

**Direct detection portals for self-interacting dark matter**Manoj Kaplinghat,<sup>1</sup> Sean Tulin,<sup>2</sup> and Hai-Bo Yu<sup>3</sup><sup>1</sup>*Department of Physics and Astronomy, University of California, Irvine, California 92697, USA*<sup>2</sup>*Michigan Center for Theoretical Physics, University of Michigan, Ann Arbor, Michigan 48109, USA*<sup>3</sup>*Department of Physics and Astronomy, University of California, Riverside, California 92507, USA*

(Received 30 November 2013; published 19 February 2014)

Dark matter self-interactions can affect the small scale structure of the Universe, reducing the central densities of dwarfs and low surface brightness galaxies in accord with observations. From a particle physics point of view, this points toward the existence of a 1–100 MeV particle in the dark sector that mediates self-interactions. Since mediator particles will generically couple to the Standard Model, direct detection experiments provide sensitive probes of self-interacting dark matter. We consider three minimal mechanisms for coupling the dark and visible sectors: photon kinetic mixing, Z boson mass mixing, and the Higgs portal. Self-interacting dark matter motivates a new benchmark paradigm for direct detection via momentum-dependent interactions, and ton-scale experiments will cover astrophysically motivated parameter regimes that are unconstrained by current limits. Direct detection is a complementary avenue to constrain velocity-dependent self-interactions that evade astrophysical bounds from larger scales, such as those from the bullet cluster.

DOI: [10.1103/PhysRevD.89.035009](https://doi.org/10.1103/PhysRevD.89.035009)

PACS numbers: 95.35.+d, 12.60.-i

**I. INTRODUCTION**

Dark and visible matter have very different distributions in the Universe: dark matter (DM) forms diffuse halos (e.g., observed via gravitational lensing maps), while visible matter undergoes dissipative dynamics and tends to clump into galaxies and stars. However, this does not preclude the possibility of new dark sector interactions beyond the usual collisionless DM paradigm. DM could have a large cross section for scattering with other DM particles, and this scenario, dubbed self-interacting DM (SIDM) [1,2], can affect the internal structure (mass profile and shape) of DM halos compared to collisionless DM. In turn, astrophysical observations of structure, compared to numerical N-body simulations, can probe the self-interacting nature of DM.<sup>1</sup> It is worth emphasizing that tests of self-interactions can shed light on the nature of DM *even if DM is completely decoupled with respect to traditional DM searches*.

In fact, there are long-standing issues on small scales that may point toward SIDM. Dwarf galaxies are natural DM laboratories since in these galaxies DM tends to dominate baryons inside the optical radius. Observations indicate that the central regions of well-resolved dwarf galaxies exhibit cored profiles [4,5], as opposed to steeper cusp profiles found in collisionless DM-only simulations [6]. Cored profiles have been inferred in a variety of dwarf halos, including within the Milky Way (MW) [7], other nearby dwarfs [8], and low surface brightness galaxies [9]. An additional problem concerns the number of massive dwarf

spheroidals in the MW. Collisionless DM simulations have a population of subhalos in MW-like halos that are too massive to host any of the known dwarf spheroidals but whose star formation should not have been suppressed by ultraviolet feedback [10]. While these apparent anomalies are not yet conclusive—e.g., baryonic feedback effects may be important [11,12]—recent state-of-the-art SIDM N-body simulations have shown that self-interactions can modify the properties of dwarf halos to be in accord with observations, without spoiling the success of collisionless DM on larger scales and being consistent with halo shape and bullet cluster bounds [13–16].

The figure of merit for DM self-interactions is cross section per unit DM mass,  $\sigma/m_X$ , where  $X$  is the DM particle. To have an observable effect on DM halos over cosmological time scales, the cross section per unit mass must be of order

$$\sigma/m_X \sim 1 \text{ cm}^2/\text{g} \approx 2 \times 10^{-24} \text{ cm}^2/\text{GeV} \quad (1)$$

or larger. From a particle physics perspective, this value is many orders of magnitude larger than expected from weak-scale physics. For a typical weakly interacting massive particle (WIMP), the cross section is  $\sigma \sim 10^{-36} \text{ cm}^2$  and the mass is  $m_X \sim 100 \text{ GeV}$ , giving  $\sigma/m_X \sim 10^{-38} \text{ cm}^2/\text{GeV}$ . Evidence for self-interactions would therefore point toward a new dark mediator particle  $\phi$  that is much lighter than the weak scale, typically  $m_\phi \sim 1\text{--}100 \text{ MeV}$  [17]. Light mediators have been widely studied within many different contexts, e.g., in hidden sector models [18], in light DM models [19–21], in Sommerfeld-enhanced models for indirect detection signals [22,23], and in connection with

<sup>1</sup>Here we assume a single DM component with *nondissipative* self-interactions. It is possible for a subdominant fraction of DM to interact via dissipative processes [3].

supersymmetry and the WIMP miracle [24–28]. The light force carrier can be probed by low energy experiments with high luminosities [29–38]. The effect of self-interactions can also be searched through various other astrophysical observations, e.g., [3,39–52]. In this paper, we focus on the case in which the SIDM candidate is a fundamental particle and the coupling constant is small. It is possible that SIDM is made up of composite states like baryons or pions in the dark sector. In this case, the strength of self-interactions is controlled by the confinement scale. To generate  $\sigma/m_X \sim 1 \text{ cm}^2/\text{g}$ , the confinement scale should be  $\sim 1\text{--}100 \text{ MeV}$  and the strong-force mediator must be lighter than the confinement scale.

Astrophysical tests of self-interactions can probe DM that is decoupled from the Standard Model (SM). However, the dark and visible sectors are likely not to be *completely* decoupled.  $\phi$  particles in the early Universe must decay; otherwise they dominate DM. Unless additional states are introduced,  $\phi$  must decay to SM particles. These decay products—typically electrons, positrons, photons, and neutrinos—must not spoil the light element abundances predicted by big bang nucleosynthesis (BBN) and/or produce too much entropy. This requires a  $\phi$  lifetime less than  $\sim 1 \text{ s}$ , although this can be relaxed slightly as we discuss later.

In this paper, we study the reach for DM direct detection experiments to probe SIDM. We consider different scenarios for how  $\phi$  might couple to SM fermions by mixing with the photon,  $Z$ , or Higgs bosons, depending on the spin of  $\phi$ . Since BBN constraints give a *lower* bound on  $\phi$  couplings to the SM, this defines a *minimum* value for the direct detection cross section. As we show, the predicted SIDM direct detection range lies within the sensitivity reach of upcoming experiments such as XENON1T [53], LUX [54], and SuperCDMS [55], although the quantitative details are model dependent. The essential physics is summarized in Fig. 1. Both self-interactions and direct detection arise through the same light mediator  $\phi$ , which can also provide an annihilation channel for setting the DM relic density. If the mediator only couples to lighter states in the dark sector, we would not expect direct detection signals generically. However, we may still be able to explore the SIDM sector through other observations. For example, the coupling between SIDM and the lighter states may delay the kinetic

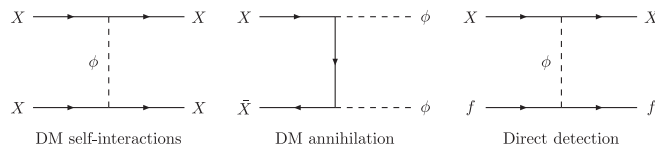


FIG. 1. Feynman diagrams arising from DM particle  $X$  coupled to a dark force mediator  $\phi$ . Self-interactions modify DM halos, while annihilation can give the observed DM relic density. Direct detection experiments are highly sensitive to potential couplings of  $\phi$  to SM fermions  $f$ , which allow decays  $\phi \rightarrow f\bar{f}$  before BBN.

decoupling of DM from the dark radiation, which may lead to signatures in the cosmic microwave background (CMB) observations [40].

Direct detection searches for SIDM are highly complementary to other astrophysical probes. For nuclear recoils, direct detection experiments are sensitive to DM masses  $m_X \gtrsim 10 \text{ GeV}$  due to threshold limitations. In this mass range, the self-interaction cross section tends to be more velocity dependent, becoming suppressed at higher velocity (like Rutherford scattering) [17]. In this case, DM can become effectively collisionless within larger DM halos, which have larger characteristic velocities. Tests for self-interactions within large DM halos, such as in merging clusters [56,57] or halo shape observables [58], are most sensitive to DM masses  $m_X \lesssim 10 \text{ GeV}$  where the self-interaction cross section tends to be more velocity independent. We also note that direct detection experiments can also probe low mass SIDM as well, via electron recoils [59,60], although the effective reach in couplings is much reduced compared to nuclear recoils.

In the remainder of this work, we first present simple particle physics models for SIDM in Sec. II. Then we discuss different portals for how the dark and visible sectors may be coupled in Sec. IV. We present our results in Sec. V, showing how direct detection and other constraints map onto SIDM parameter space. Last, we conclude in Sec. VI.

## II. PARTICLE PHYSICS OF SELF-INTERACTING DARK MATTER

As a minimal model for SIDM, we consider DM  $X$  to be a Dirac fermion which is coupled to a mediator particle  $\phi$ .  $X$  is assumed to be a SM gauge singlet with no direct coupling to SM particles. We take  $\phi$  to be either a real vector or a scalar, with an interaction given by

$$\mathcal{L}_{\text{int}} = \begin{cases} g_X \bar{X} \gamma^\mu X \phi_\mu & \text{vector mediator} \\ g_X \bar{X} X \phi & \text{scalar mediator} \end{cases}, \quad (2)$$

where  $g_X$  is a coupling constant. That is, the mediator  $\phi$  is the dark sector counterpart of force particles in the SM, analogous to a  $Z$  or Higgs boson.

Nonrelativistic DM self-scattering, shown in Fig. 1, can be described by a Yukawa potential [58,61,62]

$$V(r) = \pm \frac{\alpha_X}{r} e^{-m_\phi r}, \quad (3)$$

with a  $+$  ( $-$ ) sign for a repulsive (attractive) interaction. For the scalar  $\phi$  case, DM scattering is purely attractive. For the vector  $\phi$  case,  $XX$  or  $\bar{X}\bar{X}$  scattering is repulsive, while  $X\bar{X}$  is attractive. Other types of interactions give rise to more complicated potentials [63,64], which we do not consider. The differential scattering cross section  $d\sigma/d\Omega$  can be computed numerically through a standard partial wave analysis by solving the Schrödinger equation with  $V(r)$  for the DM two-body wave function; see Ref. [17] for details.

This method takes into account nonperturbative effects via multiple scatterings, analogous to the Sommerfeld enhancement for annihilation, which become important when the Born approximation breaks down for  $\alpha_X m_X / m_\phi \gtrsim 1$ . In the nonperturbative regime, DM self-scattering has a rich structure, including possible quantum mechanical resonances through bound state formation [65].

To compare with results from numerical N-body simulations of SIDM, we consider  $\langle \sigma_T \rangle$  as a suitable proxy for the self-interaction cross section  $\sigma$  entering into halo observables. Here, the transfer cross section is defined by  $\sigma_T \equiv \int d\Omega (1 - \cos \theta) d\sigma/d\Omega$ , where  $(1 - \cos \theta)$  is the fractional longitudinal momentum transfer for scattering angle  $\theta$ . Moreover, we average  $\sigma_T$  over the initial DM velocities  $\vec{v}_{1,2}$  by

$$\begin{aligned} \langle \sigma_T \rangle &= \int \frac{d^3 v_1 d^3 v_2}{(\pi v_0^2)^3} e^{-v_1^2/v_0^2} e^{-v_2^2/v_0^2} \sigma_T(|\vec{v}_1 - \vec{v}_2|) \\ &= \int \frac{d^3 v}{(2\pi v_0^2)^{3/2}} e^{-\frac{1}{2}v^2/v_0^2} \sigma_T(v), \end{aligned} \quad (4)$$

where  $v_0$  is the most probable velocity for a given halo and  $v = |\vec{v}_1 - \vec{v}_2|$  is the relative velocity. We choose  $v_0$  to be characteristic of different size halos, with larger halos having bigger  $v_0$ . The point is that  $d\sigma/d\Omega$  has a complicated dependence on both  $v$  and  $\theta$  in general, and by considering  $\langle \sigma_T \rangle$ , we are averaging over these dependencies in a physically meaningful way to obtain a single effective self-interaction cross section  $\sigma$  for a given halo.<sup>2</sup>

Aside from self-interactions, the mediator  $\phi$  can play an important role in setting the DM relic density in the early Universe via  $X\bar{X} \rightarrow \phi\phi$ , shown in Fig. 1. The annihilation cross section is (at the Born level)

$$(\sigma v)_{\text{ann}} \approx \frac{\pi \alpha_X^2}{m_X^2} \times \begin{cases} 1 & \text{vector mediator} \\ 3v^2/8 & \text{scalar mediator} \end{cases} \quad (5)$$

at lowest order in  $v$  and  $m_\phi$ . Annihilation is  $s$  wave for a vector mediator and  $p$  wave for a scalar mediator. For symmetric DM, where  $X$  and  $\bar{X}$  are populated equally in the early Universe, the relic density is determined by standard freeze-out, requiring a thermally averaged cross section  $\langle \sigma v \rangle_{\text{ann}} \approx 6 \times 10^{-26} \text{ cm}^3/\text{s}$ . Using Eq. (5), the DM relic density is

$$\begin{aligned} \Omega_{\text{dm}} &\sim 0.2 \times \left( \frac{6 \times 10^{-26} \text{ cm}^3/\text{s}}{\langle \sigma v \rangle_{\text{ann}}} \right) \sim 0.2 \times \left( \frac{\alpha_X}{10^{-2}} \right)^{-2} \\ &\times \begin{cases} (m_X/300 \text{ GeV})^2 & \text{vector} \\ (m_X/100 \text{ GeV})^2 & \text{scalar} \end{cases}, \end{aligned} \quad (6)$$

<sup>2</sup>In principle, since numerical simulations follow the trajectories and velocities of each “particle,” these dependencies can be straightforwardly accounted for. However, SIDM simulations thus far have been restricted to cross sections with a  $v$  dependence that is either constant [14,15] or motivated within the classical limit ( $m_X v / m_\phi \gg 1$ ) [13,16] and with isotropic angular dependence.

where we have set  $\langle v^2 \rangle = 6T/m_X \sim 0.3$  for  $p$ -wave annihilation. Equation (6) demonstrates that SIDM keeps the virtues of the WIMP miracle intact, with weak-scale mass  $m_X$  and coupling  $\alpha_X$  giving the observed value  $\Omega_{\text{dm}} \sim 0.2$ . DM masses below the weak scale are also allowed provided the ratio  $\alpha_X/m_X$  remains fixed [24].

Alternately, DM may be asymmetric, such that  $X$  and  $\bar{X}$  are populated unequally in the early Universe due to a primordial  $X$ -number asymmetry [66–68]. In this case, annihilation still plays a crucial role for depleting the symmetric density of  $X$  and  $\bar{X}$ , leaving behind the residual asymmetric component composed of  $X$  only. The annihilation cross section is required to be *larger* than for thermal freeze-out,  $\langle \sigma v \rangle_{\text{ann}} > 6 \times 10^{-26} \text{ cm}^3/\text{s}$ . Conversely, for fixed  $\alpha_X$ , this gives an upper bound on  $m_X$  above which annihilation is insufficient to avoid overclosing the Universe. Since the mediator mass is light, the Born annihilation cross section is modified by a Sommerfeld enhancement. We include the Sommerfeld enhancement in our calculation of the DM relic density [69]. Sommerfeld enhancement provides an  $\mathcal{O}(1)$  correction to the ratio  $\alpha_X/m_X$  required by the relic density for  $m_X \gtrsim 100 \text{ GeV}$ , but has virtually no impact for smaller  $m_X$  (see Ref. [17]).

### III. COSMOLOGY OF LIGHT MEDIATORS

The presence of the light mediator  $\phi$  in the early Universe has important consequences. If it is stable, its comoving number density is

$$Y_\phi \approx 0.208 \frac{g_\phi \xi^3}{g_{*s}} \sim 7 \times 10^{-2} \xi^3, \quad (7)$$

where we set the internal degrees of freedom of the mediator  $g_\phi = 3$ , the entropy degrees of freedom  $g_{*s} = 10$ , and  $\xi$  is the temperature ratio of the dark sector to the SM. Compared to the baryon number density  $Y_B \sim 10^{-10}$ , we see that  $\phi$  can easily dominate the energy of the Universe if it is stable. One option to reduce the  $\phi$  energy density is to make the dark sector very cold. From Eq. (7), one would expect that  $Y_\phi$  becomes negligible if  $\xi \lesssim 10^{-3} (10 \text{ MeV}/m_\phi)^{1/3}$ . However,  $\xi$  cannot be arbitrarily small because DM particles have to be populated in the hidden sector thermal bath, which leads to a lower bound  $\xi \gtrsim 10^{-3} (g_{*s} 1 \text{ GeV}/m_X)^{1/3}$ .

To evade the overclosure problem, we assume that  $\phi$  is not stable and decays to SM particles. There are many cosmological constraints on the late decay of the mediator particle. Since the mediator has a mass in the range of 1–100 MeV, its decay products are electrons, positrons, and neutrinos. The energy injection from decay products (if lifetime is larger than about a second) may change the BBN light element abundances in a manner inconsistent with observations. For example, for a 100 GeV mass particle decaying to electromagnetic final states, the BBN constraint on the lifetime is  $\lesssim 10^4 \text{ s}$  [70]. However, it should be noted that for  $m_\phi$  below the  $^4\text{He}$  binding energy, the effect



on BBN abundances will be restricted to destruction of deuterium and lithium isotopes and a reconsideration of the effect of energy injection in this regime with a 1–100 MeV mass decaying particle is warranted.

The decays also lead to an increase in the entropy, and this change can be estimated as [71]

$$\frac{S_f - S_i}{S_i} \approx g_*^{1/4} \frac{m_\phi Y_\phi \tau_\phi^{1/2}}{m_{\text{pl}}} \sim 0.4 \xi^3 \left( \frac{m_\phi}{10 \text{ MeV}} \right) \left( \frac{\tau_\phi}{1 \text{ s}} \right)^{1/2}, \quad (8)$$

where  $S_i$  ( $S_f$ ) is the total entropy before (after)  $\phi$  decays, and  $\tau_\phi$  is the lifetime of  $\phi$ . In deriving Eq. (8), we have assumed radiation dominates the energy density before  $\phi$  decays, which is valid if  $T_D > 0.7 \text{ MeV} (m_\phi / 10 \text{ MeV}) \xi^3$  with  $T_D$  as the temperature when  $\phi$  decays. This condition can be satisfied in our model parameter space.

If the decay occurs after BBN with  $\tau_\phi \gtrsim 1 \text{ s}$ , entropy production has to be less than  $\sim 10\%$  (assuming no change to light element abundances due to it) because the precision measurements of the baryon density through BBN and CMB observations match very well [72]. Therefore, the absence of significant entropy production after BBN will put a strong bound on the lifetime  $\tau_\phi$  and the temperature ratio  $\xi$ . If the mediator decays before BBN, entropy production constraints are nonexistent.

We do not undertake a detailed study of the BBN and entropy production constraints but simply note that  $\tau_\phi = 1 \text{ s}$  is safe from such constraints if  $\xi \sim 0.3$ –1 depending on the mediator mass. We therefore adopt this value of  $\tau_\phi$  as a benchmark to set the coupling between  $\phi$  and the SM particles, and estimate the direct detection cross section based on it. A smaller value of  $\tau_\phi$  (at fixed  $\xi$ ) will lead to a stronger signal in direct detection experiments.

For simplicity, we will present our results based on  $\xi = 1$ , which can be extended to other cases by simple rescaling. Since the DM annihilation cross section scales as  $\xi$  to obtain the correct relic density [18],  $\alpha_X$  is proportional to  $\xi^{1/2}$  at fixed  $m_X$ . This implies that the lower bound on the symmetric SIDM mass from the CMB weakens by a factor of  $\xi$ . To keep the same self-scattering cross section, the mediator mass scales as  $\xi^{1/4}$ ,  $\xi^{1/2}$ , and  $\xi^{0.1}$  in the Born, resonant, and strongly coupled classical regimes, respectively. Therefore, the direct detection cross section in the limit of zero momentum transfer ( $q^2 = 0$ ) scales as  $\xi^{-1/2}$ ,  $\xi^{-3/2}$ , and  $\xi^{0.1}$  in the Born, resonant, and classical regimes, respectively. For momentum transfer  $q \gg m_\phi$  [see Eq. (21)], the direct detection cross section scales as  $\xi^{1/2}$ . Since the favored  $\xi$  is in the range of 0.3–1, these are all  $\mathcal{O}(1)$  corrections.

#### IV. PORTALS FOR LIGHT MEDIATORS

In this section, we discuss three minimal ways that the dark mediator  $\phi$  may be coupled to the SM. Depending on its spin,  $\phi$  can mix with the photon or  $Z$  boson if it is a vector, or it can mix with the Higgs boson if it is a scalar.

This mixing allows  $\phi$  particles, produced thermally in the early Universe, to decay. That these decays do not affect BBN and dilute the baryon density puts a constraint on the mixing parameter, which in turn provides a lower bound on the DM direct detection cross section.

##### A. Vector mediator case: Photon and $Z$ mixing

Since gauge symmetries exist in the visible sector, it is natural to speculate that they may be present in the dark sector as well. The simplest example is a  $U(1)_X$  symmetry under which DM is charged, with  $\phi$  being the corresponding vector boson. We assume that the  $U(1)_X$  is spontaneously broken, say by a dark Higgs, to generate a mass  $m_\phi$ . (The massless case has also been discussed in Refs. [73,74].)

If there exist particles that are charged under both the  $U(1)_X$  and the SM  $SU(2)_L \times U(1)_Y$  gauge symmetries, this leads to mixing between  $\phi$  and the photon  $\gamma$  or  $Z$  boson. The most relevant interactions are

$$\mathcal{L}_{\text{mixing}} = \frac{\varepsilon_\gamma}{2} \phi_{\mu\nu} F^{\mu\nu} + \delta m^2 \phi_\mu Z^\mu, \quad (9)$$

where  $\phi_{\mu\nu} \equiv \partial_\mu \phi_\nu - \partial_\nu \phi_\mu$  is the  $U(1)_X$  field strength. The first term corresponds to photon kinetic mixing [75], which has been widely studied as a DM portal [22,23,76,77] and in dark photon searches [30]. The second term corresponds to mass mixing with the  $Z$  [78–80]. This term explicitly breaks gauge invariance, but can arise via higher-dimensional operators with additional Higgs and dark Higgs insertions. Defining  $\varepsilon_Z \equiv \delta m^2 / m_Z^2$ , we work in the limit of small mixing parameters,  $\varepsilon_{\gamma,Z} \ll 1$ .

At  $\mathcal{O}(\varepsilon_{\gamma,Z})$ , mixing via Eq. (9) induces a coupling of  $\phi$  to SM fermions, given by

$$\mathcal{L}_{\text{int}} = \left( \varepsilon_\gamma e J_{\text{em}}^\mu + \varepsilon_Z \frac{g_2}{c_W} J_{\text{NC}}^\mu \right) \phi_\mu, \quad (10)$$

where the electromagnetic and weak neutral currents are, respectively,

$$\begin{aligned} J_{\text{em}}^\mu &= \sum_f Q_f \bar{f} \gamma^\mu f, \\ J_{\text{NC}}^\mu &= \sum_f \bar{f} \gamma^\mu (T_{3f} P_L - Q_f s_W^2) f. \end{aligned} \quad (11)$$

Here,  $g_2$  is the  $SU(2)_L$  coupling,  $\theta_W$  is the weak mixing angle (with  $s_W \equiv \sin \theta_W$  and  $c_W \equiv \cos \theta_W$ ), and  $Q_f$  and  $T_{3f}$  denote the charge (in units of  $e$ ) and weak isospin for SM fermion  $f$ .

The decay rate for  $\phi$  can be straightforwardly computed from Eq. (10). Since  $m_\phi \sim 1$ –100 MeV is typically preferred by solving small scale structure anomalies with SIDM, only  $e^+e^-$ ,  $\nu\bar{\nu}$ , and photon final states are allowed.

For kinetic mixing, the decay is dominated by  $\phi \rightarrow e^+e^-$ , with decay rate and lifetime

$$\Gamma_\phi = \frac{\alpha_{\text{em}} m_\phi \varepsilon_\gamma^2}{3} \Rightarrow \tau_\phi \approx 3 \text{ s} \times \left( \frac{\varepsilon_\gamma}{10^{-10}} \right)^{-2} \left( \frac{m_\phi}{10 \text{ MeV}} \right)^{-1}, \quad (12)$$

where  $\alpha_{\text{em}}$  is the electromagnetic fine structure constant. For  $Z$  mixing, neutrino final states dominate since there are three families and also the electron's weak vector charge is suppressed. The total decay rate and lifetime are

$$\Gamma_\phi = \frac{\alpha_{\text{em}} m_\phi \varepsilon_Z^2 (1 - s_W^2 + 2s_W^4)}{6s_W^2 c_W^2} \Rightarrow \tau_\phi \approx 1 \text{ s} \times \left( \frac{\varepsilon_Z}{10^{-10}} \right)^{-2} \left( \frac{m_\phi}{10 \text{ MeV}} \right)^{-1}, \quad (13)$$

and the branching ratios are  $\text{BR}(\phi \rightarrow \nu\bar{\nu}) \approx 6/7$  and  $\text{BR}(\phi \rightarrow e^+e^-) \approx 1/7$ . Ensuring that decays  $\phi \rightarrow e^+e^-$  do not occur after BBN requires a lifetime  $1/\Gamma_\phi \lesssim 1 \text{ s}$ . Therefore, the mixing parameters are constrained to be  $\varepsilon_{\gamma,Z} \gtrsim 10^{-10} \times \sqrt{10 \text{ MeV}/m_\phi}$  [20].

The upper bound from the low energy beam dump experiments is  $\varepsilon_\gamma \lesssim 10^{-7}$  for  $m_\phi \lesssim 400 \text{ MeV}$  [30], while the region  $10^{-10} \lesssim \varepsilon_\gamma \lesssim 10^{-7}$  is excluded for  $m_\phi \lesssim 100 \text{ MeV}$  by energy loss arguments in supernovae [81] (although this constraint depends sensitively on assumptions about the temperature and size of the supernova core).

Next, we discuss direct detection. Given Eq. (10), direct detection signals arise from DM scattering on nuclei via  $\phi$  exchange, shown in Fig. 1. In general, we can parametrize the coupling of  $\phi$  to protons ( $p$ ) and neutrons ( $n$ ) as

$$\mathcal{L}_{\text{int}} = e\phi_\mu (\varepsilon_p \bar{p}\gamma^\mu p + \varepsilon_n \bar{n}\gamma^\mu n), \quad (14)$$

where  $\varepsilon_{p,n}$  are the effective nucleon couplings (in units of  $e$ ). For kinetic mixing or  $Z$  mixing, they are given by

$$\begin{aligned} \varepsilon_p &= \varepsilon_\gamma + \frac{\varepsilon_Z}{4s_W c_W} (1 - 4s_W^2) \approx \varepsilon_\gamma + 0.05\varepsilon_Z, \\ \varepsilon_n &= -\frac{\varepsilon_Z}{4s_W c_W} \approx -0.6\varepsilon_Z. \end{aligned} \quad (15)$$

That is, kinetic mixing couples  $\phi$  to protons only, since only protons carry an electric charge, while  $Z$  mixing couples  $\phi$  mainly to neutrons, since the weak charge of the proton is  $(1 - 4s_W^2) \ll 1$ . For arbitrary  $\varepsilon_{\gamma,Z}$ , this scenario is maximally isospin violating in the sense that any ratio of  $\varepsilon_p/\varepsilon_n$  is allowed [82]. For a given nucleus  $N$  with proton number  $Z$  and mass number  $A$ , the spin independent cross section is, in the limit of zero momentum transfer ( $q^2 = 0$ ),

$$\sigma_{XN}^{\text{SI}} = \frac{16\pi\alpha_{\text{em}}\alpha_X\mu_{XN}^2}{m_\phi^4} (\varepsilon_p Z + \varepsilon_n (A - Z))^2, \quad (16)$$

where  $\mu_{XN} = m_X m_N / (m_X + m_N)$  is the DM-nucleus reduced mass.

Cosmological constraints imply that the SIDM direct detection cross section may be within the reach of present or future searches. For kinetic mixing, the SI cross section on protons is

$$\sigma_{Xp}^{\text{SI}} \approx 1.5 \times 10^{-24} \text{ cm}^2 \times \varepsilon_\gamma^2 \times \left( \frac{\alpha_X}{10^{-2}} \right) \left( \frac{m_\phi}{30 \text{ MeV}} \right)^{-4}, \quad (17)$$

while for  $Z$  mixing, the SI cross section on neutrons is

$$\sigma_{Xn}^{\text{SI}} \approx 5 \times 10^{-25} \text{ cm}^2 \times \varepsilon_Z^2 \times \left( \frac{\alpha_X}{10^{-2}} \right) \left( \frac{m_\phi}{30 \text{ MeV}} \right)^{-4}. \quad (18)$$

If we take  $\varepsilon_{\gamma,Z} \gtrsim 10^{-10}$  as a lower bound from BBN, the predicted SI cross section per nucleon is predicted to be greater than  $\sim 10^{-45} - 10^{-44} \text{ cm}^2$ , which is in the reach of direct detection experiments.

For SIDM, the typical mediator mass is  $m_\phi \sim 1 - 100 \text{ MeV}$ , which is comparable to or less than the typical momentum transfer  $q \sim 50 \text{ MeV}$  in nuclear recoils. Therefore, unlike typical WIMPs, nuclear recoil interactions for SIDM are momentum dependent and cannot be approximated by a contact interaction [83–85]. The differential recoil rate per target nucleus is

$$\frac{dR}{dE_R} = \frac{\rho_{\text{DM}}}{m_X} \int_{v_{\min}} d^3v v f(\vec{v}) \frac{d\sigma_{XN}^{\text{SI}}(v, E_R)}{dE_R} \quad (19)$$

$$\begin{aligned} &= \frac{\rho_{\text{DM}}}{m_X} \int_{v_{\min}} d^3v v f(\vec{v}) \left( \frac{d\sigma_{XN}^{\text{SI}}(v, E_R)}{dE_R} \right)_{q^2=0} \\ &\quad \times \frac{m_\phi^4}{(m_\phi^2 + q^2)^2}, \end{aligned} \quad (20)$$

where  $E_R$  is the nuclear recoil energy,  $v_{\min} = \sqrt{m_N E_R / 2\mu_{XN}^2}$  is the minimum DM velocity for a given  $E_R$ ,  $f(\vec{v})$  is the local DM velocity distribution, and  $\rho_{\text{DM}}$  is the local DM density. The last term in Eq. (20) shows how the momentum transfer, given by  $q = \sqrt{2m_N E_R}$ , provides a suppression of the recoil rate compared to taking  $q^2 = 0$  assumed by a contact interaction. Clearly lighter target nuclei and lower thresholds are favorable for detecting SIDM.

Since cross section limits quoted by the experimental collaborations assume a contact interaction, one must perform a reanalysis of experimental data to derive constraints on SIDM. Reference [85] performed such an analysis for the kinetic mixing model we consider for several light mediator masses ( $m_\phi = 0, 10$ , and  $30 \text{ MeV}$ ), although not within an SIDM context. In the present work, we account for this suppression effect in a simplified

manner by multiplying the total  $q^2 = 0$  cross section by a  $q^2$ -dependent form factor,

$$\sigma_{XN}^{\text{SI}}(q^2) = \sigma_{XN}^{\text{SI}}(q^2 = 0) \times \frac{m_\phi^4}{(m_\phi^2 + q^2)^2}. \quad (21)$$

We take a fixed value of  $q$ , corresponding to a typical momentum transfer for a given experiment. We assume that the experimentally quoted cross section limits apply directly to Eq. (21), and then we extract what are the constraints on the particle physics parameters. We defer a complete reanalysis over the whole range of  $m_\phi$  for SIDM to future work. However, we find that this simple prescription is able to reproduce the XENON100 and CDMS reanalysis in Ref. [85], for an appropriate choice of  $q$ , with little dependence on astrophysical halo properties considered therein.

It is also useful to consider the direct detection cross section for DM scattering on electrons, which was proposed as an avenue for detecting sub-GeV DM [59,60]. The SI cross section is

$$\begin{aligned} \sigma_{Xe}^{\text{SI}} &= \frac{16\pi\alpha_{\text{em}}\alpha_X\mu_{Xe}^2\varepsilon_\gamma^2}{m_\phi^4} \approx 4 \times 10^{-29} \text{ cm}^2 \times \varepsilon_\gamma^2 \\ &\times \left(\frac{\alpha_X}{10^{-2}}\right) \left(\frac{m_\phi}{10 \text{ MeV}}\right)^{-4}. \end{aligned} \quad (22)$$

Scattering on electrons is mainly sensitive to  $\varepsilon_\gamma$ , while the  $\varepsilon_Z$  term is suppressed by the electron weak charge and we have neglected it in Eq. (22). Current sensitivities from XENON10 at the level of  $\sigma_{Xe}^{\text{SI}} \sim 10^{-38} \text{ cm}^2$  for sub-GeV mass DM therefore probe this scenario down to  $\varepsilon_\gamma \sim 10^{-4}$  [60].

Last, we discuss indirect detection. If the DM density is symmetric, i.e., populated by both  $X$  and  $\bar{X}$ , then  $X\bar{X} \rightarrow \phi\phi$  annihilation, with  $\phi$  decaying to SM states, can provide visible astrophysical signals. Observations of the ionization history of the Universe from the CMB constrain DM annihilation during recombination [86,87], in particular the process  $X\bar{X} \rightarrow \phi\phi \rightarrow e^+e^-e^+e^-$ . A recent analysis combining several cosmological data sets found  $m_X > 30 \text{ GeV}$  for  $\langle\sigma v\rangle_{\text{ann}}$  fixed by relic abundance and if  $\text{BR}(\phi \rightarrow e^+e^-) = 1$  [88,89], as in the case of kinetic mixing. Implications for indirect detection searches (e.g., Fermi and AMS-2) are also important and will be considered elsewhere [90]. On the other hand, in the case of  $Z$  mixing, these limits are weakened by a factor of  $\sim 7$  since  $\phi$  predominantly decays to neutrinos. In this scenario, DM capture and annihilation in the Sun may provide neutrino-rich signals that may be observed by IceCube. Indirect detection signals from DM annihilation are not present if DM is asymmetric.

## B. Scalar mediator case: Higgs mixing

If the dark mediator  $\phi$  is a scalar, the leading renormalizable couplings to SM particles arise through the Higgs

portal [91–101]. Assuming  $\phi$  is a real scalar singlet, the relevant terms in the scalar potential are

$$V(H, \phi) \supset (a\phi + b\phi^2)|H|^2, \quad (23)$$

where  $H$  is the Higgs doublet and  $a, b$  are coupling constants. After electroweak symmetry breaking, mixing arises between  $\phi$  and the physical Higgs boson  $h$  due to the Higgs vacuum expectation value (vev)  $v \approx 246 \text{ GeV}$ . In the limit  $a, m_\phi \ll v, m_h$ , this mixing angle is  $\varepsilon_h \approx av/m_h^2$ . This generates an effective  $\phi$  coupling to SM fermions

$$\mathcal{L}_{\text{int}} = -\frac{m_f \varepsilon_h}{v} \bar{f} f \phi. \quad (24)$$

For  $m_\phi \sim 1\text{--}100 \text{ MeV}$ , the dominant decay channel is  $\phi \rightarrow e^+e^-$ , while  $\phi \rightarrow \gamma\gamma$  is highly suppressed.<sup>3</sup> The decay rate and lifetime for  $\phi$  are

$$\Gamma_\phi = \frac{\varepsilon_h^2 m_e^2 m_\phi}{8\pi v^2} \Rightarrow \tau_\phi \approx 4 \text{ s} \times \left(\frac{\varepsilon_h}{10^{-5}}\right)^{-2} \left(\frac{m_\phi}{10 \text{ MeV}}\right)^{-1}. \quad (25)$$

Therefore, having  $\phi$  decay before BBN implies that  $\varepsilon_h \gtrsim 10^{-5}$ .

Coupling the Higgs to the dark sector can lead to invisible Higgs decays. The  $a$  term allows for  $h \rightarrow X\bar{X}$ , provided  $m_X < m_h/2$ , while the  $b$  term allows for  $h \rightarrow \phi\phi$  (with  $\phi$  escaping the detector as missing energy before decay). Assuming an otherwise SM-like Higgs boson, the resulting invisible partial widths are

$$\begin{aligned} \Gamma(h \rightarrow X\bar{X}) &= \frac{\alpha_X \varepsilon_h^2 m_h}{2} \left(1 - \frac{4m_X^2}{m_h^2}\right)^{3/2}, \\ \Gamma(h \rightarrow \phi\phi) &= \frac{b^2 v^2}{8\pi m_h}. \end{aligned} \quad (26)$$

The latter channel is utterly negligible if we require no fine-tuning in the  $\phi$  mass, such that the vev-induced mass,  $bv^2$ , is smaller than  $m_\phi^2$ . If  $|b|v^2 \lesssim m_\phi^2$ , then  $\text{Br}(h \rightarrow \phi\phi) \lesssim 10^{-10} \times (m_\phi/100 \text{ MeV})^4$ . On the other hand,  $h \rightarrow X\bar{X}$  can be phenomenologically relevant for the LHC. Invisible Higgs branching ratios larger than 10% (1%) can be achieved for  $\alpha_X \varepsilon_h^2 \gtrsim 10^{-5}$  ( $10^{-6}$ ). For  $m_X \gtrsim 10 \text{ GeV}$ , direct detection constraints strongly exclude such values (see below), but for lighter DM masses this remains an open

<sup>3</sup>For the Higgs boson of mass  $m_h \approx 125 \text{ GeV}$ , the branching ratio to  $\gamma\gamma$  is much larger than to  $e^+e^-$ , with  $\text{BR}(h \rightarrow \gamma\gamma) \approx 2 \times 10^{-3}$  and  $\text{BR}(h \rightarrow e^+e^-) \approx 5 \times 10^{-9}$ . However, a similar conclusion does not hold for  $\phi$  due to its much smaller mass entering the  $\gamma\gamma$  loop amplitude. The second and third generation fermion loop amplitudes accidentally cancel the  $W$  boson loop amplitude to  $\sim 1\%$ , and therefore the contribution from first generation fermion loops is crucial, including hadronic effects for quarks, which is beyond the scope of this paper.

possibility and is one of the few ways that the LHC can probe SIDM.

For direct detection, the SI cross section is given by Eq. (16) with

$$\varepsilon_{p,n} = \frac{m_{p,n}\varepsilon_h}{ev} \left(1 - \frac{7}{9}f_{TG}^{(p,n)}\right) \approx 3 \times 10^{-3} \times \varepsilon_h, \quad (27)$$

where we take  $f_{TG}^{(p,n)} \approx 0.943$  for the gluon hadronic matrix element [98]. As opposed to the vector case, SI scattering is isospin conserving, with equal cross sections for scattering on protons and neutrons, given by

$$\sigma_{Xp}^{\text{SI}} \approx 2 \times 10^{-29} \text{ cm}^2 \times \varepsilon_h^2 \times \left(\frac{\alpha_X}{10^{-2}}\right) \left(\frac{m_\phi}{30 \text{ MeV}}\right)^{-4}. \quad (28)$$

Since decays before BBN require  $\varepsilon_h \gtrsim 10^{-5}$ , this scenario is strongly excluded by current direct detection bounds for  $m_X \gtrsim 10 \text{ GeV}$ . As opposed to the vector mediator case, where quark and lepton couplings are universal, the scalar mediator case is comparatively more constrained by direct detection due to the larger couplings to quarks compared to electrons.

## V. SIDM MODELS FOR DIRECT DETECTION

Direct detection plays a complementary role to astrophysical observables in constraining SIDM. Our main results, presented in Figs. 2 and 3, map out the entire parameter space for different benchmark scenarios and highlight the crucial importance of direct detection experiments, especially upcoming searches, for SIDM. For illustrative purposes, we focus on limits from XENON100 [102], as well as projected limits for XENON1T [53], although other experiments such as LUX [54] and SuperCDMS [55] will also be important.

In Fig. 2, we consider four benchmark SIDM models, corresponding to different scenarios with symmetric or asymmetric DM, vector or scalar mediators  $\phi$ , and different portals for coupling  $\phi$  to the SM. First, we summarize the different observables we consider for constraining SIDM:

- (i) Self-interactions within dwarf halos can modify the DM distribution in line with observations for  $\sigma/m_X \sim 1 \text{ cm}^2/\text{g}$  [13–16]. The shaded bands in Figs. 2 and 3 show where these anomalies are ameliorated within a generous range,  $0.1 \text{ cm}^2/\text{g} \lesssim \sigma/m_X \lesssim 10 \text{ cm}^2/\text{g}$ . We assume a characteristic dwarf velocity  $v_0 = 30 \text{ km/s}$  as a compromise between smaller MW dwarf spheroidals ( $v_0 \sim 10 \text{ km/s}$ ) and larger low surface brightness galaxies ( $v_0 \sim 100 \text{ km/s}$ ).
- (ii) The ellipticity of the inner regions of DM halos, from elliptical galaxies to clusters of galaxies, suggests that  $\sigma/m_X \lesssim 1 \text{ cm}^2/\text{g}$  on these scales [15]. Of these, elliptical galaxies provide in principle the best constraint for velocity-dependent self-interaction cross

sections due to the smaller relative velocity. Based on SIDM-only simulations (no baryons), NGC 720 seems to be inconsistent with  $\sigma/m_X \gtrsim 1 \text{ cm}^2/\text{g}$  [15] and analyses of more elliptical galaxies could solidify this conclusion. Based on this expectation, the red dot-dashed contour (labeled “Halo shapes”) indicates the approximate upper bound  $\sigma/m_X = 1 \text{ cm}^2/\text{g}$  from NGC 720, adopting a characteristic velocity  $v_0 = 300 \text{ km/s}$  [15]. A similar limit on  $\sigma/m_X$  for clusters of galaxies was also suggested in Ref. [15] (based on comparison of SIDM-only simulations to some LoCuSS clusters [103]). This constraint (not shown) is weaker due to the larger relative velocity ( $v_0 \sim 1000 \text{ km/s}$ ) and would lie between the halo shapes and bullet cluster contours in Fig. 2. It is important to emphasize that these halo shapes bounds are based on SIDM-only simulations *without baryons*, and we urge caution not to interpret these limits as strict constraints. These limits (or lack thereof) depend on how much the rounded inner halos exhibited by SIDM-only simulations are modified by the aspherical baryonic component [104], which will be addressed in future work.

- (iii) For the bullet cluster constraint, we require  $\sigma/m_X \lesssim 1 \text{ cm}^2/\text{g}$  for a relative velocity  $v \approx 3000 \text{ km/s}$  [42,56].<sup>4</sup> With the discoveries of other merging cluster systems, combined analyses of many such objects have the potential to improve this bound further. We consider the potential reach of a projected constraint  $\sigma/m_X \lesssim 0.1 \text{ cm}^2/\text{g}$  for  $v \approx 2000 \text{ km/s}$ , denoted as “Merging clusters.” These limits are shown by the green dot-dashed contours.
- (iv) CMB constraints apply to symmetric DM with  $s$ -wave annihilation  $X\bar{X} \rightarrow \phi\phi$ , requiring  $m_X \gtrsim 30 \text{ GeV}$  if  $\text{BR}(\phi \rightarrow e^+e^-) \approx 1$  [88]. This is the case for  $\gamma$  kinetic mixing with  $m_\phi \sim 1\text{--}100 \text{ MeV}$ , while for  $Z$  mixing, the bound is weaker by a factor of  $\text{BR}(\phi \rightarrow e^+e^-) \approx 1/7$ . The CMB exclusion region is denoted by the brown hatched boundary.
- (v) The purple solid contours indicate constraints from XENON100 [102] for different  $\varepsilon$  parameters, while the purple short-dashed contours indicate the projected XENON1T reach with 2.2 ton  $\cdot$  yr exposure [53].

We emphasize that the quoted XENON cross section sensitivities assume an isospin-conserving contact interaction. Here, we modify these sensitivities by the xenon proton (neutron) fraction for the case of  $\gamma$  kinetic ( $Z$ ) mixing. We also include a  $q^2$ -dependent form factor, described in Eq. (21), since the typical momentum transfer  $q$  is comparable to  $m_\phi$ . We take  $q \approx 50 \text{ MeV}$  for xenon.

The first benchmark model is the case of symmetric SIDM with a vector mediator  $\phi$  that couples through  $\gamma$

<sup>4</sup>We do not average  $\sigma_T$  over a velocity distribution, but rather consider a single fixed relative velocity  $v$ .



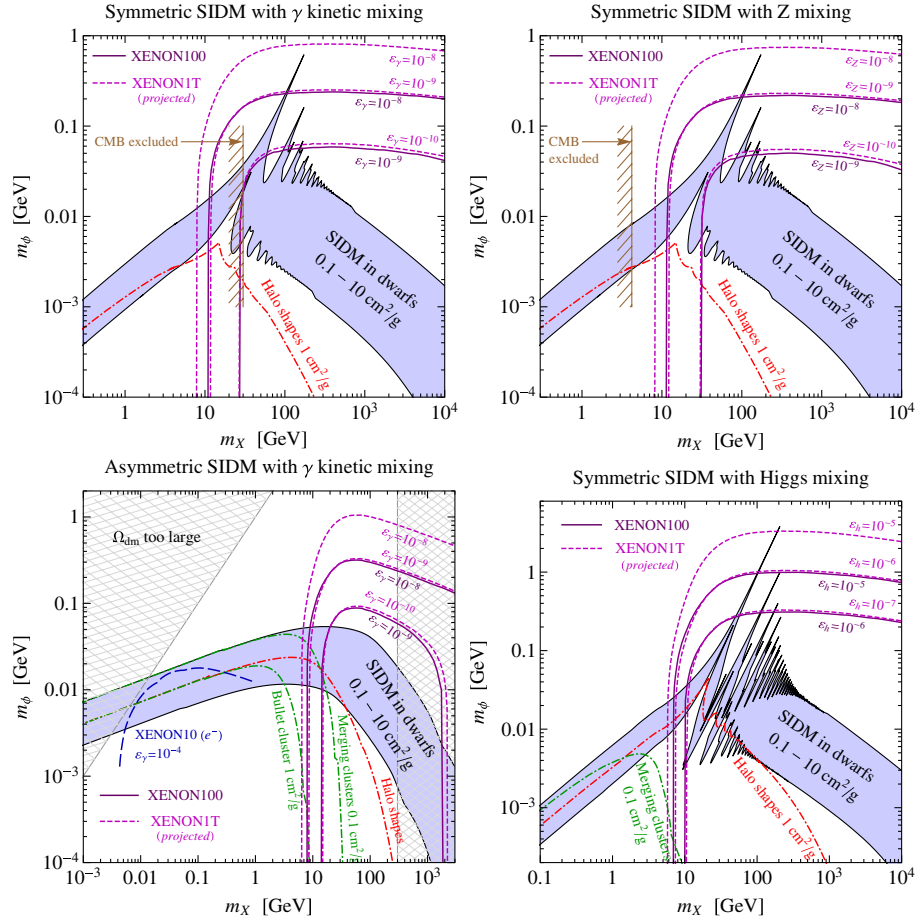


FIG. 2 (color online). Shaded band shows parameter space region for explaining small scale anomalies in different SIDM models. Dot-dashed contours show constraints from astrophysical observations on larger scales (region below is excluded). Sensitivities from XENON100 and XENON1T are shown by solid and short-dashed contours, respectively, for different  $\epsilon$  parameters. Limits from  $e^-$  recoils in XENON10 are shown by the long-dashed contour. See text for further details.

kinetic mixing. The constraints on this model are shown in Fig. 2 (upper left) as a function of  $(m_X, m_\phi)$ . The DM relic density, determined by thermal freeze-out, fixes  $\alpha_X \approx 4 \times 10^{-5} \times (m_X/\text{GeV})$ . Self-interactions are a combination of attractive and repulsive, and the peaklike features correspond to where scattering is resonantly enhanced. The halo shapes bound (with excluded regions below the red dot-dashed contour, when taken at face value) is weak due to the velocity dependence of the self-interaction cross section. However, the SIDM region below 30 GeV is excluded by CMB. The sensitivity contours for XENON100 and XENON1T are shown for different values of  $\epsilon_\gamma$ , excluding SIDM parameter space below the curves.

The second benchmark model shown in Fig. 2 (upper right) is the same as the previous case, but with  $\phi$  coupled through  $Z$  mixing. The self-interaction observables and relic density are unchanged. However, the CMB bound is weakened since  $\phi$  decays predominantly to neutrinos. Direct detection sensitivities are also affected due to the different couplings, but are quantitatively similar to the kinetic mixing scenario.

Third, we consider a scenario with asymmetric SIDM, shown in Fig. 2 (lower left), with fixed  $\alpha_X = 10^{-2}$ . For a mediator, we consider a vector  $\phi$  coupled through  $\gamma$  kinetic mixing. No resonant features are present since scattering is purely repulsive (no bound states). Also, since annihilation does not occur, indirect detection signals are absent and the CMB bound does not apply. Thus, we extend the parameter range to sub-GeV DM masses. The dot-dashed contours show how constraints from astrophysical observations can constrain the low mass regime (below the curves is excluded). Outside the range  $m_\phi < m_X \lesssim 300$  GeV there is insufficient annihilation of the symmetric density (hatched region), although larger values of  $\alpha_X$  allow for larger  $m_X$ .<sup>5</sup> In addition to the XENON100 and XENON1T contours, we also show the reach for direct detection via electron recoils for constraining sub-GeV SIDM. The current limits from XENON10 are shown by the long dashed curve for  $\epsilon_\gamma = 10^{-4}$ . (For such large values of  $\epsilon_\gamma$ , the

<sup>5</sup>We neglect a narrow parameter strip near  $m_\phi \approx 2m_X$  where resonant  $s$ -channel annihilation can occur.



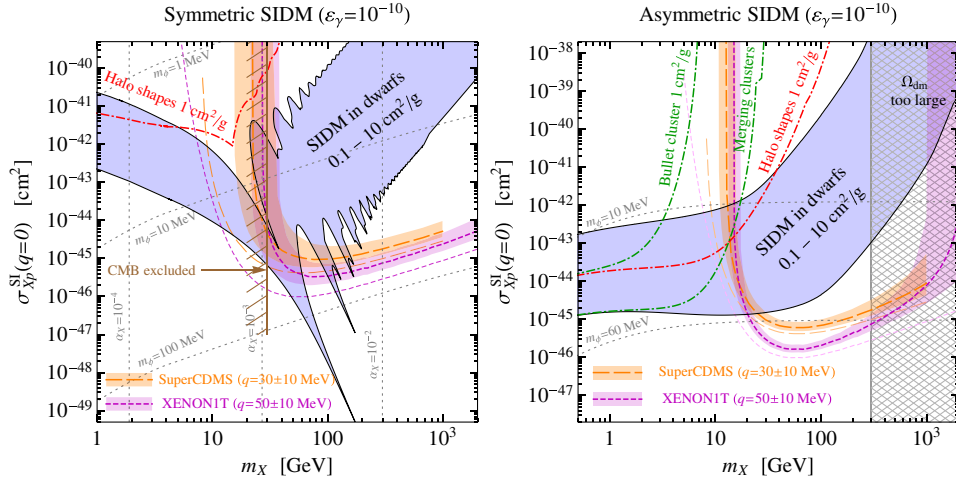


FIG. 3 (color online). Spin-independent direct detection cross section vs DM mass for symmetric (left) and asymmetric (right) SIDM with a vector mediator  $\phi$  coupled via kinetic mixing. Shaded band is where SIDM solves small scale structure anomalies. Dot-dashed curves show astrophysical constraints from larger scales (excluded region lies above). Short-dashed purple curve denotes projected limit from XENON1T for  $\varepsilon_\gamma = 10^{-10}$ . See text for details.

$m_X \gtrsim 1$  GeV region is strongly excluded by direct detection via nuclear recoils [59].)

Our final benchmark model is symmetric SIDM with a scalar mediator coupled via Higgs mixing, shown in Fig. 2 (lower right). The DM relic density fixes  $\alpha_X \approx 10^{-4} \times (m_X/\text{GeV})$ . Self-interactions are purely attractive. Indirect detection bounds do not apply since annihilation is  $p$ -wave suppressed. DM self-scattering is purely attractive, and large resonant features are evident. Since annihilation is  $p$  wave, the CMB bound does not apply.

We make several comments on the results shown in Fig. 2:

- (i) The shaded bands show that small scale structure anomalies can be explained for a wide range of SIDM mass, from sub-GeV to multi-TeV, and the mediator mass is typically in the  $\sim 1$ –100 MeV range.
- (ii) Direct detection limits constrain SIDM above  $\sim 10$  GeV; the exclusion region lies below the solid (dashed) purple curves for XENON100 (XENON1T). In this regime, self-interactions are velocity dependent, and SIDM easily evades the bullet cluster and halo shapes bounds from larger scales.
- (iii) For DM below 10 GeV, astrophysical constraints from indirect detection and structure observables are most important for constraining SIDM. In this regime, the self-interaction cross section is more velocity independent, and future merging cluster studies will be important. Direct detection via electron recoils provides another avenue to explore low mass SIDM, albeit with larger  $\varepsilon_\gamma$ . Direct detection via nuclear recoils is below threshold and provides no constraint.
- (iv) For vector mediator scenarios, ton-scale direct detection experiments will cover almost the entire SIDM parameter space for  $m_X \gtrsim 25$  GeV, down to the lower

bound  $\varepsilon_{\gamma,Z} \sim 10^{-10}$  imposed by BBN. Current experiments have no constraint for  $\varepsilon_{\gamma,Z} \sim 10^{-10}$ .

- (v) For a scalar mediator with Higgs mixing, this scenario is almost completely excluded by XENON100 unless  $m_X \lesssim 5$  GeV. This case is more strongly constrained by direct detection compared to the vector mediator case since  $\phi$  couplings to SM fermions are proportional to mass. Therefore, the effective  $\phi$ -nucleon coupling entering direction detection is enhanced compared to the  $\phi$ -electron coupling constrained by BBN.

In Fig. 3, we show how SIDM benchmark models map onto the direct detection plane of cross section vs mass. For simplicity, we focus on the case where the mediator  $\phi$  is a vector. We assume kinetic mixing with  $\varepsilon_\gamma = 10^{-10}$  motivated by BBN constraints. The different panels are for symmetric DM (left) or asymmetric DM (right), corresponding to the first two benchmarks shown in Fig. 2 (top left and right). The vertical axis is the spin-independent cross section for scattering on protons in the  $q^2 = 0$  limit. We emphasize that this cross section, however, does not appear in the direct detection rate by itself but must be multiplied by a form factor taking into account the finite  $q^2$ . The short-dashed purple curve shows the sensitivity reach of XENON1T taking  $q = 50$  MeV, while the long-dashed orange curve shows the sensitivity reach of SuperCDMS (SNOLAB) based on  $385 \text{ kg} \cdot \text{yr}$  exposure [105] taking  $q = 30$  MeV. To illustrate the importance of the  $q^2$ -dependent form factor, the surrounding band shows how the XENON1T and CDMS limits change by varying  $q$  in the range  $\pm 10$  MeV, while the corresponding thin contours show how the limits would appear for  $q = 0$ . As above, the shaded band shows where SIDM solves small scale anomalies, while the dot-dashed contours show astrophysical

constraints on self-interactions on larger scales. The hatched boundary shows the CMB exclusion limit on symmetric DM.

## VI. CONCLUSIONS

SIDM is a simple and well-motivated scenario that can explain small scale structure anomalies observed in dwarf galaxies. Self-interactions arise through a 1–100 MeV mediator particle in the dark sector. Requiring that mediator particles decay to SM states before BBN implies a minimal coupling between DM and the SM, which can be probed in direct detection experiments. In this work, we have considered three mechanisms for coupling the mediator to the SM and studied the direct detection implications of SIDM. With a light mediator favored by the small scale structure of the Universe, SIDM has direct detection features different from usual WIMPs. Current and future direct detection experiments are sensitive to SIDM candidates, even if the coupling between two sectors is extremely feeble. SIDM also interacts with target nuclei with momentum-dependent interactions, which may provide a mechanism to reconcile some of the recent dark matter direct detection hints [83,85,106]. For  $\gamma$  or  $Z$  mixing, we found that current direct detection limits provide no constraint on SIDM if the coupling strength saturates the lower bound

from the cosmological considerations. However, future ton-scale experiments will explore the entire parameter range for SIDM above  $\sim 20$  GeV. On the other hand, for Higgs mixing, the same mass range is already excluded by XENON100. Thus, astrophysical observations and direct detection experiments complement each other in the search for SIDM candidates.

## ACKNOWLEDGMENTS

We thank Sunil Golwala, Moira Gresham, Rafael Lang, David Sanford, and Andrew Zentner for useful discussions, as well as the organizers of the Harvard SIDM 2013 workshop for a stimulating meeting and hospitality during the completion of this work. M. K. is supported by NSF Grant No. PHY-1214648. H. B. Y. is supported by startup funds from UCR. S. T. is supported by the DOE under Contract No. DE-SC0007859 and NASA Astrophysics Theory Grant No. NNX11AI17G. S. T. would also like to thank Aspen Center for Physics and NSF Grant No. 1066293, as well as CETUP\* (Center for Theoretical Underground Physics and Related Areas), supported by DOE Grant No. DE-SC0010137 and NSF Grant No. PHY-1342611, for hospitality and partial support during the completion of this work.

- 
- [1] E. D. Carlson, M. E. Machacek, and L. J. Hall, *Astrophys. J.* **398**, 43 (1992).
  - [2] D. N. Spergel and P. J. Steinhardt, *Phys. Rev. Lett.* **84**, 3760 (2000).
  - [3] J. Fan, A. Katz, L. Randall, and M. Reece, *Phys. Rev. Lett.* **110**, 211302 (2013).
  - [4] B. Moore, *Nature (London)* **370**, 629 (1994).
  - [5] R. A. Flores and J. R. Primack, *Astrophys. J.* **427**, L1 (1994).
  - [6] J. F. Navarro, C. S. Frenk, and S. D. White, *Astrophys. J.* **490**, 493 (1997).
  - [7] M. G. Walker and J. Penarrubia, *Astrophys. J.* **742**, 20 (2011).
  - [8] S.-H. Oh, W. J. G. de Blok, E. Brinks, F. Walter, and R. C. Kennicutt, Jr., *Astrophys. J.* **141**, 193 (2011).
  - [9] R. Kuzio de Naray, S. S. McGaugh, and W. J. G. de Blok, *Astrophys. J.* **676**, 920 (2008).
  - [10] M. Boylan-Kolchin, J. S. Bullock, and M. Kaplinghat, *Mon. Not. R. Astron. Soc.* **422**, 1203 (2012).
  - [11] F. Governato, A. Zolotov, A. Pontzen, C. Christensen, S. H. Oh, A. M. Brooks, T. Quinn, S. Shen, and J. Wadsley, *Mon. Not. R. Astron. Soc.* **422**, 1231 (2012).
  - [12] A. M. Brooks and A. Zolotov, *arXiv:1207.2468*.
  - [13] M. Vogelsberger, J. Zavala, and A. Loeb, *Mon. Not. R. Astron. Soc.* **423**, 3740 (2012).
  - [14] M. Rocha, A. H. G. Peter, J. S. Bullock, M. Kaplinghat, S. Garrison-Kimmel, J. Onorbe, and L. A. Moustakas, *Mon. Not. R. Astron. Soc.* **430**, 81 (2013).
  - [15] A. H. Peter, M. Rocha, J. S. Bullock, and M. Kaplinghat, *arXiv:1208.3026*.
  - [16] J. Zavala, M. Vogelsberger, and M. G. Walker, *Mon. Not. R. Astron. Soc.* **431**, L20 (2013).
  - [17] S. Tulin, H.-B. Yu, and K. M. Zurek, *Phys. Rev. D* **87**, 115007 (2013).
  - [18] J. L. Feng, H. Tu, and H.-B. Yu, *J. Cosmol. Astropart. Phys.* **10** (2008) 043.
  - [19] C. Boehm and P. Fayet, *Nucl. Phys.* **B683**, 219 (2004).
  - [20] T. Lin, H.-B. Yu, and K. M. Zurek, *Phys. Rev. D* **85**, 063503 (2012).
  - [21] H. Davoudiasl and I. M. Lewis, *arXiv:1309.6640*.
  - [22] N. Arkani-Hamed, D. P. Finkbeiner, T. R. Slatyer, and N. Weiner, *Phys. Rev. D* **79**, 015014 (2009).
  - [23] M. Pospelov and A. Ritz, *Phys. Lett. B* **671**, 391 (2009).
  - [24] J. L. Feng and J. Kumar, *Phys. Rev. Lett.* **101**, 231301 (2008).
  - [25] C. Cheung, J. T. Ruderman, L.-T. Wang, and I. Yavin, *Phys. Rev. D* **80**, 035008 (2009).
  - [26] A. Katz and R. Sundrum, *J. High Energy Phys.* **06** (2009) 003.
  - [27] D. E. Morrissey, D. Poland, and K. M. Zurek, *J. High Energy Phys.* **07** (2009) 050.
  - [28] S. Andreas, M. Goodsell, and A. Ringwald, *Phys. Rev. D* **87**, 025007 (2013).
  - [29] R. Essig, P. Schuster, and N. Toro, *Phys. Rev. D* **80**, 015003 (2009).

- [30] J. D. Bjorken, R. Essig, P. Schuster, and N. Toro, *Phys. Rev. D* **80**, 075018 (2009).
- [31] B. Batell, M. Pospelov, and A. Ritz, *Phys. Rev. D* **80**, 095024 (2009).
- [32] G. Amelino-Camelia *et al.*, *Eur. Phys. J. C* **68**, 619 (2010).
- [33] H. Merkel *et al.* (A1 Collaboration), *Phys. Rev. Lett.* **106**, 251802 (2011).
- [34] S. Abrahamyan *et al.* (APEX Collaboration), *Phys. Rev. Lett.* **107**, 191804 (2011).
- [35] S. Andreas, C. Niebuhr, and A. Ringwald, *Phys. Rev. D* **86**, 095019 (2012).
- [36] J. Balewski *et al.*, [arXiv:1307.4432](#).
- [37] E. Izaguirre, G. Krnjaic, P. Schuster, and N. Toro, *Phys. Rev. D* **88**, 114015 (2013).
- [38] R. Essig, J. Mardon, M. Papucci, T. Volansky, and Y.-M. Zhong, *J. High Energy Phys.* **11** (2013) 167.
- [39] A. R. Zentner, *Phys. Rev. D* **80**, 063501 (2009).
- [40] F.-Y. Cyr-Racine and K. Sigurdson, *Phys. Rev. D* **87**, 103515 (2013).
- [41] L. G. van den Aarssen, T. Bringmann, and C. Pfrommer, *Phys. Rev. Lett.* **109**, 231301 (2012).
- [42] W. A. Dawson, *Astrophys. J.* **772**, 131 (2013).
- [43] F. Kahlhoefer, K. Schmidt-Hoberg, M. T. Frandsen, and S. Sarkar, *Mon. Not. R. Astron. Soc.* **437**, 2865 (2013).
- [44] E. V. Linder, *J. Cosmol. Astropart. Phys.* **04** (2013) 031.
- [45] D. Harvey *et al.*, [arXiv:1310.1731](#).
- [46] C. Kouvaris, *Phys. Rev. Lett.* **108**, 191301 (2012).
- [47] J. Fan, A. Katz, L. Randall, and M. Reece, *Phys. Dark Univ.* **2**, 139 (2013).
- [48] J. Bramante, K. Fukushima, J. Kumar, and E. Stopnitzky, *Phys. Rev. D* **89**, 015010 (2014).
- [49] L. Pearce and A. Kusenko, *Phys. Rev. D* **87**, 123531 (2013).
- [50] I. Goldman, R. Mohapatra, S. Nussinov, D. Rosenbaum, and V. Teplitz, *Phys. Lett. B* **725**, 200 (2013).
- [51] E. Ma, I. Picek, and B. Radovic, *Phys. Lett. B* **726**, 744 (2013).
- [52] F.-Y. Cyr-Racine, R. de Putter, A. Raccanelli, and K. Sigurdson, [arXiv:1310.3278](#).
- [53] E. Aprile (XENON1T Collaboration), [arXiv:1206.6288](#).
- [54] D. Akerib *et al.* (LUX Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **704**, 111 (2013).
- [55] P. Brink (SuperCDMS Collaboration), *J. Low Temp. Phys.* **167**, 1093 (2012).
- [56] S. W. Randall, M. Markevitch, D. Clowe, A. H. Gonzalez, and M. Bradac, *Astrophys. J.* **679**, 1173 (2008).
- [57] W. A. Dawson *et al.*, *Astrophys. J.* **747**, L42 (2012).
- [58] J. L. Feng, M. Kaplinghat, and H.-B. Yu, *Phys. Rev. Lett.* **104**, 151301 (2010).
- [59] R. Essig, J. Mardon, and T. Volansky, *Phys. Rev. D* **85**, 076007 (2012).
- [60] R. Essig, A. Manalaysay, J. Mardon, P. Sorensen, and T. Volansky, *Phys. Rev. Lett.* **109**, 021301 (2012).
- [61] A. Loeb and N. Weiner, *Phys. Rev. Lett.* **106**, 171302 (2011).
- [62] M. R. Buckley and P. J. Fox, *Phys. Rev. D* **81**, 083522 (2010).
- [63] B. Bellazzini, M. Cliche, and P. Tanedo, *Phys. Rev. D* **88**, 083506 (2013).
- [64] Z.-P. Liu, Y.-L. Wu, and Y.-F. Zhou, *Phys. Rev. D* **88**, 096008 (2013).
- [65] S. Tulin, H.-B. Yu, and K. M. Zurek, *Phys. Rev. Lett.* **110**, 111301 (2013).
- [66] H. Davoudiasl and R. N. Mohapatra, *New J. Phys.* **14**, 095011 (2012).
- [67] K. Petraki and R. R. Volkas, *Int. J. Mod. Phys. A* **28**, 1330028 (2013).
- [68] K. M. Zurek, [arXiv:1308.0338](#).
- [69] J. L. Feng, M. Kaplinghat, and H.-B. Yu, *Phys. Rev. D* **82**, 083525 (2010).
- [70] K. Jedamzik and M. Pospelov, *New J. Phys.* **11**, 105028 (2009).
- [71] E. W. Kolb and M. S. Turner, *Front. Phys.* **69**, 1 (1990).
- [72] G. Steigman, *Adv. High Energy Phys.* **2012**, 1 (2012).
- [73] L. Ackerman, M. R. Buckley, S. M. Carroll, and M. Kamionkowski, *Phys. Rev. D* **79**, 023519 (2009).
- [74] J. L. Feng, M. Kaplinghat, H. Tu, and H.-B. Yu, *J. Cosmol. Astropart. Phys.* **07** (2009) 004.
- [75] B. Holdom, *Phys. Lett. B* **166**, 196 (1986).
- [76] R. Foot, *Int. J. Mod. Phys. D* **13**, 2161 (2004).
- [77] D. Feldman, B. Kors, and P. Nath, *Phys. Rev. D* **75**, 023503 (2007).
- [78] K. Babu, C. F. Kolda, and J. March-Russell, *Phys. Rev. D* **57**, 6788 (1998).
- [79] H. Davoudiasl, H.-S. Lee, and W. J. Marciano, *Phys. Rev. D* **85**, 115019 (2012).
- [80] H. Davoudiasl, H.-S. Lee, I. Lewis, and W. J. Marciano, *Phys. Rev. D* **88**, 015022 (2013).
- [81] J. B. Dent, F. Ferrer, and L. M. Krauss, [arXiv:1201.2683](#).
- [82] M. T. Frandsen, F. Kahlhoefer, S. Sarkar, and K. Schmidt-Hoberg, *J. High Energy Phys.* **09** (2011) 128.
- [83] S. Chang, A. Pierce, and N. Weiner, *J. Cosmol. Astropart. Phys.* **01** (2010) 006.
- [84] B. Feldstein, A. L. Fitzpatrick, and E. Katz, *J. Cosmol. Astropart. Phys.* **01** (2010) 020.
- [85] N. Fornengo, P. Panci, and M. Regis, *Phys. Rev. D* **84**, 115002 (2011).
- [86] S. Galli, F. Iocco, G. Bertone, and A. Melchiorri, *Phys. Rev. D* **80**, 023505 (2009).
- [87] T. R. Slatyer, N. Padmanabhan, and D. P. Finkbeiner, *Phys. Rev. D* **80**, 043526 (2009).
- [88] L. Lopez-Honorez, O. Mena, S. Palomares-Ruiz, and A. C. Vincent, *J. Cosmol. Astropart. Phys.* **07** (2013) 046.
- [89] M. S. Madhavacheril, N. Sehgal, and T. R. Slatyer, [arXiv:1310.3815](#).
- [90] M. Kaplinghat, T. Linden and H.-B. Yu (to be published).
- [91] B. Patt and F. Wilczek, [arXiv:hep-ph/0605188](#).
- [92] J. March-Russell, S. M. West, D. Cumberbatch, and D. Hooper, *J. High Energy Phys.* **07** (2008) 058.
- [93] M. Ahlers, J. Jaeckel, J. Redondo, and A. Ringwald, *Phys. Rev. D* **78**, 075005 (2008).
- [94] S. Andreas, T. Hambye, and M. H. Tytgat, *J. Cosmol. Astropart. Phys.* **10** (2008) 034.
- [95] C. Arina, F.-X. Josse-Michaux, and N. Sahu, *Phys. Lett. B* **691**, 219 (2010).
- [96] X. Chu, T. Hambye, and M. H. Tytgat, *J. Cosmol. Astropart. Phys.* **05** (2012) 034.
- [97] A. Djouadi, O. Lebedev, Y. Mambrini, and J. Quevillon, *Phys. Lett. B* **709**, 65 (2012).
- [98] B. Bhattacharjee, S. Matsumoto, S. Mukhopadhyay, and M. M. Nojiri, *J. High Energy Phys.* **10** (2013) 032.

- [99] A. Greljo, J. Julio, J. F. Kamenik, C. Smith, and J. Zupan, *J. High Energy Phys.* **11** (2013) 190.
- [100] L. Bian, R. Ding, and B. Zhu, [arXiv:1308.3851](#).
- [101] S. Choi, C. Englert, and P. Zerwas, [arXiv:1308.5784](#).
- [102] E. Aprile *et al.* (XENON100 Collaboration), *Phys. Rev. Lett.* **109**, 181301 (2012).
- [103] J. Richard *et al.*, [arXiv:0911.3302](#) [*Mon. Not. R. Astron. Soc.* (to be published)].
- [104] M. Kaplinghat, R. E. Keeley, T. Linden, and H.-B. Yu, [arXiv:1311.6524](#).
- [105] J. Sander *et al.*, *AIP Conf. Proc.* **1534**, 129 (2013).
- [106] R. Foot, *Phys. Lett. B* **728**, 45 (2014).