 10^{-9}$ cm$^2$ $(m_1/10 \text{ GeV})$ [76]. Eq. (C9) implies that this is achieved when

$$\alpha_D e^2 \gtrsim 10^{-4} \times \left( \frac{m_1}{10 \text{ GeV}} \right)^{7/2} \left( \frac{m_{A'}}{3 m_1} \right)^4 . \quad \text{(C10)}$$

Hence, future experiments such as LUX [77] will only be competitive with existing constraints from LEP [19, 20] for $\alpha_D \gtrsim O(1) \times (m_{A'}/3 m_1)^4$. 

![Fig. 9. A representative Feynman diagram responsible for the one-loop coupling of DM ($\chi_1$) to SM fermions ($f$) (crossed diagrams not shown). There are also one-loop $Z$-exchange diagrams and two-loop diagrams couplings $\chi_1$ to the gluon field-strength. We do not include these other contributions in our estimate for the DM elastic scattering cross section.](image-url)
Appendix D: Dark Photon Mixing with Vector Mesons

According to the model of vector meson dominance (VMD), the photon couples to hadronic states through mixing with intermediate QCD vector mesons. Following the convention in which the SM photon ($A^\mu$) and vector mesons ($V^\mu$) mass-mix, the corresponding effective Lagrangian is given by

$$\mathcal{L} \subset (\epsilon m_V^2/g_V) A^\mu V_\mu, \quad \text{(D1)}$$

where $g_V$ is the vector meson-pion interaction strength. If a light dark photon ($m_{A'} \ll m_V$) kinetically mixes with SM hypercharge, the photon field is redefined as $A_\mu \rightarrow A_\mu + \epsilon A'_\mu$ (see Appendix B). This leads to an effective mass-mixing between the dark photon and vector mesons that is analogous to Eq. (D1),

$$\mathcal{L} \subset (\epsilon m_V^2/g_V) A'^\mu V_\mu \quad \text{(D2)}$$

The amplitudes for dark photon and vector meson production are related by the $A' - V$ mixing of Eq. (D2). Let us denote the amplitude to produce a vector meson as $M_V$ and the amplitude to produce a dark photon as $M_{A'}$. We can then write

$$M_{A'} \simeq \theta_V M_V, \quad \text{(D3)}$$

where the mixing parameter, $\theta_V$, is defined as

$$\theta_V \equiv \frac{\epsilon e}{g_V} \frac{m_V^2}{q^2 - m_V^2 + i m_V \Gamma_V}, \quad \text{(D4)}$$

$\Gamma_V$ is the vector meson width, and $q^2 = m_{A'}^2$ corresponds to on-shell $A'$ production.

There are multiple vector mesons that mix with $A'$ and contribute to its production, such as the $\rho$, $\omega$, and heavier hadronic resonances. In this study, we only consider the leading contribution from $A' - \rho$ mixing and neglect possible contributions from other vector mesons. This is motivated by the fact that the production rate of heavier excited states and $A' - \omega$ mixing are suppressed in comparison, e.g., $\theta_\omega/\theta_\rho \sim g_\rho/g_\omega \simeq 0.3$ where $g_\rho \simeq 5$ and $g_\omega \simeq 17$ [80].

Eq. (D3) allows us to obtain an approximate $A'$ spectrum by reweighting that of the $\rho$ meson, i.e., $d\sigma_{A'} = \theta_\rho^2 \frac{d\sigma_{\rho}}{d q^2}$. In this work, we use the inclusive $\rho$ meson spectrum provided by the Monte Carlo generator EPOS-LHC [57]. We have also checked that this procedure qualitatively matches forward Bremsstrahlung production of dark photons, as discussed in Sec. VII.

While the above method can be generalized to include additional mesons, this would require knowledge of the specific vector meson production processes in order to incorporate interference effects, which are known to be important in modeling the suppressed $A'$ production rate for $m_{A'} \gtrsim 100$ GeV [61, 62, 81]. In general, this is only available in special cases such as Bremsstrahlung, in which the dark photon is assumed to mix with vector mesons that are radiated off of the incoming proton beam.

[1] J. Alexander et al., “Dark Sectors 2016 Workshop: Community Report,” arXiv:1608.08632 [hep-ph].
[2] M. Battaglieri et al., “US Cosmic Visions: New Ideas in Dark Matter 2017: Community Report,” arXiv:1707.04591 [hep-ph].
[3] B. W. Lee and S. Weinberg, “Cosmological Lower Bound on Heavy Neutrino Masses,” Phys. Rev. Lett. 39 (1977) 165–168, [183(1977)].
[4] M. Escudero, A. Berlin, D. Hooper, and M.-X. Lin, “Toward (Finally!) Ruling Out Z and Higgs Mediated Dark Matter Models,” JCAP 1612 (2016) 029, arXiv:1609.09079 [hep-ph].
[5] J. Kearney, N. Orlofsky, and A. Pierce, “Z boson mediated dark matter beyond the effective theory,” Phys. Rev. D95 no. 3, (2017) 035020, arXiv:1611.05048 [hep-ph].
[6] Planck Collaboration, P. A. R. Ade et al., “Planck 2015 results. XIII. Cosmological parameters,” Astron. Astrophys. 594 (2016) A13, arXiv:1502.01589 [astro-ph.CO].
[7] D. Tucker-Smith and N. Weiner, “Inelastic dark matter,” Phys. Rev. D64 (2001) 043502, arXiv:hep-ph/0101138 [hep-ph].
[8] E. Izaguirre, G. Krnjaic, and B. Shuve, “Discovering Inelastic Thermal-Relic Dark Matter at Colliders,” Phys. Rev. D93 no. 6, (2016) 063523, arXiv:1508.03050 [hep-ph].
[9] C. Boehm and P. Fayet, “Scalar dark matter candidates,” Nucl. Phys. B683 (2003) 219–263, arXiv:hep-ph/0305261 [hep-ph].
[10] A. Berlin, N. Blinov, G. Krnjaic, P. Schuster, and N. Toro, “Dark Matter, Millicharges, Axion and Scalar Particles, Gauge Bosons, and Other New Physics with LDMX,” arXiv:1807.01730 [hep-ph].
[11] Particle Data Group Collaboration, C. Patrignani et al., “Review of Particle Physics,” Chin. Phys. C40 no. 10, (2016) 100001.
[12] XENON Collaboration, E. Aprile et al., “Dark Matter Search Results from a One Tonne×Year Exposure of XENON1T,” arXiv:1805.12562 [astro-ph.CO].
[13] B. Holdom, “Two U(1)’s and Epsilon Charge Shifts,” Phys. Lett. B166B (1986) 196–198.
[14] F. del Aguila, G. D. Coughlan, and M. Quiros, “Gauge Coupling Renormalization With Several U(1) Factors,” Nucl. Phys. B307 (1988) 633. [Erratum: Nucl.
Phys. B312, 751 (1989)].

[15] R. K. Leane, T. R. Slatyer, J. F. Beacon, and K. C. Y. Ng, “GeV-Scale Thermal WIMPs: Not Even Slightly Dead,” arXiv:1805.10305 [hep-ph].

[16] T. R. Slatyer, “Indirect dark matter signatures in the cosmic dark ages. I. Generalizing the bound on s-wave dark matter annihilation from Planck results,” Phys. Rev. D95 no. 2, (2016) 023527, arXiv:1506.03811 [hep-ph].

[17] BaBar Collaboration, B. Aubert et al., “Search for Invisible Decays of a Light Scalar in Radiative Transitions \( \gamma \gamma \to \gamma A_0 \),” arXiv:0808.0017 [hep-ex].

[18] BaBar Collaboration, J. P. Lees et al., “Search for Invisible Decays of a Dark Photon Produced in e+e- Collisions at BaBar,” Phys. Rev. Lett. 119 no. 13, (2017) 131804, arXiv:1702.03327 [hep-ex].

[19] A. Hook, E. Izaguirre, and J. G. Wacker, “Model Independent Bounds on Kinetic Mixing,” Adv. High Energy Phys. 2011 (2011) 859762, arXiv:1006.0973 [hep-ph].

[20] D. Curtin, R. Essig, S. Gori, and J. Shelton, “Illuminating Dark Photons with High-Energy Colliders,” JHEP 02 (2015) 157, arXiv:1412.0018 [hep-ph].

[21] K. Griest and D. Seckel, “Three exceptions in the calculation of relic abundances,” Phys. Rev. D43 (1991) 3191–3203.

[22] P. Gondolo and G. Gelmini, “Cosmic abundances of stable particles: Improved analysis,” Nucl. Phys. B360 (1991) 145–179.

[23] J. Edsjo and P. Gondolo, “Neutralino relic density including coannihilations,” Phys. Rev. D56 (1997) 1879–1894, arXiv:hep-ph/9704361 [hep-ph].

[24] E. Izaguirre, Y. Kahn, G. Krnjaic, and M. Moschella, “Testing Light Dark Matter Coannihilation With Fixed-Target Experiments,” Phys. Rev. D96 no. 5, (2017) 055007, arXiv:1703.06881 [hep-ph].

[25] “FeynRules 2.0 - A complete toolbox for tree-level phenomenology,” Comput. Phys. Commun. 185 (2014) 2250–2300, arXiv:1310.1921 [hep-ph].

[26] G. Belanger, F. Boudjema, A. Pukhov, and A. Semenov, “micrOMEGAs 3: A program for calculating dark matter observables,” Comput. Phys. Commun. 185 (2014) 960–985, arXiv:1305.0237 [hep-ph].

[27] J. L. Feng and J. Smolinsky, “Impact of a resonance on thermal targets for invisible dark photon searches,” Phys. Rev. D96 no. 9, (2017) 095022, arXiv:1707.03835 [hep-ph].

[28] M. J. Strassler and K. M. Zurek, “Echoes of a hidden valley at hadron colliders,” Phys. Lett. B651 (2007) 374–379, arXiv:hep-ph/0604261 [hep-ph].

[29] A. Berlin, S. Gori, P. Schuster, and N. Toro, “Dark Sectors at the Fermilab SeaQuest Experiment,” arXiv:1804.00661 [hep-ph].

[30] G. Inguglia, “Belle II studies of missing energy decays and searches for dark photon production,” PoS DIS2016 (2016) 263, arXiv:1607.02089 [hep-ex].

[31] Belle II Collaboration, T. Ferber, “Dark Sector Physics with Belle II,” PoS EPS-HEP2017 (2017) 061.

[32] G. De Pietro, “First data at Belle II and Dark Sector Physics,” PoS BEAUTY2018 (2018) 034, arXiv:1808.00776 [hep-ex].

[33] E. Kou et al., “The Belle II Physics book,” arXiv:1808.10567 [hep-ex].

[34] LSND Collaboration, L. B. Auerbach et al., “Measurement of electron - neutrino - electron elastic scattering,” Phys. Rev. D63 (2001) 112001, arXiv:hep-ex/0101039 [hep-ex].

[35] P. deNiverville, M. Pospelov, and A. Ritz, “Observing a light dark matter beam with neutrino experiments,” Phys. Rev. D84 (2011) 075020, arXiv:1107.4580 [hep-ph].

[36] J. D. Bjorken, S. Ecklund, W. R. Nelson, A. Abashian, C. Church, B. Lu, L. W. Mo, T. A. Nunamaker, and P. Rassmann, “Search for Neutral Metastable Penetrating Particles Produced in the SLAC Beam Dump,” Phys. Rev. D38 (1988) 3375.

[37] B. Batell, R. Essig, and Z. Surujon, “Strong Constraints on Sub-GeV Dark Sectors from SLAC Beam Dump E137,” Phys. Rev. Lett. 113 no. 17, (2014) 171802, arXiv:1406.2698 [hep-ph].

[38] MiniBooNE Collaboration, A. A. Aguilar-Arevalo et al., “Dark Matter Search in a Proton Beam Dump with MiniBooNE,” Phys. Rev. Lett. 118 no. 22, (2017) 221803, arXiv:1702.02688 [hep-ex].

[39] MiniBooNE DM Collaboration, A. Aguilar-Arevalo et al., “Dark Matter Search in Nucleon, Pion, and Electron Channels from a Proton Beam Dump with MiniBooNE,” arXiv:1807.06137 [hep-ex].

[40] CMS Collaboration, V. Khachatryan et al., “Search for dark matter, extra dimensions, and unparticles in monojet events in proton-proton collisions at \( \sqrt{s} = 8 \) TeV,” Eur. Phys. J. C75 no. 5, (2015) 235, arXiv:1408.3583 [hep-ex].

[41] CMS Collaboration, A. M. Sirunyan et al., “Search for disappearing tracks as a signature of new long-lived particles in proton-proton collisions at \( \sqrt{s} = 13 \) TeV,” arXiv:1804.07321 [hep-ex].

[42] CMS Collaboration, “Muon ID performance: low-\( p_T \) muon efficiencies.” CMS-DP-2014-020, Jun, 2014.

[43] J. Liu, Z. Liu, and L.-T. Wang, “Long-lived particles at the LHC: catching them in time,” arXiv:1805.05957 [hep-ph].

[44] LHCb Collaboration, R. Aaij et al., “Search for Dark Photons Produced in 13 TeV pp Collisions,” Phys. Rev. Lett. 120 no. 6, (2018) 061801, arXiv:1710.02867 [hep-ex].

[45] P. Itten, Y. Soreq, J. Thaler, M. Williams, and W. Xue, “Proposed Inclusive Dark Photon Search at LHCb,” Phys. Rev. Lett. 116 no. 25, (2016) 251803, arXiv:1603.08926 [hep-ph].

[46] A. Pierce, B. Shakyia, Y. Tsai, and Y. Zhao, “Searching for confining hidden valleys at LHCb, ATLAS, and CMS,” Phys. Rev. D97 no. 9, (2018) 095033, arXiv:1708.05389 [hep-ph].

[47] M. Buschmann, J. Kopp, J. Liu, and P. A. N. Machado, “Lepton Jets from Radiating Dark Matter,” JHEP 07 (2015) 045, arXiv:1506.07459 [hep-ph].

[48] A. Ibarra, E. Molinaro, and S. Vogl, “Potential for probing three-body decays of Long-Lived Particles with MATHUSLA,” arXiv:1809.01683 [hep-ph].

[49] D. Curtin et al., “Long-Lived Particles at the Energy Frontier: The MATHUSLA Physics Case,” arXiv:1806.07396 [hep-ph].

[50] C. Alpigiani et al., “A Letter of Intent for MATHUSLA: a dedicated displaced vertex detector above ATLAS or CMS.”

[51] V. V. Gligorov, S. Knapen, M. Papucci, and D. J.
