

# Self-interacting dark matter with a stable vector mediator

**Michael Duerr, Kai Schmidt-Hoberg, Sebastian Wild**

Deutsches Elektronen-Synchrotron (DESY), Notkestrasse 85, D-22607 Hamburg, Germany

E-mail: [michael.duerr@desy.de](mailto:michael.duerr@desy.de), [kai.schmidt-hoberg@desy.de](mailto:kai.schmidt-hoberg@desy.de),  
[sebastian.wild@desy.de](mailto:sebastian.wild@desy.de)

**Abstract.** Light vector mediators can naturally induce velocity-dependent dark matter self-interactions while at the same time allowing for the correct dark matter relic abundance via thermal freeze-out. If these mediators subsequently decay into Standard Model states such as electrons or photons however, this is robustly excluded by constraints from the Cosmic Microwave Background. We study to what extent this conclusion can be circumvented if the vector mediator is stable and hence contributes to the dark matter density while annihilating into lighter degrees of freedom. We find viable parts of parameter space which lead to the desired self-interaction cross section of dark matter to address the small-scale problems of the collisionless cold dark matter paradigm while being compatible with bounds from the Cosmic Microwave Background and Big Bang Nucleosynthesis observations.

**Keywords:** Dark matter detectors, recombination, big bang nucleosynthesis

**Preprint numbers:** DESY 18-051

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>A simple model</b>	<b>3</b>
<b>3</b>	<b>Annihilation channels of dark matter and freeze-out calculation</b>	<b>4</b>
<b>4</b>	<b>Observational constraints</b>	<b>6</b>
4.1	Bounds on the decay of $h_D$ from CMB spectral distortions and BBN	6
4.2	CMB constraints on energy injection during recombination	8
4.3	Self-interactions of dark matter	9
<b>5</b>	<b>Results</b>	<b>10</b>
5.1	Impact of CMB constraints	10
5.2	Viability of significant dark matter self-interactions	12
<b>6</b>	<b>Conclusions</b>	<b>14</b>
<b>A</b>	<b>Full Lagrangian</b>	<b>15</b>
<b>B</b>	<b>Relic density calculation</b>	<b>16</b>

---

## 1 Introduction

Decades of experimental efforts aiming at a discovery of dark matter (DM) via its non-gravitational interactions with particles of the Standard Model (SM) have led to stringent constraints on such couplings, in particular for the popular class of weakly interacting massive particles (WIMPs) [1–3]. In contrast, DM self-interactions are largely unconstrained, potentially leading to significant changes in the astrophysical behaviour of DM [4]. In fact large DM self-interactions may even be desirable to address a number of discrepancies found in comparing  $N$ -body simulations of collisionless cold DM with astrophysical observations at small scales (for a recent review see [5]). In light of this, scenarios in which the DM dominantly couples to particles belonging to a *dark sector* have gained significant attention over the last years (see e.g. [6–9]). Interestingly, even a fully decoupled dark sector can lead to falsifiable predictions, e.g. to a change in the primordial abundances of elements produced during Big Bang Nucleosynthesis (BBN) [10, 11] or to changes in the Cosmic Microwave Background (CMB) [12, 13].

While large DM self-interactions at small relative velocities are required to address the small-scale problems, there exist rather strong constraints on the DM self-scattering cross section in high-velocity systems such as galaxy clusters [14–20]. A scattering cross section which increases towards smaller velocities is therefore preferred observationally. This behaviour is naturally achieved if a light scalar or vector particle mediates this interaction [20–27]. At the same time the DM relic abundance can naturally be set via thermal freeze-out of DM into these mediators.

However, in their simplest forms, these light mediator scenarios are under strong pressure from observations: a vector mediator  $Z_D$  leads to  $s$ -wave annihilation and if it predominantly decays into SM states such as electrons or photons, the energy injection from late-time annihilations  $\psi\bar{\psi} \rightarrow Z_D Z_D \rightarrow \text{SM}$  generically violates the stringent bounds obtained from the CMB [28, 29]. For a scalar mediator, on the other hand, the annihilation is  $p$ -wave suppressed such that bounds from the CMB are avoided. Nevertheless, strong bounds from direct detection experiments on the coupling to SM states imply late decays of the scalar, which in turn can spoil the successful predictions of standard BBN [30–32].

A number of possibilities to circumvent these bounds have been discussed for both the vector and scalar cases. To avoid constraints for the vector mediator one possibility is to have decays into light hidden sector states such as sterile neutrinos, which do not lead to reionisation. In such a setup where DM is converted to *dark* radiation, bounds from both BBN [11] as well as the CMB [13] can be avoided. Another option would be to have asymmetric DM [33] or to avoid thermalisation of the visible and hidden sectors, in which case freeze-in production [34] can set the relic abundance and constraints can be circumvented. Suppressing the scattering cross section relevant for direct detection allows to have viable models also for scalar mediators [35, 36].

In this work we study the possibility that the vector mediator  $Z_D$  is stable, in which case the annihilation process  $\psi\bar{\psi} \rightarrow Z_D Z_D$  obviously does not lead to energy injection during recombination. The stability can be achieved either by simply postulating that the kinetic mixing of  $Z_D$  with the SM gauge fields is highly suppressed, or in fact by demanding a dark charge conjugation symmetry [37]. However, in this minimal setup  $Z_D$  freezes out while still being relativistic and, being stable, would overclose the Universe.

Recently, it has been pointed out [37] that the abundance of a stable vector mediator  $Z_D$  could be sufficiently reduced via annihilations into a lighter state long after the freeze-out of  $\psi$ . In fact, there is a natural motivation to introduce one more particle in the dark sector: if  $Z_D$  obtains its mass from the breaking of a local  $U(1)$  symmetry, the theory contains a *dark Higgs boson*  $h_D$ , which (at least at tree-level) has a mass similar to the corresponding gauge boson. For  $m_{h_D} < m_{Z_D}$ , the annihilation  $Z_D Z_D \rightarrow h_D h_D$  can then suppress the late-time  $Z_D$  abundance, and for non-zero mixing between the SM and the dark Higgs boson the latter may decay before dominating the energy density of the Universe.

By construction, the CMB constraints arising from  $\psi\bar{\psi} \rightarrow Z_D Z_D$  are avoided; furthermore, the coupling structure of the theory does not permit the annihilation of  $\psi\bar{\psi}$  into a pair of (unstable) dark Higgs bosons  $h_D$  at tree-level. However, the presence of the annihilation channel  $\psi\bar{\psi} \rightarrow Z_D h_D$  with the subsequent decay of  $h_D$  still leads to the injection of SM energy into the CMB, and depending on the values of the different couplings involved, this potentially reintroduces the corresponding constraints. Furthermore, the late-time annihilation of the subdominant DM component  $Z_D$  into a pair of dark Higgs bosons can also leave its imprint on the CMB, which is actually well-known to be highly sensitive to even very small annihilation cross sections for DM particles with masses in the MeV range [38].

In light of these considerations, we perform a detailed and comprehensive study of the phenomenological viability of this scenario, i.e. a weak-scale DM particle  $\psi$  coupled to a stable vector mediator  $Z_D$ , which itself acts as a subdominant DM component. After describing the model in section 2, we discuss the relevant annihilation channels of the two DM species and the corresponding calculation of thermal freeze-out in section 3. In particular, we point out the importance of the conversion processes between the two DM species  $\psi$  and  $Z_D$ , influencing their cosmological abundances. In section 4, we first discuss bounds from CMB spectral

distortions and BBN on the late-time decay of the dark Higgs boson  $h_D$ , before examining the impact of the energy injection during recombination induced by the annihilations of  $\psi$  and  $Z_D$ . We present our results in section 5, where we pay special attention to the question whether it is possible to have sufficiently strong self-interactions of DM to resolve the small-scale problems mentioned previously, while being consistent with all constraints from the CMB and BBN. Finally, we conclude in section 6. Additional material can be found in appendices A and B.

## 2 A simple model

We extend the SM gauge group by a ‘dark’ gauge symmetry  $U(1)_D$ , and introduce a vector-like Dirac fermion  $\psi$  as well as a complex scalar  $\sigma$  charged under this new symmetry. These dark sector particles are singlets under the SM gauge group, and all SM fields are assumed to transform trivially under  $U(1)_D$ . The dark gauge symmetry is then spontaneously broken by a vacuum expectation value (vev) of  $\sigma$ , resulting in a massive dark gauge boson  $Z_D$  as well as a real scalar  $h_D$ .

More precisely, prior to symmetry breaking of the SM and dark gauge group, the Lagrangian of the model is given by

$$\mathcal{L} = \mathcal{L}_{\widetilde{\text{SM}}} + \mathcal{L}_D(\psi, Z_D^\mu, \sigma) - V(\sigma, \Phi), \quad (2.1)$$

with  $\mathcal{L}_{\widetilde{\text{SM}}}$  denoting the SM Lagrangian excluding the Higgs potential. The term containing the fermion and gauge boson interactions is given by

$$\mathcal{L}_D(\psi, Z_D^\mu, \sigma) = i\bar{\psi}\gamma_\mu D^\mu\psi - m_\psi\bar{\psi}\psi + (D^\mu\sigma)^*(D_\mu\sigma) - \frac{1}{4}F_D^{\mu\nu}F_{\mu\nu}^D, \quad (2.2)$$

where

$$D^\mu\psi = (\partial^\mu - ig_\psi Z_D^\mu)\psi, \quad (2.3)$$

$$D^\mu\sigma = (\partial^\mu - ig_D Z_D^\mu)\sigma, \quad (2.4)$$

$$F_D^{\mu\nu} = \partial^\mu Z_D^\nu - \partial^\nu Z_D^\mu. \quad (2.5)$$

The  $U(1)_D$  charges (times the gauge coupling)  $g_\psi$  and  $g_D$  of the fields  $\psi$  and  $\sigma$  will be treated as independent parameters of the model. Notice that the mass term of the vector-like fermion  $\psi$  is gauge invariant, and is thus already present prior to symmetry breaking.

Crucially, we have not included a kinetic mixing term  $\propto F_D^{\mu\nu}B_{\mu\nu}$  in eq. (2.2), where  $B_{\mu\nu}$  denotes the SM hypercharge field strength tensor. After the breaking of  $U(1)_D$  (see below), the presence of this term would allow the massive gauge boson  $Z_D$  to decay into SM states such as  $e^\pm$  pairs or photons; as already mentioned in the introduction and explained in more detail in section 4.2, basically all of the parameter space of the model would then be excluded due to constraints on energy injection from DM annihilations during recombination. From a purely phenomenological point of view, one can thus simply postulate that the dimensionless coupling parameter controlling the kinetic mixing is sufficiently small such that  $Z_D$  has a lifetime much larger than the age of the Universe. Notice that the choice of the kinetic mixing being exactly zero is actually stable under quantum corrections: there are no fermions in the model which are charged both under  $U(1)_D$  as well as under a SM gauge symmetry, and hence all loop-induced contributions to the mixing of  $Z_D$  with the SM gauge bosons vanish.

Alternatively, as pointed out recently in [37], the kinetic mixing term can be forbidden by imposing a *dark charge conjugation symmetry*, rendering  $Z_D$  absolutely stable (as long as  $m_{Z_D} < 2m_\psi$ ). In the same way as there is the familiar charge conjugation operator  $\mathcal{C}$  associated with the SM  $U(1)_{\text{em}}$  group, the dark charge conjugation operator  $\mathcal{C}_D$  changes the signs of the  $U(1)_D$  charges  $g_\psi$  and  $g_D$ , and furthermore replaces  $\sigma$  by  $\sigma^*$ ,  $Z_D^\mu$  by  $-Z_D^\mu$  as well as  $\psi$  by the charge-conjugated spinor  $\psi^C$ . If, in contrast to  $\mathcal{C}$ , nature is symmetric with respect to dark charge conjugation, the kinetic mixing operator  $F_D^{\mu\nu} B_{\mu\nu}$  is forbidden. Notice that this symmetry is still present after the spontaneous breaking of  $U(1)_D$  via a vev of  $\sigma$ .

Finally, in the Lagrangian given by eq. (2.1),  $V(\sigma, \Phi)$  denotes the most general scalar potential involving the SM singlet  $\sigma$  and the SM Higgs doublet  $\Phi$ :

$$V(\sigma, H) = -\mu_D^2 \sigma^* \sigma + \frac{1}{2} \lambda_D (\sigma^* \sigma)^2 - \mu_h^2 \Phi^\dagger \Phi + \frac{1}{2} \lambda_h (\Phi^\dagger \Phi)^2 + \lambda_{hD} (\sigma^* \sigma) (\Phi^\dagger \Phi). \quad (2.6)$$

After spontaneous breaking of the electroweak and dark gauge symmetry, the scalar fields can be parametrised in unitary gauge as

$$\sigma = (v_D + H_D)/\sqrt{2} \text{ and } \Phi = (0, (v_h + H)/\sqrt{2})^T. \quad (2.7)$$

In the following, we eliminate  $\lambda_D$  and  $\lambda_h$  from the scalar potential (2.6) by using  $v_h \simeq 246 \text{ GeV}$  and treating the dark Higgs vev  $v_D$  as a free parameter. For a given choice of the gauge coupling  $g_D$ , the latter is in one-to-one correspondence with the gauge boson mass  $m_{Z_D} = g_D v_D$ .

The presence of the portal term proportional to  $\lambda_{hD}$  in the scalar potential leads to a mixing of  $H_D$  and  $H$ ; we denote the corresponding mass eigenstates by  $h_D$  and  $h$ . Assuming  $\lambda_{hD} \ll 1$ ,  $m_{h_D} \ll m_h$ , the mixing angle is given by  $\theta \simeq \lambda_{hD} v_D v_h / m_h^2$ , where  $m_h \simeq 125 \text{ GeV}$  is the mass of the SM Higgs boson  $h$ . While in the absence of the kinetic mixing term  $Z_D$  is stable, the dark Higgs boson  $h_D$  can decay into SM particles with a rate proportional to  $\theta^2$ . Further details, in particular the full Lagrangian including the scalar potential after symmetry breaking can be found in appendix A.

For the purpose of our phenomenological analysis, a point in the parameter space of the model after symmetry breaking is then fully specified by the free parameters

$$m_{Z_D}, m_\psi, m_{h_D}, g_D, g_\psi, \lambda_{hD}. \quad (2.8)$$

Note that as long as the dimensionless couplings  $g_D$  and  $\lambda_D$  are of order one,  $m_{Z_D}$  and  $m_{h_D}$  are expected to be of the same order of magnitude. On the other hand, the tree-level mass of  $\psi$  is not related to the breaking of  $U(1)_D$ , and thus can be naturally at a different scale.

### 3 Annihilation channels of dark matter and freeze-out calculation

The scenario introduced in the previous section involves two stable neutral particles which contribute to the observed density of DM: the Dirac fermion  $\psi$  as well as the massive gauge boson  $Z_D$ . In the following, we discuss the main qualitative aspects of the freeze-out process of these DM particles; additional technical details of our numerical implementation can be found in appendix B.

We focus our analysis on regions in parameter space where  $m_{Z_D} \ll m_\psi$ : this is a necessary condition for obtaining a self-interaction cross section of  $\psi$  which is large enough

to lead to interesting astrophysical signatures. The heavy DM particle  $\psi$  can then self-annihilate via two possible channels:

$$\psi\bar{\psi} \rightarrow Z_D Z_D \quad \text{with } (\sigma v)_{\psi\bar{\psi} \rightarrow Z_D Z_D}^{\text{tree}} \simeq \frac{g_\psi^4}{16\pi m_\psi^2}, \quad (3.1)$$

$$\psi\bar{\psi} \rightarrow Z_D h_D \quad \text{with } (\sigma v)_{\psi\bar{\psi} \rightarrow Z_D h_D}^{\text{tree}} \simeq \frac{g_D^2 g_\psi^2}{64\pi m_\psi^2}, \quad (3.2)$$

where the tree-level expressions  $(\sigma v)^{\text{tree}}$  for the annihilation cross sections are given in the limit  $m_{Z_D} \ll m_\psi$  and  $v \ll 1$ . The annihilation of  $\psi\bar{\psi}$  into a pair of dark Higgs bosons, on the other hand, is strongly suppressed as it only proceeds via a one-loop diagram and furthermore vanishes in the  $s$ -wave limit  $v \rightarrow 0$ . In our numerical calculation, we take into account Sommerfeld enhancement in the annihilation processes (3.1) and (3.2), arising from the multiple exchange of  $Z_D$  bosons in the  $\psi\bar{\psi}$  initial state (see appendix B for details). Moreover, for  $m_{h_D} < m_{Z_D}$  the massive gauge boson  $Z_D$  can annihilate via

$$Z_D Z_D \rightarrow h_D h_D \quad \text{with } (\sigma v)_{Z_D Z_D \rightarrow h_D h_D} \simeq \frac{g_D^4 \sqrt{1-r} (44 - 20r + 9r^2 - 8r^3 + 2r^4)}{9\pi m_{Z_D}^2 (8 - 6r + r^2)^2}, \quad (3.3)$$

where  $r = m_{h_D}^2 / m_{Z_D}^2$ .<sup>1</sup>

At large temperatures, these annihilation processes lead to chemical equilibrium between the dark sector particles  $\psi$ ,  $Z_D$  and  $h_D$ . Furthermore, in the following we assume the portal coupling  $\lambda_{h_D}$  to be sufficiently large such that the initial temperatures of the dark and visible sectors are identical; the precise choice for  $\lambda_{h_D}$  will be discussed in more detail in section 4.1. The cosmological evolution of the DM particles  $\psi$  and  $Z_D$  down to smaller temperatures is then described by a set of two coupled Boltzmann equations for the number densities  $n_\psi$  and  $n_{Z_D}$ . As described in more detail in appendix B, we compute the present-day abundances  $\Omega_\psi h^2$  (defined to be the sum of the abundances of  $\psi$  and  $\bar{\psi}$ ) and  $\Omega_{Z_D} h^2$  by solving these equations numerically using a modified version of `MicrOMEGAs v4.3.5` [39, 40], additionally taking into account the Sommerfeld enhancement as well as thermal decoupling of the dark and visible sector at a temperature  $T_{\text{dec}}$ .

Qualitatively, the freeze-out process can be understood as follows: at  $T \simeq m_\psi/25$ , the annihilation processes given in eqs. (3.1) and (3.2) stop being efficient, and the heavy DM particle  $\psi$  freezes out, i.e.  $n_\psi/s$  becomes constant. However, the lighter DM particle  $Z_D$  remains in chemical equilibrium with the dark Higgs boson down to much smaller temperatures  $T \simeq m_{Z_D}/x_f$ , with  $x_f \simeq 15 - 50$ . The precise value of  $x_f$ , and thus the final abundance of  $Z_D$  depends on the strength of various annihilation channels: besides the usual self-annihilation  $Z_D Z_D \rightarrow h_D h_D$ , also processes involving the already frozen-out DM particle  $\psi$  have to be taken into account, leading to additional terms in the Boltzmann equation for  $n_{Z_D}$ . Concretely, these are the conversion process  $\psi Z_D \rightarrow \psi h_D$  as well as the annihilation channels  $\psi\bar{\psi} \rightarrow Z_D Z_D$  and  $\psi\bar{\psi} \rightarrow Z_D h_D$ . Notice that even though during the freeze-out of  $Z_D$  the latter processes are already too weak in order to keep  $\psi$  in equilibrium, they nevertheless can be important for the evolution of  $n_{Z_D}$ . A more detailed discussion of this point can be found in appendix B.

<sup>1</sup>This expression differs from the one given in Ref. [37].

## 4 Observational constraints

### 4.1 Bounds on the decay of $h_D$ from CMB spectral distortions and BBN

Being in thermal equilibrium with the SM heat bath at early times, the dark Higgs boson  $h_D$  generically has a significant abundance prior to its decay. As we are interested in a scenario with  $m_{h_D} < m_{Z_D} \lesssim 100 \text{ MeV}$ , it decays either dominantly into  $e^+e^-$  (for  $m_{h_D} > 2m_e$ ) or into  $\gamma\gamma$  (for  $m_{h_D} < 2m_e$ ). If these decay products are injected at redshifts  $z \lesssim 2 \times 10^6$ , they do not fully thermalise with the background photons, and thus lead to spectral distortions in the CMB [41–43]. In the context of our scenario, this excludes all regions of parameter space with  $\tau_{h_D} \gtrsim 10^5 \text{ s}$  [44].<sup>2</sup> Even for a scalar portal coupling  $\lambda_{h_D}$  of order one, this bound is generically violated if the dark Higgs has a mass below  $2m_e$  and thus can only decay into a pair of photons at one loop. As we still want to keep  $m_{h_D} < m_{Z_D}$  in order to allow for the annihilation process  $Z_D Z_D \rightarrow h_D h_D$  to deplete the abundance of  $Z_D$ , in the following we will fix  $m_{h_D} = 1.5 \text{ MeV} > 2m_e$ , and only consider vector boson masses  $m_{Z_D} \gtrsim 2 \text{ MeV}$ . Notice that as long as  $m_{Z_D} \gtrsim m_{h_D}$ , the precise value of the dark Higgs boson mass does not impact the phenomenology elsewhere, in particular neither the CMB constraints on energy injection from DM annihilation during recombination nor the self-interaction cross section of  $\psi$ . When presenting our results in section 5, we will indicate in which regions of parameter space the lifetime of  $h_D$  for this choice of  $m_{h_D}$  nevertheless exceeds  $10^5 \text{ s}$ , and is thus excluded by the constraints on CMB spectral distortions.

The decay of  $h_D$  in the early Universe is also constrained by the excellent agreement of the observed primordial abundances of light elements with the predictions from BBN. In general, BBN can be affected by additional stable or decaying particles present at temperatures  $T \lesssim 10 \text{ MeV}$  [45, 46]. More specifically, the scenario discussed in this work potentially modifies the primordial nuclear abundances in two ways:

- (i) If the dark Higgs  $h_D$  decays well after BBN, its electromagnetic decay products can photo-disintegrate nuclei, in particular deuterium and helium.
- (ii) If  $Z_D$  and/or  $h_D$  are still in thermal equilibrium at  $T \simeq 10 \text{ MeV}$ , they provide a contribution to  $\Delta N_{\text{eff}}$  and thus enhance the expansion rate during BBN.

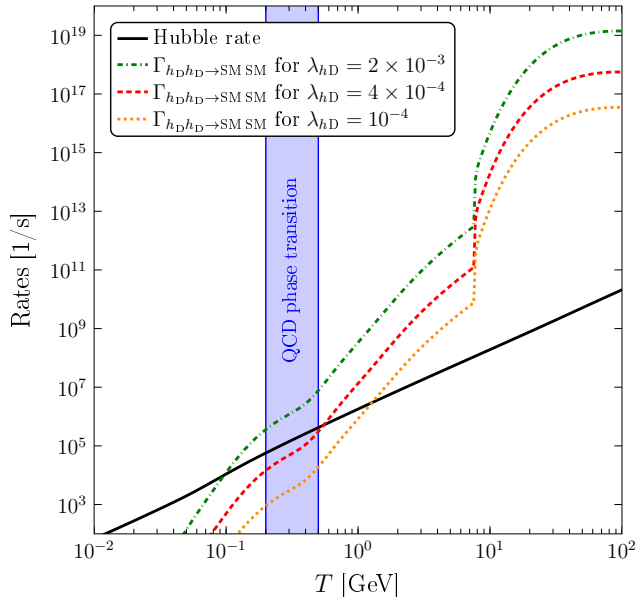
The first bound potentially constrains regions of parameter space where  $\tau_{h_D} \gtrsim 10^4 \text{ s}$ ; for smaller lifetimes, the cascade of the electromagnetic decay products caused by interactions with CMB photons leads to a cutoff of the corresponding photon spectrum below the photo-disintegration threshold  $E_{\text{dis}} = 2.2 \text{ MeV}$  of deuterium [47, 48]. However, for our choice  $m_{h_D} = 1.5 \text{ MeV}$  as motivated above from the constraints on CMB spectral distortions, the electromagnetic cascade induced by the electrons and positrons produced in the decay of  $h_D$  anyway only lead to photons with energies below  $E_\gamma \simeq 0.75 \text{ MeV}$ , which are unable to photo-disintegrate nuclei even for lifetimes  $\tau_{h_D} \gg 10^4 \text{ s}$ . Consequently, for our choice of  $m_{h_D}$ , the BBN bound (i) is automatically avoided.

The constraint (ii) from the increased Hubble rate during BBN depends critically on the temperature of the dark sector  $T_D$  at  $T \simeq 10 \text{ MeV}$ . The process most relevant for keeping the dark and visible sectors in thermal contact (leading to  $T_D = T$ ) is the annihilation of the dark

---

<sup>2</sup>This bound can be circumvented if the dark Higgs is stable on cosmological timescales and sufficiently light such that it does not contribute significantly to the present-day density of DM. In fact all CMB bounds from late time energy injection will be evaded in this case. In the following we do not further consider this part of the parameter space, and focus on the case where  $m_{h_D}$  and  $m_{Z_D}$  are of similar order of magnitude.





**Figure 1.** Reaction rate  $\Gamma_{h_D h_D \rightarrow \text{SM SM}}$  for different choices of  $\lambda_{h_D}$ . As explained in the text, choosing  $\lambda_{h_D} \lesssim 4 \times 10^{-4}$  leads to thermal decoupling of the dark and visible sectors prior to the QCD phase transition, and thus to a significantly reduced contribution of the dark sector particles to  $\Delta N_{\text{eff}}$ .

Higgs  $h_D$  into SM particles. The corresponding reaction rate  $\Gamma_{h_D h_D \rightarrow \text{SM SM}}(T)$  as a function of temperature is shown in Fig. 1 for different choices of the parameter  $\lambda_{h_D}$  appearing in the scalar potential (2.6). For  $T \gtrsim m_h/2$ , the dominant process establishing equilibrium is the production of an on-shell SM Higgs boson in the s-channel which, even for rather small values of  $\lambda_{h_D}$ , guarantees chemical equilibrium at these temperatures. For smaller  $T$ , this process gets exponentially suppressed and the annihilation rate  $\Gamma_{h_D h_D \rightarrow \text{SM SM}}$  rapidly decreases<sup>3</sup>, until eventually the dark and visible sectors decouple at a temperature  $T_{\text{dec}}$ , which we define via  $\Gamma_{h_D h_D \rightarrow \text{SM SM}}(T_{\text{dec}}) = H(T_{\text{dec}})$ . As can be seen from Fig. 1, by choosing  $\lambda_{h_D} \lesssim 4 \times 10^{-4}$ , this decoupling happens prior to the QCD phase transition, i.e.  $T_{\text{dec}} \gtrsim 500$  MeV. The visible sector is then heated with respect to the dark sector, reducing the relative contribution of the dark sector particles to the energy density. Quantitatively, the impact of  $Z_D$  and  $h_D$  on the Hubble rate during BBN can be parametrised in terms of the equivalent number of additional neutrino species:

$$\Delta N_{\text{eff}}(T \simeq 10 \text{ MeV}) \simeq 4 \cdot \left( \frac{g_{\text{SM}}(10 \text{ MeV})}{g_{\text{SM}}(T_{\text{dec}})} \right)^{4/3} \lesssim 0.27. \quad (4.1)$$

Here we conservatively assumed that both  $Z_D$  and  $h_D$  are relativistic degrees of freedom during BBN; for  $m_{Z_D} \gtrsim 10$  MeV the abundance of  $Z_D$  during BBN is already Boltzmann suppressed, and the contribution to  $\Delta N_{\text{eff}}$  is even smaller. Using the most recent information on the baryon-to-photon ratio inferred from the CMB as well as updated nuclear reaction rates, the upper limit on extra radiation during BBN is found to be  $\Delta N_{\text{eff}} < 0.2(0.36)$  at

<sup>3</sup>For  $T \lesssim 5$  GeV, the light SM quarks are no longer the appropriate degrees of freedom in the thermal bath. Following [49], in this regime the annihilation cross section for  $h_D h_D \rightarrow \text{SM SM}$  at a given center-of-mass energy  $\sqrt{s}$  can be expressed in terms of the width of a (hypothetical) scalar particle with mass  $m_* = \sqrt{s}$ , which in turn we take from [50].



$2\sigma(3\sigma)$  [45]. Given the significant impact of systematic uncertainties on deriving this limit, we conclude that the maximal contribution to  $\Delta N_{\text{eff}}$  predicted by our scenario, as given by eq. (4.1), might be in (mild) tension with BBN observations, but is certainly not robustly ruled out. A detailed analysis of BBN constraints on MeV-scale particles decaying into SM states, going beyond the simple estimate of  $\Delta N_{\text{eff}}$  via eq. (4.1) will appear elsewhere [32].

As outlined above, this conclusion holds as long as  $\lambda_{h_D} \lesssim 4 \times 10^{-4}$ , such that the dark and visible sectors decouple before the QCD phase transition. On the other hand, by choosing  $\lambda_{h_D}$  too small, the lifetime of the dark Higgs boson can get larger than  $\tau_{h_D} \gtrsim 10^5$  s, violating the bound from CMB spectral distortions as discussed at the beginning of this section. In order to weaken this constraint as much as possible, we fix  $\lambda_{h_D} = 4 \times 10^{-4}$  in the following, i.e. we choose the maximal value compatible with the constraint on the Hubble rate during BBN.<sup>4</sup>

## 4.2 CMB constraints on energy injection during recombination

The prime motivation for postulating the stability of  $Z_D$  has been to avoid the constraints arising from energy injection during recombination due to the annihilation process  $\psi\bar{\psi} \rightarrow Z_D Z_D$ . However, in our scenario the heavy DM particle can also annihilate via  $\psi\bar{\psi} \rightarrow Z_D h_D$ , potentially reintroducing the CMB constraints due to the subsequent decays of the dark Higgs boson into SM states. Moreover, also late-time annihilations  $Z_D Z_D \rightarrow h_D h_D$  lead to energy injection into the CMB, which, depending on the fraction of DM made up of  $Z_D$ , might also be in conflict with observations.

The annihilation cross section for  $\psi\bar{\psi} \rightarrow Z_D h_D$  during recombination is given by

$$(\sigma v)_{\psi\bar{\psi} \rightarrow Z_D h_D}^{\text{CMB}} \equiv S_s(v) \cdot (\sigma v)_{\psi\bar{\psi} \rightarrow Z_D h_D}^{\text{tree}}, \quad (4.2)$$

where  $(\sigma v)_{\psi\bar{\psi} \rightarrow Z_D h_D}^{\text{tree}}$  is the tree-level cross section given in eq. (3.2), and  $S_s(v)$  is the  $s$ -wave Sommerfeld enhancement factor corresponding to the multiple exchange of  $Z_D$  in the initial state, which is provided in eq. (B.1). The relative velocity  $v$  during recombination entering eq. (4.2) can be conservatively estimated to be bounded by [28]

$$v \lesssim 2 \times 10^{-7} \left( \frac{m_\psi}{100 \text{ GeV}} \right)^{-1/2}. \quad (4.3)$$

We have explicitly confirmed that the precise value of  $v$  does not affect our results as long as it satisfies this bound, since the Sommerfeld enhancement is already saturated for these velocities. A given point in parameter space is then excluded by CMB data if

$$\frac{1}{2} \cdot (\sigma v)_{\psi\bar{\psi} \rightarrow Z_D h_D}^{\text{CMB}} > (\sigma v)_{4e^\pm}^{(\text{upper bound})} (m_\psi) \cdot \left( \frac{\Omega_{\text{DM}} h^2}{\Omega_\psi h^2} \right)^2. \quad (4.4)$$

Here the factor  $1/2$  on the left hand side accounts for the fact that due to the stability of  $Z_D$  only half of the energy is transferred into electrons and positrons affecting reionisation. Furthermore,  $(\sigma v)_{4e^\pm}^{(\text{upper bound})} (m_\psi)$  is the upper bound on the annihilation cross section of DM

---

<sup>4</sup>This choice of  $\lambda_{h_D}$  leads to an invisible decay width  $\Gamma_{h \rightarrow h_D h_D} = \lambda_{h_D}^2 v_h^2 / (16\pi m_h) \simeq 3.8 \times 10^{-4} \times \Gamma_h^{\text{tot}}$  of the SM Higgs, which is well below the constraint from the latest LHC data [51]. Furthermore, depending on the vev  $v_D$  of the scalar field  $\sigma$ , the corresponding mixing angle  $\theta$  of the dark Higgs boson can be in the range where it might significantly alter the duration of the neutrino pulse from SN1987a [52, 53]. However, in view of the still large systematic uncertainties inherent in deriving the corresponding bounds, we do not consider them in the following discussion; a dedicated analysis of this point would certainly be interesting.

into a final state containing two electrons and two positrons, obtained under the assumption that  $\psi$  constitutes all of the observed DM. We take this bound as a function of  $m_\psi$  from [38], after multiplying it by a factor of two due to the Dirac nature of  $\psi$ . Finally, the last factor in eq. (4.4) takes into account the suppression of the bound if  $\psi$  does not constitute all of the observed DM, with  $\Omega_{\text{DM}}h^2 \simeq 0.12$  being the total DM abundance [54].

Similarly, the energy injection during recombination due to annihilations of  $Z_{\text{D}}$  excludes parts of the parameter space where

$$(\sigma v)_{Z_{\text{D}}Z_{\text{D}} \rightarrow h_{\text{D}}h_{\text{D}}} > (\sigma v)_{4e^\pm}^{(\text{upper bound})} (m_{Z_{\text{D}}}) \cdot \left( \frac{\Omega_{\text{DM}}h^2}{\Omega_{Z_{\text{D}}}h^2} \right)^2, \quad (4.5)$$

with  $(\sigma v)_{Z_{\text{D}}Z_{\text{D}} \rightarrow h_{\text{D}}h_{\text{D}}}$  given by eq. (3.3). Notice that in contrast to the self-annihilation of  $\psi$ , for the values of  $m_{Z_{\text{D}}}/m_{h_{\text{D}}}$  considered in this work this process is not subject to Sommerfeld enhancement, and the corresponding cross section can simply be evaluated in the limit  $v \rightarrow 0$ .

### 4.3 Self-interactions of dark matter

Via its coupling to the light mediator  $Z_{\text{D}}$ , the DM particle  $\psi$  can experience significant rates of self-scattering, even for weak couplings  $g_\psi \lesssim 1$  [23, 24]. This process can have important consequences for the distribution of DM in various astrophysical systems: it can transform cuspy profiles of DM halos into cored ones [55, 56] or more generally lead to a large diversity of DM profiles once baryonic effects are taken into account [57]. It may even lead to spectacular displacement signatures in merging galaxy clusters [18, 58, 59] if the scattering cross section is only mildly suppressed at large velocities (see [5] for a recent review on the subject).

For a large class of astrophysical objects, a good proxy for the impact of DM self-interactions is the momentum transfer cross section  $\sigma_{\text{T}}$ , defined via [18, 36]

$$\begin{aligned} \sigma_{\text{T}} &\equiv \frac{1}{2} \left( \sigma_{\text{T}}^{\psi\psi} + \sigma_{\text{T}}^{\psi\bar{\psi}} \right), \quad \text{with} \\ \sigma_{\text{T}}^{\psi\psi, \psi\bar{\psi}} &\equiv 2\pi \int_{-1}^1 \left( \frac{d\sigma}{d\Omega} \right)^{\psi\psi, \psi\bar{\psi}} (1 - |\cos \theta|) d \cos \theta. \end{aligned} \quad (4.6)$$

Here,  $(d\sigma/d\Omega)^{\psi\psi}$  and  $(d\sigma/d\Omega)^{\psi\bar{\psi}}$  denote the differential cross sections for elastic scattering of  $\psi\psi$  and  $\psi\bar{\psi}$ , respectively. We compute those by adapting the procedure outlined in [36] for DM interacting with a scalar mediator to the case of a vector mediator. In particular, we take into account non-perturbative effects related to multiple exchange of  $Z_{\text{D}}$  by solving the corresponding Schrödinger equation for a Yukawa-like scattering potential, properly taking into account the quantum indistinguishability of identical particles participating in the scattering process. For  $g_\psi^2 m_\psi / (4\pi m_{Z_{\text{D}}}) \ll 1$ , the non-perturbative effects are negligible and our results match the analytical expressions given in [36] for the Born regime (which are identical for scalar and vector mediators). On the other hand, for  $m_\psi v / m_{Z_{\text{D}}} \gtrsim 5$  solving the Schrödinger equation becomes not feasible, and we employ the results from [60] for the scattering cross section in the classical regime.

Crucially, in the regime where non-perturbative effects are important, the momentum transfer cross section  $\sigma_{\text{T}}$  typically is enhanced for small velocities  $v$  of the DM particles. Hence, one naturally expects larger effects of the DM self-scattering process in systems with small velocity dispersions such as dwarf galaxies (where  $v \simeq 30 \text{ km s}^{-1}$ ), and thus it is easier to satisfy the upper bounds on  $\sigma_{\text{T}}$  from observations of galaxy clusters (where  $v \simeq 1000 \text{ km s}^{-1}$ ). However, both the cross section required in order to transform cusps in dwarf galaxies into

cored profiles [17, 20, 27, 61, 62], as well as the largest value of  $\sigma_{\text{T}}/m_\psi$  compatible with constraints from merging galaxy clusters [15, 19, 63, 64] are still under debate. In light of this, and in order to bracket all of the potentially interesting range of momentum transfer cross sections at small scales, in section 5 we will show which regions in parameter space lead to  $0.1 \text{ cm}^2/\text{g} < \sigma_{\text{T}}/m_\psi < 10 \text{ cm}^2/\text{g}$  at  $v \simeq 30 \text{ km s}^{-1}$ , and use the rather conservative upper bound  $\sigma_{\text{T}}/m_\psi < 1 \text{ cm}^2/\text{g}$  at the scale of galaxy clusters,  $v \simeq 1000 \text{ km s}^{-1}$ .<sup>5</sup>

## 5 Results

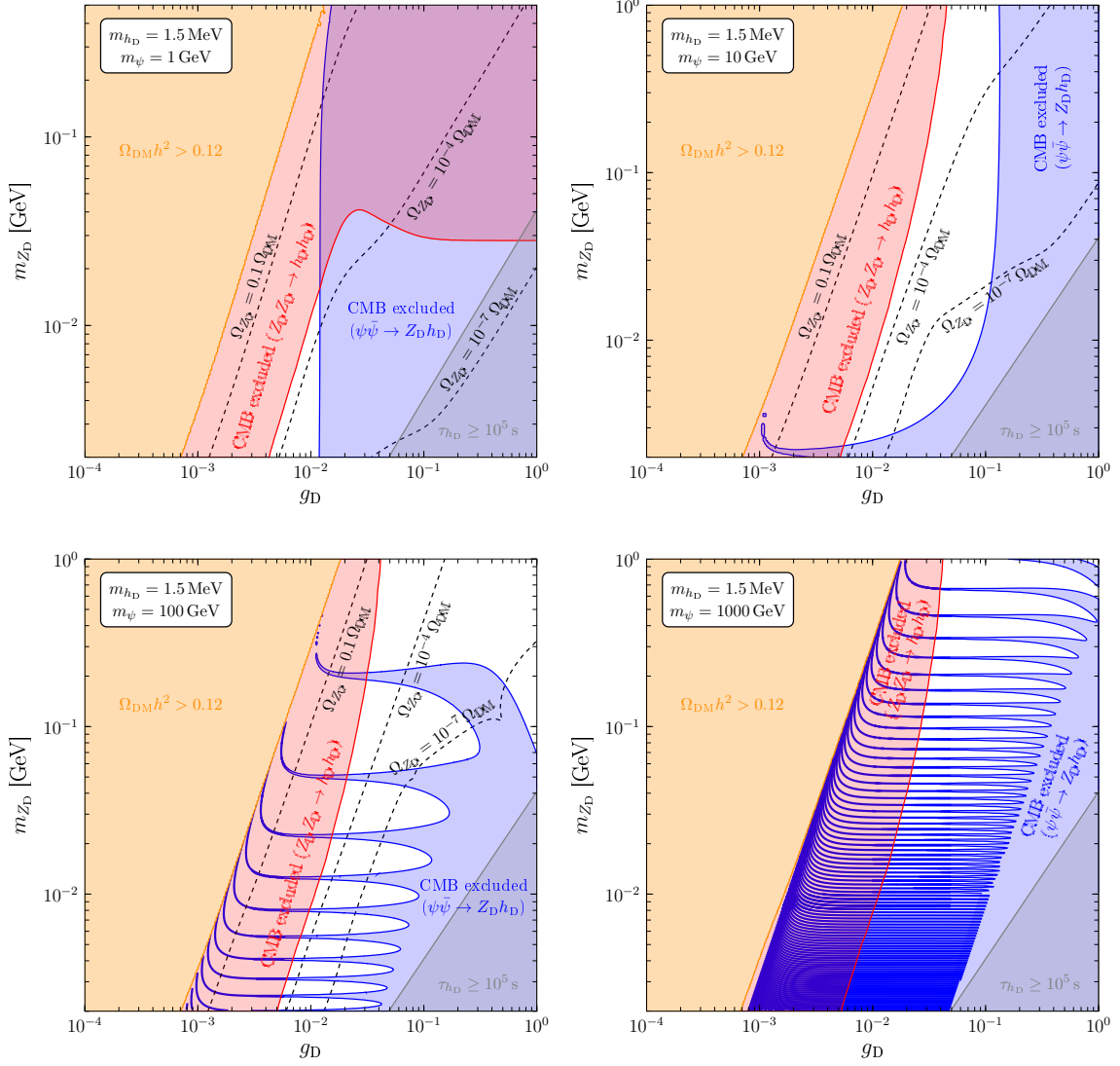
### 5.1 Impact of CMB constraints

The CMB constraints on energy injection during recombination as discussed in section 4.2 are illustrated in Fig. 2, where we show the parameter space spanned by the gauge coupling  $g_{\text{D}}$  and the light DM mass  $m_{Z_{\text{D}}}$  for different values of the mass of the heavy DM particle,  $m_\psi = 1, 10, 100$  and  $1000 \text{ GeV}$ . Following the discussion in section 4.1, in order to evade constraints from spectral distortions of the CMB as well as from BBN as much as possible, we fix the mass of the dark Higgs boson to  $m_{h_{\text{D}}} = 1.5 \text{ MeV}$ , with the precise value being irrelevant to the CMB constraints on energy injection during recombination. Notice that with this choice one has  $m_{h_{\text{D}}} < m_{Z_{\text{D}}}$  in all regions of parameter space shown in Fig. 2, as required for the annihilation channel  $Z_{\text{D}}Z_{\text{D}} \rightarrow h_{\text{D}}h_{\text{D}}$  to be kinematically allowed. Lastly, the gauge coupling  $g_\psi$  is fixed separately for each combination of  $g_{\text{D}}$ ,  $m_{Z_{\text{D}}}$  and  $m_\psi$  by the requirement that  $\psi$  and  $Z_{\text{D}}$  together account for all of the observed DM, i.e.  $\Omega_{\text{DM}}h^2 \equiv \Omega_\psi h^2 + \Omega_{Z_{\text{D}}}h^2 \simeq 0.12$ , following the discussion in section 3. The black dashed curves show contours of constant values of  $\Omega_{Z_{\text{D}}}h^2/\Omega_{\text{DM}}h^2$ , i.e. the fraction of DM composed of  $Z_{\text{D}}$ . This fraction grows towards smaller values of  $g_{\text{D}}$ , until at some point the cross section for  $Z_{\text{D}}Z_{\text{D}} \rightarrow h_{\text{D}}h_{\text{D}}$  [which scales as  $g_{\text{D}}^4$ , see eq. (3.3)] gets so small that irrespective of the choice of the gauge coupling  $g_\psi$  controlling the relic density of  $\psi$ , the abundance of  $Z_{\text{D}}$  alone overcloses the Universe. These regions of parameter space are shown as orange shaded in the different panels of Fig. 2.

In the blue shaded regions in Fig. 2, the energy injection from late-time annihilations  $\psi\bar{\psi} \rightarrow Z_{\text{D}}h_{\text{D}}$  is excluded by CMB data, as defined in eq. (4.4). Analogously, we show in red which parts of the parameter space are excluded by the CMB constraint on the annihilation process  $Z_{\text{D}}Z_{\text{D}} \rightarrow h_{\text{D}}h_{\text{D}}$ , c.f. eq. (4.5). Finally, the grey-shaded regions are excluded on the basis of the lifetime of the dark Higgs boson ( $\tau_{h_{\text{D}}} > 10^5 \text{ s}$ ), assuming  $\lambda_{h_{\text{D}}} = 4 \times 10^{-4}$  as discussed at the end of section 4.1.

For all values of  $m_\psi$  shown in the different panels of Fig. 2, we find that the annihilations from the heavy and light DM particle constrain complementary regions in parameter space: the energy injection induced by the annihilation process  $\psi\bar{\psi} \rightarrow Z_{\text{D}}h_{\text{D}}$  constrains regions of parameter space with larger values of the gauge coupling  $g_{\text{D}}$ , while the bound derived from the annihilation of the lighter DM candidate  $Z_{\text{D}}$  is most relevant for smaller  $g_{\text{D}}$ . This can be readily understood as follows: the annihilation cross section for  $\psi\bar{\psi} \rightarrow Z_{\text{D}}Z_{\text{D}}$  (which does not lead to constraints from the CMB) scales with  $g_\psi^4$ , while  $\psi\bar{\psi} \rightarrow Z_{\text{D}}h_{\text{D}}$  (which leads to the blue shaded exclusion regions in Fig. 2) is proportional to  $g_\psi^2 g_{\text{D}}^2$ . For sufficiently small values of  $g_{\text{D}}$ , the main annihilation channel of  $\psi$  both during freeze-out and recombination is then given by

<sup>5</sup>Both the preferred range for  $\sigma_{\text{T}}/m_\psi$  at small scales as well as the upper bound at scales of galaxy clusters have been derived assuming that all of the observed DM is self-interacting, while in our scenario  $Z_{\text{D}}$  does not experience significant self-interactions. However, as we will see in section 5, in all regions of the parameter space where the self-interaction cross section of  $\psi$  is within the range of interest, one has  $\Omega_{Z_{\text{D}}}h^2 \ll \Omega_\psi h^2 \simeq 0.12$ , and hence the astrophysical behaviour of DM is dominated by the properties of  $\psi$  alone.



**Figure 2.** Constraints in the  $g_{\text{D}}-m_{\text{Z}_{\text{D}}}$  plane for  $m_{\psi} = 1$  GeV (upper left panel),  $m_{\psi} = 10$  GeV (upper right panel),  $m_{\psi} = 100$  GeV (lower left panel), and  $m_{\psi} = 1000$  GeV (lower right panel). In each case we fix  $m_{h_{\text{D}}} = 1.5$  MeV, and choose the coupling  $g_{\psi}$  such that  $\Omega_{\psi} h^2 + \Omega_{\text{Z}_{\text{D}}} h^2 \simeq 0.12$ . In the orange shaded regions the DM density exceeds the observed value irrespective of the value of  $g_{\psi}$ . The regions of parameter space excluded by CMB constraints on late-time energy injection are given in blue for the process  $\psi\bar{\psi} \rightarrow \text{Z}_{\text{D}} h_{\text{D}}$  and in red for  $\text{Z}_{\text{D}} \text{Z}_{\text{D}} \rightarrow h_{\text{D}} h_{\text{D}}$ . In the grey shaded areas, the lifetime of the dark Higgs  $h_{\text{D}}$  exceeds  $10^5$  s, leading to significant spectral distortions in the CMB. Note that the range of  $m_{\text{Z}_{\text{D}}}$  shown for  $m_{\psi} = 1$  GeV (upper left plot) is smaller than in the rest of the panels.

the former process, and hence the CMB constraint from annihilations of  $\psi$  becomes less and less important. On the other hand, for small values of  $g_{\text{D}}$  the abundance of the lighter DM particle  $\text{Z}_{\text{D}}$  is dominantly set by the annihilation process  $\text{Z}_{\text{D}} \text{Z}_{\text{D}} \rightarrow h_{\text{D}} h_{\text{D}}$  (see appendix B), leading to  $\Omega_{\text{Z}_{\text{D}}} h^2 \propto g_{\text{D}}^{-4}$ . The corresponding bound from the CMB thus gets less important for larger values of  $g_{\text{D}}$ , as the suppression of the  $\text{Z}_{\text{D}}$  abundance overcompensates the rise of the cross section towards larger values of the coupling:  $(\Omega_{\text{Z}_{\text{D}}} h^2)^2 \times (\sigma v)_{\text{Z}_{\text{D}} \text{Z}_{\text{D}} \rightarrow h_{\text{D}} h_{\text{D}}} \propto g_{\text{D}}^{-8} \times g_{\text{D}}^4 = g_{\text{D}}^{-4}$ .

Interestingly, for all values of  $m_\psi$  considered in Fig. 2, there remains a region of intermediate values of  $g_D$  which is not constrained by either of the CMB constraints. Concretely, for  $m_\psi = 1$  GeV (upper left panel), couplings in the interval  $4 \times 10^{-3} \lesssim g_D \lesssim 10^{-2}$  are viable for  $m_{Z_D} \sim 2$  MeV, while all values of  $g_D$  are excluded for  $m_{Z_D} \gtrsim 20$  MeV. Note that for this value of  $m_\psi$ , the region excluded by the process  $Z_D Z_D \rightarrow h_D h_D$  becomes independent of the value of the coupling  $g_D$  for the largest  $Z_D$  masses shown. As discussed in more detail in appendix B, this is a result of additional annihilation channels significantly enhancing the abundance of  $Z_D$  in this region of parameter space.

For larger values of  $m_\psi$ , we start to observe that the CMB constraint from the annihilation of  $\psi$  reaches out to significantly smaller values of  $g_D$  for specific values of  $m_{Z_D}$ . This is due to the resonant Sommerfeld enhancement of the annihilation process  $\psi\bar{\psi} \rightarrow Z_D h_D$ , occurring for  $3g_\psi^2 m_\psi = 2\pi^3 n^2 m_{Z_D}$ , with  $n \in \mathbb{N}$  [65] (see appendix B for more details). The larger the mass ratio  $m_\psi/m_{Z_D}$ , the closer these resonances are in parameter space, which becomes particularly visible in the lower right panel of Fig. 2, corresponding to  $m_\psi = 1$  TeV. In this case the values of  $m_{Z_D}$  which are excluded or allowed by CMB constraints are extremely close to each other.<sup>6</sup> We also note that when  $g_D$  approaches the smallest value  $g_D^{(\min)}$  compatible with  $\Omega_{\text{DM}} h^2 = 0.12$ , the resonance peaks of the CMB constraint on  $\psi\bar{\psi} \rightarrow Z_D h_D$  bend upwards. This is because in the limit  $g_D \rightarrow g_D^{(\min)}$ , one has to lower the abundance of  $\psi$  to ever smaller values in order to match the total DM abundance, implying increasingly larger values of  $g_\psi$ . Thus, for fixed  $m_\psi$ , the resonance condition in this limit is satisfied for increasingly larger values of  $m_{Z_D}$ .

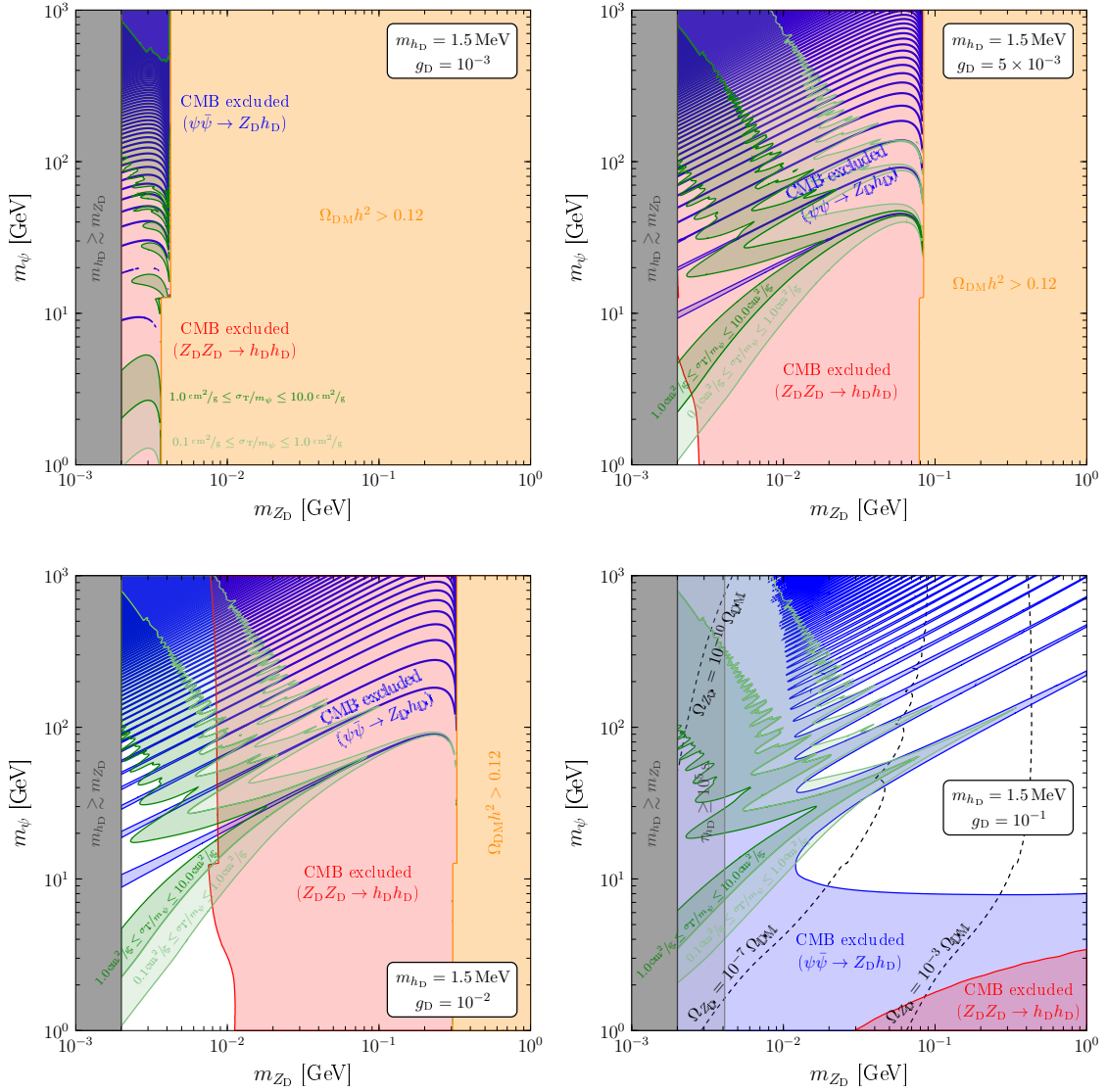
## 5.2 Viability of significant dark matter self-interactions

Finally, in Fig. 3 we present our results in the parameter space spanned by the masses  $m_{Z_D}$  and  $m_\psi$  of the two DM particles. From top left to bottom right, the four panels correspond to  $g_D = 10^{-3}$ ,  $5 \times 10^{-3}$ ,  $10^{-2}$  and  $10^{-1}$ . Again, we fix  $m_{h_D} = 1.5$  MeV and determine  $g_\psi$  in each point of the parameter space by requiring the total DM density to be equal to the observed value. As in Fig. 2, in the orange shaded regions the density of  $Z_D$  is so large that  $\Omega_{\text{DM}} h^2 > 0.12$  for all values of  $g_\psi$ . The blue and red shaded regions denote which combinations of parameters are excluded by the CMB constraint on energy injection from the annihilation of  $\psi$  and  $Z_D$ , respectively.<sup>7</sup> In addition, we show in light and dark green the regions of parameter space leading to a self-interaction cross section of  $\psi$  at the scale of dwarf galaxies in the range of  $0.1 \text{ cm}^2/\text{g} < \sigma_T/m_\psi < 1 \text{ cm}^2/\text{g}$  and  $1 \text{ cm}^2/\text{g} < \sigma_T/m_\psi < 10 \text{ cm}^2/\text{g}$ , respectively. As outlined in section 4.3, those values of  $\sigma_T/m_\psi$  can potentially address the shortcomings of collisionless cold DM at small scales. On the other hand, the bound  $\sigma_T/m_\psi \lesssim 1 \text{ cm}^2/\text{g}$  on the scale of galaxy clusters as discussed in section 4.3 is satisfied for the complete range of parameters shown in Fig. 3, and is thus not visible in the plots.

From the upper left panel of Fig. 3 (corresponding to  $g_D = 10^{-3}$ ) it follows that for sufficiently small values of  $g_D$ , all of the parameter space leading to the interesting range of DM self-interaction cross sections at the scale of dwarf galaxies is excluded by CMB

<sup>6</sup>We note that for parameter points precisely on top of one of the Sommerfeld resonances, the calculation of the DM relic abundance might be affected by late-time annihilations not taken into account in our analysis [66, 67].

<sup>7</sup>The small discontinuity of the orange and red shaded region at  $m_\psi \simeq 12$  GeV visible in some of the panels of Fig. 3 is an artefact of our approximate treatment of the impact of the chemical decoupling of the visible and dark sector on the relic density of  $Z_D$ , c.f. appendix B. A more precise treatment would lead to a smooth transition between the regions of different  $m_\psi$ , without affecting any of our conclusions.



**Figure 3.** Constraints and regions of significant DM self-interaction cross section in the  $m_{Z_D}$ - $m_\psi$  plane for  $g_D = 10^{-3}$  (upper left panel),  $g_D = 5 \times 10^{-3}$  (upper right panel),  $g_D = 10^{-2}$  (lower left panel) and  $g_D = 10^{-1}$  (lower right panel). The coupling  $g_\psi$  is fixed to reproduce the relic density where possible. As in Fig. 2, in the orange shaded regions one has  $\Omega_{\text{DM}} h^2 > 0.12$ , while the blue and red shaded regions indicate which parts of the parameter space are excluded by CMB constraints on energy injection from annihilation of  $\psi$  and  $Z_D$ , respectively. In addition, we show in light (dark) green the combination of parameters leading to a self-interaction cross section of  $\psi$  at the scale of dwarf galaxies in the range  $0.1 \text{ cm}^2/\text{g} < \sigma_T/m_\psi < 1 \text{ cm}^2/\text{g}$  ( $1 \text{ cm}^2/\text{g} < \sigma_T/m_\psi < 10 \text{ cm}^2/\text{g}$ ). The bound  $\sigma_T/m_\psi \lesssim 1 \text{ cm}^2/\text{g}$  on the scale of galaxy clusters is satisfied in all of the parameter space shown in this figure.

constraints on energy injection from the annihilation process  $Z_D Z_D \rightarrow h_D h_D$ . As already discussed above, this is a consequence of  $Z_D$  contributing in a non-negligible way to the observed amount of DM for small values of  $g_D$ ; concretely, we find  $\Omega_{Z_D}/\Omega_{\text{DM}} \gtrsim 0.2$  for  $g_D = 10^{-3}$ . On the other hand, the lower right panel of Fig. 3 corresponding to  $g_D = 0.1$  shows that if  $g_D$  is sufficiently large, the bounds from the annihilation of  $Z_D$  are irrelevant,



but then most of the parameter space leading to the desired values of the self-interaction cross section  $\sigma_{\text{T}}/m_{\psi}$  is excluded by CMB constraints arising from the annihilation process  $\psi\bar{\psi} \rightarrow Z_{\text{D}}h_{\text{D}}$ .

However, for intermediate values of the gauge coupling, such as  $g_{\text{D}} = 10^{-2}$  shown in the lower left panel of Fig. 3, we indeed find regions in parameter space leading to  $1 \text{ cm}^2/\text{g} < \sigma_{\text{T}}/m_{\psi} < 10 \text{ cm}^2/\text{g}$  on the scale of dwarf galaxies and  $\sigma_{\text{T}}/m_{\psi} < 1 \text{ cm}^2/\text{g}$  on the scale of galaxy clusters, while being consistent with the CMB bounds on the energy injection from the annihilation of  $\psi$  and  $Z_{\text{D}}$ . Concretely, for  $g_{\text{D}} = 10^{-2}$  this requires  $m_{Z_{\text{D}}} \lesssim 10 \text{ MeV}$ , as well as a combination of  $m_{\psi}$  and  $m_{Z_{\text{D}}}$  sufficiently far away from one of the resonances corresponding to the narrow blue shaded regions in the plot. Let us remark again that even though for large values of  $m_{\psi}$  the resonances are extremely dense in parameter space, the regions in between the resonance peaks are not excluded by CMB observations.

## 6 Conclusions

After years of theoretical and experimental efforts aiming at a better understanding of the astrophysical behaviour of DM at small scales, self-interacting DM remains one of the most compelling explanations for the apparent discrepancies found between observations and  $N$ -body simulations of collisionless cold DM. Realising the desired self-interaction cross section  $\sigma_{\text{T}}/m_{\psi} \simeq 1 \text{ cm}^2/\text{g}$  within a perturbative scenario of weak-scale DM requires the presence of a light mediator with a mass of  $(0.1 - 100) \text{ MeV}$ . However, two of the most basic incarnations of this general setup, a fermionic DM candidate coupled to an unstable scalar or vector mediator, are strongly disfavoured by the combination of data from direct detection experiments, CMB constraints on energy injection during recombination, as well as BBN constraints on late-time decaying particles.

In this article, we considered a scenario in which a *stable vector mediator*  $Z_{\text{D}}$  is responsible for the self-interactions of the fermionic DM particle  $\psi$ . This immediately saves the model from CMB constraints on the annihilation process  $\psi\bar{\psi} \rightarrow Z_{\text{D}}Z_{\text{D}}$ . In order to suppress the cosmological abundance of the vector mediator to a level compatible with observations, we have introduced one more particle in the dark sector, a dark Higgs boson  $h_{\text{D}}$  which is assumed to be lighter than  $Z_{\text{D}}$ . Besides being the natural by-product of the spontaneous breaking of a dark  $U(1)$  gauge symmetry giving rise to the mass of the vector mediator, we have shown that the annihilation  $Z_{\text{D}}Z_{\text{D}} \rightarrow h_{\text{D}}h_{\text{D}}$  can easily be efficient enough for  $Z_{\text{D}}$  to only constitute a subdominant fraction of the observed DM. However, also this setup is subject to constraints from the CMB: the annihilation processes  $\psi\bar{\psi} \rightarrow Z_{\text{D}}h_{\text{D}}$  as well as  $Z_{\text{D}}Z_{\text{D}} \rightarrow h_{\text{D}}h_{\text{D}}$  together with the subsequent decay of the dark Higgs can lead to significant energy injection during recombination. Interestingly, we find that these two processes constrain complementary parts of the model parameter space, with the former being important only for sufficiently large values of the dark gauge coupling  $g_{\text{D}}$  of the dark Higgs boson, and the latter for considerably smaller values. Most importantly, our results show that for a broad range of DM masses  $m_{\psi}$  and  $m_{Z_{\text{D}}}$ , intermediate values of the gauge coupling  $g_{\text{D}}$  ranging from  $\sim 5 \times 10^{-3}$  to  $\sim 10^{-1}$  are compatible with CMB constraints.

Furthermore, we have discussed the constraints arising from the late-time decays of the thermally produced dark Higgs bosons. In order to evade the stringent bounds from CMB spectral distortions, the dark Higgs has to decay with a lifetime  $\tau_{h_{\text{D}}} \lesssim 10^5 \text{ s}$ , implying a mass  $m_{h_{\text{D}}} > 2m_e$ . We have also discussed the possible impact of our scenario on the primordial abundances of light nuclei. For sufficiently small masses  $m_{h_{\text{D}}} \lesssim 4 \text{ MeV}$ , the decay products



of the dark Higgs are not energetic enough to photo-disintegrate even the most weakly bound nucleus (deuterium), and consequently there are no constraints from BBN arising from late-time changes of the nuclear abundances. In addition, by setting the scalar coupling which is responsible for the mixing of the dark and SM Higgs boson to a value below  $\simeq 4 \times 10^{-4}$ , the dark and visible sector thermally decouple before the QCD phase transition, leading to a suppressed value of  $\Delta N_{\text{eff}} \lesssim 0.27$  associated to the presence of  $Z_{\text{D}}$  and  $h_{\text{D}}$  in the thermal bath. Given all systematic uncertainties, this additional contribution to the energy density during BBN is still compatible with observations of primordial abundances.

Finally, we investigated whether the parts of parameter space which are compatible with all these constraints can lead to the range of desired values of the self-interaction cross section of DM at small scales. Indeed we find that for a gauge coupling  $g_{\text{D}} \simeq 10^{-2}$ , it is possible to obtain  $1 \text{ cm}^2/\text{g} \lesssim \sigma_{\text{T}}/m_{\psi} \lesssim 10 \text{ cm}^2/\text{g}$  at the scale of dwarf galaxies,  $\sigma_{\text{T}}/m_{\psi} \lesssim 1 \text{ cm}^2/\text{g}$  at the scale of galaxy clusters, while simultaneously being consistent with all CMB constraints on late-time energy injection as well as with BBN observations. In summary, our results thus show that if the scenario of DM interacting via an MeV-scale vector mediator is (minimally) extended by a dark Higgs boson breaking the dark gauge symmetry, it is indeed possible to restore the phenomenological viability of this setup in addressing the small-scale problems of the standard cold DM paradigm at small scales. Interestingly, the allowed range of parameters is already significantly narrowed down by current CMB and BBN observations, and could be further probed by future improvements of upper limits on the DM annihilation cross section at late times. In fact, the recent EDGES observation of an absorption feature in the 21 cm spectrum [68], if confirmed, might already be able to supersede the CMB constraints on the DM annihilation cross section [69, 70]. This leads to the exciting prospect that the idea of DM self-interactions induced by the exchange of a light vector mediator as discussed in this article can be conclusively tested in the near future.

## Acknowledgments

We thank Camilo Garcia-Cely for useful discussions and Felix Kahlhoefer for valuable comments on the manuscript. This work is supported by the German Science Foundation (DFG) under the Collaborative Research Center (SFB) 676 Particles, Strings and the Early Universe as well as the ERC Starting Grant ‘NewAve’ (638528).

## A Full Lagrangian

In this appendix, we provide details of the Lagrangian (2.1) after the breaking of the dark and SM gauge symmetries by means of eq. (2.7). The portal term  $\propto \lambda_{h_{\text{D}}}$  appearing in eq. (2.6) leads to a mixing of the scalar degrees of freedom  $H_{\text{D}}$  and  $H$ . We define the mass eigenstates  $h_{\text{D}}$  and  $h$  via

$$\begin{aligned} h_{\text{D}} &= -H \sin \theta + H_{\text{D}} \cos \theta, \\ h &= H \cos \theta + H_{\text{D}} \sin \theta, \end{aligned} \tag{A.1}$$

with the mixing angle  $\theta$  given by

$$\theta \simeq \lambda_{h_{\text{D}}} v_{\text{D}} v_h / m_h^2, \tag{A.2}$$

assuming  $\lambda_{h_D} \ll 1$  and  $m_{h_D} \ll m_h$ . Trading the parameters  $\mu_D$  and  $\mu_h$  of the Higgs potential for the physical masses  $m_{h_D}$  and  $m_h$ , and replacing  $v_D$  by  $m_{Z_D}/g_D$ , the scalar potential including only the leading terms in an expansion in  $\lambda_{h_D}$  reads

$$\begin{aligned}
V_{\text{broken}}(h, h_D) \simeq & \frac{1}{2}m_h^2 h^2 + \frac{m_h^2}{2v_h} h^3 + \frac{m_h^2}{8v_h^2} h^4 + \frac{1}{2}m_{h_D}^2 h_D^2 + \frac{g_D m_{h_D}^2}{2m_{Z_D}} h_D^3 + \frac{g_D^2 m_{h_D}^2}{8m_{Z_D}^2} h_D^4 \\
& + \frac{1}{2}\lambda_{h_D} v_h h h_D^2 - \frac{\lambda_{h_D} m_{Z_D}}{g_D} h^2 h_D + \frac{1}{4}\lambda_{h_D} h^2 h_D^2 \\
& + \frac{\lambda_{h_D} g_D v_h m_{h_D}^2}{2m_{Z_D} m_h^2} h h_D^3 - \frac{\lambda_{h_D} m_{Z_D}}{2g_D v_h} h^3 h_D. \tag{A.3}
\end{aligned}$$

The full Lagrangian after symmetry breaking is then finally given by

$$\begin{aligned}
\mathcal{L} \simeq & \mathcal{L}_{\text{SM}}|_{H \rightarrow h - \theta h_D} - \frac{1}{4}F_D^{\mu\nu} F_{\mu\nu}^D + \frac{1}{2}m_{Z_D}^2 Z_D^\mu Z_{D\mu} + i\bar{\psi}\gamma_\mu \partial^\mu \psi + g_\psi \bar{\psi}\gamma_\mu Z_D^\mu \psi - m_\psi \bar{\psi}\psi \\
& + g_D m_{Z_D} (h_D + \theta h) Z_D^\mu Z_{D\mu} + \frac{1}{2}g_D^2 (h_D + \theta h)^2 Z_D^\mu Z_{D\mu} \\
& + \frac{1}{2}(\partial^\mu h_D)(\partial_\mu h_D) - V_{\text{broken}}(h, h_D). \tag{A.4}
\end{aligned}$$

Notice that here we neglect the modifications proportional to  $\theta^2$  of couplings of  $h$  to SM fields. The couplings of  $h_D$  to the SM gauge bosons  $Z$  and  $W$ , as well as to the SM fermions  $f$  are given by

$$\begin{aligned}
\mathcal{L}_{\text{SM}}|_{H \rightarrow h - \theta h_D} \supset & \theta \left( \sum_f \frac{m_f}{v_h} \bar{f} f h_D \right) \\
& + \frac{\theta m_Z^2}{2v_h^2} \times (-2v_h h_D - 2h h_D + \theta h_D^2) (Z_\mu Z^\mu + 2 \cos^2 \theta_W W_\mu^+ W^{\mu-}), \tag{A.5}
\end{aligned}$$

with  $\theta_W$  the Weinberg angle.

## B Relic density calculation

In this appendix we describe in detail our method for calculating the relic abundances of the two DM particles  $\psi$  and  $Z_D$  for a given point in parameter space. In particular, we discuss the treatment of Sommerfeld enhancement during freeze-out, the importance of DM conversion and semi-annihilation processes, as well as the chemical decoupling of the dark and visible sector during or after DM freeze-out.

We implemented the Lagrangian of the model with `FeynRules v2.3.24` [71] and generated `CalcHEP` [72] model files to be imported into `MicrOMEGAs v4.3.5` [39, 40]. However, we find that due to the large mass hierarchy between the initial and final state particles, e.g. in the annihilation process  $\psi\bar{\psi} \rightarrow Z_D Z_D$ , the calculation of the annihilation cross sections using `CalcHEP` is facing numerical problems related to the polarisation sums over the light massive vector particles.<sup>8</sup> We therefore compute all relevant annihilation cross sections analytically and pass them to `MicrOMEGAs` for further use in the numerical solution of the Boltzmann equations. In doing so, we also take into account the Sommerfeld enhancement

<sup>8</sup>See appendix C.2 of Ref. [72] for a detailed discussion of this point.

in the annihilation processes  $\psi\bar{\psi} \rightarrow Z_D Z_D$  and  $\psi\bar{\psi} \rightarrow Z_D h_D$ , arising from the multiple exchange of the light vector boson  $Z_D$  in the initial state [73]. In practice, we compute the  $s$ - and  $p$ -wave contributions to the corresponding annihilation cross sections at tree level, and multiply them with enhancement factors  $S_s$  and  $S_p$ , respectively. Following [65, 74, 75], we approximate the Yukawa potential generated by the exchange of  $Z_D$  by a Hulthén potential, leading to

$$S_s = \frac{\pi}{a} \frac{\sinh(2\pi ac)}{\cosh(2\pi ac) - \cos(2\pi\sqrt{c - a^2 c^2})}, \quad (\text{B.1})$$

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

where  $a = 2\pi v/g_\psi^2$  and  $c = 3g_\psi^2 m_\psi / (2\pi^3 m_{Z_D})$ .

The Boltzmann equations for the number densities  $n_\psi$  and  $n_{Z_D}$  are then given by

$$\left( \frac{dn_\psi}{dt} + 3Hn_\psi \right) \Big|_{T \gg m_{Z_D}} \simeq - (\langle \sigma v \rangle_{\psi\bar{\psi} \rightarrow Z_D Z_D} + \langle \sigma v \rangle_{\psi\bar{\psi} \rightarrow Z_D h_D}) (n_\psi^2 - \bar{n}_\psi^2), \quad (\text{B.3})$$

$$\begin{aligned} \left( \frac{dn_{Z_D}}{dt} + 3Hn_{Z_D} \right) \Big|_{T \lesssim m_{Z_D} \ll m_\psi} &\simeq - \langle \sigma v \rangle_{Z_D Z_D \rightarrow h_D h_D} (n_{Z_D}^2 - \bar{n}_{Z_D}^2) \\ &+ \left( \langle \sigma v \rangle_{\psi\bar{\psi} \rightarrow Z_D Z_D} + \frac{1}{2} \langle \sigma v \rangle_{\psi\bar{\psi} \rightarrow Z_D h_D} \right) n_\psi^2 \\ &- \frac{1}{2} \langle \sigma v \rangle_{\psi Z_D \rightarrow \psi h_D} (n_{Z_D} - \bar{n}_{Z_D}) n_\psi, \end{aligned} \quad (\text{B.4})$$

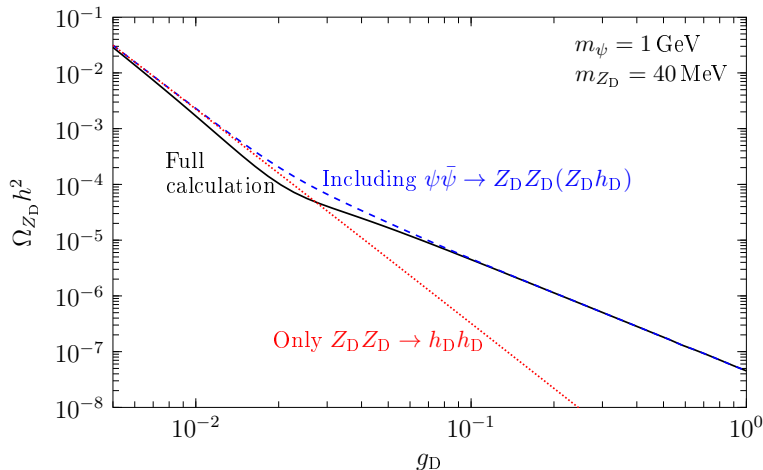
with  $\bar{n}_\psi$  and  $\bar{n}_{Z_D}$  denoting number densities in equilibrium, and  $H$  being the Hubble rate. For the sake of the following discussion, in these expressions (but not in our numerical calculation<sup>9</sup>) we have set  $\bar{n}_{Z_D} \simeq n_{Z_D}$  during freeze-out of  $\psi$ , as well as  $\bar{n}_\psi \simeq 0$  during the freeze-out process of  $Z_D$ . Under these assumptions, which are fulfilled to good accuracy as long as  $m_{Z_D} \ll m_\psi$ , the Boltzmann equation for  $n_\psi$  takes the same form as in the standard scenario of a single DM particle and can be solved independently of the evolution of  $n_{Z_D}$ .

On the other hand, the final abundance of the lighter DM particle  $Z_D$  can be significantly affected by the additional terms in eq. (B.4) involving the heavy DM particle  $\psi$ . For the case of the annihilation processes  $\psi\bar{\psi} \rightarrow Z_D Z_D$  and  $\psi\bar{\psi} \rightarrow Z_D h_D$ , this can be qualitatively understood by considering the ratio of the second and first term in the Boltzmann equation, evaluated at the temperature  $T_f$  where the annihilation process  $Z_D Z_D \rightarrow h_D h_D$  falls out of equilibrium:

$$\begin{aligned} \kappa_{\psi\bar{\psi}} &\equiv \frac{(\langle \sigma v \rangle_{\psi\bar{\psi} \rightarrow Z_D Z_D} + \frac{1}{2} \langle \sigma v \rangle_{\psi\bar{\psi} \rightarrow Z_D h_D}) \cdot n_\psi^2(T_f)}{\langle \sigma v \rangle_{Z_D Z_D \rightarrow h_D h_D} \cdot n_{Z_D}^2(T_f)} \\ &= \frac{(\langle \sigma v \rangle_{\psi\bar{\psi} \rightarrow Z_D Z_D} + \frac{1}{2} \langle \sigma v \rangle_{\psi\bar{\psi} \rightarrow Z_D h_D}) \cdot Y_\psi^2(T_f)}{\langle \sigma v \rangle_{Z_D Z_D \rightarrow h_D h_D} \cdot Y_{Z_D}^2(T_f)}, \end{aligned} \quad (\text{B.5})$$

where in the second line we replaced the number densities  $n$  by the yields  $Y = n/s$ . If  $\kappa_{\psi\bar{\psi}} \gtrsim 1$ , the standard calculation for the freeze-out of  $Z_D$  only taking into account the annihilation

<sup>9</sup>**MicrOMEGAS** solves the full Boltzmann equations in the temperature interval  $[T_{\text{start}}, T_{\text{end}}]$ . In order to make sure that the freeze-out of  $Z_D$  occurs within this range of temperatures even for the smallest values of  $m_{Z_D}$  considered in this work, we lower **Tend** from the default value  $10^{-3}$  GeV to  $10^{-6}$  GeV.



**Figure 4.** Relic abundance of  $Z_D$  as a function of  $g_D$ , assuming  $m_\psi = 1$  GeV and  $m_{Z_D} = 40$  MeV. The red dotted curve only takes into account the annihilation  $Z_D Z_D \rightarrow h_D h_D$ , the blue dashed curve in addition the self-annihilation of  $\psi$ , and the black solid curve corresponds to the full calculation including all terms of eq. (B.4).

process  $Z_D Z_D \rightarrow h_D h_D$  fails, as the residual annihilations of  $\psi$  contribute significantly to the production of  $Z_D$  around the freeze-out temperature  $T_f$ . In regions of parameter space where  $\psi$  is the dominant component of DM, the numerator of eq. (B.5) can be estimated by setting the total annihilation cross section of  $\psi$  to  $\langle\sigma v\rangle_{\text{thermal}} \simeq 4.4 \times 10^{-26} \text{ cm}^3 \text{ s}^{-1}$ , and the yield  $Y_\psi$  to the value corresponding to  $\Omega_\psi h^2 \simeq 0.12$ . Furthermore, an approximate expression for  $Y_{Z_D}(T_f)$  can be obtained from the semi-analytical solution to the standard one-particle Boltzmann equation [76]:

$$Y_{Z_D}(T_f) \simeq \frac{3.79}{\sqrt{g_\star} M_{\text{P}} m_{Z_D} \sigma_0} \log \left( \frac{0.11}{\sqrt{g_\star}} M_{\text{P}} m_{Z_D} \sigma_0 \right), \quad (\text{B.6})$$

where  $g_\star \simeq 10.75$  denotes the SM degrees of freedom at  $T_f$ ,  $M_{\text{P}}$  is the Planck mass, and  $\sigma_0 \equiv (\sigma v)_{Z_D Z_D \rightarrow h_D h_D}^{v \rightarrow 0}$ . Finally, after inserting the analytical expression for  $\sigma_0$  given in eq. (3.3) we obtain

$$\kappa_{\psi\bar{\psi}} \simeq 0.029 \cdot \left( \frac{g_D}{10^{-2}} \right)^4 \left( \frac{m_\psi}{\text{GeV}} \right)^{-2} \cdot \left( 1 + 0.16 \log \left[ \frac{g_D}{10^{-2}} \right] - 0.040 \log \left[ \frac{m_{Z_D}}{\text{MeV}} \right] \right)^{-2}, \quad (\text{B.7})$$

assuming  $m_{h_D} \ll m_{Z_D}$ . Clearly, for sufficiently large  $g_D$  and small  $m_\psi$  one has  $\kappa_{\psi\bar{\psi}} \gtrsim 1$ , indicating that the annihilation processes of the heavy DM particle  $\psi$  should indeed be taken into account in the calculation of the relic abundance of  $Z_D$ .

These simple analytical considerations are confirmed using our full numerical approach of solving the Boltzmann equation via `MicrOMEGAs`. In Fig. 4 we show the relic abundance of  $Z_D$  as a function of the coupling  $g_D$ , fixing for concreteness  $m_\psi = 1$  GeV and  $m_{Z_D} = 40$  MeV. The red dotted curve corresponds to a calculation where only the annihilation process  $Z_D Z_D \rightarrow h_D h_D$  is taken into account; as expected, the corresponding abundance scales as  $\Omega_{Z_D} h^2 \propto 1/\langle\sigma v\rangle_{Z_D Z_D \rightarrow h_D h_D} \propto g_D^{-4}$ . On the other hand, the blue dashed curve shows the abundance obtained by additionally including the terms in the Boltzmann equation accounting for the self-annihilation of  $\psi$ . The two calculations deviate significantly once  $g_D \gtrsim 2 \times 10^{-2}$ , well compatible with the simple estimate based on eq. (B.7). Lastly, the

solid black curve furthermore takes into account the conversion process  $\psi Z_{\text{D}} \rightarrow \psi h_{\text{D}}$ , which impacts the calculation mainly for intermediate values of  $g_{\text{D}}$ .

Finally, we take into account the impact of the thermal decoupling of the visible and dark sector on the abundances of  $\psi$  and  $Z_{\text{D}}$ . As explained in section 4.1, in order to evade the bounds from BBN and CMB spectral distortions as much as possible, we fix  $\lambda_{h_{\text{D}}} \simeq 4 \times 10^{-4}$  in our analysis. Then, as shown in Fig. 1, the dark and visible sector decouple at  $T_{\text{dec}} \simeq 500 \text{ MeV}$ . Assuming separate entropy conservation in both sectors for  $T < T_{\text{dec}}$ , the dark sector temperature  $T_{\text{D}}$  as a function of the photon temperature  $T$  evolves according to

$$\xi(T) \equiv \frac{T_{\text{D}}(T)}{T} = \left( \frac{g_{*S}(T)}{g_{*S}(T_{\text{dec}})} \frac{g_{*S}^{\text{D}}(T_{\text{dec}})}{g_{*S}^{\text{D}}(T_{\text{D}})} \right)^{\frac{1}{3}}, \quad (\text{B.8})$$

where  $g_{*S}(T)$  and  $g_{*S}^{\text{D}}(T_{\text{D}})$  denote the entropy degrees of freedom in the visible and dark sector at a given temperature. For the range of particle masses considered in our analysis,  $Z_{\text{D}}$  always freezes out after the decoupling of the two sectors, and so does  $\psi$  for  $m_{\psi} \lesssim 12 \text{ GeV}$ . Following [6], we take this into account by applying separate correction factors  $\xi(T_{\text{fo}})$  to the relic abundances of  $\psi$  and  $Z_{\text{D}}$  obtained from a calculation assuming equal temperatures in both sectors, where  $T_{\text{fo}}$  is the freeze-out temperature of  $\psi$  or  $Z_{\text{D}}$ , respectively. Note that we implicitly assume  $h_{\text{D}}$  to be a relativistic degree of freedom to ensure  $g_{*S}^{\text{D}}(T_{\text{D}}) > 0$ ; possible corrections to the abundance of  $Z_{\text{D}}$  in situations where all particles in the dark sector have become non-relativistic during freeze-out (see e.g. [77]) are left for future work.

## References

- [1] **XENON1T Collaboration**, E. Aprile et al., *First Dark Matter Search Results from the XENON1T Experiment*, *Phys. Rev. Lett.* **119** (2017) 181301, [[1705.06655](#)].
- [2] **Fermi-LAT Collaboration**, M. Ackermann et al., *Searching for Dark Matter Annihilation from Milky Way Dwarf Spheroidal Galaxies with Six Years of Fermi Large Area Telescope Data*, *Phys. Rev. Lett.* **115** (2015) 231301, [[1503.02641](#)].
- [3] **CMS Collaboration**, A. M. Sirunyan et al., *Search for dark matter produced with an energetic jet or a hadronically decaying W or Z boson at  $\sqrt{s} = 13 \text{ TeV}$* , *JHEP* **07** (2017) 014, [[1703.01651](#)].
- [4] D. N. Spergel and P. J. Steinhardt, *Observational evidence for selfinteracting cold dark matter*, *Phys. Rev. Lett.* **84** (2000) 3760, [[astro-ph/9909386](#)].
- [5] S. Tulin and H.-B. Yu, *Dark Matter Self-interactions and Small Scale Structure*, *Phys. Rept.* **730** (2018) 1, [[1705.02358](#)].
- [6] J. L. Feng, H. Tu, and H.-B. Yu, *Thermal Relics in Hidden Sectors*, *JCAP* **10** (2008) 043, [[0808.2318](#)].
- [7] R. Foot and S. Vagnozzi, *Dissipative hidden sector dark matter*, *Phys. Rev.* **D91** (2015) 023512, [[1409.7174](#)].
- [8] A. Berlin, D. Hooper, and G. Krnjaic, *Thermal Dark Matter From A Highly Decoupled Sector*, *Phys. Rev.* **D94** (2016) 095019, [[1609.02555](#)].
- [9] J. A. Evans, S. Gori, and J. Shelton, *Looking for the WIMP Next Door*, *JHEP* **02** (2018) 100, [[1712.03974](#)].
- [10] R. J. Scherrer and M. S. Turner, *Primordial Nucleosynthesis with Decaying Particles. 1. Entropy Producing Decays. 2. Inert Decays*, *Astrophys. J.* **331** (1988) 19.

- [11] M. Hufnagel, K. Schmidt-Hoberg, and S. Wild, *BBN constraints on MeV-scale dark sectors. Part I. Sterile decays*, *JCAP* **02** (2018) 044, [[1712.03972](#)].
- [12] V. Poulin, P. D. Serpico, and J. Lesgourgues, *A fresh look at linear cosmological constraints on a decaying dark matter component*, *JCAP* **08** (2016) 036, [[1606.02073](#)].
- [13] T. Bringmann, F. Kahlhoefer, K. Schmidt-Hoberg, and P. Walia, *Converting non-relativistic dark matter to radiation*, [1803.03644](#).
- [14] M. Markevitch, A. H. Gonzalez, D. Clowe, A. Vikhlinin, L. David, et al., *Direct constraints on the dark matter self-interaction cross-section from the merging galaxy cluster 1E0657-56*, *Astrophys. J.* **606** (2004) 819, [[astro-ph/0309303](#)].
- [15] 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, [[0704.0261](#)].
- [16] 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, [[1208.3026](#)].
- [17] M. Rocha, A. H. G. Peter, J. S. Bullock, M. Kaplinghat, S. Garrison-Kimmel, et al., *Cosmological Simulations with Self-Interacting Dark Matter I: Constant Density Cores and Substructure*, *Mon. Not. Roy. Astron. Soc.* **430** (2013) 81, [[1208.3025](#)].
- [18] 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) 2865, [[1308.3419](#)].
- [19] 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, [[1503.07675](#)].
- [20] 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) 041302, [[1508.03339](#)].
- [21] L. Ackerman, M. R. Buckley, S. M. Carroll, and M. Kamionkowski, *Dark Matter and Dark Radiation*, *Phys. Rev.* **D79** (2009) 023519, [[0810.5126](#)].
- [22] J. L. Feng, M. Kaplinghat, H. Tu, and H.-B. Yu, *Hidden Charged Dark Matter*, *JCAP* **07** (2009) 004, [[0905.3039](#)].
- [23] M. R. Buckley and P. J. Fox, *Dark Matter Self-Interactions and Light Force Carriers*, *Phys. Rev.* **D81** (2010) 083522, [[0911.3898](#)].
- [24] 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, [[0911.0422](#)].
- [25] A. Loeb and N. Weiner, *Cores in Dwarf Galaxies from Dark Matter with a Yukawa Potential*, *Phys. Rev. Lett.* **106** (2011) 171302, [[1011.6374](#)].
- [26] 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  $\Lambda$  CDM cosmology?*, *Phys. Rev. Lett.* **109** (2012) 231301, [[1205.5809](#)].
- [27] 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) 115007, [[1302.3898](#)].
- [28] T. Bringmann, F. Kahlhoefer, K. Schmidt-Hoberg, and P. Walia, *Strong constraints on self-interacting dark matter with light mediators*, *Phys. Rev. Lett.* **118** (2017) 141802, [[1612.00845](#)].
- [29] 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*, *JCAP* **05** (2017) 036, [[1612.07295](#)].



- [30] M. Kaplinghat, S. Tulin, and H.-B. Yu, *Direct Detection Portals for Self-interacting Dark Matter*, *Phys. Rev.* **D89** (2014) 035009, [[1310.7945](#)].
- [31] K. Kainulainen, K. Tuominen, and V. Vaskonen, *Self-interacting dark matter and cosmology of a light scalar mediator*, *Phys. Rev.* **D93** (2016) 015016, [[1507.04931](#)]. [Erratum: *Phys. Rev.* **D95** (2017) 079901].
- [32] M. Hufnagel, K. Schmidt-Hoberg, and S. Wild, *BBN constraints on MeV-scale dark sectors. Part II. Decays to Standard Model states, in preparation*.
- [33] I. Baldes, M. Cirelli, P. Panci, K. Petraki, F. Sala, et al., *Asymmetric dark matter: residual annihilations and self-interactions*, [1712.07489](#).
- [34] N. Bernal, X. Chu, C. Garcia-Cely, T. Hambye, and B. Zaldivar, *Production Regimes for Self-Interacting Dark Matter*, *JCAP* **03** (2016) 018, [[1510.08063](#)].
- [35] M. Blennow, S. Clementz, and J. Herrero-Garcia, *Self-interacting inelastic dark matter: A viable solution to the small scale structure problems*, *JCAP* **03** (2017) 048, [[1612.06681](#)].
- [36] F. Kahlhoefer, K. Schmidt-Hoberg, and S. Wild, *Dark matter self-interactions from a general spin-0 mediator*, *JCAP* **08** (2017) 003, [[1704.02149](#)].
- [37] E. Ma, *Inception of Self-Interacting Dark Matter with Dark Charge Conjugation Symmetry*, *Phys. Lett.* **B772** (2017) 442, [[1704.04666](#)].
- [38] 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) 023527, [[1506.03811](#)].
- [39] G. Belanger, F. Boudjema, A. Pukhov, and A. Semenov, *MicrOMEGAs 2.0: A Program to calculate the relic density of dark matter in a generic model*, *Comput. Phys. Commun.* **176** (2007) 367, [[hep-ph/0607059](#)].
- [40] G. Belanger, F. Boudjema, A. Pukhov, and A. Semenov, *micrOMEGAs4.1: two dark matter candidates*, *Comput. Phys. Commun.* **192** (2015) 322, [[1407.6129](#)].
- [41] Ya. B. Zeldovich and R. A. Sunyaev, *The Interaction of Matter and Radiation in a Hot-Model Universe*, *Astrophys. Space Sci.* **4** (1969) 301.
- [42] W. Hu and J. Silk, *Thermalization constraints and spectral distortions for massive unstable relic particles*, *Phys. Rev. Lett.* **70** (1993) 2661.
- [43] J. Chluba and R. A. Sunyaev, *The evolution of CMB spectral distortions in the early Universe*, *Mon. Not. Roy. Astron. Soc.* **419** (2012) 1294, [[1109.6552](#)].
- [44] V. Poulin, J. Lesgourgues, and P. D. Serpico, *Cosmological constraints on exotic injection of electromagnetic energy*, *JCAP* **03** (2017) 043, [[1610.10051](#)].
- [45] R. H. Cyburt, B. D. Fields, K. A. Olive, and T.-H. Yeh, *Big Bang Nucleosynthesis: 2015*, *Rev. Mod. Phys.* **88** (2016) 015004, [[1505.01076](#)].
- [46] **Particle Data Group**, C. Patrignani et al., *Review of Particle Physics*, *Chin. Phys.* **C40** (2016) 100001.
- [47] K. Jedamzik, *Big bang nucleosynthesis constraints on hadronically and electromagnetically decaying relic neutral particles*, *Phys. Rev.* **D74** (2006) 103509, [[hep-ph/0604251](#)].
- [48] J. Berger, K. Jedamzik, and D. G. E. Walker, *Cosmological Constraints on Decoupled Dark Photons and Dark Higgs*, *JCAP* **11** (2016) 032, [[1605.07195](#)].
- [49] J. M. Cline, K. Kainulainen, P. Scott, and C. Weniger, *Update on scalar singlet dark matter*, *Phys. Rev.* **D88** (2013) 055025, [[1306.4710](#)]. [Erratum: *Phys. Rev.* **D92**,no.3,039906(2015)].
- [50] 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, [[1504.04855](#)].



- [51] **CMS Collaboration**, V. Khachatryan et al., *Searches for invisible decays of the Higgs boson in pp collisions at  $\sqrt{s} = 7, 8,$  and  $13$  TeV*, *JHEP* **02** (2017) 135, [[1610.09218](#)].
- [52] G. G. Raffelt and D. S. P. Dearborn, *Bounds on Hadronic Axions From Stellar Evolution*, *Phys. Rev.* **D36** (1987) 2211.
- [53] G. Krnjaic, *Probing Light Thermal Dark-Matter With a Higgs Portal Mediator*, *Phys. Rev.* **D94** (2016) 073009, [[1512.04119](#)].
- [54] **Planck Collaboration**, P. A. R. Ade et al., *Planck 2015 results. XIII. Cosmological parameters*, *Astron. Astrophys.* **594** (2016) A13, [[1502.01589](#)].
- [55] N. Yoshida, V. Springel, S. D. M. White, and G. Tormen, *Weakly self-interacting dark matter and the structure of dark halos*, *Astrophys. J.* **544** (2000) L87, [[astro-ph/0006134](#)].
- [56] R. Dave, D. N. Spergel, P. J. Steinhardt, and B. D. Wandelt, *Halo properties in cosmological simulations of selfinteracting cold dark matter*, *Astrophys. J.* **547** (2001) 574, [[astro-ph/0006218](#)].
- [57] A. Kamada, M. Kaplinghat, A. B. Pace, and H.-B. Yu, *How the Self-Interacting Dark Matter Model Explains the Diverse Galactic Rotation Curves*, *Phys. Rev. Lett.* **119** (2017) 111102, [[1611.02716](#)].
- [58] L. L. R. Williams and P. Saha, *Light/Mass Offsets in the Lensing Cluster Abell 3827: Evidence for Collisional Dark Matter?*, *Mon. Not. Roy. Astron. Soc.* **415** (2011) 448, [[1102.3943](#)].
- [59] W. A. Dawson et al., *Discovery of a Dissociative Galaxy Cluster Merger with Large Physical Separation*, *Astrophys. J.* **747** (2012) L42, [[1110.4391](#)].
- [60] F.-Y. Cyr-Racine, K. Sigurdson, J. Zavala, T. Bringmann, M. Vogelsberger, et al., *ETHOSan effective theory of structure formation: From dark particle physics to the matter distribution of the Universe*, *Phys. Rev.* **D93** (2016) 123527, [[1512.05344](#)].
- [61] M. Vogelsberger, J. Zavala, and A. Loeb, *Subhaloes in Self-Interacting Galactic Dark Matter Haloes*, *Mon. Not. Roy. Astron. Soc.* **423** (2012) 3740, [[1201.5892](#)].
- [62] J. Zavala, M. Vogelsberger, and M. G. Walker, *Constraining Self-Interacting Dark Matter with the Milky Way’s dwarf spheroidals*, *Mon. Not. Roy. Astron. Soc.* **431** (2013) L20, [[1211.6426](#)].
- [63] 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) L54, [[1504.06576](#)].
- [64] D. Wittman, N. Golovich, and W. A. Dawson, *The Mismeasure of Mergers: Revised Limits on Self-interacting Dark Matter in Merging Galaxy Clusters*, [[1701.05877](#)].
- [65] S. Cassel, *Sommerfeld factor for arbitrary partial wave processes*, *J. Phys.* **G37** (2010) 105009, [[0903.5307](#)].
- [66] 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, [[1202.5456](#)].
- [67] T. Binder, M. Gustafsson, A. Kamada, S. M. R. Sandner, and M. Wiesner, *Reannihilation of self-interacting dark matter*, [[1712.01246](#)].
- [68] J. D. Bowman, A. E. E. Rogers, R. A. Monsalve, T. J. Mozdzen, and N. Mahesh, *An absorption profile centred at 78 megahertz in the sky-averaged spectrum*, *Nature* **555** (2018) 67.
- [69] G. D’Amico, P. Panci, and A. Strumia, *Bounds on Dark Matter annihilations from 21 cm data*, [[1803.03629](#)].
- [70] H. Liu and T. R. Slatyer, *Too Hot, Too Cold or Just Right? Implications of a 21-cm Signal for Dark Matter Annihilation and Decay*, [[1803.09739](#)].

- [71] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks, *FeynRules 2.0 - A complete toolbox for tree-level phenomenology*, *Comput. Phys. Commun.* **185** (2014) 2250, [[1310.1921](#)].
- [72] A. Belyaev, N. D. Christensen, and A. Pukhov, *CalcHEP 3.4 for collider physics within and beyond the Standard Model*, *Comput. Phys. Commun.* **184** (2013) 1729, [[1207.6082](#)].
- [73] A. Sommerfeld, *Über die Beugung und Bremsung der Elektronen*, *Annalen der Physik* **403** (1931) 207.
- [74] R. Iengo, *Sommerfeld enhancement: General results from field theory diagrams*, *JHEP* **05** (2009) 024, [[0902.0688](#)].
- [75] T. R. Slatyer, *The Sommerfeld enhancement for dark matter with an excited state*, *JCAP* **02** (2010) 028, [[0910.5713](#)].
- [76] E. W. Kolb and M. S. Turner, *The Early Universe*, *Front. Phys.* **69** (1990) 1.
- [77] D. Pappadopulo, J. T. Ruderman, and G. Trevisan, *Dark matter freeze-out in a nonrelativistic sector*, *Phys. Rev.* **D94** (2016) 035005, [[1602.04219](#)].