Dark matter self-interactions from a general spin-0 mediator

Felix Kahlhoefer, Kai Schmidt-Hoberg and Sebastian Wild

DESY, Notkestraße 85, D-22607 Hamburg, Germany

E-mail: felix.kahlhoefer@desy.de, kai.schmidt-hoberg@desy.de, sebastian.wild@desy.de

ABSTRACT: Dark matter particles interacting via the exchange of very light spin-0 mediators can have large self-interaction rates and obtain their relic abundance from thermal freeze-out. At the same time, these models face strong bounds from direct and indirect probes of dark matter as well as a number of constraints on the properties of the mediator. We investigate whether these constraints can be consistent with having observable effects from dark matter self-interactions in astrophysical systems. For the case of a mediator with purely scalar couplings we point out the highly relevant impact of low-threshold direct detection experiments like CRESST-II, which essentially rule out the simplest realization of this model. These constraints can be significantly relaxed if the mediator has CP-violating couplings, but then the model faces strong constraints from CMB measurements, which can only be avoided in special regions of parameter space.

Contents

1 Introduction

While the existence of dark matter (DM) has only been inferred from its gravitational interactions, additional couplings are usually required to explain its observed relic abundance. Couplings to Standard Model (SM) states are strongly constrained by the combination of direct, indirect and collider searches for the DM particle. Interactions within the dark sector, in contrast, are much less constrained and DM could have significant self-interactions affecting its behaviour on astrophysical and cosmological scales. Indeed, one of the original motivations to consider self-interacting DM (SIDM) scenarios was the realisation that long-standing small-scale structure problems encountered in the collisionless cold DM paradigm may be successfully addressed [1, 2]. In order to have observable effects on astrophysical scales, the DM self-scattering cross section has to be sizeable, of order $\sigma/m_\psi \sim 1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ [3–8], where m_ψ is the DM mass.

It is furthermore desirable that DM self-interactions be enhanced for small relative DM velocities, in order to generate effects at small scales (such as dwarf galaxies) while at the

same time avoiding the strong constraints from systems with large velocities (such as galaxy clusters) [9–15]. Such a velocity dependence is most naturally achieved in models where a light particle ϕ (with $m_\phi \ll m_\psi$) mediates the DM interactions [3–6, 15–18]. In such a set-up, a new DM annihilation channel becomes available and thermal DM freeze-out can proceed via $\chi\chi \to \phi\phi$ [19]. This way the observed DM relic abundance can be obtained even for very small couplings of the mediator to the SM. Nevertheless, the mediator ought to couple to the SM at some degree in order to establish thermal contact between the dark sector and the SM and to allow for the mediator to decay after thermal DM freeze-out so that it does not come to dominate the energy density of the Universe [20–22].[1] In fact, in many models there are strong upper bounds on the lifetime of the mediator from Big Bang Nucleosynthesis (BBN), which impose a lower bound on the coupling of the mediator to SM states. Conversely, a tight upper bound on the coupling can be obtained from DM direct detection experiments, because the DM-nucleon scattering cross section is strongly enhanced for very light mediators.

Regarding specific realisations of this general framework, the main focus has been on vector and scalar mediators, which both give rise to a Yukawa potential in the non-relativistic limit and can therefore induce large DM self-interactions [4, 18, 23]. It was recently shown [24] that the case of a vector mediator, for which DM annihilation into mediators is an s-wave process, is strongly constrained by indirect detection experiments and Cosmic Microwave Background (CMB) measurements due to a large Sommerfeld enhancement [25, 26] and it is not possible to obtain sizeable self-interaction cross sections. These constraints do not apply to scalar mediators, for which s-wave annihilation is forbidden by CP conservation.

For scalar mediators, on the other hand, there is significant tension between direct detection experiments and constraints from BBN [20, 27]. The reason is that scalar mediators are expected to have couplings to SM fermions proportional to their mass and therefore interact much more strongly with nucleons than with electrons. It is thus difficult to achieve sufficiently small DM-nucleon scattering cross sections to satisfy bounds from direct detection experiments and at the same time ensure that the mediator decays before BBN into electrons or photons.

In this work we consider CP violation as a possible way to alleviate the tension between the various constraints for spin-0 mediators. The central observation is that both DM self-scattering and DM-nucleon scattering are substantially suppressed for pseudoscalar couplings. Allowing for different CP phases in the dark sector and in the SM sector therefore makes it possible to obtain large self-interactions while at the same time evading constraints from direct detection experiments. However, if CP is no longer conserved, DM annihilation can proceed via s-wave processes, reintroducing the strong constraints from indirect detection and CMB observations.

To investigate whether all constraints and requirements can be satisfied simultaneously, we perform a detailed study of the full parameter space for a fermionic DM particle

---

[1]If the dark sector is much colder than the visible sector, the mediator could also be stable. Another way to evade overclosure is if the mediator decays into an even lighter dark sector state. We do not consider these possibilities further in this work.

interacting via a spin-0 mediator with general CP phases. We revisit the case where the mediator has purely scalar couplings both to DM and to SM particles and point out the importance of recent results from the CRESST-II experiment [28, 29]. This experiment is sensitive to DM masses in the GeV region and rules out the most interesting remaining parameter region. We then turn to the case where the mediator has pseudoscalar couplings to SM particles but an arbitrary CP phase in the dark sector. We show that indirect detection and CMB constraints are so strong that the CP phase must be very close to either purely scalar or purely pseudoscalar couplings. In both of these cases it is possible to obtain large self-interaction cross sections consistent with all experimental and observational constraints.

The article is structured as follows. We introduce the model in section 2 and discuss annihilation rates, Sommerfeld enhancement and the relic density calculation. Section 3 is dedicated to the various experimental and observational constraints on the parameter space, including a detailed discussion of the DM momentum transfer cross section. Our results are presented in section 4, first for the case of purely scalar interactions and then for arbitrary CP phases in the dark sector. We summarize and conclude in section 5. Additional material is provided in the appendices A–C.

2 General set-up

We consider a scenario in which the DM particle is a Dirac fermion ψ coupled to a spin-0 particle ϕ with mass $m_\phi < m_\psi$. The crucial difference to most other studies of this set-up is that we do not make the assumption that the interactions of ϕ conserve CP (see also [27, 30–33]). Its interactions with DM can therefore be written as

$$\mathcal{L}_{\mathrm{DM}} \supset - y_\psi \cos\delta_\psi \ \bar{\psi}\psi\phi \, - \, i\, y_\psi \sin\delta_\psi \ \bar{\psi}\gamma^5\psi\phi \, . \tag{2.1}$$

This coupling structure is equivalent to $y_\psi e^{i\delta_\psi} \ \bar{\psi}_R\psi_L \, \phi +$ h.c., so δ_ψ parametrizes the CP-violating phase. Note that for $\delta_\psi = 0$ ($\delta_\psi = \pi/2$) the Lagrangian conserves CP provided ϕ is an even (odd) eigenstate under CP. In this case the Lagrangian reduces to the well-studied cases of DM interacting via a scalar or pseudoscalar mediator, respectively.

In addition to mediating the interactions of DM particles with each other, the ϕ boson can also couple to SM fermions. These interactions ensure that ϕ is unstable and furthermore provide a mechanism for bringing the DM particle into thermal equilibrium with the SM states. In order not to induce unacceptably large effects in precision measurements of SM flavour observables, we require that the couplings of ϕ to SM states are consistent with the hypothesis of minimal flavour violation [34], which implies that the couplings of ϕ to the various SM fermions f should be proportional to their respective masses m_f. Nevertheless, we do allow for CP violation also in the visible sector, i.e. we consider the following interaction Lagrangian:

$$\mathcal{L}_{\mathrm{mixing}} \, = \, - y_{\mathrm{SM}} \sum_f \left[\frac{m_f}{v_{\mathrm{EW}}} \cos\delta_{\mathrm{SM}} \ \bar{f}f\phi \, + \, i\frac{m_f}{v_{\mathrm{EW}}} \sin\delta_{\mathrm{SM}} \ \bar{f}\gamma^5 f\phi \right] . \tag{2.2}$$

Note that this interaction is only apparently renormalisable because it can only be induced after electroweak symmetry breaking [35]. The parameter y_{SM} must therefore be proportional to the electroweak vacuum expectation value v_{EW} divided by the scale Λ of some (unspecified) high energy mechanism. We will be interested only in the case where $y_{\text{SM}} \ll 1$, so that it is fully sufficient to consider the effective low-energy description given in eq. (2.2).

In the following, we will adopt a phenomenological point of view, meaning that we remain mostly agnostic about the origin of the CP-violating phases in eqs. (2.1) and (2.2) and analyse the model by treating these phases as free parameters. Nevertheless, there are strong constraints on new sources of CP violation in the SM, for example from measured upper bounds on the electric dipole moments of light leptons [36, 37] and nuclei [38]. These constraints are evaded as long as y_{SM} is sufficiently small and δ_{SM} is close to either 0 or $\pi/2$. To demonstrate that such a configuration can occur quite naturally, we discuss a toy model of spontaneous breaking of CP symmetry in appendix A. In this model, one finds that δ_ψ can take any value between 0 and $\pi/2$, while δ_{SM} is always very close to $\pi/2$ as a natural consequence of $y_{\text{SM}} \ll y_\psi$.

2.1 DM annihilation processes

In contrast to y_{SM}, the coupling y_ψ between the DM particle and ϕ can be large. Since we are interested in the case where ϕ is lighter than ψ, the cross section for a pair of DM particles to annihilate into ϕ bosons can therefore be sizeable. In the limit $m_\phi \ll m_\psi$ we obtain:[2]

$$(\sigma v)_{\psi\bar\psi \to \phi\phi} \simeq \frac{y_\psi^4 \sin^2(2\delta_\psi)}{32\pi m_\psi^2} + \frac{y_\psi^4 [3 + 8\cos(2\delta_\psi) + 7\cos(4\delta_\psi)]}{768\pi m_\psi^2} \cdot v^2 \, , \qquad (2.3)$$

where we have only kept the first two terms in an expansion in the relative velocity of the two DM particles v. We note that in the CP-conserving case, i.e. for $\delta_\psi = 0$ or $\delta_\psi = \pi/2$, the velocity-independent (s-wave) contribution vanishes. This is a consequence of the fact that in this case the annihilation process must proceed with non-zero angular momentum to ensure that initial and final state have the same CP quantum numbers. This argument does not apply for other values of δ_ψ, so that the annihilation cross section for $\psi\bar\psi \to \phi\phi$ is in general non-zero in the limit $v \to 0$.

Another process of interest is the annihilation into three (very light) ϕ bosons:

$$(\sigma v)_{\psi\bar\psi \to \phi\phi\phi} \simeq \frac{(7\pi^2 - 60)y_\psi^6 \sin^6 \delta_\psi}{1536\pi^3 m_\psi^2} \, . \qquad (2.4)$$

Despite the additional phase space and coupling suppression, this process turns out to contribute significantly to the total annihilation cross section in certain regions of the parameter space, as it is s-wave even in the case of pure pseudoscalar couplings.[3]

---

[2]We note that in the limit of a scalar mediator ($\delta_\psi = 0$) we obtain an expression that is a factor 2 smaller than the corresponding one in [18]. Our expression does however agree with [39] in this limit.

[3]For values of δ_ψ different from 0 or $\pi/2$ this process exhibits an infrared divergence. This can be

If the mediator ϕ is sufficiently light and sufficiently strongly coupled, the cross sections given above must be multiplied with a Sommerfeld enhancement factor [25], which reflects the modification of the initial-state wave function due to multiple mediator exchange: $(\sigma v)_{\mathrm{enh}} = S \times \sigma v$. Approximating the Yukawa potential by a Hulthén potential, one finds for an s-wave annihilation process [40–42]

$$S_s = \frac{\pi}{a} \frac{\sinh(2\pi\,a\,c)}{\cosh(2\pi\,a\,c) - \cos(2\pi\sqrt{c - a^2 c^2})} \ , \tag{2.5}$$

where $a = v/(2\alpha_\psi)$ and $c = 6\,\alpha_\psi\,m_\psi/(\pi^2 m_\phi)$ with $\alpha_\psi = y_\psi^2 \cos^2 \delta_\psi/(4\pi)$.[4] The corresponding expression for a p-wave process is

$$S_p = \frac{(c-1)^2 + 4\,a^2 c^2}{1 + 4\,a^2 c^2} \times S_s \ . \tag{2.6}$$

Crucially, the Sommerfeld factor depends on the relative DM velocity v. For $v \gtrsim \alpha_\psi$ one obtains $S_{s,p} \approx 1$, whereas for smaller velocities S increases proportionally to $1/v$ in the s-wave case and $1/v^3$ in the p-wave case, so that effectively the annihilation cross section in both cases increases proportionally to $1/v$ with decreasing velocity. The Sommerfeld enhancement saturates for $v \lesssim m_\phi/(2m_\psi)$, so the ratio of the two masses determines the maximum possible enhancement. Note that, if the model parameters are close to a resonance, the enhancement can be even larger and saturate even later.

The annihilation processes discussed above lead to two important effects. First of all, assuming that the dark sector is initially in thermal equilibrium with the SM sector, DM particles can obtain their relic abundance from thermal freeze-out in such a way that the DM relic abundance is determined by its annihilation cross section into light ϕ bosons. Second, since ϕ can decay into SM final states, DM annihilation processes can potentially be observed indirectly. For example, DM particles annihilating at the time of recombination can inject electromagnetic energy into the plasma and thereby spoil the successful predictions of the CMB radiation [43–48]. We will first discuss the calculation of the DM relic abundance and then return to constraints from indirect measurements in the following section.

2.2 Thermal freeze-out

The DM relic abundance depends on the thermally-averaged annihilation cross section $\langle\sigma v\rangle$ in such a way that larger cross sections correspond to smaller relic abundances. The simplest way to calculate the freeze-out prediction is to compare $\langle\sigma v\rangle$ to the so-called thermal

---

attributed to the initial state radiation of a massless scalar, which can lead to an on-shell intermediate DM particle. However, as the $2 \to 3$ annihilation is expected to be only relevant for $\delta_\psi \simeq \pi/2$ (as in this region of the parameter space the lowest-order annihilation into a pair of mediators is suppressed), it is sufficient to only take into account the contribution arising from the pseudoscalar coupling $y_\psi \sin\delta_\psi$, leading to the infrared-finite expression given in eq. (2.4).

[4]Note that we only use the scalar part of the coupling to calculate the Sommerfeld enhancement. We will return to this issue in more detail in the context of DM self-interactions in section 3 and in particular in appendix B.

cross section $\langle\sigma v\rangle_\text{thermal}$, which yields the observed DM abundance $\Omega_\text{DM}h^2 = 0.12$ [49]:

$$\Omega_\psi h^2 = 0.12 \times \frac{\langle\sigma v\rangle_\text{thermal}}{\langle\sigma v\rangle} \ . \tag{2.7}$$

The thermal cross section for a Dirac DM particle is approximately given by $\langle\sigma v\rangle_\text{thermal} \approx 4.4 \cdot 10^{-26}\,\text{cm}^3\,\text{s}^{-1}$ with a slight dependence on m_ψ, which we take from the detailed calculation in [50].

A more accurate analysis of the underlying Boltzmann equation reveals that the quantity that enters in the denominator of eq. (2.7) is not exactly $\langle\sigma v\rangle$ but actually

$$x_\text{f} \int_{x_\text{f}}^\infty \frac{\langle\sigma v\rangle}{x^2}\,\mathrm{d}x \ , \tag{2.8}$$

where $x = m_\chi/T$ parametrizes the temperature and its value at freeze-out lies in the range $20 < x_\text{f} < 30$. For example, if the annihilation cross section can be written as $\sigma v = a + b v^2$, one finds $\langle\sigma v\rangle = a + 6b/x$ and hence the denominator in eq. (2.7) should read $a + 3b/x_\text{f}$.

In our case, however, σv has a more complicated dependence on the velocity due to the onset of Sommerfeld enhancement for small velocities. Assuming $\sigma v = a'/v + b'/v$, the thermal averaging yields $\langle\sigma v\rangle = (a' + b')(\pi/x)^{-1/2}$. In practise, however, Sommerfeld effects are only beginning to become important during thermal freeze-out, so it is not in general clear which of the two expressions to use. We therefore adopt an approach that interpolates between the two cases and take

$$\langle\sigma v\rangle \approx (\sigma v)_\text{enh}\big|_{v=(\pi/x_\text{f})^{1/2}} \ . \tag{2.9}$$

This yields the correct result in the case that $(\sigma v)_\text{enh} \propto 1/v$, whereas for $S = 1$ one obtains $\langle\sigma v\rangle = a + \pi b/x_\text{f}$, which differs only slightly from the correct expression $a + 3b/x_\text{f}$ obtained analytically. We find this approximation to be sufficiently accurate given the uncertainty in x_f entering eq. (2.7).

To calculate the relic abundance of ψ we take the values of x_f provided in [50] for the s-wave case. This treatment ensures that our results are exactly correct in the case that freeze-out proceeds dominantly via s-wave processes and that Sommerfeld enhancement is not relevant, which are both good approximations in most of the parameter space that we consider. By applying our approach to the well-studied case of a vector mediator (see e.g. [18]), we confirm that our approximation is still very good in the case of s-wave annihilation with sizeable Sommerfeld enhancement. Furthermore, for dominantly p-wave annihilation, as for example in the case that $\delta_\psi = 0$, we have checked that our treatment agrees with more accurate studies in the literature including Sommerfeld enhancement [18] once we account for the uncertainty in x_f inherent in the analytical approach.

We can therefore now proceed and calculate the thermal abundance of the DM particle ψ using eq. (2.7) with $\langle\sigma v\rangle$ given in eq. (2.9). Since the annihilation cross section depends monotonically on the coupling y_ψ, we can determine the value of y_ψ that yields $\Omega h^2 = 0.12$ for each combination of m_ψ and m_ϕ. Note that this value is completely independent of how quickly ϕ decays into SM particles, and it is also independent of m_ϕ as long as $m_\phi \ll m_\psi$.
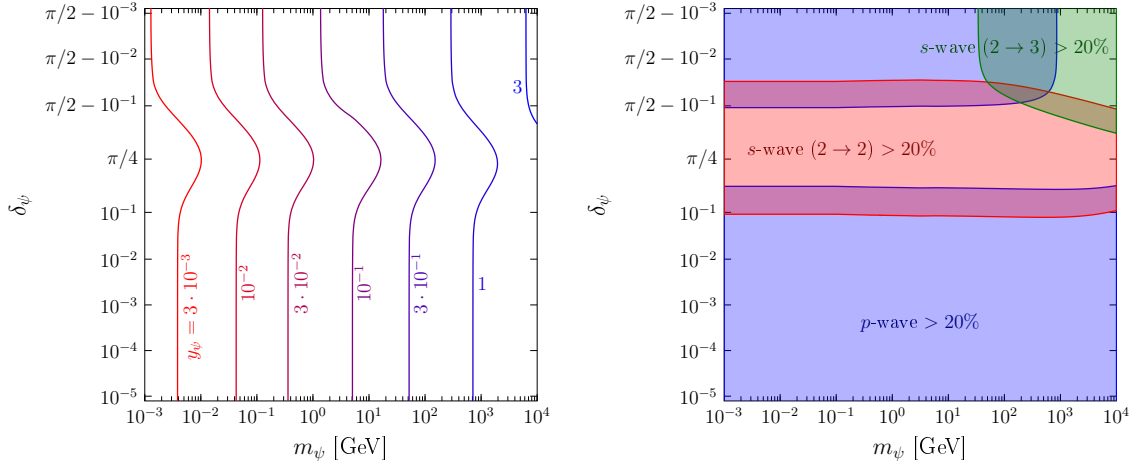
Figure 1. Values of y_ψ that give the correct relic abundance (left) and the relative importance of the different annihilation contributions (right).

The values obtained in this way are shown in the left panel of figure 1. One observes that larger couplings are required for larger m_ψ, as well as for δ_ψ close to either 0 or $\pi/2$. The right panel of figure 1 indicates in which regions of parameter space freeze-out is dominated by s-wave processes and where p-wave processes give the largest contribution. One can furthermore see the $2 \to 3$ process becoming relevant for large m_ψ and $\delta_\psi \approx \pi/2$. From now on we will always fix y_ψ to the value that reproduces the observed relic abundance and discuss the phenomenology of the remaining parameter space.

To conclude this section, let us add a number of comments on the relic density calculation. First of all, we note that for heavy DM masses and large couplings y_ψ the relic abundance calculation is modified by bound-state formation, as discussed in [39, 51, 52] for the case of scalar mediators and in [39, 53] for vector mediators. Since these effects are small for the parameter region that we are interested in, we do not discuss them further.

We also note that very close to a resonance there can be a second period of DM annihilation after kinetic decoupling [54–57], which can significantly affect the evolution of the DM abundance. As the kinetic decoupling temperature typically is in the MeV region [56], reannihilations are expected to occur around keV temperatures if the annihilation cross section scales as $1/v^2$ as is the case exactly on a resonance. In this case smaller dark sector couplings are sufficient to obtain the correct relic abundance, alleviating to some degree the constraints from direct and indirect detection. We do not consider detailed properties of these very tuned regions of parameter space further in this work.

Lastly, the expansion of (σv) in powers of v^2 given in eq. (2.3) breaks down for $m_\phi \simeq m_\psi$, requiring a more careful treatment of the dark matter freeze-out process [58]. However, as we will see below, in this case the dark matter self-interactions are in any case negligible. Hence, in the following we restrict ourselves to $m_\phi \lesssim 0.2 m_\psi$ which covers all the parameter space relevant to our study.

3 Experimental and observational constraints

We begin this section by reviewing the phenomenology of DM self-interactions in the context of the scenario that we consider. We will identify those regions of parameter space where self-interactions are so large that they are disfavoured by astrophysical observations as well as those regions where SIDM potentially provides a better explanation of astrophysical data than collisionless DM. We then continue by reviewing the various relevant experimental constraints on our model. These constraints can roughly be divided into constraints that are independent of the SM coupling y_{SM}, constraints that are independent of the DM coupling y_ψ and constraints that depend on both couplings. The first type of constraints comprises indirect detection experiments and CMB constraints while the second one consists of astrophysical constraints and constraints from searches for rare processes. Into the third category fall direct detection experiments and the requirement of thermalisation between the two sectors.

3.1 Dark matter self-interactions

DM particles interacting with each other via the exchange of a very light mediator may experience large rates of self-scattering. The resulting redistribution of momentum can reduce the central densities of DM halos (thus transforming cusps into cores), reduce halo ellipticity and even lead to the evaporation of sub-halos. The impact of DM self-interactions on astrophysical objects is quantified by the momentum transfer cross section σ_{T}, which is defined as[5]

$$\sigma_{\text{T}} = 2\pi \int_{-1}^{1} \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}(1 - |\cos\theta|)\,\mathrm{d}\cos\theta \ . \tag{3.1}$$

In a similar way as the annihilation processes discussed above, this quantity receives significant contributions from non-perturbative effects. These effects can be taken into account by solving the non-relativistic Schrödinger equation corresponding to the potential induced by the exchange of the mediator ϕ, which is given by the Fourier transformation of the matrix element $\mathcal{M}(\vec{q})$ for the scattering process:

$$V(r) = -\int \frac{\mathrm{d}^3 q}{(2\pi)^3}\, e^{i\vec{q}\cdot\vec{r}} \frac{\mathcal{M}(\vec{q})}{4m_\psi^2} \ . \tag{3.2}$$

For the model discussed in this work, both the contribution from the (CP-even) scalar coupling as well as from the (CP-odd) pseudoscalar coupling can in principle contribute to the matrix element $\mathcal{M}(\vec{q})$. The former contribution induces a well-known Yukawa potential of the form

$$V_S(r) = \alpha_S\, e^{-m_\phi r}/r \tag{3.3}$$

with $\alpha_S \equiv y_\psi^2 \cos^2\delta_\psi/(4\pi)$. The pseudoscalar coupling, on the other hand, induces a Yukawa potential with a strongly suppressed coupling strength $\propto m_\phi^2/m_\psi^2$, as well as further contributions to the potential scaling as $e^{-m_\phi r}/r^n$ with $n \geq 2$, which are of shorter range

---

[5]This expression generalizes equation (A8) from [13] to the case of particle anti-particle scattering, which was recently discussed in [59].

than the usual Yukawa potential induced by the exchange of a CP-even scalar. As discussed in detail in appendix B, we therefore expect the scalar couplings to be the dominant source of DM self-interactions and use the potential from eq. (3.3) to calculate the momentum transfer cross section. This means in particular that we do not expect any sizeable effects from DM self-interactions in the case that the mediator has purely pseudoscalar couplings to DM.

A common approach to calculate $\sigma_{\rm T}$ is to assume that the two interacting DM particles are distinguishable classical particles, which can interact with each other only via the t-channel exchange of a mediator [18, 59]. Under this assumption there is no need to differentiate between particle-particle and particle-antiparticle scattering and the differential cross section is independent of the total spin of the initial state. If the mediator is light and very weakly coupled, the differential cross section is then strongly peaked towards $\theta \to 0$. It is then a good approximation to replace the factor $1 - |\cos\theta|$ in eq. (3.1) by $1 - \cos\theta$, which significantly simplifies the integration. However, this approximation is no longer valid if the mediator is not very light, if resonances become important or if the interference between t-channel and u-channel mediator exchange are relevant. Appendix C discusses how all of these effects can be consistently included and demonstrates that the differences to the standard approach are non-negligible.

In the presence of non-perturbative effects, the momentum transfer cross section depends on the relative DM velocity $v_{\rm rel} \approx \sqrt{2}\, v_{\rm disp}$ in such a way that larger effects can be expected for systems with small DM velocity dispersion $v_{\rm disp}$ (such as dwarf galaxies) while constraints at larger velocities (e.g. from galaxy clusters) can be evaded. The strength of these constraints is a matter of ongoing debate (see e.g. [60, 61]), so we will adopt a rather conservative bound and require $\sigma_{\rm T}/m_\psi < 1\,{\rm cm}^2\,{\rm g}^{-1}$ for $v = 1000\,{\rm km\,s}^{-1}$. On the scale of dwarf galaxies ($v \approx 30\,{\rm km\,s}^{-1}$), on the other hand, the observationally interesting range of self-interaction cross sections is $0.1\,{\rm cm}^2\,{\rm g}^{-1} \lesssim \sigma_{\rm T}/m_\psi \lesssim 10\,{\rm cm}^2\,{\rm g}^{-1}$ [7, 8, 15, 20, 62]. We will compare this range to the various constraints that are discussed next.

3.2 Indirect detection experiments and CMB constraints

The first set of constraints stems from the same processes already discussed in section 2, namely the pair-annihilation of DM particles into mediators. Since these mediators are unstable, they will decay into SM final states, thus inducing potentially observable signatures. The Sommerfeld enhancement factor increases with decreasing DM velocity, so we can potentially expect strong constraints from indirect detection experiments probing $v/c \sim 10^{-4}$–10^{-3}. Even stronger constraints can result from CMB observations, which are sensitive to the DM annihilation rate at redshift $z \approx 1100$, where $v/c \lesssim 10^{-7}$ [24, 45, 46, 55, 56, 63–66].

Since we focus on the case $m_\phi < m_\psi$, only decays into SM final states are allowed and the branching ratios are independent of the magnitude of the couplings $y_{\rm SM}$ and y_ψ. The ratio of the different decay modes does however depend on the mass of the mediator. For $m_\phi < 2m_e$, only loop-induced decays into photons are kinematically accessible, while for $2m_e < m_\phi < 2m_\mu$ also decays into electrons play a relevant role. For larger mediator masses, both leptonic and hadronic decay modes can be relevant and uncertainties in the

theoretical predictions are rather large. We will return to the question which branching ratios to use below.

In our scenario, the strongest constraints on present-day annihilation of DM particles are obtained from the very precise measurements of the positron flux made by the AMS-02 experiment [67, 68]. We adopt the bounds derived in [69] for one-step cascade annihilations, including the $4e$ and 4μ final states. These bounds are independent of the mediator mass as long as $m_\phi \ll m_\chi$ and can be applied down to DM masses of 10 GeV.

To calculate CMB constraints, we assume that the Sommerfeld enhancement during recombination is fully saturated, which is a good approximation as long as the DM velocity during recombination satisfies $v_{\rm rec} \lesssim m_\phi/(2m_\psi)$ [24]. To calculate the effect on the CMB, we multiply the resulting annihilation cross section with the appropriate efficiency factors from [70]. The results are then compared with the most recent upper bound from Planck [49].

3.3 Bounds on the Standard Model coupling

It is clear that a new mediating particle with mass below the GeV scale is only phenomenologically viable if it couples very weakly to the SM. A review of the most relevant constraints can be found in [71] for the case of scalar couplings and in [31] for the case of pseudoscalar couplings (see also [72] for an experimental proposal to improve these constraints in the near future). The strongest constraints for mediator masses $m_\phi < 100\,{\rm MeV}$ typically come from searches for rare kaon decays. The reason is that W-boson loops induce effective flavour-changing interactions of the form

$$\mathcal{L}_{\rm FCNC} \supset h_{ds}^S \, \phi \, \bar{d}s + {\rm h.c.} \,, \tag{3.4}$$

where for $\delta_{\rm SM} = 0$ [73]

$$h_{ds}^S \approx \frac{3\alpha \, y_{\rm SM} \, m_s \, m_t^2}{32\pi \, m_W^2 \, \sin(\theta_W)^2 \, v_{\rm EW}} \, V_{ts} V_{td}^* \tag{3.5}$$

with m_W the W-boson mass, $\alpha \equiv e^2/(4\pi)$, θ_W the Weinberg angle and V the CKM matrix. For $\delta_{\rm SM} \neq 0$, the W-boson loop is divergent, indicating a sensitivity to the specific UV completion [74]. To estimate the magnitude of the expected effects, one can introduce a cut-off at a specific new-physics scale Λ. In this case, one obtains for $\delta_{\rm SM} = \pi/2$ [31]

$$h_{ds}^S \approx -\frac{\alpha \, y_{\rm SM} \, m_s \, m_t^2}{8\pi \, m_W^2 \, \sin(\theta_W)^2 \, v_{\rm EW}} \, V_{ts} V_{td}^* \, \log\left(\frac{\Lambda^2}{m_t^2}\right) \,. \tag{3.6}$$

These new interactions can induce new kaon decay modes [75]:

$$\Gamma(K^+ \to \pi^+ \phi) = \frac{1}{16\pi \, m_{K^+}^3} \lambda^{1/2}(m_{K^+}^2, m_{\pi^+}^2, m_\phi^2) \left(\frac{m_{K^+}^2 - m_{\pi^+}^2}{m_s - m_d}\right)^2 |h_{ds}^S|^2 \,, \tag{3.7}$$

where $\lambda(a, b, c) = (a - b - c)^2 - 4\,b\,c$ and we neglect a form factor $|f_0^{K^+}(m_\phi^2)|^2$, which is close to unity [76]. Since the light mediator is rather long-lived, it will typically escape from the

detector without decaying. Dedicated searches for this decay mode by experiments like E787 [77] and E949 [78, 79] place an upper bound of

$$\text{BR}(K^+ \to \pi^+ + \phi(\to \text{inv})) \lesssim 5 \cdot 10^{-11} . \tag{3.8}$$

This bound implies

$$y_{\text{SM}} \lesssim 1.9 \cdot 10^{-4} \qquad \text{for } \delta_{\text{SM}} = 0 , \tag{3.9}$$

$$y_{\text{SM}} \log \left(\Lambda^2 / m_t^2 \right) \lesssim 1.6 \cdot 10^{-4} \qquad \text{for } \delta_{\text{SM}} = \pi/2 \tag{3.10}$$

for $m_\phi \lesssim 50 \, \text{MeV}$. We note that the NA62 experiment [80] is expected to measure the branching ratio for $K^+ \to \pi^+ \bar{\nu} \nu$ with a precision of 10% and thereby significantly improve the bounds on new invisible decay modes.

At the same time, astrophysical and cosmological constraints can be used to place a lower bound on the coupling y_{SM}. Most importantly, if the mediators produced in the early Universe have a lifetime $\tau_\phi \gtrsim 1 \, \text{s}$, they will decay during or after BBN and can thereby potentially spoil the successful prediction of the abundance of the various elements. While it is challenging to accurately calculate the magnitude of these effects, it is clear that BBN constraints can be evaded if the mediator decays sufficiently quickly. As long as decays into leptons are kinematically allowed, a sufficiently short lifetime can always be achieved without conflicting with the upper bounds on y_{SM} mentioned before. For purely photonic decays, however, it is typically impossible for phenomenologically viable mediators to decay before BBN. A dedicated study of BBN constraints is imperative to determine whether such mediators can be consistent with the observed abundance of elements.

Strong constraints for sub-MeV mediators also come from astrophysical observations, such as the lifetime of horizontal branch stars [81] and the duration of the neutrino signal from SN1987a (see [71] for a recent re-analysis). To be safe from these constraints, we limit ourselves to $m_\phi > 30 \, \text{keV}$ and $y_{\text{SM}} > 3 \cdot 10^{-5}$ unless explicitly stated otherwise.

3.4 Direct detection experiments and thermalisation

The scenario considered in this work is very strongly constrained by direct detection experiments. The reason is that the Lagrangian given in eqs. (2.1) and (2.2) leads to spin-independent DM-nucleon scattering, which is significantly enhanced by the small mass of the mediator ϕ. The corresponding scattering cross section of dark matter off a target nucleus with mass m_T and mass number A is given by

$$\frac{\mathrm{d}\sigma_T^{\text{SI}}}{\mathrm{d}E_R} = \frac{f_p^2 \, m_p^2}{2\pi \, v_{\text{EW}}^2} \frac{m_T A^2 F^2(E_R)}{v^2} \frac{y_\psi^2 \, y_{\text{SM}}^2 \, \cos^2 \delta_\psi \, \cos^2 \delta_{\text{SM}}}{(m_\phi^2 + q^2)^2} , \tag{3.11}$$

where $f_p = f_n \approx 0.3$ is the effective nucleon coupling [82], v is the DM velocity, $F^2(E_R)$ is the usual form factor for spin-independent scattering and $q = \sqrt{2 m_T E_R} \sim (1\text{–}100) \, \text{MeV}$ denotes the momentum transfer in a nuclear recoil event. For $m_\phi \lesssim q$, the scattering process can effectively be treated as a long-range interaction.

The strongest constraints on σ_N come from LUX [83, 84] and PandaX [85] for DM masses above 5 GeV and from CRESST-II [28, 29] and CDMSlite [86] for smaller masses.

We re-analyse the 2015 results from LUX [83] following [87] to take into account the modified shape of the nuclear recoil spectrum due to the transition from contact interactions to long-range interactions. For CRESST-II, we consider the first bin ($0.3\,\mathrm{keV} \leq E_R \leq 0.4\,\mathrm{keV}$) of the data presented in [28], and compare the six observed events to the number of expected events calculated from the efficiencies provided in [29]. We conservatively only include recoil energies above 0.114 keV in our calculation, corresponding to three standard deviations in the energy resolution, and calculate the upper bound by assuming no background events.

The expression for the DM-nucleon scattering cross section in eq. (3.11) suggests that constraints from direct detection experiments become arbitrarily weak as $\delta_{\mathrm{SM}} \to \pi/2$ or $\delta_\psi \to \pi/2$. While it is true that standard spin-independent interactions vanish in this limit, a non-zero contribution arises from interaction terms that are suppressed by additional powers of q^2/m_N^2 or q^2/m_ψ^2 [88]. To include these effects, we express the scattering amplitude in terms of a complete set of non-relativistic operators [89] and calculate experimental constraints taking into account the full momentum dependence. We find that for $\delta_{\mathrm{SM}} = \pi/2$ direct detection experiments do not yield any relevant constraints on the model parameter space irrespective of the value of δ_ψ.

For $\delta_{\mathrm{SM}} = 0$, on the other hand, the constraints from direct detection experiments are so strong that they put into question one of our fundamental assumptions, namely that the dark sector was in thermal equilibrium with the SM sector at high temperatures. For this to happen, we must require that the DM production rate at some point in the early Universe exceeded the Hubble expansion rate:

$$n_f(T)\langle\sigma(f\bar{f} \to \psi\bar{\psi})v\rangle \gtrsim H(T) , \tag{3.12}$$

where $n_f(T)$ is the fermion number density as a function of the temperature T and the brackets denote thermal averaging [58]. The Hubble rate is given by $H(T) \simeq 1.66\sqrt{g_*}T^2/m_{\mathrm{Pl}}$ in terms of the effective number of degrees of freedom g_* and the Planck mass m_{Pl}. This requirement, which is most easily satisfied for $T \approx \max(m_t, m_\psi)$, can be used to obtain a lower bound on the product of the two couplings:

$$y_{\mathrm{SM}}\, y_\psi \gtrsim \begin{cases} 1.1 \cdot 10^{-6} & \text{for } m_\psi \lesssim m_t \\ \left(\dfrac{m_\psi}{\mathrm{GeV}}\right)^{1/2} \cdot \dfrac{9.4 \cdot 10^{-8}}{\sqrt{1.9 - 0.26\,(1 + \cos 2\delta_\psi)}} & \text{for } m_\psi \gtrsim m_t \end{cases} \quad (\delta_{\mathrm{SM}} = 0) . \tag{3.13}$$

Whenever this lower bound is in conflict with the upper bound obtained from direct detection experiments, the conclusion is that a different mechanism must be responsible for bringing the two sectors into thermal equilibrium in the early Universe.[6] We will return to this issue in the next section.

4 Results

Having discussed the various relevant observables and constraints, we now present the results of our analysis and show the viable regions of parameter space. We are particularly

---

[6]For example, in the toy model introduced in appendix A, thermalisation can also occur via the heavy pseudoscalar A, in which case DM production can be resonantly enhanced.

interested in finding parameter regions where large self-interactions on the scale of dwarf galaxies are consistent with all exclusion bounds. We will begin by revisiting the frequently studied case of purely scalar interactions, i.e. $\delta_{\mathrm{SM}} = \delta_\psi = 0$. We show that in this case direct detection constraints are so strong that it is almost impossible to simultaneously satisfy all phenomenological requirements. A possible way to avoid these constraints would be to consider purely pseudoscalar interactions ($\delta_{\mathrm{SM}} = \delta_\psi = \pi/2$), but in this case one does not obtain large DM self-interactions (see section 3.1). Instead, we will therefore consider the case where $\delta_{\mathrm{SM}} \approx \pi/2$ but δ_ψ is arbitrary, so that direct detection constraints are suppressed but sizeable self-interaction cross sections can be obtained.

4.1 Purely scalar interactions

For purely scalar interactions direct DM annihilation proceeds exclusively via p-wave, and hence does not lead to relevant constraints from indirect detection experiments even when taking into account Sommerfeld enhancement. It has been noted that bound-state formation (BSF) can still potentially proceed via s-wave, leading to additional constraints not included in our calculation [53]. These can potentially be important for specific combinations of m_ψ and m_ϕ in regions of the parameter space where $\alpha_S^2 m_\psi/m_\phi \gtrsim 4$, with $\alpha_S \equiv y_\psi^2 \cos^2(\delta_\psi)/4\pi$ [52, 53]. This is indicated by the gray dashed lines in Figs. 2-4.

The crucial question therefore is whether the bounds from direct detection experiments are compatible with the requirement of large DM self-interactions as well as with the lower bounds on the couplings from the mediator lifetime and the thermalisation condition for the dark sector. We show the preferred and excluded parameter regions for the case of purely scalar interactions in figure 2. At each point, we have fixed y_ψ by the requirement to reproduce the observed relic abundance as described in section 2.2. In the top row, we consider fixed values of the SM coupling y_{SM} and vary the two masses m_ψ and m_ϕ. Constraints from direct detection are shown in purple and brown, whereas the constraint $\sigma_{\mathrm{T}}/m_\psi < 1\,\mathrm{cm^2/g}$ on cluster scales is shown in green. The parameter regions corresponding to large self-interaction cross sections on dwarf galaxy scales are shown in blue: The darker region corresponds to $1\,\mathrm{cm^2 g^{-1}} < \sigma_{\mathrm{T}}/m_\psi < 10\,\mathrm{cm^2 g^{-1}}$, the lighter to $0.1\,\mathrm{cm^2 g^{-1}} < \sigma_{\mathrm{T}}/m_\psi < 1\,\mathrm{cm^2 g^{-1}}$. In addition, we show in orange the region of the parameter space where the mediator is too weakly coupled to bring the two sectors into thermal equilibrium.

For DM masses larger than about 5 GeV, we find direct detection bounds to be so strong that it is essentially impossible to obtain large self-interaction cross sections unless y_{SM} is extremely small, in which case a different mechanism must be responsible for bringing the dark sector and the visible sector into thermal equilibrium and late decays of the mediator are unavoidable. For DM masses below about 0.5 GeV, on the other hand, we find that the DM self-interaction cross section is essentially independent of the DM velocity. This parameter region is phenomenologically less interesting, as it is impossible to obtain larger effects on dwarf galaxy scales than on galaxy cluster scales.

The most interesting parameter region therefore corresponds to DM masses of around 1 GeV, which is precisely the parameter region probed by novel direct detection experiments with very low threshold, such as CRESST-II. We zoom into this parameter region in the
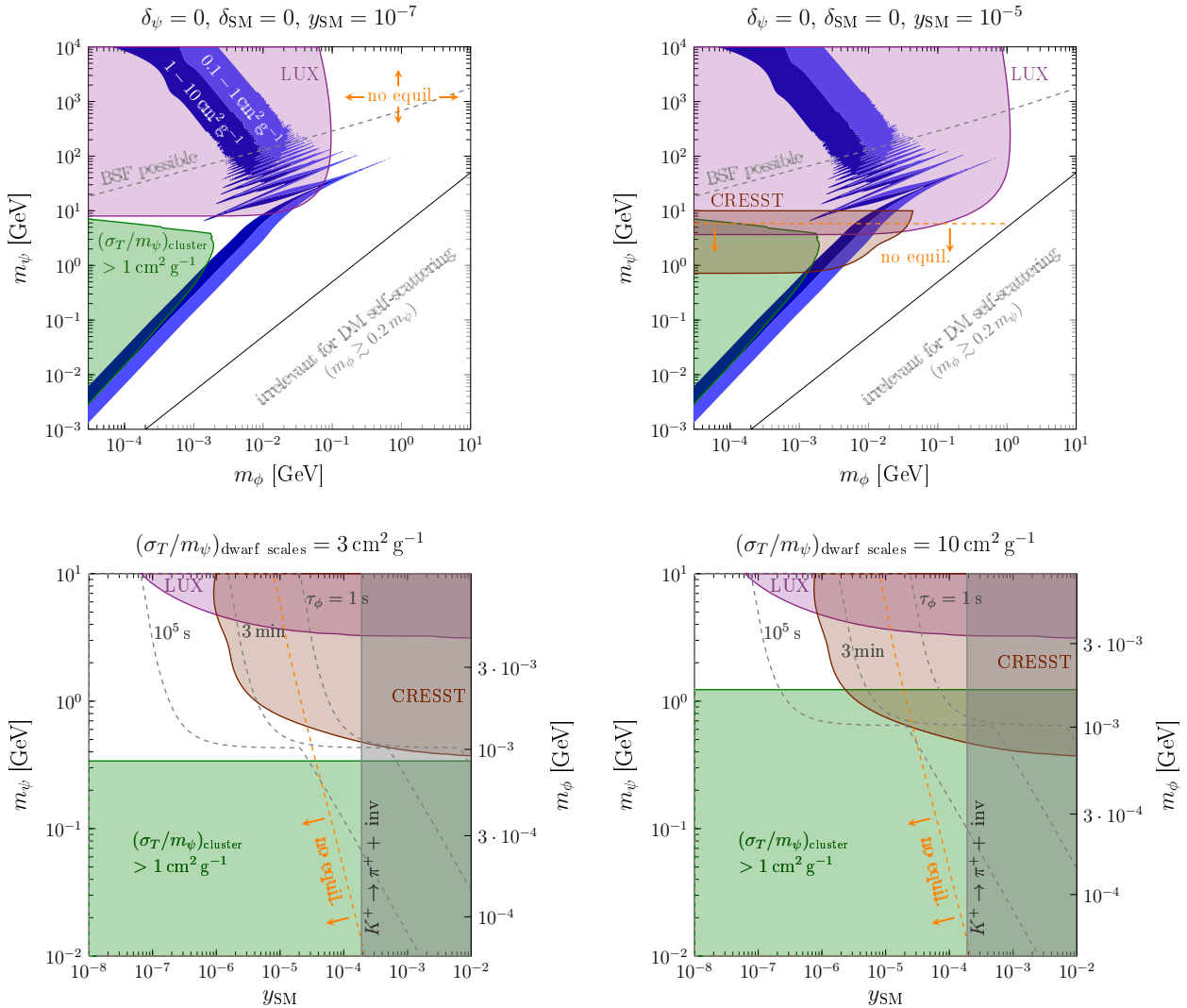
Figure 2. Constraints and interesting parameter regions for the case of purely scalar couplings both to DM and to SM fermions ($\delta_\psi = \delta_{\rm SM} = 0$) for fixed SM coupling $y_{\rm SM}$ (top row) and fixed self-interaction cross section $\sigma_{\rm T}/m_\psi$ (bottom row). In all panels y_ψ is fixed by the relic density requirement. Note that for $10^{-7} \lesssim y_{\rm SM} \lesssim 10^{-4}$ and $m_\phi \lesssim 0.1\,{\rm GeV}$ constraints from SN1987a (not shown) may also become relevant.

bottom row of figure 2 and, rather than fixing $y_{\rm SM}$, impose a fixed value of $\sigma_{\rm T}/m_\psi$ on dwarf galaxy scales. This approach effectively imposes a relation between m_ψ and m_ϕ, as indicated by the two different y-axes.[7] These plots clearly demonstrate that it is impossible to obtain $\sigma_{\rm T}/m_\psi = 10\,{\rm cm^2 g^{-1}}$ on dwarf galaxy scales for any combination of couplings and masses consistent with all other requirements. Self-interaction cross sections of the order of $3\,{\rm cm^2 g^{-1}}$ are possible in a small region of parameter space around $m_\psi \approx 0.5\,{\rm GeV}$,

---

[7]Note that this approach only works for $m_\psi \lesssim 10\,{\rm GeV}$, because for larger DM masses resonances become important and it is no longer possible to fix m_ϕ uniquely as a function of m_ψ. Moreover, for $m_\psi \lesssim 10\,{\rm GeV}$ the velocity dependence of $\sigma_{\rm T}/m_\psi$ saturates well above the scale relevant for dwarf galaxies and therefore the plots shown in the bottom row of figure 2 do not depend on the precise value assumed for $v_{\rm rel}$.

$m_\phi \approx 1\,\mathrm{MeV}$ and $y_{\mathrm{SM}} \approx 10^{-4}$. Note however that in this finely tuned parameter region the mediator has a life time $\gtrsim 1$ s, requiring a dedicated study of constraints from BBN to determine whether this corner of parameter space is still viable. Intriguingly, this window may also soon be probed by measurements of rare kaon decays at NA62 [80].

4.2 Mixed (pseudo)scalar interactions

We have shown above that direct detection constraints are now essentially so strong that it is impossible for the purely scalar case to obtain self-interaction cross sections as large as $10\,\mathrm{cm}^2\mathrm{g}^{-1}$ on dwarf galaxy scales. We therefore now turn to the case where $\delta_{\mathrm{SM}} \approx \pi/2$, i.e. the interaction of ϕ with SM fermions are nearly CP-conserving, while δ_ψ is allowed to take arbitrary values. As discussed in appendix A, this set-up can be obtained naturally from spontaneous CP violation as a consequence of the weak coupling between the dark and the visible sector. Direct detection constraints then do not impose any relevant constraints on the remaining parameters (see section 3.4).[8]

Allowing arbitrary values for δ_ψ means that indirect detection constraints become important. The origin of these constraints can be immediately understood from eq. (2.3) and figure 1. As the s-wave annihilation cross section scales as $\sin^2(2\delta_\psi)$, for $\delta_\psi \approx 0$ or $\delta_\psi \approx \pi/2$ the relic density is dominantly set via p-wave processes, which are irrelevant during recombination and for present-day DM searches. The s-wave contribution, on the other hand, is negligibly small (or in fact zero in the case of purely scalar or pseudoscalar interactions) both during thermal freeze-out and at later times, even when including Sommerfeld enhancement. As soon as we move significantly away from the two limiting cases, however, the relic density is dominantly set via s-wave processes and one obtains relevant constraints from CMB measurements and indirect detection experiments. To calculate these constraints in detail, we take the decay widths and branching ratios for a spin-0 mediator with pseudoscalar couplings to SM fermions from [31]. As shown for the case of a vector mediator in [24] these constraints are typically so strong that they completely exclude the possibility to obtain large DM self-interactions.

The most interesting parameter regions are therefore the ones corresponding to either $\delta_\psi \approx 0$ or $\delta_\psi \approx \pi/2$. The first case is shown in figure 3 for four different values of $\delta_\psi \in [0, 0.1]$, while the latter case is shown in figure 4 for four different values of $\delta_\psi \in [\pi/2 - 0.15, \pi/2]$. Since direct detection constraints are irrelevant in these figures, it is not necessary to specify y_{SM}. In particular, it is always possible for this coupling to be sufficiently large to ensure thermalisation of the two sectors and to avoid late decays of the mediator provided $m_\phi > 2m_e$.

The crucial difference between these two cases is that the Sommerfeld enhancement factor depends only on the scalar component of the DM-mediator coupling, $y_\psi \cos\delta_\psi$, and is therefore much larger for $\delta_\psi \approx 0$ than for $\delta_\psi \approx \pi/2$. We therefore expect much larger self-interaction cross sections – and much stronger constraints from indirect detection experiments – in the former case than in the latter. Indeed, the parameter regions corresponding to large self-interactions shown in figure 3 are very similar to the ones previously

---

[8]We have checked that this statement remains true even if we set $\delta_{\mathrm{SM}} = \pi/2 - y_{\mathrm{SM}}/y_\psi$, corresponding to the amount of CP violation expected in the model discussed in appendix A.
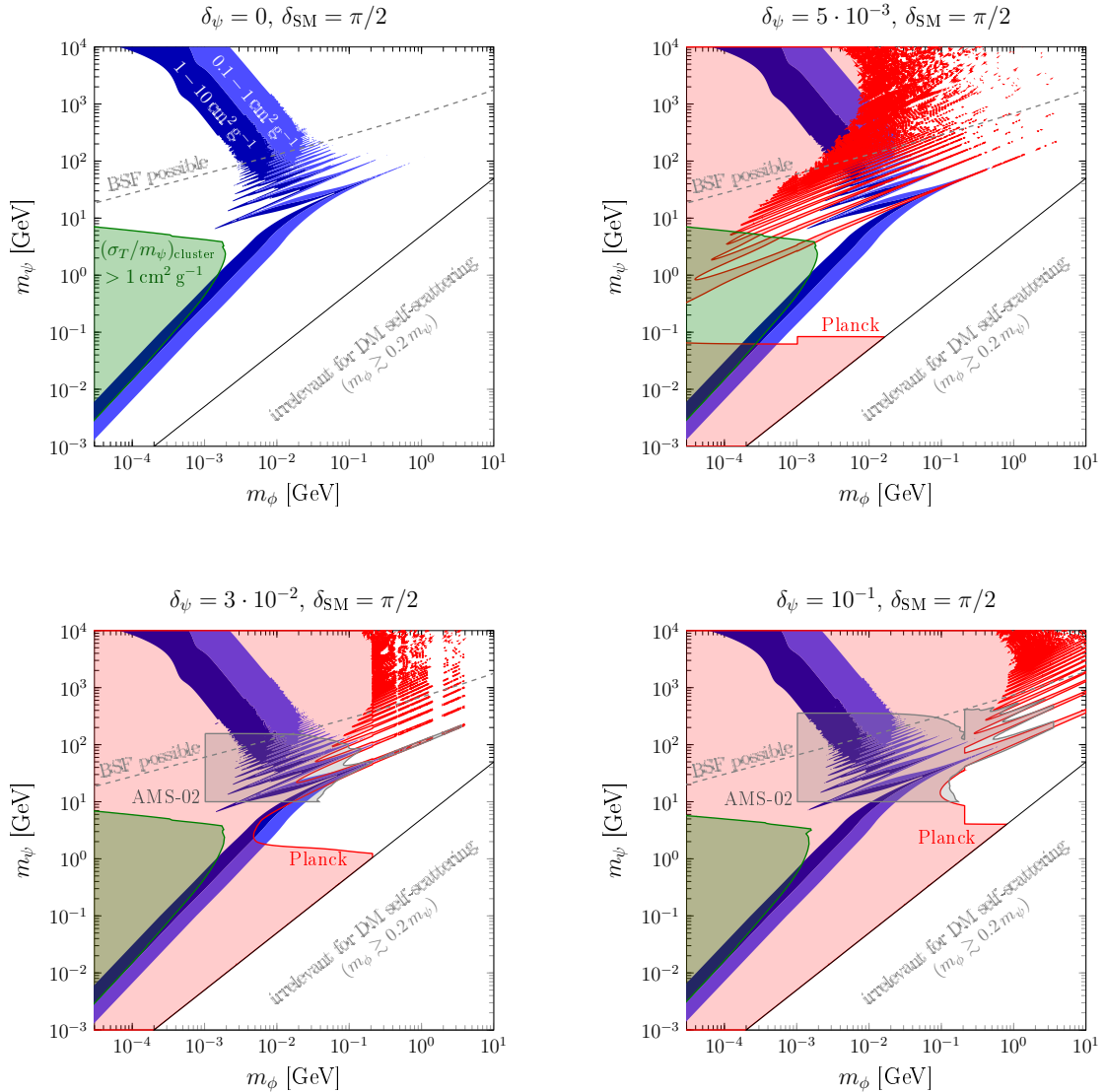
Figure 3. Constraints and interesting parameter regions for the case $\delta_{\text{SM}} = \pi/2$ and $\delta_\psi \approx 0$. In all panels y_ψ is fixed by the relic density requirement.

shown in figure 2. In the case $\delta_\psi = 0$, shown in the top-left panel, there are no relevant indirect detection constraints, so large DM self-interactions are phenomenologically viable. However, since we consider $\delta_{\text{SM}} = \pi/2$ we cannot invoke CP symmetry to ensure that $\delta_\psi = 0$ holds exactly. It is therefore crucial to understand how large δ_ψ can be before the scenario is ruled out by indirect detection constraints. Indeed, for $\delta_\psi = 0.1$ the entire parameter region corresponding to large self-interactions is in conflict with CMB constraints (bottom-right panel). In order to find allowed parameter space we must require δ_ψ to be of order 10^{-2} or smaller (see top-right and bottom-left panel).

For $\delta_\psi \approx \pi/2$ we encounter a different situation, because large self-interactions are only possible if δ_ψ is appreciably different from the purely pseudoscalar case. In particular,
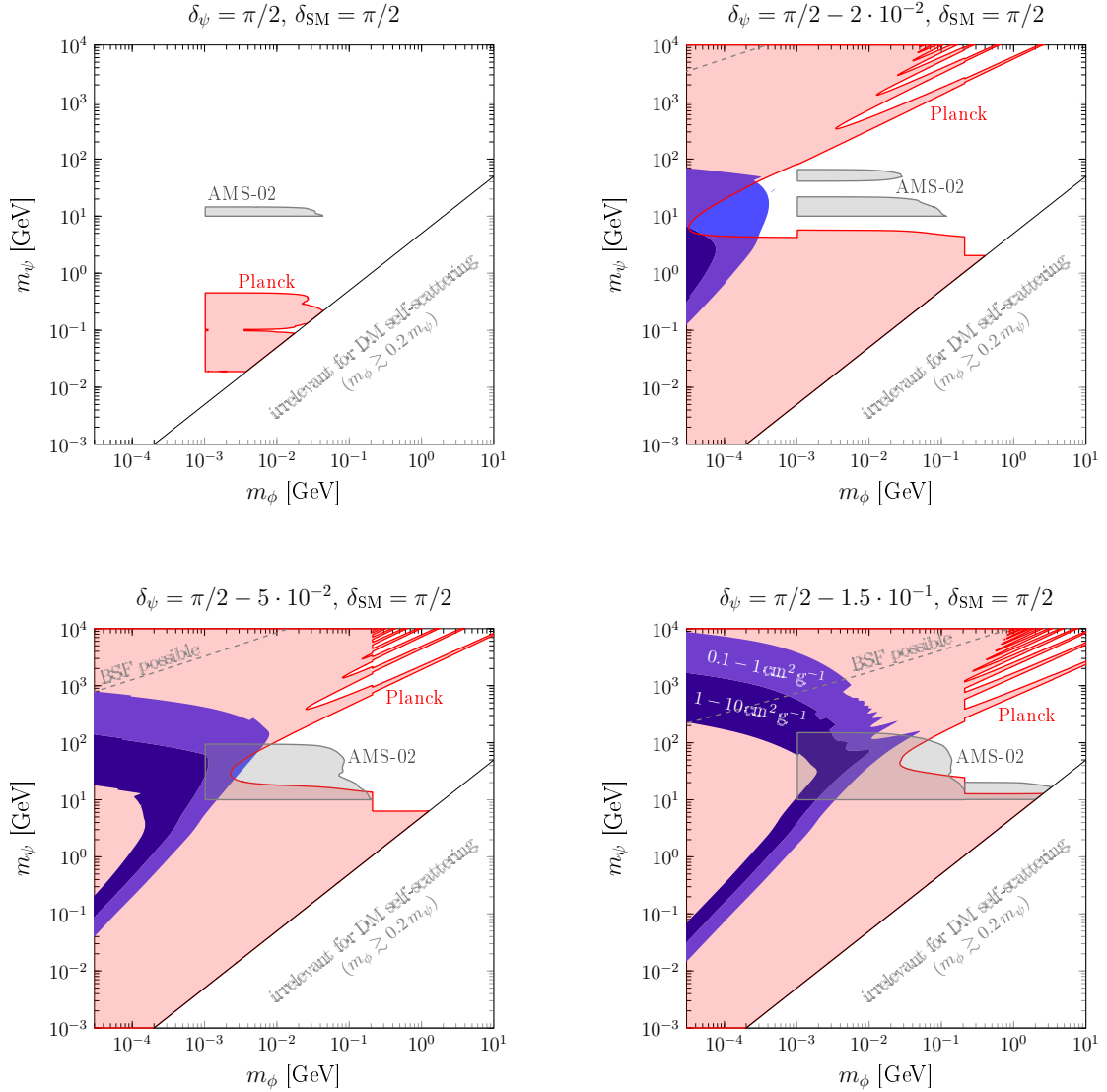
Figure 4. Same as figure 3 but for $\delta_\psi \approx \pi/2$.

the resonant enhancement of self-interactions found in the case $\delta_\psi \approx 0$ is absent unless $\delta_\psi \lesssim \pi/2 - 0.1$. For such values of δ_ψ, however, there are significant constraints from CMB measurements (see bottom-right panel of figure 4). Indeed, in none of the four panels shown in figure 4 is it possible to obtain DM self-interaction cross sections significantly larger than $1\,\mathrm{cm}^2\mathrm{g}^{-1}$ without being excluded by CMB constraints. Note, however, that for $\delta_\psi \approx \pi/2$ the velocity dependence of the self-interactions has typically not saturated at dwarf galaxy scales, so that the precise value of the self-scattering cross section depends on the assumed DM relative velocity.

To understand whether our conclusions depend on the specific choices of δ_ψ made in figure 3 and 4, we can show the constraints as a function of δ_ψ for fixed DM masses. This is done in figure 5 for $m_\psi = 5\,\mathrm{GeV}$ and $m_\psi = 20\,\mathrm{GeV}$, varying δ_ψ in the complete
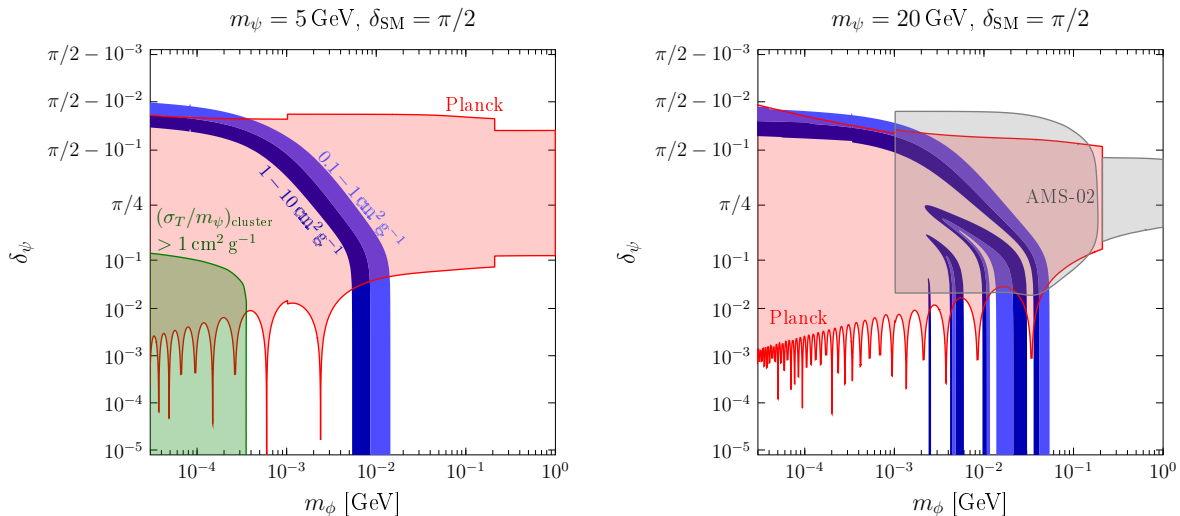
Figure 5. Same as figure 3 and 4 but showing the constraints as a function of δ_ψ for fixed m_ψ. Note the special scaling of the y-axis intended to emphasize the interesting limiting cases $\delta_\psi = 0$ and $\delta_\psi = \pi/2$.

range between 0 and $\pi/2$. As expected, indirect detection constraints are overwhelming in the range $0.1 \lesssim \delta_\psi \lesssim \pi/2 - 0.1$. For larger δ_ψ, self-interaction cross sections as large as $\gtrsim 1\,\mathrm{cm}^2/\mathrm{g}$ are marginally compatible with CMB measurements. Values of δ_ψ smaller than 0.1, on the other hand, make it possible to reconcile the requirement of large self-interactions on dwarf galaxy scales with all indirect detection constraints.

To conclude this section, let us briefly discuss how our results would change if we were to allow arbitrary values of δ_{SM}. First of all, varying δ_{SM} would affect the branching ratios of the mediator for $m_\phi > 2m_\pi$, as hadronic decay modes are more important for scalar than for pseudoscalar couplings. The resulting changes in the CMB constraints do however not modify any of our central conclusions. The more important difference clearly comes from the fact that direct detection constraints become relevant again. For example, for $m_\psi = 5\,\mathrm{GeV}$, $m_\phi = 10\,\mathrm{MeV}$, $\delta_\psi = 10^{-2}$ and $y_{\mathrm{SM}} = 10^{-4}$, direct detection constraints will become relevant as soon as $\delta_{\mathrm{SM}} \lesssim \pi/2 - 6 \cdot 10^{-3}$. In other words, for this solution to be viable a mechanism like the one discussed in appendix A must ensure that δ_{SM} is indeed very close to $\pi/2$.

5 Conclusions

In this work we have discussed the detailed phenomenology of DM particles interacting with each other via the exchange of a very light spin-0 mediator. Such a set-up is motivated by the observation that the resulting DM self-interactions can be sufficiently large to affect astrophysical systems and at the same time exhibit a characteristic velocity dependence which makes it possible to consistently describe astrophysical observations over a range of different scales. Furthermore, thermal freeze-out can be naturally realised in models

with a light mediator and the observed DM relic abundance can be obtained via direct annihilation of a DM pair into light mediators.

In a minimal realisation of the set-up, the mediator must also couple to SM states in order to bring the dark sector into thermal equilibrium with the SM sector and to ensure that the mediator decays before BBN. In this case, one can expect a number of relevant constraints from a variety of experimental and observational probes. Specifically, we have considered direct and indirect detection experiments, as well as CMB measurements and bounds from searches for rare meson decays, and compared the resulting constraints to the parameter regions relevant for DM self-interactions.

For the case of purely scalar interactions, we have pointed out the important role of low-threshold direct detection experiments like CRESST-II. These searches strongly constrain the viable regions of parameter space and essentially make it impossible to obtain self-interaction cross sections larger than $5\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ on dwarf galaxy scales and smaller than $1\,\mathrm{cm}^2\,\mathrm{g}^{-1}$ on galaxy cluster scales in the simplest set-up. For purely pseudoscalar interactions on the other hand no relevant DM self-interactions are expected to arise.

Large self interactions with substantially suppressed constraints can be realised however if the mediator couples to DM and SM states with different CP phases δ_ψ and δ_{SM}, respectively. In particular, for $\delta_{\mathrm{SM}} \approx \pi/2$, direct detection constraints are largely absent. In the absence of CP conservation, however, indirect detection constraints and CMB constraints are very strong and impose $\delta_\psi \lesssim 0.1$ or $\delta_\psi \gtrsim \pi/2 - 0.1$. In particular in the former case it is possible to obtain large DM self-interactions on dwarf galaxy scales consistent with all other requirements.

There are a number of ways in which the tension between the different constraints and requirements discussed in this work can be ameliorated. For example, for the case of purely scalar interactions constraints from direct detection experiments may be reconciled with BBN constraints in models where the mediator has suppressed couplings to nucleons but can still decay sufficiently quickly into light leptons and photons (e.g. leptophilic DM [90]). In such a model one would likely need a different mechanism for bringing the dark sector into thermal equilibrium with the SM, which offers the interesting possibility to have different temperatures in the two sectors. In fact, it is also possible to obtain SIDM from a dark sector that was never in thermal equilibrium with the SM, for example via the freeze-in mechanism [22], so that direct and indirect detection constraints are absent. Finally, it is worth pointing out that both BBN and CMB constraints can be significantly weakened if DM is asymmetric [39] or if the mediator decays into inert particles such as sterile neutrinos [17]. Exploring cosmological constraints for such set-ups offer an interesting avenue for future research.

In conclusion, while the idea of sizeable DM self-interactions remains very attractive, the simplest attempts to construct specific models based on light mediators face a number of strong constraints. Following the recent observation that vector mediators are in strong tension with CMB constraints and indirect detection experiments [24], we have extended previous analyses in this work to show that also the case of scalar mediators is essentially ruled out in its simplest realization. While extended models, such as mediators with CP-violating couplings, are presently still viable, the exciting interplay of astrophysical

observables and particle physics experiments means that we can hope to comprehensively explore the idea of DM particles interacting via the exchange of light mediators in the near future.

Acknowledgments

A A toy model

In this appendix we discuss how the coupling structure that we consider can arise from a theory with spontaneous breaking of CP symmetry. The starting point is a CP conserving theory that contains a Dirac fermion ψ and a pseudoscalar P:

$$\mathcal{L}_{\text{DM}} = \bar{\psi}(i\slashed{\partial} - m_0)\psi - iy_\psi P\bar{\psi}\gamma^5\psi - V(P) \,. \tag{A.1}$$

The fact that the theory is CP-conserving is reflected in the fact that the coupling y_ψ is real and that under a CP transformation $P \to -P$ so that the Yukawa interaction remains invariant.

Since P is a SM singlet, it cannot couple directly to either $\bar{q}_L u_R$ or $\bar{q}_L d_R$, both of which are not invariant under $SU(2)$. For a CP-even real scalar singlet S such interactions can be generated after electroweak symmetry breaking (EWSB) via mixing with the SM Higgs. If the scalar potential contains the term $\mu_S H^\dagger H S$, EWSB will induce an off-diagonal mass term of the form $\frac{1}{2}\mu_S v_{\text{EW}} h S$, where v_{EW} is the electroweak vacuum expectation value (vev) and $H = \frac{1}{\sqrt{2}}(0, h + v_{\text{EW}})^T$. For a CP-odd real scalar P, however, the analogous term $\mu_P H^\dagger H P$ would violate CP and is therefore absent.

To generate couplings of a CP-odd real scalar singlet to SM quarks, one typically assumes the presence of two Higgs doublets, H_1 and H_2, which contain one physical pseudoscalar degree of freedom A after EWSB [91–94]. One can then consider a CP-conserving mixing of the form $i\mu_P H_1^\dagger H_2 P + \text{h.c.}$, which after EWSB leads to the mixing term $\frac{1}{2}v_{\text{EW}} \mu_P A P$.

The mixing between the pseudoscalar singlet P and the pseudoscalar component of the Higgs doublet then leads to couplings of P to SM fermions f:[9]

$$\mathcal{L}_{\text{mixing}} \supset -i\sin\theta \sum_f \frac{y_f}{\sqrt{2}} P\bar{f}\gamma^5 f \,, \tag{A.2}$$

---

[9]We note that in generic two Higgs doublet models it is possible for the pseudoscalar to have e.g. enhanced couplings to down-type fermions, depending on the detailed structure of the Yukawa sector and the ratio $\tan\beta$ of the vevs of the two Higgs doublets. We assume here that $\tan\beta \approx 1$ and hence the pseudoscalar couples in the same fashion as the SM-like Higgs boson.

where y_f denotes the SM Yukawa couplings and the mixing angle θ is given by

$$\tan 2\theta = \frac{\mu_P \, v_{\text{EW}}}{m_P^2 - m_A^2} \ . \tag{A.3}$$

If P has a mass below the GeV scale, SM precision measurements (for example of rare meson decays) constrain $\sin\theta$ to be very small, typically of order 10^{-4} or less (see section 3.3). Such small mixing angles can easily be achieved if $\mu_P \ll v_{\text{EW}} \ll m_A$. In this case the mass eigenstates are almost identical to the interaction eigenstates P and A (and we will hence not make a distinction between the two).

Let us now assume that the potential of the pseudoscalar P is given by

$$V(P) = -\mu^2 P^2 + \lambda_P P^4 \ , \tag{A.4}$$

so that P obtains a vev $v_P = \mu/\sqrt{2\lambda_P}$, which spontaneously breaks the CP symmetry. Writing $P = v_P + \phi$, we obtain

$$\mathcal{L}_{\text{DM}} = \bar\psi \left[i \slashed\partial - (m_0 + i y_\psi v_P \gamma^5) \right] \psi - i y_\psi \phi \bar\psi \gamma^5 \psi - V(\phi) \ . \tag{A.5}$$

Defining $\tan\alpha = y_\psi v_P / m_0$, we can perform a chiral rotation of the DM field ψ to absorb the complex phase in the DM mass: $\psi \to \exp(i\gamma^5 \alpha/2)\psi$:

$$\mathcal{L}_{\text{DM}} = \bar\psi(i\slashed\partial - m_\psi)\psi - y_\psi \phi \bar\psi(\cos\delta_\psi + i\sin\delta_\psi \gamma^5)\psi \ , \tag{A.6}$$

where we have defined $m_\psi = \sqrt{y_\psi^2 v_P^2 + m_0^2}$ and $\delta_\psi = \pi/2 - \alpha$. The fact that ϕ obtains both scalar and pseudoscalar couplings to DM makes the spontaneous CP breaking explicit. In fact, for $y_\psi v_P \gg m_0$ we find $\delta_\psi \approx 0$, i.e. the CP violation in the dark sector is maximal.

In a similar way the spontaneous symmetry breaking will induce a complex mass term for the SM fermions:

$$\mathcal{L}_{\text{mass}} \supset - \sum_f \bar{f} \frac{y_f}{\sqrt{2}} (v_{\text{EW}} + i\sin\theta \, v_P \gamma^5) f \ . \tag{A.7}$$

In complete analogy to the DM field ψ we can now define a phase $\alpha_{\text{SM}} = \sin\theta \, v_P / v_{\text{EW}}$ and perform chiral rotations of the fermion fields to recover real mass terms. It then becomes apparent that the spontaneous CP breaking in the dark sector potentially also induces CP violation in SM observables.

The crucial point is however that as discussed above the mixing angle must be very small, $\sin\theta \lesssim 10^{-4}$, and hence the CP-violating phase is also small: $\alpha_{\text{SM}} \lesssim 10^{-4} v_P/v_{\text{EW}}$.[10] In other words, because P is only very weakly coupled to the SM, the spontaneous breaking of CP does not induce any large effects in SM observables. In particular, we can write the couplings of ϕ to SM fermions as

$$\mathcal{L}_{\text{mixing}} = -y_{\text{SM}} \sum_f \frac{y_f}{\sqrt{2}} \phi \bar{f}(\cos\delta_{\text{SM}} + i\sin\delta_{\text{SM}}\gamma^5)f \tag{A.8}$$

---

[10]We note that another source of CP violation arises from quartic couplings of the form $\lambda_{P1} H_1^\dagger H_1 P^2$, which induce mixing between the SM Higgs and the pseudoscalar singlet once P obtains a vev. We must require that these quartic interactions are sufficiently small that P does not obtain unacceptably large couplings to SM states. In this case, the resulting CP-violating effects will also be small.

with $y_{\mathrm{SM}} = \sin\theta$ and $\delta_{\mathrm{SM}} = \pi/2 - \alpha_{\mathrm{SM}}$. We then find $\cos\delta_{\mathrm{SM}} \approx \alpha_{\mathrm{SM}} \ll 1$ and $\sin\delta_{\mathrm{SM}} \approx 1$, i.e. ϕ has almost exclusively pseudoscalar couplings to SM states.

To conclude this discussion, let us briefly review experimental bounds on light spin-0 bosons with CP-violating couplings. For $m_\phi < m_e$ such a light scalar can potentially induce a sizeable electric dipole moment of the electron, which is constrained to be $|d_e| < 8.7 \cdot 10^{-29}\, e\,\mathrm{cm}$ [95]. The one-loop contribution is given by [36]

$$d_e = y_{\mathrm{SM}}^2 \frac{m_e}{v_{\mathrm{EW}}^2} \frac{e}{16\pi^2} \sin 2\delta_{\mathrm{SM}}\,, \tag{A.9}$$

whereas the two-loop Barr-Zee contribution is approximately given by [37]

$$d_e \sim y_{\mathrm{SM}}^2 \frac{m_e}{v_{\mathrm{EW}}^2} \frac{e}{16\pi^2} \frac{\alpha}{\pi} \sin 2\delta_{\mathrm{SM}} \mathcal{F}(m_\phi) \log(\Lambda/m_\phi)\,. \tag{A.10}$$

Here $\mathcal{F}(m_\phi)$ denotes the form factor for the effective vertex $\phi F^{\mu\nu}\tilde{F}_{\mu\nu}$ obtained from integrating out heavy quarks, mesons and leptons. For $m_\phi \sim m_e$ and $\Lambda \sim 1\,\mathrm{TeV}$, one obtains $\mathcal{F}(m_\phi)\log(\Lambda/m_\phi) \sim 10^2$, meaning that the one-loop and two-loop contributions are of similar magnitude.[11] The experimental bounds is thus satisfied for $y_{\mathrm{SM}}^2 \sin 2\delta_{\mathrm{SM}} \lesssim 10^{-4}$, which is always the case in the parameter region that we consider. Experimental bounds on the electric dipole momenta of the muon, the neutron or mercury nuclei give comparable or weaker constraints.

B   Dark matter self-interactions from pseudoscalar exchange

Tree-level analysis

The self-interactions of DM particles induced by the exchange of a spin-0 boson ϕ can be calculated by solving the Schrödinger equation for ψ in the presence of a non-relativistic scattering potential $V(r)$. This potential is obtained by taking the Fourier transform of the matrix element $\mathcal{M}_{\psi\psi \to \psi\psi}$ with respect to the exchanged three-momentum $\vec{q}$. For the scalar coupling $\bar{\psi}\psi\phi$ between DM and the mediator, at lowest order in $\vec{q}^2$ this procedure gives rise to an attractive Yukawa potential,

$$V_{\mathrm{S}}(r) = -\alpha_{\mathrm{S}} \frac{e^{-m_\phi r}}{r} \tag{B.1}$$

with $\alpha_{\mathrm{S}} \equiv (y_\psi \cos\delta_\psi)^2/(4\pi)$, which can induce strong self-interactions of DM. For the pseudoscalar coupling $\bar{\psi}\gamma^5\psi\phi$, at lowest non-vanishing order in $\vec{q}^2$ one obtains

$$\begin{aligned}
V_{\mathrm{PS}}(r) = &\frac{\alpha_{\mathrm{PS}}}{12} \frac{m_\phi^2}{m_\psi^2} \left(\frac{e^{-m_\phi r}}{r} - \frac{4\pi}{m_\phi^2} \delta^{(3)}(\vec{r}) \right) \vec{\sigma}_1 \cdot \vec{\sigma}_2 \\
&+ \frac{\alpha_{\mathrm{PS}}}{12} \frac{m_\phi^2}{m_\psi^2} \left(1 + \frac{3}{m_\phi r} + \frac{3}{m_\phi^2 r^2} \right) \frac{e^{-m_\phi r}}{r} S_{12}(\vec{r})\,,
\end{aligned} \tag{B.2}$$

---

[11]Higher-order contributions from light-by-light scattering and vacuum polarisation are found to be subdominant [37].

with $\alpha_{\rm PS} \equiv (y_\psi \sin \delta_\psi)^2/(4\pi)$ and $S_{12}(\vec{r}) \equiv 3(\vec{\sigma}_1 \cdot \hat{r})(\vec{\sigma}_2 \cdot \hat{r}) - \vec{\sigma}_1 \cdot \vec{\sigma}_2$, which agrees with [96], including the overall sign. The Yukawa-like part of this potential is suppressed by m_ϕ^2/m_ψ^2, and hence is irrelevant for the self-scattering of DM. The part proportional to $1/r^3$, on the other hand, is not suppressed in the limit $m_\phi \ll m_\psi$, and could potentially lead to strong self-interactions. However, it is well-known that the singular behaviour of a $1/r^3$ potential leads to ill-defined solutions of the Schrödinger equation; in particular a naive calculation of the scattering amplitude for DM self-scattering diverges, making it impossible to directly extract information about the self-scattering cross section of DM [97, 98].

It is however by no means clear that the $1/r^3$ singularity is actually physical. The potential given by eq. (B.2) is derived by taking the Fourier transform of the matrix element $\mathcal{M}_{\psi\psi\to\psi\psi}$ and expand the resulting expression for small relative momenta $\vec{q}$, corresponding to large values of r. Hence, to investigate the behaviour of $V(r)$ for $r \to 0$ more and more powers of q have to be taken into account, leading to additional terms $1/r^n$ in the potential, which are expected to regulate the behaviour of the potential at the origin.

In fact, exactly the same argument applies also to the simpler case of scalar or vector exchange. For example, the famous term leading to the spin-orbit coupling in the hydrogen atom is given by $V_{\rm spin\text{-}orbit} \propto (\vec{L} \cdot \vec{S})/r^3$. In the familiar context of the fine splitting of the hydrogen energy levels, the singularity of this potential is not a problem, as in first order perturbation theory one is only interested in the finite expectation value of this operator with respect to the unperturbed wave function. However, one would encounter the same problems with the singular behaviour at $r \to 0$ when attempting to actually solve the Schrödinger equation for the hydrogen atom including this correction term. In other words, when worrying about the singular behaviour of the pseudoscalar potential $V_{\rm PS}(r)$, one should also worry about the corresponding singularities in the scalar potential $V_{\rm S}(r)$. Intuitively, the higher-order terms are then expected to give rise to only subdominant contributions to the scattering cross section of DM, suggesting that the exchange of a pseudoscalar does not give rise to strong self-interactions of DM compared to the cross section induced by scalar exchange. Hence, in the following we will only take into account the scattering of DM off the lowest-order potential $V_{\rm S}(r)$ given by eq. (3.3).[12]

Possible impact of one-loop corrections

It has been proposed that the exchange of two pseudoscalars via a box diagram could effectively give rise to a scalar coupling between the DM particles and hence to larger self-interaction rates [99]. We have repeated the calculation of these loop diagrams, and we find that the momentum and velocity dependence of the resulting scattering amplitude *cannot* be described by the effective exchange of a scalar particle. In particular, in the limit $v \to 0$ the scattering cross section is given by

$$\sigma_{\psi\psi\to\psi\psi}^{(\rm box)} = \frac{\alpha_{\rm PS}^4}{64\pi^2 m_\psi^2} \left[-3 + \log\left(\frac{m_\psi^2}{m_\phi^2}\right) \right]^2 \tag{B.3}$$

---

[12]For the same reason we also neglect the monopole-dipole potential which is present for CP-violating phases δ_ψ [96]: All terms in that potential are suppressed either by m_ϕ or by additional powers of $1/r$.

with α_{PS} defined as above. This is drastically different from the tree-level exchange of a scalar particle, which gives rise to a cross section $\propto 1/m_{\mathrm{scalar}}^4$ in the limit $v \to 0$. In particular, we find numerically that the one-loop induced cross section is always significantly too small in order to give rise to relevant self-interactions of DM particles. This finding agrees with similar conclusions obtained for loop-induced direct detection cross sections [100].

C Dark matter self-interactions from scalar exchange

In this appendix we summarize the formalism for the calculation of the momentum transfer cross section σ_{T} arising from the scalar coupling of the DM particle to the spin-0 mediator ϕ. In contrast to previous works, we fully take into account effects arising from the indistinguishability of the scattered particles. Under the assumption that there is no asymmetry in the abundances of ψ and $\bar{\psi}$, the averaged momentum transfer cross section is given by

$$\sigma_{\mathrm{T}} = \frac{1}{2} \left(\sigma_{\mathrm{T}}^{\mathrm{PP}} + \sigma_{\mathrm{T}}^{\mathrm{PA}} \right) , \tag{C.1}$$

where PP (PA) denotes particle-particle (particle-antiparticle) scattering and

$$\sigma_{\mathrm{T}}^{\mathrm{PP,PA}} = \int \mathrm{d}\Omega \, (1 - |\cos\theta|) \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \right)^{\mathrm{PP,PA}} . \tag{C.2}$$

The factor $1/2$ for two identical particles in the final state has been included implicitly in the definition of $(\mathrm{d}\sigma/\mathrm{d}\Omega)^{\mathrm{PP}}$.

Born regime

Non-perturbative effects in the scattering process are negligible as long as $\alpha_S m_\psi/m_\phi \ll 1$ (the Born regime), where $\alpha_S \equiv y_\psi^2 \cos^2 \delta_\psi/(4\pi)$ denotes the coupling strength relevant for DM self-interactions. We find

$$\sigma_{\mathrm{T}}^{\mathrm{PP}}\big|_{\mathrm{Born}} = \frac{4\pi\alpha_S^2}{m_\psi^2 v^4} \left[6\log\left(\frac{m_\psi^2 v^2}{2m_\phi^2} + 1 \right) - \frac{4m_\psi^2 v^2 + 6m_\phi^2}{m_\psi^2 v^2 + 2m_\phi^2} \log\left(\frac{m_\psi^2 v^2}{m_\phi^2} + 1 \right) \right] , \tag{C.3}$$

$$\sigma_{\mathrm{T}}^{\mathrm{PA}}\big|_{\mathrm{Born}} = \frac{16\pi\alpha_S^2}{m_\psi^2 v^4} \left[\log\left(\frac{m_\psi^2 v^2}{2m_\phi^2} + 1 \right) - \frac{m_\psi^2 v^2}{m_\psi^2 v^2 + 2m_\phi^2} \right] . \tag{C.4}$$

Compared to [18], where only the t-channel contribution is included and the momentum transfer cross section is defined via $\int \mathrm{d}\Omega\,(1-\cos\theta)\,(\mathrm{d}\sigma/\mathrm{d}\Omega)$, our full calculation for particle-particle scattering in the Born regime gives the same result in the limit $m_\psi v \gg m_\phi$, while it is a factor of 4 smaller for $m_\psi v \ll m_\phi$. For particle-antiparticle scattering, the expressions above are larger by a factor of 2 for $m_\psi v \gg m_\phi$ and smaller by a factor of 2 for $m_\psi v \ll m_\phi$ compared to the results obtained in [18].

Non-perturbative regime

For $\alpha_S m_\psi/m_\phi \gtrsim 1$, non-perturbative effects corresponding to the multiple exchange of the mediator ϕ have to be taken into account. In practice, one numerically solves the

Schrödinger equation corresponding to the scattering potential using the standard techniques of partial wave decomposition (see also appendix B). The scattering amplitude can be written as

$$f(\theta) = \frac{2}{m_\psi v} \sum_{l=0}^{\infty} (2l+1) e^{i\delta_l} \sin \delta_l P_l(\cos\theta) , \qquad (C.5)$$

where we obtain the phase shifts δ_l by employing the numerical technique described in [18]. For scattering of identical particles, the differential scattering cross section then follows from

$$\left(\frac{d\sigma}{d\Omega}\right)_\xi^{PP} = |f(\theta) + \xi f(\pi - \theta)|^2 \qquad (C.6)$$

with $\xi = +1$ (-1) if the spatial wave function is symmetric (antisymmetric) under particle exchange. If scattering occurs via the spin-singlet channel, the spin wave function is antisymmetric, and hence the spatial wave function has to be symmetric; correspondingly, for scattering via the spin-triplet channel the spatial wave function is antisymmetric. Assuming unpolarized DM particles, we obtain

$$\left(\frac{d\sigma}{d\Omega}\right)^{PP} = \frac{1}{4}\left(\frac{d\sigma}{d\Omega}\right)_{\xi=+1}^{PP} + \frac{3}{4}\left(\frac{d\sigma}{d\Omega}\right)_{\xi=-1}^{PP} \qquad (C.7)$$

On the other hand, since particle and anti-particle are distinguishable, one simply obtains

$$\left(\frac{d\sigma}{d\Omega}\right)^{PA} = |f(\theta)|^2 . \qquad (C.8)$$

The momentum transfer cross sections σ_T then follows from eq. (C.2) by numerically integrating over the angular variable θ.[13]

Classical regime

For $m_\psi v \gg m_\phi$ (denoted as classical regime), more and more partial waves l have to be taken into account in the calculation of the scattering amplitude in eq. (C.5), in order to obtain a sufficiently convergent series. At some point (in our case typically at $l_{max} \sim 100$), the numerical approach becomes infeasible, and instead we use an approach based on the fitting functions for σ_T provided in [23]. However, these results have been obtained using only t-channel exchange and adopting the definition $\int d\Omega (1 - \cos\theta) (d\sigma/d\Omega)$ for the momentum transfer cross section. We find that in order to match our results based on the numerical solution of the Schrödinger equation sufficiently smoothly onto the fitting functions in the regime where $m_\psi v \gg m_\phi$, we need to multiply the expressions given in [23] by a factor $1/2$. This factor of $1/2$ can be understood from the fact that $d\sigma/d\Omega$ is approximately independent of $\cos\theta$ in the classical regime [18] and that

$$\int_{-1}^{1} d\cos\theta (1 - |\cos\theta|) = \frac{1}{2} \int_{-1}^{1} d\cos\theta (1 - \cos\theta) . \qquad (C.9)$$

---

[13]Due to the presence of the weighting factor $1 - |\cos\theta|$ instead of $1 - \cos\theta$ in the definition of σ_T, there is no simple analytical expression for the momentum transfer cross section in terms of the phase shifts δ_l.
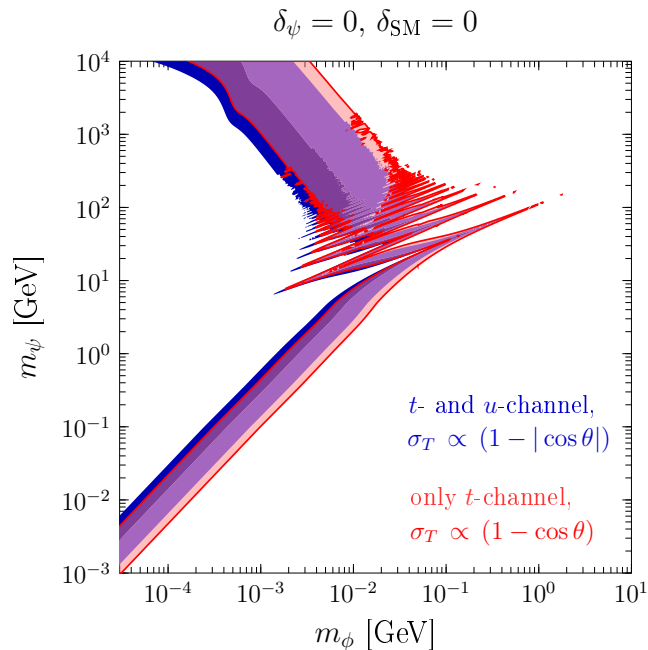
Figure 6. Regions of the parameter space leading to sufficiently large DM self-interactions on dwarf scales, obtained via the full calculation of σ_T (blue shaded bands) and via the approximation of classical distinguishability in the scattering process (red curves).

To conclude this discussion, we illustrate in figure 6 the impact of using the correct definition for the momentum transfer cross section. The blue regions correspond to the full calculation, which takes into account the indistinguishability of two DM particles. The red lines, on the other hand, illustrate the results that one obtains from the simpler calculation, including only t-channel scattering and defining the momentum transfer cross section via $\int d\Omega \, (1 - \cos\theta) d\sigma/d\Omega$.

References

[1] A. A. de Laix, R. J. Scherrer, and R. K. Schaefer, *Constraints of selfinteracting dark matter*, *Astrophys. J.* **452** (1995) 495, [astro-ph/9502087].

[2] D. N. Spergel and P. J. Steinhardt, *Observational evidence for selfinteracting cold dark matter*, *Phys. Rev. Lett.* **84** (2000) 3760–3763, [astro-ph/9909386].

[3] M. R. Buckley and P. J. Fox, *Dark Matter Self-Interactions and Light Force Carriers*, *Phys. Rev.* **D81** (2010) 083522, [arXiv:0911.3898].

[4] J. L. Feng, M. Kaplinghat, and H.-B. Yu, *Halo Shape and Relic Density Exclusions of Sommerfeld-Enhanced Dark Matter Explanations of Cosmic Ray Excesses*, *Phys. Rev. Lett.* **104** (2010) 151301, [arXiv:0911.0422].

[5] J. L. Feng, M. Kaplinghat, H. Tu, and H.-B. Yu, *Hidden Charged Dark Matter*, *JCAP* **0907** (2009) 004, [arXiv:0905.3039].

[6] A. Loeb and N. Weiner, *Cores in Dwarf Galaxies from Dark Matter with a Yukawa Potential*, Phys. Rev. Lett. **106** (2011) 171302, [arXiv:1011.6374].

[7] J. Zavala, M. Vogelsberger, and M. G. Walker, *Constraining Self-Interacting Dark Matter with the Milky Way's dwarf spheroidals*, Monthly Notices of the Royal Astronomical Society: Letters **431** (2013) L20–L24, [arXiv:1211.6426].

[8] M. Vogelsberger, J. Zavala, and A. Loeb, *Subhaloes in Self-Interacting Galactic Dark Matter Haloes*, Mon. Not. Roy. Astron. Soc. **423** (2012) 3740, [arXiv:1201.5892].

[9] M. Markevitch, A. H. Gonzalez, D. Clowe, A. Vikhlinin, L. David, W. Forman, C. Jones, S. Murray, and W. Tucker, *Direct constraints on the dark matter self-interaction cross-section from the merging galaxy cluster 1E0657-56*, Astrophys. J. **606** (2004) 819–824, [astro-ph/0309303].

[10] S. W. Randall, M. Markevitch, D. Clowe, A. H. Gonzalez, and M. Bradac, *Constraints on the Self-Interaction Cross-Section of Dark Matter from Numerical Simulations of the Merging Galaxy Cluster 1E 0657-56*, Astrophys. J. **679** (2008) 1173–1180, [arXiv:0704.0261].

[11] A. H. G. Peter, M. Rocha, J. S. Bullock, and M. Kaplinghat, *Cosmological Simulations with Self-Interacting Dark Matter II: Halo Shapes vs. Observations*, Mon. Not. Roy. Astron. Soc. **430** (2013) 105, [arXiv:1208.3026].

[12] M. Rocha, A. H. G. Peter, J. S. Bullock, M. Kaplinghat, S. Garrison-Kimmel, J. Onorbe, and L. A. Moustakas, *Cosmological Simulations with Self-Interacting Dark Matter I: Constant Density Cores and Substructure*, Mon. Not. Roy. Astron. Soc. **430** (2013) 81–104, [arXiv:1208.3025].

[13] F. Kahlhoefer, K. Schmidt-Hoberg, M. T. Frandsen, and S. Sarkar, *Colliding clusters and dark matter self-interactions*, Mon. Not. Roy. Astron. Soc. **437** (2014), no. 3 2865–2881, [arXiv:1308.3419].

[14] D. Harvey, R. Massey, T. Kitching, A. Taylor, and E. Tittley, *The non-gravitational interactions of dark matter in colliding galaxy clusters*, Science **347** (2015) 1462–1465, [arXiv:1503.07675].

[15] M. Kaplinghat, S. Tulin, and H.-B. Yu, *Dark Matter Halos as Particle Colliders: Unified Solution to Small-Scale Structure Puzzles from Dwarfs to Clusters*, Phys. Rev. Lett. **116** (2016), no. 4 041302, [arXiv:1508.03339].

[16] L. Ackerman, M. R. Buckley, S. M. Carroll, and M. Kamionkowski, *Dark Matter and Dark Radiation*, Phys. Rev. **D79** (2009) 023519, [arXiv:0810.5126]. [,277(2008)].

[17] L. G. van den Aarssen, T. Bringmann, and C. Pfrommer, *Is dark matter with long-range interactions a solution to all small-scale problems of ΛCDM cosmology?*, Phys. Rev. Lett. **109** (2012) 231301, [arXiv:1205.5809].

[18] S. Tulin, H.-B. Yu, and K. M. Zurek, *Beyond Collisionless Dark Matter: Particle Physics Dynamics for Dark Matter Halo Structure*, Phys. Rev. **D87** (2013), no. 11 115007, [arXiv:1302.3898].

[19] M. Pospelov, A. Ritz, and M. B. Voloshin, *Secluded WIMP Dark Matter*, Phys. Lett. **B662** (2008) 53–61, [arXiv:0711.4866].

[20] M. Kaplinghat, S. Tulin, and H.-B. Yu, *Direct Detection Portals for Self-interacting Dark Matter*, Phys. Rev. **D89** (2014), no. 3 035009, [arXiv:1310.7945].

[21] E. Del Nobile, M. Kaplinghat, and H.-B. Yu, *Direct Detection Signatures of Self-Interacting Dark Matter with a Light Mediator*, *JCAP* **1510** (2015), no. 10 055, [arXiv:1507.04007].

[22] N. Bernal, X. Chu, C. Garcia-Cely, T. Hambye, and B. Zaldivar, *Production Regimes for Self-Interacting Dark Matter*, *JCAP* **1603** (2016), no. 03 018, [arXiv:1510.08063].

[23] F.-Y. Cyr-Racine, K. Sigurdson, J. Zavala, T. Bringmann, M. Vogelsberger, and C. Pfrommer, *ETHOSan effective theory of structure formation: From dark particle physics to the matter distribution of the Universe*, *Phys. Rev.* **D93** (2016), no. 12 123527, [arXiv:1512.05344].

[24] T. Bringmann, F. Kahlhoefer, K. Schmidt-Hoberg, and P. Walia, *Strong constraints on self-interacting dark matter with light mediators*, arXiv:1612.00845.

[25] A. Sommerfeld, *Über die Beugung und Bremsung der Elektronen*, *Annalen der Physik* **403** (1931) 207.

[26] N. Arkani-Hamed, D. P. Finkbeiner, T. R. Slatyer, and N. Weiner, *A Theory of Dark Matter*, *Phys. Rev.* **D79** (2009) 015014, [arXiv:0810.0713].

[27] K. Kainulainen, K. Tuominen, and V. Vaskonen, *Self-interacting dark matter and cosmology of a light scalar mediator*, *Phys. Rev.* **D93** (2016), no. 1 015016, [arXiv:1507.04931].

[28] **CRESST Collaboration**, G. Angloher *et. al.*, *Results on light dark matter particles with a low-threshold CRESST-II detector*, *Eur. Phys. J.* **C76** (2016), no. 1 25, [arXiv:1509.01515].

[29] **CRESST Collaboration**, G. Angloher *et. al.*, *Description of CRESST-II data*, arXiv:1701.08157.

[30] K. R. Dienes, J. Kumar, B. Thomas, and D. Yaylali, *Overcoming Velocity Suppression in Dark-Matter Direct-Detection Experiments*, *Phys. Rev.* **D90** (2014), no. 1 015012, [arXiv:1312.7772].

[31] M. J. Dolan, F. Kahlhoefer, C. McCabe, and K. Schmidt-Hoberg, *A taste of dark matter: Flavour constraints on pseudoscalar mediators*, *JHEP* **03** (2015) 171, [arXiv:1412.5174]. [Erratum: JHEP07,103(2015)].

[32] E. Ma, *Self-Interacting Dark Matter with Naturally Light Mediator*, *Mod. Phys. Lett.* **A32** (2017), no. 06 1750038, [arXiv:1608.08277].

[33] C. Arina, M. Backovic, J. Heisig, and M. Lucente, *Solar γ-rays as a Complementary Probe of Dark Matter*, arXiv:1703.08087.

[34] J. Abdallah *et. al.*, *Simplified Models for Dark Matter Searches at the LHC*, *Phys. Dark Univ.* **9-10** (2015) 8–23, [arXiv:1506.03116].

[35] N. F. Bell, Y. Cai, J. B. Dent, R. K. Leane, and T. J. Weiler, *Dark matter at the LHC: Effective field theories and gauge invariance*, *Phys. Rev.* **D92** (2015), no. 5 053008, [arXiv:1503.07874].

[36] C.-Y. Chen, H. Davoudiasl, W. J. Marciano, and C. Zhang, *Implications of a light dark Higgs solution to the g-2 discrepancy*, *Phys. Rev.* **D93** (2016), no. 3 035006, [arXiv:1511.04715].

[37] W. J. Marciano, A. Masiero, P. Paradisi, and M. Passera, *Contributions of axionlike particles to lepton dipole moments*, *Phys. Rev.* **D94** (2016), no. 11 115033, [arXiv:1607.01022].

[38] S. Mantry, M. Pitschmann, and M. J. Ramsey-Musolf, *Distinguishing axions from generic light scalars using electric dipole moment and fifth-force experiments*, *Phys. Rev.* **D90** (2014), no. 5 054016, [arXiv:1401.7339].

[39] I. Baldes and K. Petraki, *Asymmetric thermal-relic dark matter: Sommerfeld-enhanced freeze-out, annihilation signals and unitarity bounds*, arXiv:1703.00478.

[40] S. Cassel, *Sommerfeld factor for arbitrary partial wave processes*, *J. Phys.* **G37** (2010) 105009, [arXiv:0903.5307].

[41] R. Iengo, *Sommerfeld enhancement: General results from field theory diagrams*, *JHEP* **05** (2009) 024, [arXiv:0902.0688].

[42] T. R. Slatyer, *The Sommerfeld enhancement for dark matter with an excited state*, *JCAP* **1002** (2010) 028, [arXiv:0910.5713].

[43] J. A. Adams, S. Sarkar, and D. W. Sciama, *CMB anisotropy in the decaying neutrino cosmology*, *Mon. Not. Roy. Astron. Soc.* **301** (1998) 210–214, [astro-ph/9805108].

[44] X.-L. Chen and M. Kamionkowski, *Particle decays during the cosmic dark ages*, *Phys. Rev.* **D70** (2004) 043502, [astro-ph/0310473].

[45] T. R. Slatyer, N. Padmanabhan, and D. P. Finkbeiner, *CMB Constraints on WIMP Annihilation: Energy Absorption During the Recombination Epoch*, *Phys. Rev.* **D80** (2009) 043526, [arXiv:0906.1197].

[46] S. Galli, F. Iocco, G. Bertone, and A. Melchiorri, *CMB constraints on Dark Matter models with large annihilation cross-section*, *Phys. Rev.* **D80** (2009) 023505, [arXiv:0905.0003].

[47] J. M. Cline and P. Scott, *Dark Matter CMB Constraints and Likelihoods for Poor Particle Physicists*, *JCAP* **1303** (2013) 044, [arXiv:1301.5908]. [Erratum: JCAP1305,E01(2013)].

[48] H. Liu, T. R. Slatyer, and J. Zavala, *Contributions to cosmic reionization from dark matter annihilation and decay*, *Phys. Rev.* **D94** (2016), no. 6 063507, [arXiv:1604.02457].

[49] **Planck Collaboration**, P. A. R. Ade *et. al.*, *Planck 2015 results. XIII. Cosmological parameters*, *Astron. Astrophys.* **594** (2016) A13, [arXiv:1502.01589].

[50] G. Steigman, B. Dasgupta, and J. F. Beacom, *Precise Relic WIMP Abundance and its Impact on Searches for Dark Matter Annihilation*, *Phys. Rev.* **D86** (2012) 023506, [arXiv:1204.3622].

[51] B. von Harling and K. Petraki, *Bound-state formation for thermal relic dark matter and unitarity*, *JCAP* **1412** (2014) 033, [arXiv:1407.7874].

[52] M. Cirelli, P. Panci, K. Petraki, F. Sala, and M. Taoso, *Dark Matter's secret liaisons: phenomenology of a dark $U(1)$ sector with bound states*, arXiv:1612.07295.

[53] H. An, M. B. Wise, and Y. Zhang, *Strong CMB Constraint On P-Wave Annihilating Dark Matter*, arXiv:1606.02305.

[54] J. B. Dent, S. Dutta, and R. J. Scherrer, *Thermal Relic Abundances of Particles with Velocity-Dependent Interactions*, *Phys. Lett.* **B687** (2010) 275–279, [arXiv:0909.4128].

[55] J. Zavala, M. Vogelsberger, and S. D. M. White, *Relic density and CMB constraints on dark matter annihilation with Sommerfeld enhancement*, *Phys. Rev.* **D81** (2010) 083502, [arXiv:0910.5221].

[56] J. L. Feng, M. Kaplinghat, and H.-B. Yu, *Sommerfeld Enhancements for Thermal Relic Dark Matter*, *Phys. Rev.* **D82** (2010) 083525, [arXiv:1005.4678].

[57] L. G. van den Aarssen, T. Bringmann, and Y. C. Goedecke, *Thermal decoupling and the smallest subhalo mass in dark matter models with Sommerfeld-enhanced annihilation rates*, *Phys. Rev.* **D85** (2012) 123512, [`arXiv:1202.5456`].

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

[59] P. Agrawal, F.-Y. Cyr-Racine, L. Randall, and J. Scholtz, *Make Dark Matter Charged Again*, `arXiv:1610.04611`.

[60] F. Kahlhoefer, K. Schmidt-Hoberg, J. Kummer, and S. Sarkar, *On the interpretation of dark matter self-interactions in Abell 3827*, *Mon. Not. Roy. Astron. Soc.* **452** (2015), no. 1 L54–L58, [`arXiv:1504.06576`].

[61] D. Wittman, N. Golovich, and W. A. Dawson, *The Mismeasure of Mergers: Revised Limits on Self-interacting Dark Matter in Merging Galaxy Clusters*, `arXiv:1701.05877`.

[62] M. Vogelsberger, J. Zavala, C. Simpson, and A. Jenkins, *Dwarf galaxies in CDM and SIDM with baryons: observational probes of the nature of dark matter*, *Mon. Not. Roy. Astron. Soc.* **444** (2014), no. 4 3684–3698, [`arXiv:1405.5216`].

[63] L. Bergstrom, G. Bertone, T. Bringmann, J. Edsjo, and M. Taoso, *Gamma-ray and Radio Constraints of High Positron Rate Dark Matter Models Annihilating into New Light Particles*, *Phys. Rev.* **D79** (2009) 081303, [`arXiv:0812.3895`].

[64] J. Mardon, Y. Nomura, D. Stolarski, and J. Thaler, *Dark Matter Signals from Cascade Annihilations*, *JCAP* **0905** (2009) 016, [`arXiv:0901.2926`].

[65] S. Hannestad and T. Tram, *Sommerfeld Enhancement of DM Annihilation: Resonance Structure, Freeze-Out and CMB Spectral Bound*, *JCAP* **1101** (2011) 016, [`arXiv:1008.1511`].

[66] D. P. Finkbeiner, L. Goodenough, T. R. Slatyer, M. Vogelsberger, and N. Weiner, *Consistent Scenarios for Cosmic-Ray Excesses from Sommerfeld-Enhanced Dark Matter Annihilation*, *JCAP* **1105** (2011) 002, [`arXiv:1011.3082`].

[67] **AMS Collaboration**, L. Accardo *et. al.*, *High statistics measurement of the positron fraction in primary cosmic rays of 0.5˘500 gev with the alpha magnetic spectrometer on the international space station*, *Phys. Rev. Lett.* **113** (Sep, 2014) 121101.

[68] **AMS Collaboration**, M. Aguilar *et. al.*, *Electron and positron fluxes in primary cosmic rays measured with the alpha magnetic spectrometer on the international space station*, *Phys. Rev. Lett.* **113** (Sep, 2014) 121102.

[69] G. Elor, N. L. Rodd, T. R. Slatyer, and W. Xue, *Model-Independent Indirect Detection Constraints on Hidden Sector Dark Matter*, *JCAP* **1606** (2016), no. 06 024, [`arXiv:1511.08787`].

[70] T. R. Slatyer, *Indirect dark matter signatures in the cosmic dark ages. I. Generalizing the bound on s-wave dark matter annihilation from Planck results*, *Phys. Rev.* **D93** (2016), no. 2 023527, [`arXiv:1506.03811`].

[71] G. Krnjaic, *Probing Light Thermal Dark-Matter With a Higgs Portal Mediator*, *Phys. Rev.* **D94** (2016), no. 7 073009, [`arXiv:1512.04119`].

[72] S. Alekhin *et. al.*, *A facility to Search for Hidden Particles at the CERN SPS: the SHiP physics case*, *Rept. Prog. Phys.* **79** (2016), no. 12 124201, [`arXiv:1504.04855`].

[73] B. Batell, M. Pospelov, and A. Ritz, *Multi-lepton Signatures of a Hidden Sector in Rare B Decays*, *Phys. Rev.* **D83** (2011) 054005, [arXiv:0911.4938].

[74] M. Freytsis, Z. Ligeti, and J. Thaler, *Constraining the Axion Portal with $B \to Kl^+l^-$*, *Phys. Rev.* **D81** (2010) 034001, [arXiv:0911.5355].

[75] N. G. Deshpande, G. Eilam, and J. Jiang, *On the possibility of a new boson X0 (214-MeV) in $\Sigma^+ \to p\mu^+\mu^-$*, *Phys. Lett.* **B632** (2006) 212–214, [hep-ph/0509081].

[76] W. J. Marciano and Z. Parsa, *Rare kaon decays with missing energy*, *Phys. Rev.* **D53** (1996), no. 1 R1.

[77] **E787 Collaboration**, S. Adler *et. al.*, *Further search for the decay $K^+ \to \pi^+\nu\bar{\nu}$ in the momentum region P ¡ 195-MeV/c*, *Phys. Rev.* **D70** (2004) 037102, [hep-ex/0403034].

[78] **E949 Collaboration**, V. V. Anisimovsky *et. al.*, *Improved measurement of the $K^+ \to pi^+\nu\bar{\nu}$ branching ratio*, *Phys. Rev. Lett.* **93** (2004) 031801, [hep-ex/0403036].

[79] **BNL-E949 Collaboration**, A. V. Artamonov *et. al.*, *Study of the decay $K^+ \to \pi^+\nu\bar{\nu}$ in the momentum region $140 < P_\pi < 199$ MeV/c*, *Phys. Rev.* **D79** (2009) 092004, [arXiv:0903.0030].

[80] **NA62 Collaboration**, C. Lazzeroni *et. al.*, *Study of the $K^\pm \to \pi^\pm\gamma\gamma$ decay by the NA62 experiment*, *Phys. Lett.* **B732** (2014) 65–74, [arXiv:1402.4334].

[81] G. G. Raffelt and D. S. P. Dearborn, *Bounds on Hadronic Axions From Stellar Evolution*, *Phys. Rev.* **D36** (1987) 2211.

[82] J. M. Cline, K. Kainulainen, P. Scott, and C. Weniger, *Update on scalar singlet dark matter*, *Phys. Rev.* **D88** (2013) 055025, [arXiv:1306.4710]. [Erratum: Phys. Rev.D92,no.3,039906(2015)].

[83] **LUX Collaboration**, D. S. Akerib *et. al.*, *Improved Limits on Scattering of Weakly Interacting Massive Particles from Reanalysis of 2013 LUX Data*, *Phys. Rev. Lett.* **116** (2016), no. 16 161301, [arXiv:1512.03506].

[84] **LUX Collaboration**, D. S. Akerib *et. al.*, *Results from a search for dark matter in the complete LUX exposure*, *Phys. Rev. Lett.* **118** (2017), no. 2 021303, [arXiv:1608.07648].

[85] **PandaX Collaboration**, A. Tan *et al.*, *Dark Matter Results from First 98.7 Days of Data from the PandaX-II Experiment*, *Phys. Rev. Lett.* **117** (2016), no. 12 121303, [arXiv:1607.07400].

[86] **SuperCDMS**, R. Agnese *et. al.*, *New Results from the Search for Low-Mass Weakly Interacting Massive Particles with the CDMS Low Ionization Threshold Experiment*, *Phys. Rev. Lett.* **116** (2016), no. 7 071301, [arXiv:1509.02448].

[87] R. Catena, A. Ibarra, and S. Wild, *DAMA confronts null searches in the effective theory of dark matter-nucleon interactions*, *JCAP* **1605** (2016), no. 05 039, [arXiv:1602.04074].

[88] J. Fan, M. Reece, and L.-T. Wang, *Non-relativistic effective theory of dark matter direct detection*, *JCAP* **1011** (2010) 042, [arXiv:1008.1591].

[89] A. L. Fitzpatrick, W. Haxton, E. Katz, N. Lubbers, and Y. Xu, *The Effective Field Theory of Dark Matter Direct Detection*, *JCAP* **1302** (2013) 004, [arXiv:1203.3542].

[90] P. J. Fox and E. Poppitz, *Leptophilic Dark Matter*, *Phys. Rev.* **D79** (2009) 083528, [arXiv:0811.0399].

[91] S. Ipek, D. McKeen, and A. E. Nelson, *A Renormalizable Model for the Galactic Center Gamma Ray Excess from Dark Matter Annihilation*, Phys. Rev. **D90** (2014), no. 5 055021, [arXiv:1404.3716].

[92] J. M. No, *Looking through the pseudoscalar portal into dark matter: Novel mono-Higgs and mono-Z signatures at the LHC*, Phys. Rev. **D93** (2016), no. 3 031701, [arXiv:1509.01110].

[93] D. Goncalves, P. A. N. Machado, and J. M. No, *Simplified Models for Dark Matter Face their Consistent Completions*, Phys. Rev. **D95** (2017), no. 5 055027, [arXiv:1611.04593].

[94] M. Bauer, U. Haisch, and F. Kahlhoefer, *Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators*, arXiv:1701.07427.

[95] **ACME Collaboration**, J. Baron *et. al.*, *Order of Magnitude Smaller Limit on the Electric Dipole Moment of the Electron*, Science **343** (2014) 269–272, [arXiv:1310.7534].

[96] R. Daido and F. Takahashi, *The sign of the dipole-dipole potential by axion exchange*, arXiv:1704.00155.

[97] P. F. Bedaque, M. I. Buchoff, and R. K. Mishra, *Sommerfeld enhancement from Goldstone pseudo-scalar exchange*, JHEP **11** (2009) 046, [arXiv:0907.0235].

[98] B. Bellazzini, M. Cliche, and P. Tanedo, *Effective theory of self-interacting dark matter*, Phys. Rev. **D88** (2013), no. 8 083506, [arXiv:1307.1129].

[99] S. Oshima, T. Fujita, N. Kanda, and A. Yoshimi, *Two Pion versus σ-Meson Exchange Potentials*, Journal of Modern Physics **6** (2015) 927–936.

[100] M. Freytsis and Z. Ligeti, *On dark matter models with uniquely spin-dependent detection possibilities*, Phys. Rev. **D83** (2011) 115009, [arXiv:1012.5317].