Kinetic decoupling and small-scale structure in effective theories of dark matter

Jonathan M. Cornell, Stefano Profumo, and William Shepherd

Kinetic Decoupling and Small-Scale Structure
in Effective Theories of Dark Matter

Jonathan M. Cornell,[1,2,3,∗] Stefano Profumo,[2,3,†] and William Shepherd[2,3,‡]

[1]*Oskar Klein Centre for Cosmoparticle Physics, Department of Physics,*

Stockholm University, SE-106 91 Stockholm, Sweden

[2]*Department of Physics, University of California,*

1156 High St., Santa Cruz, CA 95064, USA

[3]*Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA*

(Dated: July 9, 2013)

Abstract

The size of the smallest dark matter collapsed structures, or protohalos, is set by the temperature at which dark matter particles fall out of kinetic equilibrium. The process of kinetic decoupling involves elastic scattering of dark matter off of Standard Model particles in the early universe, and the relevant cross section is thus closely related to the cross section for dark matter scattering off of nuclei (direct detection) but also, via crossing symmetries, for dark matter pair production at colliders and for pair-annihilation. In this study, we employ an effective field theoretic approach to calculate constraints on the kinetic decoupling temperature, and thus on the size of the smallest protohalos, from a variety of direct, indirect and collider probes of particle dark matter.

[∗]Electronic address: `jcornell@ucsc.edu`

[†]Electronic address: `profumo@ucsc.edu`

[‡]Electronic address: `wshepher@ucsc.edu`

I. INTRODUCTION

In the paradigm of weakly interacting massive particles (WIMPs) as dark matter candidates, the abundance of dark matter observed in the Universe stems from thermal decoupling of the dark matter particles in the early universe. This process involves the pair-annihilation and pair-creation of WIMPs going out of *chemical* equilibrium, with the resulting number density freezing out and remaining approximately constant per comoving volume to the present age. WIMP models possess the right range of masses and pair-annihilation/creation cross sections to produce a thermal relic density in the same ballpark as the observed dark matter density, a feat often dubbed the "WIMP miracle" [1].

After chemical decoupling, WIMPs do not cease to interact with the surrounding thermal bath. It is simply their number density which is no longer affected by particle-number-changing processes. WIMPs (χ) continue to scatter off of (Standard Model) particles in the thermal bath (f), thus remaining in *kinetic* equilibrium with the thermal bath, up until the relevant elastic processes ($\chi f \leftrightarrow \chi f$) go out of equilibrium, i.e. the rate for such processes falls below the Hubble expansion rate. At that point, WIMPs completely decouple from the thermal bath, free-streaming and slowing down as the Universe keeps expanding. To a first approximation, this is the age when the first gravitationally collapsed dark matter structures form, with typical size on the same order as an Hubble length at that epoch. WIMP *kinetic* decoupling thus sets the small-scale cut-off to the dark matter power spectrum (for a recent review see e.g. Ref. [2]).

Given a WIMP model, it is thus in principle a well-posed question to ask what the small-scale cutoff to dark matter halos (which we hereafter refer to with the symbol M_{cut}) is. The cut-off scale is an important quantity in cosmology: if large enough, it could affect significantly how many "visible" small-scale structures, such as dwarf galaxies, form, perhaps being relevant to the question of the "missing satellite problem" [3] or to other issues associated with small scales in cold dark matter cosmologies [4]. In principle, the small-scale cutoff sets the size of the most numerous dark matter "mini-halos", or protohalos, which might be detectable today either with direct [5] or indirect [6] dark matter search experiments. Finally, the cutoff scale is highly relevant to the question of the so-called "boost factor", as it literally sets the integration cutoff in the calculation of this factor (in practice, the enhancement to the annihilation rate from a given dark matter halo from sub-structure

within the halo).

The calculation of the kinetic decoupling temperature $T_{\rm kd}$, and thus of the small-scale cutoff $M_{\rm cut}$ has been carried out in a variety of model-dependent contexts, including supersymmetry [7–9], universal extra-dimensions [9, 10], and models with Sommerfeld enhancement [11, 12]. It has become clear that WIMP models accommodate a broad variety of kinetic decoupling temperatures, with resulting cutoff scales ranging from $10^3\ M_\odot$ to much less than $10^{-6}\ M_\odot$ even only within the limited framework of the minimal supersymemtric extension of the Standard Model [9], where the symbol $M_\odot$ indicates the mass of the Sun. Particle physics details of the WIMP model affect in a highly model-dependent way the kinetic decoupling, producing a wide array of outcomes, but for many particle physics models there is still a decent correlation between certain experimentally accessible quantities such as the direct detection scattering cross section, as explored in Ref. [13], and $M_{\rm cut}$.

A possible model-independent route to evaluating ranges for the expected small-scale cutoff is to consider an effective theory description of interactions between Standard Model and dark matter particles, as pursued, recently, in Ref. [14, 15]. For example, assuming the dark matter is a spin 1/2 fermion, it is simple to write down the complete set of lowest dimensional operators that mediate such interactions. In turn, by assuming that only one single operator is dominating the relevant dark matter interactions, crossing symmetry allows to draw stringent constraints on the allowed effective energy scale associated with the operator, for example from direct dark matter detection or from collider searches. As a result we can robustly set *upper limits* to the size of the small-scale cutoff, for each class of operators, as a function of the relevant operator's effective energy scale. This upper limit is quite significant, as cosmologically relevant effects occur only for sufficiently large such cutoffs.

While rather sophisticated codes now exist to reliably calculate $T_{\rm kd}$ (see e.g. [2]), two potentially important ingredients have been only marginally studied thus far:

(i) scattering off of quarks only, for example in "lepto-phobic" theories with suppressed couplings to leptons (this was first partly addressed in Ref. [14]), and

(ii) the role of scattering off of pions, for the same class of theories, for kinetic decoupling temperatures below the QCD confinement phase transition.

In addition, a third aspect that remains entirely unexplored to date is (iii) the relevance of loop-mediated scattering off of leptons, again notably in leptophobic theories.

In the present study, in addition to the general program of setting upper limits to the small-scale cutoff in the context of the mentioned effective theory description of interactions between Standard Model and dark matter particles, we address in detail the three novel issues listed above. We show that for leptophobic theories there exists an interesting interplay between loop-mediated scattering off of leptons and scattering off of pions, and that the two effects are generically comparable. We find that for WIMP models that can be described to a good approximation by an effective operator belonging to the class we consider here, there are stringent upper limits on the cut-off scale to the matter power spectrum, typically on the order of 10^{-3} $M_\odot$. This scale hints at the fact that WIMP effective theories are not likely to have any impact on small-scale structure issues in cold dark matter cosmology. On the other hand, since the predicted protohalos are typically very small, sizable boost factors from substructure enhancements are a rather generic prediction of effective theories of dark matter.

The reminder of this paper is organized as follows: we outline the class of effective operators we consider in the following section II; we then discuss how we calculate the kinetic decoupling and how we estimate the size of the small-scale halo size cutoff in section III; section IV presents all of our results; and the final section V summarizes our findings and concludes.

II. CLASSIFICATION OF EFFECTIVE OPERATORS

The effective operator framework has been explored as a method for comparing experimental bounds coming from various types of experiments on dark matter couplings to Standard Model fields [16, 17]. Within this framework, one writes down higher-dimensional operators which couple dark matter to quarks, leptons, or Standard Model bosons, requiring that (i) the operator contain at least two dark matter particles to ensure stability, and that (ii) Standard Model gauge symmetries are respected. One operator from the list of possible operators is then assumed to be the dominant one for the physics being investigated, and its effects are explored assuming the other operators are suppressed and, thus, do not contribute to the observables in question. Each operator of interest is investigated separately in this way, and any interference effects from having multiple operators active simultaneously are assumed to be small. Generally these interference effects are equivalent to changing the

assumed chirality structure of the operator in question, e.g. interfering a vector and an axial operator with equal suppression scales is equivalent to considering an operator which only couples to one chirality.

The basic assumption of this parametrization of dark matter interactions is that dark matter is the only new field light enough to be kinematically relevant, and these operators are suppressed by a mass scale which is related to the expected mass of the additional particles which mediate the interactions in some more complete model underlying the effective theory. Within the region of parameter space where this assumption is valid, a given complete model can be mapped into the space of these operators by integrating out the additional heavy fields. This assumption is a fairly weak one for elastic scattering of dark matter off of Standard Model particles, where the momentum exchange is typically on the order of the MeV, but is a fairly strong assumption for LHC searches, where the center of mass energy of the created dark matter pair can be quite large compared to the dark matter mass. We therefore encourage caution when considering the collider bounds on these operators, but expect that the results for kinetic decoupling and the bounds arising from direct detection should be robust.

We also calculate the thermal relic density of WIMPs under the assumption that the same operator dominates dark matter interactions with Standard Model particles in the early as well as in the late universe. Of course, this is a rather strong assumption, as it entails for example the absence of processes such as coannihilation, the presence of thresholds or resonances that could exist at finite temperature but not in the late universe, and the absence of temperature-suppressed operators that might dominate the chemical freeze-out while being irrelevant at the later kinetic freeze-out. We note, however, that this assumption is largely equivalent to other assumptions discussed above, where it is presumed that dark matter is the only kinematically relevant new particle in the theory and that one operator is dominant in all of the observables being searched for.

We consider here a subset of all possible operators which conserve parity in addition to

the Standard Model gauge symmetries. The operators of interest are

$$\mathcal{O}_\mathcal{S} = \frac{m_f}{\Lambda_S^3} \bar{\chi}\chi \bar{f}f \tag{1}$$

$$\mathcal{O}_\mathcal{P} = \frac{m_f}{\Lambda_P^3} \bar{\chi}\gamma^5\chi \bar{f}\gamma^5 f \tag{2}$$

$$\mathcal{O}_\mathcal{V} = \frac{1}{\Lambda_V^2} \bar{\chi}\gamma^\mu\chi \bar{f}\gamma_\mu f \tag{3}$$

$$\mathcal{O}_\mathcal{A} = \frac{1}{\Lambda_V^2} \bar{\chi}\gamma^\mu\gamma^5\chi \bar{f}\gamma_\mu\gamma^5 f \tag{4}$$

$$\mathcal{O}_\mathcal{T} = \frac{m_f}{\Lambda_T^3} \bar{\chi}\sigma^{\mu\nu}\chi \bar{f}\sigma_{\mu\nu} f, \tag{5}$$

where Λ_I is the suppression scale for operator $\mathcal{O}_\mathcal{I}$. Note that the operators which are chirality-violating are assumed to be proportional to the fermion mass to preserve $SU(2)_L$ and avoid inducing large effects in low-energy flavor observables. The first four operator normalizations are standard within the effective dark matter literature, but previous searches for contact operators have not included the mass suppression for the tensor operator to better make contact with direct detection bounds. We choose to consider the theoretically better motivated normalization of the tensor operator which does include a quark mass suppression, as the operator is chirality-violating and thus would require an insertion of the Higgs field to respect the SM gauge symmetries. Previous analyses have considered the operator without a quark mass dependence to make better contact with direct detection searches, as the unsuppressed tensor induces a coupling to the spin of the quarks composing the nucleon minus the spin of the antiquarks in the nucleon, but it is not clear how a tensor operator with that normalization would be alligned with the mass basis of the quarks so well as to avoid inducing unacceptably large corrections to flavor observables. The choice to include the quark mass suppression of the tensor operator leaves us without collider and direct detection bounds to compare to, and therefore we will only plot the early universe curves for these operators.

For each operator, we specify which Standard Model fermions the dark matter particle couples to. Generically, leptons are the most significant contributors to keeping the dark matter in kinetic equilibrium with the Standard Model thermal bath, while many of the key experimental searches constrain primarily the couplings to quarks. We choose here to consider explicitly three cases, wherein the dark matter couples only to leptons, only to quarks, or to both with equal suppression scales. For cases including quark couplings we plot the strongest available experimental bounds from LHC searches and direct detection searches,

6

and in cases including lepton couplings we will additionally plot LEP search bounds. In the special case of the lepton-only vector operator we will in addition plot the direct detection bounds induced at one-loop order, as discussed in Ref. [17, 18].

III.   THE FORMATION OF PROTOHALOS

A.   Temperature of kinetic decoupling

To calculate the temperature of kinetic decoupling, we use the numerical routine described in Ref. [2], which has been integrated into the DarkSUSY code [19]. An effective WIMP temperature parameter is defined in the following form:

$$T_\chi \equiv \frac{2}{3} \left\langle \frac{\mathbf{p}^2}{2m_\chi} \right\rangle = \frac{1}{3m_\chi n_\chi} \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \mathbf{p}^2 f(\mathbf{p}). \tag{6}$$

In the equation above m_χ is the WIMP mass and n_χ is its number density. To determine the time evolution of this parameter, we consider the Boltzmann equation for a flat Friedmann-Robertson-Walker metric:

$$E(\partial_t - H\mathbf{p} \cdot \nabla_{\mathbf{p}})f = C[f]. \tag{7}$$

Here f is the WIMP phase space density, E and $\mathbf{p}$ are the comoving energy and 3-momentum respecitively, and H is the Hubble parameter. $C[f]$ is the collision term for a scattering process between a non-relativistic WIMP and a relativistic Standard Model scattering partner. This was shown in Ref. [2] to be of the form

$$C[f] = c(T)m_\chi^2 \left[m_\chi T \nabla_{\mathbf{p}}^2 + \mathbf{p} \cdot \nabla_{\mathbf{p}} + 3 \right] f(\mathbf{p}), \tag{8}$$

where

$$c(T) = \sum_i \frac{g_{\mathrm{SM}}}{6(2\pi)^3 m_\chi^4 T} \int dk \, k^5 \omega^{-1} \, g^\pm \left(1 \mp g^\pm \right) \overline{|\mathcal{M}|}^2 \Big|_{\substack{t=0 \\ s=m_\chi^2 + 2m_\chi\omega + m_\ell^2}}. \tag{9}$$

In Equation (9), the sum is taken over all possible Standard Model scattering partners, g_{SM} is the number of associated spin degrees of freedom, ω is the energy of the Standard Model particle and k its momentum, and $g^\pm$ is the distribution for Fermi or Bose statistics, $g^\pm(\omega) = (e^{\omega/T} \pm 1)^{-1}$. In all expressions above, the upper sign is for fermions and the lower is for bosons. $\overline{|\mathcal{M}|}^2$ represents the scattering amplitude squared, summed over final and averaged over initial spin states. Detailed calculations of $\overline{|\mathcal{M}|}^2$ for all relevant cases for our results are included in the appendices.

As kinetic decoupling can take place either before or after the QCD phase transition at $T_c \approx 170$ MeV, we need to consider carefully the effects of quark confinement on the above sum. At temperatures before $4\,T_c$, we follow the convention of Ref. [2], where the WIMPs scatter off leptons and, to be conservative, the three lightest quarks. After $4\,T_c$, we no longer consider scattering off quarks. We however extend the treatment of Ref. [2] by including scattering of the dark matter off pions after the QCD phase transitions for the cases in which this process occurs at leading order, i.e. for the scalar, Eq. (1), and vector, Eq. (3), operator cases. Also, it is important to note in the above expression for $c(T)$, the scattering amplitude is evaluated in the $t = 0$ limit, where t is the squared difference between the incoming and outgoing 4-momenta of a scattering particle. This limit is reasonable because the average momentum transfer in a scattering event between a relativistic particle and a heavy WIMP should be quite small. However, for the pseudoscalar case, Eq. (2), the scattering amplitude vanishes for forward scattering, so we need to consider the scattering amplitude when the momentum transfer is not zero. Ref. [14] introduced a method to average over all possible values of t, in which $c(T)$ now takes the form:

$$c(T) = \sum_i \frac{g_{\mathrm{SM}}}{6(2\pi)^3 m_\chi^4 T} \int dk\, k^5 \omega^{-1}\, g^\pm \left(1 \mp g^\pm\right) \frac{1}{(4k^2)^2} \int_{-4k^2}^{0} dt(-t) \left.\overline{|\mathcal{M}|}^2\right|_{s=m_\chi^2+2m_\chi\omega+m_\ell^2} .$$

(10)

Returning now to T_χ in Eq. (6), to find the differential equation which describes its evolution, we multiply Eq. 7 by $\mathbf{p}^2/E$ and integrate over $\mathbf{p}$ to find

$$\left(\partial_t + 5H\right) T_\chi = 2m_\chi c(T)\left(T - T_\chi\right).$$

(11)

The behavior of T_χ has two limiting cases: at high temperatures when $T_\chi = T$ and at low temperatures when T_χ changes only because of the expansion of the universe, i.e. $T_\chi \propto a^{-2}$, and the kinetic decoupling temperature is when there is a rapid change between these two behaviors. As described in Ref. [2], a code has been developed to numerically integrate Eq. (11) and find this transition temperature, and we use this routine to calculate T_{kd}.

B. Protohalo Size

There are two mechanisms which independently set a limit on the smallest possible protohalo mass, M_{cut}: (i) the free streaming of WIMPs after kinetic decoupling and (ii) the

coupling of the WIMP fluid to acoustic oscillations in the SM particle heat bath. In determining our limit on the protohalo mass, we use the outcome of these two processes giving the largest (hence dominant) M_{cut}.

1. Viscosity and Free Streaming

At kinetic decoupling, the decoupling of the WIMP fluid from the SM particle fluid leads to viscosity between the two fluids that cause density perturbations in the WIMP fluid to be damped out [7]. After T_{kd}, the WIMPs free stream from areas of high to low density, causing further damping of the perturbations. The net result of these processes is an exponential damping of the perturbations with a characteristic comoving wavenumber [20, 21]:

$$k_{\text{fs}} \approx \left(\frac{m}{T_{\text{kd}}}\right)^{1/2} \frac{a_{\text{eq}}/a_{\text{kd}}}{\ln(4a_{\text{eq}}/a_{\text{kd}})} \frac{a_{\text{eq}}}{a_0} H_{\text{eq}}. \tag{12}$$

In the Equation above the "eq" subscript signifies that the quantity should be evaluated at matter-radiation equality. To find the resulting mass cutoff from these effects, one just calculates the mass contained in a sphere of radius π/k_{fs}, i.e. [2]:

$$M_{\text{fs}} \approx \frac{4\pi}{3}\rho_\chi \left(\frac{\pi}{k_{\text{fs}}}\right)^3 = 2.9 \times 10^{-6} M_\odot \left(\frac{1 + \ln\left(g_{\text{eff}}^{1/4} T_{\text{kd}}/30 \text{ MeV}\right)/18.56}{(m_\chi/100 \text{ GeV})^{1/2} g_{\text{eff}}^{1/4} (T_{\text{kd}}/50 \text{ MeV})^{1/2}}\right)^3. \tag{13}$$

In the above equation g_{eff} is the number of effective degrees of freedom in the early universe evaluated at T_{kd}.

2. Acoustic Oscillations

It has also been noted that the density perturbations in the WIMP fluid, coupled to the SM particle fluid before T_{kd}, should oscillate with the acoustic oscillations in the heat bath. At kinetic decoupling, modes of oscillation with k values large enough that they have entered the horizon are damped out, while modes with k values corresponding to scales larger than the horizon size grow logarithmically [22, 23]. Therefore, the characteristic damping scale is just the size of the horizon at kinetic decoupling ($k_{\text{ao}} \approx \pi H_{\text{kd}}$), and there is another cutoff mass from this process of the form [2]:

$$M_{\text{ao}} \approx \frac{4\pi}{3}\frac{\rho_\chi}{H^3}\bigg|_{T=T_{\text{kd}}} = 3.4 \times 10^{-6} M_\odot \left(\frac{T_{\text{kd}} g_{\text{eff}}^{1/4}}{50 \text{ MeV}}\right)^{-3}. \tag{14}$$

9

IV. RESULTS

As discussed above, for each effective operator in Eq. (1-5) we consider three cases as far as the relevant Standard Model particle class the dark matter couples to:

1. Universal couplings to all SM fermions;

2. Couplings to leptons only;

3. Couplings to quarks only.

Each case presents distinct behaviors in the early as well as in the late universe, and leads to different constraints and conclusions for the effective cutoff scale. Leptonic couplings, when present, tend to dominate the process of kinetic decoupling, as a simple result of the fact that at the relevant temperatures leptons (especially electron/positron and neutrinos) are in a relativistic state and the number densities are not Boltzmann-suppressed. On the other hand, quark couplings lead to stronger bounds from colliders and direct detection. If lepton couplings are absent then the contributions of quark couplings to kinetic decoupling must be treated with care due to the QCD confinement phase transition. Before the phase transition there is a thermal bath of quarks and the calculation of the scattering rate proceeds analogously with that for the leptonic couplings, but after the phase transition pions are the dominant hadrons and the matrix element of the quark bilinear in the pion must be evaluated. In addition, loop-induced scattering off of leptons arises generically even for vanishing direct couplings to leptons. This effect, which has never been considered in this context before, competes with scattering off of pions, and becomes more and more relevant as pions become less and less abundant at decreasing temperatures due to Boltzmann suppression. We will discuss each operator's coupling to pions individually in presenting our results.

For each case we also present all relevant bounds on effective dark matter interactions from collider searches both at the LHC [24, 25] and at LEP [17]. These constraints are subject to the concerns described in section II regarding the possibility of probing additional particles at colliders due to the large center of mass energies involved. For all operators which lead to appreciable direct detection cross sections we also plot the current leading bounds from those experiments. For spin-independent scattering the current leading bounds come from

the Xenon 100 experiment [26], while for spin-dependent scattering they are set by the SIMPLE [27] and PICASSO [28] experiments.

For all plots we also present relic density constraints. The line on the plots corresponds to when $\Omega_\chi h^2 = 0.1189$, the best fit value quoted by the Planck collboration [29] when combining their CMB results, WMAP polarization results, high-ℓ CMB data from ground telescopes and baryon acoustic oscillation measurements. For all operators except the tensor case (which has no simple, tree-level UV completion) we use the micrOMEGAs code [30, 31] to calculate the relic density. This was checked analytically to correspond with setting the annihilation cross section to the appropriate value $\langle \sigma v \rangle \approx 3 \times 10^{-26}$ cm^3/s, and this analytical requirement was used to calculate the relic density requirement for the tensor operator case.

A. Scalar Operator

The scalar-type coupling of dark matter to SM fermions contributes to direct detection in the case of quark couplings, and has been constrained by collider searches for both quark and lepton couplings. The collider constraints are relatively weak in this case, however, because of the mass-suppression of this chirality-violating operator. While pair annihilation, direct detection, and scattering responsible for kinetic decoupling all have access to the heavier SM fermion generations, the initial state, for collider studies, is dominated by the lighter states, and therefore collider bounds are weakened relative to the other dark matter probes.

For the case of universal coupling to SM fermions through the scalar operator the results are presented in figure 1a. We note that for dark matter above approximately 10 GeV in mass the bounds from direct detection, which are strongest in that region, indicate that the kinetic decoupling temperature must be on the order of 1 GeV. The resulting cutoff scale for the smallest protohalos is on the order of the Earth mass (about 10^{-6} $M_\odot$) for WIMP masses above 10 GeV. The relic density matches the observed dark matter only for masses above 200 GeV. Models that possess the right thermal relic density have extremely suppressed cutoff scales, smaller than 10^{-9} $M_\odot$ (see the right panel of figure 1a).

For the case of lepton-only couplings, the only relevant bound on this operator is from LEP, and the resulting bound is weak enough to not significantly constrain the process of kinetic decoupling. The corresponding results are shown in figure 1b, which indicates that kinetic decoupling temperatures below 10 MeV are possible in this case, resulting in small-

11

FIG. 1: *Plots of contours of constant T_{kd} (left) and M_{cut} (right) for the case of a scalar operator interaction between WIMPs and SM particles. Shaded regions and the dashed curve represent regions of parameter space excluded by collider and direct detection results, while the solid gray curve represents the correct dark matter relic density.*
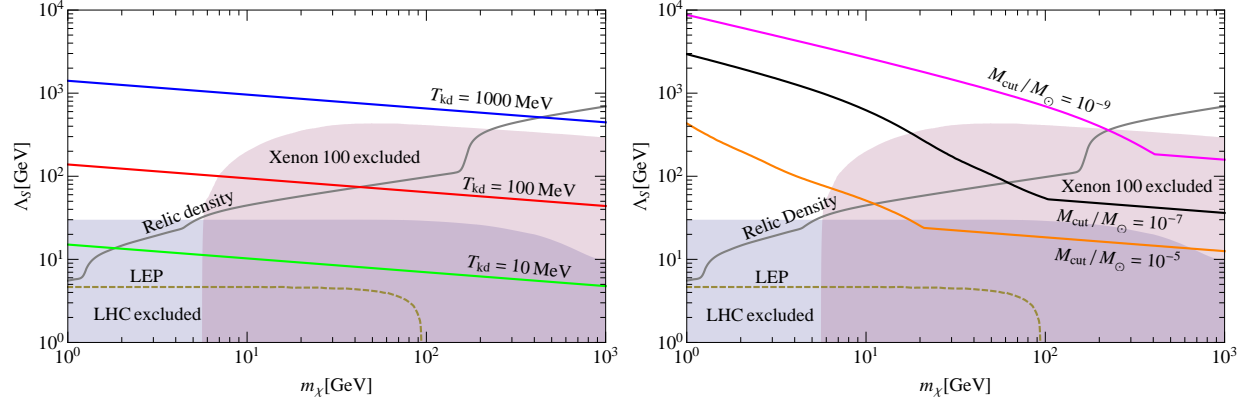
(a) *DM scatters off all SM fermions before the QCD phase transition and leptons and pions after.*
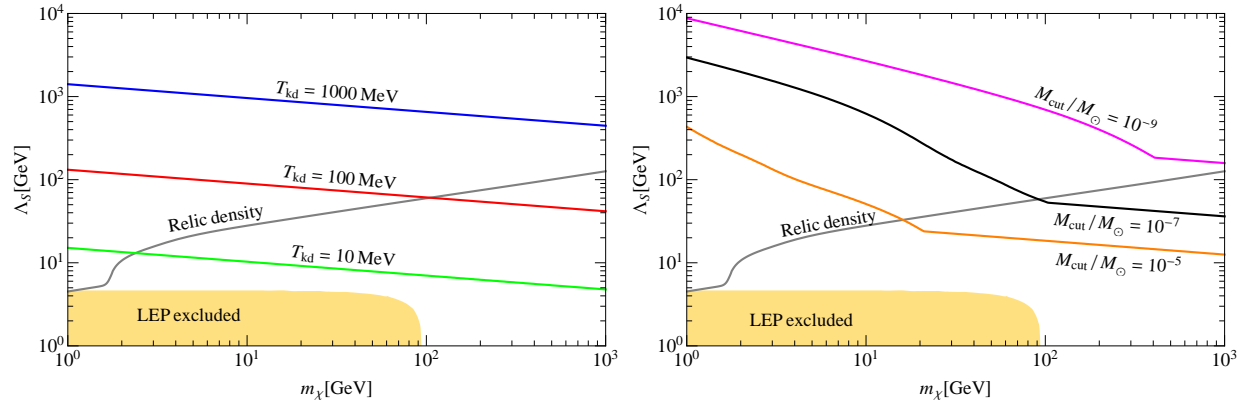
(b) *DM scatters off leptons only.*
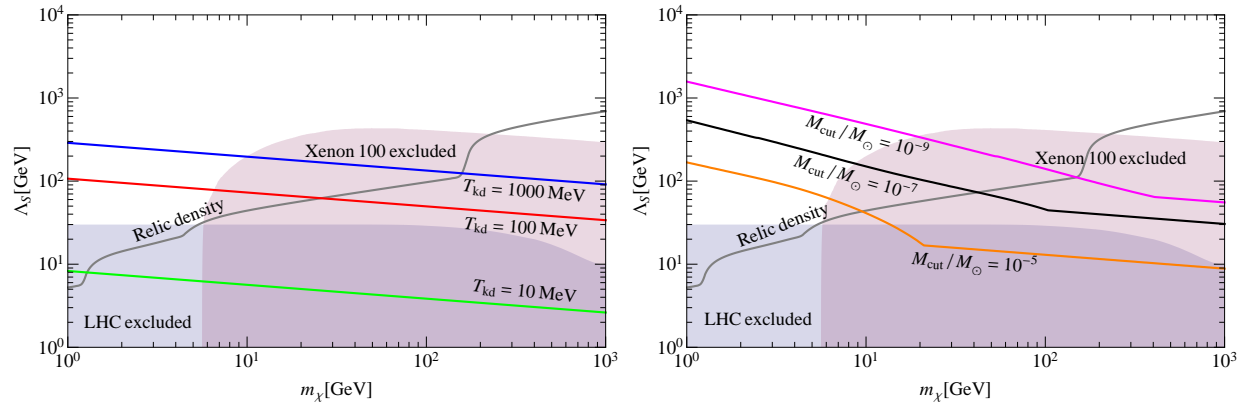
(c) *DM scatters off quarks before the QCD phase transition and pions after.*

12

scale cutoffs exceeding the Earth mass. We estimate in this case that the largest possible cutoff mass scale is of about $M_{\text{cut}} \sim 10^{-3} \, M_\odot$. We also note that the LEP limits do not impact the cutoff scales Λ_S needed to produce the correct thermal relic density. For WIMP masses of about 10 GeV we find that models that have the correct thermal relic density produce a small scale cutoff of $10^{-5} \, M_\odot$, while those with a mass of 100 GeV of about $10^{-7} \, M_\odot$ and those with a mass of 1 TeV of approximately $10^{-9} \, M_\odot$.

Finally, for quark-only couplings the matrix element $\langle \pi | \bar{q}q | \pi \rangle$, implicitly summed over quark flavors, is needed to evaluate the coupling to pions after the QCD phase transition. This has been evaluated previously in the context of contributions to direct detection by [32] using soft-pion techniques to be

$$\langle \pi^a | \bar{q}q | \pi^a \rangle = \frac{m_\pi^2}{2} \langle \pi^a | \vec{\pi} \cdot \vec{\pi} | \pi^a \rangle, \tag{15}$$

where $\vec{\pi}$ is a pion iso-vector, Eq. (B2). We have implemented this scattering amplitude for interactions after the QCD phase transition with pions, which are the dominant components of the thermal bath at the relevant temperatures. The results for quark-only couplings are shown in figure 1c. Direct detection forces in this case the size of the smallest protohalos to values well below $10^{-9} \, M_\odot$ for dark matter particle masses larger than 20 GeV. Models with the correct relic density must have masses above 200 GeV, and small scale cutoff smaller than $10^{-11} \, M_\odot$ in this case.

B. Vector Operator

Vector operators are generically better constrained by collider searches than scalar operators are, but are also very tightly constrained by direct dark matter detection. For universal couplings, direct detection is again the dominant constraint for dark matter masses above about 10 GeV, and those constraints again force us to conclude that the kinetic decoupling temperature must be of order 1 GeV. We show our results in figure 2. We do not find any sub-TeV WIMP models with a viable thermal relic density, if this operator is the only important contributor to chemical freeze-out. For masses above 100 GeV, we find that the cutoff scale is always smaller than approximately $10^{-9} \, M_\odot$.

Our results for vector interactions only with leptons are shown in figure 3a. We have plotted bounds from LEP and from direct detection, which arise at one-loop level by effec-
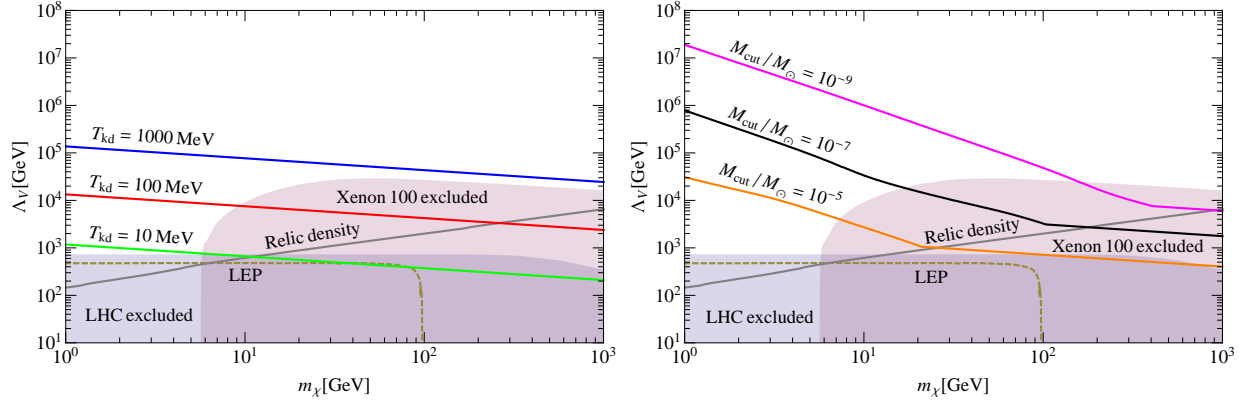
FIG. 2: *Plots of contours of constant T_{kd} (left) and M_{cut} (right) for the case of a vector operator interaction between WIMPs and SM fermions in which WIMPs couple with the same strength to all SM fermions. Shaded areas signify the regions of parameter space excluded by LHC and direct detection results, and the dashed line corresponds to a limit on the parameters from an analysis of LEP data. The solid gray curve represents the correct relic density.*

tively inducing a mixing between the integrated-out heavy vector boson and the SM photon. This mechanism was first discussed by Fox et. al. [17], and the bounds we plot are updates of those they derived from the first release of Xenon 100 data to take in to account the full 2012 data set. Even with a loop suppression, direct detection is still the dominant bound on dark matter models with masses above about 8 GeV, and the decoupling temperature is required to be of order 100 MeV. The resulting smallest possible protohalos are smaller than about 10^{-5} $M_\odot$ for WIMP masses below 100 GeV, and are generically of order 10^{-7} $M_\odot$ or smaller for masses above 100 GeV.

Considering couplings to quarks only below the QCD phase transition, we now must evaluate $\langle \pi^a | \bar{q} \gamma^\mu q | \pi^a \rangle$. This also was shown in [32], using the conservation of the vector current, to be

$$\langle \pi^a | \bar{q} \gamma^\mu q | \pi^a \rangle = (a_u - a_d) \langle \pi^a | \vec{\pi} \times \partial^\mu \vec{\pi} | \pi^a \rangle, \tag{16}$$

where a_q is the coupling to quarks of type q. This clearly vanishes for the coupling structure we have chosen of universal couplings to all quark flavors.

For a vector interaction coupling only to quarks, the induced direct detection amplitude in the former case can be effectively inverted to give an induced, loop-level coupling to leptons, which can be important as leptons are generically greater contributors to kinetic

FIG. 3: *Plots of contours of constant T_{kd} (left) and M_{cut} (right) for the case of a vector operator interaction between WIMPs and SM particles in which WIMPs couple directly to only leptons or quarks. Shaded regions represent regions of parameter space excluded by collider and direct detection results, while the solid gray curve represents the correct relic density.*
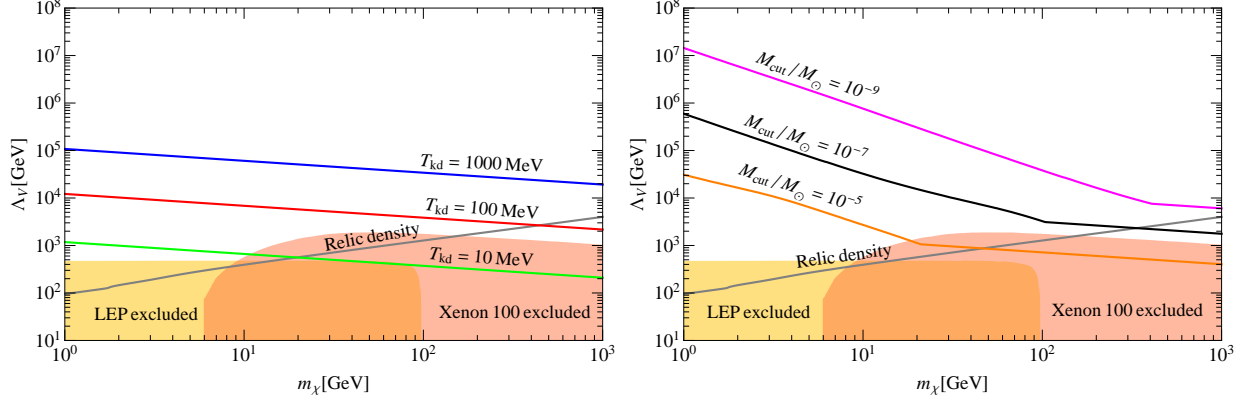
(a) *As in figure 2, but for DM scattering off leptons only.*
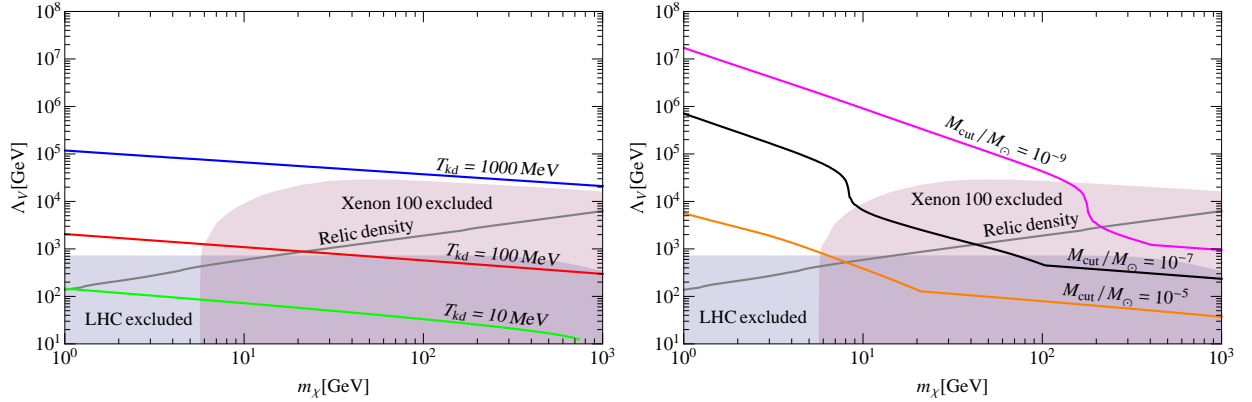
(b) *As in figure 2, but for DM coupling directly only to quarks and to leptons via a loop process.*
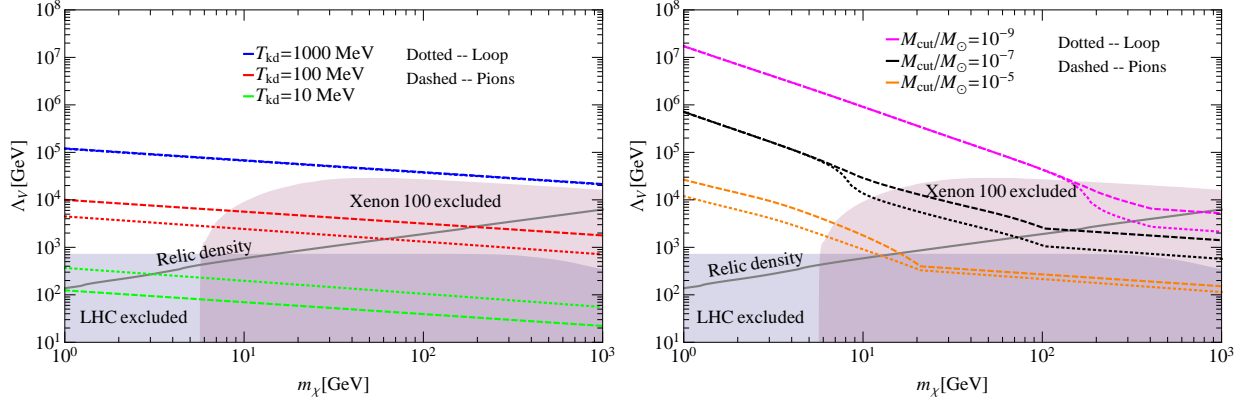
(c) *As in figure 2, but for DM coupling directly to quarks and to either leptons via a loop process (dotted lines) or to pions (dashed lines).*

15

decoupling than quarks. Since kinetic decoupling is dominated by scatterings at low dark matter velocities, the loop induced coupling to leptons from quarks can be considered as a simple rescaling of the suppression scale involved, and can be calculated in the "leading-log" approximation as discussed in Ref. [18]. The formula for this rescaling in the case where we consider identical couplings to all quark flavors is

$$\Lambda_\ell = \sqrt{\frac{3\pi}{2\alpha}} \frac{\Lambda_q}{\sqrt{\sum_{d,s,b} \ln\left(m_q/\Lambda_q\right) - 2\sum_{u,c,t} \ln\left(m_q/\Lambda_q\right)}}, \tag{17}$$

where α is the electromagnetic fine structure constant and all running quantities are evaluated at the renormalization scale of Λ_q, the scale of the effective operator. While this does not minimize the logarithms involved, it does allow us to neglect renormalization running and mixing of different operators, such that it is well-defined to assume one operator is dominant. We enforce perturbativity of this loop expansion by truncating our results when the induced coupling to leptons is not weaker than the initial coupling to quarks. The results for this coupling structure are presented in figure 3b. We find, as in fig. 2, that no models with sub-TeV masses have the right thermal relic density, and that the predicted cutoff for masses above 10 GeV is always smaller than $10^{-7}\ M_\odot$, while it is smaller than $10^{-9}\ M_\odot$ for masses above 100 GeV.

To explore the relative importance of pion scattering to that of loop-induced lepton scattering, we also consider a quarks-only vector-like operator which couples with opposite sign to up- and down-type quarks. This doesn't change the bounds from colliders or the scattering amplitudes above the QCD phase transition, but it allows for pion scattering below the QCD phase transition and alters the bounds from direct detection and the loop-induced coupling to leptons by changing the relative sign of the up- and down-type contributions in Eq. (17). We have presented the results for this coupling structure in figure 3c. This plot only shows results for including the coupling to pions or the loop coupling to leptons, but including both contributions leads to a curve that lies along the curve of larger suppression scale: e.g. for $T_{\rm kd} = 100\,{\rm MeV}$ the curve with both effects included would lie along the pion only curve. From this plot, we observe that following the QCD phase transition, pion scattering is the dominant process in setting $T_{\rm kd}$, but as T decreases, the loop coupling to leptons becomes more important as the pions become non-relativistic and their interaction rate with dark matter is Boltzmann suppressed. We thus note that the relative importance of scattering off of pions versus loop-mediated scattering off of leptons below the QCD phase

16

transition is generically comparable, with one process dominating over the other depending upon the kinetic decoupling temperature: for large decoupling temperatures, hence closer to the QCD confinement phase transition, scattering off of pions dominates, while lepton loop-induced scattering dominates as the number density of pions declines at lower temperatures.

C. Pseudoscalar Operator

Pseudoscalar operators present a unique complication among all parity-conserving operators, in that the scattering amplitude vanishes in the limit $t \to 0$ even when the center-of-mass velocity is large. This necessitates a summation over angles which can be neglected in the case of the other operators, as described in section III A.

Pseudoscalar operators lead to strongly suppressed direct detection scattering, so the only relevant bounds are from collider searches. Here, when the coupling is universal, the largest possible value for M_{cut} is $10^{-6} M_\odot$ when $m_\chi \geq 20 \, \text{MeV}$, as shown in figure 4a. Models with the correct relic density have masses of a few GeV and higher and increasingly suppressed cutoff scales as a function of mass: from $10^{-5} M_\odot$ at 5 GeV to $10^{-7} M_\odot$ at 20 GeV, and downward to $10^{-9} M_\odot$ and smaller for any mass larger than 200 GeV.

With lepton only couplings constrained by just LEP data, M_{cut} is again much less constrained, as the next figure, 4b, shows. Focusing again on models with the correct relic density, we find kinetic decoupling temperatures from slightly more than 10 MeV at WIMP masses in the GeV range, up to 1 GeV for 400 GeV WIMPs. The inferred cutoff mass scale ranges from $10^{-5} M_\odot$ at 6 GeV to $10^{-7} M_\odot$ at 30 GeV, to $10^{-9} M_\odot$ and smaller for any mass larger than 200 GeV, again for models with the correct thermal relic density.

For quark-only pseudoscalar couplings there exists a minimum value of T_{kd}, irrespective of how strongly the dark matter couples, which is the QCD phase transition temperature. After the phase transition the only hadronic state available with cosmologically-relevant number densities are the pions. Since QCD is a parity-conserving theory, we can require that the parity behavior of the quark bilinear match that of the pion state which the dark matter would couple to. However, with only two pions it is impossible to construct any pseudoscalar invariant. This indicates that elastic scattering off of two pions is completely forbidden by the symmetries of the theory for this operator. Other scattering processes are possible, however. Inelastic scatterings, whether changing the number of particles or changing, for

FIG. 4: *Plots of contours of constant T_{kd} (left) and M_{cut} (right) for the case of a pseudoscalar operator interaction between WIMPs and SM particles. Shaded regions and the dashed curve represent regions of parameter space excluded by collider results, while the solid gray curve represents the correct relic density.*
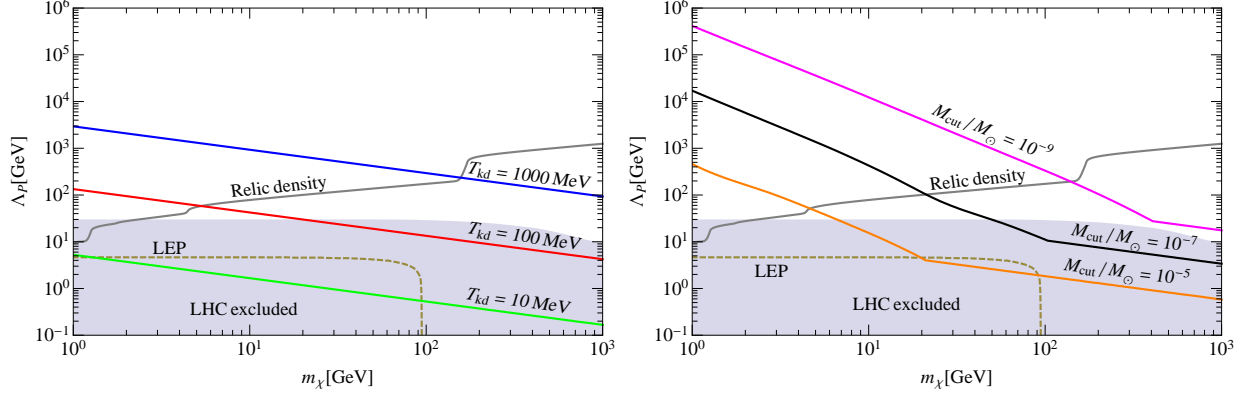
(a) *DM couples to all SM fermions.*
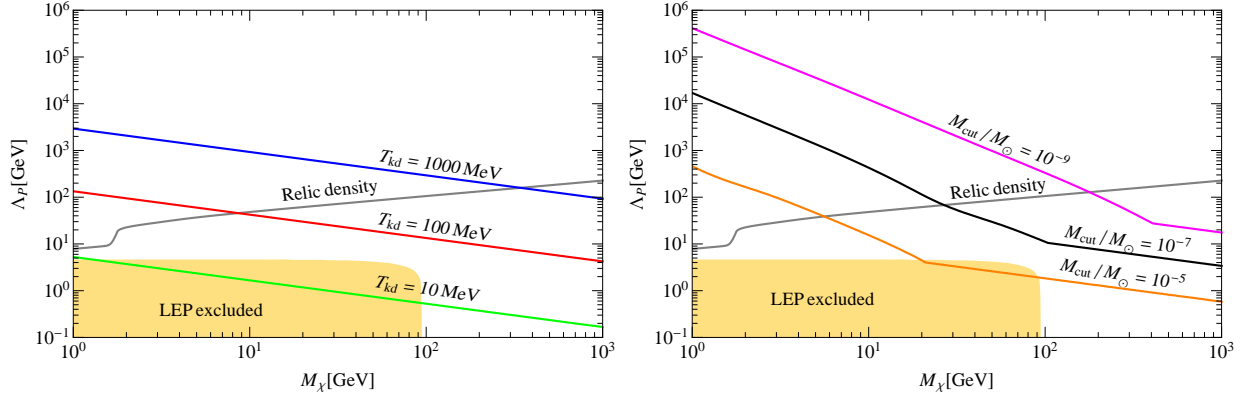
(b) *DM couples only to leptons.*

example, a pion into a sigma meson, are allowed by the symmetries of the problem. These nonetheless do not contribute efficiently to the continued thermalization of the dark matter kinematics, because the thermal bath is not energetic enough to produce the more exotic (i.e. higher-mass) QCD states or to provide sufficient energy to produce additional pions in scattering. Thus, the leading contribution is a one-loop-suppressed process requiring two insertions of the operator, which is very strongly suppressed.

18

D. Axial-Vector Operator

Axial-vector couplings are constrained at levels comparable to vector couplings by colliders, but lead to spin-dependent rather than -independent scattering in direct detection, so the collider bounds are generically stronger than the direct detection bounds for these interactions. Universal couplings to SM fermions are presented in figure 5a. We find that with universal couplings existing bounds generically require T_{kd} to be above 10 MeV, and the cutoff is smaller than 10^{-5} $M_\odot$ for any mass above 20 GeV. This class of operators produces the right thermal relic density for WIMPs above 100 GeV, leading in all cases to cutoff masses smaller than 10^{-5} $M_\odot$

For couplings to leptons only there are no appreciable direct detection bounds, as any loop-induced scattering akin to that in the vector case would have to proceed through Z-boson exchange, and the additional suppression of t/M_Z^2 makes such contributions negligible. Thus, only LEP bounds are shown along with our results in figure 5b. The figure indicates that cutoff scales as small as about 10^{-3} $M_\odot$ are in principle possible for very light WIMPs. The thermal relic density and LEP bounds put the dark matter mass in the 100 GeV and up range, with cutoff scales at most of 10^{-4} $M_\odot$, as before suppressed with increasing WIMP mass.

For the same reason that there are no bounds from direct detection on leptonic axial-vector couplings, there is no induced lepton coupling in the case of a quark-only interaction. Additionally, elastic scattering of dark matter off of pions vanishes in this model, as there is no axial invariant which can be constructed from the kinematics of two pions. Once again, inelastic scattering, whether producing or destroying an additional pion or scattering a pion into a different QCD state, is possible, but the low temperature below the QCD phase transition makes these possibilities contribute negligibly to the kinetic decoupling. Thus axial interactions with quarks only also have a minimum $T_{kd} = T_c$, analogously with the case of pseudoscalar couplings.

E. Tensor Operator

The tensor operator normalization which we consider preserves the SM gauge group where other normalizations do not, but is not particularly well studied because it does

FIG. 5: *Plots of contours of constant T_{kd} (left) and M_{cut} (right) for the case of a axial-vector operator interaction between WIMPs and SM particles. Shaded regions represent regions represent regions excluded by collider results and the non-solid curves represent regions of parameter space excluded by collider and direct detection results. The solid gray curve curve represents the correct relic density.*
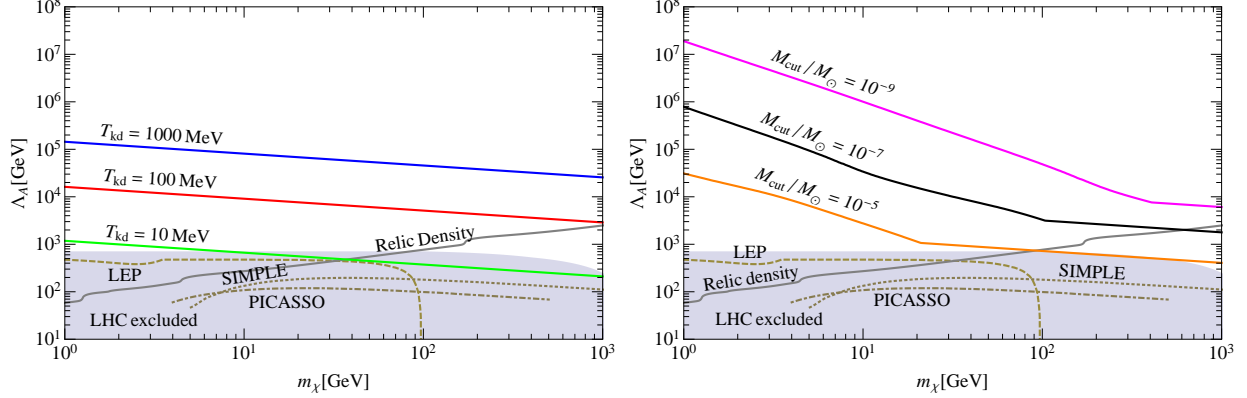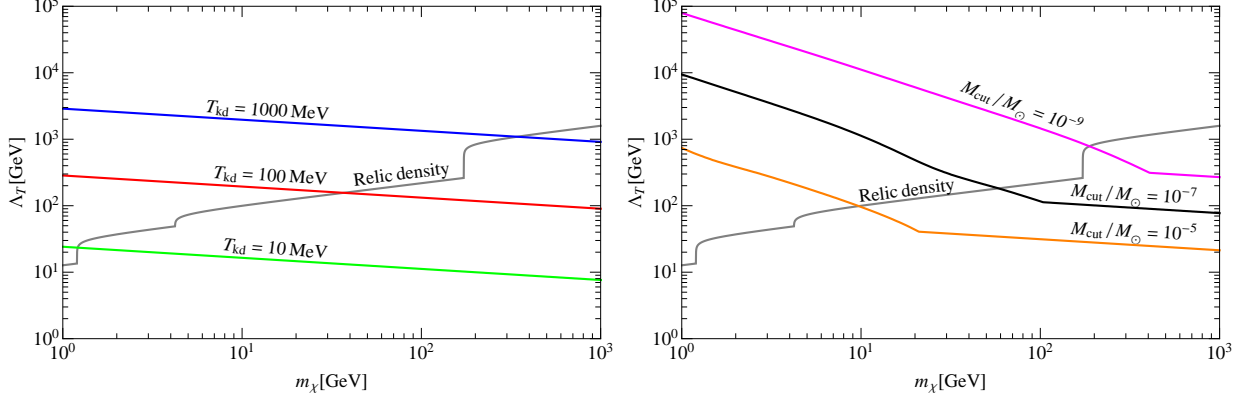

(a) *DM couples to all SM fermions.*
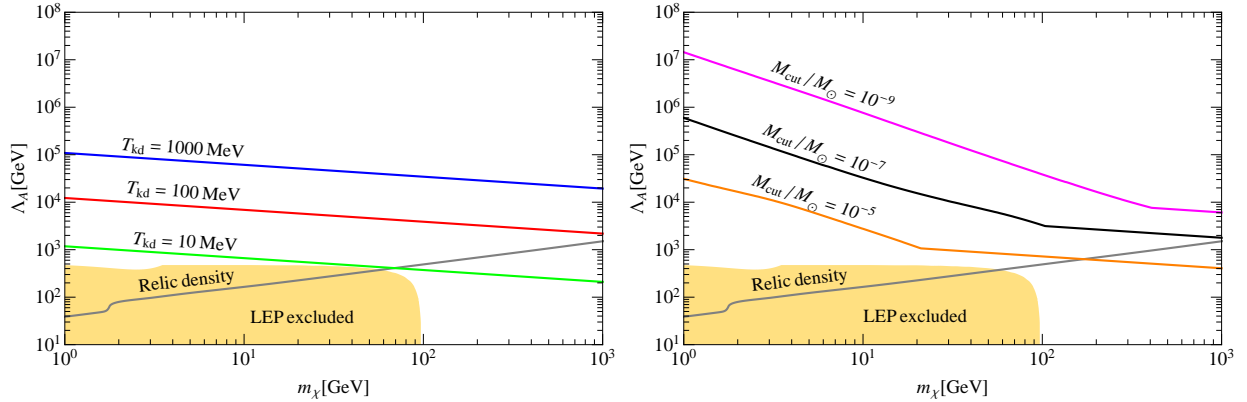
(b) *DM couples only to leptons.*

not correspond to the QCD matrix element which is probed in direct detection. Thus, we cannot present bounds from direct detection on this operator. Additionally, current collider searches have been normalized to correspond to direct detection, so we can't compare directly to those results either. The closest approximation to collider searches which could be considered would be the constraints on the other chirality-suppressed operators, in this paper the scalar and pseudoscalar cases. For the direct detection comparison the theoretical picture is a bit more muddled, as the quark mass which appears in this operator should

FIG. 6: *Plots of contours of constant* T_{kd} *(left) and* M_{cut} *(right) for the case of a tensor operator interaction between WIMPs and SM particles, where the solid gray curve represents the correct relic density.*

(a) *DM couples to all SM fermions.*
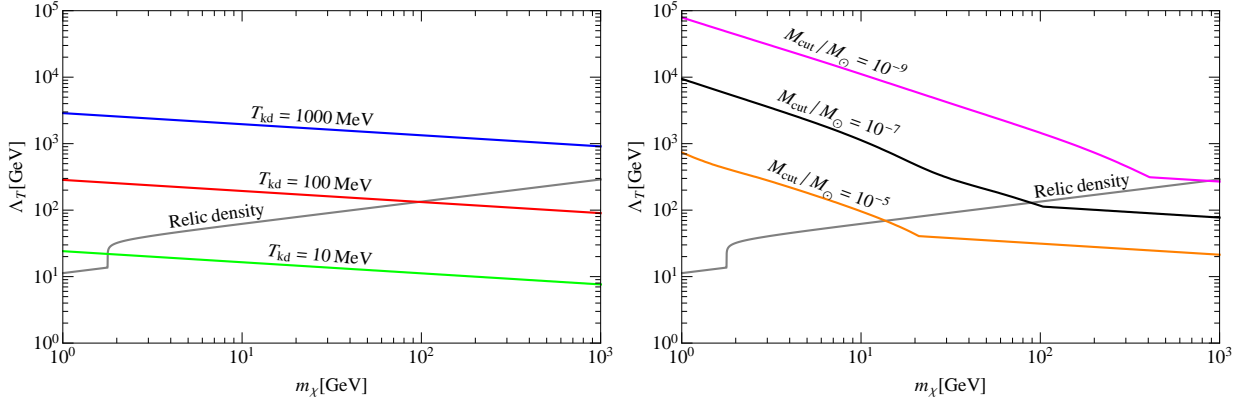
(b) *DM couples only to leptons.*

be taken to be related to the yukawa coupling, which will be affected by renormalization running in the strong phase of QCD, and is therefore nontrivial to factor out of the operator and find a meaningful bound. Since neither comparison technique yields a perfect mapping, we will only discuss the early-universe behavior of the operator.

The results for universal couplings are given in figure 6a, while those for leptons only are in figure 6b. As a tensor mediated interaction cannot be implemented in CalcHEP, we do not use micrOMEGAs to calculate the relic density, but rather require the velocity averaged cross section to equal the canonical value for s-wave annihilation that gives the correct relic

density, i.e

$$\langle \sigma v_{\rm rel} \rangle = \sum_f \frac{9}{2\pi} \frac{m_f^2 m_\chi^2}{\Lambda_T^6} \left(1 - \frac{m_f^2}{m_\chi^2} \right)^{1/2} \approx 3 \times 10^{-26} \frac{\rm cm^3}{\rm s}, \tag{18}$$

where the sum is over all kinematically accessible fermion annihilation products. Once again, after QCD confines in the early universe there is no pion configuration which has the Lorentz transformation properties of a tensor, and thus quark couplings become irrelevant below T_c.

Fig. 6a indicates that for good thermal relics, the expected kinetic decoupling temperatures are of 10 MeV in the few GeV mass range, up to 100 MeV for a 30 GeV WIMP, and to 1 GeV for a 300 GeV WIMP. The resulting small-scale cutoff masses are of 10^{-5} $M_\odot$ for a 10 GeV WIMP mass, decreasing to below 10^{-9} $M_\odot$ for masses above 200 GeV. For the exclusively leptonic coupling tensor case, we find a qualitatively similar behavior, with good thermal relics producing slightly lower kinetic decoupling temperatures, and slightly larger cutoff scales.

V. DISCUSSION AND CONCLUSIONS

In this study, we addressed the question of establishing the small-scale cutoff of the cosmological matter power spectrum in a variety of particle dark matter models where WIMP coupling to Standard Model fermions is described by effective operators. We included cases where the dark matter separately couples exclusively to leptons, exclusively to quarks, or universally to both leptons and quarks. We also used collider searches and dark matter direct detection to set model-independent limits on the largest experimentally viable value of the small-scale cutoff resulting from kinetic decoupling for each class of operators, and we calculated the dark matter thermal relic abundance on the same parameter space.

The largest possible cutoffs are found for theories where the dark matter exclusively couples to leptons, as a result of the absence of limits from hadron colliders. For the case of coupling to quarks, in some cases direct dark matter searches squeeze the maximal cutoff scale for protohalos to very small values, in some instances much smaller than the Earth mass. Insisting on setups that produce a thermal relic density in accordance with the observed dark matter density we universally find increasingly suppressed small scale cutoffs with increasing dark matter particle masses.

For theories with quark-only couplings, if the kinetic decoupling falls below the QCD

22

confinement phase transition, two effects exhibit an interesting interplay: scattering off of the lightest available hadronic bound states (pions) and loop-mediated scattering off of leptons, this latter process never having been considered before. We showed that depending on the operator's effective energy scale one or the other effect can dominate the kinetic decoupling process.

While there exist many instances of dark matter models where the effective operator description we adopted here does not apply, the present study has wide applicability to a broad range of WIMP models. In addition, our findings provide a model-independent framework where the relevant range for the small-scale cutoff to the matter power spectrum can effectively be predicted. Finally, this study highlights the complementarity of collider and direct detection of dark matter with questions pertaining to the cosmology of dark matter and the formation of structure in the universe.

Appendix A: Standard Model Fermion Scattering Matrix Elements

1.   Scalar

For the effective operator describing the scalar interaction between a SM fermion f and a Dirac fermion χ

$$\mathcal{O}_S = \frac{m_f}{\Lambda_S^3} \bar{\chi}\chi \bar{f}f \,, \tag{A1}$$

the matrix element for the scattering process between f and χ squared and summed over initial spin states and averaged over final spin states is of the form

$$\frac{1}{4} \sum_{\text{Spin States}} |\mathcal{M}|^2 = 4 \frac{m_f^2}{\Lambda_S^6} \left(\mathbf{p} \cdot \mathbf{p}' + m_\chi^2\right) \left(\mathbf{k} \cdot \mathbf{k}' + m_f^2\right) \,, \tag{A2}$$

23

where $\mathbf{p}$ and $\mathbf{p}'$ are the incoming and outgoing 4-momentum of the χ particle respectively and $\mathbf{f}$ and $\mathbf{f}'$ are the same for the SM fermion. Setting $t = 0$, this becomes

$$\frac{1}{4} \sum_{\text{Spin States}} |\mathcal{M}| = 16 \frac{m_f^4 m_\chi^2}{\Lambda_S^6} . \tag{A3}$$

2. Psuedoscalar

We now consider the effective operator describing pseudoscalar interactions,

$$\mathcal{O}_P = \frac{m_f}{\Lambda_P^3} \bar{\chi} \gamma^5 \chi \bar{f} \gamma^5 f . \tag{A4}$$

For a scattering process when the interaction is described by this operator, we find

$$\frac{1}{4} \sum_{\text{Spin States}} |\mathcal{M}|^2 = 4 \frac{m_f^2}{\Lambda_P^6} \left(m_\chi^2 - \mathbf{p} \cdot \mathbf{p}' \right) \left(m_f^2 - \mathbf{f} \cdot \mathbf{f}' \right) = \frac{m_f^2}{\Lambda_P^6} t^2 . \tag{A5}$$

3. Vector

Now considering the operator

$$\mathcal{O}_V = \frac{1}{\Lambda_V^2} \bar{\chi} \gamma^\mu \chi \bar{f} \gamma_\mu f , \tag{A6}$$

$$\frac{1}{4} \sum_{\text{Spin States}} |\mathcal{M}|^2 = \frac{8}{\Lambda_V^4} \left[(\mathbf{p} \cdot \mathbf{k}) (\mathbf{p}' \cdot \mathbf{k}') + (\mathbf{p} \cdot \mathbf{k}') (\mathbf{p}' \cdot \mathbf{k}) - (\mathbf{p} \cdot \mathbf{p}') m_f^2 - (\mathbf{k} \cdot \mathbf{k}') m_\chi^2 + 2m_\chi^2 m_f^2 \right] . \tag{A7}$$

As before, we consider only forward scattering, so $t = 0$. Working in the frame where the dark matter particle is stationary, $s = m_\chi^2 + 2m_\chi \omega + m_f^2$ and the matrix element becomes:

$$\frac{1}{4} \sum_{\text{Spin States}} |\mathcal{M}|^2 = 16 \frac{m_\chi^2}{\Lambda_V^4} \omega^2 . \tag{A8}$$

4. Pseudovector

The axial vector operator is of the form

$$\mathcal{O}_A = \frac{1}{\Lambda_V^2} \bar{\chi} \gamma^\mu \gamma^5 \chi \bar{f} \gamma_\mu \gamma^5 f , \tag{A9}$$

$$\frac{1}{4} \sum_{\text{Spin States}} |\mathcal{M}|^2 = \frac{8}{\Lambda_A^4} \left[(\mathbf{p} \cdot \mathbf{k}) (\mathbf{p}' \cdot \mathbf{k}') + (\mathbf{p} \cdot \mathbf{k}') (\mathbf{p}' \cdot \mathbf{k}) + (\mathbf{p} \cdot \mathbf{p}') m_f^2 + (\mathbf{k} \cdot \mathbf{k}') m_\chi^2 + 2m_\chi^2 m_f^2 \right] . \tag{A10}$$

Once again, taking the two limits $t = 0$ and $s = m_\chi^2 + 2m_\chi \omega + m_f^2$, this becomes

$$\frac{1}{4} \sum_{\text{Spin States}} |\mathcal{M}|^2 = 16 \frac{m_\chi^2}{\Lambda_A^4} \left(\omega^2 + 2m_f^2\right) . \tag{A11}$$

As $m_f \approx 0$ for relativistic fermions, this is essentially the same as the result for the vector operator.

5. Tensor

Finally, the tensor operator takes the form

$$\mathcal{O}_T = \frac{m_f}{\Lambda_T^3} \bar{\chi} \sigma^{\mu\nu} \chi \bar{f} \sigma_{\mu\nu} f \tag{A12}$$

where $\sigma^{\mu\nu} = (i/2)[\gamma^\mu, \gamma^\nu]$.

$$\frac{1}{4} \sum_{\text{Spin States}} |\mathcal{M}|^2 = 32 \frac{m_f^2}{\Lambda_T^6} \left(2 \left(\mathbf{p}' \cdot \mathbf{k}\right) \left(\mathbf{p} \cdot \mathbf{k}'\right) - \left(\mathbf{p} \cdot \mathbf{p}'\right) \left(\mathbf{k} \cdot \mathbf{k}'\right) + 2 \left(\mathbf{p} \cdot \mathbf{k}\right) \left(\mathbf{p}' \cdot \mathbf{k}'\right) + 3 m_\chi^2 m_f^2\right) ,$$
$$\tag{A13}$$

and then when $t = 0$ and $s = m_\chi^2 + 2m_\chi \omega + m_f^2$, this becomes

$$\frac{1}{4} \sum_{\text{Spin States}} |\mathcal{M}|^2 = 64 \frac{m_f^2 m_\chi^2}{\Lambda_T^6} \left(2\omega^2 + m_f^2\right) . \tag{A14}$$

Appendix B: Pion Scattering Matrix Elements

1. Scalar

For the scalar pion coupling we have a Lagrangian term [32]

$$\mathcal{L} \supset \frac{m_\pi^2}{2\Lambda_S^3} \bar{\chi} \chi \vec{\pi} \cdot \vec{\pi} , \tag{B1}$$

where

$$\vec{\pi} = \begin{pmatrix} \frac{1}{\sqrt{2}} \left(\pi^+ + \pi^-\right) \\ \frac{i}{\sqrt{2}} \left(\pi^+ - \pi^-\right) \\ \pi^0 \end{pmatrix} . \tag{B2}$$

Simplifying the dot product, this gives

$$\mathcal{L} \supset \frac{m_\pi^2}{2\Lambda^3} \bar{\chi} \chi \left(\pi^0 \pi^0 + 2\pi^+ \pi^-\right) . \tag{B3}$$

Note that this leads to a Feynman rule which is identical for all pion charges, and the scattering amplitude which we calculate is

$$i\mathcal{M} = \frac{m_\pi^2}{\Lambda_S^3} \bar{\chi}\chi \, . \tag{B4}$$

Squaring and averaging over initial spins, then choosing the zero relative velocity limit, gives the final result

$$\frac{1}{2} \sum_{\text{Spin States}} |\mathcal{M}|^2 = \frac{4m_\pi^4 m_\chi^2}{\Lambda_S^6} \, . \tag{B5}$$

2. Vector

The coupling from the vector operator has the form [32]

$$\mathcal{L} \supset \frac{2i}{\Lambda_V^2} \bar{\chi}\gamma_\mu\chi \left(\vec{\pi} \times \partial^\mu \vec{\pi}\right)_3 \tag{B6}$$

when we introduce a negative sign in front of the operator in Eq. A6 for down type quark interactions, as otherwise this term is zero. The relevant component of the cross product is $\pi_1 \partial^\mu \pi_2 - \pi_2 \partial^\mu \pi_1$, which can be rewritten in terms of the physical fields to give

$$\mathcal{L} \supset \frac{2i}{\Lambda_V^2} \bar{\chi}\gamma_\mu\chi \left(\pi^+ \partial^\mu \pi^- - \pi^- \partial^\mu \pi^+\right) \, . \tag{B7}$$

Thus, this operator does not couple to neutral pions, and the scattering amplitude off of a charged pion is equal to

$$i\mathcal{M} = \frac{2}{\Lambda_V^2} \bar{\chi}\gamma_\mu\chi \left(\mathbf{k} + \mathbf{k}'\right)^\mu \, . \tag{B8}$$

Squaring and averaging over incoming spins, we have

$$\frac{1}{2} \sum_{\text{Spin States}} |\mathcal{M}|^2 = \frac{8}{\Lambda_V^4} \left(\left(\mathbf{k} + \mathbf{k}'\right)^2 \left(m_\chi^2 - \mathbf{p}\cdot\mathbf{p}'\right) + 2\mathbf{p}\cdot\left(\mathbf{k} + \mathbf{k}'\right)\mathbf{p}'\cdot\left(\mathbf{k} + \mathbf{k}'\right)\right) \, . \tag{B9}$$

Simplifying this in terms of Mandelstam variables we find

$$\frac{1}{2} \sum_{\text{Spin States}} |\mathcal{M}|^2 = \frac{4}{\Lambda_V^4} \left(4m_\pi^2 t - t^2 + s^2 + u^2 - 2su\right) \, , \tag{B10}$$

and working in the limit where $t \to 0$, this becomes

$$\frac{1}{2} \sum_{\text{Spin States}} |\mathcal{M}|^2 = \frac{64 m_\chi^2 \omega^2}{\Lambda_V^4} \, . \tag{B11}$$

[1] Ya. B. Zeldovich, Adv. Astron. Astrophys. **3**, 241 (1965); H.Y. Chiu, Phys. Rev. Lett. **17**, 712 (1966); G. Steigman, Ann. Rev. Nucl. Part. Sci. **29**, 313 (1979); R.J. Scherrer and M.S. Turner, Phys. Rev. D **33**, 1585 (1986).

[2] T. Bringmann, New J. Phys. **11**, 105027 (2009). [arXiv:0903.0189 [astro-ph.CO]].

[3] L. E. Strigari, J. S. Bullock, M. Kaplinghat, J. Diemand, M. Kuhlen and P. Madau, Astrophys. J. **669**, 676 (2007) [arXiv:0704.1817 [astro-ph]].

[4] J. R. Primack, New J. Phys. **11**, 105029 (2009) [arXiv:0909.2247 [astro-ph.CO]].

[5] J. Diemand, M. Kuhlen, P. Madau, M. Zemp, B. Moore, D. Potter and J. Stadel, Nature **454**, 735 (2008) [arXiv:0805.1244 [astro-ph]]; M. Kuhlen, N. Weiner, J. Diemand, P. Madau, B. Moore, D. Potter, J. Stadel and M. Zemp, JCAP **1002**, 030 (2010) [arXiv:0912.2358 [astro-ph.GA]].

[6] J. Diemand, M. Kuhlen and P. Madau, Astrophys. J. **657**, 262 (2007) [astro-ph/0611370].

[7] S. Hofmann, D. J. Schwarz, H. Stoecker, Phys. Rev. **D64**, 083507 (2001). [astro-ph/0104173].

[8] T. Bringmann and S. Hofmann, JCAP **0407**, 016 (2007) [hep-ph/0612238].

[9] S. Profumo, K. Sigurdson, M. Kamionkowski, Phys. Rev. Lett. **97**, 031301 (2006). [astro-ph/0603373].

[10] D. Hooper and S. Profumo, Phys. Rept. **453**, 29 (2007) [hep-ph/0701197].

[11] L. G. van den Aarssen, T. Bringmann and Y. CGoedecke, Phys. Rev. D **85**, 123512 (2012) [arXiv:1202.5456 [hep-ph]].

[12] L. G. van den Aarssen, T. Bringmann and C. Pfrommer, Phys. Rev. Lett. **109**, 231301 (2012) [arXiv:1205.5809 [astro-ph.CO]].

[13] J. M. Cornell and S. Profumo, JCAP **1206**, 011 (2012) [arXiv:1203.1100 [hep-ph]].

[14] P. Gondolo, J. Hisano and K. Kadota, Phys. Rev. D **86**, 083523 (2012) [arXiv:1205.1914 [hep-ph]].

[15] I. M. Shoemaker, arXiv:1305.1936 [hep-ph].

[16] P. J. Fox, R. Harnik, R. Primulando and C. -T. Yu, Phys. Rev. D **86**, 015010 (2012) [arXiv:1203.1662 [hep-ph]]. A. Rajaraman, W. Shepherd, T. M. P. Tait and A. M. Wijangco, Phys. Rev. D **84**, 095013 (2011) [arXiv:1108.1196 [hep-ph]], P. J. Fox, R. Harnik, J. Kopp and Y. Tsai, Phys. Rev. D **85**, 056011 (2012) [arXiv:1109.4398 [hep-ph]]. J. Good-

man, M. Ibe, A. Rajaraman, W. Shepherd, T. M. P. Tait and H. -B. Yu, Phys. Rev. D **82**, 116010 (2010) [arXiv:1008.1783 [hep-ph]], Y. Bai, P. J. Fox and R. Harnik, JHEP **1012**, 048 (2010) [arXiv:1005.3797 [hep-ph]]. J. Goodman, M. Ibe, A. Rajaraman, W. Shepherd, T. M. P. Tait and H. -B. Yu, Phys. Lett. B **695**, 185 (2011) [arXiv:1005.1286 [hep-ph]].

[17] P. J. Fox, R. Harnik, J. Kopp and Y. Tsai, Phys. Rev. D **84**, 014028 (2011) [arXiv:1103.0240 [hep-ph]].

[18] J. Kopp, V. Niro, T. Schwetz and J. Zupan, Phys. Rev. D **80**, 083502 (2009) [arXiv:0907.3159 [hep-ph]].

[19] P. Gondolo, J. Edsjo, P. Ullio, L. Bergstrom, M. Schelke and E. A. Baltz, JCAP **0407**, 008 (2004) [astro-ph/0406204].

[20] A. M. Green, S. Hofmann, D. J. Schwarz, Mon. Not. Roy. Astron. Soc. **353**, L23 (2004). [astro-ph/0309621].

[21] A. M. Green, S. Hofmann, D. J. Schwarz, JCAP **0508**, 003 (2005). [arXiv:astro-ph/0503387 [astro-ph]].

[22] A. Loeb, M. Zaldarriaga, Phys. Rev. **D71**, 103520 (2005). [astro-ph/0504112].

[23] E. Bertschinger, Phys. Rev. **D74**, 063509 (2006). [astro-ph/0607319].

[24] ATLAS-CONF-2012-147.

[25] CMS-PAS-EXO-12-048.

[26] E. Aprile *et al.* [XENON100 Collaboration], Phys. Rev. Lett. **109**, 181301 (2012) [arXiv:1207.5988 [astro-ph.CO]].

[27] M. Felizardo, T. A. Girard, T. Morlat, A. C. Fernandes, A. R. Ramos, J. G. Marques, A. Kling and J. Puibasset *et al.*, Phys. Rev. Lett. **108**, 201302 (2012) [arXiv:1106.3014 [astro-ph.CO]].

[28] S. Archambault *et al.* [PICASSO Collaboration], Phys. Lett. B **711**, 153 (2012) [arXiv:1202.1240 [hep-ex]].

[29] P. A. R. Ade *et al.* [Planck Collaboration], arXiv:1303.5076 [astro-ph.CO].

[30] G. Belanger, F. Boudjema, A. Pukhov and A. Semenov, arXiv:1005.4133 [hep-ph].

[31] G. Belanger, F. Boudjema, A. Pukhov and A. Semenov, Comput. Phys. Commun. **176** (2007) 367 [arXiv:hep-ph/0607059]

[32] G. Prezeau, A. Kurylov, M. Kamionkowski and P. Vogel, Phys. Rev. Lett. **91**, 231301 (2003) [astro-ph/0309115].