Title

Astrophysical and Phenomenological Implications of Bound States in Extensions of the Standard Model of Particle Physics

Author

Wozniak Pearce, Lauren Marie

Publication Date

2014

Peer reviewed|Thesis/dissertation

University of California
Los Angeles

Astrophysical and Phenomenological Implications of Bound States in Extensions of the Standard Model of Particle Physics

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Physics

by

Lauren Marie Wozniak Pearce

2014

Abstract of the Dissertation

Astrophysical and Phenomenological Implications of Bound States in Extensions of the Standard Model of Particle Physics

by

Lauren Marie Wozniak Pearce

Doctor of Philosophy in Physics

University of California, Los Angeles, 2014

Professor Alexander Kusenko, Chair

While the Standard Model of particle physics has undoubtedly been an experimental success, several questions remain unresolved. In particular, the Standard Model cannot account for the observed cosmological preference for matter over dark matter, nor does it provide a viable candidate for dark matter. This motivates us to consider extensions to the Standard Model; in this thesis, we will focus on several extensions of the Standard Model in which the formation of bound states is a significant factor. We will argue that the formation of bound states produces new phenomena that can address these unsettled questions.

First, we consider a strongly-coupled version of the Minimal Supersymmetric Standard Model. We demonstrate that in this model, electroweak symmetry breaking may be triggered by the presence of squark bound states which mix with the fundamental Higgs boson. Next, we show that this model has a viable phenomenology (e.g., it does not have large flavor-changing-neutral-currents or break $SU_C(3)$ symmetry). Additionally, this strongly-coupled version of the MSSM can relatively easily accommodate electroweak scale baryogenesis.

Following this, we turn our attention to the possibility of dark matter bound states in asymmetric dark matter models. We first consider a simplistic scalar model and demonstrate that bound state formation can produce a detectable gamma ray excess in certain regions

of parameter space. This signal is produced through the decay of the dark force mediator whose emission necessarily accompanies bound state formation. Next, we consider models in which the dark matter self-interactions are described by a broken $U_D(1)$ gauge group. We argue that in such models dark matter is generically multi-component, consisting of two species of ions along with dark atoms. We then investigate the possibility of using these self-interactions between the different species to alleviate tension between the cold dark matter paradigm and observations of dwarf galaxies, while retaining the ellipticity of larger halos.

Finally, we consider the formation and growth of Q-balls (non-topological solitons) in a simplified model inspired by the MSSM. In particular models, Q-balls can trigger a phase transition once they reach a critical size. In certain regions of parameter space, small charge Q-balls can be approximated using the Bethe-Salpeter equation. This allows us to study the growth of small Q-balls; by joining this to the semi-classical regime at large charges, we can analyze their growth from individual squarks to critical size. In our simplistic model, we show that Q-balls can indeed reach critical size on cosmological time scales.

The dissertation of Lauren Marie Wozniak Pearce is approved.

Lawrence Lyons

Graciela Gelmini

Zvi Bern

Alexander Kusenko, Committee Chair

University of California, Los Angeles

2014

To my husband and my family, and additionally, Mrs. King and Mr. Duncan, without whom I would not have pursued science.

Table of Contents

List of Figures

List of Tables

Acknowledgments

First, I would like to thank my collaborators: Professor Alex Kusenko, Professor Roberto Peccei, Professor John (Mike) Cornwall, and Dr. Kalliope Petraki. I would like to particularly thank Alex for his helpful guidance and suggestions throughout this entire project, and especially for his patience with my incessant questions. Mike deserves special thanks for his patience in teaching me how to use the Bethe-Salpeter equation.

The debt of gratitude I owe to my parents is enormous; they were my earliest science teachers who endued me with a sense of curiosity about the world. I'd like to especially thank Dad for leaving his college astronomy book where I could find it, and for patiently trying to answer the resulting questions at a level my elementary-school-age brain could understand. Science is the search for patterns in nature; consequently I'd like to thank Pap, who taught me to recognize patterns over many hours of solving jigsaw puzzles. I'd also like to thank Bub, who taught me to love nature through gardening. Words cannot express my gratitude to Mrs. King and Mr. Duncan; without their support I would not be in science at all. Finally, above all, I would like to thank my husband Kenny for his continual patience, support, and encouragement.

Chapter 3 is a version of J. M. Cornwall, A. Kusenko, L. Pearce and R. D. Peccei, "Can supersymmetry breaking lead to electroweak symmetry breaking via formation of scalar bound states?," Phys. Lett. B **718**, 951 (2013) [arXiv:1210.6433 [hep-ph]].

Chapter 4 is a version of L. Pearce, A. Kusenko and R. D. Peccei, "Phenomenology of Supersymmetric Models with a Symmetry-Breaking Seesaw Mechanism," Phys. Rev. D **88**, 075011 (2013 [arXiv:1307.6157 [hep-ph]].

Chapter 5 is a version of L. Pearce and A. Kusenko, "Indirect Detection of Self-Interacting Asymmetric Dark Matter," Phys. Rev. D **87** no. 12, 123531 (2013) [arXiv:1303.7294 [hep-ph]].

Chapter 6 is a version of K. Petraki, L. Pearce and A. Kusenko, "Self-interacting asymmetric dark matter coupled to a light massive dark photon", which is under submission at the Journal of Cosmology and Astroparticle Physics (JCAP). It is publicly available as arXiv:1403.1077 [hep-ph].

Chapter 7 is a version of L. Pearce, "Solitosynthesis induced phase transitions", Phys. Rev. D **85**, 125022 (2012) [arXiv:1202.0873 [hep-ph]].

Vita

2008	B.A. (Physics, Mathematics, and Philosophy Majors), University of Pennsylvania.
	M.S. (Physics), University of Pennsylvania.
2008–2013	Teaching Assistant, Physics Department, UCLA.
2013	Research Assistant, Physics Department, UCLA.

Publications

L. Pearce, "Solitosynthesis induced phase transitions," Phys. Rev. D **85**, 125022 (2012).

J. M. Cornwall, A. Kusenko, L. Pearce and R. D. Peccei, "Can supersymmetry breaking lead to electroweak symmetry breaking via formation of scalar bound states?," Phys. Lett. B **718**, 951 (2013).

L. Pearce and A. Kusenko, "Indirect Detection of Self-Interacting Asymmetric Dark Matter," Phys. Rev. D **87**, 123531 (2013).

L. Pearce, A. Kusenko and R. D. Peccei, "Phenomenology of Supersymmetric Models with a Symmetry-Breaking Seesaw Mechanism," Phys. Rev. D **88**, 075011 (2013).

K. Petraki, L. Pearce, and A. Kusenko, "Self-interacting asymmetric dark matter coupled to a light massive dark photon," Submitted to JCAP.

CHAPTER 1

The Standard Model of Particle Physics and Its Shortcomings

1.1 The Standard Model of Particle Physics

The Standard Model of Particle Physics is a quantum field theory that describes the strong, weak, and electromagnetic interactions at the subatomic level. A pedagogical introduction of the Standard Model can be found in [6]; in this chapter, I will give a quick summary of the Standard Model and outline several of its shortcomings which will motivate the research presented in this dissertation. Historically, the Standard Model is an amalgam of the work of several researchers, most notably Sheldon Glashow, Stephen Weinberg, and Abdus Salem [7–9].

Interactions in the Standard Model are described by gauge theories, which are closely connected to symmetries of the Lagrangian [10]. The Standard Model gauge symmetry group is $\mathrm{SU_C}(3) \times \mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$. The $\mathrm{SU_C}(3)$ part describes the strong interactions, while the $\mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$ part describes electroweak interactions; this will be related to the observed electromagnetic and the weak interactions in the following section. For simplicity, we will illustrate gauge theories with a simpler example: the transformations of the Lagrangian under the Abelian $\mathrm{U_{em}}(1)$ gauge group which gives rise to quantum electrodynamics (QED).

We begin by considering the action of non-gauged group on the fields which appear in the Lagrangian. The fields are organized into multiplets which have well-defined transformation

properties under the symmetry; these are characterized by the charge carried by the members of the multiplet. In our $\mathrm{U_{em}}(1)$ example, the electron Dirac spinor transforms as

$$\psi \to \psi' = e^{-ie\chi}\psi, \tag{1.1}$$

where e is the charge of the electron and χ is a constant parameter that specifies which transformation is being performed; for U(1), it can be thought of as an angle of rotation. Similarly, the conjugate field $\bar{\psi}$ transforms as[1]

$$\bar{\psi} \to \bar{\psi}' = \bar{\psi}e^{+ie\chi}. \tag{1.2}$$

We see that $\bar{\psi}'\psi' = \bar{\psi}\psi$, and thus the term $\bar{\psi}\psi$ is invariant. If, after transforming all of the fields in the Lagrangian, the Lagrangian maintains the same form, we say that the Lagrangian is invariant under the symmetry, or that the symmetry is a "good" symmetry.

For a symmetry that has not been gauged, the parameter(s) which describe the transformation (e.g., χ in our $\mathrm{U_{em}}(1)$ example) are constants. This type of symmetry is called a global symmetry, as the same transformation is applied everywhere in spacetime. To promote a theory into a gauge theory, we turn the global symmetries into local symmetries, in which we allow the transformation parameters to vary as a function of spacetime location. (E.g., χ is replaced with $\chi(x^\mu)$.) We now demand that the Lagrangian be invariant under these local transformations. Some terms, such as $\bar{\psi}\psi$, are invariant under both local and global symmetries. However, other terms are not; this is particularly true of the kinetic terms (required for each field), which involve derivatives. As an example, the kinetic term for an electron, $\bar{\psi}\gamma^\mu\partial_\mu\psi$, is invariant under local but not global $\mathrm{U_{em}}(1)$ transformations, as

$$\bar{\psi}'\gamma^\mu\partial_\mu\psi' = \bar{\psi}\gamma^\mu\partial_\mu\psi - i\bar{\psi}\gamma^\mu\psi\partial_\mu\chi. \tag{1.3}$$

Consequently, to maintain the invariance, derivatives generally must be replaced with gauge-covariant derivatives. A gauge-covariant derivative involves a gauge field, whose transformation properties are determined by the requirement that the Lagrangian be invariant.

[1]Throughout this work, the symbol † will be used for the Hermitian conjugate. For Dirac spinors, an overhead bar is used to designate $\psi^\dagger\gamma_0$.

For our example, the derivative acting on the electron is replaced with

$$\partial^\mu \to D^\mu = \partial^\mu + ieA_\mu,$$

where A_μ is the gauge field. As this is a quantum field theory, the gauge fields have their associated quanta, which are the particles which mediate the force represented by the gauge field. In our $\mathrm{U_{em}}(1)$ example, the quanta of A_μ field are identified as photons, and the force mediated by this field is electromagnetism. Generically, promoting a symmetry to a gauge symmetry introduces interactions between the fields which are members of a multiplet with a nonzero charge; these interactions are mediated by the gauge fields introduced in the gauge-covariant derivative.

These gauge symmetries are internal symmetries of the theory, in contrast with spacetime symmetries. These are the more familiar symmetries of rotations and Lorentz boosts, which are represented by the Lorentz group. Including translations gives the Poincaré group, which is the most general group of spacetime isometries of Minkowski spacetime. (We will return to this point in Ch. 3, when we introduce supersymmetry.) As with gauge symmetries, quantum fields are also assigned to multiplets of the Poincaré group.

Finally, we briefly classify the particle content of the Standard Model. The gauge fields include the photon, the gluons, the $W^\pm$ bosons, and the Z boson, all of which are spin-1. A fifth boson, the spin-0 Higgs boson, will be discussed in more detail below. There are also spin-1/2 fermions which can be divided into two classes: quarks and leptons. Quarks carry color charge and thus participate in the $\mathrm{SU_C}(3)$ strong interaction, while leptons do not. Fermions can also be organized into generations; each includes two quarks and two leptons. (On a technical level, this is advantageous because it is not sufficient that the Lagrangian be invariant under a symmetry; quantum corrections known as anomalies can still ruin the symmetry. However, if the fermions come in complete generations, then the Standard Model is free from anomalies.)

Three generations of fermions are currently known. The first generation consists of the up quark, the down quark, the electron, and the electron neutrino; the second generation consists of the charm quark, strange quark, muon, and muon neutrino; the third generation consists of the top quark, bottom quark, tau, and tau neutrino. The typical energy scale of each generation is greater than the previous one, although as yet there is no known explanation of this phenomenon.

In Ch. 6, we will introduce a new gauge symmetry regarding dark matter. Furthermore, this understanding of interactions as gauge theories is foundation to the Higgs mechanism, which will appear several times in this work, and which we discuss next.

1.2 Spontaneous Symmetry Breaking

The above introduction to the Standard Model emphasized the role of symmetries; however, not all the symmetries underlying a theory may be evident in the quantized theory (even if the theory is anomaly-free). It is possible for the vacuum state to not be invariant under a symmetry, even if the Lagrangian is invariant under that same symmetry. Consequently, the symmetry will not be evident in the particle spectrum of the resulting theory, and we say that the symmetry is spontaneously broken. This can give mass to the otherwise massless gauge bosons through the Higgs mechanism, which was first presented in Refs. [11–13].

As an example, we first discuss a simplified model of spontaneous symmetry breaking. Consider a theory of a single complex field, φ, with the potential

$$V = -m^2\varphi^2 + \lambda\varphi^4. \tag{1.4}$$

This potential is invariant under the U(1) symmetry $\varphi \to \varphi' = e^{i\alpha}\varphi$, where α is an arbitrary phase. However, $\varphi = 0$ is a local maximum, not minimum, of the potential, and the theory should be quantized about a minimum. The minimum of the potential is at the vacuum expectation value $\langle\varphi_0\rangle = e^{i\beta}m/\sqrt{2\lambda}$, with the phase β arbitrary. (That is, there is a set of

degenerate minima parameterized by β.) Without a loss of generality, we may choose $\beta = 0$ and expand the potential about this point, which is an appropriate vacuum state. As the U(1) transformation alters β, the vacuum state is not invariant under the U(1) symmetry.

We will now consider the real and imaginary components of the field ϕ separately; the expansion amounts to the replacement

$$\varphi \to \varphi_0 + \frac{1}{\sqrt{2}} (\varphi_1 + i\varphi_2) , \tag{1.5}$$

where the $\sqrt{2}$ accounts for the different number of degrees of freedom in real and complex fields; both φ_1 and φ_2 are real fields. Performing the expansion and dropping unphysical constants gives

$$V' = m^2\varphi_1^2 + m\sqrt{\lambda}\left(\varphi_1\varphi_2^2 + \varphi_1^2\varphi_2\right) + \frac{\lambda}{4}\varphi_1^4 + \frac{\lambda}{4}\varphi_2^4 + \frac{\lambda}{2}\varphi_1^2\varphi_2^2, \tag{1.6}$$

which is no longer invariant under the symmetry $\varphi \to \varphi' = e^{i\alpha}\varphi$ (which takes φ_1 to $\cos(\alpha)\varphi_1$ and φ_2 to $\sin(\alpha)\varphi_2$). Thus, the symmetry present in the original theory is broken by the choice of vacuum.

We also note that there is also no quadratic term of the form φ_2^2; since such a term would represent the mass of the φ_2 field, this field is massless. This is a consequence of Goldstone's Theorem [14]; this is a general result which predicts a massless boson for each generator of the symmetry group which is "broken" (that is, for each generator that, when it acts on the vacuum, does not return that vacuum). The other field φ_1 does have a mass term.

As the above example illustrates, a symmetry is spontaneously broken when the origin is not the global minimum of the potential. Since quantum field theories do not include terms linear in the fields (such terms can always be eliminated by a field redefinition), the behavior of the potential near the minimum is generally determined by the quadratic (mass) term. Thus, spontaneous symmetry breaking occurs when the potential has a quadratic term of the form $-m^2\phi^2$. (Although naively such a theory appears to contain fields with an

imaginary mass, the final quantized theory expanded about the proper vacuum contains no such physically unreasonable mass values.)

In the Standard Model, the electromagnetic interaction and the weak interaction are united into one fundamental interaction, the electroweak interaction, corresponding to the gauge group $\mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$. Experimental evidence, though, shows that this symmetry appears not to be realized; the Standard Model accommodates this by having the symmetry be spontaneously broken through the Higgs mechanism. This mechanism also gives mass to the $W^{\pm}$ and Z bosons, and possibly all of the fermions.

The Standard Model Higgs field has an associated potential, which has the same form as Eq. (1.4), although the Higgs field φ is a complex doublet, not a complex scalar. The elements of the $\mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$ symmetry group act on the complex doublet field as

$$\varphi \to \varphi' = e^{i\boldsymbol{\alpha}\cdot\boldsymbol{\sigma}/2+i\beta}\varphi, \tag{1.7}$$

where $\boldsymbol{\sigma}$ are the Pauli matrices. The doublet φ can be written in terms of four real components as

$$\varphi = \frac{1}{\sqrt{2}} \begin{pmatrix} \varphi_1 + i\varphi_2 \\ \varphi_3 + i\varphi_4 \end{pmatrix}. \tag{1.8}$$

When this is expanded about a minimum of the potential, the resulting theory is no longer invariant under the $\mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$ transformations, although it remains invariant under a residual $\mathrm{U_{em}}(1)$ gauge symmetry, which is identified with the electromagnetic interaction.

Following Goldstone's Theorem, we would expect the theory to contain three massless bosons, corresponding to the three generators of the $\mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$ group that do not leave the vacuum invariant; however, no such bosons are observed. This apparent conflict is resolved through the Higgs mechanism, in which these degrees of freedom are "eaten up" giving mass to the $W^{\pm}$ and Z bosons, which are necessarily massless in the original $\mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$ gauge theory. Mathematically, when one expands φ about the minimum

φ_0, the gauge-covariant derivatives produce the terms $m^2W^+W^-$ and m'^2Z^2. The massless Goldstone bosons become the longitudinal degrees of freedom for these now-massive bosons.

The fourth degree of freedom in the complex doublet φ remains in the final theory; like φ_1 in our example, it is a massive scalar boson. This Higgs boson was recently discovered by the ATLAS and CMS experiments at the Large Hadron Collider (LHC) [15, 16]. In the research presented here, the Standard Model Higgs mechanism appears in two places: in Ch. 3, we consider modifications to electroweak symmetry breaking in a strongly-coupled supersymmetric theory, and secondly, in Ch. 5 the Higgs boson will enable the decay of dark sector particles to Standard Model particles. Additionally, in Ch. 6 we introduce a Higgs mechanism in the dark sector to give mass to a dark mediator boson; this has important effects on the cosmology of the resulting model.

1.3 Shortcomings of the Standard Model

Despite its immense experimental success, the Standard Model alone cannot be a complete description of the subatomic world. In this section, we will focus on some of its failings, which lead us to consider extensions of the Standard Model which address these issues. In particular, the Standard Model cannot explain the observed matter-antimatter asymmetry of the universe, nor does it provide a candidate particle for dark matter.

As is well known, matter and antimatter annihilate with each other. However, antimatter is mysteriously missing in the cosmos. Large regions of antimatter in contact with matter would produce copious amounts of electromagnetic radiation, but astronomers do not see signals of significant matter anti-matter annihilation. Furthermore, the presence of additional antimatter in the early universe would have altered the abundances of light nuclei produced in big bang nucleosynthesis. All the evidence points to a cosmological abundance of baryons over antibaryons, which is described by the parameter $\eta = (n_B - n_{\bar{B}})/s$, where n_B is the number density of baryons, $n_{\bar{B}}$ the number density of antibaryons, and s the entropy density

of the universe. η remains approximately constant as the universe evolves; our current observational evidence supports $\eta \approx 10^{-10}$.

In 1967, Sakharov laid out three conditions necessary to generate a baryonic asymmetry in the early universe [17]. First, there must be baryon number violation; that is, some process must treat baryons differently than antibaryons. Secondly, CP-symmetry must be broken; otherwise, equal numbers of left-handed baryons and right-handed antibaryons will be produced, leading to no net asymmetry.[2] Finally, the interactions must occur out of thermal equilibrium; otherwise CPT-symmetry ensures that processes that increase baryon number and processes that decrease baryon number will occur at equal rates[3].

Since these three conditions are necessary, it is reasonable to ask whether they occur in the Standard Model. While QCD and QED do not have processes that treat baryons differently than antibaryons, transitions involving electroweak sphalerons do not conserve baryon number. Furthermore, the vertex for a $W^{\pm}$ boson interacting includes an element from the CKM matrix, which has a single nonzero phase which breaks CP-symmetry. Thus, two of the three ingredients are present, but the third, thermal non-equilibrium, is not. The only opportunity for weak interactions to fall out of thermal equilibrium is during the electroweak phase transition, in which the Higgs boson acquires its nonzero vacuum expectation value. However, for a Higgs mass of about 125 GeV (as has recently been measured by the ATLAS and CMS collaborations at the Large Hadron Collider [15, 16]), this phase transition is second order [18]. Thus, the system does not depart from thermal equilibrium, and the Standard Model alone cannot account for the observed preponderance of matter over antimatter. In Ch. 4 and Ch. 7, we will discuss baryogenesis in two extensions of the Standard Model.

[2] CP symmetry is the product of applying C and P symmetry operators to a system; C symmetry is in the interchange of particle and antiparticles, while P (parity) inverts the spatial axes.

[3] CPT is the product of the C and P operators, described in the previous footnote, with T, which is the time-reversal operator.

In the early 1930's, astronomers Jan Oort and Francis Zwicky observed that the speeds of certain astrophysical objects were too fast for the amount of luminous matter present [19,20]; Oort considered orbital velocities of stars in the Milky Way, while Zwicky studied the speeds of galaxies in clusters. This work was expanded upon in the 1970's by Vera Rubin, who studied the motion of stars in distant galaxies [21]. In all of these systems, the radial velocity of an object at a particular radius is determined by the amount of mass enclosed by a sphere at that radius. However, while the luminous material dropped off significantly as a function of radius, these observations showed that the mass did not. Consequently, one is forced to conclude that much of the mass within an individual galaxy or cluster is non-luminous. Today, we believe that between 20 percent and 30 percent of the total mass-energy of the universe is this non-luminous material, known as dark matter [22].

Even without extending the Standard Model, one can propose several possibilities for this non-luminous material, such as dark holes, neutrinos, or massive compact halo objects (e.g., a relatively large number of planets of the size of Jupiter). Each of these possibilities is discussed in the review [23], but none of the the Standard Model candidates prove viable. For example, measurements of gravitational microlensing disfavor (normal) black holes and massive compact halo objects, and neutrinos inhibit rather than assist early structure formation. While astronomical observations eliminate the Standard Model possibilities, they do not provide much guidance as to how to extend the Standard Model.

Perhaps the best-motivated candidates are WIMPs- weakly interacting massive particles. As the universe expands and cools, interactions may not be sufficient to maintain a particular species in thermal equilibrium. If that particle is sufficiently long-lived, then the number density of that species of particle evolves only through the expansion of the universe. Generically, a heavy particle with interactions near the weak scale freezes out at approximately the correct abundance to account for dark matter; this is known as the WIMP miracle. Supersymmetry (which will be discussed in Ch. 3 and Ch. 4) naturally provides a WIMP dark matter candidate, the lightest supersymmetric particle. An overview of WIMPs, and

supersymmetric WIMPs in particular, is given in Ref. [24].

However, models outside the WIMP paradigm have also been considered. While collisionless cold dark matter is frequently taken as a generic baseline model, self-interactions between dark matter may be desirable. Such interactions are necessary if dark matter is a thermal relic (such as a WIMP), but they may also be motivated on other grounds. They enable momentum-transfer within dark matter halos, which can establish cored profiles, as opposed to a cusps [25]. Furthermore, they may resolve the "too-big-to-fail" problem, in which the observed number of dwarf halos is well beneath the value expected from collisionless dark matter simulations [26–28]. These may be weak interactions, or there may be an entire dark sector with its own set of interactions.

In models of asymmetric dark matter, dark matter is assumed to carry its own baryon-number-like asymmetry; high-energy operators produce this asymmetry along with the observed baryon asymmetry (for a review, see Ref. [29]). In Ch. 5 and Ch. 6, we will consider self-interactions in asymmetric dark matter models, with a particular focus on the formation of dark matter bound states. Further motivation for these models will be discussed in these chapters.

CHAPTER 2

The Bethe-Salpeter Equation

Bound states are a generic feature of strongly-interacting systems, which are notoriously difficult to study. One popular technique to study such systems is to perform numerical analysis on a lattice; in the final step, the lattice spacing is taken to zero. In addition to being computationally intensive, this approach requires one to choose numerical values for most or all the parameters of a theory. This is not beneficial for exploring the generic features of some extension of the Standard Model with several unknown and lightly constrained parameters. Such a numerical analysis could miss certain features due to a poor choice of parameters; conversely, if a promising result is obtained, one must question whether this is a fortuitous result from a lucky choice of numerical values. Thus, there is a need for analytical tools to study bound states; one such tool is the Bethe-Salpeter equation.

The Bethe-Salpeter equation will be used extensively in Ch. 3, in which we use it to analyze supersymmetry bound states, and in Ch. 7, in which we use it to approximate small charge Q-balls. It will also be mentioned in Ch. 5, in the discussion of the mass of dark matter bound states. Therefore, in this chapter we will introduce the Bethe-Salpeter equation in some generality and illuminate the points that will be important to the subsequent chapters. A pedagogical introduction to the Bethe-Salpeter equation as used in particle physics, and the Wick-Cutkosky model in particular, can be found in [30].

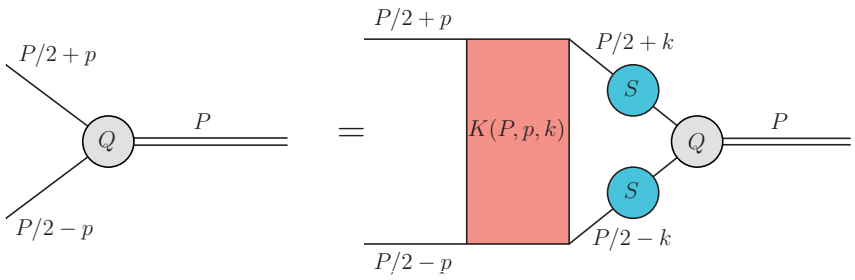

Figure 2.1: The exact Bethe-Salpeter equation in vertex form. The double line represents the bound state which carries momentum P, and the single lines represent the constituent particles.

The Bethe-Salpeter equation gives a relativistically covariant formalism for describing a bound state in a quantum field theory [31]. Mathematically, it is frequently written as:

$$Q(P,p) = \int \frac{d^4k}{(2\pi)^4} K(P,p,k) S(k - P/2) Q(P,k) S(k + P/2), \tag{2.1}$$

where Q is the Bethe-Salpeter amplitude, K is the kernel, and S are the propagators of the two legs. This is represented symbolically in Fig. 2.1.[1] This form of the equation is applicable when the bound state is made of two identical particles, or a particle and its antiparticle.

We can transform the Bethe-Salpeter equation from the vertex form to an equation for the wavefunction using

$$Q(P,p) = S(P - p/2)^{-1} S(P + p/2)^{-1} \Psi(P,p), \tag{2.2}$$

which gives

$$S(P + p/2)^{-1} S(P - p/2)^{-1} \Psi(P,p) = \int \frac{d^4k}{(2\pi)^4} K(P,p,k) \Psi(P,k). \tag{2.3}$$

In these equations, K is the total kernel, which includes (with appropriate weightings) all of the possible particle exchanges that contribute to the bound state. Similarly, S represents

[1] A similar figure appears in Ch. 3, where we identify the particles involved as squarks and consider one specific kernel.

the complete propagators, summed to all orders of perturbation theory. However, to make practical use of the Bethe-Salpeter equation, one typically considers only the lowest order contributions to K and S. For example, if the constituent particles are bosons, one may take

$$S(q) = \frac{i}{q^2 - m^2}, \tag{2.4}$$

while for fermions one may use

$$S(q) = \frac{i(\not{q} + m)}{q^2 - m^2}. \tag{2.5}$$

The kernel depends on the particular characteristics of the interactions in the model; perhaps the most commonly used kernel is the exchange of a single massless scalar boson. If additionally the constituent particles ("legs") are identical spin-0 bosons, then this is the well-known Wick-Cutkosky model [32, 33]. In this case, to lowest order, the kernel is

$$K(P, p, k) = \frac{A^2}{(p - k)^2}, \tag{2.6}$$

where A is the coupling constant between the legs and the exchanged massless scalar particle. If we define $\lambda = A^2/16\pi^2$, then the Bethe-Saltpeter equation for the wavefunction is

$$\left[m^2 + \left(\frac{P}{2} + p\right)^2\right]\left[m^2 + \left(\frac{P}{2} - p\right)^2\right]\Psi(P, p) = \frac{\lambda}{\pi^2}\int d^4q \frac{\Psi(P, q)}{(p - q)^2}, \tag{2.7}$$

after a Wick rotation.

In Ch. 3, we will consider a similar model, also involving scalars, but with a crossed-graph kernel with massive bosons. We will specifically be interested in the case in which the bound state has zero mass; we will now consider this case in the Wick-Cutkosky model. (These results will inspire our trial wavefunctions in the variational approach used in Ch. 3.) If $P^2 = M^2 = 0$, the above equation simplifies to

$$(m^2 + p^2)^2\Psi(P, p) = \frac{\lambda}{\pi^2}\int d^4q \frac{\Psi(P, q)}{(p - q)^2}. \tag{2.8}$$

This equation has a spectrum of solutions, each with a corresponding eigenvalue λ. One particular solution of interest is $\Psi(P, q) = (p^2 + m^2)^{-3}$; we now proceed to demonstrate

that this does indeed satisfy the integral equation. Let us name the right hand side $\mathcal{I}$ and introduce a Feynman parameter:

$$\begin{aligned}\mathcal{I} &= \frac{\lambda}{\pi^2}\int d^4q \frac{1}{(p-q)^2(q^2+m^2)^3} \\ &= \frac{\lambda}{\pi^2}\int d^4q \int_0^1 dx \frac{3(1-x)^2}{[x(p-q)^2+(1-x)(q^2+m^2)]^4} \\ &= \frac{\lambda}{\pi^2}\int d^4q \int_0^1 dx \frac{3(1-x)^2}{[(q-xp)^2+(1-x)(m^2+xp^2)]^4}. \end{aligned} \tag{2.9}$$

The denominator in the integrand can be rewritten as a Gaussian using the identity

$$B^{-4} = \frac{1}{3!}\int_0^\infty \alpha^3 e^{-\alpha B}\, d\alpha, \tag{2.10}$$

which gives

$$\mathcal{I} = \frac{\lambda}{2\pi^2}\int d^4q \int_0^1 dx \int_0^\infty d\alpha\, (1-x)^2\alpha^3 \exp\left(\alpha[(q-xp)^2+(1-x)(m^2+xp^2)]\right). \tag{2.11}$$

Although this appears complicated, we can now perform the d^4q integral, as it is a Gaussian integral; the result is

$$\mathcal{I} = \frac{\lambda}{2\pi^2}\int_0^1 dx \int_0^\infty d\alpha\, \frac{\pi^2}{\alpha^2}\cdot \alpha^3(1-x)^2 \exp\left(\alpha(1-x)(m^2+xp^2)\right). \tag{2.12}$$

Next the $d\alpha$ integral gives

$$\mathcal{I} = \frac{\lambda}{2}\int_0^1 \frac{(1-x)^2\, dx}{(1-x)^2[m^2+xp^2]^2} = \frac{\lambda}{2}\int_0^1 \frac{dx}{[m^2+xp^2]^2}, \tag{2.13}$$

and the remaining integral is

$$\mathcal{I} = \frac{\lambda}{2m^2}\frac{1}{m^2+p^2}. \tag{2.14}$$

This does indeed equal the left hand side of the Bethe-Salpeter equation

$$(m^2+p^2)(m^2+p^2)^{-3} = \frac{1}{m^2+p^2}, \tag{2.15}$$

if the eigenvalue is $\lambda = 2m^2$. As mentioned, this result in the Wick-Cutkosky model will be used to motivate our choice of trial wavefunctions in Ch. 3. As we will work with the vertex form of the Bethe-Salpeter equation, we note that the corresponding vertex solution is $Q(0,p) = 1/(p^2+m^2)$.

In the general case in which the bound state mass is not zero, one can transform the integral equation into a differential equation [33]. One way of writing this differential equation is [30]

$$\frac{d^2\Psi_n}{dt^2} + \left[\frac{\lambda}{t[m^2(1+t)^2 - M^2t^2]} - \frac{n(n+2)}{4t^2}\right]\Psi_n = 0. \tag{2.16}$$

Although we will not repeat the derivation here, for sufficiently small λ, the eigenvalues of this differential equation are approximately [33]

$$\lambda = \frac{2}{\pi}\sqrt{1 - M^2/4}. \tag{2.17}$$

Inverting this equation for the mass of the bound state, we find

$$M^2 = 4m^2 - \lambda^2\pi^2m^2 = 4m^2 - \frac{A^4}{256m^2\pi^2}, \tag{2.18}$$

which, in the limit that $A \ll 1$, gives

$$M = 2m - \frac{A^4}{1024m^3\pi^2}, \tag{2.19}$$

which we will see again in Ch. 7 as Eq. (7.22).

Thus far, we have only considered the scenario in which the two constituent particles are identical; this is sufficient for the strongly-interacting MSSM scenario we investigate in Ch. 3 and Ch. 4. However, when we discuss the growth of small Q-balls in Ch. 7, we will have different particles on the upper and lower legs. The more general Bethe-Salpeter equation appropriate to this scenario is (after a Wick transformation)

$$\left[(m+\Delta)^2 + (p - i\eta_1 P)^2\right]\left[(m-\Delta)^2 + (p - i\eta_2 P)^2\right]\Psi(P,p) = \frac{\lambda}{\pi^2}\int d^4q \frac{\Psi(P,q)}{(p-q)^2}, \tag{2.20}$$

where $m_1 = m + \Delta$, $m_2 = m - \Delta$, and

$$\eta_1 = \frac{m_1}{m_1 + m_2}, \qquad \eta_2 = \frac{m_2}{m_1 + m_2}. \tag{2.21}$$

We have inserted a kernel which corresponds to the exchange of a single massless boson, and again the legs are also scalars. The eigenvalue is now $\lambda = A_1A_2/16\pi^2$, where A_1 and A_2 are the couplings at the top leg and bottom leg respectively. As before, this integral equation

can be transformed into a differential equation, and a change of variables transforms it to a form reminiscent of the equal mass case (see the extended derivation in [30]). The resulting differential equation is

$$\frac{d^2\Psi_n}{d\tilde{t}^2} + \left[\frac{\lambda/(1-\Delta^2/m^2)}{\tilde{t}[m^2(1+\tilde{t})^2 + ((M^2-4\Delta^2)/(1-\Delta^2/m^2))\tilde{t}]} - \frac{n(n+2)}{4\tilde{t}^2}\right]\psi_n = 0. \tag{2.22}$$

The similarities in form to Eq. (2.16) are evident. We conclude that if $\lambda = F(M^2)$ is the eigenvalue spectrum in the equal mass case, then in the unequal mass case,

$$\frac{\lambda}{1-\Delta^2/m^2} = F\left(\frac{M^2-4\Delta^2}{1-\Delta^2/m^2}\right). \tag{2.23}$$

Using Eq. (2.17), we write

$$\frac{\lambda}{1-\Delta^2/m^2} = \frac{2}{\pi}\sqrt{1-\frac{1}{4}\left(\frac{M^2-4\Delta^2}{1-\Delta^2/M^2}\right)}, \tag{2.24}$$

which, when solved for M^2, gives

$$M^2 = 4\Delta^2 + 4m^2\left(1-\frac{\Delta^2}{m^2}\right)\left[1-\frac{\lambda^2\pi^2}{4(1-\Delta^2/m^2)}\right]. \tag{2.25}$$

This relation will be used extensively in Ch. 7, where it appears as Eq. (7.25).

CHAPTER 3

Strongly-Interacting Supersymmetry

3.1 A Brief Introduction to Supersymmetry

The introductory chapter discussed the important role of symmetries in the Standard Model; consequently, a natural question is to what extent the symmetries of the Standard Model can be extended. While the gauge symmetry group can be embedded in larger groups, the same cannot be said of the Poincaré group which describes spacetime symmetries. This important "no-go" theorem was proved by Coleman and Mandula in 1967 [34].

More precisely, they showed that, for a theory with non-trivial interactions between fields, the underlying Lie algebra must be a direct product of the Lorentz algebra and an algebra describing an internal (that is, gauge) symmetry group. The Lorentz algebra underlies the Lorentz group and its non-simply-connected extension, the Poincaré group. Thus, the Coleman-Mandula "no-go" theorem seemed to prove that the only possible way to extend the Standard Model is to introduce new gauge interactions.

However, a loophole was discovered in 1971: one can build a quantum field theory based on a superalgebra instead of an algebra [35, 36]. Superalgebras are a more general mathematical concept than algebras; the defining relations between generators can include anti-commutation relations. (Formally, superalgebras are algebras with a $\mathbb{Z}_2$ grading.) Supersymmetric models extend the spacetime symmetries of the Standard Model using these superalgebras; the simplest of these is the Minimal Supersymmetric Standard Model (MSSM).

The new symmetry introduced by the superalgebra relates bosons (integer spin particles) to fermions (half-integer spin particles). All supersymmetric models necessarily predict the existence of an equal mass fermion for each boson and vice versa. These "superpartners" have not been observed experimentally, and so it was immediately evident that supersymmetry cannot possibly be a "good" symmetry. However, just as the Standard Model $\mathrm{SU_L}(2)$ gauge theory is spontaneously broken, supersymmetry may be a broken symmetry. Thus, although supersymmetry is, by some measure, one of the most elegant extensions of the Standard Model, its elegance is dramatically decreased by the ad hoc introduction of supersymmetry-breaking terms.

There are two primary ways of breaking supersymmetry, gauge-mediated supersymmetry breaking and gravity-mediated supersymmetry breaking. For our purposes, it is important that supersymmetry breaking generically introduces trilinear terms into the Lagrangian of the form:

$$A_{ij} Q^\dagger_{Ui} H_u q_{Uj} + A'_{ij} Q^\dagger_{Di} H_d q_{Dj} + h.c. \tag{3.1}$$

where Q is the squark $\mathrm{SU_L}(2)$ doublet; the subscript U designates the up-type doublets and D designates the down-type doublets. A lower-case q denotes $\mathrm{SU_L}$ singlets, and again, U and D subscripts distinguish up and down-type quarks. i and j are flavor indices, and the Higgs doublets H_u and H_d are the up- and down-type doublets respectively. By strongly-interacting supersymmetry, we will mean a model in which the scale of these couplings A_{ij} and/or A'_{ij} is large compared to the scale of the squark masses.

Given that supersymmetry must be a broken symmetry, one may question the motivation of such extensions of the Standard Model. Here, we briefly give two. Although the Standard Model is renormalizable, the renormalization procedure for the Higgs mass requires certain "bare" parameters to be fine-tuned to many orders of magnitude, which is considered unnatural. In unbroken supersymmetric models, the fermionic and bosonic loop corrections to the Higgs mass exactly cancel, and no fine-tuning is necessary. Since supersymmetry is broken,

the cancellation is not perfect, but a Higgs mass near or just above the electroweak scale remains natural. In general, the divergences which require renormalization are far less severe in supersymmetric models due to the cancellation of fermionic and bosonic contributions.

Secondly, many supersymmetric models include a charge which distinguishes supersymmetric particles from Standard Model particles. If this charged is conserved, the lightest supersymmetric particle cannot decay. As supersymmetric particles will generally be produced in the early universe and eventually decay into the lightest supersymmetric particle, we expect a large number of these remain at present. These may interact weakly with Standard Model particles, and therefore, this lightest supersymmetry particle is a good dark matter candidate.

3.2 Motivation for Strongly-Coupled Supersymmetry

The ATLAS and CMS experiments at the Large Hadron Collider have recently discovered a Higgs boson with a mass of 125 GeV [15, 16]. This mass is larger than the range predicted for the lowest mass Higgs boson in the MSSM, and additionally, these experiments have not yet observed any superpartners [37]. There have been a number of efforts to reconcile these observations with low-energy supersymmetry (e.g., Refs. [38–40]). At least one of these models [40] assumes strong couplings in the supersymmetry breaking sector. This is perhaps somewhat unusual, because models with gauge-mediated supersymmetry breaking generically predict a small trilinear supersymmetry-breaking coupling A. However, a large value of this coupling is helpful in raising the range of the Higgs boson masses toward 125 GeV [38, 41]. Large trilinear terms can appear in gauge-mediated supersymmetry breaking models, albeit some fine-tuning of parameters may be required in a realistic model [38].

It has already been established that the exchange of the (lighter) Higgs boson between (heavier) squarks can lead to formation of bound states [42, 43]. Thus, large trilinear couplings lead to a new strongly coupled realization of the Minimal Supersymmetric Standard

Model (MSSM). In this chapter, we focus particularly on the possibility that supersymmetry breaking may trigger electroweak symmetry breaking via the formation of squark bound states with the same quantum gauge numbers as the fundamental Higgs boson. These new states can mix with the Higgs boson and thus they also can acquire a nonzero vacuum expectation value (VEV). Consequently, the resulting multi-Higgs low-energy effective theory may have a very different appearance from the usual weakly coupled MSSM.

Let us consider a simplified version of MSSM, in which we will focus only on the third generation of squarks. For simplicity, we will further assume that only one of the trilinear terms is large,

$$\mathcal{L} = A_t(\tilde{t}_L \cdot H_u)\tilde{t}_R, \tag{3.2}$$

where $\tilde{t}_L$ is the $Y = 1/3$ stop doublet under $\mathrm{SU_L}(2)$, $\tilde{t}_R$ is the $Y = 4/3$ stop singlet, and H_u is the $Y = -1$ Higgs doublet. (For clarity, we have not written the full potential; we assume that it also includes the usual quartic H_u^4 term and mass terms.) To further simplify matters, we additionally assume the squarks have a common mass M of a few TeV, considerably larger than the Higgs mass m.

We have suppressed the $\mathrm{SU_C}(3)$ indices in Eq. (3.2), and we will concentrate on the color-singlet bound state. We acknowledge that the same exchange of Higgs bosons will generally also produce $\mathrm{SU_C}(3)$ non-singlet bound states. However, the singlet is the only bound state that can mix with the fundamental Higgs boson, and as will be discussed in below, consequently there is a range of parameters in which this bound state alone has a non-zero VEV. In this case, the $\mathrm{SU_C}(3)$ symmetry underlying quantum chromodynamics remains unbroken.

We have referred to the bound states as being made of squarks exchanging Higgs bosons. However, this terminology requires some care. The trilinear coupling constant A_t will, in general, evolve with scale. If it remains large at all scales, then there is no regime in which

it is proper to speak of individual squarks, as they are never the appropriate degrees of freedom to use in an effective theory. Furthermore, as mentioned in Sec. 3.1, one motivation for supersymmetry is its high-energy behavior, in which the quadratic divergences present in the Standard Model are weakened due to the (near) cancellation between fermionic loops and bosonic loops. If the degrees of freedom that appear in the weakly-coupled MSSM do not appear at high energy, this result is called into question.

We note that the analysis presented below will, in fact, assume only that $|A_t|$ is large; the specific sign of the coupling will not appear. If A_t is large and positive, the analysis of [38] suggests that perturbativity is not regained as the energy scale increases. However, if $|A_t|$ is large while A_t itself is negative, then perturbativity may be regained, since A_t increases as the renormalization scale increases. In this scenario, the desired loop cancellations are certainly retained; the formation of bound states at low energies has no impact on the high energy behavior. The difference with the MSSM is in the low-energy effective theory, which contains different degrees of freedom: fewer squarks and more Higgs bosons, whose VEVs produce a more complicated vacuum. In particular, in this vacuum the usual MSSM relations between the gauge couplings and the scalar self-coupling do not hold, and, therefore, there is no reason for the upper bound on the lightest Higgs boson to be the same as in the usual version of MSSM. Hence, the tension between the MSSM and the observed Higgs mass is ameliorated.

Thus, strongly-coupled supersymmetry may retain the beneficial aspects of supersymmetry while addressing the relatively large observed Higgs mass; we are therefore motivated to investigate the low-energy behavior of this theory. Due to the presence of additional bound state Higgs-like doublets, a particular interest is electroweak symmetry breaking.

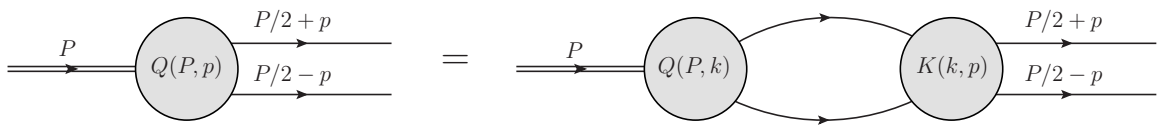

Figure 3.1: The Bethe-Salpeter equation in vertex form. The double line represents the bound state Φ with momentum P, and the single lines represent the constituent squarks.

3.3 Electroweak Symmetry Breaking By the Bound-State Higgs

The up-type Higgs boson that breaks electroweak symmetry in the MSSM must be a CP+ scalar doublet with $Y = 1$; additional doublets, of any source, with the same quantum numbers will generically modify electroweak-symmetry-breaking. Such a doublet does arise in the strongly-coupled realization of supersymmetry considered here, as a $(\tilde{t}_R \tilde{t}_L^\dagger)$ bound state. We might ask whether this bound state can possibly supplant the fundamental Higgs boson in triggering electroweak symmetry breaking; this question will be investigated in this section. (Note, however, that even if it is possible to trigger electroweak symmetry breaking with the bound state, the fundamental Higgs boson is not superfluous; the bound state itself is held together through the exchanged of members of the fundamental Higgs doublet.)

The bound state can be described by a Euclidean Bethe-Salpeter equation, as shown in Fig. 3.1. This Bethe-Salpeter equation is written in vertex form, where the internal lines represent propagators and the usual Bethe-Salpeter wave function $\Psi(P, p)$ is related to $Q(P, p)$ by

$$\left[\left(\frac{P}{2} + p\right)^2 + M^2\right]\left[\left(\frac{P}{2} - p\right)^2 + M^2\right]\Psi(P, p) = Q(P, p). \tag{3.3}$$

A general discussion of the Bethe-Salpeter equation and its solutions is given in Ch. 2. We present the Bethe-Salpeter equation in this form because it is reminiscent of the chiral gap equation; indeed, we will derive a similar gap equation below. In chiral symmetry breaking, the gap equation describes the mass splitting between the chiral states; our analogous gap

equation will likewise describe the mass splitting between left- and right-handed stops. Since a mass splitting between left- and right-handed states (or as it is more commonly described, a mass term connecting left- and right-handed states) is a hallmark sign of electroweak symmetry breaking, we can use our gap equation to determine when electroweak symmetry breaking has occurred. Namely, if we find the condition for which the gap equation has a non-trivial solution, we will have found the condition for the bound state Higgs doublet to break electroweak symmetry, given some caveats to be specified below.

We have mentioned that the bound state will generically mix with the Higgs doublet, and so the mass eigenstates will generally be mixtures of the two. We will incorporate this mixing in the following section; in this section, we assume no mixing between the fundamental and bound state Higgs doublets. This allows us to determine whether the bound state alone can drive electroweak symmetry breaking.

We have already represented the Euclidean Bethe-Salpeter equation diagrammatically in Fig. 3.1; mathematically, this equation is

$$Q(P,p) = \frac{A_t^4}{(2\pi)^4} \int d^4k \; Q(P,k) \frac{1}{[(P/2+k)^2+M^2][(P/2-k)^2+M^2]} K(P,p,k). \qquad (3.4)$$

For simplicity, we have suppressed an $\mathrm{SU_L}(2)$ spinor index on Q and a corresponding factor δ_{ij} on the kernel. The lowest-order contribution to the kernel is a crossed graph, as shown in Fig. 3.2. In this figure, the line labeled $(P/2)+p$ represents an outgoing $\tilde{t}_R$, and the line labeled $(P/2)-p$ represents an outgoing $\tilde{t}_L^\dagger$. The lines labeled L, R are the stops, and the exchanged dashed lines are the fundamental up-type Higgs fields of the MSSM.

As discussed in Ch. 2, the Bethe-Salpeter equation has been solved exactly in the case of the exchange of a single massless scalar particle; this is the well-known Wick-Cutkosky model [32, 33]. Unfortunately, our kernel is complicated by two factors: first, we have a crossed graph involving the exchange of two bosons, not one, and secondly, we cannot approximate the exchanged bosons as massless. If this were the case, then we would be

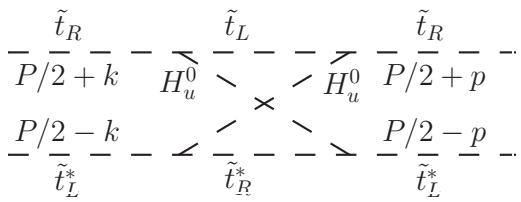

Figure 3.2: The lowest-order kernel for the Higgs-like bound state.

at the onset of spontaneous electroweak symmetry breaking due to the fundamental Higgs boson; however, we are interesting in the possibility of breaking electroweak symmetry with the bound state Higgs doublet. Due to these two complications, we have not been able to perform a general analysis of the Bethe-Salpeter equation. If one were able to solve this integro-differential equation for a general momentum P, the result would be a relationship between the values of P at which the equation has a non-trivial solution and the coupling A_t.

However, for our purpose of investigating electroweak symmetry breaking, the general solution is unnecessary. Since $P^2 = m_{BS}^2$, where m_{BS} is the mass of the bound state, and electroweak symmetry breaking occurs when the mass-squared become negative, we can look at the specific value $P^2 = m_{BS}^2 = 0$, which corresponds to the onset of symmetry breaking. Strictly speaking, at $P = 0$, we have four degenerate massless bound states, which correspond to the two complex elements of the bound states $\mathrm{SU_L}(2)$ doublet. As in the Standard Model Higgs mechanism, at the onset of symmetry breaking, these states can also be considered as states of a broken $\mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$ theory: here we have a zero-mass composite Higgs boson, plus three zero-mass Nambu-Goldstone bosons.

The simplifications that result from setting $P = 0$ allow us to investigate the kernel, and the resulting gap equation, with decent quantitative accuracy. We introduce the notation $K(p, k) = K(k, p)$ for the original kernel evaluated at $P = 0$. We first summarize our goal: Eq. (3.4), with $P = 0$, is an integro-differential eigenvalue equation. These eigenvalues are

proportional to A_t^4; if we imagine slowly increasing $|A_t|$, the equation will at first have no non-trivial solutions. This does not mean that bound states do not exist; rather, they do not satisfy the Bethe-Salpeter equation with $P^2 = 0$ (nor should they, since they have a positive mass-squared). As $|A_t|$ increases, it reaches the lowest eigenvalue, at which point the gap equation has a non-trivial solution (namely, the corresponding eigenfunction). This eigenfunction, which gives the bound state wavefunction, signals the presence of the massless states discussed above. Hence, our goal is to find the lowest eigenvalue of the $P = 0$ equation, which we can solve in turn for the value of the coupling $|A_t|$ at which the bound state mass-squared becomes zero.

In the general case of a Bethe-Salpeter equation evaluated at $P = 0$, the Bethe-Salpeter wave function $\Psi(P, p)$ is the Fourier transform of the matrix element[1]

$$\psi(X, x) = \left\langle 0 \middle| T(\tilde{t}_L^\dagger(x_1)\tilde{t}_R(x_2)) \middle| P \right\rangle \tag{3.5}$$

with P conjugate to the center-of-mass coordinate $X = (1/2)(x_1 + x_2)$ and p conjugate to $x_1 - x_2$; the state $|P\rangle$ is the bound state. At $P = 0$ this looks like a vacuum-to-vacuum propagator $\Delta_{LR}(x_1 - x_2)$, but one must be careful when using this language. The true vacuum is a non-perturbative construct, which may in general be quite different from the bare vacuum (as in superconductivity, for example). In our case, the true vacuum has matrix elements connecting L and R stop squarks, which represent the usual mass terms. This connection comes from a symmetry-breaking order parameter that is a mass splitting $\delta M^2(p^2)$ found in this LR propagator which mixes L and R squarks. The order parameter obviously vanishes in the symmetric case. (This is analogous to the evolution of the $\psi\psi$ propagator in superconductivity as the temperature increases [44].)

To lowest order in this order parameter (that is, the matrix element connecting L and R stop squarks), the diagonal propagators of the $t_{L,R}$ fields are just those already shown in the

[1]We continue to use $\tilde{t}_L$ and $\tilde{t}_R$ for the legs; however, in the general case the particles are not necessarily squarks.

Bethe-Salpeter equation,

$$\Delta_{LL}(p) = \Delta_{RR}(p) = \frac{1}{p^2 + M^2}, \tag{3.6}$$

while the LR mixing propagator is

$$\Delta_{LR}(p) = \frac{1}{p^2 + M^2} \delta M^2(p^2) \frac{1}{p^2 + M^2}. \tag{3.7}$$

Above, we discussed the interpretation of the Bethe-Salpeter equation evaluated at $P = 0$ at the onset of spontaneous symmetry breaking. Let us now consider the behavior one expects as $|A_t|$ is further increased. As mentioned, at the onset of symmetry breaking we can think of the four massless states, previously considered as the elements of the complex bound state Higgs doublets, as one (real) bosonic degree of freedom (the Higgs boson) and three other (real) massless states. At larger $|A_t|$ we expect a tachyonic ($P^2 = m_{BS}^2 < 0$) solution [45]. When the theory is considered in the correct vacuum, one finds one massive real boson and three massless bosons, which are "eaten up" giving mass to the $W^{\pm}$ and Z bosons. This is completely analogous to the Standard Model, in which the Mexican-hat-shaped potential is stabilized by the non-perturbative vacuum, which results in a condensate and a massive Higgs boson.

Several times we have made reference to a gap equation. The gap equation is an integral equation whose solution is the symmetry-breaking order parameter; hence, a non-trivial solution corresponds to a non-trivial order parameter and hence symmetry breaking. To connect this to the picture of symmetry breaking already presented, we will show explicitly that the gap equation is essentially the Bethe-Salpeter equation at $P = 0$. Hence, the basic equation we use in our analysis can be motivated either by considering the appearance of a scalar field with negative mass-squared as the hallmark of electroweak symmetry breaking as, or one can take the development of a mass term which mixes left and right state as the hallmark sign. In the first case, one begins with the Bethe-Salpeter equation, and in the second case, one begins the gap equation, but both of these ultimately give the same equation. (Strictly speaking, the first scenario should be phrased in terms of the appearance

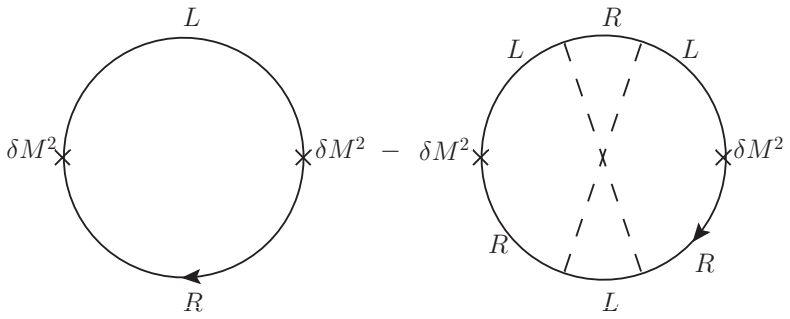

Figure 3.3: Diagrams for the effective potential quadratic in δM^2 (weights not shown).

of the massless Nambu-Goldstone bosons in the true vacuum, but these Nambu-Goldstone bosons are also a consequence of a non-trivial solution of the gap equation [44]. See [46] for a proof of this Nambu-Goldstone mechanism in gauge theories such as QCD.)

Let us derive the gap equation. We express the dynamics of symmetry breaking through the usual two-particle irreducible (2PI) effective potential Γ [47], in which Γ is a functional of $\delta M^2(p^2)$. As mentioned, in this section we ignore a number of interesting phenomena, including the VEV of the elementary Higgs field and its mixing with the bound-state Higgs; therefore, the effective potential (in the notation of [47]) is

$$\Gamma = \frac{1}{2}\mathrm{Tr}\left(\ln(G) + [1 - GG_0^{-1}]\right) + \text{2PI graphs}, \tag{3.8}$$

where the trace is over space-time as well as the other relevant indices. G is the exact propagator, and G_0 is the free propagator (when relevant; the term in square brackets is omitted for the LR propagator). The Schwinger-Dyson equations of the theory can be found by considering the extrema of Γ as the G are varied.

To lowest order in δM^2, the effective action Γ is given by the diagrams shown in Fig. 3.3, which give the following expression

$$\Gamma = \frac{1}{2}\int d^4p\, \rho(p^2)[\delta M^2(p^2)]^2 - \frac{A_t^4}{2(2\pi)^4}\int d^4p \int d^4k\, \rho(p^2)\delta M^2(p^2)K(p,k)\rho(k^2)\delta M^2(k^2), \tag{3.9}$$

where

$$\rho(k^2) = \frac{1}{(k^2+M^2)^2}. \tag{3.10}$$

Variation of the quadratic terms in Γ with respect to δM^2 yields

$$\delta M^2(p^2) = \frac{A_t^4}{(2\pi)^4}\int d^4k\,\frac{1}{k^2+M^2}\delta M^2(k^2)\frac{1}{k^2+M^2}K(p,k). \tag{3.11}$$

This is the promised gap equation, and just as for chiral gap equations, it is the original Bethe-Salpeter equation (in vertex form) at $P=0$. (To see that this is identical to the Bethe-Salpeter approach, set $P=0$ in Eq. (3.3), and then identify $Q(0,p)$ with $\delta M^2(p^2)$.) This illustrates again the necessary existence of composite Nambu-Goldstone bosons when symmetries are broken without elementary Higgs fields [44,46]. This gap equation differs from chiral symmetry breaking gap equations because the kernel is well-behaved in the ultraviolet limit. The kernel falls like $1/p^4$ (modulo logarithms) at large momentum, implying the same falloff for δM^2; thus Eq. (3.11) is finite.

To analyze Eq. (3.11) we need to first study the kernel $K(p,k)$, which has the form

$$K(p,k) = \frac{1}{(2\pi)^4}\int d^4\ell \frac{1}{[\ell^2+M^2][(p+\ell)^2+m^2][(k+\ell+p)^2+M^2][(k+\ell)^2+m^2]}. \tag{3.12}$$

Next we introduce Feynman parameters in order to perform the spacetime integral. Then

$$K(p,k) = \int\frac{d^4\ell}{(2\pi)^4}\int dx_1\,dx_2\,dx_3\,dx_4\frac{3!\delta(1-x_1-x_2-x_3-x_4)}{D_0^4}, \tag{3.13}$$

where

$$D_0 = x_4(\ell^2+M^2)+x_3[(\ell+p)^2+m^2]+x_2[(\ell+p+k)^2+M^2]+x_1[(k+\ell)^2+m^2], \tag{3.14}$$

$$= (x_1+x_2+x_3+x_4)\ell^2 + 2\ell\cdot[x_3p+x_2(k+p)+x_1k] + x_3p^2+x_2(k+p)^2+x_1k^2$$
$$+ M^2(x_2+x_4)+m^2(x_1+x_3). \tag{3.15}$$

We define a new variable of integration, $\ell' = \ell + x_3p + x_2(k+p) + x_1k$. The integral over ℓ' can be performed, which yields

$$K(p,k) = \frac{1}{16\pi^2}\int\prod_{i=1\ldots4}dx_i\;\delta\left(1-\sum_i x_i\right)\frac{1}{D^2}, \tag{3.16}$$

where

$$D = k^2(x_1x_2+x_3x_4)+p^2(x_1x_4+x_2x_3)+(p+k)^2x_2x_4+(p-k)^2x_1x_3+M^2(x_2+x_4)+m^2(x_1+x_3). \tag{3.17}$$

We assume that the squarks are significantly heavier than the fundamental Higgs boson; that is, $M^2 \gg m^2$. In this case, x_2, x_4 must be small compared to the other Feynman parameters, and so we write $x_2 = \lambda x$ and $x_4 = \lambda(1-x)$, with the new integration variables running from 0 to 1. The integral over λ will be dominated by small λ, so we can drop this variable judiciously. Then approximately

$$\prod_{i=1\ldots4} dx_i\,\delta\left(1-\sum_i x_i\right) = \lambda\,d\lambda\,dx\,dx_1\,dx_3\,\delta(1-x_1-x_3) \tag{3.18}$$

and

$$D = x_1x_3(p-k)^2 + m^2(x_1+x_3) + \lambda[ak^2+(1-a)p^2+M^2], \tag{3.19}$$

where

$$a = x_1x + x_3(1-x), 0 \le a \le 1, \tag{3.20}$$

and we dropped a term $\sim \lambda^2$ in D.

With this approximation, we can do the integral over λ explicitly, with the result

$$K(p,k) \approx \frac{1}{16\pi^2}\int dx_1\,dx_3\,dx\,\delta(1-x_1-x_3)\left[\frac{1}{A^2}\ln\left[\frac{A+B}{B}\right]-\frac{1}{A(A+B)}\right], \tag{3.21}$$

where

$$A = [ak^2+(1-a)p^2+M^2], \qquad B = x_1x_3(p-k)^2+m^2. \tag{3.22}$$

Let us discuss the high-energy behavior of the kernel. If $k^2 \gg p^2$, we may drop the terms in A and B which are proportional to p^2. Then B no longer depends on x and the x integral can be done analytically, giving

$$K \to \frac{\ln(k^2)}{8\pi^2k^4}. \tag{3.23}$$

Thus, the kernel vanishes rapidly at high momentum k; this confirms our earlier claim that the kernel was well-behaved in the ultraviolet regime. (This is important because the kernel goes inside an integral over d^4k in the gap equation, Eq. (3.11). If the kernel became infinite at large momenta k, the d^4k integral in the gap equation would diverge, as does the chiral gap equation.)

Now that we have reduced our kernel to the simpler form of Eq. (3.21), we return to studying the gap equation (or $P = 0$ Bethe-Salpeter equation) given by Eq. (3.11). The next step is to reduce it to a one-dimensional integral by integrating over the angles of k. We observe that the angle between the four-momenta p and k appears only in logarithms or in a term parametrically small with respect to M^2; this allows us to make the approximation

$$\int d\Omega_k F[(p-k)^2] \approx 2\pi^2 \left[\Theta(p^2-k^2)F(p^2) + \Theta(k^2-p^2)F(k^2)\right]. \tag{3.24}$$

This is exactly true for $F = 1/(p-k)^2$ and for constant F, and it is acceptable for the logarithmic functions we encounter.

Formally, this approximation is equivalent to projecting out the s-wave contribution to the kernel. To lowest (quadratic) order in δM^2, the relevant part of the effective potential Γ is

$$\Gamma = \frac{1}{2}\int dp^2\, p^2 \rho(p^2)[\delta M^2(p^2)]^2 - \frac{A_t^4}{2(2\pi)^4}\int dp^2\, p^2\, dk^2\, k^2 \rho(p^2)\delta m^2(p^2)\hat{K}(p^2,k^2)\rho(k^2)\delta M^2(k^2), \tag{3.25}$$

where $\hat{K}$ is the s-wave projection of the kernel, which is Eq. (3.21) after the approximation (3.24) has been made. Variation of this equation gives the s-wave projection of Eq. (3.11), which is a standard one-dimensional homogeneous Fredholm integral equation with a discrete spectrum of eigenvalues A_t^4. From either perspective, one arrives at the equation

$$\delta M^2(p^2) = \frac{A_t^4}{(2\pi)^4}\int dk\, \frac{k^3\, \delta M^2(k^2)}{(k^2+M^2)^2} K(p,k). \tag{3.26}$$

We seek the lowest eigenvalue of this equation using a variational principle: we insert a trial function into Eq. (3.25) and perform the integrals numerically. We then minimize the resulting eigenvalue with respect to the variational parameter that describes the family of trial functions. To motivate our trial functions, we consider first the known exact results for a kernel in which only a single, massless scalar is exchanged [32],

$$K \sim \frac{1}{(p-k)^2}. \tag{3.27}$$

For the vertex form of the Bethe-Salpeter equation, given in Eq. (3.3), the lowest eigenfunction at $P = 0$ is

$$Q(0,p) \equiv \delta M^2(p^2) \sim \frac{1}{p^2 + M^2}. \tag{3.28}$$

(A proof of this is given in the discussion of the Bethe-Salpeter equation in Ch. 2.)

In the present problem the asymptotic behavior of the kernel is different (as shown in Eq. (3.23)), so we chose the following as a zeroth-order trial function:

$$\delta M_0^2(p^2) \sim \frac{1}{p^4 + \mu^4}. \tag{3.29}$$

We have studied other trial functions, such as $1/(p^2 + \mu^2)^2$, with similar results. We improve this first variational estimate by using δM_0^2 as input to the right-hand side of Eq. (3.25) and numerically calculating a new output δM_1^2. We made a simple but accurate fit to δM_1^2, amounting to adding a term $\sim p$ to the denominator of Eq. (3.29). Then we used the average $\delta M_2^2 \equiv (1/2)(\delta M_0^2 + \delta M_1^2)$ as a trial function, and calculated the output again. This yields excellent agreement between the new input and output, as shown in Fig. 3.4, for the specific value $m/M = 0.05$.

At this specific mass ratio, the critical coupling resulting from our numerical calculations is

$$\frac{A_t}{M} \approx 15.14. \tag{3.30}$$

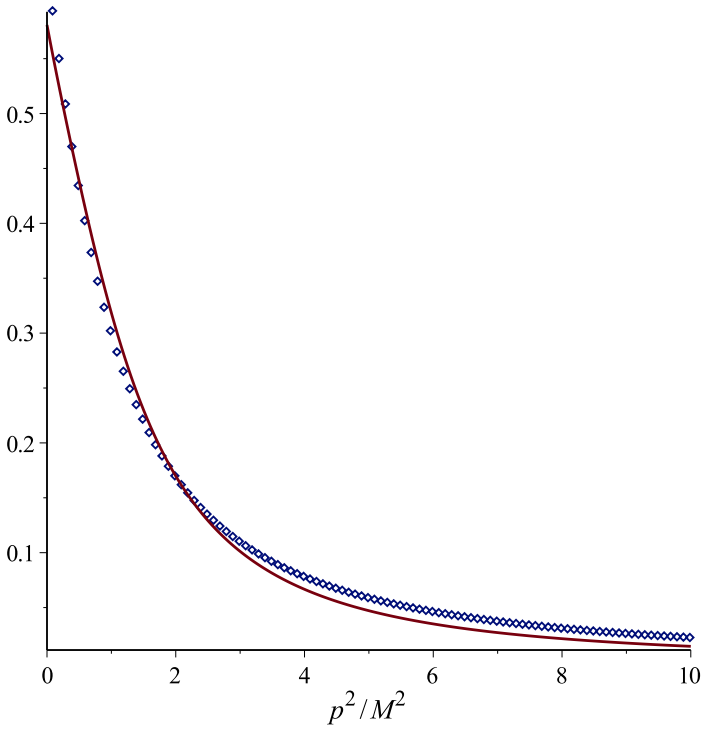

Figure 3.4: A comparison of the input and output using the second-order eigenfunction δM^2, as a function of p^2/M^2, calculated numerically as described in the text. In the case of the exact solution, the two curves would be identical.

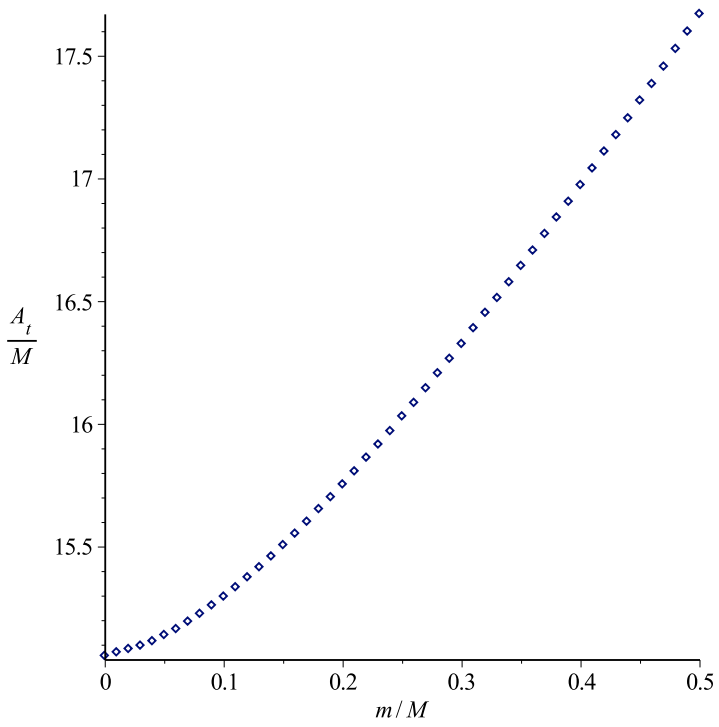

Figure 3.5: Behavior of the critical coupling as a function of the elementary Higgs mass.

Given our motivation of accommodating the observed Higgs boson mass of 125 GeV, it is useful to compare this result with the value $A_t/M = \sqrt{6}$, which maximizes the lightest Higgs mass in an approximate one-loop calculation [38, 48, 49]. We see that the coupling which breaks electroweak symmetry is substantially larger. However, this is not the final conclusion, since we have neglected mixing between the fundamental Higgs boson and the bound state Higgs boson; this will be discussed below. Before we do so, though, the behavior of the critical coupling as a function of the mass ratio is of interest. This is shown in Fig. 3.5. As expected, the critical coupling increases with increasing Higgs mass. We see that even as m/M approaches zero, the critical coupling remains well above $\sqrt{6}$.

3.4 Seesaw Symmetry Breaking: Effects of Including Mixing

The critical value determined above, $A_t/M \approx 15$, has another significant drawback, from a phenomenological perspective. To have a viable model, we must produce a Standard Model-like vacuum state, in which $\mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$ is broken but $\mathrm{SU_C}(3)$ remains a good symmetry. Even in its traditional realization, the MSSM has a number of dangerous color breaking minima, and even minima in which electric charge is not conserved [50–57].

If the Standard-Model-like vacuum is only a local minimum, one has to examine cosmological evolution; in many cases the vacuum that is populated first is the one in which the color symmetry is preserved. Thus, cosmological evolution may favor the vacuum with unbroken $\mathrm{SU_C}(3)$ even in some cases where it is not the global minimum of the potential [56, 58]. An approximate criterion for the absence of dangerous color and charge breaking minima, based on the analysis of [56], is

$$\frac{A_t}{M} \lesssim 2.7, \tag{3.31}$$

This is well-beneath the value of $A_t/M \approx 15$ found above.

In fact, the problem is even more pronounced in our model. In addition to the color-singlet states, the same trilinear scalar interactions can cause colored bound states to form. Since the scalar exchange forces are essentially color-blind, the bound states with different $\mathrm{SU_C}(3)$ properties can have similar binding energies. Therefore, the analysis performed above applies equally well to both the states carrying color and without color, and so we conclude that at $A_t/M \approx 15$ all acquire a nonzero vacuum expectation value, breaking $\mathrm{SU_C}(3)$. Instead, we prefer a vacuum in which $\mathrm{SU_C}(3)$ is preserved, in which the colored bound states are confined into color singlet through interactions with gluons, squarks, or quarks.

We will now show that by considering mixing between the fundamental Higgs boson and the bound state Higgs boson, electroweak symmetry breaking can occur for smaller ratios of A_t/M, which at least appears to allow us to escape the above problems. This alone is not

sufficient to ensure that the resulting model is phenomenologically viable; these issues, along with a more extended treatment of protecting color symmetry, are addressed in Ch. 4.

In our case, we must now re-examine the same issue of electroweak symmetry breaking, taking into account a number of new effective degrees of freedom, including those carrying color charge. While the full analysis is obviously complicated and the results will inevitably be model-dependent, there is one feature of the color-singlet states that sets them apart from the rest: they can have a mixing with the fundamental Higgs bosons. In fact, this mixing proceeds via the same coupling as that which enters the Bethe-Salpeter equation, and so the mass matrix in the bound-state-fundamental-Higgs basis necessarily has both diagonal and off-diagonal terms.

As the trilinear coupling increases, the mass-squared of all of the bound states decreases, as discussed above. However, due to the off-diagonal terms, a linear combination of the fundamental Higgs boson and the colorless bound state can develop a nonzero vacuum expectation value while the diagonal terms are positive. Since one of these diagonal terms represents the mass-squared of the colorless bound state, which is expected to be similar to the mass-squared of the color-charge-carrying bound states, we conclude that this leads to a vacuum in which $\mathrm{SU_C}(3)$ is not broken by a colored bound state VEV.

Let us now discuss this possibility in more detail. In this case, electroweak symmetry breaking occurs via a *symmetry-breaking seesaw*, when the following mass matrix has a negative eigenvalue:

$$\mathcal{M}_H^2 = \begin{pmatrix} \mu_H^2 & m_{HB}^2 \\ m_{HB}^2 & m_{BS}^2 \end{pmatrix}. \tag{3.32}$$

Here μ_H^2 is the Higgs mass-squared parameter, which is negative in the Standard Model, but which we take to be positive; m_{HB} is the mixing parameter calculated from the diagram shown in Fig. 3.6, and m_{BS} is the mass of the bound state. Our calculations above dealt with $m_{BS}^2 = P^2 = 0$, which required the large value of A_t/M quoted in Eq. (3.30). However, in

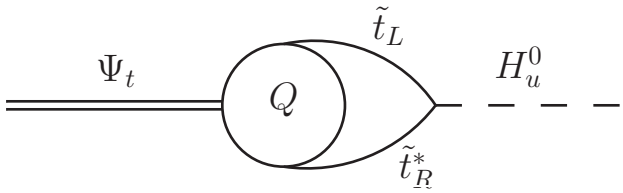

Figure 3.6: Mixing of the composite and fundamental Higgs fields.

the presence of mixing, the symmetry breaking can occur for a non-zero m_{BS} and a positive μ_H^2, as long as

$$\det(\mathcal{M}_H^2) = \mu_H^2 m_{BS}^2 - m_{HB}^4 < 0. \tag{3.33}$$

The value of the mixing parameter m_{HB} is determined by the strong dynamics; theoretically it can be extracted from a general solution of the Bethe-Salpeter equation (with nonzero bound state mass), with a proper wave function normalization. However, due to the complicated kernel, this calculation remains infeasible. Therefore, we parameterize the mixing as

$$m_{HB} = A_t \xi, \tag{3.34}$$

where ξ is dimensionless number (which may depend on A_t and on other parameters). The critical coupling (the value of A_t at which the mass matrix develops a zero eigenvalue) is generally different from that in Eq. (3.30); we have:

$$\frac{A_t}{M} = \min\left\{\frac{1}{\xi}\sqrt{\frac{\mu_H m_{BS}}{M^2}}, 15\right\}. \tag{3.35}$$

Given that μ_H can be arbitrarily close to zero (and, in fact, $\mu_H^2 < 0$ in the Standard Model), the symmetry breaking can occur even for smaller values of A_t/M. Of course, if μ_H^2 has to be fine-tuned to be small, the scenario in question becomes rather contrived. For reasonable values of parameters and for the binding energy that is not as large as the mass, which is the case for a smaller A_t, there is a self-consistent set of values, for example,

$$\mu_H \sim 10^2 \text{ GeV}, \qquad M_B \sim 2M \sim 2 \text{ TeV}, \qquad A_t \sim \frac{\sqrt{\mu_H m_{BS}}}{\xi} \sim \frac{450 \text{ GeV}}{\xi}, \tag{3.36}$$

and, thus the critical coupling is

$$\frac{A_t}{M} \sim \frac{0.45}{\xi}. \tag{3.37}$$

We see that this is consistent with Eq. (3.31) for some reasonable range of parameters. Thus, we conclude that electroweak symmetry breaking can occur for smaller values of the trilinear coupling, when one includes the mixing between the Higgs doublets.

In this chapter, we have shown that the presence of squark bound states in a strongly-interacting implementation of the MSSM can substantially modify electroweak symmetry breaking. At sufficiently large couplings, the bound state itself can break the symmetry; at smaller couplings, its presence is important due to the off-diagonal elements in the fundamental-bound-state-Higgs mass matrix. These elements allow a linear combination of the two states to break electroweak symmetry while protecting $\mathrm{SU_C}(3)$.

This scenario is clearly different from the widely discussed technicolor models [59], including walking technicolor [60, 61], models with color singlets [62], and the models in which supersymmetry and technicolor are combined [63, 64]. This scenario has the potential to relate the scales of supersymmetry breaking and electroweak breaking in a novel way, and it is evident that the effects of squark bound states should not be ignored in models with large trilinear supersymmetry breaking terms. However, we have not shown that strongly-coupling supersymmetry produces a phenomenologically viable model; this will be addressed in the following chapter.

CHAPTER 4

Phenomenology of Strongly-Interacting Supersymmetry

In the previous chapter, we introduced the seesaw-symmetry-breaking mechanism, in which a symmetry is spontaneously broken through the mixing of a fundamental Higgs field with a bound state. This was illustrated in a strongly-interacting version of the Minimal Supersymmetric Standard Model (MSSM), which differs from the usual MSSM in that the trilinear supersymmetry-breaking coupling is taken to be large [42, 43]. In this regime, squarks form bound states via the exchange of fundamental Higgs bosons, creating additional composite states, some of which can have nonzero vacuum expectation values (VEVs). In this chapter we will explore the phenomenology of this strongly-coupled model.

This model has very different particle content in the low-energy effective theory as compared to the usual MSSM; in particular, the model predicts more Higgs bosons and fewer superpartners at the electroweak scale. Secondly, as mentioned in Ch. 3, the usual MSSM bounds on the lightest Higgs mass are not applicable, because the bound-state quartic coupling is not related to the gauge coupling. These point to the possibility of a very different low-energy phenomenology than is typically associated with the MSSM, which is quite relevant for the ongoing supersymmetric searches at the LHC.

In Ch. 3, we did address one phenomenological issue: preserving $SU_C(3)$ symmetry, which motivated the introduction of the seesaw-symmetry-breaking mechanism. For the self-completeness of this chapter on phenomenology, we will briefly review how this mecha-

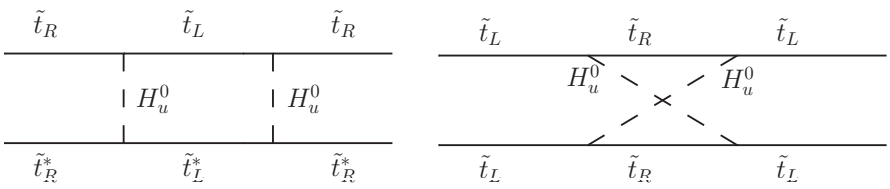

Figure 4.1: Kernels for squark bound states which produce $\mathrm{SU_L}(2)$ singlets and triplets.

nism protects against color-breaking minima. Because squarks carry color charge, the bound states come in both color octet and color singlet versions. If an octet state acquired a nonzero VEV, it would break $\mathrm{SU_C}(3)$, making the model phenomenologically unacceptable. However, to acquire a nonzero VEV through the seesaw-symmetry breaking mechanism, the bound state must mix with the fundamental Higgs boson, which is only possible for an $\mathrm{SU_C}(3)$ singlet state. Thus, it is the quantum numbers of the fundamental Higgs boson of the MSSM that determine the pattern of symmetry breaking. Color symmetry is protected because the fundamental Higgs doublet of the MSSM is colorless. (Please see Ch. 3 for more details.)

4.1 Description of Bound States

In Ch. 3, our primary focus was electroweak symmetry breaking, not phenomenology. Therefore we did not give a complete accounting of all the possible squark bound states. However, generically all of these states will appear in the low-energy effective theory. In addition to the $\mathrm{SU_L}(2)$ doublet which influences electroweak symmetry breaking, there are $\mathrm{SU_L}(2)$ singlets and $\mathrm{SU_L}(2)$ triplets; examples of the relevant kernels are shown in Fig. 4.1. The vertices in these bound states are all proportional to A_t; because Yukawa interactions are attractive in all channels, all of these states exist if the doublet bound state exists. We will assume that only $|A_t|$ (and possibly $|A_b|$) is large enough to produce bound states; these bound states appear in the up and down Higgs sectors respectively. We have summarized the possible bound states, along with their quantum numbers, in Table 4.1.

	Squarks	$SU_L(2)$ Charge	$SU_C(3)$ Charge	$U_Y(1)$ Charge
1	$\tilde{t}_L\tilde{t}_L$	$2 \otimes 2 = 3 \oplus 1$	$3 \otimes 3 = 6 \oplus \bar{3}$	$1/3$
2	$\tilde{t}_L^*\tilde{t}_L^*$	$\bar{2} \otimes \bar{2} = \bar{3} \oplus 1$	$\bar{3} \otimes \bar{3} = \bar{6} \oplus 3$	$-1/3$
3	$\tilde{t}_R\tilde{t}_R$	$1 \otimes 1 = 1$	$3 \otimes 3 = 6 \oplus \bar{3}$	$4/3$
4	$\tilde{t}_R^*\tilde{t}_R^*$	$1 \otimes 1 = 1$	$\bar{3} \otimes \bar{3} = \bar{6} \oplus 3$	$-4/3$
5	$\tilde{t}_L\tilde{t}_R^*$	$2 \otimes 1 = 2$	$3 \otimes \bar{3} = 1 \oplus 8$	$-1/2$
6	$\tilde{t}_R\tilde{t}_L^*$	$\bar{2} \otimes 1 = \bar{2}$	$3 \otimes \bar{3} = 1 \oplus 8$	$1/2$
7	$\tilde{t}_L\tilde{t}_L^*$	$2 \otimes \bar{2} = 3 \oplus 1$	$3 \otimes \bar{3} = 1 \oplus 8$	0
8	$\tilde{t}_R\tilde{t}_R^*$	$1 \otimes 1 = 1$	$3 \otimes \bar{3} = 1 \oplus 8$	0

Table 4.1: Quantum numbers of bound states; note that in SU(2), the anti-fundamental representation $\bar{2}$ is identical to the fundamental representation 2. Line numbers are provided for ease of reference in the text.

Let us summarize the states present in the model and their salient features. We observe that rows 1 and 2 in Table 4.1 are Hermitian conjugates of each other; consequently, these may be combined to form complex representations. Rows 3 and 4 may also be similarly combined. Thus, the first two rows in the table describe a complex $SU_L(2)$ triplet and a complex $SU_L(2)$ singlet, while the next two rows describe another complex $SU_L(2)$ singlet. All of these states carry color charge, and they also have fractional electric charge. The implications of these states will be discussed further in Sec. 4.6 on collider phenomenology.

Similarly, the next two rows (5 and 6) are also Hermitian conjugates of each other, which may be combined into a complex $SU_L(2)$. This is the doublet discussed in Ch. 3, which is responsible for the seesaw-symmetry-breaking mechanism. There are additionally eight colored states which do not generally acquire vacuum expectation values. The last two rows (7 and 8) describe two real singlets and one real triplet; these come in both colored and colorless versions. As we will show in Sec. 4.3, the electroweak phase transition is generally first order due to the two colorless singlets.

Thus, strongly-interacting supersymmetry has a rather extended Higgs sector. With so many degrees of freedom, it is important to ensure that none of the colored states acquire a nonzero VEV, and not just the color octet $\mathrm{SU_L}(2)$ states considered above. We note that there are no terms which involve only a single colored field, because the Lagrangian must be invariant under $\mathrm{SU_C}(3)$. Thus, even if all color singlet states acquire nonzero VEVs, there is no term linear in a colored field; such a term would necessarily induce spontaneous symmetry breaking of $\mathrm{SU_C}(3)$. Therefore, the effects of color singlet VEVs are limited to corrections to the masses and couplings of the colored fields. Provided that the pre-electroweak symmetry breaking masses are sufficiently large, the mass corrections will not drive the mass-squared values negative, and $\mathrm{SU_C}(3)$ will remain unbroken.

The full model, including both up and down sectors, is rather complicated. Therefore, we will make the simplifying assumption that the sectors are relatively decoupled, and we will only consider the up sector. It is also possible that the bound states form only in the up sector, if $|A_t|$ is large but $|A_b|$ is not.

Now that we have characterized the bound states, we will consider the resulting phenomenology. First, we will discuss flavor-changing neutral currents; this is a concern in any model with additional Higgs doublets. Next, we will investigate the electroweak phase transition in this model. Finally, we will make some remarks regarding collider phenomenology.

4.2 Flavor-Changing Neutral Currents

Any model which introduces additional Higgs doublets must address flavor-changing neutral currents (FCNCs), which are highly constrained experimentally and generically large when additional $\mathrm{SU_L}(2)$ doublets are introduced. One well-known method of suppressing FCNCs is to have the doublets couple to different types of quarks; for example, in the MSSM the Higgs doublet H_u couples to only up-type quarks and the doublet H_d couples to only down-type quarks. This suppresses FCNCs provided that the mixing between the two

doublets (after supersymmetry is spontaneously broken) is not too large; this assumption that the sectors are relatively decoupled is made both by the MSSM and our model.

However, even if the up sector and down sector are decoupled, our model potentially has large FCNCs because we have two doublets within each sector: both the fundamental doublet H_u and the bound-state doublet Φ_u couple to up-type quarks. Therefore, we consider a second way of suppressing FCNCs: if diagonalizing the quark matrix with respect to interactions with one of the doublets also approximately diagonalizes the quark matrix with respect to interactions with the other doublet, then FCNCs are small, because they are proportional to the off-diagonal elements. Equivalently, FCNCs are suppressed if the Yukawa couplings between the quarks and the first doublet are approximately proportional to the Yukawa couplings between the quarks and the second doublet. This approach is frequently disfavored because it typically requires fine-tuning, but we will demonstrate that this condition is naturally satisfied in strongly-coupled supersymmetry.

The bound-state doublet is comprised of up squarks exchanging H_u bosons. We note that there is no tree-level coupling between $u\bar{u}$ and $\tilde{u}\bar{\tilde{u}}$; however, there is a tree-level coupling between H_u and the quarks: the Yukawa coupling y_q. Therefore, to lowest order, an up-type quark sees only the H_u contribution in the bound state, and so the coupling between the quark and the bound state is $y'_q = \beta y_q$. This argument is shown diagrammatically in Fig. 4.2; the lowest order contribution to the Yukawa coupling between a quark and the bound-state Higgs doublet comes through the exchange of a true Higgs boson H_u. Therefore, it is proportional to the Yukawa coupling y_t, and the proportionality constant β describes the mixing between H_u and the bound state Φ_u. This mixing is clearly independent of the quark involved on the right-hand side of the diagram. Therefore, the Yukawa couplings satisfy $y'_q = \beta y_q$ with a single proportionality constant β. As claimed, FCNCs are indeed suppressed naturally, without fine-tuning, at tree-level.

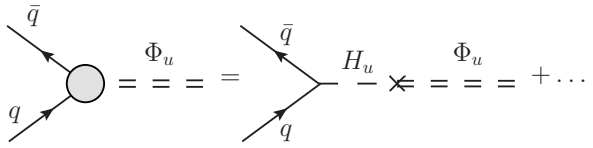

Figure 4.2: The lowest order diagram for the Yukawa coupling between quarks and the bound state Higgs doublet.

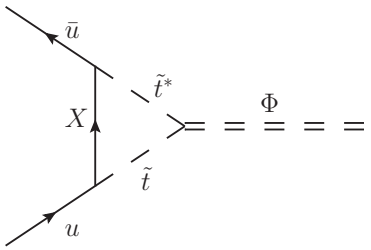

Figure 4.3: The next order diagram for the Yukawa coupling between up quarks and the bound state Higgs doublet.

We may be concerned about higher-order corrections, particularly to the extremely small up quark Yukawa coupling. The next order corrections to this coupling come from diagrams like those shown in Fig. 4.3; here, we take the squarks and quarks to be super-CKM eigenstates. The X fermion must convert a stop squark into an up quark; therefore, it must be either a gaugino or a higgsino. However, diagrams with a gaugino are suppressed by the MSSM alignment between squark mass matrices and Yukawa matrices, while diagrams with a higgsino are suppressed by off-diagonal CKM elements. Therefore, to suppress FCNCs at one-loop level, strongly-interacting supersymmetry requires no more fine-tuning than is present in the weakly-coupled MSSM.

For future reference, let us relate y_q and y'_q to the Yukawa couplings between the quark and the mass eigenstates; this will be relevant in our discussion of baryogenesis in below.

We write the mass eigenstates in terms of H_u and Φ_u by

$$\begin{aligned}\Psi_1 &= \cos(\theta)H_u + \sin(\theta)\Phi_u \\ \Psi_2 &= -\sin(\theta)H_u + \cos(\theta)\Phi_u.\end{aligned} \tag{4.1}$$

Then the relevant Yukawa couplings are given by

$$\begin{aligned}y_{1q} &= \cos(\theta)y_q + \sin(\theta)y'_q = (\cos(\theta) + \beta\sin(\theta))\, y_q \\ y_{2q} &= -\sin(\theta)y_q + \cos(\theta)y'_q = (-\sin(\theta) + \beta\cos(\theta))\, y_q.\end{aligned} \tag{4.2}$$

In particular we note that

$$y_{2q} = \frac{-\sin(\theta) + \beta\cos(\theta)}{\cos(\theta) + \beta\sin(\theta)} y_{1q}, \tag{4.3}$$

and we expect $\beta \sim \sin(\theta)$ due to its close relation to the mixing.

4.3 Temperature Evolution and Effective Potential

In this section, we will discuss the low-energy effective potential. Before we do so, it is useful to consider the evolution of the model with temperature. We recall from Ch. 3 that it is $|A_t|$, and not A_t, which is relevant found bound state formation and electroweak symmetry breaking. Since a Yukawa interaction is always attractive, the sign of the trilinear coupling was completely irrelevant to our analysis. Ref. [38] shows that if A_t is negative (with $|A_t|$ large, as required for the bound states to exist), then A_t increases as the renormalization scale increases. Consequently, at sufficiently high temperatures, the model behaves as the standard (weakly interacting) MSSM. At lower temperatures, the model undergoes a phase transition to a strongly interacting phase, and electroweak symmetry breaking takes place. This is in somewhat analogous to QCD, which is described by quarks and gluons at high temperatures, but at lower temperatures the bound-state baryons and mesons are the appropriate degrees of freedom.

If A_t is large and positive, the same reference suggests that there is no energy scale in which perturbativity is regained. Therefore, there is no regime in which it is appropriate to

discuss individual squarks, and one may argue that the "bound states" should instead be considered fundamental. For ease of exposition, we will assume that we are in the former scenario, although the features of the baryogenesis phase transition to be discussed below do not depend on this assumption. Additionally, as mentioned in Ch. 3, in this case it is evident that the theory retains the ultraviolet behavior of the MSSM.

In the strongly-coupled phase, the model should be described by an effective Lagrangian written in terms of the low-energy degrees of freedom, which include the bound states. Ideally, one would like to calculate all the parameters in the effective potential in terms of the parameters of high-energy weakly-coupled MSSM. However, because the theory is strongly coupled, it is not feasible to calculate these parameters explicitly. A calculation on the lattice may be possible [65], but no detailed results are available at present. In the absence of such a calculation, one can only estimate the low-energy couplings using generic values consistent with symmetries. (This is similar to the approach used to study strong interactions before QCD was discovered and understood.) Obviously, the predictive power of this approach is limited. However, since the values of "fundamental" MSSM parameters in the high-energy Lagrangian are unknown and not strongly constrained, this approach appears to be well justified.

We note that the colorless $SU_L(2)$ gauge singlets are not associated with any symmetry breaking; therefore, they may acquire nonzero VEVs in the strongly-coupled phase. Such vacuum expectation values have no physical meaning and may be removed with a field redefinition that makes the tadpole diagrams vanish order by order in perturbation theory. We will assume that this has been done in writing the effective potential. We have already shown that the colored fields do not acquire nonzero VEVs, and we will neglect them for the remainder of this section.

The notation used in this section is as follows: the complex $SU_L(2)$ doublet mass eigenstates are Ψ_1 and Ψ_2, where Ψ_2 has a negative mass-squared eigenvalue due to the seesaw

symmetry-breaking mechanism, the real $\mathrm{SU_L}(2)$ singlet mass eigenstates are S_1 and S_2, and the real $\mathrm{SU_L}(2)$ triplet field is V. Then the effective potential, including all the color singlet fields mentioned in 4.1, can be written as

$$\begin{aligned} V(S_1, S_2, \Psi_1, \Psi_2, V) &= V_2(S_1, S_2, \Psi_1, \Psi_2, V) \\ &+ V_{3a}(\Psi_1, \Psi_2, S_1) + V_{3b}(\Psi_1, \Psi_2, S_2) + V_{3c}(S_1, S_2, V) + V_{3d}(\Psi_1, \Psi_2, V) + V_{3e}(S_1, S_2) \\ &+ V_{4a}(\Psi_1, \Psi_2) + V_{4b}(\Psi_1, \Psi_2, S_1, S2) + V_{4c}(\Psi_1, \Psi_2, V) + V_{4d}(S_1, S_2, V) + V_{4e}(S_1, S_2), \end{aligned} \tag{4.4}$$

where the mass terms are

$$V_2(S_1, S_2, \Psi_1, \Psi_2, V) = -m_1^2 \Psi_1^\dagger \Psi_1 + m_2^2 \Psi_2^\dagger \Psi_2 + m_{S1}^2 S_1^2 + m_{S2}^2 S_2^2 + m_V^2 V^T V. \tag{4.5}$$

One of the doublet mass eigenvalues is negative due to the seesaw symmetry breaking mechanism, and we emphasize that the triplet, V, is real. The cubic terms are

$$\begin{aligned} V_{3a}(\Psi_1, \Psi_2, S_1) &= A_{S1} S_1 \Psi_1^\dagger \Psi_1 + A_{S2} S_1 \Psi_2^\dagger \Psi_2 + A_{S12} S_1 \Psi_1^\dagger \Psi_2 + h.c., \\ V_{3b}(\Psi_1, \Psi_2, S_2) &= \tilde{A}_{S1} S_2 \Psi_1^\dagger \Psi_1 + \tilde{A}_{S2} S_2 \Psi_2^\dagger \Psi_2 + \tilde{A}_{S12} S_2 \Psi_1^\dagger \Psi_2 + h.c., \\ V_{3c}(S_1, S_2, V) &= A_{SV} S_1 V^T V + \tilde{A}_{SV} S_2 V^T V, \\ V_{3d}(\Psi_1, \Psi_2, V) &= A_{V1} \Psi_1^\dagger (\boldsymbol{\sigma} \cdot \boldsymbol{V}) \Psi_1 + A_{V2} \Psi_2^\dagger (\boldsymbol{\sigma} \cdot \boldsymbol{V}) \Psi_2 + A_{V12} \Psi_1^\dagger (\boldsymbol{\sigma} \cdot \boldsymbol{V}) \Psi_2 + h.c., \\ V_{3e}(S_1, S_2) &= A_S S_1^3 + \tilde{A}_S S_2^3 + A' S_1^2 S_2 + A'' S_1 S_2^2. \end{aligned} \tag{4.6}$$

Finally, the possible quartic terms are

$$\begin{aligned}
V_{4a}(\Psi_1,\Psi_2) &= \lambda_1(\Psi_1^\dagger\Psi_1)^2+\lambda_2(\Psi_2^\dagger\Psi_2)^2+\lambda_{12}(\Psi_1^\dagger\Psi_1)(\Psi_2^\dagger\Psi_2)+\lambda'_{12}(\Psi_1^\dagger\Psi_2)^2\\
&\quad+\lambda_1(\Psi_1\boldsymbol{\tau}\Psi_1)\cdot(\Psi_1\boldsymbol{\tau}\Psi_1)+\lambda_2(\Psi_2\boldsymbol{\tau}\Psi_2)\cdot(\Psi_2\boldsymbol{\tau}\Psi_2)\\
&\quad+\lambda_3(\Psi_1\boldsymbol{\tau}\Psi_1)\cdot(\Psi_2\boldsymbol{\tau}\Psi_2)+\lambda_4(\Psi_1\boldsymbol{\tau}\Psi_2)\cdot(\Psi_1\boldsymbol{\tau}\Psi_2),\\
&\quad+\lambda''_{12}(\Psi_1^\dagger\Psi_2)(\Psi_2^\dagger\Psi_1)+h.c.\\
V_{4b}(\Psi_1,\Psi_2,S_1) &= \lambda_{S1}S_1^2\Psi_1^\dagger\Psi_1+\lambda_{S2}S_1^2\Psi_2^\dagger\Psi_2+\tilde\lambda_{S1}S_2^2\Psi_1^\dagger\Psi_1+\tilde\lambda_{S2}S_2^2\Psi_2^\dagger\Psi_2\\
&\quad+\lambda_{S12}S_1^2\Psi_1^\dagger\Psi_2+\tilde\lambda_{S12}S_2^2\Psi_1^\dagger\Psi_2+\lambda'_{S12}S_1S_2\Psi_1^\dagger\Psi_2+h.c.,\\
V_{4c}(\Psi_1,\Psi_2,V) &= \lambda_{V1}\Psi_1^\dagger\Psi_1V^TV+\lambda_{V2}\Psi_2^\dagger\Psi_2V^TV+\lambda_{V12}\Psi_1^\dagger\Psi_2V^TV+h.c.,\\
V_{4d}(S_1,S_2,V) &= \lambda_{SV}S_1^2V^TV+\tilde\lambda_{SV}S_2^2V^TV+\lambda_{S12}S_1S_2V^TV+\lambda_V(V^TV)^2,\\
V_{4e}(S_1,S_2) &= \lambda_SS_1^4+\tilde\lambda_SS_2^4+\lambda'_SS_1^3S_2+\tilde\lambda'_SS_1S_2^3+\lambda_{SS}S_1^2S_2^2.
\end{aligned}\tag{4.7}$$

We note that the following 8 parameters are generally complex: A_{S12}, $\tilde A_{S12}$, A_{V12}, λ'_{12}, λ_{S12}, $\tilde\lambda_{S12}$, λ'_{S12}, and λ_{V12}.

Due to the negative mass squared of Ψ_1, the origin of the potential is not a local minimum and the neutral component of doublet Ψ_1 acquires a nonzero vacuum expectation value. The terms $A_{S1}S_1\Psi_1^\dagger\Psi_1$ and $\tilde A_{S1}S_2\Psi_1^\dagger\Psi_1$ produce terms linear in S_1 and S_2 respectively, and so consequently neither $\langle S_1\rangle=0$ nor $\langle S_2\rangle=0$ can be a local minimum. Hence, once Ψ_1 acquires a nonzero vacuum expectation value, the singlets also acquire a nonzero vacuum expectation value.

When both Ψ_1 and the singlets have acquired nonzero vacuum expectation values, the neutral component of the other doublet, Ψ_2, must also acquire a nonzero VEV due to terms such as $S_1\Psi_1^\dagger\Psi_2$; however, the charged component does not acquire a nonzero VEV. Next let us consider the triplet; we parameterize it as $V^a=(V_1,V_2,V_3)$. The cubic terms in Eq. 4.6 include

$$\Psi_1^\dagger\sigma^aV^a\Psi_1=\Psi_1^\dagger\begin{pmatrix}V_3 & V_1-iV_2\\ V_1+iV_2 & V_3\end{pmatrix}\Psi_1=\Psi_1^\dagger\begin{pmatrix}V^0/2 & -V^+/\sqrt{2}\\ -V^-/\sqrt{2} & -V^0/2\end{pmatrix}\Psi_1,\tag{4.8}$$

where we have identified the charge states. When the neutral component of Ψ_1 acquires a nonzero VEV, $\Psi_1^\dagger \sigma^a V^a \Psi_1$ produces a term linear in V^0; consequently, this field also acquires a nonzero vacuum expectation value. (The consequences of this for the ρ_0 parameter and neutrino masses are discussed in Sec. 4.5 below.)

The presence of gauge singlet Higgs states and the existence of tree-level cubic couplings generically make the phase transition strongly first order. Because of these singlets, the potential includes cubic terms. When the fields are expanded about their vacuum expectation values, the cubic terms give terms linear in the VEVs. Such linear terms produce a barrier which results in a strongly first order phase transition. This is identical to the manner in which singlets produce a barrier in the Next-to-Minimal Supersymmetric Standard Model (NMSSM) [66–69]. This is in contrast with the Standard Model and MSSM, in which the cubic terms are forbidden by the $\mathrm{SU_L}(2)$ symmetry. In the Standard Model, the transition is not first order for a Higgs mass above 45 GeV. In the MSSM, the transition is weakly first order, and only for such parameters for which the two-loop corrections generate a sufficient barrier in the potential [70–73].

We note that in strongly-coupled supersymmetry, finite temperature corrections produce terms proportional to $T^2 M^2$ in the potential, where M^2 are the mass eigenvalues of the shifted fields as functions of the VEVs; this is the same as in the Standard Model [74, 75]. Hence, increasing the temperature will restore electroweak symmetry. However, a comprehensive study of the temperature evolution of the potential is infeasible due to the large number of parameters in the cubic and quartic terms. However, simply because there are so many parameters, we expect there to be a region of parameter space in which the phase transition occurs at a temperature of the order of the electroweak scale and with $v_u = \sqrt{|\langle \Psi_1 \rangle|^2 + |\langle \Psi_2 \rangle|^2} \lesssim 247$ GeV, with the difference to be provided by the down sector.

4.4 Baryogenesis

The strongly first-order phase transition reopens the possibility of baryogenesis at the electroweak scale [76, 77]. In addition to the first-order phase transition, additionally CP violation beyond that in the Standard Model CKM matrix is necessary. The full potential written in Sec. 4.3 has eight complex parameters; two of these may be eliminated by rotating the complex doublets Ψ_1 and Ψ_2. This leaves six physical phases representing new sources of CP violation. If $|A_b|$ is large, there are similar phases in the down sector, and mixing between the sectors can give additional CP-violating phases.

This CP violation in the Higgs sector must be communicated to the matter sector for successful baryogenesis. This can be accomplished through interactions in the bubble wall with the top quark; only the top quark Yukawa coupling is sufficiently large for the interactions to be in thermal equilibrium during the phase transition. Our analysis will follow that of Ref. [78], which considered a simpler model with two Higgs doublets and a complex singlet, of which only one doublet and the singlet acquired a nonzero VEV.

In general, both mass eigenstates Ψ_1 and Ψ_2 couple to up-type fermions, and thus the effective Lagrangian contains terms of the form

$$-y_q \epsilon_{ab} T_L^a \Psi_1^b \bar{t}_R - y'_q \epsilon_{ab} T_L^a \Psi_2^b \bar{t}_R + h.c., \tag{4.9}$$

where $T_L = (t_L, b_L)$ is the doublet which includes the left-handed top and bottom quarks, and a, b are $\mathrm{SU_L}(2)$ indices. We recall that y_q and y'_q are proportional to each other, as described by Eq. 4.3, which suppresses FCNCs. If we write the vacuum expectation values of the doublets after spontaneous symmetry breaking as

$$\langle \Psi_1 \rangle = \begin{pmatrix} 0 \\ \xi_1 e^{i\theta_1} \end{pmatrix}, \qquad \langle \Psi_2 \rangle = \begin{pmatrix} 0 \\ \xi_2 e^{i\theta_2} \end{pmatrix}, \tag{4.10}$$

then these terms become

$$\left(y_q \xi_1 e^{i\theta_1} + y'_q \xi_2 e^{i\theta_2} \right) t_L \bar{t}_R + h.c.. \tag{4.11}$$

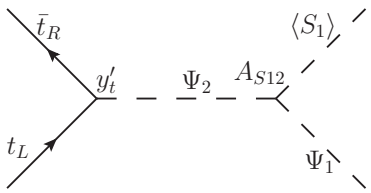

Figure 4.4: One of the diagrams that modifies the phase of the top quark Yukawa coupling.

This is the usual top quark mass term; the nonzero phase is absorbed in a rotation of the top quark. To simplify our analysis, we will assume that $\xi_1 \gg \xi_2$; this is expected since the mass term of Ψ_1 in the effective potential is negative but the mass term of Ψ_2 is positive. Therefore, the dominant contribution to the top quark's mass is from the Yukawa coupling between Ψ_1 and the top quark, and thus we take $y_t \approx \sqrt{2} m_t / v = 0.996$. We will define the other vacuum expectation values to be $\langle S_1 \rangle = \xi_3$, $\langle S_2 \rangle = \xi_4$, and $\langle V_3 \rangle = 2 \langle V^0 \rangle = \xi_5$.

When we consider the nonzero vacuum expectation values of the other fields, this top quark Yukawa coupling is modified, and these modifications can introduce a nonzero physical phase into the coupling. An example of one such contribution is shown in Fig. 4.4. The tree-level corrections to the top quark Yukawa coupling after spontaneous symmetry breaking are given by

$$y_{tf} = y_t + \tilde{y}_{tf} + \frac{y'_t}{m_2^2} \Big(A_{S12}\xi_3 + \tilde{A}_{S12}\xi_4 + A_{V12}\xi_5 + \lambda''_{12}\xi_1\xi_2 e^{i(\theta_1-\theta_2)} + \lambda'_{12}\xi_1\xi_2 e^{i(\theta_2-\theta_1)} + \lambda'_{S12}\xi_3\xi_4 + \lambda_{V12}\xi_5^2 \Big) , \tag{4.12}$$

where $\tilde{y}_{tf}$ summarizes the contributions from diagrams that do not contribute a net phase. (We have used the freedom to rotate Ψ_1 and Ψ_2 to make the parameters λ_{S12} and $\tilde{\lambda}_{S12}$ real; we also recall that we have set up our effective Lagrangian such that the singlets have zero VEV before spontaneous symmetry breaking.)

Let us assume that the corrections are small with respect to y_t; then this corrected Yukawa coupling may be written as

$$y_{tf} \approx y_t e^{i\phi_f}, \tag{4.13}$$

where

$$\phi_f = \frac{y_t'}{y_t m_2^2} \Im \Big(A_{S12}\xi_3 + \tilde{A}_{S12}\xi_4 + A_{V12}\xi_5 + \lambda_{V12}\xi_5^2 + \lambda_{12}''\xi_1\xi_2 e^{i(\theta_1-\theta_2)} + \lambda_{12}'\xi_1\xi_2 e^{i(\theta_2-\theta_1)} + \lambda_{S12}'\xi_3\xi_4 \Big) \tag{4.14}$$

This change of phase in the Yukawa coupling can be transformed according to the standard techniques [79, 80]; the quarks are rotated by an amount proportional to their hypercharge to eliminate the phase, which introduces a new kinetic term for the top quark which violates CP. The phase ϕ_t can be approximated as space independent, although time dependent, because the mean free path of the top quarks and gauge bosons is small compared to the scale on which ϕ_t varies (which is approximately the thickness of the electroweak bubble walls). Then this additional term in the Lagrangian has the form of a chemical potential for baryon number. Consequently, during the transition the free energy is minimized for nonzero baryon number. The equilibrium baryon density is

$$n_{B,eq} = \alpha \frac{T^2}{6} \dot{\phi}_t, \tag{4.15}$$

where α is a constant of order 1; for a simple two-doublet model, it is 72/111 [79]. During the phase transition, the sphaleron-induced $B - L$ violation processes drive the system toward this equilibrium value at the rate [81]:

$$\frac{dn_B}{dt} = 18 \frac{\Gamma_{sp}}{T^3} n_{B,eq}, \tag{4.16}$$

but the minimum of the free energy is not reached because the transition occurs too quickly. The sphaleron transition rate is [76, 77, 80]

$$\Gamma_{sp} = \begin{cases} \kappa(\alpha_W T)^4, & m_W \leq \sigma \alpha_W T, \\ \gamma(\alpha_W T)^{-3} M_W^7 e^{-E_{sp}/T} \approx 0, & m_W > \sigma \alpha_W T, \end{cases} \tag{4.17}$$

where κ, σ, and γ are dimensionless constants. For the Standard Model, κ is expected to be between 0.1 and 1 [82], while σ is expected to be between 2 and 7 [81]. Integrating the above rate gives the baryon asymmetry produced during the phase transition,

$$n_B = 3\alpha\kappa\alpha_W^4 T^3 \Delta\phi_t, \tag{4.18}$$

where $\Delta\phi_t$ is the change in the phase of the top quark Yukawa coupling during the phase transition; this is not necessarily ϕ_f because the sphaleron $B+L$-violating interactions may go out of thermal equilibrium before the phase transition is complete. The entropy density is

$$s = \frac{2\pi^2}{45} g_S(T) T^3, \tag{4.19}$$

and so baryon-to-entropy ratio after the phase transition is

$$\frac{n_B}{s} = \frac{135\alpha}{2\pi^2 g_S(T_{EW})} \kappa\alpha_W^4 \Delta\phi_t. \tag{4.20}$$

To match the observed value of $n_B/s \sim 10^{-10}$, the change in phase must be of order 10^{-2} (assuming $g_S(T_{EW}) \sim 100$). This is a reasonable number; given the form of Eq. (4.14), we expect this to be satisfied for a relatively large region of the parameter space of the effective potential. Thus, we conclude that the electroweak phase transition in the strongly-coupled MSSM can account for the observed matter asymmetry.

4.5 Implications of the Triplet Vacuum Expectation Value

We have noted in Sec. 4.3 that it is an unavoidable consequence of this model that the neutral component of the hypercharge $Y = 0$ Higgs triplet acquires a nonzero vacuum expectation value. In this section, we discuss the phenomenological consequences of this. Models in which a single $Y = 0$ Higgs triplet acquires a vacuum expectation value have been considered [83–87]; the low-energy behavior of this theory was described in detail in Ref. [88]. Such models are quite constrained by precision measurements of the ρ_0 parameter, which is experimentally measured to be $\rho_0 = 1.0004^{+0.0003}_{-0.0004}$ [89]. A $Y = 0$ triplet nonzero VEV $\langle V^0 \rangle$

modifies ρ_0 by [88]

$$\Delta\rho_0 = \frac{4|\,\langle V^0\rangle\,|^2}{v_u^2}. \tag{4.21}$$

We recall that we must have $v_u \leq 246$ GeV; this means that we must have $|\,\langle V^0\rangle\,| = |\,\langle V_3\rangle\,|/2 \lesssim 2.5$ GeV, or equivalently,

$$\frac{|\,\langle V_3\rangle\,|}{v_u} \leq 10^{-2}. \tag{4.22}$$

If $v_u \approx |\,\langle \Psi_1\rangle\,|$, we expect $|\,\langle V_3\rangle\,| \approx A_{V1}v_u^2/m_V^2$; the above condition becomes

$$\frac{A_{V1}v_u}{m_V^2} \leq 10^{-2}. \tag{4.23}$$

We expect A_{V1} and m_V, like the other parameters in the effective Lagrangian, to be near the electroweak scale. The exact values of A_{V1} and m_V are determined from the high-energy (MSSM) Lagrangian through strong dynamics, and it is infeasible to calculate analytic expressions for them. It may be that a lattice calculation shows that triplet states are less strongly bound than the singlet states, and thus have larger masses, or it may be that this model requires some fine-tuning to satisfy this condition. We note that if A_{V1} and v_u are both on the 100 GeV scale, then we only require $m_V \sim \mathcal{O}(\text{TeV})$.

Another concern with the triplet acquiring a nonzero vacuum expectation value is that generically such vacuum expectation values may produce large neutrino masses [90–93]. We recall the seesaw mass matrix of the neutrino sector,

$$M_\nu = \begin{pmatrix} \delta & M \\ M & D \end{pmatrix}, \tag{4.24}$$

where δ and D are Majorana mass terms for the left- and right-handed neutrinos respectively, and M is the Dirac mass term. Phenomenologically, it is required for δ to be small; since D is assumed to include a large contribution from high energy physics, any large corrections to M can be offset by requiring a larger contribution to D. We argue that the nonzero triplet VEV does not alter any of these terms.

We note that lepton number is conserved in our high-energy theory (MSSM), and forming squark bound states will not break lepton number conservation. Consequently, lepton number is conserved in our low-energy effective theory, and this symmetry forbids our low-energy theory from generating Majorana mass terms, including δ. (Of course, as in the MSSM, if one wishes to use the seesaw mechanism to produce small neutrino masses, one must insert lepton-number breaking terms "by hand" to create the Majorana masses; in this respect, our model is no different than the MSSM. Our point is that despite having a nonzero triplet VEV, we naturally do not have large contributions to the Majorana masses.) Furthermore, regarding the Dirac mass term, we note that $V \to \bar{\nu}_L \nu_R$ diagrams are forbidden by $\mathrm{SU_L}(2)$ symmetry, and therefore the triplet VEV does not alter M either. Thus, although our triplet does acquire a nonzero vacuum expectation value, it does not affect neutrino masses in any way.

4.6 Collider Phenomenology

Finally, we make some qualitative remarks regarding the collider phenomenology of this model. As we have shown in Table 4.1, this model has a rather extended Higgs sector, with numerous states. As a result, it will be difficult to discern individual states at an experiment such as the Large Hadron Collider, although broad excess may be detected. There may also be additional observable consequences, although further study (including numerical simulations) would be necessary to make concrete predictions.

The gauge singlet states can be detected via deviations of the Higgs decay branching ratios from the predictions of the Standard Model [94, 95]. Many of the Higgs states present in this model carry color charge; this is in contrast to the Standard Model and the weakly interacting MSSM, in which the Higgs sector contains only colorless states. Again, due to the large number of such states they may be difficult to discern individually; however, these states may influence the number and structure of jets observed in high-energy scattering

processes.

Secondly, we have noted the presence of a $Y = 4/3$ triplet which carries color charge. Since $\mathrm{SU_C}(3)$ symmetry is preserved, these states must form a colorless combination by joining with other colored particles, most frequently by pulling quarks from the quantum vacuum. This process produces jets with an integer charge. However, some of these jets will carry charge ± 2, for example, if a $\tilde{t}_L \tilde{t}_L$ bound state combines with an up quark. These unusual jets may be experimentally detectable.

Additionally, this model predicts numerous charge ± 1 states; these arise from the extra doublet as well as the triplets. The Standard Model, in contrast, only has an electrically neutral Higgs boson, while the MSSM has one set of ± 1 charged Higgs bosons. Furthermore, some of the singly charged states carry color charge, again in contrast to the MSSM. Therefore, searches for charged scalar bosons may produce evidence for strongly-coupled supersymmetry.

4.7 Conclusions

As an example of a strongly-interacting extension of the Standard Model with novel features, we have considered strongly-interacting supersymmetry. In Ch. 3, we showed a particularly interesting theoretical feature: that the resulting bound states can mix with the fundamental Higgs doublet, driving electroweak symmetry breaking. Although we illustrated this seesaw-symmetry-breaking mechanism with strongly-interacting supersymmetry, the mechanism has broader applications to any model in which bound states can mix with other scalar states. As applied to supersymmetry, it has the theoretical advantage of uniting supersymmetry breaking, which produces the large trilinear couplings that lead to bound states, and electroweak symmetry breaking.

However, it is not sufficient for a model to be theoretically interesting; it also must pass a series of rigorous constraints set by decades of experimental observations. In Ch. 4, we studied the phenomenology of strongly-interacting supersymmetry. In particular, we found that in addition to protecting $\mathrm{SU_C}(3)$ symmetry, the strongly-interacting MSSM naturally avoided large flavor-changing neutral currents. Furthermore, this model can easily accommodate electroweak scale baryogenesis, in contrast to the Standard Model and the MSSM.

Thus, we have seen that going beyond the perturbative limit can lead to interesting theoretical possibilities and novel phenomenology that may be absent in weakly-coupled versions.

CHAPTER 5

Indirect Detection Signals in Asymmetric Dark Matter Models

5.1 A Brief Introduction to Asymmetric Dark Matter

We now turn our attention away from bound states in strongly-interacting supersymmetry and to dark matter bound states. As was discussed in Ch. 1, one shortcoming of the Standard Model is that it fails to provide a candidate particle for dark matter, and so numerous extensions have been proposed. In this chapter and the next, we consider asymmetric dark matter models. In these models, dark matter carries its own conserved particle number, and interactions between the dark sector and the Standard Model gave rise to the observed baryon asymmetry of the universe along with a corresponding asymmetry in the dark sector [96–116]. (For a recent review, see Ref. [29].)

Asymmetric dark matter models typically include self-interactions; this is necessary for particles and antiparticles to annihilate in the universe, leaving a remnant determined by the asymmetry. Such self-interactions may be also be motivated by several inconsistencies between numerical simulations of collisionless cold dark matter and astronomical observations [25–28, 117–128]. Self-interactions between dark matter particles can facilitate momentum transfer and angular momentum transfer in halos, creating cored rather than cuspy density profiles in both dwarf spheroidal galaxies and in larger halos. Secondly, the number of subhalos of Milky-Way-sized dark matter halos produced in the simulations exceeds the number of observed dwarf galaxies of the Milky Way, and furthermore, the most massive

subhalos produced in the simulations are too dense to fail to produce bright stars. Self-interactions can significantly modify small halos, although too much interaction can destroy the ellipticity of large halos.

These self-interactions can take a number of different forms in asymmetric dark matter models. Examples include hidden-sector particles with gauge bosons [124], Yukawa interactions [129], and even non-topological solitons with a large enough geometrical size [120]. We will consider two different types of interactions. In this chapter, we study a simplistic model consisting only of scalars. In Ch. 6, we introduce a spontaneously broken gauge symmetry in the dark sector.

In this chapter, we will demonstrate that although the typical indirect detection signal from self-annihilation is absent in asymmetric dark matter models, dark matter bound state formation can provide a detectable signal. We will also discuss potential signals from bremsstrahlung emission of the dark mediator, although we will the it is difficult to produce a detectable signal above the large astrophysical backgrounds.

While this chapter is narrowly focused on the question of indirect detection signals, the subsequent chapter considers a variety of topics. In particular, we will argue that in this atomic dark matter model (with a broken U(1) gauge symmetry), dark matter in generally multi-component, including dark atoms. The cosmology is consequently more involved, and the possibility of several species with multiple inter- and intra-species interactions alters the parameter space in which self-interactions modify halos. Thus, these two chapters will establish that dark matter bound states have implications for indirect detection searches, cosmology, and astrophysics.

5.2 Scalar Self-Interacting Asymmetric Dark Matter Model

In most asymmetric dark matter models, indirect detection signals from dark matter self-annihilation are absent at present, as it is assumed that dark matter today consists of an asymmetric population of dark particles with few, if any, antiparticles. We will investigate the possibility of generating indirect detection signals via bound state formation (Sec. 5.3) and bremsstrahlung emission of a dark mediator (Sec. 5.4). We will consider a fairly generic model of scalar dark matter S interacting through the exchange of a lighter scalar field σ, both of which are singlets of the Standard Model gauge group. The mediator field σ can have a nonzero mixing with the Higgs boson, and, therefore the σ boson can decay into photons and other Standard Model particles, even if its couplings to Standard Model particles are otherwise highly suppressed. This decay can ultimately produce a detectable signal.

We begin this section by introducing the minimal particle physics model that we will use. Following this, we discuss the dark matter halo within the Milky Way. We then discuss the relevant constraints on the parameters present in the model, although detailed discussion of the implementation of these constraints is contained in later sections where the signal from bremsstrahlung and bound state formation is explicitly calculated. Then finally we discuss the decay of the dark force mediator in flight in the dark matter halo.

We supplement the Standard Model with a complex scalar $\mathrm{SU_C}(3) \times \mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$ singlet S; we also introduce a global $\mathrm{U}_S(1)$ symmetry, analogous to baryon number, under which

$$S \to e^{i\alpha} S, \qquad S^\dagger \to e^{-i\alpha} S^\dagger. \tag{5.1}$$

Without a loss of generality we assume S particles carry unit $\mathrm{U_S}(1)$ charge, and because of conservation of this charge, the S particles are completely stable. We assume that dark matter today is composed of S particles, and the correct abundance is generated in some process similar to or combined with baryogenesis. We do not assume that dark matter is necessarily a thermal relic.

The symmetries of the theory allows the term $\lambda_S(S^\dagger S)^4$; which results in a velocity-independent scattering cross section for the S particles. It is difficult to have such a cross section large enough to resolve the problems with small halos without violating astronomical bounds [25, 28]. A velocity-dependent cross section, which arises when an interaction is mediated by a light particle, can more easily reproduce our observations [27,130]. Therefore, we also introduce an additional singlet scalar field σ, which is also neutral under the global $\mathrm{U_S}(1)$. The most general potential, after the Standard Model electroweak symmetry is spontaneously broken, is

$$\begin{aligned} V = {} & M^2 h^0\sigma + m_S^2 S^\dagger S + \frac{m_\sigma^2}{2}\sigma^2 + \frac{m_h^2}{2}(h^0)^2 + A_h(h^0)^3 + A_\sigma\sigma^3 + A_{\sigma S}S^\dagger S\sigma + A_{\sigma h}(h^0)^2\sigma \\ & + A_{h\sigma}\sigma^2 h^0 + A_{hS}S^\dagger S h^0 + \lambda_S(S^\dagger S)^2 + \frac{\lambda_\sigma}{4}\sigma^4 + \frac{\lambda_{hS}}{2}S^\dagger S(h^0)^2 + \frac{\lambda_h}{4}(h^0)^4 \\ & + \frac{\lambda_{\sigma S}}{2}\sigma^2 S^\dagger S + \frac{\lambda_{\sigma h}}{4}\sigma^2(h^0)^2, \end{aligned} \tag{5.2}$$

where h^0 is the Standard Model Higgs boson. The Higgs boson h^0 and σ carry identical quantum numbers and therefore mix. The mass eigenvalues are

$$m_{1,2}^2 = \frac{1}{2}\left(m_h^2 + m_\sigma^2 \pm \sqrt{(m_h^2 - m_\sigma^2)^2 + M^4}\right), \tag{5.3}$$

and the eigenstates are

$$\begin{aligned} \phi_1 &= \cos(\theta_M/2)h^0 + \sin(\theta_M/2)\sigma, \\ \phi_2 &= -\sin(\theta_M/2)h^0 + \cos(\theta_M/2)\sigma, \end{aligned} \tag{5.4}$$

where the mixing angle is

$$\tan(\theta_M) = \frac{M^2}{m_h^2 - m_\sigma^2}. \tag{5.5}$$

We will require that the mixing between the Higgs field and the σ field be small; this can be accomplished by setting the free parameter M appropriately. Then we may speak of the σ field and the Higgs field as approximate mass eigenstates with masses m_h and m_σ respectively; this allows the mass m_σ to be small even though no light scalar boson has been observed. We further assume that any interactions between the dark sector particles (S, $S^\dagger$, and σ) and Standard Model particles are highly suppressed.

We note that the relevant unitless coupling for the attractive Yukawa interaction between the S and σ particles is $\alpha = A_{\sigma S}^2/16\pi m_S^2$. This can be established in two ways. First, one may consider the non-relativistic limit of two particle scattering. We recall that quantum-field-theoretic wavefunctions include a normalization factor of $1/\sqrt{2m_S}$ for each S particle. Therefore, the relevant prefactor before the overlap integral for one particle exchange is $4\pi\alpha = A_{\sigma S}^2/4m_S^2$, which gives $\alpha = A_{\sigma S}^2/16\pi m_S^2$. Secondly, it is well known that the Bethe-Salpeter equation reproduces the ground state energy of the hydrogen atom. Setting the binding energy from the Bethe-Salpeter equation, $A_{\sigma S}^4/1024\pi^2 m_S^3$, equal to the hydrogen-like ground state energy for identical particles, $\alpha^2 m_S/4$, again gives $\alpha = A_{\sigma S}^2/16\pi m_S^2$. If one instead defines $\alpha = A_{\sigma S}^2/4\pi m_S^2$ (as in e.g. [131]), then additional factors of 4 must be introduced in other equations, e.g., the bound state mass.

Now that we have specified the particle physics of our dark matter model, let us consider the Milky Way dark matter halo. We assume that the correct abundance of S particles is determined by some process that is similar to baryogenesis or related to baryogenesis, as in the models reviewed in Ref. [116]. The absence of antiparticles in today's universe eliminates the possibility of a signal from $SS^\dagger$ annihilation. We use the Navarro-Frenk-White profile [132] to approximate the spatial mass distribution of dark matter,

$$\rho(r) = \frac{\rho_0}{(r/R_s)(1 + r/R_s)^2}. \tag{5.6}$$

We do not expect this profile to be accurate near the center of the galaxy; indeed, one of the motivations of self-interacting dark matter is to remove the cusp at $r = 0$ in the NFW profile. Therefore, we will cut off our integrals at scales of 1 kpc, and we emphasize that our results are not dependent on the sharp cusp present in the NFW profile. The parameters ρ_0 and R_s are related to the virial mass, virial radius, and concentration by

$$\begin{aligned} R_s &= \frac{r_{\rm vir}}{C}, \\ \rho_0 &= \frac{M_{\rm vir}}{\ln(1+C) - C/(1+C)} \frac{1}{4\pi R_s^3}. \end{aligned} \tag{5.7}$$

For the Milky Way, we use the parameters $M_{\rm vir} = 1.0 \cdot 10^{12} M_\odot$, $r_{\rm vir} = 258$ kpc, and $C = 12$ [133]. This gives $R_s = 3.4 \cdot 10^{36}$ GeV^{-1} and $\rho_0 = 1.4 \cdot 10^{-42}$ GeV^4.

In this chapter, we will ignore the cosmological history of the dark sector. This is especially questionable in the strongly-interacting regime, in which one generally expects significant bound state formation in the early universe. In general, the cosmology of self-interacting asymmetric dark matter models may be quite non-trivial; this will be explored in the next chapter for a slightly different model. If desired, we can imagine the following scenario: in the early universe, dark matter consists of one or more heavy species, with interactions sufficient to annihilate to the necessary dark asymmetry. At a later time, once the recombination process has frozen out, these particles decay into the particles considered here. These may form bound states only when gathered into dense halos.

In calculating the cross sections for bremsstrahlung emission of σ particles and bound state formation, we will need to average over the relative velocities of the S particles. Therefore, we need the velocity distribution $P(v(r))$ as a function of the distance from the center of the galaxy. If the dark matter has virialized, then its average circular velocity should decrease near the galactic center, except for a small region near the super-massive black hole. However, the dark matter radial velocity profile and dispersion are currently unknown. Because of these uncertainties, we will instead use a Maxwellian distribution with the effective temperature $T_{\rm eff}$ chosen such that the average velocity is 220 km/s. We note that simulations support the assumption of a locally Gaussian velocity distribution even for cold dark matter [134], and the isothermal approximation is better for self-interacting dark matter [25, 130]. The velocity distribution for two non-relativistic S particles is

$$P(v_1, v_2)\, dv_1\, dv_2 = (4\pi)^2 \left(\frac{m_S}{2\pi T_{\rm eff}}\right)^3 e^{-m(v_1^2+v_2^2)/2T_{\rm eff}} v_1^2 v_2^2\, dv_1\, dv_2. \tag{5.8}$$

In terms of the total velocity $\boldsymbol{v}_T = \boldsymbol{v}_1 + \boldsymbol{v}_2$ and the relative velocity $\boldsymbol{v}_{\rm rel} = \boldsymbol{v}_1 - \boldsymbol{v}_2$, the distribution is

$$P(v_{\rm rel}, v_T)\, dv_{\rm rel}\, dv_T = \frac{(4\pi)^2}{8} \left(\frac{m_S}{2\pi T_{\rm eff}}\right)^3 e^{-m_S(v_{\rm rel}^2+v_T^2)/4T_{\rm eff}} v_{\rm rel}^2 v_T^2\, dv_{\rm rel}\, dv_T. \tag{5.9}$$

We integrate over the total velocity to find the relative velocity distribution (in a reference frame at rest with respect to the Milky Way),

$$P(v_{\rm rel})\,dv_{\rm rel} = \frac{4\pi}{\sqrt{8}}\left(\frac{m_S}{2\pi T_{\rm eff}}\right)^{3/2} e^{-m_S v_{\rm rel}^2/4T_{\rm eff}} v_{\rm rel}^2\,dv_{\rm rel}. \tag{5.10}$$

Because the S particles are moving non-relativistically, this distribution also applies in their center of momentum frame. We observe that this is peaked at slightly larger velocities than the velocity distribution of a single particle.

The potential Eq. (5.2) contained several parameters, which cannot be set arbitrarily if dark matter is to be comprised entirely of S particles as outlined above. There are numerous constraints on these parameters, from both astrophysics and particle physics. In this section, we will give only a general discussion of these constraints; the specifics of how we implement them will be discussed when particular values for the parameters are chosen, which will be done separately for bremsstrahlung emission (Sec. 5.4) and bound state formation (Sec. 5.3). Since we aim to produce an observable indirect detection signal, we must satisfy all experimental constraints except those from indirect detection experiments.

First, we demand that this model make only insignificant modifications to the branching ratio for the decays of the Higgs boson. We forbid the decay $h^0 \to SS^\dagger$ by requiring $m_S > m_{h^0}/2 \approx 63$ GeV, using the recent Higgs mass measurements [15, 16]. The decay $h^0 \to \sigma\sigma$ can be arbitrarily suppressed by taking $A_{\sigma H}$ to be sufficiently small; this parameter is not used elsewhere in our analysis. We also demand that the mixing angle θ_M be small enough that the apparent branching ratio for $h^0 \to \sigma$ is less than the branching ratio for the $h^0 \to \gamma\gamma$ decay.

There are many well-known bounds on the self-interaction cross section of dark matter. As explained in [129], these constraints more appropriately restrict σ_T, the momentum-transfer cross section. (For identical particles, the closely-related viscosity cross section should be used instead [135]. In the limit $m_S\bar{v}/m_\sigma \gg 1$, which will be valid for our parameters, these

differ by $\mathcal{O}(1)$ and so we will ignore this complication.) The bullet cluster bound requires $\sigma_{SS}/m_S \lesssim 0.7$ cm^2/g [136], and bounds from the evaporation of galactic halos favor $\sigma_{SS}/m_S \lesssim 0.1$ cm^2/g [28]. Naively, these bounds appear to be in conflict with the preferred range to eliminate cuspy profiles, 0.56 cm^2/g $\lesssim \sigma_{SS}/m_S \lesssim 5.6$ cm^2/g [118, 119].

However, because these bounds affect vastly different scales, they may be resolved by considering a velocity-dependent cross section, as naturally arises in the Yukawa exchange of a light boson [137], [129]. Furthermore, such a cross section may additionally solve the "too big to fail" problem [26–28]. (It was this desire for a velocity-dependent cross section motivated the introduction of the σ boson above.) For an attractive Yukawa interaction, as we have introduced above, the bounds are consistent for provided that the masses m_S and m_σ satisfy particular relations given in [27, 138]. The precise constraint is a function of $v_{\rm max} = \sqrt{2\alpha m_\sigma/\pi m_S}$, the velocity at which $v\sigma_T$ peaks at a transfer cross section equal to $\sigma_T^{\rm max} = 22.7/m_\sigma^2$.

Additional bounds on the self-interaction cross section arise from observations of halo ellipticity, as introduced in Ref. [139]. Self-interactions are desirable in resolving the above-mentioned problems with small halos, but too much self-interaction can isotropize large halos, which are observed to be elliptical. The specific bounds are model-independent; Yukawa couplings are discussed in Ref. [125]. This reference considered dark matter masses up to 4 TeV; however, we will consider larger masses, which will require us to extend these bounds. Additionally, Ref. [28] has observed that these bounds may be somewhat weaker due corrections from the triaxial distribution of dark matter outside of the core; however, as they note, more detailed simulations are required to firmly establish this conclusion. Therefore, we will parameterize this uncertainty by the coefficient F; the numerical simulations discussed in Ref. [28] could be interpreted as favoring $F \sim 0.1$ for the particular halo model considered.

To extend the analysis of Ref. [125], we note that a dark matter halo will be spherical, as opposed to elliptical, if collisions which change the particle velocities by factors of order 1

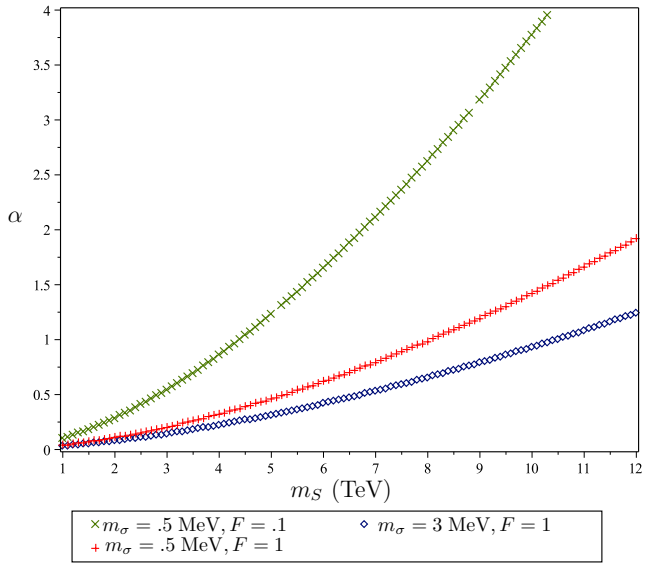

Figure 5.1: A plot of the critical coupling $\alpha = A_{\sigma S}^2/16\pi m_S^2$ as a function of m_S. Couplings below the critical coupling are consistent with the elliptical shape of large dark matter halos.

happen frequently enough. The rate at which these collisions occur is

$$\Gamma_k = \int d^3v_1\, d^3v_2\, f(v_1) f(v_2) (n_S v_{\rm rel} F \sigma_T)(v_{\rm rel}^2/v_0^2), \tag{5.11}$$

where σ_T is the momentum-transfer cross section, given by $\sigma_T = \int d\Omega\, (d\sigma/d\Omega)(1-\cos(\theta))$, and $f(v)$ is the dark matter velocity distribution. The analytic fit for σ_T, the distribution functions, and the relevant parameters for NGC 720 are all available in Ref. [125]. We note that quantum corrections to the cross section become important if the limit $m_S\bar{v}/m_\sigma \gg 1$ is violated; however, all of our parameters will be in this regime. If not, corrections such as those discussed in Ref. [135] should be included. We extrapolate the plot of the critical coupling $\alpha = A_{\sigma S}^2/16\pi m_S^2$ to $m_S = 12$ TeV in Fig. 5.1. We show the results for $m_\sigma = 0.5$ MeV and for $m_\sigma = 3$ MeV, considering both $F = 1$ and $F = 0.1$. As an example of the effect of this uncertainty, the critical coupling is decreased from 0.93 to 3.6 for $m_\sigma = 0.5$ MeV and $m_S = 10$ TeV.

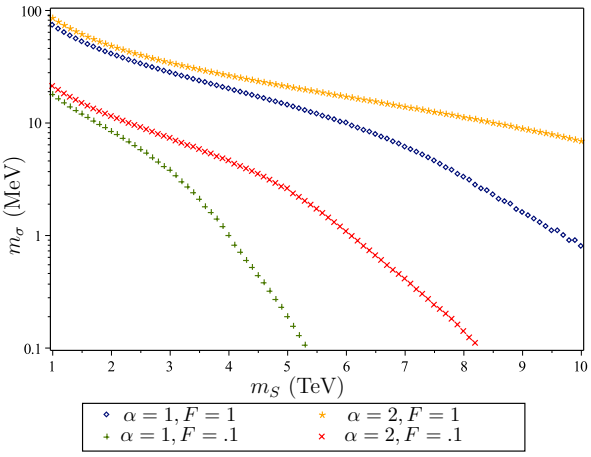

Figure 5.2: The minimum value of m_σ for which $\alpha = 2$ or $\alpha = 1$ is consistent with large elliptical halos.

We will see in Sec. 5.3 that a large coupling is desirable to produce a significant signal through bound state formation. This is generally a problem with halo ellipticity bounds; however, these become weaker as m_σ increases. Therefore, we also calculate the minimum m_σ for which $\alpha = 2$ or $\alpha = 1$ is consistent with the observed ellipticity of large halos, as a function of m_S. The results are shown in Fig. 5.2. Again we see that taking $F \sim 0.1$ dramatically weakens the bound.

Next we consider constraints from direct detection experiments. Experiments such as XENON100 [140] and CDMS [141] have set an upper bound on the cross section for the interaction between S particles and nucleons; because this interaction primarily occurs through Higgs boson exchanges, this constrains A_{hS}. The stability of neutron stars generally imposes stronger constraints on A_{hS} [142–145] but these constraints do not apply to scalar dark matter with masses at the TeV scale or above [146]. We will not use A_{hS} in our analysis; therefore, it can be set arbitrarily small. These constraints can also restrict the quartic couplings [147]; we may also set these arbitrarily small because they will not be used in our analysis.

We note that while we can arbitrarily suppress the S-nucleon interaction which occurs through the exchange of a Higgs boson, there is an additional diagram in which the S boson emits a σ boson, which transforms into a Higgs boson via the mixing, which then interacts with the nucleon. This diagram involves the coupling $A_{\sigma S}$ and the mixing angle θ_M, and we cannot take either of these parameters to zero without eliminating the signal we are trying to produce. (Although we will see that the mixing angle may be quite small while still giving a detectable signal.) Thus, in principle, one cannot arbitrarily decrease the S-nucleon cross section while maintaining a detectable indirect detection signal; there is a minimum value set by this diagram. We will now show that the contribution of this diagram is indeed quite small and is not in tension with direct detection bounds.

The oscillation time scale $\tau_{\text{osc}} = 2\pi E/\Delta m^2$ is generally many orders of magnitude smaller than the interaction time scale, which can be estimated by considering the overlap of the wavefunctions. Consequently, averaging over the "detector scale" (nucleon size), along with the source location, will simply give a factor of 1/2. (This is in contrast with neutrino oscillation experiments, for which τ_{osc} is large in comparison to other experimental scales, due to the small Δm^2. In this scenario, $\Delta m^2 \sim m_h^2$.) The S particles under consideration are generally much heavier than the protons; we will consider masses between 4 and 10 TeV. Therefore, in the center of momentum reference frame the S particles are approximately stationary, while the protons approach at speeds of approximately 220 km/s. The momentum transfer is approximately $2m_p v = 1.5$ MeV, which is far below the scale at which the nucleon form factors must be included.

In calculating the matrix element, the particle exchanged should be taken to be one of the mass eigenstates given in Eq. (5.4). The mass eigenstate will carry a 4-momentum on the scale of $p^\mu = (m_S \bar{v}^2/2, m_S \bar{\boldsymbol{v}})$, which is relatively small. Therefore, the diagram in which the exchanged eigenstate is mostly comprised of σ (that is, ϕ_2) will dominate. We will calculate this diagram in the laboratory reference frame, in which the denominator of the propagator

is $m_2^2 - m_S^2\bar{v}^4/4 + m_S^2\bar{v}^2 \approx m_S^2\bar{v}^2$ and the relevant matrix element is

$$-i\mathcal{M} \approx 3\bar{u}\frac{i}{m_S^2\bar{v}^2}A_{\sigma S}\cos\left(\frac{\theta_M}{2}\right)\frac{m_q}{v}\sin\left(\frac{\theta_M}{2}\right)u,$$

where v is the vacuum expectation value of the Higgs boson and u, $\bar{u}$ are spinors for the proton. This yields

$$|\mathcal{M}|^2 \approx \frac{9}{m_S^4\bar{v}^4}\frac{A_{\sigma S}^2 m_q^2 m_n^2}{v^2}\cos^2\left(\frac{\theta_M}{2}\right)\sin^2\left(\frac{\theta_M}{2}\right).$$

Because the velocities are non-relativistic, the initial energy squared is approximately $(m_S + m_n)^2 \approx m_S^2$, which gives an approximate cross section

$$\sigma \approx \frac{1}{16\pi m_S^2}\cdot\frac{9}{m_S^4\bar{v}^4}\frac{A_{\sigma S}^2 m_q^2 m_n^2}{v^2}\cos^2\left(\frac{\theta_M}{2}\right)\sin^2\left(\frac{\theta_M}{2}\right).$$

Let us consider one choice of parameters (which we will use below): $m_S = 4$ TeV and $A_{\sigma S} = 20$ TeV, corresponding to $\alpha = 2$. For the average effective mass of a quark, we use 3 MeV, and we choose $\theta_M = 10^{-3}$. This gives $\sigma \sim 10^{-18}$ GeV^{-2}. All of our other choices for parameters give a cross section below this value. This is well beneath the limits from direct detection experiments, which are 10^{-43} cm^2 (10^{-16} GeV^{-2}) [140, 141].

At present, there is significant uncertainty in the contribution of the s-quark to the effective quark mass, and since the Higgs coupling to s is much larger than the couplings to u and d, even a relatively small contribution of the sea quarks with higher masses can dominate the cross section. The measured s quark contribution, manifest as the nuclear pion-nucleon sigma term, is uncertain, and the resulting uncertainty in the cross section can be as large as an order of magnitude [148]. However, the cross section calculated above is at least two orders of magnitude beneath the lowest cross sections experimentally accessible. Therefore, direct detection experiments do not currently constrain the vast majority of the available parameter space.

Thus, in conclusion, the primary constraints that we must impose on the masses and coupling $A_{\sigma S}$ consist of the velocity-dependent constraints from astronomical observations

(e.g., the bullet cluster) and constraints from halo ellipticity. We will explicitly verify that we satisfy these when we consider specific parameter choices in Sec. 5.3 and Sec. 5.4.

Thus far, we have not addressed how we are going to produce an indirect detection signal. We will consider emission of σ particles from either bremsstrahlung or bound state formation. Although this particle's interactions with Standard Model particles are assumed to be small, it can decay through its (small) mixing with the Standard Model Higgs boson. It therefore has the same decay modes as the Higgs boson (provided they are kinematically allowed), although the amplitudes are suppressed by the σ-Higgs mixing parameters. Since we have seen that astrophysical constraints require a small mass m_σ, we consider the decays $\sigma \to \gamma\gamma$ and $\sigma \to e^+e^-$. For $m_\sigma \sim$ a few MeV, the dominant decay is $\sigma \to e^+e^-$. The decay rate in the rest frame of the σ boson is

$$\Gamma_{e^+e^-} = \frac{g_W^2 m_e^2 m_\sigma \sin^2(\theta_M/2)}{32\pi m_W^2}\left(1 - \frac{4m_e^2}{m_\sigma^2}\right)^{3/2}, \tag{5.12}$$

where g_W is the weak coupling constant. If $m_\sigma < 2m_e$, the decay $\sigma \to e^+e^-$ is kinematically forbidden, and instead the dominant decay is $\sigma \to \gamma\gamma$. In the σ particle's rest frame, the decay rate is

$$\Gamma_{\gamma\gamma} = \frac{\sin^2\left(\dfrac{\theta_M}{2}\right)\alpha_{\rm em}^2 g_W^2}{1024\pi^3}\frac{m_\sigma^3}{m_W^2}\left|\sum_i N_{ci} e_i^2 F_i\right|^2. \tag{5.13}$$

In this equation, $\alpha_{\rm em} \approx 1/137$, N_{ci} is the number of color states of the particle in the loop, and e_i is this particle's electric charge. The dominant contributions to the loop will be from electrons, up quarks, and down quarks, for which

$$F = -\tau\left(1 + (1-\tau)f(\tau)\right), \tag{5.14}$$

where $\tau = 4m_i^2/m_h^2$, and

$$f(\tau) = \begin{cases} \left(\arcsin\left(\sqrt{1/\tau}\right)\right)^2 & \tau \geq 1, \\ -\left(\ln(\eta_+/\eta_-) - i\pi\right)^2 & \tau < 1, \end{cases} \tag{5.15}$$

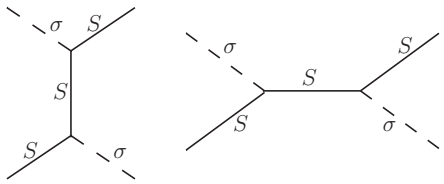

Figure 5.3: Diagrams for the scattering of σ particles with dark matter.

and $\eta_\pm = 1 \pm \sqrt{1-\tau}$. In a reference frame in which the σ particle is moving with speed v, the lifetime is $\tau = \gamma/\Gamma$ due to time dilation. For the decay products to be observed, the σ particles must decay in flight before they travel the ~ 8 kpc separating Earth from the galactic center. For our choice of parameters, the mean distance traveled in one lifetime is significantly shorter than this distance.

However, constant scattering can act like a quantum Zeno experiment that prevents the decay. Therefore, we must also show that the mean free path for σ particles in the Milky Way is greater than the distance they travel in one lifetime (measured in the Milky Way reference frame). Since the quartic coupling $\lambda_{\sigma S}$ and the cubic coupling A_σ may be made arbitrarily small, we assume that the scattering is dominated by the $S\sigma$ interaction mediated by an S boson; there are two diagrams that contribute, which are shown in Fig. 5.3.

In the Milky Way reference frame, the σ particle is moving relativistically with energy E_σ, while the S particle is moving nonrelativistically with velocity v of order 10^{-3}. We do not assume any relation between E_σ and the kinetic energy of the S particle. Since the cross section is a relativistic invariant, we may evaluate it in the center of momentum frame, which is attained by boosting by $\beta = E_\sigma/(E_\sigma + m_S)$. Keeping only the largest terms, we find that

the initial and final four-momenta in the CM frame are

$$\begin{aligned}
p^{\mu}_{\sigma,i} &= (\gamma\beta m_S, 0, 0, \gamma\beta m_S),\\
p^{\mu}_{S,i} &= (\gamma m_S, 0, 0, -\gamma\beta m_S),\\
p^{\mu}_{\sigma,f} &= \gamma\beta m_S, \gamma\beta m_S \sin(\theta), 0, \gamma\beta m_S \cos(\theta)),\\
p^{\mu}_{S,f} &= (\gamma m_S, -\gamma\beta m_S \sin(\theta), 0, -\gamma\beta m_S \cos(\theta),
\end{aligned} \tag{5.16}$$

where we have used the fact that the collision is elastic. The matrix element is

$$\begin{aligned}
-i\mathcal{M} &= -\frac{A^2_{\sigma S}}{m_S^2 - (p_{S,i} - p_{\sigma,f})^2} - \frac{A^2_{\sigma S}}{m_S^2 - (p_{S,i} + p_{\sigma,i})^2},\\
&= -\frac{A^2_{\sigma S}}{m_S^2 - m_S^2 + 2p_{S,i}\cdot p_{\sigma,f}} - \frac{A^2_{\sigma S}}{m_S^2 - m_S^2 - 2p_{S,i}\cdot p_{\sigma,i}},\\
&= -\frac{A^2_{\sigma S}}{2\gamma^2\beta m_S^2 + 2\gamma^2\beta^2 m_S^2\cos(\theta)} - \frac{A^2_{\sigma S}}{-2\gamma^2\beta m_S^2 - 2\gamma^2\beta^2 m_S^2},\\
&= -\frac{A^2_{\sigma S}}{2\gamma^2 m_S^2}\left(\frac{1+\cos(\theta)}{(1+\beta)(1+\beta\cos(\theta))}\right),
\end{aligned} \tag{5.17}$$

which yields the cross section

$$\begin{aligned}
\sigma &= \frac{1}{64\pi^2}\int \frac{|\mathcal{M}|^2}{\gamma^2(1+\beta)^2 m_S^2} d\Omega,\\
&= \frac{A^4_{\sigma S}}{128\pi\gamma^6(1+\beta)^4 m_S^6}\int_0^{\pi}\left(\frac{1+\cos(\theta)}{1+\beta\cos(\theta)}\right)^2 \sin(\theta)\, d\theta,\\
&= \frac{A^4_{\sigma S}}{64\pi m_S^6}\frac{2\beta + (1-\beta^2)\ln((1-\beta)/(1+\beta))}{\gamma^6\beta^3(1+\beta)^5}.
\end{aligned} \tag{5.18}$$

The mean free path is $\ell = (\sigma n_S)^{-1}$, where n_S can be found using Eq. (5.6). Since n_S depends on r, the mean free path will also depend on r; it is the smallest as we approach the galactic center. Although we will avoid small radii due to the central cusp, let us evaluate the mean free path at 1 pc as an extreme example. For $m_S = 5$ TeV, $A_{\sigma S} = 3$ TeV, and $E_\sigma = 1$ TeV, the mean free path of order 10^{36} m; if we decrease E_σ to 1 MeV, it increases only to $\prime(10^{35}$ m). These are much greater than the 10^{20} m between the galactic center and the Solar System; therefore, scattering will not decohere the oscillations.

We are now certain that the σ bosons produced in the Milky Way will decay to $\gamma\gamma$ or e^+e^-, potentially providing an indirect detection signal. To determine the signal, we also need to know the width of the energy distribution of the decay products. The decay channels considered are two body decays; therefore, the energy spectrum of the decay products is a sharp line at $m_\sigma/2$ in the rest frame of the m_σ particle. This energy spectrum must be boosted into the Milky Way reference frame, in which the σ particles are moving with speed $v = \sqrt{E_\sigma^2 - m_\sigma^2}$. Because the σ boson is spinless, the resulting energy distribution is flat. Consequently, for decays to an electron and positron, the energy distribution is

$$P(E_e) = \frac{1}{\sqrt{(E_\sigma^2 - m_\sigma^2)(1 - 4m_e^2/m_\sigma^2)}} \tag{5.19}$$

for E_e between the values of

$$E_{\mathrm{e,max}}, E_{\mathrm{e,min}} = \frac{E_\sigma}{2} \pm \frac{\sqrt{(E_\sigma^2 - m_\sigma^2)(1 - 4m_e^2/m_\sigma^2)}}{2}. \tag{5.20}$$

Similarly, for decays to two photons, the energy distribution is

$$P(E_\gamma) = \frac{1}{\sqrt{E_\sigma^2 - m_\sigma^2}} \tag{5.21}$$

for E_e between the values of

$$E_{\gamma,\mathrm{max}}, E_{\gamma,\mathrm{min}} = \frac{E_\sigma}{2} \pm \frac{\sqrt{E_\sigma^2 - m_\sigma^2}}{2}. \tag{5.22}$$

In this section, we have outlined a minimal scalar self-interacting asymmetric dark matter model. We have also discussed the constraints we must impose on the parameter space, and the decays of the dark force mediator that may produce a detectable indirect detection signal. However, we have not considered the production of the σ bosons, nor have we considered the spectrum of the signal that would be observed at the Solar System. We will next complete the analysis for the emission of σ bosons in bound state formation; following this, we consider bremsstrahlung emission in Sec. 5.4.

5.3 Signal From Dark Matter Bound State Formation

It has previously been noted that many models of self-interacting dark matter, including supersymmetric models, permit the existence of dark matter bound states [131]. The same reference notes that the decay of emitted force carrier particles could, in theory, produce a signal for indirect detection experiments. In this section, we will by explicitly calculate this signal for the above asymmetric scalar dark matter model. We will show that it is indeed possible to produce a signal above current observational bounds, establishing the possibility of indirect detection of self-interacting asymmetric dark matter. However, we will further show that the limit $\alpha \ll 1$, as taken in Ref. [131], does not produce a detectable signal.

For illustrative purposes, we choose specific points in parameter space, taking care to satisfy the constraints discussed in Sec. 6.1. To facilitate the formation of bound states, we desire a large coupling $A_{\sigma S}$; we will choose $\alpha = A_{\sigma S}^2/16\pi m_S^2 = 2$. Although calculations in strongly-interacting regimes are notoriously difficult, the astrophysical bounds discussed in Refs. [125, 138] may be extrapolated to large α; the transfer cross section used in both references includes corrections for strong interactions. We will also consider $\alpha = 1$ and show that this is not sufficient to produce an observable signal.[1]

First, we ensure that our parameters are consistent with large elliptical halos; as discussed in Ref. [125], the restriction on α becomes weaker as m_σ is increased. Fig. 5.2 shows the minimum m_σ as a function of m_S for which we may consistently take $\alpha = 2$. The choice $m_S = 4$ TeV requires $m_\sigma \gtrsim 30$ MeV; we will choose $m_\sigma = 40$ MeV. Similarly, for $m_S = 4$ TeV and $\alpha = 1$, we must satisfy $m_\sigma \gtrsim 20$ MeV; we choose $m_\sigma = 25$ MeV. We must also ensure that our parameters are consistent with the velocity-dependent bounds based on the bullet cluster and other astrophysical observations. The velocity for which $\langle v\sigma_T \rangle = \sigma_T^{\max}$ is $v_{\max} = \sqrt{2\alpha m_\sigma/\pi m_S}$; this is 1100 km/s and 2000 km/s respectively for the two sets

[1] Readers familiar with atomic dark matter models may be concerned because for a gauge vector mediator, the ground state has nonzero angular momentum at such large coupling constants. This does not hold for the exchange of scalars considered here.

of parameters above. To be consistent with astrophysical data, we then must have $\sigma_T^{\max}/m_S \leq 100\ \mathrm{GeV}^{-3}$ [138]; the above numbers correspond to $0.2\ \mathrm{GeV}^{-3}$ and $9\ \mathrm{GeV}^{-3}$. Finally, bound state formation is hindered if a real σ boson cannot be emitted. Therefore, we must also take $m_\sigma \ll B$, where the binding energy $B = \alpha^2 m_S/4$. For the first set of parameters chosen, the binding energy is 4 TeV, and for the second set of parameters, $B = 1$ TeV. For both, $m_\sigma \ll B$. Hence, these two points in parameter space satisfy all the requisite constraints.

The rate of formation of bound states, neglecting charge depletion, is given by

$$\frac{dN_{\mathrm{BS}}}{dt} = \int n_S(r)^2 \sigma_{\mathrm{BS}} v_{\mathrm{rel}}\, dV, \tag{5.23}$$

where $n_S(r) = \rho(r)/m_S$ is the number density of S particles and σ_{BS} is the cross section for bound state formation. Because the S particles do not escape to infinity, this cross section cannot be approximated using the Born approximation. However, we are in the classical regime ($m_S \bar{v}/m_\sigma \gg 1$) even though the coupling is strong; therefore we do not need to include additional quantum corrections such as those calculated numerically in Ref. [135]. We now proceed to calculate this cross section in the classical limit.

We will approximate the σ boson as massless. The cross section for non-relativistic electrons and positrons to form a bound state through photon exchange was calculated in Ref. [149]; we adapt this derivation for scalar fields. The matrix element is

$$\mathcal{M} = -i \int \Psi_f^*(\boldsymbol{r}_1, \boldsymbol{r}_2) \left(\sum_{n=1,2} A_n e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_n} \right) \Psi_i(\boldsymbol{r}_1, \boldsymbol{r}_2) d^3 r_1\, d^3 r_2 (2\pi) \delta(E_i - E_f - E_\sigma). \tag{5.24}$$

In this equation, $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are the locations of the two S particles respectively, Ψ_f is the wavefunction of the bound state, and Ψ_i is the wavefunction for the two incoming S particles. The factor $e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_n}$ represents the wavefunction of the σ particle, and the sum is over the two S particles it can couple to. In this equation, the wavefunctions have the standard normalization in quantum field theory; however, since we are interested in the

non-relativistic limit, let us use wavefunctions that are normalized to one. Then the matrix element is

$$\mathcal{M} = -i\frac{A_{\sigma S}}{2m_S}\int \Psi_f^*(\boldsymbol{r}_1,\boldsymbol{r}_2)\left(e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_1}+e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_2}\right)\Psi_i(\boldsymbol{r}_1,\boldsymbol{r}_2)d^3r_1\,d^3r_2(2\pi)\delta(E_i-E_f-E_\sigma). \tag{5.25}$$

Next we define

$$\boldsymbol{R} = \frac{\boldsymbol{r}_1+\boldsymbol{r}_2}{2}, \qquad \boldsymbol{r} = \boldsymbol{r}_1-\boldsymbol{r}_2, \tag{5.26}$$

and write the wavefunctions as

$$\Psi_i(\boldsymbol{r}_2,\boldsymbol{r}_1) = e^{i\boldsymbol{Q}\cdot\boldsymbol{R}}\Psi_i(\boldsymbol{r}), \qquad \Psi_f(\boldsymbol{r}_2,\boldsymbol{r}_1) = e^{i\boldsymbol{P}\cdot\boldsymbol{R}}\Psi_f(\boldsymbol{r}), \tag{5.27}$$

where $\boldsymbol{Q} = \boldsymbol{p}_1+\boldsymbol{p}_2$ is the total momentum of the initial particles. Similarly, $\boldsymbol{P}$ is the momentum of the bound state. Performing the d^3R integral yields

$$\mathcal{M} = -i\frac{A_{\sigma S}}{2m_S}\int \Psi_f^*(\boldsymbol{r})\left(e^{i\boldsymbol{k}\cdot\boldsymbol{r}/2}+e^{-i\boldsymbol{k}\cdot\boldsymbol{r}/2}\right)\Psi_i(\boldsymbol{r})d^3r(2\pi)^4\delta(E_i-E_f-E_\sigma)\delta^3(\boldsymbol{Q}-\boldsymbol{k}-\boldsymbol{P}). \tag{5.28}$$

The reduced matrix element is

$$\bar{\mathcal{M}} = \int \Psi_f^*(\boldsymbol{r})\left(e^{i\boldsymbol{k}\cdot\boldsymbol{r}/2}+e^{-i\boldsymbol{k}\cdot\boldsymbol{r}/2}\right)\Psi_i(\boldsymbol{r})d^3r, \tag{5.29}$$

and the differential probability is

$$dW = \frac{TV}{(2\pi)^2 2E_\sigma}\frac{A_{\sigma S}^2}{4m_S^2}\delta(E_i-E_f-E_\sigma)\delta^3(\boldsymbol{Q}-\boldsymbol{k}-\boldsymbol{P})|\bar{\mathcal{M}}|^2|\boldsymbol{k}|^2\,d|\boldsymbol{k}|\,d\Omega\,d^3P, \tag{5.30}$$

where V is the normalized volume, T is the interaction time, and $d\Omega$ is the solid angle for the σ particle. The remaining integrals enforce momentum and energy conservation; we may perform them by directly imposing these constraints in our calculation. The transition probability per unit volume and unit time is

$$dw = \frac{A_{\sigma S}^2}{4m_S^2}\frac{|\boldsymbol{k}|^2 d\Omega}{2E_\sigma(2\pi)^2}|\bar{\mathcal{M}}|^2. \tag{5.31}$$

If $m_\sigma \ll B$, then $E_\sigma \approx |\boldsymbol{k}|$ and this simplifies to

$$dw = \frac{A_{\sigma S}^2}{4m_S^2}\frac{|\boldsymbol{k}| d\Omega}{2(2\pi)^2}|\bar{\mathcal{M}}|^2. \tag{5.32}$$

The differential cross section is $d\sigma = dw/v_{\rm rel}$ where $v_{\rm rel}$ is the relative velocity of the particles in the initial state. We define the relative momentum by $\boldsymbol{p} = \mu \boldsymbol{v}_{\rm rel}$ where $\mu = m_S/2$ is the reduced mass. $|\boldsymbol{p}|$ is also the momentum of one of the incoming particles in the center of momentum frame; we now specialize to this frame. (We note that the cross section is Lorentz invariant, and therefore still applicable to other reference frames.) Then

$$d\sigma = \frac{A_{\sigma S}^2}{4m_S} \frac{|\boldsymbol{k}| d\Omega}{|\boldsymbol{p}|(2\pi)^2} |\bar{\mathcal{M}}|^2. \tag{5.33}$$

The free S particles do not escape to infinity; they exist only in the initial state. Therefore, at large $\boldsymbol{r}$, $\Psi_i(\boldsymbol{r})$ must be a superposition of a plane wave and an outgoing spherical wave, which we approximate as a Coulomb wave. This approximation should be reasonable in the classical regime. The appropriate wavefunction with these limits is [150] (also discussed in Ref. [149])

$$\Psi_i(\boldsymbol{r}) = e^{\pi\zeta/2}\Gamma(1 - i\zeta)F(i\zeta, 1, i(pr - \boldsymbol{p}\cdot\boldsymbol{r}))e^{i\boldsymbol{p}\cdot\boldsymbol{r}}, \tag{5.34}$$

where $\zeta = A_{\sigma S} m_S/4|\boldsymbol{p}|m_S = A_{\sigma S}/4|\boldsymbol{p}|$, and F is the confluent hypergeometrical function. This has the same normalization as a plane wave. We note that the cross section will be very sensitive to the ratio $A_{\sigma S}/|\boldsymbol{p}| \sim A_{\sigma S}/m_S$ due to the first exponential.

For the final state $\Psi_f(\boldsymbol{r})$, we adapt the hydrogen ground state wavefunction; this is reasonable in the approximation that m_σ is negligible. For the actual hydrogen atom, in which the interaction is mediated by a vector gauge boson, the ground state carries non-zero angular momentum for $\alpha \geq 1/2$. However, the derivation of this result depends on the fact that the mediator is a vector gauge boson, while our mediator is a scalar boson.[2] Thus we take

$$\Psi_f = \sqrt{\frac{\eta^3}{\pi}} e^{-r\eta}, \tag{5.35}$$

[2]In QED, the gauge derivatives in the field equation give rise to $(E - e^2/r)^2$, which gives a term $\sim e^4/r^2$. This combines with $\ell(\ell+1)/r^2$ and leads to the $\alpha > 1/2$ requirement. In the scalar exchange considered here, there are no gauge derivatives and we simply insert the potential $\sim \alpha \exp(-mr)/r$ into the field equation. No term like α^2/r^2 appears.

where $\eta = \zeta|\boldsymbol{p}| = A_{\sigma S}/4$; this is the radius of the bound state. The reduced matrix element is

$$\bar{\mathcal{M}} = \sqrt{\frac{\eta^3}{\pi}} e^{\pi\zeta/2}\Gamma(1-i\zeta) \int e^{i\boldsymbol{p}\cdot\boldsymbol{r}-r\eta} \left(e^{i\boldsymbol{k}\cdot\boldsymbol{r}/2} + e^{-i\boldsymbol{k}\cdot\boldsymbol{r}/2}\right) \cdot F(i\zeta, 1, i(pr - \boldsymbol{p}\cdot\boldsymbol{r}))\, d^3r. \quad (5.36)$$

To evaluate the integral, we differentiate the identity [151]

$$\int e^{i(\boldsymbol{p}-\boldsymbol{\kappa})\cdot\boldsymbol{r}-\eta r} F(i\zeta, 1, i(pr - \boldsymbol{p}\cdot\boldsymbol{r})) \frac{d^3r}{r} = 4\pi \frac{[|\boldsymbol{\kappa}|^2 + (\eta - i|\boldsymbol{p}|)^2]^{-i\zeta}}{[(\boldsymbol{p}-\boldsymbol{\kappa})^2 + \eta^2]^{1-i\zeta}} \quad (5.37)$$

with respect to η. The result is

$$\begin{aligned}
&\int e^{i(\boldsymbol{p}-\boldsymbol{\kappa})\cdot\boldsymbol{r}-\eta r} F(i\zeta, 1, i(pr - \boldsymbol{p}\cdot\boldsymbol{r}))\, d^3r \\
&\quad = 8\pi \frac{[|\boldsymbol{\kappa}|^2 + (\eta - i|\boldsymbol{p}|)^2]^{-i\zeta}}{[(\boldsymbol{p}-\boldsymbol{\kappa})^2 + \eta^2]^{2-i\zeta}} \left[\zeta \frac{(\eta - i|\boldsymbol{p}|)[(\boldsymbol{p}-\boldsymbol{\kappa})^2 + \eta^2]}{[|\boldsymbol{\kappa}|^2 + (\eta - i|\boldsymbol{p}|)^2]} - i\eta(1-i\zeta)\right] \\
&\quad \equiv g(\kappa, \chi), \qquad (5.38)
\end{aligned}$$

where χ is the angle between $\boldsymbol{p}$ and $\boldsymbol{\kappa}$. We observe that $g(\kappa, \pi - \chi) = g(-\kappa, \chi)$. If the angle between $\boldsymbol{k}$ and $\boldsymbol{p}$ is Υ, the reduced matrix element is

$$\bar{\mathcal{M}} = \sqrt{\frac{\eta^3}{\pi}} e^{\pi\zeta/2}\Gamma(1-i\zeta) \left(g\left(\frac{|\boldsymbol{k}|}{2}, \Upsilon\right) + g\left(-\frac{|\boldsymbol{k}|}{2}, \Upsilon\right)\right). \quad (5.39)$$

This can be evaluated numerically. The last remaining unknown quantity in (5.33) is $|\boldsymbol{k}|$, which can be found from the energy conservation equation

$$2m + \frac{|\boldsymbol{p}|^2}{m} = (2m - B) + \frac{|\boldsymbol{k}|^2}{2(2m-B)} + |\boldsymbol{k}|, \quad (5.40)$$

where we have noted that in the center of momentum reference frame, the bound state also has momentum $|\boldsymbol{k}|$. We find the total cross section by numerically integrating (5.33). We must also average over a relative momentum distribution $P(|\boldsymbol{p}|)$; the total averaged cross section is given by

$$\sigma_{BS} = \iint \frac{A_{\sigma S}^2}{4m_S} \frac{|\boldsymbol{k}||\bar{\mathcal{M}}|^2}{|\boldsymbol{p}|(2\pi)^2} P(|\boldsymbol{p}|)\, d|\boldsymbol{p}|\, 2\pi \sin(\Upsilon)\, d\Upsilon. \quad (5.41)$$

We note that in the non-relativistic limit the momentum difference of the two particles is independent of reference frame; therefore we can calculate $P(|\boldsymbol{p}|)$ in any convenient frame

even though we specialized to the center of momentum reference frame above. The total cross section is, of course, Lorentz invariant. Using Eq. (5.10), we find the relative momentum distribution

$$P(|\boldsymbol{p}|)\, d|\boldsymbol{p}| = \frac{4\pi\sqrt{8}}{m_S^3}\left(\frac{m_S}{2\pi T_{\text{eff}}}\right)^{3/2} e^{-|\boldsymbol{p}|^2/m_S T_{\text{eff}}} |\boldsymbol{p}|^2\, d|\boldsymbol{p}|. \tag{5.42}$$

This gives us the desired cross section for bound state formation.

In fact, we are interested in the rate of bound state formation, which we recall is given by Eq. (5.23). We should average the whole equation over the relative momentum (instead of simply substituting the average cross section), using

$$\frac{dN_{\text{BS}}}{dt} = \int n_S(r)^2 dV \iint \frac{2|\boldsymbol{p}_{\text{rel}}|}{m_S}\sigma(|\boldsymbol{p}_{\text{rel}}|)P(|\boldsymbol{p}|)\, d|\boldsymbol{p}_{\text{rel}}|. \tag{5.43}$$

The first set of parameters discussed above corresponds to a cross section of $4.11{\cdot}10^{-2}$ GeV^2, which gives the rate $dN_{\text{BS}}/dt = 2.1{\cdot}10^{14}$ GeV. In one year, $9.8{\cdot}10^{45}$ bound states are formed, which means that during the lifetime of the Milky Way, $1.3 \cdot 10^{56}$ would have formed. This is indeed negligible in comparison to the total number of S particles between 1 kpc and 8 kpc, which is $7.2 \cdot 10^{63}$. This justifies our neglect of charge depletion. If we decrease α to 1, then the cross section drops by two orders of magnitude, to $5.76 \cdot 10^{-4}$ GeV^2. The rate is also two orders of magnitude smaller, $dN_{\text{BS}}/dt = 2.9 \cdot 10^{12}$ GeV. Again, we may neglect charge depletion.

The S particles do not interact electromagnetically. Therefore, when a bound state is formed, the excess energy is carried off by emission of a σ particle. Although the binding energy is large enough that a Higgs boson could be emitted instead, the σ particle dominates because it is lighter and has a stronger coupling to the S particles. Because of the nonzero σ-Higgs mixing, this σ particle may decay into Standard Model particles. Given our choices for m_σ, the dominant decay is $\sigma \to e^+e^-$ for both $\alpha = 2$ and $\alpha = 1$.

In the rest frame of the Milky Way, the σ particles will have a typical energy equal to the binding energy; the additional energy the σ particle may carry from the kinetic energy of the nonrelativistic S particles is negligible. Using Eq. (5.13), we find that the lifetime of the σ boson, in the Milky Way's rest frame, is $1.46 \cdot 10^{25}$ GeV, or 9.61 s, for the first set of parameters. The distance that they travel before decaying is 10^9 m, which is significantly less than the distance from the galactic center to the Solar System. For the second set of parameters, the lifetime is $9.36 \cdot 10^{24}$ GeV ≈ 6.17 s, and they similarly decay in flight.

As explained in Sec. 5.2, the resulting electrons and positrons have a flat energy distribution; their spectrum is

$$\frac{dN_e}{dE_e\, dt} = \frac{1}{\sqrt{(B^2 - m_\sigma^2)(1 - 4m_e^2/m_\sigma^2)}} \frac{dN_{\rm BS}}{dt}. \tag{5.44}$$

These particles have typical energies on or just below the TeV scale; they lose energy through synchrotron radiation and inverse Compton scattering rapidly, within about 1 kpc [152]. Therefore, few of these high-energy electrons and positrons will be observed near Earth.

There are three sources of background photons: the cosmic microwave background (CMB) radiation, starlight, and the starlight reprocessed by dust (including the extra-galactic background light, which is the starlight re-emitted by dust outside Milky Way). Outside the central molecular zone, the cosmic microwave background radiation dominates the photon number density [153]. For the signal from distances between 1 and 8 kpc from the galactic center, one may safely neglect scattering from photons other than CMB photons.

Because of the smallness of the IC mean free path, other propagation effects are not significant; we will show explicitly that we can neglect the energy loss due to synchrotron radiation, which is described by

$$\frac{dE_e}{dt} = -b_{\rm sync} E_e^2, \tag{5.45}$$

where the unitless coefficient $b_{\rm sync}$ is given by

$$b_{\rm sync} = \frac{4\sigma_T}{3m_e} \frac{B^2}{8\pi}. \tag{5.46}$$

σ_T is the Thomson cross section. Since we consider a spherical region extending from 1 kpc to 8 kpc, very few of the fermions will be created in the galactic plane. Therefore, the appropriate magnetic field is 1 μG [154, 155], which gives $b_{\text{sync}} = 6 \cdot 10^{-43}$.

This is to be compared to the energy loss of a single fermion due to inverse Compton scattering, which is described by the equation

$$\frac{dE_e}{dt} = -b_{\text{ICS}} E_e^2, \tag{5.47}$$

where now the unitless coefficient is

$$b_{\text{ICS}} = \frac{4\sigma_{\text{KN}} w_{ph}}{3m_e^2}. \tag{5.48}$$

σ_{KN} is the Klein-Nishina cross section, which reduces to the Thomson cross section when relativistic corrections are negligible. For scattering with scattering CMB photons, $b_{\text{ICS}} = 5.2 \cdot 10^{-41}$ and is approximately independent of energy, for the set of parameters with $\alpha = 2$. For the parameters with $\alpha = 1$, we have $b_{\text{ICS}} = 5.3 \cdot 10^{-41}$ instead. Since this is two orders of magnitude larger than the corresponding value for synchrotron radiation, we may neglect energy loss due to synchrotron radiation.

Now we proceed to calculate the photon energy spectrum from inverse Compton scattering with CMB photons. The cosmic microwave background radiation is a blackbody at $T_{\text{CMB}} = 2.73$ K; therefore the photon density per unit energy is

$$n_{\text{ph}}(\epsilon) \equiv \frac{d^2 N_{\text{ph,CMB}}}{dV\, d\epsilon} = \frac{1}{\pi^2} \frac{\epsilon^2}{\exp(\epsilon / T_{\text{CMB}}) - 1}, \tag{5.49}$$

where ϵ is the energy of the unscattered photon. For inverse Compton scattering, the number of scattered photons per unit energy per unit time produced by an electron or positron with Lorentz factor γ is given by [156, 157]

$$\frac{d^2 N_\gamma}{dE\, dt_s}(E, \gamma) = \int_0^\infty d\epsilon\, n_{\text{ph}}(\epsilon) \sigma_{\text{KN}}(E, \epsilon, \gamma), \tag{5.50}$$

where $\sigma_{\rm KN}(E,\epsilon,\gamma)$ is

$$\sigma_{\rm KN}(E,\epsilon,\gamma) = \frac{3\sigma_T}{4\epsilon\gamma^2} G(q,\Gamma) \tag{5.51}$$

and

$$G(q,\Gamma) = 2q\ln(q) + (1+2q)(1-q) + 2\eta q(1-q),$$
$$\Gamma = \frac{4\epsilon}{m_e}, \qquad \eta = \frac{\epsilon E}{m_e^2}, \qquad q = \frac{E}{\Gamma(m_e - E)}.$$

We have put a subscript on t_S in Eq. (5.50) to denote that this variable measures the time during which the fermion scatters against CMB photons. We use the symbol E for the final energy of the scattered photon. The Thomson limit corresponds to $\Gamma \ll 1$ which is applicable here. By energy conservation, only energies E between the following values are allowed:

$$E_{\rm min}(\gamma,\epsilon) = \frac{\gamma m_e \Gamma}{4\gamma^2 + \Gamma}, \qquad E_{\rm max}(\gamma,\epsilon) = \frac{\gamma m_e \Gamma}{1+\Gamma}, \tag{5.52}$$

which we enforce by writing

$$\frac{d^2N_\gamma}{dE\, dt_S}(E,\gamma) = \int_0^\infty d\epsilon\; n_{\rm ph}(\epsilon)\sigma_{\rm KN}(E,\epsilon,\gamma)\Theta(E_{\rm max}(\gamma,\epsilon) - E)\Theta(E - E_{\rm min}(\gamma,\epsilon)). \tag{5.53}$$

This equation gives the number of photons per unit energy per unit time scattered by an electron or positron with energy γm_e. From the fermion energy distribution given in Eq. (5.19), the corresponding distribution is

$$P(\gamma) = \frac{m_e}{\sqrt{(B^2 - m_\sigma^2)(1 - 4m_e^2/m_\sigma^2)}} \tag{5.54}$$

for γ between the values

$$\gamma_{\rm max}, \gamma_{\rm min} = \pm\frac{\sqrt{(B^2 - m_\sigma^2)(1 - 4m_e^2/m_\sigma^2)}}{2m_e}. \tag{5.55}$$

Averaging $d^2N_\gamma/dE\, dt$ over the γ distribution gives

$$\frac{d^2N_\gamma}{dE\, dt_S}(E) = \int_{\gamma_{\rm min}}^{\gamma_{\rm max}} P(\gamma) \int_0^\infty d\epsilon\; n_{\rm ph}(\epsilon)\sigma_{\rm KN}(E,\epsilon,\gamma)\Theta(E_{\rm max}(\gamma,\epsilon) - E)\Theta(E - E_{\rm min}(\gamma,\epsilon)). \tag{5.56}$$

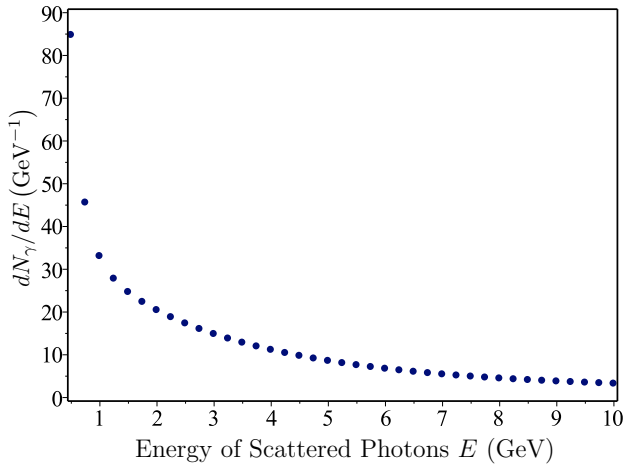

Figure 5.4: dN_γ/dE, given by Eq. (5.58), evaluated for the first set of parameters ($\alpha = 2$, $m_S = 4$ TeV, $m_\sigma = 40$ MeV).

This equation gives us the number of photons scattered per electron (or positron) per unit time; however, we require the total number of photons scattered by one electron before it loses all of its energy. Properly, we should integrate over t_S; this is complicated because γ is a function of t_S. Therefore, we will approximate

$$\frac{dN_\gamma}{dE} \approx \frac{d^2N}{dE\, dt_S} \cdot T, \tag{5.57}$$

where $T = 1/b_{\mathrm{ICS}}E_e = 1/b_{\mathrm{ICS}}\gamma m_e$ is the relevant time-scale for energy loss. This gives

$$\frac{dN_\gamma}{dE}(E) = \frac{1}{b_{\mathrm{ICS}}m_e}\int_{\gamma_{\min}}^{\gamma_{\max}} \frac{P(\gamma)}{\gamma}\int_0^\infty d\epsilon\, n_{\mathrm{ph}}(\epsilon)\sigma_{\mathrm{KN}}(E,\epsilon,\gamma)\Theta(E_{\max}(\gamma,\epsilon)-E)\Theta(E-E_{\min}(\gamma,\epsilon)). \tag{5.58}$$

This equation describes the total number of scattered photons of a particular energy, for a single electron or positron. In Fig. 5.4 we have evaluated this equation for the $\alpha = 2$ parameters for energies between 1 GeV and 10 GeV. We observe that the number of photons drops off rapidly as a function of energy. To find the total number of photons per unit energy per unit time, we must multiply by the rate of production of the high energy fermions, which

yields

$$\frac{dN_{\gamma,\mathrm{tot}}}{dE\,dt}(E) = \frac{dN_{\mathrm{BS}}}{dt}\frac{2}{b_{\mathrm{ICS}}m_e}\int_{\gamma_{\mathrm{min}}}^{\gamma_{\mathrm{max}}}\frac{P(\gamma)}{\gamma}\int_0^\infty d\epsilon\; n_{\mathrm{ph}}(\epsilon)\sigma_{\mathrm{KN}}(E,\epsilon,\gamma) \\ \cdot\,\Theta(E_{\mathrm{max}}(\gamma,\epsilon)-E)\Theta(E-E_{\mathrm{min}}(\gamma,\epsilon)). \tag{5.59}$$

From this result, the number of photons produced per unit energy per unit time, we can find the flux of gamma rays at the Solar System.

The scalar σ particles emitted in bound state formation are not emitted preferentially in any direction; therefore we expect an isotropic flux of these particles about the galactic center. Similarly, as a scalar particle, its decay products are emitted isotropically. Since the CMB is also isotropic, the distribution of scattered photons will be isotropic about the galactic center. Therefore, the photon flux per unit area may be approximated by an equivalent point source at the galactic center. This a rather crude approximation to the true diffuse flux is only meant to demonstrate that a detectable signal is possible. If such a signal were to be observed, more careful analysis should be done before attempting to fit this scenario to the data. Integrating out to $r = 8$ kpc will include the flux (as seen from Earth) from the hemisphere centered on the galactic center. Thus, we divide by 2π st to find the average flux per unit area per unit angle. We emphasize that this is an average; as a function of solid angle, we expect the signal to be greater near the galactic center and smaller further away from it.

Furthermore, we note that the production of the dark force mediator bosons σ scales as the density squared, which increases towards the galactic center. Since the σ bosons only travel 10^9 m before decaying into the fermions which scatter the CMB photons, the signal is dominated by the innermost region we consider. Since we have cut off our calculation at an inner radius of 1 kpc to avoid the known cusp in the NFW profile, the signal comes predominantly from the region near this cut. Therefore, the point source approximation is better than one may naively expect.

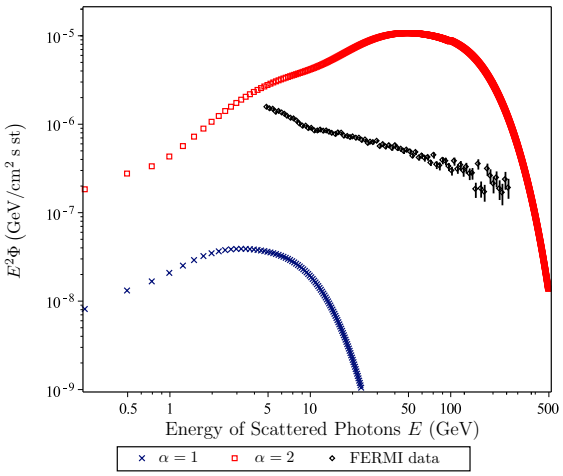

Figure 5.5: Depending on the model parameters, the signal can range from undetectable to already excluded. The signal is shown for $m_S = 4$ TeV, $m_\sigma = 40$ MeV for $\alpha = 2$, and $m_\sigma = 25$ MeV for $\alpha = 1$. For comparison, we also show data from Fermi LAT space telescope [158].

The average flux over the hemisphere centered on the galactic center, neglecting the galactic center itself, is

$$\Phi = \frac{dN_{\gamma,\text{tot}}}{dE\,dt} \cdot \frac{1}{2\pi \text{ st}} \cdot \frac{1}{4\pi(8 \text{ kpc})^2}. \tag{5.60}$$

To compare with the the sensitivity of the Fermi-LAT Gamma Ray Telescope, we evaluate $E^2\Phi$; this function is plotted in Fig. 5.5. The signal for $\alpha = 1$ is peaked at a lower energy and falls off more sharply, as we would expect because the binding energy is smaller. We see that $\alpha = 2$ produces a signal that is one order of magnitude larger than the values measured by Fermi-LAT, but the signal produced by $\alpha = 1$ is two orders of magnitude too small. Therefore, we conclude that sufficiently large couplings may produce a detectable signal. This suggests that WIMPonium models [131], which assume $\alpha \ll 1$, will not produce a detectable signal through bound state formation.

One can also show that the resulting signal is rather insensitive to the precise value of m_σ, provided that $\sigma \to e^+e^-$ remains the dominant decay. However, this parameter is highly

constrained by the astrophysical bounds discussed in Sec. 5.2. Finally, we also note that this signal depends relatively weakly on the cutoff we imposed to avoid the center cusp of the NFW profile. If we cut off the integral at 1 pc instead of 1 kpc, the signal would only be about 20 percent greater. Hence, including a central core is not likely to be raise the $\alpha = 1$ signal to detectable levels.

Next we will briefly discuss the possibility of producing a detectable positron or electron signal. This is a particularly interesting question because of the positron excess observed by PAMELA [159], which was confirmed by Fermi-LAT [160] and more recently AMS-II [161]. In order to travel from the galactic center to the Solar System relatively unimpeded, the fermions would need to be lower energy than those discussed above, which lost significant energy due to inverse Compton scattering. Energy loss due to inverse Compton scattering is somewhat suppressed for energies on the GeV scale.

The energy of the fermions produced by the decay of the dark force mediator depends only on the mass of the S particles and α, and to produce a significant number of dark force mediator particles, we must keep α relatively large. (Smaller values of α suppress the bound state formation cross section.) Therefore, to produce 10 GeV-scale positrons and electrons, we must decrease m_S to the scale of a few hundred GeV. This is below the scale typically discussed in the WIMPonium literature.

Naively, these parameters appear to run into difficulties with the halo ellipticity bounds [125]; for example, $m_S = 100$ GeV with $\alpha = 2$ appears to require m_σ be at least 232 MeV, for which the dominant decay is to muons instead of e^+e^-. (However, it should be noted that the analytic approximation for the cross section begins to break down at $m_\sigma \sim 100$ MeV). This appears to eliminate the possibility of an observable electron or positron excess. We have noted that a more detailed analysis of the bounds by [28] suggests that these bounds should be about an order of magnitude weaker. When we considered these bounds in Sec. 5.2, we parameterized this uncertainty with the parameter F, which is one if the consider-

ations of [28] are ignored. If one assumes $F \sim 0.1$, then a small region of parameter space remains which is consistent with the halo ellipticity bounds and m_σ is small enough (50-80 MeV) that the decay to e^+e^- dominates.

The analysis proceeds as above, up to the calculation of the inverse Compton scattering. For these lower energy electrons and positron, we do not expect inverse Compton scattering to be a significant effect. However, other effects can influence the shape of the spectrum observed at Earth. For example, positron annihilation may dramatically decrease the detected positron fraction, and we expect these particles to lose energy due to bremsstrahlung. A detailed analysis could be run using cosmic ray propagation software such as GALPROP.

However, we believe that it is unlikely that the resulting spectrum could reproduce the observed positron excess observed in PAMELA, Fermi-LAT, and AMS-II. The energy spectrum of fermions produced through the decay of dark force mediator particles is flat, and while this spectrum will no doubt be modified by a detailed analysis of propagation from the galactic center to the Solar System, we think the resulting $E^3\Phi$ is unlikely to be as flat as that observed by PAMELA. Furthermore, the positron excess extends to higher energies beyond that which can be accommodated by our model. Hence, we conclude that bound state formation in self-interacting asymmetric dark matter models is unlikely to account for this observed excess.

Finally, we note that it may be possible to adjust the parameters so that an excess of positrons or electrons above PAMELA's observations is produced, although again careful analysis of the propagation of the fermions would be necessary. If such an excess can be produced, we would expect it decrease at or before the TeV scale, at which point the spectrum would be limited due to inverse Compton scattering energy losses. Since no such behavior is observed in PAMELA's spectrum, one would translate this into bounds on the parameters. However, since the parameter space in which such a signal is potentially possible is already quite small, any resulting constraints (if any) would be quite weak.

5.4 Signal From Bremsstrahlung Emission of Dark Force Mediators

In this section, we discuss the signal produced by bremsstrahlung emission of a σ boson which decays to photons. We will see that the resulting signal is significantly beneath observational limits due to the substantial background at lower energies. In bremsstrahlung emission, the energy of the emitted σ boson is comparable to the kinetic energy of the S particles. Since these have a velocity of order 10^{-3}, the typical energy scale of bremsstrahlung emission will be 6 orders of magnitude below m_S. We will require $m_\sigma \ll \bar{v}^2 m_S/2$ because we do not want an additional suppression from the difficulty of emitting real σ bosons. This will influence the implementation of the constraints discussed in Sec. 5.2. (We note, however, that due to the contribution of the tail of the relative velocity distribution, we do not necessarily expect a sharp cutoff at the average kinetic energy).

Additional constraints come from the astrophysical considerations discussed in Sec. 5.2. As has been noted, Ref. [138] explores bounds for velocity-dependent cross sections; they present their results as a function of $v_{\max} = \sqrt{2\alpha m_\sigma/\pi m_S}$. If $v_{\max} \sim 10$ km/s, then the astrophysical constraints are consistent if $22.7/m_\sigma^2 m_S \lesssim 35$ cm^2/g $= 16000$ GeV^{-3} [27]. Combining this with $m_\sigma \ll \bar{v}^2 m_S/2$ yields

$$m_S \gg \left(\frac{22.7 \cdot 4}{\bar{v}^4 \cdot 16000 \text{ GeV}^{-3}}\right)^{1/3} = 1.3 \text{ TeV}. \tag{5.61}$$

Let us choose $m_S = 10$ TeV and $m_\sigma = 0.5$ MeV.

This does not include observations of large elliptical halos; now that we have set the masses, this will constrain the coupling of $A_{\sigma S}$. From our results in Sec. 5.2 extending the analysis of Ref. [125], we must take $\alpha = A_{\sigma S}^2/16\pi M_S^2 \lesssim 0.93$ (for the masses chosen above). If we choose to saturate this bound, we find $A_{\sigma S} = 68$ TeV. These values give $v_{\max} \approx 50$ km/s, consistent with our initial assumption that $v_{\max} \sim 10$ km/s.

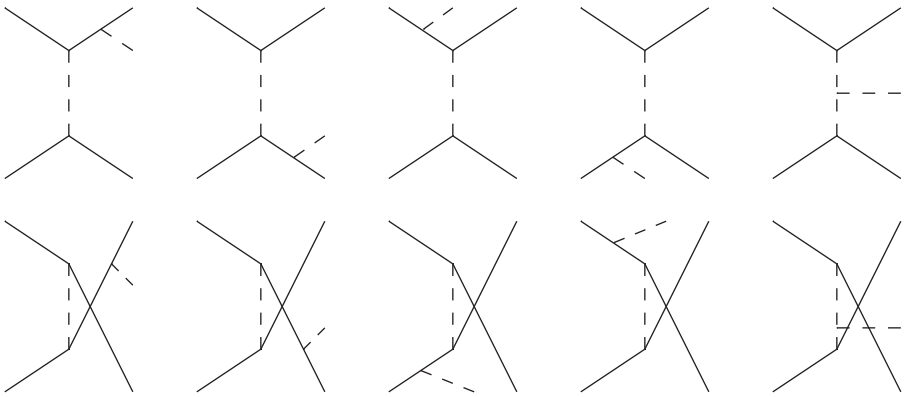

Figure 5.6: Diagrams contributing to the emission of a bremsstrahlung σ boson. The solid lines represent S bosons, while the dashed lines represent σ bosons. The top line shows t-channel scattering, while the bottom line shows u-channel scattering.

The σ boson can be emitted by either of the S particles, or it could also be emitted by the σ boson exchanged between the S particles through the A_σ coupling, which is thus far unconstrained. In order to enhance the signal, we will saturate the perturbativity bound, taking $A_\sigma = 3.5$ MeV. On the other hand, we will also consider $A_\sigma = 0$, which is equivalent to neglecting the two diagrams on the right of Fig. 5.6. While this certainly will not help to increase our signal, the properties of the signal will be qualitatively different in the two cases in interesting ways.

Next, we find the cross section for bremsstrahlung emission of a soft σ boson, which involves evaluating the 10 diagrams shown in Fig. 5.6. Note that the t- and u-channel diagrams cancel to lowest order in the $m_\sigma \to 0$ limit. Therefore, the resulting cross section is smaller than what one may naively expect. Let us denote the incoming four-momenta as p_1 and p_2, the outgoing momenta of the two S particles as p_3 and p_4, and the outgoing

momentum of the bremsstrahlung σ particle as p_5. The matrix element is then

$$\begin{aligned}-i\mathcal{M} = &-\frac{A_{\sigma S}^3}{(m_\sigma^2-(p_3-p_1)^2)(m_S^2-(p_4+p_5)^2)} - \frac{A_{\sigma S}^3}{(m_\sigma^2-(p_2-p_4)^2)(m_S^2-(p_3+p_5)^2)} \\ &- \frac{A_{\sigma S}^3}{(m_\sigma^2-(p_3-p_1)^2)(m_S^2-(p_2-p_5)^2)} - \frac{A_{\sigma S}^3}{(m_\sigma^2-(p_2-p_4)^2)(m_S^2-(p_1-p_5)^2)} \\ &- \frac{A_{\sigma S}^2 A_\sigma}{(m_\sigma^2-(p_2-p_4)^2)(m_\sigma^2-(p_3-p_1)^2)} + (p_3 \leftrightarrow p_4), \end{aligned} \tag{5.62}$$

where the last term, in which the momenta p_3 and p_4 are switched, represents the contribution of the bottom row of diagrams. We will specialize to the center of mass frame; we note that the total cross section is a relativistic invariant and therefore it is irrelevant what frame it is calculated in. Without a loss of generality we write the momenta as

$$\begin{aligned} p_1^\mu &= \left(m_S + \frac{|\boldsymbol{p}_I|^2}{2m_S}, 0, 0, |\boldsymbol{p}_I|\right), \\ p_2^\mu &= \left(m_S + \frac{|\boldsymbol{p}_I|^2}{2m_S}, 0, 0, -|\boldsymbol{p}_I|\right), \\ p_3^\mu &= \left(m_S + \frac{|\boldsymbol{p}_3|^2}{2m_S}, |\boldsymbol{p}_3|\sin(\theta_3)\cos(\phi_3), |\boldsymbol{p}_3|\sin(\theta_3)\sin(\phi_3), |\boldsymbol{p}_3|\cos(\theta_3)\right), \\ p_4^\mu &= \left(m_S + \frac{p_3^2}{2m_S}, |\boldsymbol{p}_4|\sin(\theta_4)\cos(\phi_4), .|\boldsymbol{p}_4|\sin(\theta_4)\sin(\phi_4), |\boldsymbol{p}_4|\cos(\theta_4)\right), \\ p_5^\mu &= \left(\sqrt{m_\sigma^2+|\boldsymbol{p}_5|^2}, |\boldsymbol{p}_5|\sin(\theta_5), 0, |\boldsymbol{p}_5|\cos(\theta_5)\right). \end{aligned} \tag{5.63}$$

The cross section is given by

$$\sigma_{\text{brem}} = \int \frac{|\mathcal{M}|^2}{4(E_1+E_2)^2}(2\pi)^4\delta^4(p_1+p_2-p_3-p_4-p_5)\; d\text{Lips}, \tag{5.64}$$

where the extra $1/2$ comes from the two identical particles in the final state and dLips is the well-known Lorentz-invariant phase space for the final state particles,

$$d\text{Lips} = \prod_{i=3}^{5} \frac{d^3\boldsymbol{p}_i}{(2\pi)^3 2E_i}. \tag{5.65}$$

In the phase space denominators, we may make the approximation $E_1 = E_2 = E_3 = E_4 = mS$, and we integrate over the three-momentum delta function, setting $\boldsymbol{p}_3 = -\boldsymbol{p}_4 - \boldsymbol{p}_5$. When the S particles are nonrelativistic, the energy delta function is

$$\delta\left(\frac{|\boldsymbol{p}_I|^2}{m_S} - \frac{|\boldsymbol{p}_4|^2}{2m_S} - \frac{|\boldsymbol{p}_4+\boldsymbol{p}_5|^2}{2m_S} - \sqrt{m_\sigma^2-|\boldsymbol{p}_5|^2}\right). \tag{5.66}$$

Let us define θ_{45} to be the angle between $\boldsymbol{p}_4$ and $\boldsymbol{p}_5$. The delta function enforces

$$\frac{|\boldsymbol{p}_I|^2}{m_S} - \frac{|\boldsymbol{p}_4|^2}{m_S} - \frac{|\boldsymbol{p}_5|^2}{2m_S} - \frac{|\boldsymbol{p}_4||\boldsymbol{p}_5|\cos(\theta_{45})}{m_S} - \sqrt{m_\sigma^2 + |\boldsymbol{p}_5|^2} = 0, \tag{5.67}$$

which can be solved for $|\boldsymbol{p}_4|$ in terms of $|\boldsymbol{p}_5|$ and θ_{45}:

$$|\boldsymbol{p}_4| = -\frac{|\boldsymbol{p}_5|\cos(\theta_{45})}{2} + \frac{1}{2}\sqrt{|\boldsymbol{p}_5|^2\cos^2(\theta_{45}) - 2|\boldsymbol{p}_5|^2 + 4|\boldsymbol{p}_I|^2 - 4m_S\sqrt{m_\sigma^2 + |\boldsymbol{p}_5|^2}}. \tag{5.68}$$

We must of course ensure that the result is positive. By our choice of coordinates, the $d\phi_5$ integral is trivial; this leaves the integrals over θ_4, ϕ_4, θ_5, and $d|p_5|$ to be done numerically. This integral is not infrared divergent due to the nonzero mass of the σ boson. Since the initial momentum in the center of momentum frame is $p_I = v_{\rm rel}/2m_S$, the above calculation gives the cross section as a function of $v_{\rm rel}$. We can then average over the relative velocity

$$\sigma_{\rm brem} = \int P(v_{\rm rel})\sigma(v_{\rm rel})\, dv_{\rm rel} \tag{5.69}$$

using Eq. (5.10).

Finally, we address Sommerfeld factors, which multiply the cross section and naively can have a large impact at low velocities. (Note that this is a multiplicative factor in addition to the typical $1/v$ behavior of the cross section). These describe the formation of a quasi-bound state during the interaction; the modified cross section is

$$\sigma_{\rm Somm} = \frac{\pi\alpha/v}{1 - \exp(-\pi\alpha/v)}\sigma. \tag{5.70}$$

For the parameters under consideration, these factors can be extremely large, of order 10^3 or 10^4. However, it has been argued that in this regime the Sommerfeld factor as written above is unreliable; additional diagrams beyond the ladder diagrams implicitly summed in the above equation must be taken into account and a proper resummation suggests the factors are of order $\mathcal{O}(1)$ to $\mathcal{O}(10)$ [162]. This is supported by some experimental evidence [163,164], including more recent observations at BABAR [165]. We will argue below that even these large Sommerfeld factors (if correct) would not be sufficient to produce a detectable signal through bremsstrahlung emission.

Evaluating the cross section numerically, including averaging over the relative velocity of the incoming particles, gives $\sigma_{\rm brem} = 0.0108\ \text{GeV}^{-2}$. This is the same order of magnitude as the $\alpha = 2$ cross section for bound state formation; we note, however, that we did not have to increase the coupling α into the non-perturbative regime in order to reach this value. In general, for comparable values of α, the bremsstrahlung cross sections are significantly larger than the bound state formation cross section.

The rate of production of bremsstrahlung σ bosons is

$$\frac{dN_\sigma}{dt} = \iint n_S(r)^2 v_{\rm rel} \sigma_{\rm brem}(v_{\rm rel}) P(v_{\rm rel})\, dV\, dv_{\rm rel}, \tag{5.71}$$

where $n_S(r) = \rho(r)/m_S$ is the number density of dark matter S particles and $\rho(r)$ is given by Eq. (5.6). We have also averaged over the relative velocity of the S bosons, and the integration extends from 1 kpc to 8 kpc, the distance from the solar system to near the galactic center, ignoring the central cusp. When we saturate the perturbativity bound with A_σ, we find $dN_\sigma/dt = 3.21 \cdot 10^{13}$ GeV, or $4.87 \cdot 10^{37}\ \text{s}^{-1}$. As might be expected, for $A_\sigma = 0$ we find the lower rate $dN_\sigma/dt = 1.45 \cdot 10^{11}\ \text{GeV}^{-1} = 2.20 \cdot 10^{35}\ \text{s}^{-1}$. The substantial difference (two orders of magnitude) shows that the first case is dominated by the diagrams in which the emitted σ boson comes from the exchanged σ boson. Since the two cases are dominated by different diagrams, we expect qualitative differences in their spectra.

We will show that bremsstrahlung emission does not produce a detectable signal, while we found that bound state formation can make such a signal for sufficiently large couplings. Since this is perhaps a surprising result, one may find it beneficial to compare the calculations. We emphasize, however, that such comparisons must be made carefully, since the bound state calculations were performed in a different region of parameter space. We wish to emphasize that for any fixed perturbative value of α, the rate of bremsstrahlung production will always be much greater than the rate of bound state formation, as one would expect. However, if one compares the value of dN_σ/dt found above with $dN_{\rm BS}/dt$ given in the previous section, which are evaluated at different parameters, one finds that $dN_{\rm BS}/dt$ is larger by about an

order of magnitude, even though we have chosen parameters such that the cross sections are comparable. This is a result of taking $m_S = 10$ TeV here as opposed to 4 TeV above; increasing m_S decreases the number density $n_S(r)$.

The spectrum of the emitted σ bosons (per $SS \to SS\sigma$ event) is

$$\frac{dN_\sigma}{dE_\sigma} = \frac{1}{\sigma_{\text{brem}}} \frac{d\sigma_{\text{brem}}}{dE_\sigma}. \tag{5.72}$$

As we might expect, this spectrum is peaked at 600 keV, which is on the same scale as the kinetic energy. The resulting spectrum of the total produced σ bosons per unit time is

$$\frac{d^2N_\sigma}{dt\, dE_\sigma} = \left(\int n_S(r)^2 \, dV \right) \int v_{\text{rel}} \frac{d\sigma_{\text{brem}}}{dE_\sigma} P(v_{\text{rel}}) \, dv_{\text{rel}}. \tag{5.73}$$

For $m_\sigma = 0.5$ MeV, the dominant decay mode of the σ boson is $\sigma \to \gamma\gamma$, which is described by Eq. (5.21). If we assume the mixing angle between the σ boson and the Higgs boson is 10^{-3}, then the typical lifetime of the produced σ bosons is 10^5 s, during which they travel about 10^{14} m, which is significantly less than the 10^{20} m between the galactic center and the Solar System. Therefore, they will decay in flight, and the spectrum of the resulting photons is given by

$$\frac{d^2N_\gamma}{dE_\gamma\, dt} = 2 \int \frac{d^2N_\sigma}{dt\, dE_\sigma} P(E_\gamma, E_\sigma) \, dE_\sigma, \tag{5.74}$$

where the distribution of photon energies, as a function of the initial σ boson energies, is given by Eq. (5.21). (The E_γ dependence appears in evaluating the Heaviside step functions.) As we would expect, this spectrum is peaked around 300 keV. We note that the tail of the spectrum is larger for smaller A_σ values. As a result, the signal for $A_\sigma = 0$ will be skewed towards higher energies.

The production of dark force mediator particles results in an isotropic flux of these particles about the galactic center; similarly, we expect the flux of their decay products to be isotropic about the galactic center. Again we estimate the flux using an equivalent point source at

the galactic center,

$$\Phi = \frac{1}{4\pi d^2} \cdot \frac{1}{2\pi \text{ st}} \frac{d^2 N_\gamma}{dE_\gamma \, dt}, \tag{5.75}$$

where $d = 8$ kpc is the distance from the galactic center to the solar system. As before, since we have calculated the number of produced σ bosons out to a radius of 8 kpc, the signal included here comes from the hemisphere centered on the galactic center, which explains the 2π st in the denominator. The true flux will be somewhat greater towards the galactic center and somewhat less towards the edges; however, this is a relatively small effect, contributing perhaps an order of magnitude increase as we approach the center. (We remind our reader that we are already neglecting the contribution from the cusp at the galactic center itself.)

Again, with the same caveats as above, let us compare with the bound state case. The photon energies here are spread out over the scale of 100 keV, whereas the photon signal from CMB scattering we produced from the bound state production is spread over the scale of 100 GeV. However, a single high energy fermion produces about 10^2-10^3 GeV-scale photons through scattering off with CMB photons, while each σ boson produced through bremsstrahlung produces a mere two photons. As a result, the estimated ratio of fluxes is $\Phi_{\text{brem}}/\Phi_{\text{BS}} \sim 10^4$ or 10^5. We note that since the two scenarios are in different regions in parameter space, this cannot be interpreted as the ratio of actual bremsstrahlung-produced photons to bound state produced photons in the galaxy.

As noted, the relevant energy scale for bremsstrahlung emission is on the scale of hundreds of keV. Astrophysical backgrounds are significantly larger at this smaller scale; the SPI instrument on the INTEGRAL experiment records $E_\gamma^2 \Phi$ on the order of 1-10 keV/cm^2 s st for energies between 20 keV and 1000 keV [166]. The flux of bremsstrahlung-emission produced photons cannot be distinguished from this large background. The resulting signal is shown in Fig. 5.7; the larger $A_\sigma = 3.5$ MeV signal is about 7 orders of magnitude beneath INTEGRAL's observations. We also can see the qualitative difference in the signal shapes alluded to above; this is because for $A_\sigma = 3.5$ MeV, the rightmost two diagrams of Fig.

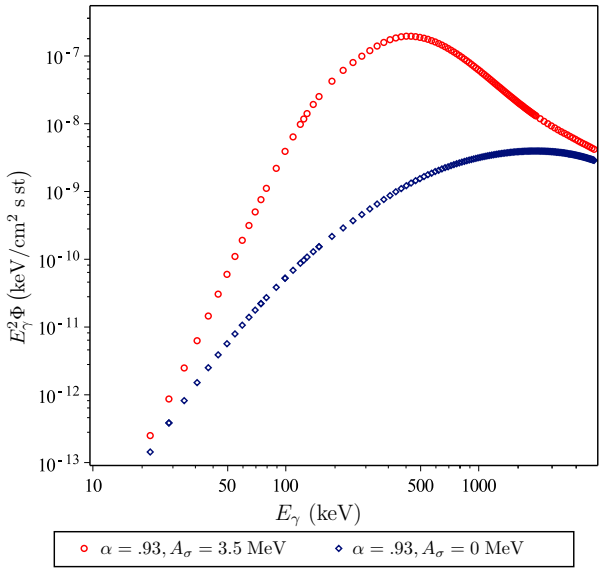

Figure 5.7: The flux of gamma rays produced by bremsstrahlung emission of σ particles and their subsequent decay for $m_S = 10$ TeV, $m_\sigma = 0.5$ MeV.

5.6 dominate, whereas these are absent if $A_\sigma = 0$. Without these diagrams, the signal is significantly smaller, but it is peaked at higher energies.

We have noted above we have not included the Sommerfeld enhancement. Even if we assume that the naive Sommerfeld factor given by Eq. (5.70) is accurate to arbitrarily large scales, this enhancement is not sufficient to produce a detectable signal. For the parameters in the range discussed, the naive enhancement is of order 10^3 or 10^4, which is still too small to produce the seven orders of amplification required for the signal to be detectable.

The signal can be increased by increasing the couplings; and indeed, as discussed in Sec. 5.2, there is some uncertainty in the halo ellipticity bounds (which set the upper bound on $A_{\sigma S}$). However, producing a detectable signal requires increasing the coupling α to $\sim 10^3$, well outside the perturbative regime and far beyond what can be made consistent with the halo ellipticity bounds. It is true that A_σ is unrestricted by astrophysical bounds, but in order to amplify the two relevant diagrams to the scale of INTEGRAL's observations, we

would need to take $A_\sigma/m_\sigma \sim 10^3$, which is again unreasonably large.

Finally, even if the signal produced near the Milky Way's galactic center is not detectable, perhaps such processes enhance the gamma ray or x-ray emission of nearby dwarf galaxies sufficiently to be detectable; however, a simple estimate reveals that this is not the case. Even if the signal calculated above, for the Milky Way galaxy, was somehow shrunk into a dwarf galaxy 40 kpc from us which covered a 3°-by-3° patch of the sky, the number of counts expected in an ideal 1 m^2 detector is of order 10^{-5} keV^{-1} s^{-1}, which is again well below the background emission. Therefore, we conclude that bremsstrahlung emission of dark force mediator particles cannot produce significant indirect detection signals.

5.5 Summary

We have considered indirect detection signals produced by a minimal scalar asymmetric self-interacting dark matter. Due to the asymmetry, dark matter today cannot self-annihilation, and this typical indirect detection signal is absent. However, the decays of a mediator particle can potentially produce high energy cosmic rays, reopening the possibility of indirect detection. We considered two processes that can produce a significant number of dark mediators: bremsstrahlung emission and dark bound state formation. The first process produced a large flux; however, it was at lower energies where the background is larger. We saw that for sufficiently large couplings, bound state formation can potentially produce a signal above current observational limits.

CHAPTER 6

Atomic Dark Matter with a Massive Gauge Mediator

The previous chapter was concerned with the possibility of indirect detection in a minimal scalar self-interacting asymmetric dark matter model. As noted, we ignored issues of cosmology, which allowed us to assume that dark matter today is single-component. Here, we return to these questions in a slightly different model, inspired by the Standard Model. The fields that make up normal matter are the fermionic quarks and electrons; therefore, we likewise take dark matter to be fermionic. As discussed in Ch. 1, the Standard Model is defined in terms of its gauge group; therefore, instead of considering Yukawa interactions in the dark sector, we will introduce a gauge interaction. In this chapter, we will explore the nature of dark matter in the resulting model, showing that it is generally multi-component, with two species of ions and dark atoms. Therefore, there are several types of intra- and inter-species interactions, and so we return to the issues surrounding scattering in dark matter halos.

6.1 Atomic Dark Matter Model

Thermal relic dark matter (whether asymmetric or not) assumes the existence of interactions which efficiently annihilate dark matter until it reaches the observed density. In asymmetric dark matter models, the existence of a particle-antiparticle asymmetry reduces the annihilation rate, and therefore, to reach a particular relic density a larger annihilation cross section is required. The cross section does not need to be much larger; if the annihilation cross section is about 2.4 times larger than the canonical thermal value, then antiparticles contribute less than 1 percent of the total dark matter density [168]. The annihilation cross section to Standard Model particles is constrained by direct detection experiments and col-

lider experiments; for asymmetric dark matter, these constraints push the interaction scale above the weak scale [170–173]. To avoid these constraints, one can consider a model in which dark matter does not annihilate directly into Standard Model particles; one possibility is that it annihilates into other dark sector particles, which may or may not subsequently decay into Standard Model particles.

As means of generating these interactions, we consider a model with a new Abelian gauge group in the dark sector (see e.g. [113, 114, 174]). An advantage of this type of interaction is that it can result in an accidental particle number symmetry at low energies, which is necessary in asymmetric dark matter models. We will refer to this conserved particle number as the dark baryon number B_D. In this section, we argue that the dark matter in this model is multi-component in much of the relevant parameter space.

6.1.1 Massless Mediator Boson

First, we argue for the multi-component nature of dark matter in the case of a perfect $\mathrm{U_D}(1)$ gauge symmetry, in which case the dark photon is exactly massless. The dark baryon number asymmetry must be generated by interactions which are invariant under this gauge symmetry. Therefore, of the relic particles with the net dark baryon number, half must be positively charged under $\mathrm{U_D}(1)$, and half negatively charged. In general, we expect dark matter to consist of the lightest positively charged dark particle and the lightest negatively charged dark particle. A comparison can be made to the Standard Model; the baryonic asymmetry carried by protons is inevitably associated a net positive electric charge. However, the universe as a whole is electrically neutral due to a corresponding negative charge carried by electrons. As the lightest positively charged and negatively charged particles in the Standard Model, both protons and electrons are stable. Consequently, we will call the two stable dark ions the dark proton p_D and the dark electron e_D. We will designate their masses by m_p and m_e, their $\mathrm{U_D}(1)$ charges by q_p and q_e, etc. We will further assume that these particles are fermions, and we assume $m_p \geq m_e$.

The low-energy effective Lagrangian of this model is

$$\mathcal{L} = \bar{p}_D(i\not{\partial} - m_p)p_D + \bar{e}_D(i\not{\partial} - m_e)e_D - \frac{1}{4}F_{D\mu\nu}F_D^{\mu\nu}, \tag{6.1}$$

where $F_D^{\mu\nu} = \partial^\mu A_D^\nu - \partial^\nu A_D^\mu$. (As in QED, we will use A_D^μ for the field in the Lagrangian, and γ_D for the dark photon when discussing processes such as $e_D p_D \to \gamma_D \gamma_D$.) We note that there is no problem in writing mass terms for the dark electron and dark proton, as they are singlets under the Standard Model gauge group. The covariant derivative is $D^\mu = \partial^\mu + iq_i g A_D^\mu$, where g is the $\mathrm{U_D}(1)$ coupling constant; we will generally use $\alpha_D = g^2/4\pi$ instead. For this to be an asymmetric dark matter model, we must have a dark matter population which carries a net dark baryon number B_D; this implies the existence of high-energy interactions which generate gauge-invariant and B_D-violating effective operators, which are suppressed at low energies.

Since the relic dark matter population consists of oppositely charged dark protons and dark electrons, they may form dark hydrogen atoms, of mass $m_H = m_p + m_e - \Delta$, where $\Delta = \mu_D \alpha_D^2/2$ is the binding energy. (The dark reduced mass is $\mu_D = m_p m_e/(m_p + m_e)$.) This must satisfy the consistency condition

$$4\mu_D \leq m_H + \Delta. \tag{6.2}$$

Dark matter today generically consists of dark protons, dark electrons, and dark hydrogen atoms; thus it is naturally multi-component. The cosmology of this scenario is discussed in Ref. [175], and we will use some of these results in Sec. 6.2 below.

6.1.2 Massive Mediator Boson

While long-range self-interactions are desirable insofar as they ameliorate disagreements between simulations of cold non-interacting dark matter and astrophysical observations, these self-interactions cannot be infinitely long-ranged; as one example, such interactions would influence the clustering of matter at very large distances. Consequently, dark matter

self-interactions must be screened, for example, due to the mediator having a nonzero mass. This is the case we consider; we assume the dark photon has a small but nonzero mass M_D. Since this requires the $\mathrm{U_D}(1)$ symmetry to be broken, we introduce a dark Higgs field, a complex scalar ϕ_D with charge q_ϕ. The additional terms in the Lagrangian are

$$\mathcal{L}_\phi = D_\mu \phi_D^\dagger D^\mu \phi_D - \lambda(\phi_D^2 - v_D^2)^2. \tag{6.3}$$

At zero temperature, the field ϕ_D has a nonzero vacuum expectation value $\langle \phi_D \rangle = v_D$. Consequently, the dark photon acquires a mass

$$M_D = (8\pi q_\phi^2 \alpha_D)^{1/2} v_D. \tag{6.4}$$

(Since v_D will not appear elsewhere in our analysis, we will treat M_D as a free parameter.) We expand about the vacuum in the unitary gauge, in which $\phi_D = v_D + \varphi_D/\sqrt{2}$. The physical Higgs boson has a mass $m_\varphi^2 = 4\lambda v_D^2$, and the interaction terms involving the dark Higgs boson and the dark photon are

$$\mathcal{L}_{\text{int}} = (2\pi q_\phi^2 \alpha_D)\varphi_D^2 A_\mu A^\mu + (4\pi q_\phi^2 \alpha_D)^{1/2} M_D \varphi_D A_\mu A^\mu - \frac{(\pi q_\phi^2 \alpha_D)^{1/2} m_\varphi^2}{M_D}\varphi_D^3 - \frac{\pi q_\phi^2 \alpha_D m_\varphi^2}{2M_D^2}\varphi_D^4. \tag{6.5}$$

Now that the dark photon is massive, it may be able to decay. We augment our model with a kinetic mixing between the dark photon and the hypercharge mediator field, involving the renormalizable operator [176, 177]

$$\mathcal{L}_{\text{mix}} = \frac{\epsilon}{2} F_{Y\mu\nu} F_D^{\mu\nu}. \tag{6.6}$$

This allows the decay into Standard Model fermions if $M_D > 1.022$ MeV at a rate [178]

$$\Gamma_{\gamma_D \to f^+ f^-} = \frac{f_{EM}}{3}\epsilon^2 \alpha_D M_D, \tag{6.7}$$

where f_{EM} counts the number of available decay channels. If this coupling exists but the dark photon is lighter, it may still decay into neutrinos via the resulting mixing with the Z boson, or to three photons through a loop diagram. However, these decay rates are severely suppressed. For completeness, we also note that the dark Higgs may mix with the Standard Model Higgs boson through the term $\lambda_{\phi H}\phi_D^2 H_{SM}^2$.

As in the Standard Model, finite temperature corrections will restore the $\mathrm{U_D}(1)$ symmetry at high temperatures; therefore, the cosmology of this model includes an additional phase transition, discussed in detail in Sec. 6.2 below. This $\mathrm{U_D}(1)$-breaking phase transition occurs at a dark sector temperature of

$$T_{D,\mathrm{PT}} \sim v_D = M_D/(8\pi q_\phi^2 \alpha_D)^{1/2}. \tag{6.8}$$

As long as this phase transition takes place after the B_D-asymmetry generation, our previous arguments from the massless model apply, and this establishes that equal asymmetries of p_D and e_D must be generated.

B_D-asymmetry generation must take place before annihilations diminish the dark matter abundance below the observed dark matter density, that is, when

$$Y_p = n_p/s > \Omega_{\mathrm{DM}}\rho_c/(s_0 m_p) \approx 10^{-11}(100\ \mathrm{GeV}/m_p). \tag{6.9}$$

Here n_p is the number density of dark protons, s is the entropy density of the universe, ρ_c and s_0 are the critical density and entropy density of the universe today, and $\Omega_{DM} \sim 0.2$. For thermal dark matter, this implies $m_D/T_{\mathrm{asym}} < 25 + \ln(m_D/100\ \mathrm{GeV})$, where T_{asym} is the dark-sector temperature at the time of the dark asymmetry generation. In realistic models one expects B_D-asymmetry generation to occur at larger temperatures, but let us consider the lower bound of this constraint (replacing the thermal relic with the dark proton, m_p). This is still necessarily above the $\mathrm{U_D}(1)$ phase transition temperature if

$$M_D \lesssim (8\pi q_\phi^2 \alpha_D)^{1/2} m_p/25. \tag{6.10}$$

(Note that this assumes that the dark sector and Standard Model sectors are in thermal equilibrium when the asymmetry is generated, as expected if the same process is responsible for the baryon asymmetry as well as the B_D-asymmetry. If this is not the case, there would be logarithmic corrections involving the ratio of the sectors' temperatures.)

If the $\mathrm{U_D}(1)$ symmetry is unbroken when the asymmetry is generated, the gauge invariance argument in Sec. 6.1.1 applies; thus this condition implies the creation of an e_D asymmetry

along with the p_D asymmetry. (It does not automatically ensure the survival of a significant e_D abundance at late times. After the breaking $\mathrm{U_D}(1)$, the dark electron no longer carries a conserved quantum number, and the produced asymmetry may be washed out. To determine whether there is a significant abundance of dark electrons today, we will need to consider the sequence of various cosmological events, and therefore, we postpone a discussion of the relic dark electron abundance until Sec. 6.2.4.)

We note that it is not necessary to use the Higgs mechanism to give the dark photon a mass; instead one may use the Stückelberg mechanism. In this case, the particles charged under the $\mathrm{U_D}(1)$ gauge group still couple to the dark photon through a conserved current, and the above arguments from gauge invariance apply. Hence, equal asymmetries of p_D and e_D are again generated.

Although we will consider halo dynamics more carefully in Sec. 6.3, we will present some crude estimates here considering only dark proton-dark proton scattering. This will show that condition 6.10 is satisfied in much of the parameter space of interest, in which ion collisions are significant but compatible with the ellipticity of larger halos. These desiderata fit together most comfortably if the self-scattering is long range, which occurs if the momentum transfer dominates over the mass of the mediator, or $m_p v/2 \gtrsim M_D$. For dwarf-galaxy-size halos (with $v \sim 10\ \mathrm{km}/s$) and larger, this implies

$$M_D \lesssim 2 \cdot 10^{-5}\ m_p. \tag{6.11}$$

The region of parameter space that fulfills condition (6.10) generally contains the long-range scattering regime where Eq. (6.11) is satisfied; specifically, this holds for couplings $\alpha_D \gtrsim q_\phi^2 \cdot 7 \cdot 10^{-9}$. Taking into account the minimum value of α_D required for efficient annihilation in the early universe, this corresponds to $m_p \gtrsim q_\phi^{-2} \cdot 45\ \mathrm{keV}$, which covers all of the range of interest for $q_\phi \sim \mathcal{O}(1)$. In other words, if one wants significant scattering in dark matter halos in even this rather minimal model, one must consider the regime in which dark

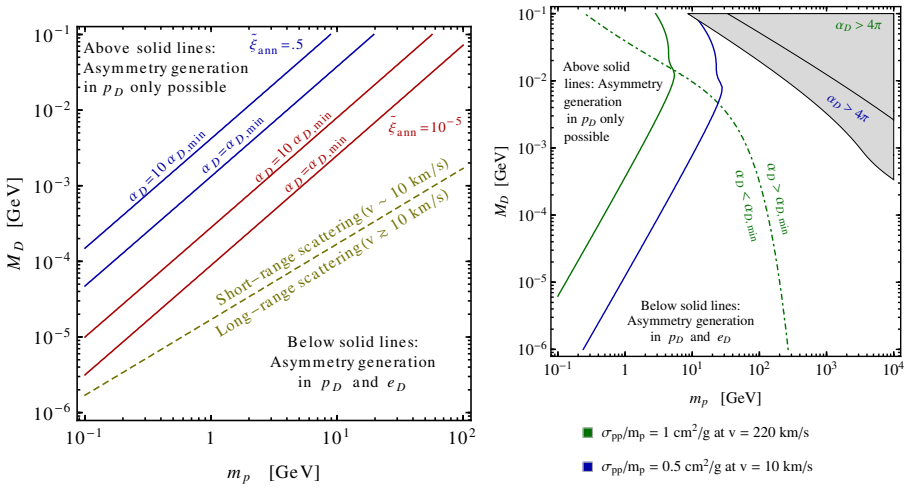

Figure 6.1: The left plot demonstrates the region of parameter space in which dark asymmetry is generated in two species; the right plot shows the same for parameter choices relevant to halo structure. $\alpha_{\min}$ is the minimum coupling for efficient annihilation of the symmetric dark matter component. (These plots are for $q_\phi = 1$; $\alpha_D > 4\pi$ exceeds the perturbativity bound.)

matter is multi-component. Even if one doesn't necessarily require the $p_D p_D$ interaction to be long ranged, condition (6.10) encompasses much of the parameter space ion which this scattering can affect the dynamics of smaller halos. We illustrate this in Fig. 6.1.

The first plot shows clearly that dark matter is generally multi-component across a wide range of dark proton masses, although the condition depends somewhat weakly on the coupling α_D and $\tilde{\xi}_{\text{ann}}$, the ratio of the temperature between the dark and visible sectors when dark proton self-annihilations freeze out. Increasing the coupling α_D increases the parameter space in which dark matter is multi-component. To the right of the golden dashed line, dark proton collisions are long range in halos with $v \sim 10$ km/s; in this regime, self-interactions may modify the shape of smaller halos. As claimed, this region is entirely contained in the multi-component regime.

The second plot focuses on choices of parameters relevant for large and small halos. For the blue lines, we have chosen the minimum value of α_D for which (under the assumption of single-component dark matter) there can be a significant effect on the dynamics of small halos, corresponding to $\sigma_{pp}/m_p = 0.5$ cm^2/g at $v = 10$ km/s. (Because the cross section decreases with speed, the ellipticity of larger halos is retained at this value.) The green lines refer to the opposite limit; we have taken $\sigma_{pp}/m_p = 1$ cm^2/g at $v = 220$ km/s, which gives the maximum value of α_D that is currently considered compatible with the observed ellipticity of large halos. To the left of the dot-dashed line, $\alpha_D < \alpha_{D,\min}$ and the scenario does not appear viable, that is, naively the dark proton self-interaction in halos is too strong. As we discuss in Sec. 6.3, this region of parameter space may produce viable scenarios when the formation of dark atoms is taken into account. ($\alpha_{D,\min}$ will be defined in Sec. 6.2.)

Thus, we see that in much of the parameter space an abundance of both dark electrons and dark protons is produced, and if these ions survive until late times, they may form $\mathrm{U_D}(1)$ neutral bound states. In the non-relativistic regime and for a massive dark photon, the interaction between the p_D and e_D ions is given by a Yukawa potential

$$V = -\frac{\alpha_D}{r} e^{-M_D r}, \tag{6.12}$$

which has bound state solutions if $M_D < \mu_D \alpha_D$. These bound states can be found by solving Schrödinger's equation using a Hulthén potential [179,180]. (Note that as α_D becomes large, the Dirac equation must be solved instead; if $\alpha_D \geq 1/2$, the ground state has nonzero angular momentum.) The binding energy of the bound state is approximately

$$\Delta \approx \frac{\alpha_D \mu_D^2}{2} \left(1 - \frac{M_D}{\alpha_D \mu_D}\right)^2. \tag{6.13}$$

Dark atoms form via the process $p_D + e_D \to H_D + \gamma_D(\omega)$, where ω is the energy of the emitted dark photon γ_D, which is given by

$$\omega + \frac{\omega^2 - M_D^2}{2m_H} = \Delta + \frac{\mu_D v_{\mathrm{rel}}^2}{2}, \tag{6.14}$$

where $v_{\rm rel}$ is the relative velocity of the p_D and e_D ions in the center-of-momentum frame. This process only occurs if $\omega > M_D$, and note that for $\Delta + \mu_D v_{\rm rel}^2/2 \ll m_H$, the energy is approximately $\omega \approx \Delta + \mu_D v_{\rm rel}^2/2$.

Although the cosmological evolution of this model will be discussed more thoroughly in the subsequent section, for now we note that bound states form when the temperature drops beneath the binding energy. Thus the condition for their formation in the early universe is

$$M_D < \Delta \approx \frac{1}{2}\alpha_D^2 \mu_D. \tag{6.15}$$

This constraint is stronger than the condition for bound states to exist, and it is also stronger than Eq. (6.10), the condition for the asymmetry to be generated in the dark sector before $\mathrm{U_D}(1)$ is broken.

6.2 Cosmology

Although this atomic dark matter model is rather minimal, with only a single Abelian gauge group necessarily broken in order to screen large-scale effects, it has a complex cosmology. We generically expect two phase transitions: one when the $\mathrm{U_D}(1)$ symmetry is broken and another when the ions form bound states at $T \sim \Delta$. Additionally, there are several other cosmologically important events, such as dark asymmetry generation and the thermal freeze-out of annihilation reactions. A generic sequence of events in this model is given in Table 6.1; however, other orderings are possible.

6.2.1 Kinetic Equilibrium Between the Standard Model and Dark Sectors

The cosmological sequence outlined below will depend in some detail on the ratio of temperatures in the dark and Standard Model sectors. We assume, as is typical in asymmetric dark matter scenarios, that the baryon number and dark baryon number asymmetries are generated when the two sectors are in thermal equilibrium at a common temperature $T_{\rm asy}$

Event	Temperature Scale
Dark asymmetry generation	$T_{\rm asy}$
Thermal decoupling of dark and Standard Model sectors	$T_{\rm dec}$
Freeze-out of $p_D\bar{p}_D$ annihilations	$T_{D,p,\rm fo} \approx m_p/30$
Freeze-out of $e_D\bar{e}_D$ annihilations	$T_{D,e,\rm fo} \approx m_e/30$
Dark recombination	$\Delta \gtrsim T_{D,\rm DR} \gtrsim \Delta/50$
$\mathrm{U_D}(1)$-breaking phase transition	$T_{D,\rm PT} \sim M_D/(8\pi q_\phi^2 \alpha_D)^{1/2}$
Dark photon chemical decoupling	$T_{D,\gamma_D} \sim M_D/[40(8\pi q_\phi^2 \alpha_D)^{1/2}]$
Big Bang Nucleosynthesis	$T_{V,\rm BBN} \approx 1$ MeV

Table 6.1: One possible sequence of cosmological events and temperature scales. The subscripts D and V refer to the dark and Standard Model sectors; for asymmetry generation and thermal decoupling of the two sectors we assume a common temperature.

and that the sectors later decoupled at $T_{\rm dec}$. Afterwards, the comoving entropy is separately conserved in each sector, which means that the two temperatures T_D and T_V are related by

$$\frac{g_D T_D^3}{g_V T_V^3} = \frac{g_{D,\rm dec}}{g_{V,\rm dec}}, \tag{6.16}$$

where g_D and g_V are the number of relativistic degrees of freedom in thermal equilibrium in the dark and Standard Model (visible) sectors respectively. We define the ratio $\xi \equiv T_D/T_V$; this can take values $\lessgtr 1$ depending on the number of degrees of freedom in each sector and the order in which they decouple. ξ will evolve in time; as an example, if the mass scale of the dark sector is large, then these degrees of freedom will decouple first, resulting in $\xi > 1$. However, the later decoupling of degrees of freedom in the visible sector will push ξ to lower values. In the following, we will find it useful to define

$$\tilde{\xi} = \min[1, \xi]. \tag{6.17}$$

We note that if the dark photon (or the dark Higgs boson ϕ_D) is sufficiently long-lived, it can contribute to the relativistic energy density of the universe, which is constrained by

observations. In this case, these observations require $\xi < 1$. However, our considerations include scenarios in which the dark photon can decay into Standard Model particles through the kinetic mixing introduced in Sec. 6.1; hence we consider both cases with $\xi < 1$ and $\xi > 1$. A more thorough discussion of the fate of the dark sector photons and scalar bosons is given in Sec. 6.2.4 below.

If other processes bring the dark sector and visible sector into equilibrium, we will again have $\xi = 1$ at later times. The only such process in this minimal model is the kinetic mixing between the dark photon and the hypercharge boson. The energy transfer between the two sectors occurs predominantly via scattering of ordinary electrons on the lightest charged species of the dark sector, which is either e_D or ϕ_D, with rate

$$\frac{d\rho}{dt} \approx n_e n_d \int \delta E_L v \frac{d\sigma}{d\Omega} d\Omega, \tag{6.18}$$

where $\delta E_L \sim k(1-\cos(\theta))$ with $k \sim 3\sqrt{2}T$ is the longitudinal energy (and momentum) transfer per collision in a plasma at temperature T; here we have assumed we are in the relativistic regime. n_e and n_d are the number densities of (ordinary) electrons and the lightest charged species of the dark sector. The momentum-transfer cross section is

$$\int d\Omega (1-\cos(\theta)) \frac{d\sigma}{d\Omega} \approx \frac{4\pi\epsilon^2 \alpha_D \alpha_{EM} \mu_{ed}^2}{k^4} \ln\left[\csc(\theta_{\min}/2)\right], \tag{6.19}$$

where μ_{ed} is the reduced mass of the scattering particles (an electron and either ϕ_D or e_D). The mass of the Higgs boson, m_ϕ, is a temperature-dependent quantity; we are considering the situation before $\mathrm{U_D}(1)$-symmetry breaking. After the dark phase transition, the scattering of the physical scalar φ_D with ordinary electrons is loop-suppressed, as can be seen from Eq. (6.5). $\theta_{\min}$ is the minimum scattering angle, which can be estimated as $\csc(\theta_{\min}/2) = 1 + (2\lambda_{\mathrm{Debye}} k)^2/(\epsilon^2 \alpha_{EM} \alpha_D)$, where $\lambda_{\mathrm{Debye}} = \min(\lambda_V, \lambda_D)$ is the smallest of the Debye screening lengths of the ordinary and dark plasma, with $\lambda_D = T/[4\pi \max(n_e \alpha_{EM}, n_D \alpha_D)]$. This scattering process does not bring the sectors into equilibrium if

$$\frac{1}{\rho}\frac{d\rho}{dt} < H, \tag{6.20}$$

where H is the Hubble parameter and $\rho \sim (\pi^2/30) g_* T^4$ is the energy density of either sector. For relativistic number densities of charged particles in the two sectors, $(1/\rho) d\rho/dt \propto 1/T$; due to the long-range nature of the interaction, the energy exchange rate becomes larger as the temperature drops. Consequently, this condition should be evaluated at the latest time when both sectors have significant number densities of relativistic charged particles, at $T \sim \max[m_{e,\mathrm{SM}}/3, \min(m_e/3, v_D)] \sim \max[m_{e,\mathrm{SM}}, m_d]/3$, where we have defined $m_d = \min(m_\phi, m_e)$. We have used $T_{\mathrm{PT}} \sim v_D$ as the latest time the ϕ_D particles can participated in the energy exchange between the two sectors. After one of the species becomes non-relativistic, the energy transfer is further suppressed. This yields roughly the condition

$$\epsilon^2 \alpha_D \lesssim 10^{-20} \left(\frac{\max(m_{e,\mathrm{SM}}, m_d)^3}{m_{e,\mathrm{SM}} \min(m_{e,\mathrm{SM}}, m_d)^2} \right), \tag{6.21}$$

where we have approximated $\ln[\csc(\theta_{\min}/2)] \sim 20$. Obviously, the term in the parentheses is greater than 1 and the bound becomes more relaxed the heavier the dark-sector charged particles are.

We emphasize that it is improper to consider this a constraint on the model; if satisfied, it simply allows for $\xi \neq 1$. Furthermore, this is a sufficient but not necessary condition; the sectors would also be hindered from reaching equilibrium if the dark photons decay or become non-relativistic before this process would bring the sectors into equilibrium.

6.2.2 Efficient Annihilation of Dark Matter

To develop an significantly asymmetric dark matter population, it is necessary that the processes $p_D \bar{p}_D \to \gamma_D \gamma_D$ and $e_D \bar{e}_D \to \gamma_D \gamma_D$ annihilate away the symmetric dark matter components; the cross section for this process is

$$(\sigma v)_{\mathrm{ann}} = \frac{\pi \alpha_D^2}{m_i^2} S, \tag{6.22}$$

where $m_i = m_p, m_e$ and S is the Sommerfeld enhancement factor, which is relevant only at large couplings α_D. (Note that a similar cross section holds for $p_D \bar{p}_D \to e_D \bar{e}_D$.)

The self-interaction cross section for thermal relic dark matter is constrained by the necessity of producing the correct abundance at freeze-out; specifically, $(\sigma v)_{\rm sym} \approx \tilde{\xi}_{\rm ann} \cdot 6 \cdot 10^{-26}$ cm^3/s, where $\tilde{\xi}_{\rm ann} = 1$ if the dark sector dominates the energy density of the universe when annihilations freeze out, and $\tilde{\xi}_{\rm ann} = \xi_{\rm ann} = (T_D/T_V)_{\rm ann}$ if the ordinary-sector energy density dominates [125]. (This is for non-self-conjugate dark matter, and considering only s-wave annihilation.) For symmetric dark matter, this fixes the coupling as a function of the dark matter mass m_{DM}:

$$\alpha_{D,\rm sym}(m_{DM}) = (4 \cdot 10^{-3}) \left(\frac{\tilde{\xi}_{\rm ann}}{S_{\rm sym}}\right)^{1/2} \left(\frac{m_{DM}}{10^2 \text{ GeV}}\right). \tag{6.23}$$

The Sommerfeld enhancement factor $S_{\rm sym}$ becomes important for $m_{DM} \gtrsim 800$ GeV (see e.g. [135]). Asymmetric dark matter models require efficient annihilation in the early universe to eliminate the symmetric component of dark matter, leaving only the asymmetric component. Thus, if one parametrizes the annihilation cross section as $(\sigma v)_{\rm ann} = f \times (\sigma v)_{\rm sym}$, then one must require $f > 1.4$ for dark antiparticles to be less than 10 percent of the dark matter density; for antiparticles to contribute less than 1 percent, one must have $f > 2.4$ [168]. Thus, instead of an equality, we have a minimum value of α_D set by the dark proton mass. Taking into account the two annihilation channels and taking $f = 1.4$, the condition for efficient annihilation is

$$\alpha_D > \alpha_{D,\rm min} = (3.4 \cdot 10^{-3}) \left(\frac{\tilde{\xi}_{\rm ann}}{S_{\rm sym}}\right)^{1/2} \left(\frac{m_p}{10^2 \text{ GeV}}\right). \tag{6.24}$$

This constraint can be relaxed if more annihilation channels exist, although this would require a more complicated dark sector.

The annihilation of the dark fermions freezes-out around $T_{D,\rm fo} \sim m_i/x_{\rm fo}$, with [167–169]

$$x_{\rm fo} \approx 30 + \ln\left(\frac{m_i}{100 \text{ GeV}} \frac{\tilde{\xi}^2_{\rm ann}(\sigma v)_{\rm ann}}{10^{-24} \text{ cm}^3/\text{s}}\right). \tag{6.25}$$

This typically occurs before dark recombination.

6.2.3 Dark Recombination and Residual Ionization Fraction

Once the temperature of the dark sector drops below the binding energy of dark atoms, it becomes energetically favorable for the dark ions to form atoms. There are three regimes describing this dark recombination process [175]. For large α_D or small masses, the recombination process is quite efficient and can be described well by the Saha equation until this process freezes out. On the other hand, for small α_D or large masses, recombination is very weak and most dark matter remains ionized. In the intermediate regime, recombination occurs in quasi-equilibrium and the details of atomic transport are important. (Dark matter also remains ionized if the dark photon mass is greater than the binding energy.) The residual ionization fraction can be approximated by [175]

$$x_D \approx \begin{cases} \min\left[1, 10^{-10}\dfrac{\tilde{\xi}_{\rm DR}}{\alpha_D^4}\left(\dfrac{m_H \mu_D}{\rm GeV^2}\right)\right], & M_D < \dfrac{\alpha_D^2 \mu_D}{2} \\ 1, & M_D > \dfrac{\alpha_D^2 \mu_D}{2}, \end{cases} \tag{6.26}$$

where $\tilde{\xi}_{\rm DR} = \min[1, \xi_{\rm DR}]$ and $\xi_{\rm DR} = (T_D/T_V)_{\rm DR}$ is the temperature ratio at dark recombination. This equation is less than satisfactory when $x_D \lesssim 1$; however, due to the α_D^4 dependency, the region near but below 1 is only a small slice of parameter space. Using thermodynamic equilibrium equations, we can estimate the temperature at which the recombination process freezes out, $T_{D,\rm fo} = \Delta/x_{\rm rec,fo}$, where

$$x_{\rm rec,fo} \approx 53 + \ln\left[\tilde{\xi}_{\rm DR}\left(\frac{\alpha_D}{0.1}\right)^5\left(\frac{(10\text{ GeV})^3}{m_H \mu_D^2}\right)\right]. \tag{6.27}$$

This temperature may be either above or below the temperature of the $\rm U_D(1)$-breaking phase transition. If it is above the temperature of this phase transition, everything proceeds as above. Even if the $\rm U_D(1)$-symmetry is broken, we expect Eq. (6.26) and Eq. (6.27) to be reasonable approximations since they are based primarily on equilibrium thermodynamics, which is nearly unchanged for a small photon mass.

The dark atoms produced by recombination may bind into molecules, particularly H_D^2. While the process $H_D + H_D \to H_D^2 + \gamma_D$ is slow, it can be catalyzed by the presence

of a small ionized component, through the process $H_D + p_D \rightarrow H_D^{2+} + \gamma_D$ followed by $H_D^{2+} + H_D \rightarrow H_D^2 + p_D$. Because the binding energy of dark molecules is (slightly) smaller than the binding energy of dark atoms, this process requires a slightly tighter condition on the mass of the dark photon. In the following, we neglect the possibility of dark molecule formation, which merits a dedicated study. Some discussion of the scattering properties of dark molecules is given in [181].

6.2.4 The Dark Phase Transition and Late-Time Dark Electron Abundance

We have determined the condition to generate an asymmetry in both dark electrons and dark protons, Eq. (6.10), and we have noted that if dark ions remain until late times, they may form dark atoms. The dark protons are, by assumption, the lightest particles that carry the conserved dark baryon number; consequently, they cannot decay and their abundance must remain. However, after the $\mathrm{U_D}(1)$-breaking phase transition, the dark electron no longer carries any conserved quantum numbers. Thus, it is possible for this asymmetry to be washed out and the dark electron abundance to be diminished by $e_D\bar{e}_D$ annihilations.

The survival of a significant abundance of dark electrons depends particularly on q_ϕ, the charge of Higgs field. For specific values of q_ϕ, dark electrons may acquire a Majorana mass or Majorana-type mass mixing with other species. As a specific example, we consider $q_\phi = 2$, which allows for the coupling

$$\mathcal{L} \supseteq -\frac{y_L}{2}\phi_D \bar{e}^c_{D,L} e_{D,L} - \frac{y_R}{2}\phi_D \bar{e}^c_{D,R} e_{D,R} + h.c., \tag{6.28}$$

where $e_{D,L}$ and $e_{D,R}$ are the left- and right-chirality components of e_D, and y_L and y_R are dimensionless Yukawa couplings. (We recall that e_D is a singlet under the Standard Model gauge group and so the L and R indices do not imply any $\mathrm{SU_L}(2)$ charge assignments.) These terms give a Majorana mass to dark electrons after $\mathrm{U_D}(1)$ is broken, which allows oscillations between e_D and $\bar{e}_D$, potentially washing out the observed asymmetry. As e_D is a singlet under the Standard Model gauge group, it may also have a normal Dirac mass

term without breaking $\mathrm{SU_L}(2)$. If the Majorana contribution is small with respect to the Dirac mass (that is, $|y_L - y_R|v_D/m_e \ll 1$), the mass eigenstates are approximately the self-conjugate fields

$$e_{D,1} \approx \frac{i}{\sqrt{2}}\left(e_{D,L} + e^c_{D,L} - e_{D,R} - e^c_{D,R}\right), \qquad e_{D,2} \approx \frac{1}{\sqrt{2}}\left(e_{D,L} + e^c_{D,L} + e_{D,R} + e^c_{D,R}\right), \tag{6.29}$$

with masses $m_{1,2} = m_e \mp y v_D$ where $y = (y_L + y_R)/2$. These eigenstates have the interactions

$$\begin{aligned}\mathcal{L}_e = &\frac{1}{2}\bar{e}_{D,1} i\not{\partial} e_{D,1} + \frac{1}{2}\bar{e}_{D,2} i\not{\partial} e_{D,2} - \frac{1}{2}(m_e - y v_D)\bar{e}_{D,1} e_{D,1} - \frac{1}{2}(m_e + y v_D)\bar{e}_{D,2} e_{D,2} \\ &- \frac{y}{2\sqrt{2}}\varphi_D\left(\bar{e}_{D,2} e_{D,2} - \bar{e}_{D,1} e_{D,1}\right) + \frac{i q_e g}{2} A^\mu_D \bar{e}_{D,1}\gamma_\mu e_{D,2} + h.c.,\end{aligned} \tag{6.30}$$

where we have set $\phi_D = (v_D + \varphi_D/\sqrt{2})\exp(i\theta)$. As this is a case of near maximal mixing, $e_D - e^c_D$ oscillations are rapid and may erase the $\mathrm{U_D}(1)$ asymmetry carried by the dark electrons completely. (For typical parameters, $H \lesssim \omega_{\mathrm{osc}}$ is easily satisfied.)

However, the formation of $p_D e_D$ bound states can severely hinder the $e_D - \bar{e}_D$ oscillations. The oscillation of a dark electron bound in a dark atom is energetically forbidden if the energy gain from oscillation is insufficient to unbind the dark electron. The energy difference between the mass eigenstates is $2 y v_D$, while the expectation values of the kinetic and potential energies of a dark atom are $\langle E_K\rangle = \Delta$, $\langle E_P\rangle = -2\Delta$ and the total energy is $-\Delta$. For an $p^+_D \bar{e}^+_D$ bound state with the same wavefunction (or the same superposition of plane waves), $\langle E_K\rangle = \Delta$, $\langle E_P\rangle = 2\Delta$, giving a total energy of 3Δ. Thus if $2 y v_D < 4\Delta$, dark atoms are energetically stable. This sets an upper bound on the Yukawa coupling

$$y \lesssim 20\alpha_D^{1/2}(\Delta/M_D), \tag{6.31}$$

where we have take $q_\phi = 2$. The right hand side cannot be arbitrarily small; for bound state formation we must have $\Delta/M_D > 1$ and requiring sufficient annihilation in the early universe sets a lower limit on α_D. Of course, for bound states to hinder wash-out due to oscillation, dark recombination must occur before the dark phase transition which generates the e_D Majorana mass term. Equivalently, this requires $T_{D,\mathrm{PT}} < \Delta/x_{\mathrm{rec,fo}}$, or

$$M_D < (32\pi\alpha_D)^{1/2}\,\Delta/x_{\mathrm{rec,fo}}. \tag{6.32}$$

When this is satisfied, dark matter today remains at least partially atomic, even if the asymmetry in free dark electrons is washed out and they self-annihilate. Note that this condition is stronger than Eq. 6.15, the condition to form bound states.

If dark atoms have not formed before $e_D - \bar{e}_D$ oscillations begin, then these oscillations can erase the $\mathrm{U_D}(1)$ asymmetry carried by the dark electrons. However, they cannot alter the total abundance of dark electrons and anti-electrons if $e_D\bar{e}_D$ annihilations are inefficient when the oscillations regenerate the $\bar{e}_D$ population; then the abundance of e_D ions can only decrease by a factor of at most 1/2. In this scenario, the abundance of dark electrons after the freeze-out of annihilations and before oscillations occur is

$$Y_{e,\mathrm{fo}} = n_{e,\mathrm{fo}}/s = n_{p,\,\mathrm{fo}}/s = \Omega_{DM}\rho_c/(m_p s_0), \tag{6.33}$$

where ρ_c is the critical energy density and s_0 the critical entropy density today. Annihilations are inefficient if $\Gamma_{\mathrm{ann}} < H_{\mathrm{PT}}$, where $\Gamma_{\mathrm{ann}} = s_{\mathrm{PT}} Y_{e,\mathrm{fo}} (\sigma v)_{e,\mathrm{ann}}$ and $H_{\mathrm{PT}} \approx 1.66\sqrt{g_{\mathrm{PT}}} T_{\mathrm{PT}}^2/m_{Pl}$. These are the annihilation rate and Hubble parameter at the time of the dark phase transition; also g_{PT}, $T_{\mathrm{PT}} = T_{D,\mathrm{PT}}/\bar{\xi}_{\mathrm{PT}}$ and $s_{\mathrm{PT}} = (2\pi^2/45) g_{\mathrm{PT}} T_{\mathrm{PT}}$ are the number of effective relativistic degrees of freedom, the temperature, and the entropy density at the same time. $(\sigma v)_{e,\mathrm{ann}}$ is the self-annihilation cross section times relative velocity of dark electrons, given by Eq. (6.22). Substituting these expressions and using condition (6.10) yields

$$M_D \lesssim 10^{-11} \tilde{\xi}_{\mathrm{PT}} (32\pi\alpha_D)^{1/2} \alpha_D^{-2} m_p m_e^2/\mathrm{GeV}^2. \tag{6.34}$$

If this is satisfied, then the relic e_D abundance changes by at most a factor of 2. For a more detailed treatment of the coupled effect of oscillations, annihilations and scatterings, see Ref. [182].

In Fig. 6.2, we plot conditions (6.32) and (6.34), which shows the region of parameter space where dark matter at late times is necessarily multi-component, either because dark atoms are present along with dark protons or a significant density of dark electrons remain. We see that generally the region where Eq. (6.34) is satisfied is within the region where Eq. (6.32)

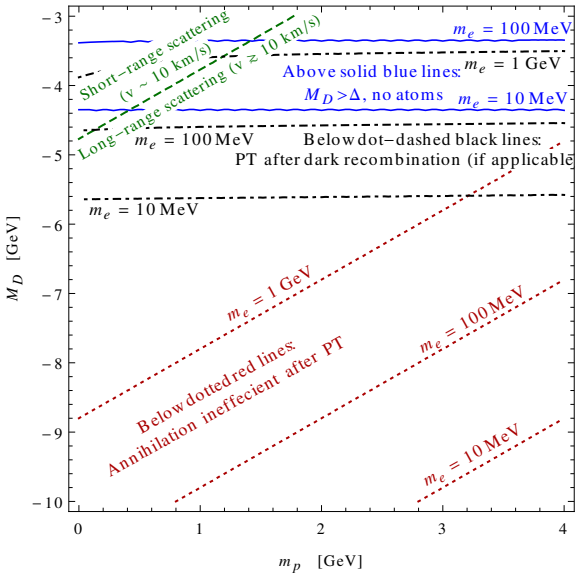

Figure 6.2: Condition (6.2) is satisfied beneath the black lines; then dark matter includes dark atoms. Condition (6.34) is satisfied beneath the red lines; then the dark electrons abundance remains significant even without the formation of dark atoms. (We have set $\alpha_D = 0.1$ and $q_\phi = 2$).

is satisfied. Although the parameter space in which dark matter is multi-component is more limited than the general case without Majorana masses (Fig. 6.1), a significant portion of the region with long-range scattering will be multi-component for light dark electron masses. (The entire region shown in the plot is in the regime in which the asymmetry is generated in both dark electrons and dark protons.)

As seen in this figure, the range of dark photon masses encompassed by either condition (6.32) or (6.34) decreases as m_e decreases. However, the dark electrons cannot be arbitrarily light without implications. If $m_e < M_D$, then the annihilation of non-relativistic electrons in dark photons is forbidden. Then dark electrons decouple while relativistic, and their relic density is generically large, which influences later cosmology. Consequently, there would be significant cosmological bounds on the temperature ratio between the sectors. (It is possible that additional annihilation channels allow the electrons to decay into lighter degrees

of freedom, but similar reasoning would constrain the cosmological abundances of these particles; additionally, this would be a less minimal model.) Therefore, even if Majorana masses are generated, a significant region of parameter space remains in which dark matter is multi-component.

Again, though, we emphasize that these conditions only apply for a limited range of q_ϕ values, namely those for which a Majorana mass or Majorana-type mixing is generated for the dark electrons. In a generic model, this is not the case and the only condition for dark matter to be multi-component at late times is that the asymmetry be generated in both dark electrons and dark protons (Eq. (6.10)). The absence of these terms corresponds to the conservation of a global U(1) remnant symmetry under which e_D is charged. In analogy to the Standard Model, we will call this symmetry dark lepton number L_D. By a linear transformation, we can define the quantum numbers of the dark ions to be $B_D(p_D) = 1$, $B_D(e_D) = 0$ and $L_D(p_D) = -L_D(e_D) = 1$. In this scenario, the high-energy processes which generate the B_D asymmetry conserve L_D, and equal asymmetries in p_D and e_D are generated. As the dark electron is the lightest particle charged under L_D, it is stable and thus its asymmetry is preserved.

If condition (6.10) is satisfied, and either Eq. (6.32) or Eq. (6.34) if necessary, then dark matter is multi-component at late times. Consequently, considering only dark proton collisions in halos does not accurately represent the dynamics of dark matter self-interaction; one should include the different intra- and inter-species collisions. We give such an analysis in Sec. 6.3.

6.2.5 The Fate of Dark Photons

Although dark protons and dark electrons decouple from the dark photons at the end of dark recombination, the dark photons and the scalar field φ_D remain chemically coupled via the annihilations $\varphi_D^* \varphi_D \leftrightarrow \gamma_D \gamma_D$ until one of these species becomes non-relativistic. We will

assume that φ_D is heavier than γ_D; this is true for $\lambda > 2\pi q_\phi^2 \alpha_D$. This leads to constraints on the abundance of dark photons after the chemical decoupling. It is straightforward to reverse this assumption; then the constraints apply to the φ_D abundance instead.

The $\mathrm{U_D}(1)$-breaking phase transition takes place at a temperature $T = T_{D,\mathrm{PT}} \sim v_D$. Immediately after the phase transition, γ_D and φ_D are moving at least quasi-relativistically, if not relativistically. The heavier φ_D bosons become non-relativistic at $T_D \lesssim m_\varphi/3$. The dark photons become non-relativistic at

$$T_D \lesssim \frac{M_D}{3} = \frac{m_\varphi \sqrt{2\pi q_\phi^2 \alpha_D/\lambda}}{3}. \tag{6.35}$$

Over a substantial region of parameter space M_D is within one order of magnitude of m_φ. Therefore, we expect the dark photons to become non-relativistic shortly after the φ_D bosons do so. The annihilation cross section for $\varphi_D\varphi_D \to \gamma_D\gamma_D$ is

$$(\sigma v)_{\varphi_D\varphi_D \leftrightarrow \gamma_D\gamma_D} \approx \frac{44\pi q_\phi^4 \alpha_D^2}{m_\varphi^2} \left(1 - \frac{20 M_D^2}{11 m_\varphi^2} + \frac{12}{11}\frac{M_D^4}{m_\varphi^4}\right) \left(1 - \frac{M_D^2}{m_\varphi^2}\right)^{1/2}, \tag{6.36}$$

The φ_D bosons freeze-out with an abundance

$$x_{\varphi,\mathrm{fo}} \equiv m_\varphi/T_D \approx 41 + \ln\left[q_\phi^4 \alpha_D^2 \bar{\xi}^2 (\mathrm{GeV}/m_\varphi)\right]. \tag{6.37}$$

In the parameter space of interest, $m_\varphi \ll m_p$, and the φ_D annihilation cross section is comparatively large. Consequently, the relic abundance is small in comparison to that of the dark protons, and so it is cosmologically insignificant. Additionally, if $m_\varphi > 2M_D$, then the φ_D bosons decay rapidly into dark photons after they decouple; if $m_D + 1.022\ \mathrm{MeV} < m_\varphi < 2M_D$, they may decay into $\varphi_D \to \gamma_D e^+ e^-$ via a virtual dark photon and its kinetic mixing with hypercharge.

The $\varphi_D\gamma_D$ chemical decoupling also determines the abundance of dark photons. During decoupling the temperature is $T_{D,\gamma_D} = m_\varphi/x_{\varphi,\mathrm{fo}}$, and therefore

$$x_{\gamma_D} \equiv \frac{M_D}{T_{D,\gamma_D}} = \frac{x_{\varphi,\mathrm{fo}} M_D}{m_\varphi} \sim x_{\varphi,\mathrm{fo}} \left(\frac{2\pi q_\phi^2 \alpha_D}{\lambda}\right)^{1/2}. \tag{6.38}$$

Once decoupled, the dark photons may decay into Standard Model charged fermions via their hypercharge mixing (Eq. 6.6) at the rate given by Eq. (6.7) if $M_D > 1.022$ MeV. For smaller masses, the dark photons may be long-lived (although they can still decay into neutrinos, or to three photons via a loop diagram). Either case has potential implications for cosmology. The decay of the dark photon abundance into Standard Model particles injects relativistic energy density, which can affect Big Bang Nucleosynthesis (BBN) and the Cosmic Microwave Background Radiation (CMB). On the other hand, if a significant relic abundance survives until very late, it can affect the time of matter-radiation equality or contribute to the matter density of the universe. Thus, we must require either that the dark photons decay before BBN or that their energy density be sufficient small (to be specified below).

Decay Before BBN: A basic requirement for this scenario is that the dark photons must have acquired their mass before BBN (that is, the dark phase transition must occur before BBN): $T_{V,\mathrm{PT}} > T_{V,\mathrm{BBN}} \sim 3$ MeV. This occurs if

$$\xi_{\mathrm{PT}} < 6.6 \left(\frac{10^{-2}}{q_\phi^2 \alpha_D}\right)^{1/2} \left(\frac{M_D}{10\ \mathrm{MeV}}\right). \tag{6.39}$$

Requiring $T_{V,\mathrm{decay}} > T_{V,\mathrm{BBN}}$ gives the condition

$$\epsilon > \frac{10^{-10}}{f^{1/2}} \left(\frac{10\ \mathrm{MeV}}{M_D}\right)^{1/2}. \tag{6.40}$$

All of these must be satisfied for dark photons to decay before BBN.

Survive through BBN: If either of the above inequalities does not hold, or $M_D <$ 1.022 MeV, then dark protons survive through BBN. If they are relativistic at this time, observations require $\xi_{\mathrm{BBN}} \lesssim 0.6$; this corresponds to the relativistic energy density of one extra neutrino species, as allowed by current data [183]. Dark photons will be relativistic during BBN if $M_D < 3\xi_{\mathrm{BBN}} T_{V,\mathrm{BBN}} \approx \xi_{\mathrm{BBN}} \cdot 9$ MeV; if the dark photons are heavier than this, they are non-relativistic during BBN and no constraint is set on ξ_{BBN}.

Furthermore, we must require that the abundance of dark photons does not alter the time of matter-radiation equality, and if dark photons remain until today, they must be a sub-dominant component of the dark matter. The number density of dark photons, normalized with respect to the entropy density s, is

$$Y \equiv \frac{n_{\gamma_D}}{s} \approx \frac{\xi_{\gamma_D}^3}{g_{*,S}(t_{\gamma_D})} h(x_{\gamma_D}), \tag{6.41}$$

where $g_{*,S}(t_{\gamma_D})$ is the effective number of relativistic degrees of freedom at the time of dark photon decoupling, and x_{γ_D} is given by Eq. 6.38. The function $h(x)$ is

$$h(x) = \frac{135}{4\pi^4} \int_x^\infty dy \frac{y\sqrt{y^2 - x^2}}{e^x - 1} \approx \begin{cases} 0.8, & x < 3, \\ 0.4 x^{3/2} e^{-x}, & x \gtrsim 3. \end{cases} \tag{6.42}$$

Depending on when (and if) the dark photons become non-relativistic, there are three possible cases:

1. Dark photons alter the time of matter-radiation equality if they become non-relativistic and dominate the energy density of the universe at some temperature $T_{V,\mathrm{dom}} \geq T_{V,\mathrm{eq}}$, where $T_{V,\mathrm{eq}} \sim 5$ eV is the temperature of matter-radiation equality. In this case, when $T_V = T_{V,\mathrm{dom}}$, $sY(\gamma_D)M_D \approx \rho_U = (\pi^2/90) g_* T_V^4$, where ρ_U is the energy density of the universe. This gives $T_{V,\mathrm{dom}} \approx 4Y(\gamma_D) m_D$. Putting together the constraints $M_D / T_{V,\mathrm{dom}} > 3$ and $T_{V,\mathrm{dom}} \geq T_{V,\mathrm{eq}}$ yields

   $$\frac{1.25 \text{ eV}}{M_D} < \frac{\xi_{\gamma_D}^3 h(x_{\gamma_D})}{g_{*,S}(t_{\gamma_D})} < 0.08. \tag{6.43}$$

   In this case we must require that photons decay before they would dominate, or $T_{V,\mathrm{decay}} > T_{V,\mathrm{dom}}$, which requires

   $$\epsilon > 6 \cdot 10^{-9} f^{-1/2} \frac{\xi_{\gamma_D}^3}{g_{*,S}(t_{\gamma_D})} \left(\frac{M_D}{10 \text{ MeV}} \right)^{1/2} h(x_{\gamma_D}), \tag{6.44}$$

   in addition to $M_D > 1.022$ MeV.

2. If Eq. (6.43) is not satisfied, then the dark photons do not dominate the energy density of the universe before matter-radiation equality. We need only require that if the dark

photons have become non-relativistic today, their relic abundance is not the dominant component of the dark matter. Their contribution to the matter density of the universe is $\Omega(\gamma_D) = s_0 Y(\gamma_D) M_D/\rho_c$, and to have $\Omega(\gamma_D) < 0.01$ requires

$$\xi_{\gamma_D} < 0.02 \left(\frac{g_{*,S}(t_{\gamma_D})}{10}\right)^{1/3} \left(\frac{100 \text{ keV}}{M_D}\right)^{1/3} h(x_{\gamma_D})^{-1/3}, \tag{6.45}$$

which also implies that $T_{V,\text{dom}} < T_{V,\text{eq}}$. This condition holds if

$$M_D > 3\xi_0 T_{V,0} \approx \xi_0 \cdot 7 \cdot 10^{-4} \text{ eV}, \tag{6.46}$$

where the subscript 0 refers to the present.

3. If Eq. 6.46 is not satisfied, then dark photons are still relativistic today, and the only applicable bound is $\xi_{\text{BBN}} \lesssim 0.6$. This case includes the limit of a massless dark photon.

The various constraints are illustrated in Fig. 6.3 for the case $\epsilon \to 0$. If $\epsilon \neq 0$, these constraints can only be relaxed due to the possibility of dark photon decay, although condition (6.21) provides an upper bound. Additionally, ϵ is constrained by experimental results; for a compilation of bounds, see Ref. [184], and for minicharged particles, see Ref. [185]. Bounds on very light dark photons mixing with hypercharge are given in Ref. [186]. Additionally, this kinetic mixing opens direct detection channels, although a study of direct detection requires considering the both the inelastic and elastic scattering of each different species with matter [187–189]. Related studies of direct detection of multi-component dark matter with long-range interactions can be found in Refs. [190–193]. Furthermore, if a significant fraction of dark matter remains ionized today, dark atoms can potentially form in the dense environment of halos [131, 194]. In the previous chapter, we discussed how the decay of the emitted mediator may produce observable indirect detection signals.

6.2.6 Dark-Matter Kinetic Decoupling and Large-Scale Structure

If dark matter remains coupled to a thermal bath of dark photons until later times, this can affect the matter power spectrum and gravitational clustering. In particular, the acoustic

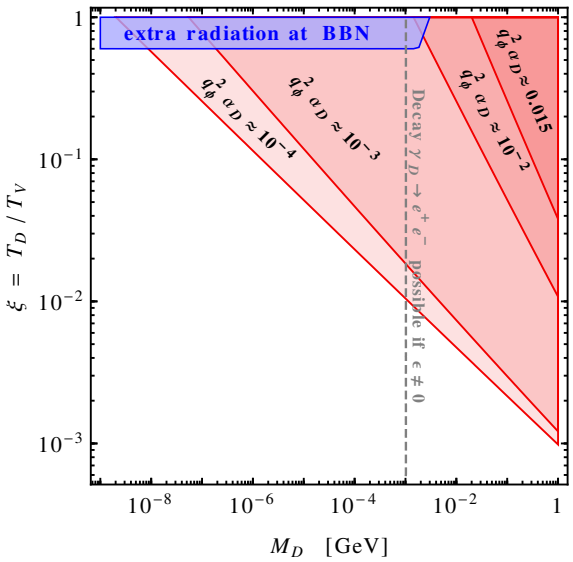

Figure 6.3: Bounds on the ordinary-to-dark temperature ratio, $\xi = T_D/T_V$, assuming $\epsilon \to 0$. In the blue region, the extra radiation due to relativistic dark photons exceeds the BBN limit; this constraint applies to $\xi_{\rm BBN}$. In the red regions, the relic abundance of dark photons may alter the time of matter-radiation equality or dominate the dark matter density. These apply to ξ_{γ_D}. To the right of the grey line, the dark photons may decay into Standard Model fermions, which can relax or eliminate the bounds.

oscillations of the coupled dark matter and dark-radiation system can imprint a new characteristic scale on the matter power spectrum, which can affect the CMB temperature and polarization spectra [124, 175, 195, 196]. Thus, observations of the CMB and galaxy surveys give additional constraints on α_D.

In this model, dark matter and dark radiation remain coupled mostly via Compton scattering of dark photons on dark ions and Rayleigh scattering of dark photons on neutral dark atoms [175]. The scale of dark acoustic oscillations is [196]

$$\Sigma_{\rm DAO} \equiv \alpha_D \left(\frac{\rm eV}{\Delta}\right) \left(\frac{\rm GeV}{m_H}\right)^{1/6}, \tag{6.47}$$

assuming a significant fraction of the dark ions are in bound states. Observations of the CMB and galaxy surveys require $\Sigma_{\rm DAO} \leq 10^{-4.5}$ if the inequality $\xi_{\rm BBN} \lesssim 0.6$ is saturated; lower values of $\xi_{\rm BBN}$ relax this bound [196]. For our model, with a massive dark photon, we note that these constraints are applicable only if the dark photons are relativistic when the CMB is generated; that is, $M_D/T_{D,\rm CMB} \lesssim 3$, or $M_D \lesssim 3\xi_{\rm CMB} T_{V,\rm CMB} \approx \xi_{\rm CMB} \cdot 15$ eV. Thus, these constraints limit only a rather small range of dark photon masses, and we will see that constraints from the ellipticity of large halos are generally stronger.

6.3 Dark Matter Self-Interaction in Halos

As we have noted in the introduction to this chapter, one motivation for considering self-interacting dark matter is that such interactions can modify the dynamics of small halos, alleviating the tension between cold dark matter simulations and observations. For definiteness, we will assume that in the entire parameter space, dark matter is made of equal amounts of p_D^+ and e_D^-, with no relic antiparticles present, and that these dark matter ions may be bound into atoms, with an ionization fraction given by Eq. (6.26). This is inevitable only if condition (6.10) holds (and, if $q_\phi = 2$, also (6.32) or (6.34)); as noted above, these are sufficient but not necessary for all models. As this covers much of our parameter space, we focus on the multi-component case; for a study of halos for a single-component dark matter

coupled to a vector boson, see Refs. [128, 135].

6.3.1 Dark Matter Scattering Rates

To determine the effect of dark matter self-interaction on halo shapes, we first need the scattering rates for the various dark interactions (i.e., ion-ion, ion-atom, atom-atom). We assume that dark matter has a velocity distribution which is locally Maxwellian

$$f(v,\bar{v}) = \left(\frac{3}{2\pi\bar{v}^2}\right)^{3/2} e^{-3v^2/2\bar{v}^2}, \tag{6.48}$$

where $\bar{v} = \bar{v}(r)$ is the rms average velocity, which is generally a function of position inside the halo. Let Γ_p, Γ_e, and Γ_H be the average rates of momentum-changing collisions for dark protons, dark electrons, and dark hydrogen atoms respectively; each of these rates includes contributions from scattering with each other species,

$$\begin{aligned}
\Gamma_p &= \Gamma_{pp} + \Gamma_{pe} + \Gamma_{pH}, \\
\Gamma_e &= \Gamma_{ep} + \Gamma_{ee} + \Gamma_{eH}, \\
\Gamma_H &= \Gamma_{Hp} + \Gamma_{He} + \Gamma_{HH}.
\end{aligned} \tag{6.49}$$

Γ_{ij} is the average momentum-loss rate from species i to species j [175],

$$\Gamma_{ij} = \frac{\dot{p}_{ij}}{\bar{p}_i}, \tag{6.50}$$

where $\bar{p}_i = m_i\bar{v}$ is the average rms momentum of species i and $\bar{p}_{ij}$ is the average momentum-loss rate of species i due to collisions with species j in the halo. We estimate this as

$$\begin{aligned}
\dot{p}_{ij} &= n_j(r) \int d^3v_i\, f(v_i,\bar{v}) \int d^3v_j\, f(v_j,\bar{v}) |\boldsymbol{v}_i - \boldsymbol{v}_j| \int d\Omega\, \frac{d\sigma_{ij}}{d\Omega} \delta p_{ij}, \\
&= n_j(r) \int d^3v\, f(v,\bar{v}_{\rm rel}) v \int d\Omega\, \frac{d\sigma_{ij}}{d\Omega} \delta p_{ij},
\end{aligned} \tag{6.51}$$

where $n_j(r)$ is the number density of species j, $d\sigma_{ij}/d\Omega$ is the differential cross section for ij scattering, δp_{ij} is the momentum transfer from i to j, and finally $\bar{v}_{\rm rel} = \sqrt{2}\bar{v}$. Both the cross section and momentum transferred depend depend only on the relative velocity $|\boldsymbol{v}_i - \boldsymbol{v}_j|$ and

the scattering angle θ. The momentum transfer is

$$\delta p_{ij} = \frac{m_i m_j}{m_i + m_j} v g(\theta). \tag{6.52}$$

The angular function $g(\theta)$ depends on whether the total, longitudinal, or transverse momentum transfer is considered,

$$\begin{aligned}
g_{\rm tot}(\theta) &= \begin{cases} 2\sin(\theta/2), & i \neq j, \\ 2\min[\sin(\theta/2), \cos(\theta/2)], & i = j, \end{cases} \\
g_{\rm long}(\theta) &= \begin{cases} 1 - \cos(\theta), & i \neq j, \\ \min[1 - \cos(\theta), 1 + \cos(\theta)], & i = j, \end{cases} \\
g_{\rm tr}(\theta) &= |\sin(\theta)|.
\end{aligned} \tag{6.53}$$

The case $i = j$ takes into account that forward and backward scattering of identical particles are equivalent. In recent self-interacting dark matter simulations, the longitudinal momentum transfer is used and distinguishable particles are assumed, with the resulting momentum-transfer cross section defined as

$$\sigma_T = \int d\Omega (1 - \cos(\theta)) \frac{d\sigma}{d\Omega}. \tag{6.54}$$

For the ion-ion collisions, governed by a Yukawa potential, we use existing analytical formulas for σ_T. In the small coupling regime, $r_{ij} \equiv 4\alpha_D \mu_{ij}/M_D \ll 1$ where $\mu_{ij} = m_i m_j/(m_i + m_j)$, the Born approximation is valid and the momentum-transfer cross section is [197–199]

$$\sigma_{T,ij}^{\rm born} = \frac{2\pi\beta_{ij}^2}{M_D^2} \left[\ln\left(1 + \frac{r_{ij}}{\beta_{ij}}\right) - \frac{r_{ij}}{\beta_{ij} + r_{ij}} \right], \tag{6.55}$$

where $\beta_{ij} = \alpha_D M_D / v^2 \mu_{ij}$ and v is the relative velocity of interacting pair. For larger couplings, the Born approximation is inapplicable, but the classical approximation becomes

valid. In this regime, the attractive $e_D p_D$ cross section is

$$\sigma_{T,ep}^{\rm class} \approx \begin{cases} \dfrac{4\pi}{M_D^2}\beta_{ep}^2 \ln\left(1+\dfrac{1}{\beta_{ep}}\right), & \beta_{ep} \lesssim 10^{-1}, \\ \dfrac{8\pi}{M_D^2}\dfrac{\beta_{ep}^2}{1+1.5\beta_{ep}^{1.65}}, & 10^{-1} \lesssim \beta_{ep} \lesssim 10^{3}, \\ \dfrac{0.81\pi}{M_D^2}\left[\ln^2(\beta_{ep}) + 2\ln(\beta_{ep}) + 2.5 + \dfrac{4}{\ln(\beta_{ep})}\right], & \beta_{ep} \gtrsim 10^{3}. \end{cases} \tag{6.56}$$

This classical approximation is valid only if $r_{ep}/\beta_{ep} > 1$; for $1 < r_{ep} < \beta_{ep}$, the scattering cross section exhibits resonances due to the contribution of virtual bound states [128, 135]. These resonances affect only a small part of parameter space, and we will ignore them for simplicity. Hence, we use the classical approximation everywhere outside of the Born approximation's range of validity.

For the repulsive $p_D p_D$ and $e_D e_D$ scattering in the classical regime we use [125, 199]

$$\sigma_{T,ii}^{\rm class} = \begin{cases} \dfrac{2\pi}{M_D^2}\beta_{ii}^2 \ln\left(1+\dfrac{1}{\beta_{ii}^2}\right), & \beta_{ii} \lesssim 1 \\ \dfrac{\pi}{M_D^2}\left(\ln(2\beta_{ii}) - \ln(\ln(2\beta_{ii}))\right)^2, & \beta_{ii} \gtrsim 1. \end{cases} \tag{6.57}$$

While the ion-ion collisions are significantly affected by the nonzero dark photon mass, the atom-atom and atom-ion collisions are less affected. The nonzero mass modifies the interatomic potential at distances $r \gtrsim 1/M_D$, while even in the massless case the interatomic potential is significant only within a Bohr radius $a_0 = (\mu_D \alpha_D)^{-1}$ (See Ref. [181] and references therein.) In the regime in which an asymmetry is generated in dark electrons and dark protons (i.e., which satisfies Eq. (6.10)), the Bohr radius is smaller than $1/M_D$ and so the effects of the finite photon mass are minimal. Consequently, we will use the atom-atom and atom-ion scattering rates determined using a massless dark photon.

However, even in the limit of a massless dark photon, this is currently considerable uncertainty in the existing literature about the atom-atom and atom-ion collision rates. We

present results following two different approaches, developed in [181] and [175]. First, though, we summarize the differences in the approaches.

The authors of Ref. [181] calculated the low-energy atom-atom scattering cross section by the direct computation of the phase shifts induced by interatomic potentials. The atom-atom cross section exhibits a rich resonant structure due to the multiple states of hydrogen-like atoms, but away from the resonance the energy dependence of the transverse energy-transfer cross section can be fit across a wide energy range by

$$\sigma_{T_E} \approx (\mu_D \alpha_D)^{-2} \left[b_0 + b_1 \left(\frac{m_H v^2}{4\mu_D \alpha_D^2} \right) + b_2 \left(\frac{m_H v^2}{4\mu_D \alpha_D^2} \right)^2 \right]^{-1}, \tag{6.58}$$

where the energy-transfer cross section is defined as

$$\sigma_{T_E} \equiv \int d\Omega \, \sin^2(\theta) \frac{d\sigma}{d\Omega}. \tag{6.59}$$

The use of σ_{T_E} rather than σ_T may be well-motivated by physical arguments; however, as simulations use σ_T, this choice makes it difficult to compare with simulations. The parameters b_0, b_1, and b_2 are determined by numerical fits and depend mildly on the ratio m_p/m_e. To cover a continuum range of the ratio m_p/m_e, we interpolate the parameters between the values provided in Ref. [181].

However, Ref. [181] does not address atom-ion scattering; thus we ignore atom-ion collisions when adopting this approach. This is justified because atom-ion collisions are either subdominant or insignificant throughout the entire parameter space. Atomic interactions are screened on the scale of the Bohr radius, while ion-ion interactions are screened on the scale of the smaller photon mass; hence ion-ion interactions dominate over atomic interactions when present (i.e., for $x_D > 0.5$). While the atom-ion cross section may be comparable or larger than the atom-atom cross section, it plays no role if there are few ions present ($x_D \ll 1$). For moderate values of x_D it possible that atom-ion collisions dominate the momentum-transfer rate, but in this fairly limited regime the gauge coupling is typically not

large enough for the interactions to affect the halo dynamics. This assertion is confirmed using the analysis of Ref. [175], described next.

The authors of Ref. [175] estimate both the atom-atom and atom-ion collision rates by rescaling the experimentally measured rates for ordinary atoms and ions. They consider the momentum-transfer cross section and average over a Maxwellian velocity distribution. The results are

$$\begin{aligned}
\Gamma_{HH} &\approx n_H \left[15\pi(4/3)^{3/8}\Gamma(19/8)\right] \left(\frac{\alpha_D \bar{v}^{3/4}}{\Delta^2}\right) \left(\frac{m_{e,\mathrm{SM}} m_H}{\mu_D \Delta}\right)^{-1/8} \left[1 + \frac{m_{e,\mathrm{SM}} m_H}{\mu_D \Delta} \frac{\bar{v}^2}{225}\right]^{-19/8}, \\
\Gamma_{pH} &\approx n_H \frac{30\sqrt{3\pi^3}\alpha_D^2 \bar{v}}{\Delta^2} \frac{(m_H m_p)^{1/2}}{m_H + m_p} \left[1 + \frac{m_{e,\mathrm{SM}} m_p}{(\mu_D + m_p)\Delta} \frac{\bar{v}^2}{150}\right]^{-5/2}, \\
\Gamma_{eH} &\approx n_H \frac{30\sqrt{3\pi^3}\alpha_D^2 \bar{v}}{\Delta^2} \frac{(m_H m_e)^{1/2}}{m_H + m_e} \left[1 + \frac{m_{e,\mathrm{SM}} m_e}{(\mu_D + m_)\Delta} \frac{\bar{v}^2}{150}\right]^{-5/2}, \\
\Gamma_{Hp} &\approx \frac{n_p m_p}{n_H m_H} \Gamma_{pH}, \\
\Gamma_{He} &\approx \frac{n_e m_e}{n_H m_H} \Gamma_{eH},
\end{aligned} \tag{6.60}$$

where $m_{e,\mathrm{SM}} = 511$ keV is the (normal) electron mass. These are valid in the interval $10^{-3} \lesssim E_{cm}/\Delta \lesssim 10$, where $E_{cm} = \mu_{ij} v^2/2$ is the center-of-mass energy of the colliding particles. As noted in [175], this is expected to over-estimate the atom-atom collision rate at low-energies; in fact it diverges as $v \to 0$, in contrast to the approach of [181], which finds that the s-wave contribution dominates. Our numerical analysis shows that [175]'s approach gives significantly larger cross sections than [181]'s approach even in its assumed range of validity (by about an order of magnitude). Part of this difference can be attributed to using the momentum-transfer cross section instead of the energy-transfer cross section.

6.3.2 Effects of Scattering on Halos

We have established that the scenario considered here generically results in multi-component dark matter, with several types of inter- and intra-species interactions. Consequently, dark matter simulations of single-component dark matter are not directly applicable; but we will

use insights from these simulations to determine reasonable conditions to study the impact of dark matter interactions on the dynamics of dark matter halos. We have two complementary goals: first, to determine rough constraints that ensure that scattering in Milky-Way-sized halos does not destroy their observed ellipticity, and secondly, to identify regions of parameter space in which scattering could affect the dynamics of dwarf-galaxy-sized halos, potentially bringing predictions into better agreement with observations.

We define an effective average momentum-transfer rate

$$\Gamma_{\rm eff} = h_p \min\left(\Gamma_p, \Gamma_{\rm crit}/h\right) + h_e \min\left(\Gamma_e, \Gamma_{\rm crit}/h\right) + h_H \min\left(\Gamma_H, \Gamma_{\rm crit}/h\right), \tag{6.61}$$

where h_p, h_e, and h_H are the mass fractions carried by dark protons, dark electrons, and dark atoms respectively,

$$\begin{aligned} h_p &\equiv \frac{x_D m_p}{x_D(m_p+m_e)+(1-x_D)m_H} \approx \frac{x_D m_p}{m_H}, \\ h_e &\equiv \frac{x_D m_e}{x_D(m_p+m_e)+(1-x_D)m_H} \approx \frac{x_D m_e}{m_H}, \\ h_H &\equiv \frac{(1-x_D)m_H}{x_D(m_p+m_e)+(1-x_D)m_H} \approx (1-x_D). \end{aligned} \tag{6.62}$$

$\Gamma_{\rm eff}$ depends on the position in the dark matter halo, through its dependence on the densities and the velocity dispersions of the dark matter species. (The velocity dispersion dependence arises mostly through the strong velocity dependence of the cross sections given above.)

$\Gamma_{\rm crit}$ is an estimate (to be specified below) for the magnitude of the effective momentum-transfer rate above which there is a significant effect on the halo; it is this value that we compare to $\Gamma_{\rm eff}$. It also generically depends on the position inside the halo. In Eq. 6.61, we cap the contribution of each contribution at $\Gamma_{\rm crit}/h$, with $h < 1$. The physical reasoning for this is that if the momentum-transfer rate of a species is large, but the species contributes only a fraction of the mass in the halo, then the effect of the momentum loss by this species on halo dynamics is unimportant. This species' contribution should not drive $\Gamma_{\rm eff}$ above the critical value. Only the interactions of species which carry a significant portion of the

halo mass should determine the dynamics of the halo. This cap is an albeit somewhat crude way of incorporating this; h is the fraction of dark matter which, if strongly interacting, can drive $\Gamma_{\rm eff}$ to its critical value. We somewhat arbitrarily choose $h = 0.5$. (This choice is partly informed by the dynamics of dark matter and ordinary matter in halos. Ordinary matter contributes 15 percent of the mass of the universe and is quite self-interacting and dissipative, but it does not significantly affect the clustering of dark matter at most scales.)

The strongest constraints on dark matter self-interactions arise from the observed ellipticity of halos of the size of the Milky Way or larger; these correspond to distances $r \sim (4 - 50)$ kpc from the center of the halo [200]. Therefore, we evaluate $\Gamma_{\rm eff}$ at $\rho_{DM} = 1$ GeV/cm^3; this is estimated to occur at $r \sim 4.5$ kpc for both an NFW and an isothermal profile. We also set $\bar{v} = 220$ km/s. We define $\Gamma^{\rm MW}_{\rm crit}$ as follows: For the chosen values of the dark matter density and velocity dispersion, and in the limit of single-component dark matter of mass m with a velocity independent scattering cross section, one must require $\sigma_T/m \lesssim 1$ cm^2/g [28]. For single component dark matter, $\Gamma_{\rm eff} = \rho_{DM}(\sigma_T/m)\bar{v}$; thus we set

$$\Gamma^{\rm MW}_{\rm crit} = (1 \text{ GeV/cm}^2)(1 \text{ cm}^2/\text{g})(220 \text{ km/s}) \approx 1.2 \text{ Gyr}^{-1} \approx 17 H_0. \tag{6.63}$$

We require $\Gamma^{\rm MW}_{\rm eff} < \Gamma^{\rm MW}_{\rm crit}$, where $\Gamma^{\rm MW}_{\rm eff}$ is the effective momentum-transfer rate evaluated at the above density and average velocity. This is a reasonable bound for preventing the thermalization and isotropization of the halo.

Secondly, we want to identify regions of parameter space in which self-interaction can affect the dynamics of smaller halos; since the dwarf spheroidal galaxies of the Milky Way are consistent with isothermal isotropic profiles, these considerations set only a lower bound on $\Gamma_{\rm eff}$. We evaluate the momentum-transfer rates at $\rho_{DM} = 0.5$ GeV/cm^3 and $\bar{v} = 10$ km/s. (For a review of the kinematics of dwarf spheroidal galaxies, see Ref. [201].) For single component dark matter, the desired cross section is $\sigma_T/m > 0.5$ cm^2/g [25,27,28,130]. Thus we set

$$\Gamma^{\rm DW}_{\rm crit} = (0.5 \text{ GeV/cm}^3)(0.5 \text{ cm}^2/\text{g})(10 \text{ km/s}) \approx 0.014 \text{ Gyr}^{-1} \approx 0.2 H_0. \tag{6.64}$$

We desire $\Gamma_{\rm eff}^{\rm DW} > \Gamma_{\rm crit}^{\rm DW}$.

Note that the above approach in choosing $\Gamma_{\rm crit}^{\rm MW}$ and $\Gamma_{\rm crit}^{\rm DW}$ renders our bounds independent of the dark matter density at which the momentum-transfer rates are evaluated, and thus can be reasonably compared to estimated constraints from dark matter simulations. Evaluating the momentum-transfer rates requires knowing the spatial distributions of the different species of dark matter; however these depend on the strength of the various interactions and the relative abundances of each species. Ultimately, the clustering of multi-component and self-interacting dark matter must be resolving through simulations; here we will make the simplifying assumption that all species follow the same density profile

$$n_H(r) \approx (1 - x_D)\rho_{DM}(r)/m_H \tag{6.65}$$

$$n_p(r) = n_e(r) \approx x_D \rho_{DM}(r)/m_H. \tag{6.66}$$

We note that we do expect $n_p(r) = n_e(r)$ due to the attractive $p_D e_D$ interaction and intra-species repulsion. However, since the ion-atom and atom-atom interactions are typically much weaker than the ion-ion interaction, it is possible that atoms and ions have separate profiles [174]. However, in the limits $x_D \approx 1$ or $x_D \ll 1$, we do expect all the dark matter species to follow the same profile, determined primarily by the gravitational pull of the dominant species.

6.3.3 Discussion of Results

In this section we present graphs showing the parameter space that satisfies the criteria outlined above. Our first constraint is a strict bound from requiring that Milky-Way-sized halos retain their ellipticity; the region where this fails is colored pink in the plots below. The second, a desiderata, shows the regime in which self-scattering can affect smaller halos; this region is enclosed by a blue line. We also hatch the region where there is insufficient annihilation in the early universe in this simplistic model, although this bound can be loosened by a less minimal dark sector. In the grey region, the self-consistency bound $4\mu_D > m_H + \Delta$

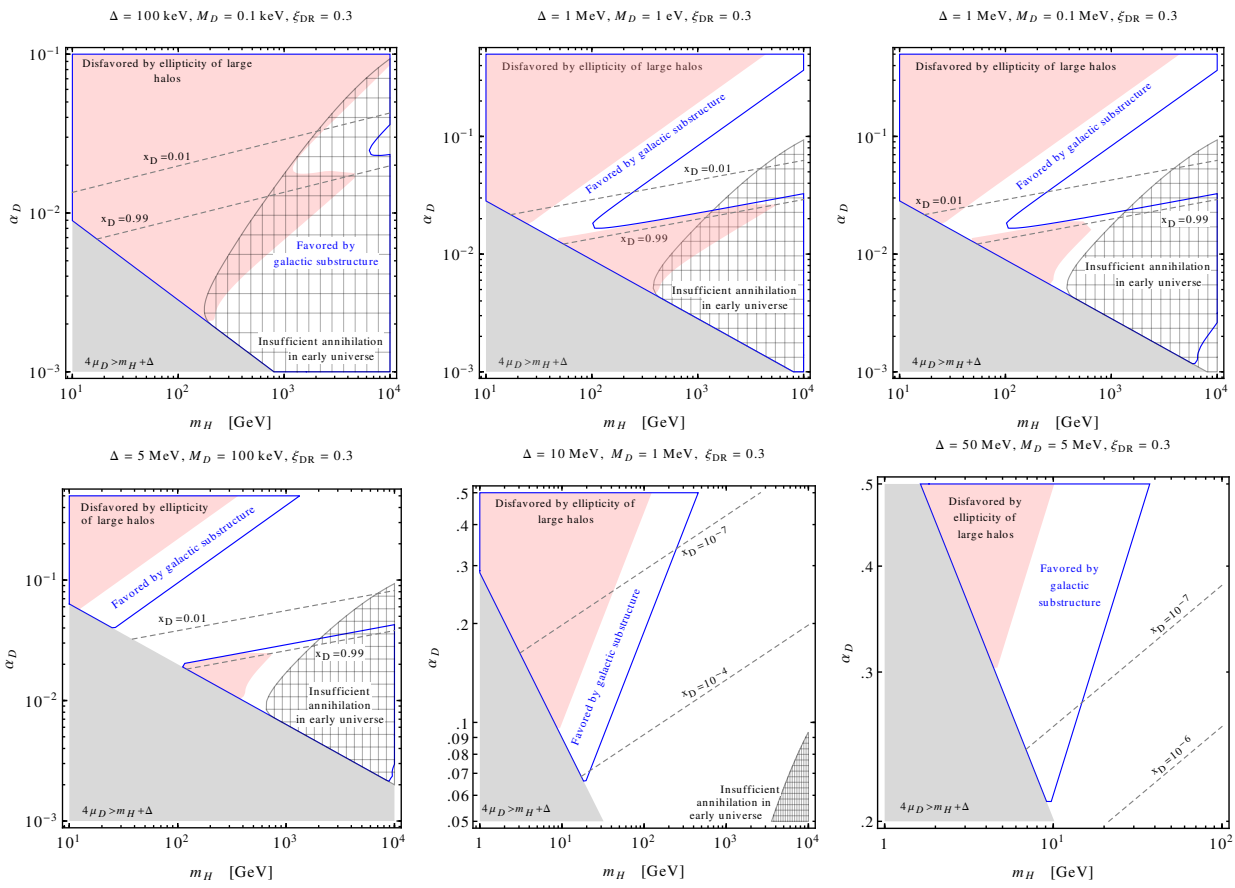

Figure 6.4: Halo bounds as a function of α_D and m_H (fixed Δ and M_D) for $\xi_{\rm DR} = 0.3$, following the methods of [175].

is violated. The broken lines in plots following [181] result from the numerical interpolation of b_0, b_1, and b_2.

There are several different quantities one can use to describe the relevant parameter space. One can describe the dark atoms using the coupling α_D and atomic mass m_H along with either the binding energy Δ or the reduced mass μ_D. The full parameter space is then described by one of these sets of three parameters along with M_D, the mass of the dark photon. The two resulting sets of parameters are useful for different tasks; considering the binding energy is useful in looking for indirect detection signals (as outlined in the previous chapter), although fixing the reduced mass is more helpful when analyzing the cosmological

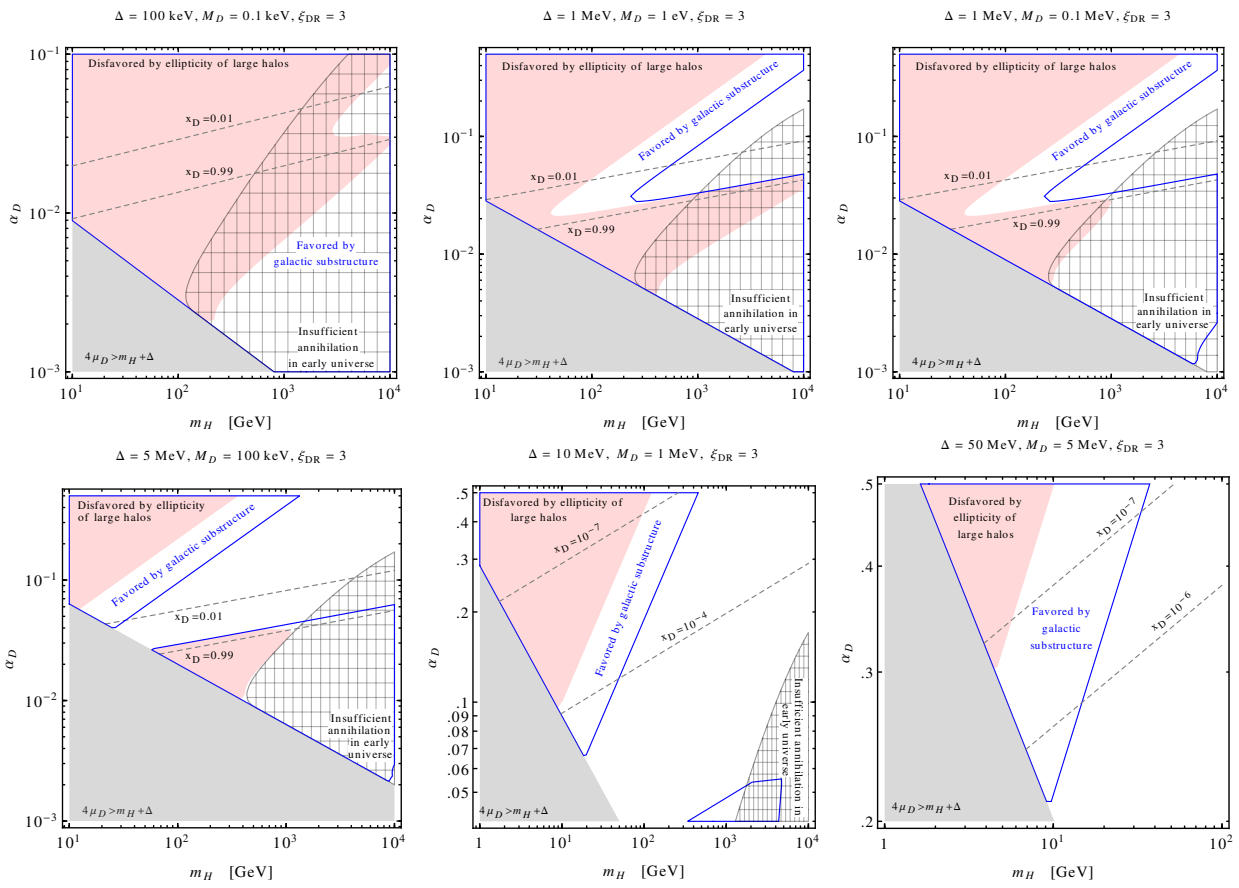

Figure 6.5: Halo bounds as a function of α_D and m_H (fixed Δ and M_D) for $\xi_{\rm DR} = 3$, following the methods of [175].

history. Consequently, we present plots both with the binding energy Δ and the reduced mass μ_D considered as an independent parameter.

First, we look at the effects of scattering on halos as a function of the dark coupling α_D and the dark atom mass m_H; we fix the dark mediator mass. In Figs. 6.4, 6.5, and 6.6, we also fix Δ, while in the following three, Figs. 6.7, 6.8, and 6.9, we fix μ_D. In each of these sets, the first two (i.e., 6.4 and 6.5, 6.7 and 6.8) use the approach of [175]; the remaining two (6.9 and 6.9) follow [181]. Figs. 6.4, 6.5, 6.7, and 6.8 additionally demonstrate the weak $\xi_{\rm DR}$ dependence, as each choice of parameters is evaluated at both $\xi_{\rm DR} = 0.3$ and $\xi_{\rm DR} = 3$.

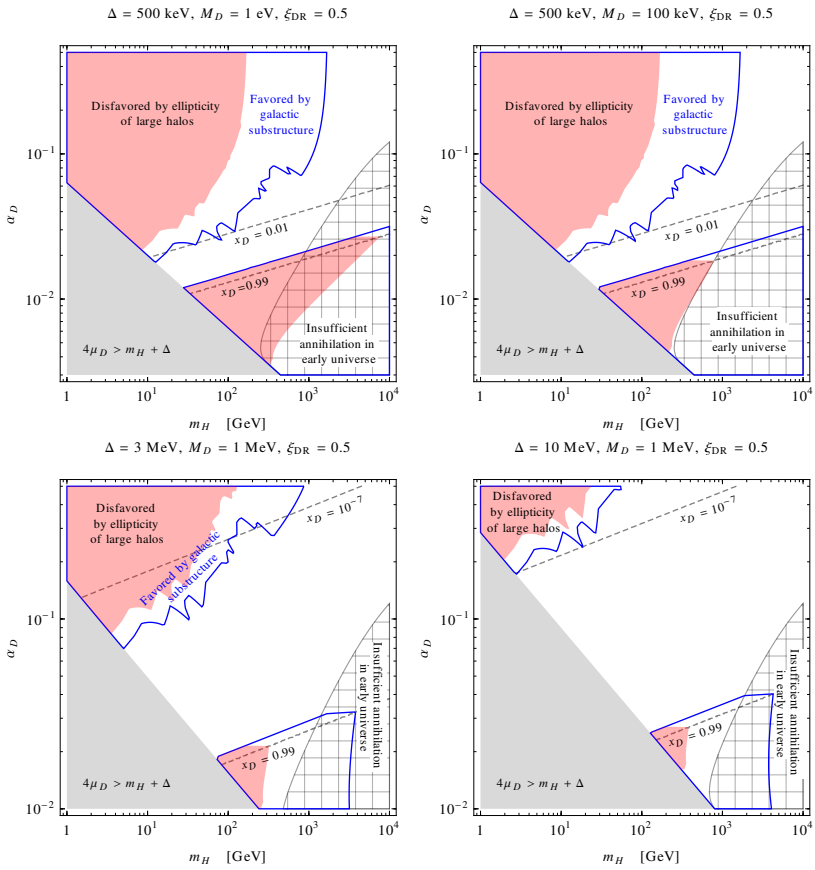

Figure 6.6: Halo bounds as a function of α_D and m_H (fixed Δ and M_D) for $\xi_{\rm DR} = 0.5$, following the methods of [181].

The most pronounced feature of this plot is the "wedge", which demonstrates that the scattering rate varies non-monotonically with α_D and m_H. This can be understood as follows: for small α_D, dark recombination is inefficient and dark matter today consists mostly of ions ($x_D \approx 1$). At very low values of α_D, there is negligible self-interaction among the ions, but increasing α_D increases the scattering rate, which can become sizable before recombination becomes very efficient. At the other extreme, dark atoms are efficiently formed in the early universe, and the atom-atom scattering cross section increases the dark matter self-interaction in the halos, leading to $\Gamma_{\rm eff} > \Gamma_{\rm crit}$. This wedge feature becomes more pronounced as $\xi_{\rm DR}$ is increased. It is also present whether we follow the analysis of [175] or [181]. As may be expected from the above explanation, the precise edges of the wedge, but not its

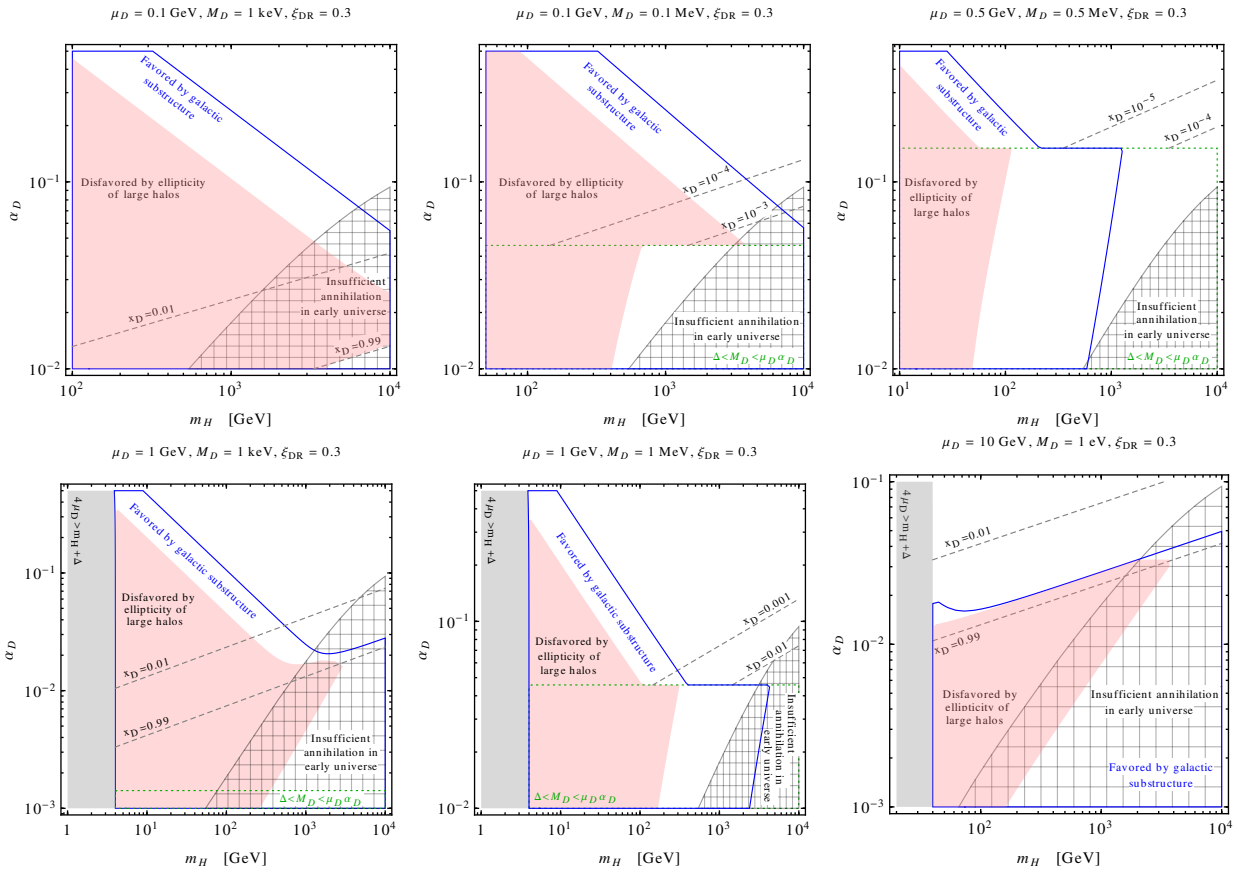

Figure 6.7: Halo bounds as a function of α_D and m_H (fixed μ_D and M_D) for $\xi_{\rm DR} = 0.3$, following the methods of [175].

existence, depend somewhat weakly on the choice of $\Gamma_{\rm crit}$.

We emphasize that this behavior exemplifies the importance of carefully considering the cosmology of models in which dark matter couples to a light mediator. Failing to consider dark recombination would overestimate ion scattering, which would give inaccurate bounds on α_D and m_H.

These plots also clearly demonstrate that the scenario considered provides a region of parameter space in which dark matter interactions can affect smaller halos without endangering the ellipticity of large halos; this is the region enclosed by the blue line which is not colored

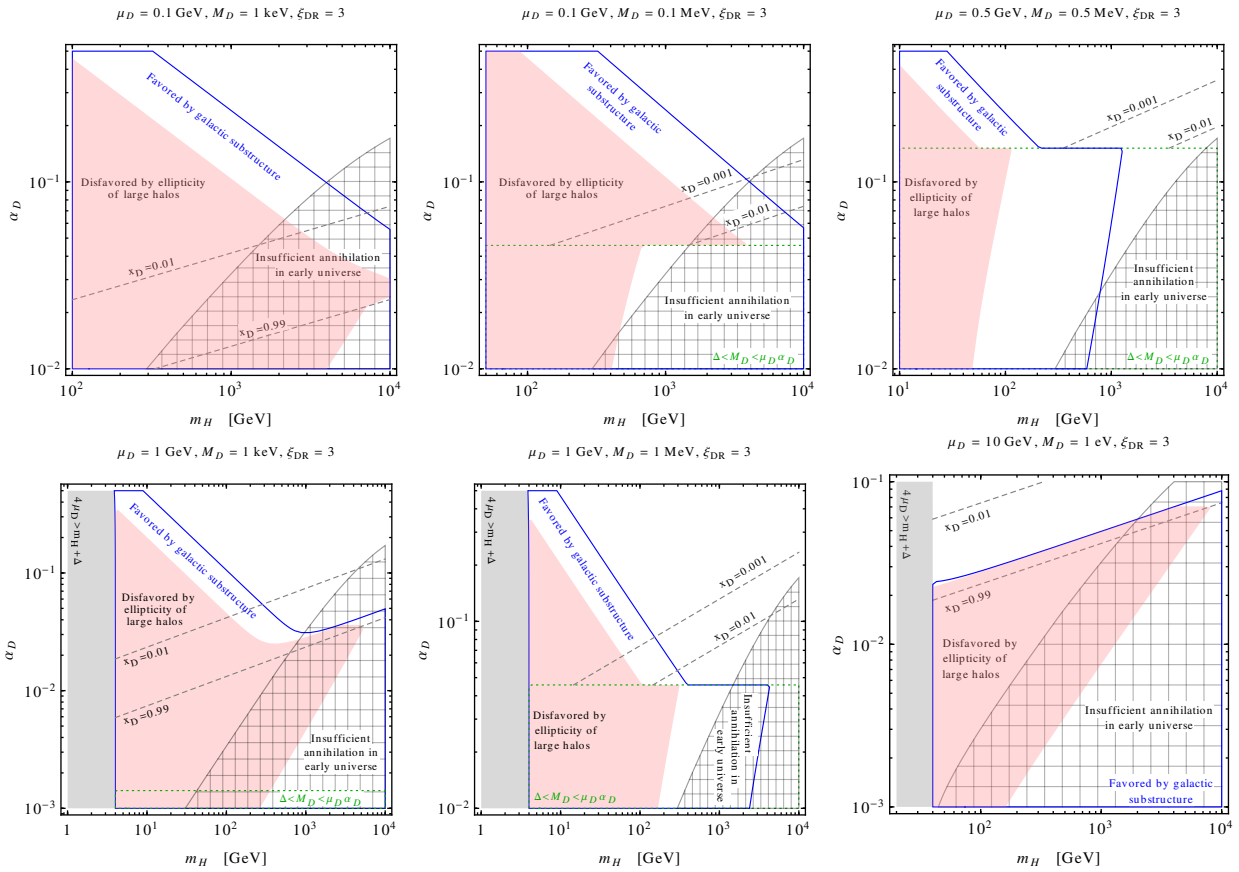

Figure 6.8: Halo bounds as a function of α_D and m_H (fixed μ_D and M_D) for $\xi_{\rm DR} = 3$, following the methods of [175].

pink. Furthermore, this region exists both at small are large ionization fractions, although the large ionization case may be in tension with having sufficient annihilation in the early universe. (However, the fourth plot in Fig. 6.4 demonstrates that this is not necessarily the case.)

The existence of a region satisfying both $\Gamma_{\rm eff}^{\rm DW} > \Gamma_{\rm crit}^{\rm DW}$ and $\Gamma_{\rm eff}^{\rm MW} < \Gamma_{\rm crit}^{\rm MW}$ can be traced to the velocity dependence of the cross sections. Both the atom-atom and ion-ion cross sections decrease with increasing velocity; for ion-ion scattering, $\sigma_{ij} \propto 1/v^4$ as $M_D \to 0$; larger values of M_D weaken this dependence. The sensitivity of σ_{HH} to velocity varies according to the two approaches considered. According to Ref. [181], the atom-atom contribution is velocity-

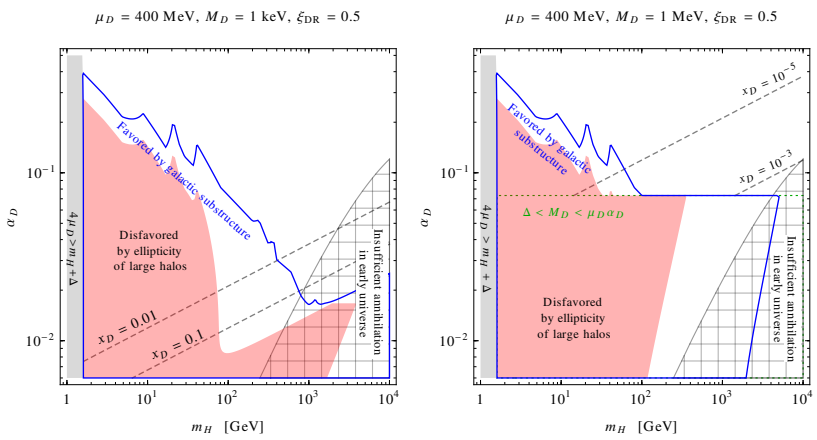

Figure 6.9: Halo bounds as a function of α_D and m_H (fixed μ_D and M_D) for $\xi_{\rm DR} = 0.5$, following the methods of [181].

independent at low energies, while at higher energies it can scale as $1/v^2$ or even as $1/v^4$. Thus, the different velocity scales relevant for dwarf halos and Milky-way-sized halos enable us to satisfy both conditions simultaneously in some regions of parameter space, consistent with the motivation for considering self-interacting dark matter given at the beginning of this chapter.

In the μ_D plots (Figs. 6.7, 6.8, and 6.9), we also explore the region in which $\Delta < M_D < \mu_D \alpha_D$. In this region, bound state solutions exist (and hence m_H and Δ are well-defined), but bound state formation is hindered due to the impossibility of emitting an on-shell dark photon. In this region, therefore, we take $x_D = 1$. We see interesting behavior around the dividing line $M_D = \Delta$; in the upper middle plots of Figs. 6.7 and 6.8, the pink shaded region extends to larger m_H values above the diving line, which is the region in which bound states form. Conversely, in the bottom middle plots of 6.7 and 6.8, the pink shaded region extends to larger m_H values below the dividing line, in the region where bound states do not form. (Recall that the pink shaded region denotes the regime in which $\Gamma_{\rm eff}^{\rm MW} > \Gamma_{\rm crit}^{\rm MW}$.)

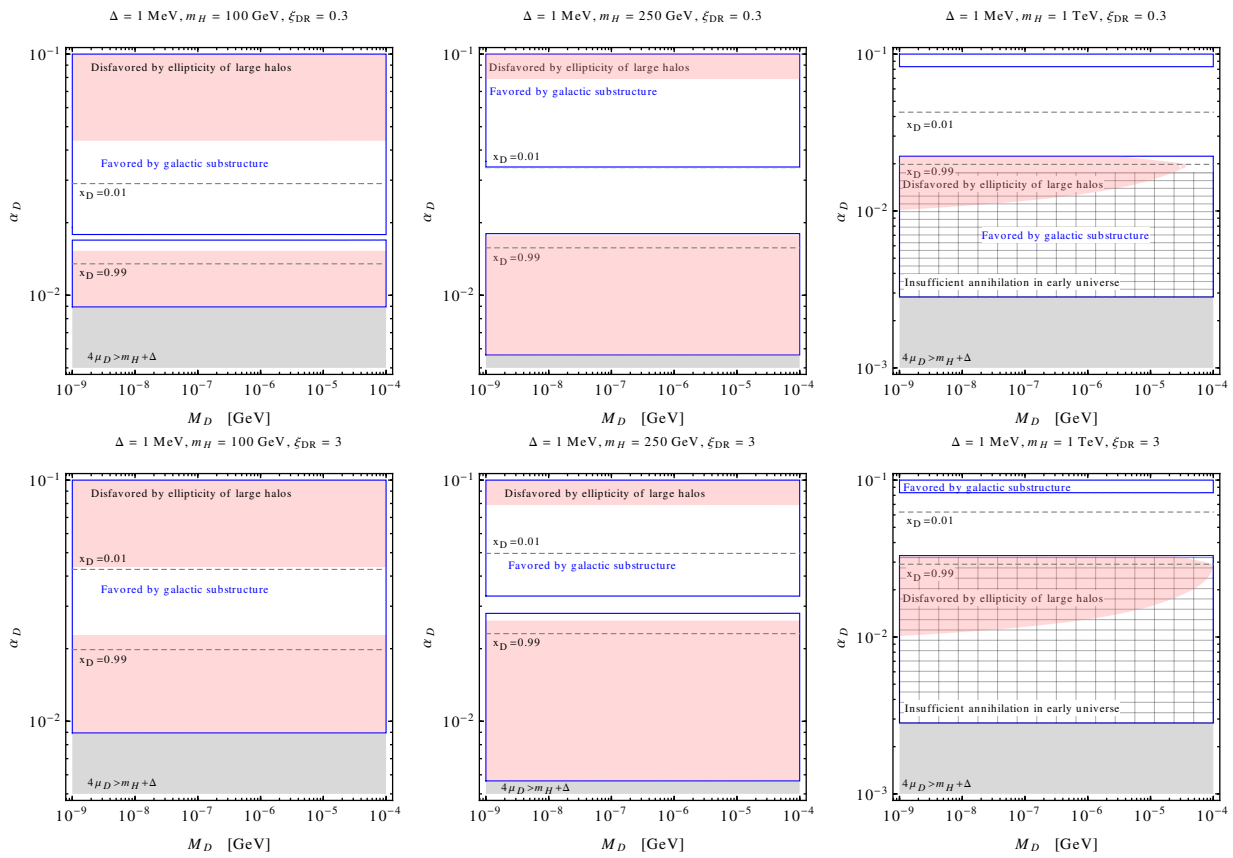

Figure 6.10: Halo bounds as a function of α_D and M_D (fixed Δ and m_H) for $\xi_{\rm DR} = 0.3$ (top) and $\xi_{\rm DR} = 3$ (bottom), following the methods of [175].

This can be understood as follows: If $\Gamma_{\rm eff}^{\rm MW}$ is determined by atom-atom scattering on the $M_D < \Delta$ side of the $M_D = \Delta$ line, then in the $M_D > \Delta$ region the absence of atoms may be sufficient to prevent $\Gamma_{\rm eff}^{\rm MW}$ from exceeding the critical value. Hence, $\Gamma_{\rm eff}^{\rm MW}$ will exceed its critical value for a larger range of dark atom masses where dark atoms can form. Conversely, if we are in a regime in which ion-ion scattering determines $\Gamma_{\rm eff}^{\rm MW}$, then the bound $\Gamma_{\rm eff}^{\rm MW} > \Gamma_{\rm crit}^{\rm MW}$ is more stringent in the $M_D > \Delta$ region, because in this region $x_D = 1$ and more ions are present.

Next, we present plots in which m_H and either Δ or μ_D are held constant, while α_D and M_D vary. These are shown following the approach of [175] in Figs. 6.10 and 6.11,

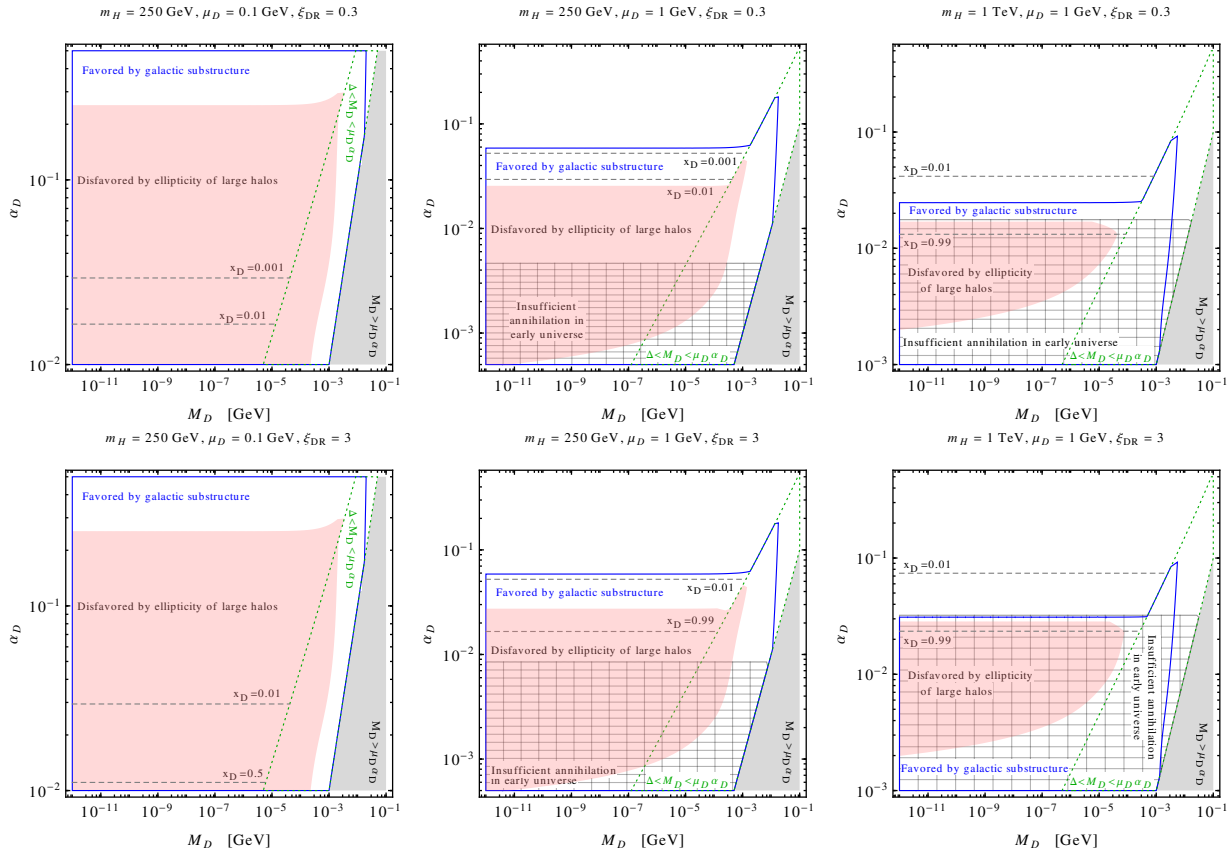

Figure 6.11: Halo bounds as a function of α_D and M_D (fixed μ_D and m_H) for $\xi_{\rm DR} = 0.3$ (top) and $\xi_{\rm DR} = 3$ (bottom), following the methods of [175].

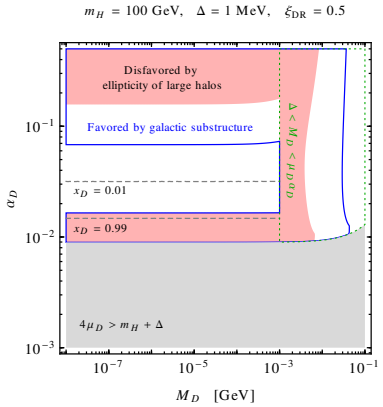

Figure 6.12: Halo bounds as a function of α_D and M_D (fixed Δ and m_H) for $\xi_{\rm DR} = 0.5$, following the methods of [181].

and following the approach of [181] in Fig. 6.12. The wedge seen in the previous plots is converted to a gap between the pink-shaded region and/or the region enclosed in blue. The most notable feature is how little the bounds evolve with the mass of the dark mediator. We do see that the bound from Milky-Way-sized halos weakens as the dark mediator becomes heavy; this is as we expect due to the increased screening. The importance of the mediator mass near the $M_D = \Delta$ bound, beyond which it is difficult to form bound states although these solutions exist in the theory, is evident in Fig. 6.11. Generally, the constraint is stronger on the $M_D > \Delta$ side due to the increased number of ions with large scattering cross sections, although the left-most plots of Fig. 6.11, in which the momentum transfer is dominated by atom-atom scattering, is a notable exception. Again, the formation of dark atoms significantly alters the allowed and preferred regions of parameter space.

Figs. 6.13 and 6.14 show the parameter space as a function of m_H and M_D following the methods of Ref. [175]; Fig. 6.15 follows Ref. [181]. Consistent with Figs. 6.10, 6.11, and 6.12, minimal evolution with the dark photon mass is seen until near $\Delta = M_D$. The regime in which dark matter self-interaction can influence small halos while preserving the ellipticity of large halos can extend over nearly a whole decade in α_D. (In Fig. 6.14, the region $M_D > \mu_D \alpha_D$, is grayed out because there are no bound state solutions.)

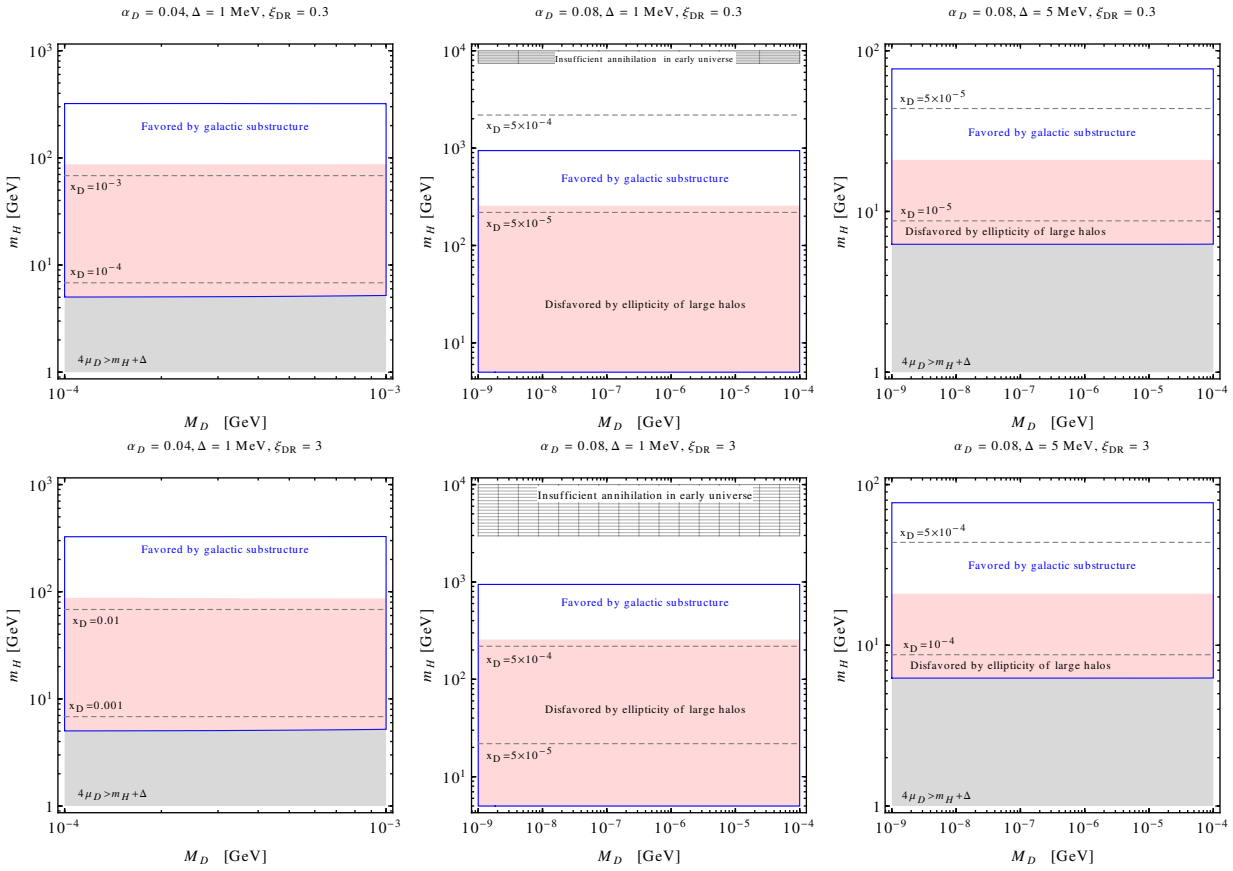

Figure 6.13: Halo bounds as a function of m_H and M_D (fixed Δ and α_D) for $\xi_{\rm DR} = 0.3$ (top) and $\xi_{\rm DR} = 3$ (bottom), following the methods of [175].

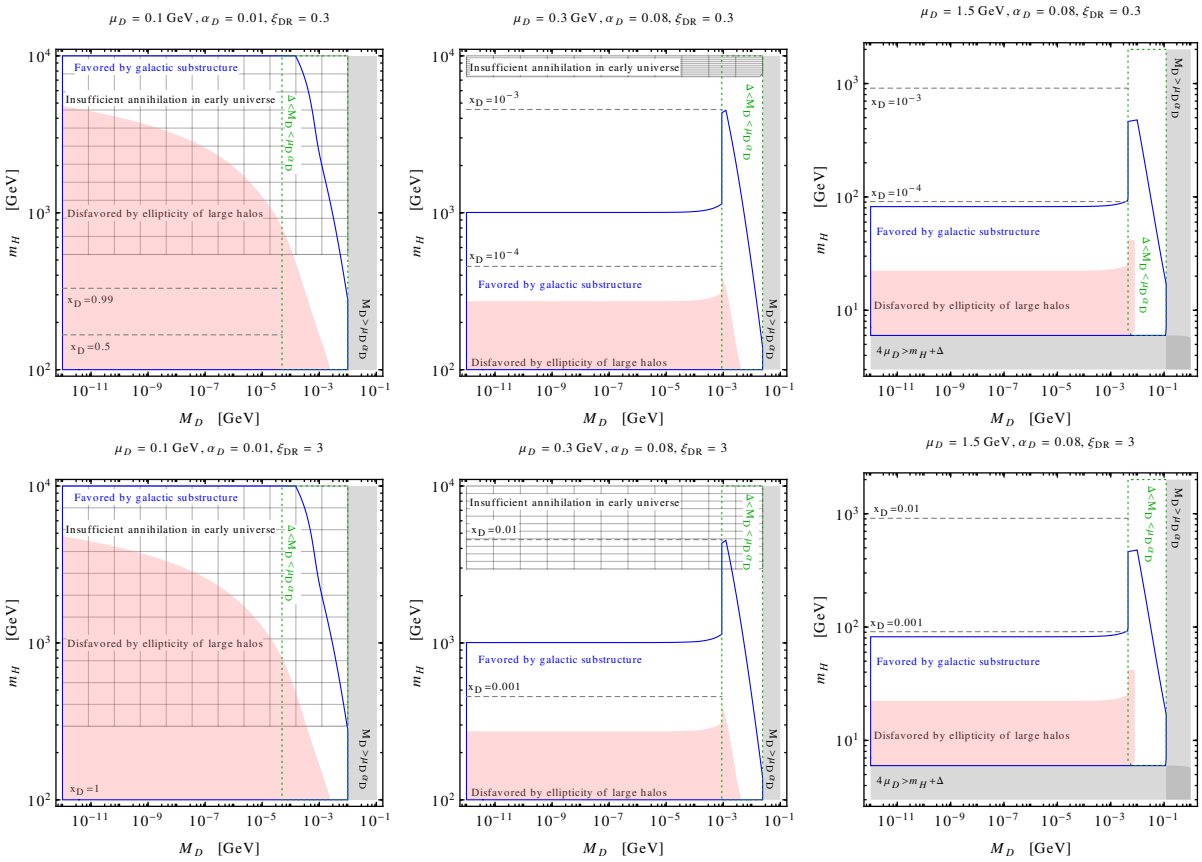

Figure 6.14: Halo bounds as a function of m_H and M_D (fixed μ_D and α_D) for $\xi_{\rm DR} = 0.3$ (top) and $\xi_{\rm DR} = 3$ (bottom), following the methods of [175].

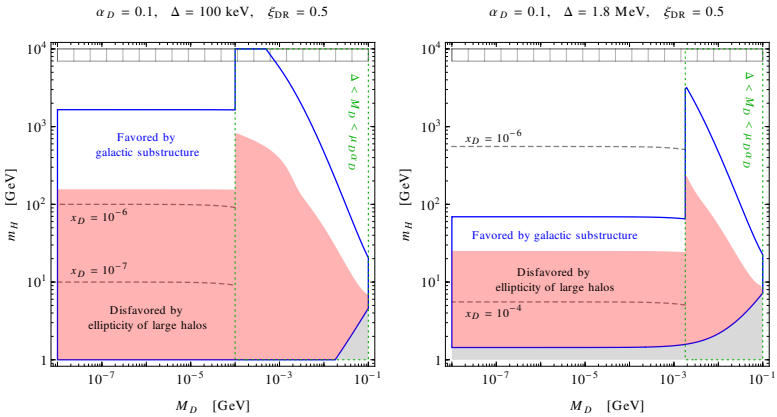

Figure 6.15: Halo bounds as a function of m_H and M_D (fixed α_D and Δ) for $\xi_{\rm DR} = 0.5$, following the methods of [181].

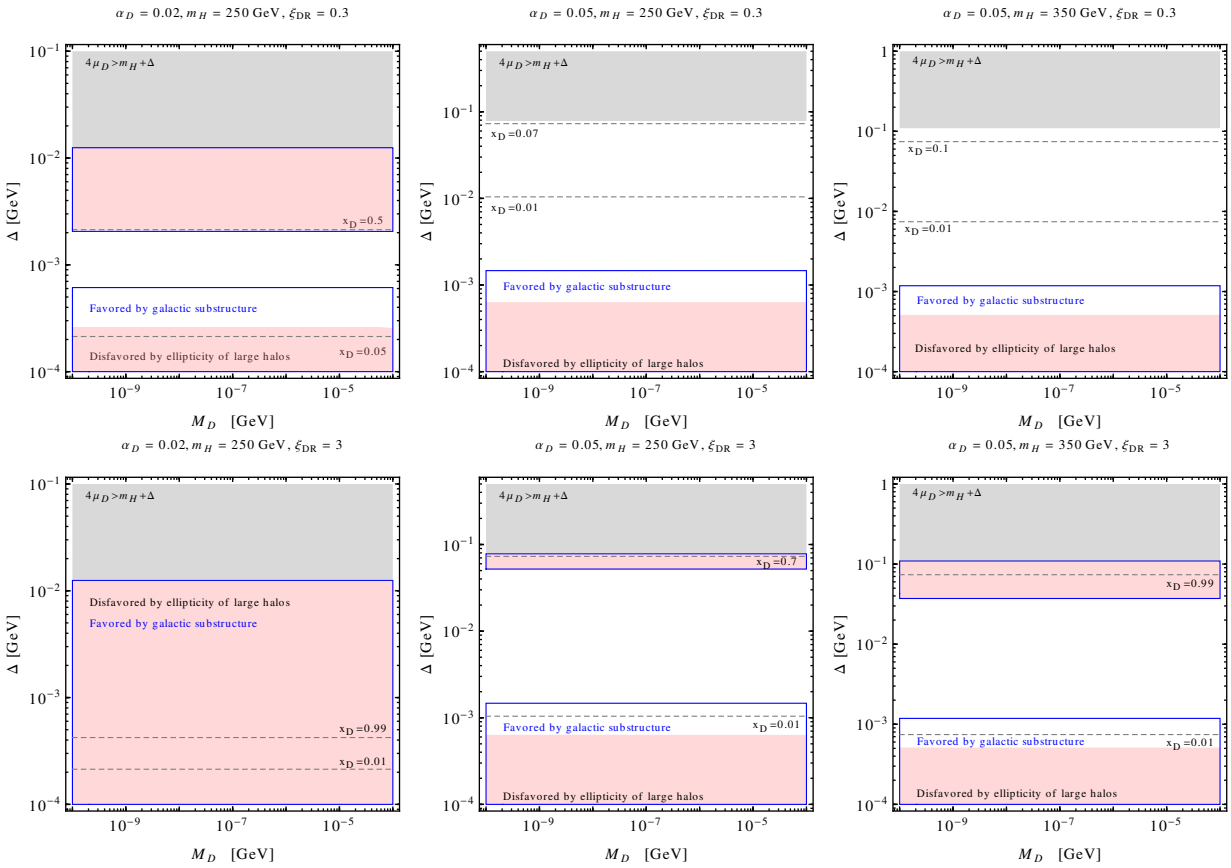

Figure 6.16: Halo bounds as a function of Δ and M_D (fixed m_H and α_D) for $\xi_{\rm DR} = 0.3$ (top) and $\xi_{\rm DR} = 3$ (bottom), following the methods of [175].

In the next set of plots (Fig. 6.16, Fig. 6.17, and Fig. 6.18), we fix the dark coupling α_D and the dark atom mass m_H, while allowing M_D along with either Δ or μ_D to vary. Again, we see very little evolution with the dark photon mass. The wedge feature again results in two bands. We see that it is generally easier to alter the shape of small halos while retaining the ellipticity of large halos at small ionization fractions, when most of the charge is contained in dark atoms.

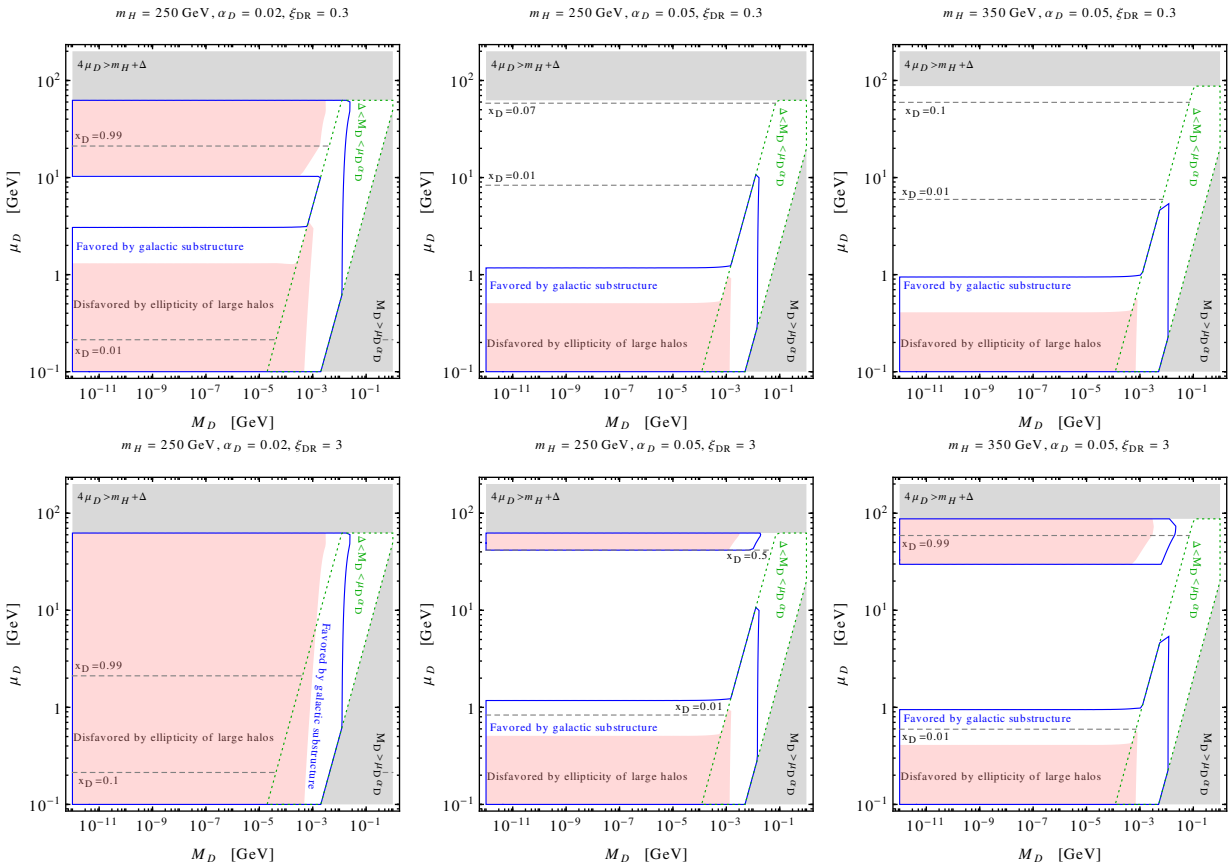

Figure 6.17: Halo bounds as a function of μ_D and M_D (fixed m_H and α_D) for $\xi_{\rm DR} = 0.3$ (top) and $\xi_{\rm DR} = 3$ (bottom), following the methods of [175].

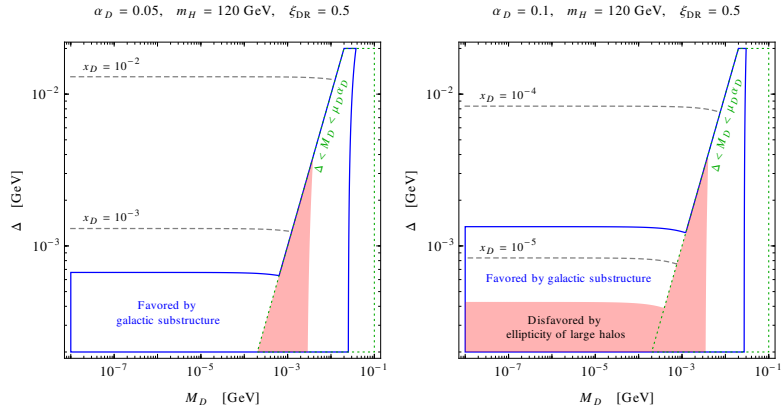

Figure 6.18: Halo bounds as a function of Δ and M_D (fixed m_H and α_D) for $\xi_{\rm DR} = 0.5$, following the methods of [181].

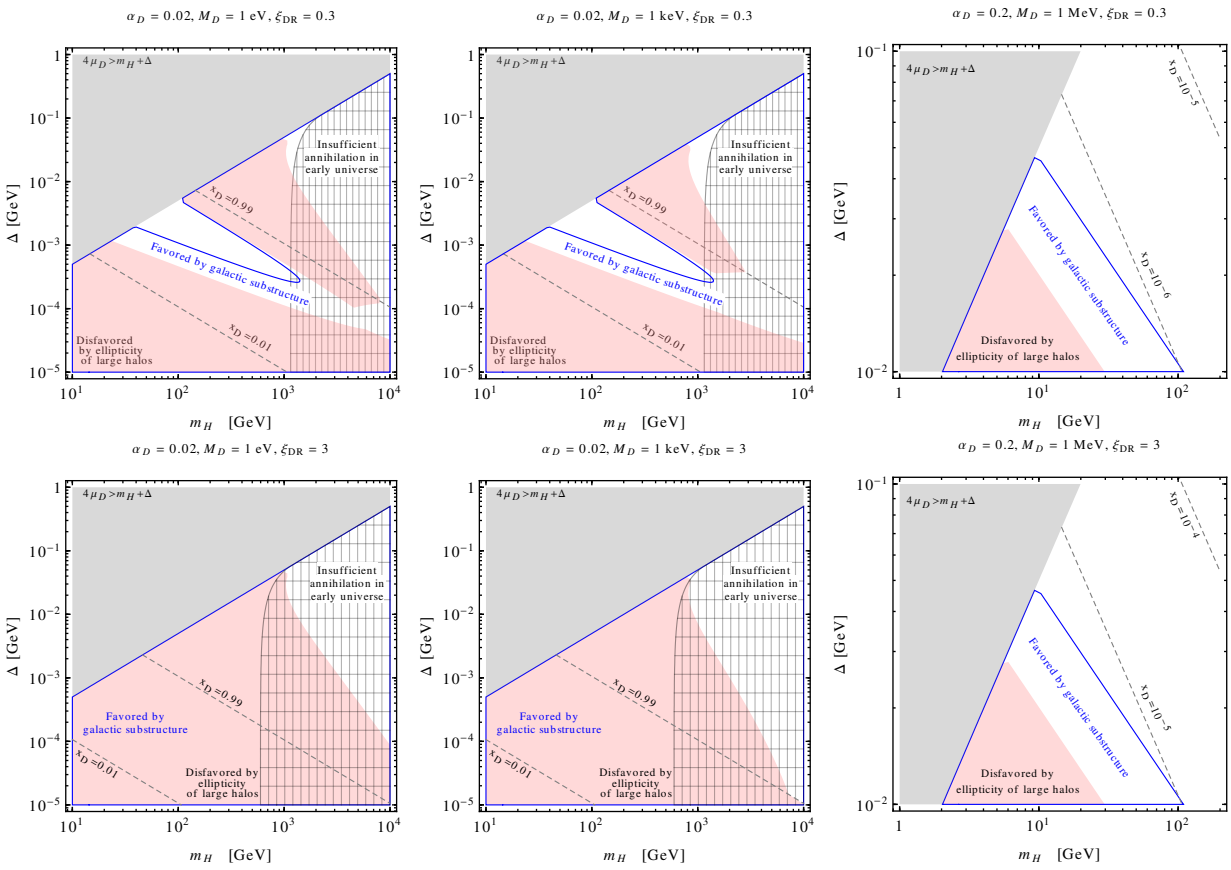

Figure 6.19: Halo bounds as a function of Δ and m_H (fixed M_D and α_D) for $\xi_{\rm DR} = 0.3$ (top) and $\xi_{\rm DR} = 3$ (bottom), following the methods of [175].

Next we hold the dark coupling α_D and the dark photon mass M_D constant while allowing m_H along with the binding energy Δ or reduced mass μ_D to vary. This allows us to investigate the evolution of the wedge region with binding energy/reduced mass, which we were not able to do effectively in the first set of plots. These are shown in Fig. 6.19 as a function of Δ following the approach of [175], in Fig. 6.20 as a function of μ_D following the approach of [175], and in Fig. 6.21 as a function of Δ following the approach of [181]. In particular, we note that the wedge extends for a much larger range of dark atom masses taking the approach of [181] as opposed to [175]. We also see that, when considering this set of fixed parameters, the existence of the wedge occurs only for a narrow range of binding energies or reduced masses, and it is also rather sensitive to $\xi_{\rm DR}$. Consistent with the above remarks,

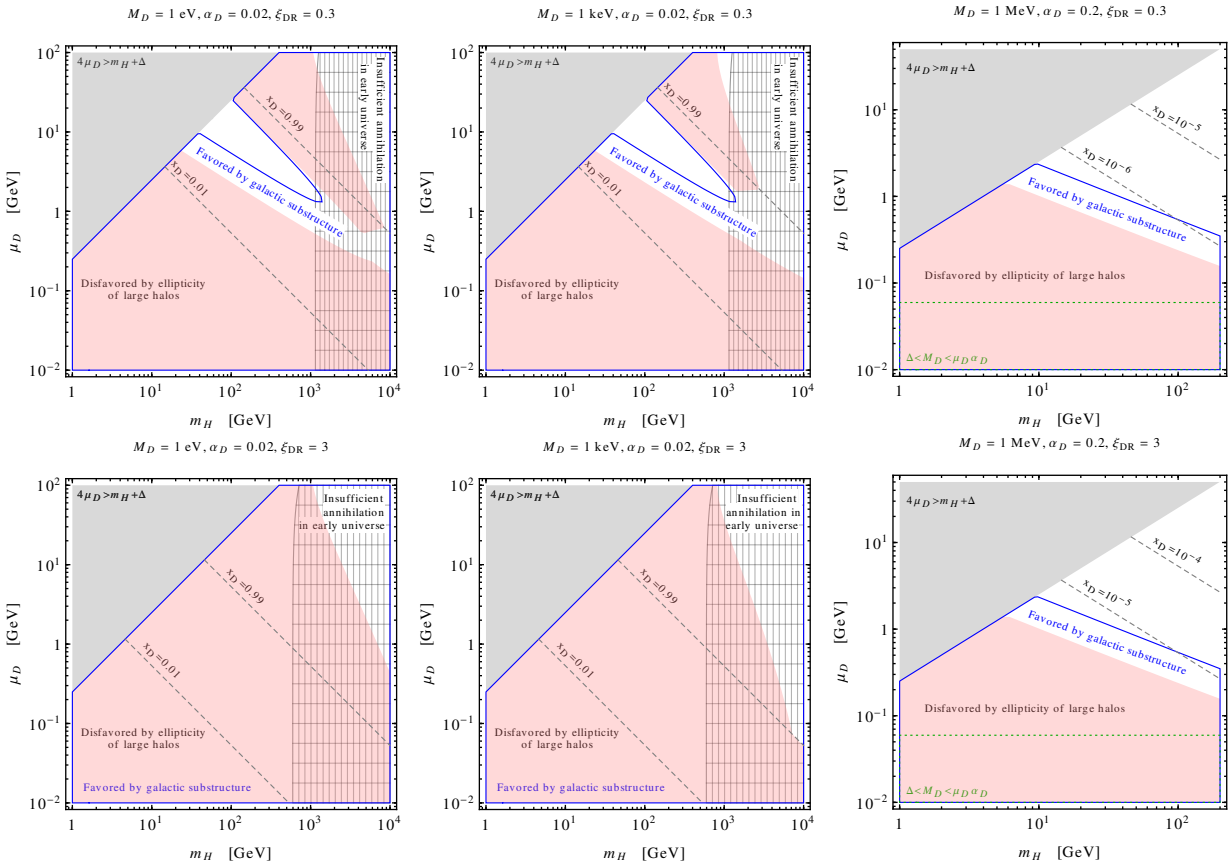

Figure 6.20: Halo bounds as a function of μ_D and m_H (fixed M_D and α_D) for $\xi_{\rm DR} = 0.3$ (top) and $\xi_{\rm DR} = 3$ (bottom), following the methods of [175].

we see little evolution of the parameter space until the dark photon mass becomes relatively large.

Additionally, these plots show both small and large ionization regimes in which these desiderata hold; see particularly the first two plots in Fig. 6.19. As noted previously, the favored region at high ionization fraction is frequently in tension with having sufficient annihilation, but this region should not necessarily be considered forbidden, as an extended dark sector with additional annihilation channels may loosen this bound.

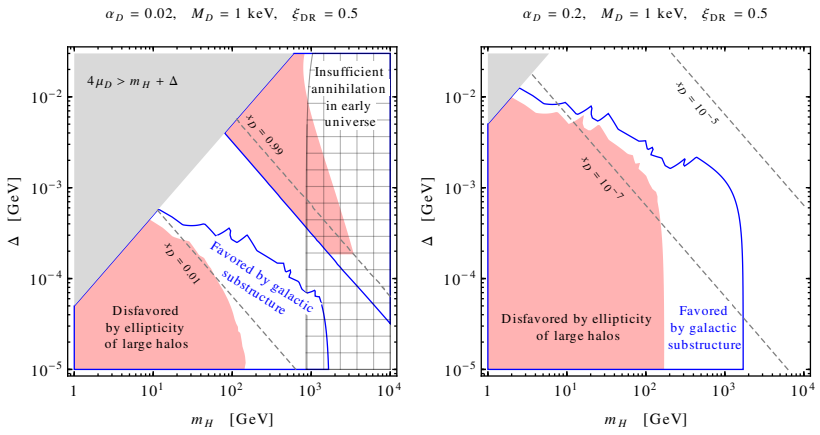

Figure 6.21: Halo bounds as a function of Δ and m_H (fixed M_D and α_D) for $\xi_{\rm DR} = 0.5$, following the methods of [181].

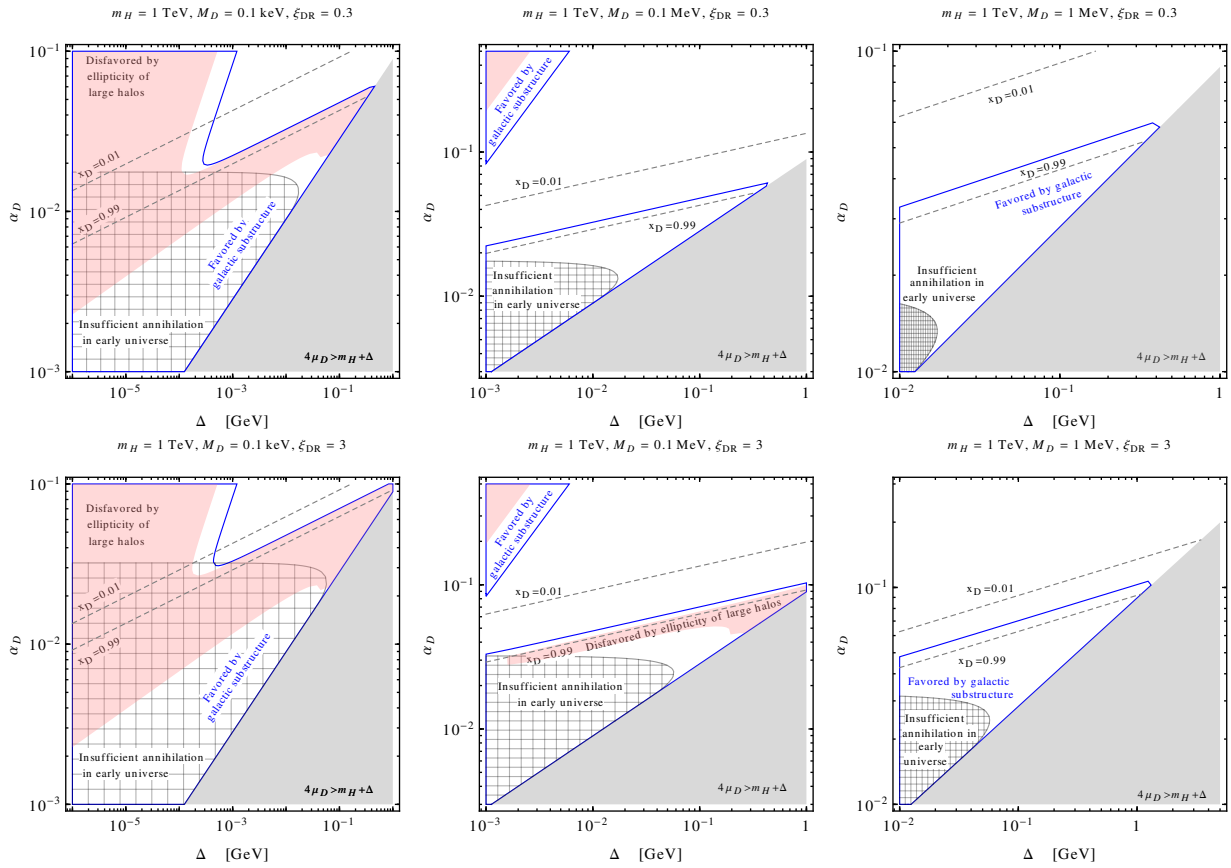

Figure 6.22: Halo bounds as a function of α_D and Δ (fixed m_H and M_D) for $\xi_{\rm DR} = 0.3$ (top) and $\xi_{\rm DR} = 3$ (bottom), following the methods of [175].

Finally, we consider the remaining possibility: fixing m_H and M_D; Fig. 6.22 shows these as a function of Δ and Fig. 6.23 shows these as a function of μ_D. As these are rather similar to previously shown plots, we show the analysis only for following [175]. These plots particularly illuminate the evolution of the preferred region at large ionization fractions.

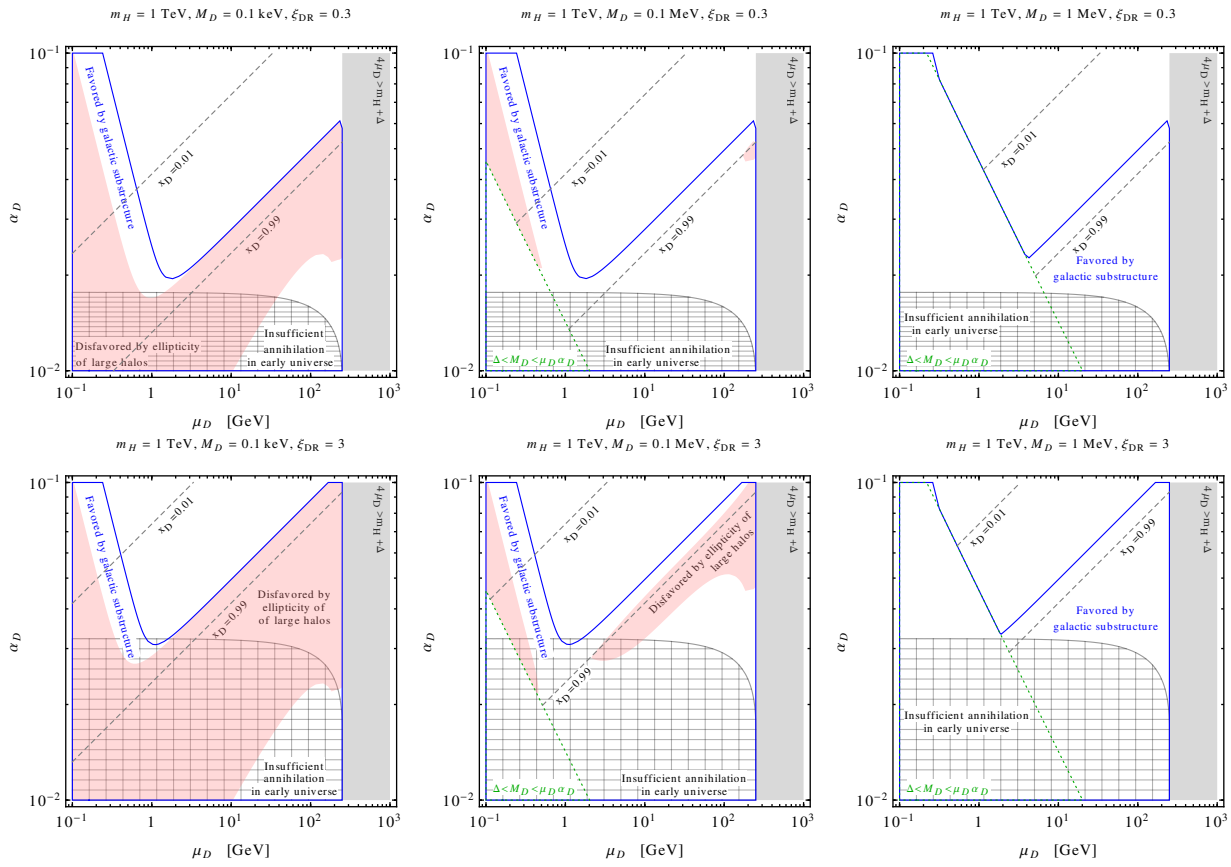

Figure 6.23: Halo bounds as a function of α_D and μ_D (fixed m_H and M_D) for $\xi_{\rm DR} = 0.3$ (top) and $\xi_{\rm DR} = 3$ (bottom), following the methods of [175].

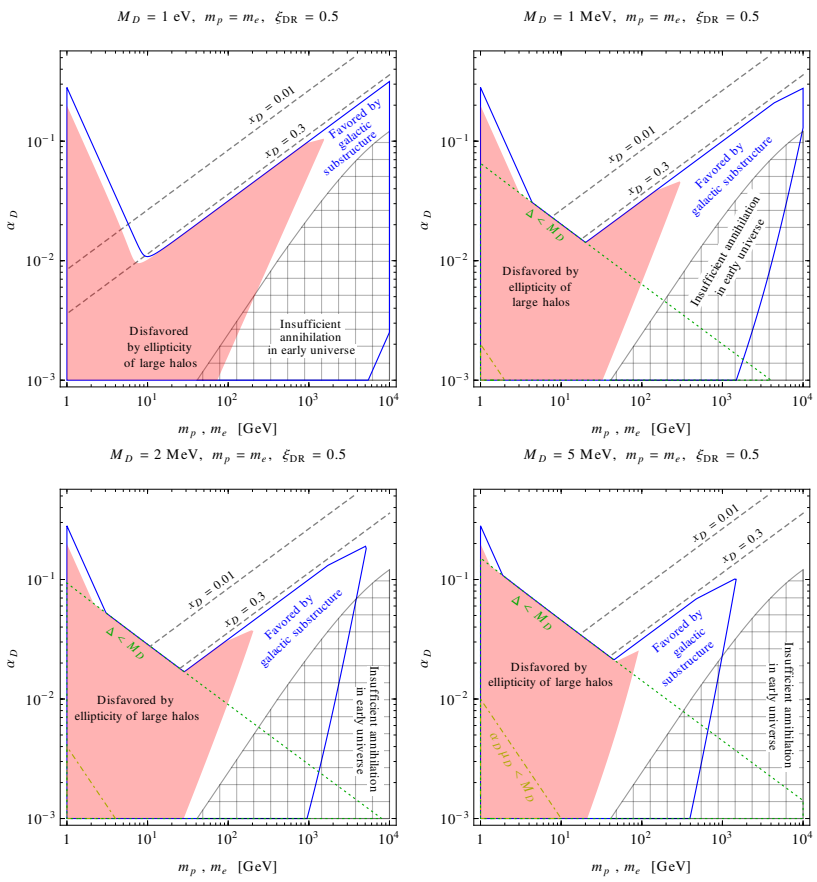

Figure 6.24: Halo bounds as a function of α_D and $m_p = m_e$ (fixed M_D) for $\xi_{\rm DR} = 0.5$, following the methods of [181].

We have seen that arbitrarily small dark photon masses produce viable scenarios, even at moderately large values of the dark coupling (e.g., Fig. 6.12). This is due to the formation of dark atoms, which generally suppresses the dark matter self-scattering rate. If the dark photon has mass, this mass must necessarily be small for existence dark atoms to form; for $M_D > \Delta$, the formation of atoms is kinematically suppressed due to the emitted dark mediator being off mass-shell, and for $M_D > \mu_D \alpha_D$, the spin-zero bound state solution does not exist. The above plots did explore the $M_D > \Delta$ regime by setting the ionization fraction to one; however, to explore the $M_D > \mu_D \alpha_D$ regime one should use a different set of independent variables, as the parameters m_H and Δ are not well-defined in this regime.

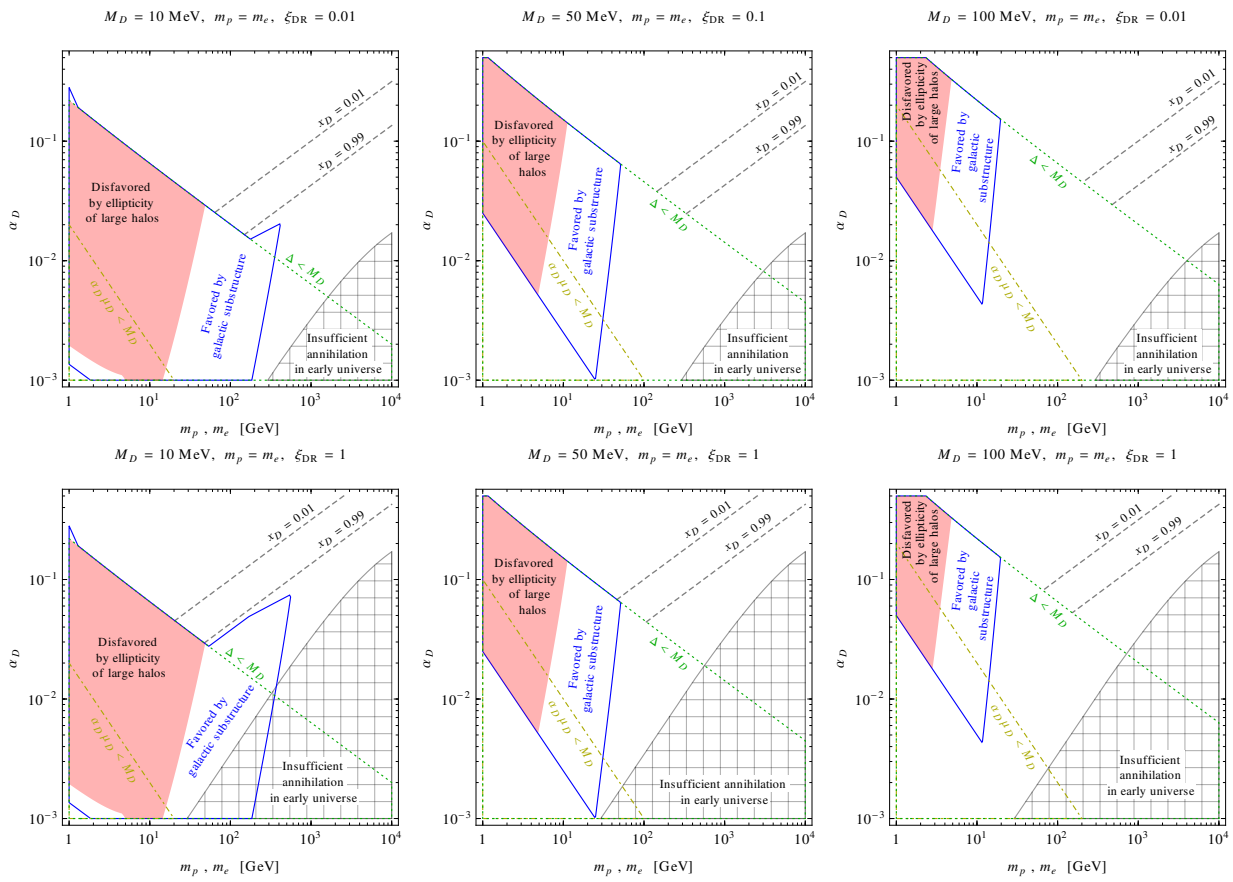

Figure 6.25: Halo bounds as a function of α_D and $m_p = m_e$ (fixed larger M_D) for $\xi_{\rm DR} = 0.01$ (top) and $\xi_{\rm DR} = 1$ (bottom), following the methods of [181].

Consequently, we next present plots in which we use m_p, m_e, M_D, and α_D as independent parameters; these are all made following the approach of [181]. In most of these plots, we will set $m_p = m_e$ for simplicity; this also most closely approximates single component self-interacting dark matter models.

In the first set of plots, Fig. 6.24, we investigate the evolution of the wedge seen in the plots above. We see that it generically becomes larger as M_D is increased, consistent with the plots shown above. The two bottom plots explore the region with no bound states; we observe that in this regime, the large ion-ion scattering cross section results in spherical large halos, as we expect from our observation that dark atom formation tends to relax this

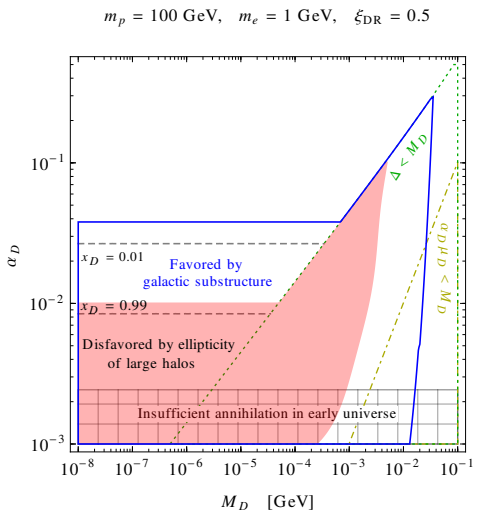

Figure 6.26: Halo bounds as a function of α_D and M_D (fixed larger $m_p \gg m_e$) for $\xi_{\rm DR} = 0.5$, following the methods of [181].

constraint.

Next we consider larger mediator masses, at both $\xi_{\rm DR} = 0.01$ and $\xi_{\rm DR} = 1$; these are shown in Fig. 6.25. In these plots, we can see that both the $\Gamma^{\rm MW}_{\rm eff} > \Gamma^{\rm MW}_{\rm crit}$ (pink) region and the $\Gamma^{\rm DW}_{\rm eff} > \Gamma^{\rm DW}_{\rm crit}$ (enclosed in blue) region extend into the large mediator mass regime, in which there are no bound state solutions. However, as the dark photon mass continues to increase with respect to α_D, the screening effect does eventually reduce the effective momentum transfer beneath the critical value.

In this limit, there is again a viable regime in which the scattering remains sufficient to affect smaller halos without destroying the ellipticity of larger halos. We emphasize that in this scenario this is due entirely due to the screening from the dark photon mass, in contrast with the preferred parameter space due to the formation of dark atoms. However, for cosmologically stable dark photons, a large mass implies rather strong upper limits on ξ, as shown in Fig. 6.2. This bound can be weakened by increasing $q_\phi^2 \alpha_D$; however, this cannot be taken too large without weakening the ionization fraction. However, if the kinetic mixing

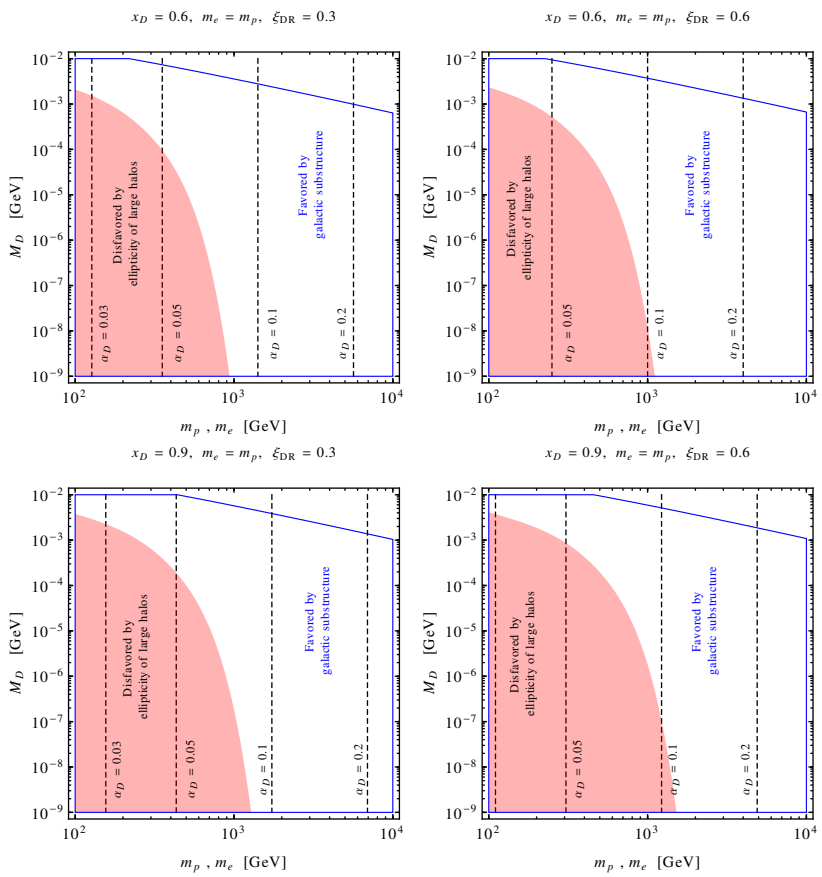

Figure 6.27: Halo bounds as a function of M_D and $m_p = m_e$ for fixed x_D and $\xi_{\rm DR}$, following the methods of [181].

between the dark photon and hypercharge is sufficiently large, then the dark photons will decay rapidly, relaxes or eliminating the constraints shown in Fig. 6.2. We emphasize that even in this minimal model, there are a variety of viable regions of parameter space, with very different characteristics.

While Fig. 6.25 shows this additional possibility, it does also illustrate the dramatic effect that the formation of dark atoms has on halo ellipticity bounds, which we have explored above. When the ionization fraction is small, the effective momentum rate does not exceed its critical value due to the weakness of atom-atom scattering. We also observe that the effect of changing $\xi_{\rm DR}$ is rather minimal, mostly altering the location of the lines of constant

ionization.

These figures have all investigated the case $m_p = m_e$; however, the limit $m_p \gg m_e$ is also of interest. This is shown in 6.26. This provides an even larger region in which the mass of the dark photon screens the interactions sufficiently to retain the ellipticity of large halos while allowing scattering to modify the shape of smaller halos.

The above discussion has emphasized the importance of the ionization fraction. At small ionization fractions, scattering may be suppressed by atom-atom scattering, which is weaker than ion-ion scattering (however, this may still be strong enough to cause $\Gamma_{\rm eff}$ to exceed its critical value). At lower ionization fractions, screening can only be accomplished by a larger dark photon mass; this may or may not require one to go into the regime in which dark atoms do not form. Consequently, it is of interest to fix the ionization fraction x_D; this is shown in Fig. 6.27. The preferred region tends to be at relatively large dark proton/dark electron masses, unless one considers relatively large dark photon masses, consistent with the above discussion.

6.4 Summary

In this chapter, we have considered a rather minimal model of self-interacting asymmetric dark matter: a dark sector interacting through a broken $\mathrm{U_D}(1)$ gauge group. We have argued that for small dark photon masses, the dark asymmetry is necessarily produced in two species oppositely charged under $\mathrm{U_D}(1)$; furthermore, these species generically can survive to form dark atoms. Therefore, even in this simplistic model, dark matter is generally multi-component, comprised of negative dark ions, positive dark ions, and dark atoms.

This model necessarily has a more involved cosmology, which involves a dark phase transition (if the symmetry is broken using a dark Higgs field) and a period of dark recombination. The different annihilation processes may freeze out at different temperatures; the decay of

the dark Higgs boson and/or the dark photon can affect the visible sector. These factors lead to a rich cosmological history, with many possibilities; we have discussed this in Sec. 6.2.

Finally, single-component halo bounds are generally not applicable to this model; in the final section, we have considered such bounds taking full account of the multi-component nature of the dark matter. We have seen that scattering in halos may be suppressed due to dark atom formation or due to large dark photon masses, both of which result in regions of parameter space in which scattering can modify small halos while large ones remain elliptical.

This chapter complements the previous chapter, which did not consider a $U_D(1)$ gauge group, and thus was able to consider single-component dark matter. However, the indirect detection signals studied there are also generally applicable to this model, and can probe regions of parameter space in which the coupling is rather large.

CHAPTER 7

Solitosynthesis-Induced Phase Transitions

Now we turn our attention to our third, and final, topic. We return again to the question of baryogenesis. In Ch. 4, we explored the possibility of producing a baryonic asymmetry during the electroweak phase transition in a strongly-coupled supersymmetric model. However, this is not the only possibility; other phase transitions in the early universe can also produce the necessary asymmetry. In this chapter, we consider a different possibility, a phase transition induced by Q-balls. Q-balls [202] are non-topological solitons [203–206] that are stable because they carry a conserved global charge. They arise in a number of models, and, in particular, in supersymmetric extensions of the Standard Model, where they carry baryon and/or lepton number [207]. Stable supersymmetric Q-balls can form in the early universe from the fragmentation of an Affleck-Dine condensate [100, 208–212] or in other processes [213, 214], and they can play the role of cosmological dark matter [100, 101, 208–211, 215–217].

Furthermore, it has been suggested that Q-balls can facilitate phase transitions even when the tunneling rate is too small for the phase transition to occur otherwise; the Q-balls accumulate charge until they reach a critical charge, at which point they expand, causing a phase transition [218, 219]. Such phase transitions have been considered as a potential baryogenesis mechanism. While the possibility of such a phase transition has been explored in the literature [220,221], a complete model of it has not yet been demonstrated. This is due to difficulties with the quantum nature of small charge Q-balls and also with the properties of Q-balls in the false vacuum. This chapter will demonstrate, from beginning to end, a model in which a phase transition is induced by solitosynthesis of Q-balls.

This chapter is organized as follows: first, we specify the potential that gives rise to our Q-balls and show that it has the requisite properties; then we consider the properties of the non-topological solitons in the false vacuum. There are primarily four regimes to consider. For large charges, the thin wall semi-classical approximation is valid, while for smaller charges, the thick wall semi-classical approximation is valid. For intermediate charges, we interpolate between these two regimes. For extremely small charges, quantum effects are important and the semi-classical approximation is invalid; instead, we apply the Bethe-Salpeter equation.

After we have described the radii and energies of the Q-balls, we proceed to consider the properties of the phase transition; in particular, the critical charge and the critical radius. Then we discuss solitosynthesis, the process by which Q-balls grow by accreting charge. We find the temperature at which such growth begins, and then we calculate the rate of growth in each regime. We demonstrate that the growth is not hindered by charge depletion and freeze out, which could end solitosynthesis before critically sized Q-balls form. Finally, we discuss explicit numerical examples to show that such a phase transition is a theoretical possibility.

In all this, we use a simplified toy model inspired by the Minimal Supersymmetric Standard Model (MSSM). In the last section, we discuss the application of this analysis to the MSSM, and in particular we consider phase transitions of cosmological interest.

7.1 The Potential

For this analysis, we will use an MSSM-inspired potential [217, 220]

$$V(\tilde{q}, H) = \frac{m_q^2}{2}\tilde{q}^2 + \frac{m_h^2}{2}h^2 - A_0 h\tilde{q}^2 + \frac{\lambda_1}{4}\tilde{q}^2 h^2 + \frac{\lambda_2}{4}\tilde{q}^4 + \frac{\lambda_3}{4}h^4, \tag{7.1}$$

in which $\tilde{q}$ is a squark field and h is the lightest Higgs boson. For simplicity, we use real fields. In general, renormalization effects, including the effects of Q-balls if they exist, can be significant [222]; we take the couplings to be the renormalized couplings. We assume

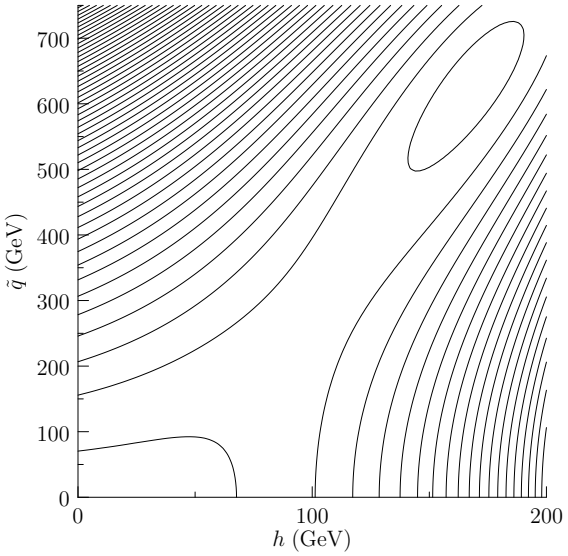

Figure 7.1: A contour plot of the potential $V(\tilde{q}, h)$. There is a local minimum at the origin and a global minimum at $\langle h \rangle = 169$ GeV and $\langle \tilde{q} \rangle = 631$ GeV.

that any other particles which carry baryon number are heavier than the squark, to ensure the squark's stability. The origin is always a local minimum of this potential; however, for particular values of the coupling constants, a different global minimum exists. For example, if $m_q = 200$ GeV, $m_h = 10$ GeV, $A_0 = 240$ GeV, $\lambda_1 = \lambda_2 = 0.1$, and $\lambda_3 = 19$, the origin is a local minimum and there are global minima at $\langle h \rangle = 169$ GeV and $\langle \tilde{q} \rangle = \pm 631$ GeV. Fig. 7.1 shows a contour plot of this potential.

Phase transitions involving multiple fields are difficult to solve exactly; a reasonable approximation is that they occur along the line connecting the two minima. This is a valid approximation when the potential does not have an unusually shaped barrier; we see no sign of an unusual barrier in the contour plot. The potential along this line may be found by substituting $\tilde{q} = \phi \sin(\theta)$ and $h = \phi \cos(\theta)$ with $\theta = 1.309$:

$$V(\phi) = 18700 \text{ GeV}^2 \phi^2 - 58.0 \text{ GeV} \phi^3 + 0.0447 \phi^4. \tag{7.2}$$

This effective potential is shown in Fig. 7.2. We see that there is a large barrier between the minima; this suppresses phase transitions induced by thermal fluctuations.

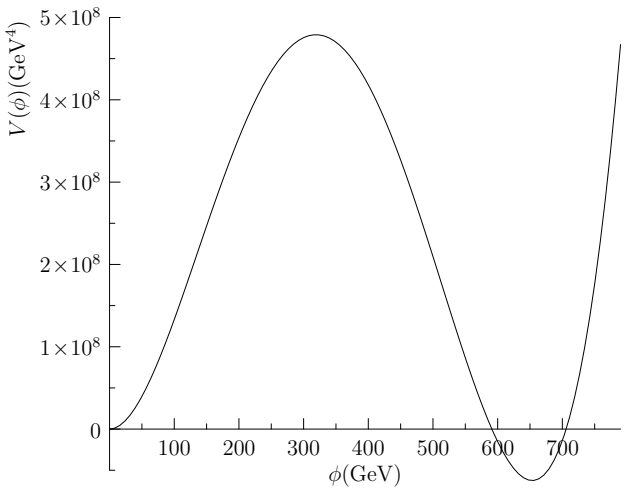

Figure 7.2: The effective potential along the line connecting the false vacuum to the true vacuum. The large barrier suppresses phase transitions driven by thermal fluctuations.

Next, we must demonstrate the existence of Q-balls in the false vacuum (located at the origin). The conserved charge carried by the Q-balls is baryon number; this is conserved in the false vacuum because squarks do not have a vacuum expectation value. The condition for the existence of Q-balls involving multiple fields is that $2V/\sum_k Q_k\phi_k^2$ is minimized at a nonzero value of the fields, where Q_k is the charge of the field ϕ_k [202, 207]. This condition is derived using the semi-classical approximation, and therefore is applicable to states with large charge (as will be clarified below). Because the baryon number carried by the squark field is 1/3, we consider the minimum of $6V/\tilde{q}^2$. Because of the $\tilde{q}h$ term, the origin will not be even a local minimum of $V/\tilde{q}^2$; in fact, the global minimum of this function is located at $\tilde{q}_0 = 624$ GeV and $h_0 = 168$ GeV.

Therefore, Q-balls carrying baryon number exist in this vacuum. Furthermore, $V(\tilde{q}_0, h_0) = -6.18\cdot10^7$ GeV^4; this is beneath the value in the false vacuum, $V(0,0) = 0$. This is necessary for a Q-ball induced phase transition to occur; such a phase transition converts the fields everywhere to their vacuum expectation values inside the Q-ball, $\tilde{q}_0$ and h_0. Since this is not quite the true minimum of the potential (which is located at $\tilde{q} = 631$ GeV and $h = 169$ GeV),

the system will then slide classically into the true minimum.

Again, we consider the potential along the line connecting the initial vacuum and the final state (the interior of the Q-ball), which, with $\tilde{q} = \phi \sin(\theta)$, $h = \phi \cos(\theta)$, and $\theta = 1.308$, is

$$V(\phi) = 18700 \text{ GeV}^4 \phi^2 - 58.2 \text{ GeV}^3 \phi^3 + 0.0450 \phi^4. \tag{7.3}$$

We can also describe the Q-balls in terms of the effective ϕ field. The field inside such a Q-ball is the value of ϕ that minimizes $V(\phi)/\phi^2$, which is $\sqrt{\tilde{q}_0^2 + h_0^2} = \phi_0 = 646$ GeV. Thus, at least for large charge, we can consider the Q-balls as coherent oscillations of ϕ quanta, but when the classical approximation is not valid, we should remember that the charge is really carried by bound states of squarks exchanging Higgs bosons.

Because squarks carry charge 1/3, the charge present in a particular field configuration (including a Q-ball) is

$$Q = \frac{\omega}{3} \int d^3x \, \tilde{q}^2 = \frac{\sin^2(1.308)}{3} \omega \int d^3x \, \phi^2, \tag{7.4}$$

where ω describes the oscillatory time dependence of the fields. While the first equality is generally true, the second holds along the line connecting the interior of the Q-ball to the origin of the potential. For later convenience, we measure the charge in units such that $Q = \sin^2(1.308) Q'/3$; then,

$$Q' = \omega \int d^3x \, \phi^2. \tag{7.5}$$

Physically, Q' is the charge in a single ϕ quanta. In terms of Q', a single squark carries charge 3.107; this will be even closer to 3 the nearer the global minimum is to the $\tilde{q}$ axis. Speaking loosely, since $\theta \approx \pi/2$, the field ϕ is "almost" the squark field, and the charge Q' is approximately the number of squarks present.

Generically, then, we consider a potential of the form

$$V(\phi) = \frac{m_0^2}{2} \phi^2 - A \phi^3 + \frac{\lambda}{4} \phi^4, \tag{7.6}$$

in which the field ϕ is made of squarks and Higgs bosons, and the general form of the relation between the charges is

$$Q' = \frac{3Q}{\sin^2(\theta)}. \tag{7.7}$$

Let us make a few brief remarks about other sources of correction to the potential. At the scale of color confinement, we expect the squarks to arrange into color singlets of the form $\epsilon_{abc}\epsilon_{\alpha\beta}\tilde{Q}_a^\alpha \tilde{Q}_b^\beta \tilde{q}_c$, where Greek indices denote $\mathrm{SU_L}(2)$ indices and Latin letters denote color indices [223]. To avoid complications from these corrections, we will ensure that the relevant temperatures for our final analysis are above the scale of the QCD phase transition. We note, however, that even when this is not satisfied, the scalar binding interaction mediated by Higgs bosons is much stronger than the strong interaction, and so modifications to the Q-balls are expected to be rather minimal. A second source of corrections to the potential are the finite temperature corrections $\lambda^2 T^2 \phi^2$ and $\lambda T \phi^3$. Although the temperatures considered will be large in comparison to the QCD scale, they will still be significantly smaller than m_0 and A, and so the finite temperature corrections will not in fact be significant in the numerical examples we will consider.

7.2 Properties of Q-Balls in the False Vacuum

The exact and general equation for the energy of a Q-ball of arbitrary charge has three terms,

$$E(Q') = \int d^3x \left(\frac{1}{2}|\dot{\phi}|^2 + \frac{1}{2}|\nabla\phi|^2 + V(\phi) \right). \tag{7.8}$$

The field oscillates in time as $e^{i\omega t}\bar{\phi}(x)$ where the frequency is related to the charge by

$$Q' = \frac{1}{2i}\int d^3x\, \phi^* \overleftrightarrow{\partial_t} \phi = \omega \int \phi^2\, d^3x. \tag{7.9}$$

After some manipulation, one can write [224]

$$\begin{aligned} E &= \int d^3x \left(\frac{1}{2}|\nabla\bar{\phi}|^2 + \hat{V}_\omega(\bar{\phi}) \right) + \omega Q', \\ &\equiv S_3[\bar{\phi}(x)] + \omega Q', \end{aligned} \tag{7.10}$$

where the first term is the three-dimensional Euclidean action of the "bounce" solution tunneling between the two minima of the effective potential $\hat{V}_\omega(\phi) = V(\phi) - \omega^2\phi^2/2$.

In our analysis below, we will need to describe the Q-balls for a variety of charges; in particular, we would like to know their mass and radius as a function of charge. For sufficiently large Q-balls we may use the well-known thin wall approximation; for smaller Q-balls, the thick wall approximation can be made. For the region between these two limits, we can use an interpolation to approximate the behavior of the Q-balls. However, there is a regime of very small charges in which even the thick wall approximation, which remains a semi-classical approximation, is inapplicable. We will use the Bethe-Salpeter equation to approximate the Q-ball properties at these small charges.

First, let us review the thin wall regime. In this regime, the energy of a Q-ball may be calculated in a different manner, due to [225]. Beginning again with Eq. (7.8), we use the oscillatory time dependence to write

$$\begin{aligned} E(Q') &= \frac{Q'^2}{2\int \bar{\phi}^2\, d^3x} + \int \frac{1}{2}\left(\nabla\bar{\phi}\right)^2\, d^3x + \int V(\bar{\phi})\, d^3x, \\ &= \frac{Q'^2}{2\int \bar{\phi}^2\, d^3x} + T + V, \end{aligned} \tag{7.11}$$

where $T = \int (\nabla\bar{\phi})^2\, d^3x/2$ and $V = \int V(\bar{\phi})\, d^3x$. In the thin wall approximation, $\bar{\phi} \approx \phi_0$ for $r < R - \delta/2$ and $\bar{\phi} \approx 0$ for $r > R + \delta/2$, where δ is the width of the surface of the Q-ball. The thin wall approximation is valid if $\delta \ll R$. The "volume" term then has two contributions, one from the interior of the Q-ball and one from the surface. In the interior of the Q-ball, the potential is $V(\phi_0)$, while in the surface of the Q-ball, it is $\beta m_0^2\phi_0^2$, where β is a positive constant. Therefore, the volume term is

$$V = \int V(\bar{\phi})\, d^3x = \frac{4}{3}\pi V(\phi_0) R^3 + 4\pi\delta R^2 \cdot \beta m_0^2\phi_0^2. \tag{7.12}$$

In the surface, the field changes by $\delta\bar{\phi} = \phi_0$ in the distance $\Delta r = \delta$, thus $d\bar{\phi}/dr \approx \phi_0/\delta$. Introducing a constant α to account for the uncertainty, the surface term is

$$T = 4\pi\delta R^2 \cdot \alpha\frac{\phi_0^2}{\delta^2}. \tag{7.13}$$

The energy must be a minimum with respect to both δ and R; minimizing with respect to δ gives $\delta = \sqrt{\alpha/\beta m_0^2}$ and

$$E = \frac{3Q'^2}{8\pi\phi_0^2 R^3} + 8\pi m_0\sqrt{\alpha\beta}\cdot R^2\phi_0^2 + \frac{4}{3}V(\phi_0)R^3. \tag{7.14}$$

By manipulating Eq. (7.10), one can relate $\sqrt{\alpha\beta}$ to the one-dimensional Euclidean action for the true potential

$$S_1 = \int_0^{\phi_0}\sqrt{2V(\phi)}\,d\phi = 2m_0\sqrt{\alpha\beta}\phi_0^2, \tag{7.15}$$

which is related to the three-dimensional action through $S_3 = 4\pi R^2 V(\phi_0)/3 + 4\pi R^2 S_1$ [224, 226]. Minimizing the energy with respect to R results in a constraint between the charge and the radius:

$$0 = -\frac{9Q'^2}{8\pi\phi_0^2} + 16\pi m_0\sqrt{\alpha\beta}\phi_0^2 R^5 + 4\pi V(\phi_0)R^6. \tag{7.16}$$

$V(\phi_0) > 0$, the last term dominates over the second term; neglecting this term allows us to solve for R in terms of the charge. This gives the familiar $R \propto Q^{1/3}$ behavior [202]. However, for Q-balls to induce a phase transition we must have $V(\phi_0) < 0$, and so one cannot neglect the second term. This sixth order equation has no closed form solution.

For the thick wall approximation, we consider Eq. (7.10). As the charge becomes small, the frequency ω becomes large; then the asymmetric minimum in $V_\omega(\phi)$ is significantly lower than the symmetric minimum. This is true independent of the sign of $V(\phi_0)$; in fact, since we have $V(\phi_0) < 0$, the asymmetric minimum is lower than the symmetric minimum even at $\omega = 0$. Therefore, the Euclidean action S_3 is the same in both cases, and the relations between the energy, radius, and charge are unchanged. These are [224]

$$E = Q'm_0\left(1 - \frac{\epsilon^2}{6} - \frac{\epsilon^4}{8} - \dots\right), \tag{7.17}$$

$$R^{-1} = \epsilon m_0\left(1 + \frac{1}{2}\epsilon^2 + \frac{7}{8}\epsilon^4 + \dots\right), \tag{7.18}$$

where $\epsilon = Q'A^2/3S_\psi m_0^2$ and $S_\psi \approx 4.85$ was determined numerically. This is valid when:

$$Q' \ll \frac{3S_\psi m_0}{\sqrt{\lambda}A}, \qquad Q' < \frac{3S_\psi m_0^2}{2A^2}. \tag{7.19}$$

The thick wall approximation, like the thin wall approximation, neglects quantum corrections; thus it breaks down when these are large, which occurs around $Q' \lesssim 7$ [227].

No general solutions are known for the regime between when the thin wall approximation is valid and when the thick wall approximation applies. Therefore, we use a linear interpolation between the two regimes. We will need only the radius in terms of the charge Q', for which we use

$$R = \frac{Q'-7}{Q'} R_{\text{thin}} + \frac{Q'}{7} R_{\text{thick}}. \tag{7.20}$$

We have already written R_{thick} in terms of Q'; however, as found above there is no closed-form equation for R_{thin} in terms of Q'. Therefore, we use a numerical approximation for the thin wall radius at small charge of the form $R \approx a + bQ'^{2/5}$. While this can only be justified by its numerical accuracy, it can be motivated by neglecting the third term in the constraint Eq. 7.16. Then in the intermediate regime the linear interpolation becomes

$$R = \frac{Q'-7}{Q'} \left(a + bQ'^{\,2/5}\right) + \frac{3S_\psi}{7A}. \tag{7.21}$$

The stability of states with very low charge is vital to building critically sized Q-balls, but for these states, one cannot use the approximations already discussed because quantum effects are important. At small charge, it is furthermore incorrect to think of the Q-balls as made of ϕ quanta; instead we should remember that they are bound states made of squarks and Higgs bosons. Unlike the full supersymmetric model discussed in Ch. 3, in our simplified model we have not given put the squarks or Higgs bosons in $\mathrm{SU_L}(2)$ doublets; nor have we assigned them hypercharge quantum numbers. Consequently, in this simple model, the squarks exchange Higgs bosons directly; the basic kernel is the single particle exchange that appears in the Wick-Cutkosky model, which has been discussed in some detail in Ch. 2. We may use this model in the limit that the Higgs mass is much lighter than the squark mass. As this approximation is not perfect, we use an effective coupling $\tilde{A}$ in our analysis, which we tune to ensure that the energy at large charge matches the result from the thick wall approximation.

In this section, we will use the label $\mathsf{n} = 3Q$, the number of squarks present in the state, to designate the bound states. At the end, we will relate this to the charge Q' that we have been using as a label in the other regimes using $Q' = 3Q/\sin^2(\theta)$, Eq. (7.7). The lowest state is a single squark, and then the first step is the relatively simple case of two individual squarks forming one $\mathsf{n} = 2$-ball; in this equal mass and equal coupling case the bound state masses given by the Bethe-Salpeter equation (in the Wick-Cutkosky model) are

$$M_n = 2m\left(1 - \frac{\alpha^2}{8n^2}\right) = 2m_0\left(1 - \frac{\tilde{A}^4}{2048\pi^2 m_0^4 n^2}\right), \tag{7.22}$$

if $\alpha = \tilde{A}^2/16\pi m_0^2 < 1$. We may find the binding energy for the ground state by taking $n = 1$.

For the remaining states, the masses and couplings at the top and bottom of the ladder are unequal; after a Wick rotation, the Bethe-Salpeter equation in the ladder approximation is

$$\left[(m+\Delta)^2 + (p - i\eta_1 P)^2\right]\left[(m-\Delta)^2 + (p + i\eta_2 P)^2\right]\Phi(p) = \frac{\lambda}{\pi^2}\int dq \frac{\Phi(q)}{(p-q)^2}, \tag{7.23}$$

where the masses of the particles, $m_{\rm top}$ and $m_{\rm bottom}$, are $m \pm \Delta$. We have discussed this case in Ch. 2; however, we will repeat the main results here. The coupling constant in the above equation is $\lambda = g_{\rm top} g_{\rm bottom}/16\pi^2 = (\mathsf{n}-1)\tilde{A}^2/16\pi^2$, where $g_{\rm top} = (\mathsf{n}-1)\tilde{A}$ is the coupling at the top of the ladder and $g_{\rm bottom} = \tilde{A}$ is the coupling at the bottom of the ladder. The total charge Q of the resulting Q-ball is $\mathsf{n}/3$. The energy-momentum four-vector of the bound state, P, is given by $(M, 0)$ where M is the bound state mass. η_1 and η_2 come from transforming to a "center of momentum" reference frame; these are

$$\eta_1 = \frac{m_{\rm top}}{m_{\rm top} + m_{\rm bottom}}, \qquad \eta_2 = \frac{m_{\rm bottom}}{m_{\rm top} + m_{\rm bottom}}. \tag{7.24}$$

The binding energies are [30]

$$M^2 = 4\Delta^2 + 4m^2\left(1 - \frac{\Delta^2}{m^2}\right)\left(1 - \frac{A'^2\pi^2}{4m^4}\frac{1}{(1-\Delta^2/m^2)^2}\right). \tag{7.25}$$

We use this equation iteratively to find the masses and binding energies of the small charge Q-balls; the results are shown in Table 7.1.

n (No. squarks)	Mass of Q-ball	Binding Energy
1	m_0	0
2	$2m_0 - 0.0000989\tilde{A}^4/m_0^3$	$0.0000989\tilde{A}^4/m_0^3$
3	$3m_0 - 0.0002659\tilde{A}^4/m_0^3$	$0.0001670\tilde{A}^4/m_0^3$
4	$4m_0 - 0.0005298\tilde{A}^4/m_0^3$	$0.0002639\tilde{A}^4/m_0^3$
5	$5m_0 - 0.0009163\tilde{A}^4/m_0^3$	$0.0002865\tilde{A}^4/m_0^3$
6	$6m_0 - 0.0014506\tilde{A}^4/m_0^3$	$0.0005343\tilde{A}^4/m_0^3$
7	$7m_0 - 0.0021577\tilde{A}^4/m_0^3$	$0.0007071\tilde{A}^4/m_0^3$

Table 7.1: Energies of small Q-balls from Bethe-Salpeter equation.

We have calculated this until $\mathsf{n} = 7$, or $Q = 7/3$. Using $\theta = 1.308$, this corresponds to $Q' = 7.51$ Since this is greater than 7, the thick wall approximation is applicable. In the thick wall regime, a Q-ball with this charge has energy

$$M = 7.51m_0 + 0.333A^4/m_0^3. \tag{7.26}$$

The difference in the first terms (about 7.2 percent) comes from the fact that states described by the Bethe-Salpeter equation do not have exactly the same proportion of squarks and Higgs bosons as the Q-balls described in the thick wall regime. This can be further improved by moving the global minimum closer to the $\tilde{q}$ axis. This is because

$$Q'm_0 = \mathsf{n}\sqrt{m_{\tilde{q}}^2 \sin^2(\theta) + m_h^2 \cos^2(\theta)/\sin^2(\theta)}, \tag{7.27}$$

as seen by comparing potentials (7.1) and (7.2), and using the relation between n and Q'.

Comparing the second terms in the mass equations gives $\tilde{A} = 4.51A$. Using the value of A in the potential in (7.2) gives an effective coupling of $\tilde{A} = 263$ GeV for the coupling between the squarks and the Higgs boson, while the value that we put into our original potential in (7.1) is 240 GeV. These differ by 9.6 percent; we attribute this difference to a combination these two factors: first, the difference in squark-Higgs proportions already mentioned above,

and the inherent inaccuracy of the ladder approximation in the Bethe-Salpeter equation. We note that the difference does not improve if we iterate the Bethe-Salpeter equation to larger charges. Eq. (7.25) for the bound state masses is valid provided that $\lambda \ll m^2$. Now that we have set $\tilde{A}$, and thus λ, we can verify that we satisfy this bound. For the largest A', at $\mathsf{n} = 7$, this ratio is approximately $\lambda/m^2 \approx 6\tilde{A}/(16\pi^2 \cdot 3.5^2 m_{\tilde{q}}^2) = 0.0053$.

7.3 Critical Values for the Phase Transition

Now that we have described the Q-balls in each of the four regimes, we next discuss the solitosynthesis-induced phase transition. In the thin wall regime, the interior of the Q-ball is nearly in the true vacuum; it is at least in a state which has negative energy density. If charge continues to increase, the Q-ball expands, converting more space into the negative energy state. At a particular value of the charge and radius, it expands uncontrollably, thereby converting all space into the negative energy state; following this, the system will roll down to its true vacuum [218]. This Q-ball-induced phase transition can occur even when such a phase transition cannot be induced by thermal fluctuations.

As will be demonstrated in our numerical example, the critical charge is of order 10^5, which is within the thin wall regime. At the critical point, not only is $dE/dR = 0$, but also $d^2E/dR^2 = 0$, which gives the additional constraint

$$0 = \frac{9Q_c'^2}{2\pi\phi_0^2} + 16\pi m_0\sqrt{\alpha\beta}\phi_0^2 R_c^5 - 8\pi V_0 R_c^6, \tag{7.28}$$

where $V(\phi_0) = -V_0$ with $V_0 > 0$ in the phase transition case. We solve the two constraint equations, (7.16) and (7.28), for the critical charge and radius:

$$R_c = \frac{10 m_0\sqrt{\alpha\beta}\phi_0^2}{3V_0}, \tag{7.29}$$

$$Q_c' = 2\pi\phi_0\sqrt{\frac{8V_0}{45}}\left(\frac{10 m_0\sqrt{\alpha\beta}\phi_0^2}{3V_0}\right)^3. \tag{7.30}$$

We aim to show that Q-balls of this size will grow from individual squarks.

7.4 Solitosynthesis

7.4.1 Solitosynthesis Temperature, Rate of Diffusion, and Freeze-Out

Now let us discuss the growth of Q-balls. In thermal equilibrium, the number density of Q-balls of a particular charge is given by a Saha equation:

$$n_{Q'} = \frac{g_{Q'}}{g_\phi g_{Q'-1}} n_\phi n_{Q'-1} \left(\frac{2\pi}{m_0 T}\right)^{3/2} e^{B_{Q'}/T}, \tag{7.31}$$

where $B_{Q'}$ is the binding energy of a soliton (Q-ball) of charge Q' and $g_{Q'}$ is the internal partition function of the soliton. g_ϕ is the number of degrees of freedom associate with the ϕ field. While we have chosen to work with real fields which are not arranged in $\mathrm{SU_L}(2)$ doublets as a simplification, we will set $g_\phi = 3$, to account for the possible color states, to make the model somewhat more realistic. (Taking $g_\phi = 1$ makes only minor changes to the analysis.) The charge density n_ϕ is the number of free squarks; since $\theta \approx \pi/2$, conceptually we can think of this as the number of ϕ-quanta (1-balls) present. This is given by

$$n_\phi = \eta n_\gamma - \sum_{Q' \geq 2} Q' n_{Q'}, \tag{7.32}$$

where the baryon asymmetry is η and in the radiation-dominated era the photon density is $2\zeta(3)T^3/\pi^2$.

The typical approach would be to solve these coupled equations numerically. However, the critical charge is generically of order 10^3 to 10^5, which leads to at least 10^3 coupled equations. It is infeasible to solve these simultaneously. Therefore, we take a different approach following [228] and consider the evolution of a single Q-ball. We will see that the Q-balls grows sufficiently fast that we can ignore charge depletion; then

$$n_\phi \approx \eta n_\gamma = \eta \frac{2.404 T^3}{\pi^2}. \tag{7.33}$$

A single Q-balls grows or shrinks according to

$$\frac{dQ'}{dt} = r_{\mathrm{abs}}(Q') - r_{\mathrm{evap}}(Q'), \tag{7.34}$$

where $r_{\rm abs}$ is the absorption rate and $r_{\rm evap}$ is the evaporation rate; by detailed balance,

$$n_{Q'} r_{\rm abs}(Q') = n_{Q'+1} r_{\rm evap}(Q'+1). \tag{7.35}$$

The rate of absorption is $r_{\rm abs}(Q') n_\phi v_\phi \sigma_{\rm abs}(Q')$, where $\sigma_{\rm abs}(Q')$ is the cross section for a Q'-charged Q-ball to absorb a ϕ quanta. For large charges, $\sigma_{\rm abs}(Q') \approx \pi R^2$. As we will verify numerically below, the radius does not change rapidly as a function of charge, at least in the large charge regime. Consequently, $\sigma_{\rm abs}(Q') \approx \sigma_{\rm abs}(Q'-1)$, and together, these approximations yield

$$\begin{aligned} \frac{dQ'}{dt} &= r_{\rm abs}(Q') - \frac{n_{Q'-1}}{n_{Q'}} r_{\rm abs}(Q'-1), \\ &\approx n_\phi v_\phi \sigma_{\rm abs}(Q') \left(1 - \frac{n_{Q'-1}}{n_{Q'}}\right). \end{aligned} \tag{7.36}$$

Thus the determining factor is $n_{Q'-1}/n_{Q'}$: if this is less than 1, absorption dominates, but if it is greater than 1, evaporation dominates. From the Saha equations (7.31), this important ratio is

$$\frac{n_{Q'-1}}{n_{Q'}} = g_\phi \frac{\pi^2}{2.404\eta} \left(\frac{2\pi T}{m_0}\right)^{-3/2} e^{-B_{Q'}/T}, \tag{7.37}$$

where we have used $g_{Q'} \approx g_{Q'-1}$ since our Q-balls are large objects made of scalar particles (and hence have no internal spin degrees of freedom). At large temperatures, the exponential is negligible and this scales as $T^{3/2}$. However, this ratio is less than 1 only if $T > \eta^{-2/3} m_0$, which is typically quite large due to the smallness of η; it may even be above the reheat temperature. Therefore, relatively quickly in the evolution of the universe, we expect evaporation to dominate.

As the temperature decreases, the exponential term is no longer negligible. Because $B_{Q'} < 0$, this term decreases $n_{Q'-1}/n_{Q'}$. Therefore, at some solitosynthesis temperature T_S absorption will dominate and the Q-ball will grow; this temperature is determined by

$$T_S = \frac{B_{Q'}}{\ln(g_\phi) - \ln(\eta) + (3/2)\ln(m_0/T_S) - 1.34}. \tag{7.38}$$

It is possible for the binding energy to be sufficiently large that solving this equation for temperature results in an imaginary value; returning to Eq. (7.37), this happens when $B_{Q'}$ is so large that the right-hand side is always less than 1, which entails $n_{Q'} > n_{Q'-1}$ always. Physically, whenever a $(Q'-1)$-ball forms, it will always grow into a larger ball; we may say that the solitosynthesis temperature for these charges is infinite. Smaller charges have smaller binding energies, and so we will find finite solitosynthesis temperatures for small Q-balls. Therefore, we will need to wait for these smaller Q-balls to form, and then wait for these to grow into the larger ones which can always grow. As this suggests, $T_S s$ is greater for larger charges, which we will verify numerically. Therefore, Q-ball growth is a winner-take-all situation, and the solitosynthesis temperature cannot cut off a growing Q-ball.

The above calculation tells us when it is more favorable for a Q-ball to grow rather than evaporate charge. However, since a Q-ball grows by absorbing the nearby charge, if the nearby charge is not be replenished sufficiently quickly through diffusion, there may be a local depletion of charge near the Q-ball which limits its growth. If this occurs, the rate of growth will be given by $r_{\rm diff}$, the rate that free squarks diffuse into the surface of the Q-ball, instead of $r_{\rm abs}$. Ref. [229] is concerned with the related process of the diffusion of evaporating squarks away from a Q-ball. The particle flux through the Q-ball surface is given by

$$r_{\rm diff} = \frac{dQ'}{dt} = -4\pi k R D n_\phi^{\rm eq}, \tag{7.39}$$

where $D \approx aT^{-1}$, $a \approx 4$ for relativistic squarks, and $k \approx 1$ was determined numerically. We need to adjust this equation because we are concerned with particles diffusing towards the Q-ball; the rate has the opposite sign and we multiply this by the velocity of the nonrelativistically squarks. Thus,

$$r_{\rm diff} = v_\phi 16\pi R T^{-1} n_\phi^{\rm eq}. \tag{7.40}$$

This gives the ratio $r_{\rm diff}/r_{\rm abs} = 4T^{-1}/R$. Perhaps surprisingly, this is small for high temperatures and large for low temperatures. This occurs because the rate of diffusion is

proportional to $T^{5/2}$ while the rate of absorption is proportional to $T^{7/2}$. Even though diffusion decreases as the temperature decreases, the rate of absorption drops faster; therefore, diffusion will limit the growth of Q-balls for temperatures above $4/R$. For radii of order 0.01 GeV^{-1} to 0.1 GeV^{-1}, this temperature is of order 40 to 400 GeV, which is significantly above the solitosynthesis temperatures. Therefore, diffusion will replenish the charge sufficiently quickly at the relevant temperatures for Q-ball growth. This, combined with the winner-take-all behavior, demonstrates that global depletion of charge is not an issue, provided that most of the charge is in free squarks during solitosynthesis. We will verify this last point numerically.

The growth of a Q-ball can also be ended (prior to it reaching the critical size) by the freeze-out of the accretion interactions. The reactions responsible for Q-ball growth freeze-out when their time scale is greater than the Hubble time scale, $\tau_H = H^{-1}$. While the universe is radiation dominated, the Hubble parameter is T^2/M_{Pl}, and so the Hubble time scale is $\tau_H = 2.43 \cdot 10^{18}\ \text{GeV}/T^2$. The time scale of Q-ball growth is $\tau_{\text{abs}} = 1/r_{\text{abs}} = 1/n_\phi \sigma v_\phi$. We consider the later reactions in the sequence; then the heavy Q-balls are effectively at rest and the $Q' = 1$-balls are moving nonrelativistically in thermal equilibrium, with $v_\phi \approx 2\sqrt{2T/\pi m_0}$. Use the geometric cross section $\sigma = \pi R^2$ yields

$$\tau_{\text{abs}} = \frac{\pi^2}{4.808 \eta T^3 R^2} \sqrt{\frac{2T}{m_0}}. \tag{7.41}$$

Setting these time scales equal and solving for T gives

$$T = \left(\frac{\sqrt{2}\pi^2}{2.43 \cdot 10^{18}\ \text{GeV} \cdot 4.808\eta} \cdot \frac{1}{R^2 \sqrt{m_0}} \right)^2 \tag{7.42}$$

When we do our numerical analysis below, we will find that these are orders of magnitude smaller than the temperatures relevant to the phase transition.

Thus, we have determined the temperature at which a Q-ball of particular charge begins growing, for sufficiently large charges. (Small charge Q-balls will be discussed separately

below.) We have also determined the rate of growth as a function of radius; since this is different in each of the four regimes considered, we will now focus on each regime separately.

7.4.2 Rates of Growth

Our goal is to show that we can grow a critically-sized Q-ball in the early universe. Therefore, we must find the rate of growth for a single Q-ball in each of the four regimes. For the thin wall, thick wall, and intermediate regimes, we have sufficiently large charge that we can use the absorption rate discussed above. For the Bethe-Salpeter regime, a different analysis will be used.

Let us begin with the thin wall regime. For temperatures below the solitosynthesis temperature, the rate of evaporation is small, and we may approximate $dQ'/dt = n_\phi v_\phi \sigma_{\rm abs}(Q')$ from Eq. (7.36); we note that we have large charges in the thin wall regime, and so we can use the geometrical area for the cross section. Additionally, since charge depletion is negligible, we may assume $n = \eta n_\gamma$. The $Q' = 1$-balls being absorbed are in thermal equilibrium at $T \ll m_0$ with average velocity $v_\phi = 2\sqrt{2T/\pi m_0}$. (As we have noted, an individual ϕ quanta is almost interchangeable with individual squarks for $\theta \approx \pi/2$.) In the radiation-dominated era, the temperature and the time are not independent; they are related by

$$t = \frac{1}{T^2}\sqrt{\frac{3}{16\pi G\mathcal{N}}} + \text{const}, \tag{7.43}$$

where $\mathcal{N}$ is the effective number of degrees of freedom of the particles in thermal equilibrium, with fermionic degrees of freedom weighted by 7/8. In our toy model with only the Higgs boson, squarks, and photons, we have $\mathcal{N} = 6$. Then we have

$$dt = -1.34 \cdot 10^{18}\ \text{GeV}\, dT/T^3. \tag{7.44}$$

Thus, the differential equation for the growth of one Q-ball as a function of temperature is

$$-\frac{1}{1.34 \cdot 10^{18}\ \text{GeV}}\frac{dQ'}{dT} = \pi R^2 \eta \frac{4.808}{\pi^2}\sqrt{\frac{2T}{\pi m_0}}, \tag{7.45}$$

using $n_\gamma = 2.404T^3/\pi^2$.

The right-hand side involves the radius which is not independent of the charge; however, in the thin wall approximation the radius cannot be written in terms of the charge because of the form of the 6th order equation relating them, given by Eq. (7.16). Fortunately, one can write the charge in terms of the radius, and then we consider the rate of the growth of the radius of the Q-ball until it reaches the critical radius:

$$\begin{aligned} Q' &= \sqrt{\frac{128\pi^2 m_0\sqrt{\alpha\beta}\phi_0^4}{9}R^5 + \frac{32\pi^2 U_0\phi_0^2}{9}R^6}, \\ &\equiv \sqrt{a_6R^6 + a_5R^5}. \end{aligned} \tag{7.46}$$

This gives

$$dQ' = \frac{6a_6R^5 + 5a_5R^4}{2\sqrt{a_6R^6 + a_5R^5}}\,dR = R^{3/2}\frac{a_6R + a_5}{2\sqrt{a_6R + a_5}}, \tag{7.47}$$

and the resulting differential equation is

$$R^{-1/2}\frac{6a_6R + 5a_5}{2\sqrt{a_6R + 5a_5}}\,dR = -\frac{1.64\cdot 10^{18}\text{ GeV}\eta}{\sqrt{m_0}}T^{1/2}\,dT. \tag{7.48}$$

Both sides of this equation can be integrated explicitly,

$$\begin{aligned} &3\sqrt{a_6R^2 + a_5R} + \frac{a_5}{\sqrt{a_6}}\ln\left(\frac{a_5 + 2a_6R + 2\sqrt{a_6R\,(a_6R + a_5)}}{2\sqrt{a_6}}\right) - 3\sqrt{a_6R_i^2 + a_5R_i} \\ &- \frac{a_5}{\sqrt{a_6}}\ln\left(\frac{a_5 + 2a_6R_i + 2\sqrt{a_6R_i\,(a_6R_i + a_5)}}{2\sqrt{a_6}}\right) = \frac{2}{3}\cdot\frac{1.64\cdot 10^{18}\eta\text{ GeV}}{\sqrt{m_0}}\left(T_{\text{start}}^{3/2} - T^{3/2}\right), \end{aligned} \tag{7.49}$$

where R_i is the radius of the smallest Q-ball at which the thin wall approximation is valid and T_{start} is the temperature at which this Q-ball starts to grow. This can be less than T_s (for the Q' corresponding to R_i) if these Q-balls do not form until a lower temperature. If we set $R = R_c$, this equation can be solved for the temperature at which the Q-ball becomes critically sized.

Now we turn our attention to the thick wall regime. We begin with the differential equation (7.45) which is valid in the thick wall regime also. We use $R = 3S_\psi m_0/Q'A^2$ to write the

differential equation as

$$\frac{dQ'}{dT} = -1.47 \cdot 10^{19}\ \text{GeV} \frac{S_\psi^2 m_0^{3/2} \eta}{Q'^2 A^4} \sqrt{T}, \tag{7.50}$$

whose solution is

$$Q_f'^3 - Q_i'^3 = 2.95 \cdot 10^{19}\ \text{GeV} \frac{S_\psi^2 m_0^{3/2} \eta}{A^4} \left(T_{\text{start}}^{3/2} - T_f^{3/2} \right), \tag{7.51}$$

where T_{start} is the starting temperature for thick wall growth. This is either the $Q' = 7$ solitosynthesis temperature, or the temperature at which $Q' = 7$-balls form, whichever is smaller.

Now that we have addressed both the thin and thick wall regimes, let us consider the intermediate regime. Again, we begin with the differential equation (7.45), but this time we use the linear interpolation for the radius, Eq. (7.21), which yields

$$\int_7^{Q_f'} \frac{Q'^2\, dQ'}{((Q'-7)(a+bQ'^{2/5}) + 21 S_\psi / Q'^2 A^2)^2} = \frac{1.09 \cdot 10^{18}\ \text{GeV} \eta}{\sqrt{m_0}} \left(T_{\text{start}}^{3/2} - T_f^{3/2} \right). \tag{7.52}$$

The left-hand side of this equation must be integrated numerically.

Now we turn to the primary issue addressed in this chapter, the growth of very small Q-balls. In this regime, cross sections cannot be approximated by the geometrical area and so Eq. (7.45) is not valid. However, because these are the first steps of solitosynthesis, nearly all of the charge will be in these lowest seven states (which we again label with n, the number of squarks present in the state). Therefore, one can return to the initial method of considering the evolution of the number densities as a function of temperature; we have 8 equations to solve numerically, instead of 10^5.

The number densities of the Q-balls are given by the Saha equations like (31), which we write in terms of fractional densities $X_\mathsf{n} = n_\mathsf{n} \mathsf{n}/N$, where N is the total number of squarks,

$\eta \cdot 2.404T^3/\pi^2$. We find

$$\begin{aligned}
\frac{n_{\mathsf{n}}}{N} &= \frac{n_1}{g_\phi}\frac{n_{\mathsf{n}-1}}{N}\left(\frac{2\pi}{m_0 T}\right)^{3/2} e^{B_{\mathsf{n}}/T}, \\
\frac{X_{\mathsf{n}}}{\mathsf{n}} &= \frac{N}{g_\phi} X_1 \frac{X_{\mathsf{n}-1}}{\mathsf{n}-1}\left(\frac{2\pi}{m_0 T}\right)^{3/2} e^{B_{\mathsf{n}}/T}, \\
X_{\mathsf{n}} &= \frac{\mathsf{n}}{\mathsf{n}-1}\frac{2.404\eta}{3\pi^2}\left(\frac{2\pi T}{m_0}\right)^{3/2} X_{\mathsf{n}-1} X_1 e^{B_{\mathsf{n}}/T},
\end{aligned} \tag{7.53}$$

where we have used $X_1 = n_1/N$ and $g_\phi = 3$. We also have the additional constraint $X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7 = 1$.

Note that it is not necessary for every Q-ball to reach the critical size to induce the phase transition; only one per Hubble volume, $1/H^3$ where $H = T^2/2.43 \cdot 10^{18}$ GeV, is needed. Therefore, as a conservative approach, we numerically find temperature at which there are of order 10 $\mathsf{n} = 7$-balls per Hubble volume and begin the thick wall analysis at this temperature. We also verify that at this temperature most of the charge remains in $\mathsf{n} = 1$-balls (free squarks); otherwise, the analysis above is invalid because we assumed negligible charge depletion.

Due to the exponential, X_7 is large at low temperatures; it is also large at high temperatures due to the $T^{3/2}$ factor. In between these two extremes it reaches a minimum. (This description of the behavior is identical to the balancing of evaporation and absorption we discussed above for larger charge values.) Typically, the solitosynthesis temperature for $Q' = 7$-balls is after the number of $\mathsf{n} = 7$-balls per Hubble volume has dropped beneath 1; that is, most of the $\mathsf{n} = 7$-balls have evaporated away. Then we need to wait until $\mathsf{n} = 7$-balls form again at lower temperatures before any Q-balls may grow to the critical size. However, there exist scenarios in which the number of $\mathsf{n} = 7$-balls per Hubble volume is still greater than 1 when the universe cools to the solitosynthesis temperature of $Q' = 7$-balls; in this case, they may begin to grow immediately.

In such cases, our analysis underestimates the temperature at which the phase transition occurs. Since larger Q-balls begin accreting charge earlier, it is likely that there are even larger Q-balls that have not evaporated away when the temperature reaches their (higher) solitosynthesis temperature. If one of these $\mathsf{n} = 7$-balls has time to induce a phase transition, then we can be certain that any larger ones that had the opportunity to grow earlier would also induce a phase transition; thus, in these cases our analysis establishes only that the phase transition occurs. It should be noted that in these scenarios it is especially important to verify that most of the charge is in $\mathsf{n} = 1$-balls. (We will discuss this more below.)

7.5 Numerical Analysis

Finally, we demonstrate that the potential above is one in which all of these processes can work together to result in a phase transition.[1] As an reminder, the numbers selected above give $m_0 = 193$ GeV, $A = 58.2$ GeV, and $\lambda = 0.0450$, which gives a potential where the thin wall approximation is valid for large charge. Above, we also found that the value of ϕ inside the Q-ball is $\phi_0 = 646$ GeV; at this value, the potential is $-V_0$ with $V_0 = 6.18 \cdot 10^7$ GeV^4.

The other constant that must be set is η, the baryon asymmetry. In the actual universe, this is about $5 \cdot 10^{-10}$; to induce baryogenesis via this phase transition, one must of course take $\eta = 0$.[2] In this chapter, our emphasis is on demonstrating that such a phase transition, induced through Q-balls accreting charge, is a theoretical possibility even when small charge states are included in the analysis. We are not trying to establish that such a phase transition definitely happened during the evolution of our universe; we leave the problem of building a phenomenologically acceptable model for a later work, although we will discuss some issues regarding MSSM phenomenology below. Therefore, in this first numerical example, we will take $\eta = 8 \cdot 10^{-7}$. This value illustrates the phase transition well, although the final

[1] Note that I have chosen a slightly different set of parameters than in Phys. Rev. D **85** 125022.

[2] In these models, the Q-balls are typically formed via the fragmentation of an Affleck-Dine condensate and not through charge accretion, e.g., [100, 208–212]

Charge Q'	Radius (GeV^{-1})	$B_{Q'}(\text{GeV})$	T_s (GeV)
$Q_c = 5.36 \cdot 10^5$	0.341	$9.31 \cdot 10^7$	∞
5000	0.0375	$6.20 \cdot 10^5$	$1.72 \cdot 10^5$
1000	0.0195	$9.60 \cdot 10^4$	$1.28 \cdot 10^4$
500	$\sim$0.0148	$4.46 \cdot 10^3$	3950
200	$\sim$0.0102	$1.03 \cdot 10^3$	932
7	0.118	0.482	0.0155

Table 7.2: Solitosynthesis temperatures for several charge values. Observe that because the temperature rises with increasing charge, it cannot cut off a growing Q-ball. The top three were calculated in the thin wall regime, while the last one was calculated in the thick wall approximation. The other two are in the intermediate regime; to approximate their solitosynthesis temperatures, we used the thin wall result. By infinity, we mean that such a Q-ball always grows.

temperature will be below the QCD confinement temperature; therefore, we will also give a second, although more complicated, numerical example which avoids this.

With these parameters, the critical charge is $5.36 \cdot 10^5$ and the critical radius is 0.341 GeV^{-1}. We present a table of the radii and solitosynthesis temperatures for various charges in Table 7.2. With these radii and our chosen value of η, the freeze-out temperature is of the order 10^{-24} GeV, which is significantly smaller than any of the temperatures we will consider.

We begin with solitosynthesis in the Bethe-Salpeter regime. At the $Q' = 7$ solitosynthesis temperature, there are of order 10^{-3} n = 7-balls per Hubble volume. Therefore, we need to wait until these small Q-balls form again at low temperatures before thick wall growth can begin. Numerically, we find that there are order 10 $\mathsf{n} = 7$-balls at $T = 0.0121$ GeV.

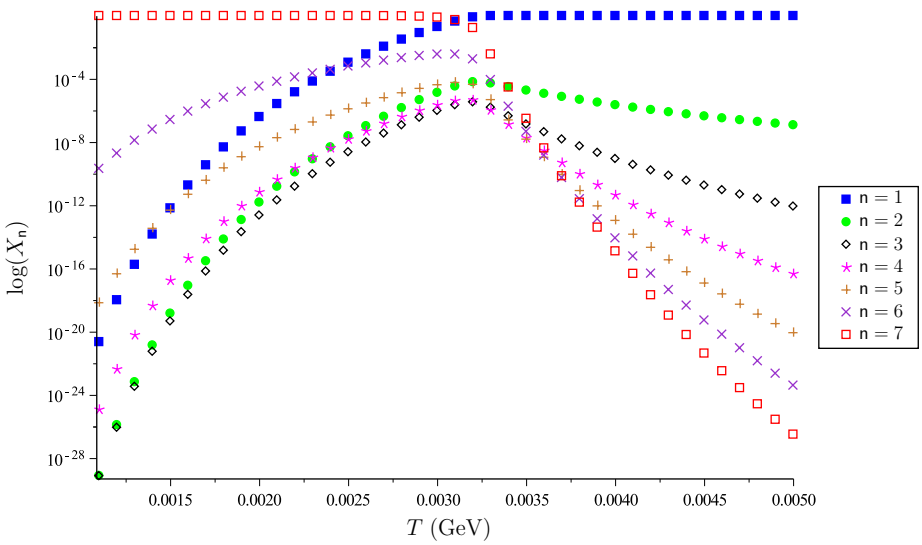

Figure 7.3: Fractional densities of small Q-balls, which shows the growth of very small Q-balls. Notice that as temperature decreases, the number of $\mathsf{n} = 7$-balls increases, while the number of $\mathsf{n} = 1$-balls decreases.

Next we address charge depletion. At the starting temperature of 0.0121 GeV, less that $4 \cdot 10^{-10}$ of the charge is in any of the bound states; free squarks are dominant. However, we also need to know what is the lowest temperature for which this assumption is valid; if our final temperature is beneath this, then our analysis is untrustworthy. To approximate this, we consider at what temperature the majority of the charge is no longer in $\mathsf{n} = 1$-balls (individual squarks) if we ignore all of the states above $\mathsf{n} = 7$. These charge densities are shown in Fig. 7.3. We see that the majority of the charge is no longer in individual squarks around $T = 0.0031$ GeV. As long as our final temperature is above this, we are justified in neglecting charge depletion. (As has been noted noted, charge diffuses fast enough that local charge depletion is not problematic.)

The thick wall approximation is valid until $Q' \approx 80$; the two constraints (7.19) give $Q' < 80$ and $Q' < 228$. The growth in this regime is relatively rapid. The growth in this regime is virtually instantaneous; the temperature drops by less than 1 part in 10^9. Thus, the starting

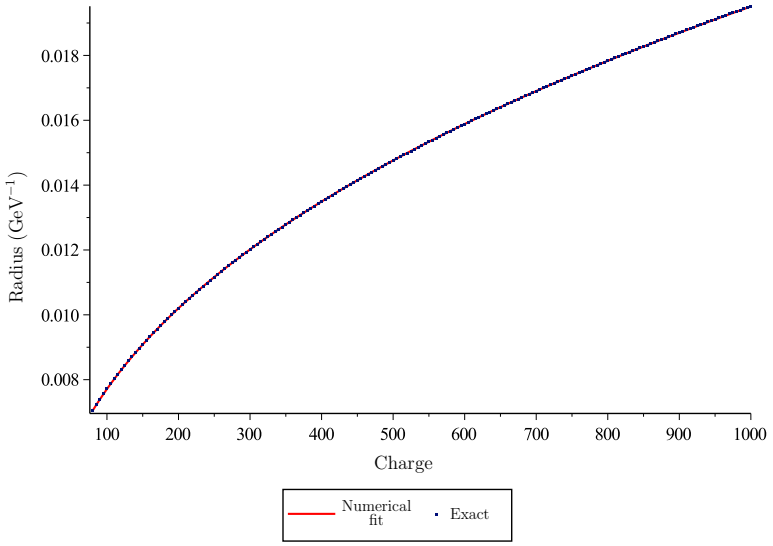

Figure 7.4: Numerical fit for the radius as a function of charge Q' in the thin wall approximation, for small charges. This fit is used in the interpolation in the intermediate regime.

temperature is still 0.0154 GeV for growth in the intermediate regime. The thin wall regime becomes applicable for $Q' \gg 228$; therefore, we will use the intermediate regime for charges between 80 and 1000. First we determine the constants a and b for Eq. (7.21); we numerically fit the function $a + bQ'^{2/5}$ to the radius for small values of Q'. The result is plotted in Fig. 7.4; this yields $a = 9.21 \cdot 10^{-5}$ and $b = 1.24 \cdot 10^{-3}$. Substituting this into differential equation (7.52) and solving numerically gives $T_f = 0.0116$ GeV for the temperature when the Q-ball reaches $Q' = 1000$. This is less than the solitosynthesis temperature for such a Q-ball (which is larger for larger charge), so thin wall growth begins immediately.

Before proceeding to the thin wall analysis, we first verify that $R_{Q'+1} \approx R_{Q'}$, as was assumed in the derivation of Eq. (7.49) (and also the solitosynthesis temperature). The mass (energy) and radius for the thin wall approximation are shown in Fig. 7.5; we see that this approximation is reasonable. Using $T = 0.0116$ GeV as the starting temperature in Eq. (7.49), we find that the Q-ball grows to critical size at $T = 0.00550$ GeV. This temperature is greater than 0.0031 GeV, and furthermore, it is greater than the freeze-out temperature

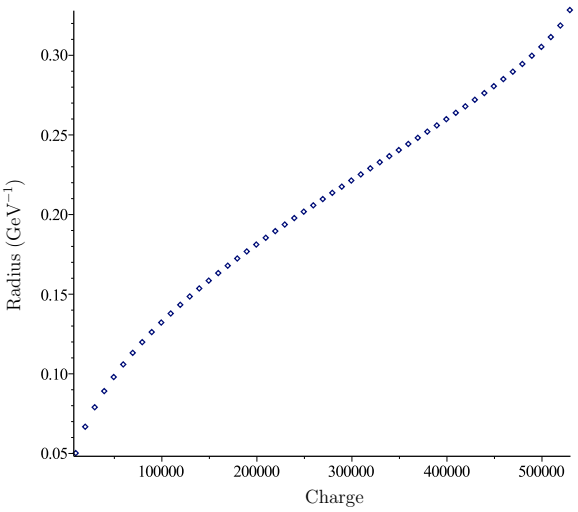

Figure 7.5: Radii as a function of charge in the thin wall regime.

scale; therefore, we conclude that such phase transitions are indeed theoretically possible.

As mentioned, the previous numerical example is not strictly speaking acceptable, because the final temperature is less than the QCD scale, $\Lambda = 0.217$ GeV, at which confinement introduces additional complications. Therefore, we consider a second example, which avoids this problem. We use the same potential as above; thus m_0, A, and λ are unchanged. The critical charge and radius are also unchanged, as are the ranges where the thick and thin wall approximations are applicable. Similarly, the numerical fit for the radius at small charges used in the intermediate regime is unchanged.

We will, however, choose an exceptionally large $\eta = 3 \cdot 10^3$. Then the solitosynthesis temperature for $Q' = 0.258$ GeV, above the QCD scale. Furthermore, at this temperature, there are still of order 10^{55} $\mathsf{n} = 7$-balls per Hubble volume; thick wall growth may begin immediately. This is one of the exceptional cases mentioned at the end of Sec. 7.4.2; even though these Q-balls are evaporating away, there are sufficiently many of them for growth to begin immediately. For clarity, we plot the number density of $\mathsf{n} = 7$-balls per Hubble volume in Fig. 7.6 to show it has the expected behavior. Furthermore, at this temperature, around

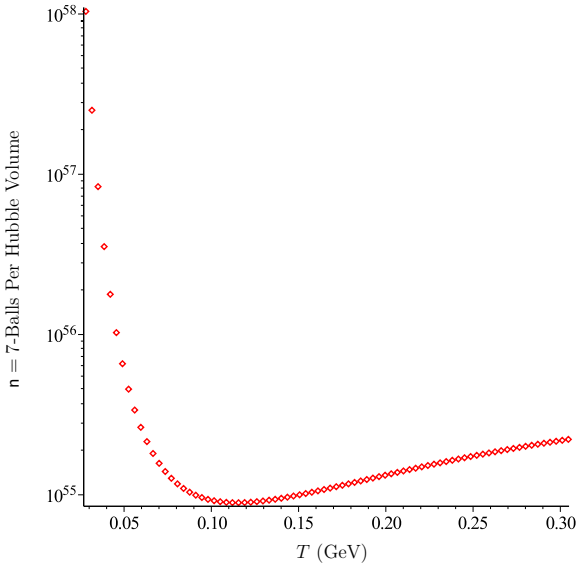

Figure 7.6: Number of $\mathsf{n} = 7$-balls per Hubble volume as a function of temperature.

65 percent of the charge is in individual squarks, and so we may ignore the charge depletion if the phase transition occurs sufficiently rapidly. (The large percentage of charge in squarks should not be surprising, since the larger Q-balls are in the process of evaporating.)

We will show that these Q-balls can grow into a critically charged Q-ball and induce the desired phase transition. However, as emphasized in Sec. 7.4.2, this underestimates the temperature at which the phase transition occurs. This is because larger Q-balls may likewise have evaporated sufficiently slowly that there are still sufficiently many of them to induce the phase transition when they begin accreting charge at their larger solitosynthesis temperatures. Since there are so many 7-balls at their solitosynthesis temperature, this is extremely probable. If larger Q-balls began accreting charge earlier, they would induce the phase transition before the 7-balls we analyze here. However, if these 7-balls can induce a phase transition, then we can be certain that any larger ones that began growing earlier would also induce the phase transition, and so the phase transition certainly occurs.

As above, the thick wall growth is virtually instantaneous; there is no appreciable drop in the temperature. Similarly, the growth in the intermediate regime is also extremely fast; the temperature changes less than 1 part in 10^9 . Finally, growth in the thin wall approximation is also extremely fast; again the temperature changes by less than 1 part in 10^9. Thus, such a Q-ball becomes critically charged within a temperature change of 10^{-9} GeV, and so the final temperature is still above the QCD scale. Again, though, we can only conclude that the phase transition occurs, and not that it occurs at this temperature (chosen to be near the QCD scale). This is why we additionally presented the first numerical example, which is a more typical case in which our analysis also determines the phase transition temperature.

7.6 Potential Applications to the MSSM

While the theoretical possibility of such a phase transition is in itself interesting, one would also like to know whether such a phase transition could occur in extensions of the Standard Model such as the MSSM, which naturally provides squarks carrying baryon number. This analysis suggests that, provided that the requisite vacuum structure can be found, such phase transitions are indeed possible. Indeed, one can ask whether such a phase transition is possible within the evolution of our own universe. If squarks do exist, they must be significantly heavier than quarks. Therefore, they will decay rapidly, and thus we cannot build critically charged Q-balls out of them in our current vacuum. However, such a phase transition could have occurred in the past, if the vacuum structure has these requisite properties:

1. A global minimum in which no squarks or sleptons develop vacuum expectation values, so that baryon number and lepton number are conserved. (This minimum corresponds to our current vacuum.)

2. A local minimum in which no squarks develop vacuum expectation values, so that baryon number is conserved.

3. In the local minimum, quarks must be heavier than squarks, so that they are stable

against decay into quarks.

4. In the local minimum, one of the bosons that mediates an interaction between squarks must be lighter than the squarks; this is required for bound states to develop.

5. The potential expanded in the local minimum must allow the creation of Q-balls through the squark fields.

6. The barrier between the local minimum and the global minimum must be sufficiently large to suppress tunneling between the minima by thermal fluctuations.

There will be additional complications in the Bethe-Salpeter stage of the analysis from the crossed-graph kernel, discussed in Ch. 3; however, the techniques of that chapter can be extended to cover this scenario. Therefore, our primary concern is about the vacuum structure of the theory.

Such a vacuum structure can indeed be found; as an example, consider an MSSM potential of the form

$$\begin{aligned} V = &-m_h^2 H^* H + m_{\tilde{Q}}^2 \tilde{Q}^* \tilde{Q} + m_{\tilde{q}}^2 \tilde{q}^* \tilde{q} + m_{\tilde{L}}^2 \tilde{L}^* \tilde{L} + m_{\tilde{l}}^2 \tilde{\ell}^* \tilde{\ell} + \frac{\lambda}{4}(H^* H)^2 - A_S \left(H \tilde{Q}^* \tilde{q} + H^* \tilde{Q} \tilde{q}^* \right) \\ &- A_L \left(H \tilde{L}^* \tilde{l} + H^* \tilde{L} \tilde{l}^* \right) + y^2 \left(H^* H \tilde{Q}^* \tilde{Q} + H^* H \tilde{q}^* \tilde{q} + \tilde{Q}^* \tilde{Q} \tilde{q}^* \tilde{q} \right) \\ &+ y^2 \left(H^* H \tilde{L}^* \tilde{L} + H^* H \tilde{l}^* \tilde{l} + \tilde{L}^* \tilde{L} \tilde{l}^* \tilde{l} \right) + \frac{g_1^2}{8} \left(H^* H - \tilde{Q}^* \tilde{Q} \right)^2 + \frac{g_1^2}{8} \left(H^* H - \tilde{L}^* \tilde{L} \right)^2 \\ &+ \frac{g_2^2}{8} \left(H^* H + \tilde{Q}^* \tilde{Q} - 2 \tilde{q}^* \tilde{q} \right)^2 + \frac{g_2^2}{8} \left(H^* H + \tilde{L}^* \tilde{L} - 2 \tilde{l}^* \tilde{l} \right)^2, \end{aligned} \tag{7.54}$$

where $\tilde{Q}$ and $\tilde{q}$ are squarks, $\tilde{L}$ and $\tilde{\ell}$ are sleptons, and H is a Higgs boson. One local minimum of this potential is at $\left\langle \tilde{Q} \right\rangle = \langle \tilde{q} \rangle = \left\langle \tilde{L} \right\rangle = \left\langle \tilde{\ell} \right\rangle = 0$ and $\langle H \rangle = m_h / \sqrt{\lambda}$. For these values of the coupling constants:

$$\begin{aligned} &m_{\tilde{L}} = m_{\tilde{l}} = 10\sqrt{2}\ \text{GeV}, && m_H = 0.5\sqrt{2}\ \text{GeV}, \\ &m_{\tilde{Q}} = m_{\tilde{q}} = 15\sqrt{2}\ \text{GeV}, && A_L = 23\ \text{GeV}, \\ &A_S = 31\ \text{GeV}, && \lambda = 0.006, \\ &g_1 = g_2 = 0.6, && y = 1, \end{aligned} \tag{7.55}$$

the global minimum is at the minimum mentioned above, while a local minimum occurs at $\left\langle \tilde{Q} \right\rangle = \langle \tilde{q} \rangle = 0$, $\left\langle \tilde{L} \right\rangle = 5.025$ GeV, $\left\langle \tilde{l} \right\rangle = 5.136$ GeV, and $\langle H \rangle = 8.292$ GeV. Since the squarks do not acquire a VEV in this vacuum, baryon number is conserved and can be used to construct Q-balls.

In the false vacuum, the lightest squark $\tilde{q}'$ has a mass $m_{\tilde{q}'}$ of 7.80 GeV. Because lepton number is not conserved, the sleptons mix with the Higgs boson; the lightest of these eigenstates is h' with a mass of 6.83 GeV. We assume that the quark acquires a mass from the term $y\bar{q}Hq$ in the Lagrangian; then it has mass $y \langle H \rangle = 8.29$ GeV, and so the lightest squark is stable against decay into a quark.

The potential along the line connecting the minima is

$$V(\phi) = .324\phi^4 - 7.38 \text{ GeV } \phi^3 + 41.2 \text{ GeV}^2 \phi^2, \tag{7.56}$$

which is shown in Fig. 7.7. We notice that the barrier separating the local minimum from the global minimum is quite large, which dramatically suppresses tunneling through thermal fluctuations. By expanding the potential in terms of the appropriate eigenstates in the false vacuum, one can show that Q-balls constructed of $\tilde{q}'$ and h' fields do exist. Thus we have all of the necessary ingredients for a solitosynthesis-induced phase transition.

While most of our analysis could be straightforwardly applied to this model, the main contribution of the chapter, the analysis of the Bethe-Salpeter regime, cannot. We have noted in Sec. 7.2 that the field content of the states described by the Bethe-Salpeter equation does not necessarily match the field content of Q-balls described by the thick wall approximation. In our numerical example, we made the difference small by choosing the global minimum such that $\sin(\theta) \approx 1$. However, in this scenario we must have $\sin(\theta) \ll 1$ because we must tunnel to a state near the global minimum, which like the false vacuum has $\langle \tilde{q}' \rangle = 0$. Therefore, the analysis should be modified to account for the fact that the Q-balls consist almost entirely of the h' field, with very little of the $\tilde{q}'$ field.

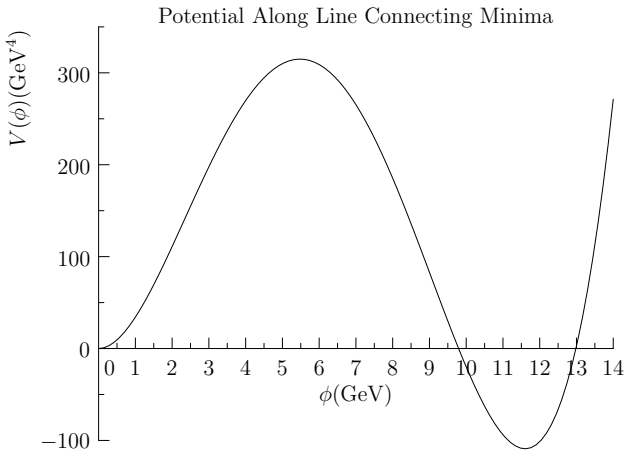

Figure 7.7: The potential along the line connecting the false vacuum to the true vacuum, in our MSSM potential.

However, even if such a phase transition could not have occurred in the evolution of our universe, it is still important to study the regions of parameter space in the MSSM in which such a phase transition could occur. Such a phase transition could destabilize a vacuum previously thought to be stable on cosmological timescales, leading to further constraints beyond those of [58].

In this chapter, we have used the Bethe-Salpeter equation to analyze small Q-balls. At least for certain models, this allows a smooth transition to the semi-classical regime described by the thick wall approximation. As an application of this technique, we have considered the growth of Q-balls through solitosynthesis, starting from individual squarks. We have considered their growth in a radiation-dominated universe like ours, although we have chosen unreasonably large values of η. We have established that such Q-balls can indeed reach critical size and induce a phase transition.

CHAPTER 8

Conclusions

The primary aim of this thesis was to demonstrate that bound state formation in extensions of the Standard Model results in novel features which can potentially resolve well-known problems in the Standard Model. We have illustrated this with three examples: strongly-interacting supersymmetry, dark matter bound states, and Q-balls.

In Ch. 3 we introduced the strongly-coupled version of the Minimal Supersymmetric Standard Model. We motivated the large trilinear couplings by the recently measured Higgs boson mass; an inevitable consequence of large couplings is the formation of squark bound states. These bound states can influence electroweak symmetry breaking through the seesaw-symmetry-breaking mechanism, connecting supersymmetry breaking and electroweak symmetry breaking in a unique manner. In the subsequent chapter, we investigated the phenomenology of this strongly-interacting model. We saw that it avoids common problems such as large flavor-changing neutral currents or breaking $\mathrm{SU_C}(3)$. Beyond this, it can easily accommodate electroweak-scale baryogenesis, in contrast to the Standard Model and the MSSM.

In Ch. 5 and Ch. 6, we considered dark matter bound states in self-interacting asymmetric dark matter models. In Ch. 5, we showed that ongoing formation of bound states in the dense Milky Way halo can produce a detectable indirect detection signal. This was illustrated in a minimal scalar model; in Ch. 6, we introduced a broken $\mathrm{U_D}(1)$ gauge symmetry into the dark sector. We established that in much of the possible parameter space, the dark asymmetry is necessarily generated in two oppositely charged ions, which may form dark atoms. This has

significant implications for the cosmological history of such models. Additionally, since dark matter is generally multi-component with different inter- and intra-species interactions, the effects of scattering in dark matter halos cannot be estimated by a single-species model. We have investigated the parameter space which is disfavored due to large elliptical halos and the parameter space in which dark matter self-interaction can modify small halos.

Finally, we considered Q-balls in a simplified model inspired by the MSSM. We showed that in certain regions of parameter space, we are able to describe small charge Q-balls using the Bethe-Salpeter equation. This allowed us to study the growth of Q-balls from single squarks to large charges; in particular, we were able to establish that, in at least some models, Q-balls can reach their critical size and induce a phase transition. Such phase transitions are one potential baryogenesis mechanism.

These three examples concretely demonstrate the unique opportunities present in extensions of the Standard Model which include bound states.

References

[1] J. M. Cornwall, A. Kusenko, L. Pearce and R. D. Peccei, Phys. Lett. B **718**, 951 (2013) [arXiv:1210.6433 [hep-ph]].

[2] L. Pearce, A. Kusenko and R. D. Peccei, Phys. Rev. D **88**, 075011 (2013) [arXiv:1307.6157 [hep-ph]].

[3] L. Pearce and A. Kusenko, Phys. Rev. D **87**, no. 12, 123531 (2013) [arXiv:1303.7294 [hep-ph]].

[4] K. Petraki, L. Pearce and A. Kusenko, Submitted to JCAP. arXiv:1403.1077 [hep-ph].

[5] L. Pearce, Phys. Rev. D **85**, 125022 (2012) [arXiv:1202.0873 [hep-ph]].

[6] T.-P. Cheng and L.-F. Li, Gauge Theory of Elementary Particle Physics (Clarendon, Oxford, 1984)

[7] S. L. Glashow, Nucl. Phys. **22**, 579 (1961).

[8] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

[9] A. Salam, Conf. Proc. C **680519**, 367 (1968).

[10] E. Noether, Gott. Nachr. **1918**, 235 (1918) [Transp. Theory Statist. Phys. **1**, 186 (1971)] [physics/0503066].

[11] F. Englert and R. Brout, Phys. Rev. Lett. **13**, 321 (1964).

[12] P. W. Higgs, Phys. Rev. Lett. **13**, 508 (1964).

[13] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, Phys. Rev. Lett. **13**, 585 (1964).

[14] J. Goldstone, A. Salam and S. Weinberg, Phys. Rev. **127**, 965 (1962).

[15] G. Aad *et al.* [ATLAS Collaboration], Phys. Lett. B **716**, 1 (2012) [arXiv:1207.7214 [hep-ex]].

[16] S. Chatrchyan *et al.* [CMS Collaboration], Phys. Lett. B **716**, 30 (2012) [arXiv:1207.7235 [hep-ex]].

[17] A. D. Sakharov, Pisma Zh. Eksp. Teor. Fiz. **5**, 32 (1967) [JETP Lett. **5**, 24 (1967)] [Sov. Phys. Usp. **34**, 392 (1991)] [Usp. Fiz. Nauk **161**, 61 (1991)].

[18] W. Buchmuller and O. Philipsen, Nucl. Phys. B **443**, 47 (1995) [hep-ph/9411334].

[19] J.H. Oort, Bull. Astr. Inst. Netherlands **6**, 249 (1932).

[20] F. Zwicky, Helv. Phys. Acta **6**, 110 (1933).

[21] V. C. Rubin, N. Thonnard and W. K. Ford, Jr., Astrophys. J. **238**, 471 (1980).

[22] P. A. R. Ade *et al.* [Planck Collaboration], arXiv:1303.5062 [astro-ph.CO].

[23] K. A. Olive, astro-ph/0301505.

[24] G. Jungman, M. Kamionkowski and K. Griest, Phys. Rept. **267**, 195 (1996) [hep-ph/9506380].

[25] M. Rocha, A. H. G. Peter, J. S. Bullock, M. Kaplinghat, S. Garrison-Kimmel, J. Onorbe and L. A. Moustakas, Mon. Not. Roy. Astron. Soc. **430**, 81 (2013) [arXiv:1208.3025 [astro-ph.CO]].

[26] M. Boylan-Kolchin, J. S. Bullock and M. Kaplinghat, Mon. Not. Roy. Astron. Soc. **415**, L40 (2011) [arXiv:1103.0007 [astro-ph.CO]].

[27] M. Vogelsberger, J. Zavala and A. Loeb, Mon. Not. Roy. Astron. Soc. **423**, 3740 (2012) [arXiv:1201.5892 [astro-ph.CO]].

[28] A. H. G. Peter, M. Rocha, J. S. Bullock and M. Kaplinghat, arXiv:1208.3026 [astro-ph.CO].

[29] K. Petraki and R. R. Volkas, Int. J. Mod. Phys. A **28**, 1330028 (2013) [arXiv:1305.4939 [hep-ph]].

[30] Z. K. Silagadze, hep-ph/9803307.

[31] E. E. Salpeter and H. A. Bethe, Phys. Rev. **84**, 1232 (1951).

[32] G. C. Wick, Phys. Rev. **96**, 1124 (1954).

[33] R. E. Cutkosky, Phys. Rev. **96**, 1135 (1954).

[34] S. R. Coleman and J. Mandula, Phys. Rev. **159**, 1251 (1967).

[35] J. L. Gervais and B. Sakita, Nucl. Phys. B **34**, 632 (1971).

[36] Y. A. Golfand and E. P. Likhtman, JETP Lett. **13**, 323 (1971) [Pisma Zh. Eksp. Teor. Fiz. **13**, 452 (1971)].

[37] S. Chatrchyan *et al.* [CMS Collaboration], JHEP **1211**, 147 (2012) [arXiv:1209.6620 [hep-ex]].

[38] P. Draper, P. Meade, M. Reece and D. Shih, Phys. Rev. D **85**, 095007 (2012) [arXiv:1112.3068 [hep-ph]].

[39] J. L. Evans, M. Ibe, S. Shirai and T. T. Yanagida, Phys. Rev. D **85**, 095004 (2012) [arXiv:1201.2611 [hep-ph]].

[40] T. T. Yanagida, N. Yokozaki and K. Yonekura, JHEP **1210**, 017 (2012) [arXiv:1206.6589 [hep-ph]].

[41] Y. Okada, M. Yamaguchi and T. Yanagida, Phys. Lett. B **262**, 54 (1991).

[42] A. Kusenko, V. Kuzmin and I. I. Tkachev, Phys. Lett. B **432**, 361 (1998) [hep-ph/9801405].

[43] G. F. Giudice and A. Kusenko, Phys. Lett. B **439**, 55 (1998) [hep-ph/9805379].

[44] Y. Nambu, Phys. Rev. **117**, 648 (1960).

[45] T. Kugo, Phys. Lett. B **76**, 625 (1978).

[46] R. Delbourgo and M. D. Scadron, J. Phys. G **6**, 649 (1980).

[47] J. M. Cornwall, R. Jackiw and E. Tomboulis, Phys. Rev. D **10**, 2428 (1974).

[48] H. E. Haber, R. Hempfling and A. H. Hoang, Z. Phys. C **75**, 539 (1997) [hep-ph/9609331].

[49] C. Wymant, Phys. Rev. D **86**, 115023 (2012) [arXiv:1208.1737 [hep-ph]].

[50] L. Alvarez-Gaume, J. Polchinski and M. B. Wise, Nucl. Phys. B **221**, 495 (1983).

[51] J. M. Frere, D. R. T. Jones and S. Raby, Nucl. Phys. B **222**, 11 (1983).

[52] M. Drees, M. Gluck and K. Grassie, DO-TH 85/3.

[53] J. F. Gunion, H. E. Haber and M. Sher, Nucl. Phys. B **306**, 1 (1988).

[54] H. Komatsu, Phys. Lett. B **215**, 323 (1988).

[55] P. Langacker and N. Polonsky, Phys. Rev. D **50**, 2199 (1994) [hep-ph/9403306].

[56] A. Kusenko, P. Langacker and G. Segre, Phys. Rev. D **54**, 5824 (1996) [hep-ph/9602414].

[57] J. A. Casas, A. Lleyda and C. Munoz, Phys. Lett. B **389**, 305 (1996) [hep-ph/9606212].

[58] A. Kusenko and P. Langacker, Phys. Lett. B **391**, 29 (1997) [hep-ph/9608340].

[59] E. Farhi and L. Susskind, Phys. Rept. **74**, 277 (1981).

[60] T. Akiba and T. Yanagida, Phys. Lett. B **169**, 432 (1986).

[61] T. W. Appelquist, D. Karabali and L. C. R. Wijewardhana, Phys. Rev. Lett. **57**, 957 (1986).

[62] T. A. Ryttov and R. Shrock, Phys. Rev. D **84**, 056009 (2011) [arXiv:1107.3572 [hep-ph]].

[63] M. Dine, W. Fischler and M. Srednicki, Nucl. Phys. B **189**, 575 (1981).

[64] A. J. Buras and T. Yanagida, Phys. Lett. B **121**, 316 (1983).

[65] P. Hernandez, N. Rius and V. Sanz, Nucl. Phys. Proc. Suppl. **95**, 272 (2001).

[66] M. Pietroni, Nucl. Phys. B **402**, 27 (1993) [hep-ph/9207227].

[67] A. T. Davies, C. D. Froggatt and R. G. Moorhouse, Phys. Lett. B **372**, 88 (1996) [hep-ph/9603388].

[68] S. J. Huber and M. G. Schmidt, in *Proceedings of SEWM 2000*, edited by C. P. Korthals Altes (World Scientific, Singapore, 2001), p. 272 [hep-ph/0011059].

[69] S. J. Huber and M. G. Schmidt, Nucl. Phys. B **606**, 183 (2001) [hep-ph/0003122].

[70] M. E. Carrington, Phys. Rev. D **45**, 2933 (1992).

[71] M. Dine, R. G. Leigh, P. Huet, A. D. Linde and D. A. Linde, Phys. Lett. B **283**, 319 (1992) [hep-ph/9203201].

[72] M. Dine, R. G. Leigh, P. Y. Huet, A. D. Linde and D. A. Linde, Phys. Rev. D **46**, 550 (1992) [hep-ph/9203203].

[73] G. F. Giudice, Phys. Rev. D **45**, 3177 (1992).

[74] L. Dolan and R. Jackiw, Phys. Rev. D **9**, 3320 (1974).

[75] S. Weinberg, Phys. Rev. D **9**, 3357 (1974).

[76] V. A. Kuzmin, V. A. Rubakov and M. E. Shaposhnikov, Phys. Lett. B **155**, 36 (1985).

[77] V. A. Rubakov and M. E. Shaposhnikov, Usp. Fiz. Nauk **166** (1996) 493 [Phys. Usp. **39** (1996) 461] [hep-ph/9603208].

[78] J. McDonald, Phys. Lett. B **323**, 339 (1994).

[79] A. G. Cohen, D. B. Kaplan and A. E. Nelson, Phys. Lett. B **263**, 86 (1991).

[80] M. Dine, P. Huet, R. L. Singleton, Jr and L. Susskind, Phys. Lett. B **257**, 351 (1991).

[81] P. B. Arnold and L. D. McLerran, Phys. Rev. D **36**, 581 (1987).

[82] J. Ambjorn, T. Askgaard, H. Porter and M. E. Shaposhnikov, Nucl. Phys. B **353**, 346 (1991).

[83] G. Passarino, Phys. Lett. B **231**, 458 (1989).

[84] G. Passarino, Phys. Lett. B **247**, 587 (1990).

[85] B. W. Lynn and E. Nardi, Nucl. Phys. B **381**, 467 (1992).

[86] T. Blank and W. Hollik, Nucl. Phys. B **514**, 113 (1998) [hep-ph/9703392].

[87] M. Aoki, S. Kanemura and K. Yagyu, Phys. Rev. D **85**, 055007 (2012) [arXiv:1110.4625 [hep-ph]].

[88] R. S. Chivukula, N. D. Christensen and E. H. Simmons, Phys. Rev. D **77**, 035001 (2008) [arXiv:0712.0546 [hep-ph]].

[89] J. Beringer *et al.* [Particle Data Group Collaboration], Phys. Rev. D **86**, 010001 (2012).

[90] R. N. Mohapatra and G. Senjanovic, Phys. Rev. D **23**, 165 (1981).

[91] J. Schechter and J. W. F. Valle, Phys. Rev. D **22**, 2227 (1980).

[92] M. Magg and C. Wetterich, Phys. Lett. B **94**, 61 (1980).

[93] T. P. Cheng and L. -F. Li, Phys. Rev. D **22**, 2860 (1980).

[94] D. O'Connell, M. J. Ramsey-Musolf and M. B. Wise, Phys. Rev. D **75**, 037701 (2007) [hep-ph/0611014].

[95] I. M. Shoemaker, K. Petraki and A. Kusenko, JHEP **1009**, 060 (2010) [arXiv:1006.5458 [hep-ph]].

[96] S. Dodelson and L. M. Widrow, Phys. Rev. D **42**, 326 (1990).

[97] S. M. Barr, R. S. Chivukula and E. Farhi, Phys. Lett. B **241**, 387 (1990).

[98] D. B. Kaplan, Phys. Rev. Lett. **68**, 741 (1992).

[99] V. A. Kuzmin, Phys. Part. Nucl. **29**, 257 (1998) [Fiz. Elem. Chast. Atom. Yadra **29**, 637 (1998)] [Phys. Atom. Nucl. **61**, 1107 (1998)] [arXiv:hep-ph/9701269].

[100] A. Kusenko and M. E. Shaposhnikov, Phys. Lett. B **418**, 46 (1998) [arXiv:hep-ph/9709492].

[101] A. Kusenko, V. Kuzmin, M. E. Shaposhnikov and P. G. Tinyakov, Phys. Rev. Lett. **80**, 3185 (1998) [arXiv:hep-ph/9712212].

[102] M. Laine and M. E. Shaposhnikov, Nucl. Phys. B **532**, 376 (1998) [arXiv:hep-ph/9804237].

[103] R. Kitano and I. Low, Phys. Rev. D **71**, 023510 (2005) [arXiv:hep-ph/0411133].

[104] Z. Berezhiani, D. Comelli and F. L. Villante, Phys. Lett. B **503**, 362 (2001) [arXiv:hep-ph/0008105].

[105] R. Foot and R. R. Volkas, Phys. Rev. D **68**, 021304 (2003) [arXiv:hep-ph/0304261].

[106] R. Foot and R. R. Volkas, Phys. Rev. D **69**, 123510 (2004) [arXiv:hep-ph/0402267].

[107] D. E. Kaplan, M. A. Luty and K. M. Zurek, Phys. Rev. D **79**, 115016 (2009) [arXiv:0901.4117 [hep-ph]].

[108] L. J. Hall, J. March-Russell and S. M. West, arXiv:1010.0245 [hep-ph].

[109] R. Allahverdi, B. Dutta and K. Sinha, Phys. Rev. D **83**, 083502 (2011) [arXiv:1011.1286 [hep-ph]].

[110] B. Dutta and J. Kumar, Phys. Lett. B **699**, 364 (2011) [arXiv:1012.1341 [hep-ph]].

[111] N. F. Bell, K. Petraki, I. M. Shoemaker and R. R. Volkas, Phys. Rev. D **84**, 123505 (2011) [arXiv:1105.3730 [hep-ph]].

[112] C. Cheung and K. M. Zurek, Phys. Rev. D **84**, 035007 (2011) [arXiv:1105.4612 [hep-ph]].

[113] B. von Harling, K. Petraki and R. R. Volkas, JCAP **1205**, 021 (2012) [arXiv:1201.2200 [hep-ph]].

[114] K. Petraki, M. Trodden and R. R. Volkas, JCAP **1202**, 044 (2012) [arXiv:1111.4786 [hep-ph]].

[115] J. J. Heckman and S. J. Rey, JHEP **1106**, 120 (2011) [arXiv:1102.5346 [hep-th]].

[116] H. Davoudiasl, D. E. Morrissey, K. Sigurdson and S. Tulin, Phys. Rev. Lett. **105**, 211304 (2010) [arXiv:1008.2399 [hep-ph]].

[117] D. N. Spergel and P. J. Steinhardt, Phys. Rev. Lett. **84**, 3760 (2000) [arXiv:astro-ph/9909386].

[118] R. Dave, D. N. Spergel, P. J. Steinhardt and B. D. Wandelt, Astrophys. J. **547**, 574 (2001) [arXiv:astro-ph/0006218].

[119] N. Yoshida, V. Springel, S. D. M. White and G. Tormen, Astrophys. J. **544**, L87 (2000) [arXiv:astro-ph/0006134].

[120] A. Kusenko and P. J. Steinhardt, Phys. Rev. Lett. **87**, 141301 (2001) [arXiv:astro-ph/0106008].

[121] D. E. Holz and A. Zee, Phys. Lett. B **517**, 239 (2001) [hep-ph/0105284].

[122] S. Andreas, T. Hambye and M. H. G. Tytgat, JCAP **0810**, 034 (2008) [arXiv:0808.0255 [hep-ph]].

[123] N. Arkani-Hamed, D. P. Finkbeiner, T. R. Slatyer and N. Weiner, Phys. Rev. D **79**, 015014 (2009) [arXiv:0810.0713 [hep-ph]].

[124] J. L. Feng, M. Kaplinghat, H. Tu and H. B. Yu, JCAP **0907**, 004 (2009) [arXiv:0905.3039 [hep-ph]].

[125] J. L. Feng, M. Kaplinghat and H. B. Yu, Phys. Rev. Lett. **104**, 151301 (2010) [arXiv:0911.0422 [hep-ph]].

[126] M. Boylan-Kolchin, J. S. Bullock and M. Kaplinghat, Mon. Not. Roy. Astron. Soc. **422**, 1203 (2012) [arXiv:1111.2048 [astro-ph.CO]].

[127] M. Gonderinger, H. Lim and M. J. Ramsey-Musolf, Phys. Rev. D **86**, 043511 (2012) [arXiv:1202.1316 [hep-ph]].

[128] S. Tulin, H. -B. Yu, and K. M. Zurek, arXiv:1210.0900 [hep-ph].

[129] M. R. Buckley and P. J. Fox, Phys. Rev. D **81**, 083522 (2010) [arXiv:0911.3898 [hep-ph]].

[130] M. Vogelsberger and J. Zavala, Mon. Not. Roy. Astron. Soc. **430**, 1722 (2013) [arXiv:1211.1377 [astro-ph.CO]].

[131] W. Shepherd, T. M. P. Tait and G. Zaharijas, Phys. Rev. D **79**, 055022 (2009) [arXiv:0901.2125 [hep-ph]].

[132] J. F. Navarro, C. S. Frenk and S. D. M. White, Astrophys. J. **462**, 563 (1996) [astro-ph/9508025].

[133] A. Klypin, H. Zhao and R. S. Somerville, Astrophys. J. **573**, 597 (2002) [astro-ph/0110390].

[134] M. Vogelsberger, A. Helmi, V. Springel, S. D. M. White, J. Wang, C. S. Frenk, A. Jenkins and A. D. Ludlow *et al.*, Mon. Not. Roy. Astron. Soc. **395**, 797 (2009) [arXiv:0812.0362 [astro-ph]].

[135] S. Tulin, H. -B. Yu and K. M. Zurek, Phys. Rev. D **87**, 115007 (2013) [arXiv:1302.3898 [hep-ph]].

[136] S. W. Randall, M. Markevitch, D. Clowe, A. H. Gonzalez and M. Bradac, Astrophys. J. **679**, 1173 (2008) [arXiv:0704.0261 [astro-ph]].

[137] P. Colin, V. Avila-Reese, O. Valenzuela and C. Firmani, Astrophys. J. **581**, 777 (2002) [arXiv:astro-ph/0205322].

[138] A. Loeb and N. Weiner, Phys. Rev. Lett. **106**, 171302 (2011) [arXiv:1011.6374 [astro-ph.CO]].

[139] J. Miralda-Escudé, Astrophys. J. **564**, 60 (2002).

[140] E. Aprile *et al.* [XENON100 Collaboration], Phys. Rev. Lett. **107**, 131302 (2011) [arXiv:1104.2549 [astro-ph.CO]].

[141] Z. Ahmed *et al.* [CDMS Collaboration], Science **327**, 1619 (2010).

[142] S. D. McDermott, H. -B. Yu and K. M. Zurek, Phys. Rev. D **85**, 023519 (2012) [arXiv:1103.5472 [hep-ph]].

[143] C. Kouvaris, Phys. Rev. Lett. **108**, 191301 (2012) [arXiv:1111.4364 [astro-ph.CO]];

[144] T. Guver, A. E. Erkoca, M. H. Reno and I. Sarcevic, arXiv:1201.2400 [hep-ph];

[145] N. F. Bell, A. Melatos and K. Petraki, Phys. Rev. D **87**, 123507 (2013) [arXiv:1301.6811 [hep-ph]].

[146] C. Kouvaris and P. Tinyakov, Phys. Rev. D **87**, 123537 (2013) [arXiv:1212.4075 [astro-ph.HE]

[147] J. Bramante, K. Fukushima and J. Kumar, Phys. Rev. D **87**, 055012 (2013) [arXiv:1301.0036 [hep-ph]].

[148] G. B. Gelmini, P. Gondolo and E. Roulet, Nucl. Phys. B **351**, 623 (1991);

[149] A. Akhiezer and N. Merenkov, J. Phys. B: At. Mol. Opt. Phys. 29 (1996).

[150] A. Sommerfeld, Atombau and Spektrallinien vol II 1951 (Braunschweig: Vieger).

[151] A. Akhiezer and V. Berestetsky, Quantum Electrodynamics 1969 (Moscow: Nauka).

[152] T. Kobayashi et al., Astrophys. J. **601**, 340 (2004).

[153] J. S. Mathis, P. G. Mezger, and N. Panagia, A&A, **128**, 212 (1983).

[154] Ya. B. Zeldovich, A.A. Ruzmaikin, and D.D. Sokoloff, *Magnetic fields in astrophysics*, Gordon and Breach, New York, 1983;

[155] A. Kandus, K. E. Kunze and C. G. Tsagas, Phys. Rept. **505**, 1 (2011) [arXiv:1007.3891 [astro-ph.CO]].

[156] G. Blumenthal & R. Gould, Rev. Mod. Phys. **42**, 237 (1970).

[157] F. Jones, 1968, Phys. Rev., **167**, 1159 (1968).

[158] M. Ackermann *et al.* [LAT Collaboration], Phys. Rev. D **86**, 022002 (2012) [arXiv:1205.2739 [astro-ph.HE]].

[159] O. Adriani *et al.* [PAMELA Collaboration], Nature **458**, 607 (2009) [arXiv:0810.4995 [astro-ph]].

[160] M. Ackermann *et al.* [Fermi LAT Collaboration], Phys. Rev. Lett. **108**, 011103 (2012) [arXiv:1109.0521 [astro-ph.HE]].

[161] M. Aguilar *et al.* [AMS Collaboration], Phys. Rev. Lett. **110**, no. 14, 141102 (2013).

[162] M. Backovic and J. P. Ralston, Phys. Rev. D **81**, 056002 (2010) [arXiv:0910.1113 [hep-ph]].

[163] G. Elwert Ann Phys 34, **178** (1939).

[164] G. Elwert and E. Haug, Phys Rev 183, **90** (1969).

[165] R. B. Ferroli, S. Pacetti and A. Zallo, Eur. Phys. J. A **48**, 33 (2012) [arXiv:1008.0542 [hep-ph]].

[166] J. Knodlseder, G. Weidenspointner, P. Jean, R. Diehl, A. Strong, H. Halloin, B. Cordier and S. Schanne *et al.*, arXiv:0712.1668 [astro-ph].

[167] E. W. Kolb and M. S. Turner, Front. Phys. **69**, 1 (1990).

[168] M. L. Graesser, I. M. Shoemaker and L. Vecchi, JHEP **1110**, 110 (2011) [arXiv:1103.2771 [hep-ph]].

[169] J. L. Feng, H. Tu and H. -B. Yu, JCAP **0810**, 043 (2008) [arXiv:0808.2318 [hep-ph]].

[170] Y. Bai, P. J. Fox and R. Harnik, JHEP **1012**, 048 (2010) [arXiv:1005.3797 [hep-ph]].

[171] M. R. Buckley, Phys. Rev. D **84**, 043510 (2011) [arXiv:1104.1429 [hep-ph]].

[172] P. J. Fox, R. Harnik, R. Primulando and C. -T. Yu, Phys. Rev. D **86**, 015010 (2012) [arXiv:1203.1662 [hep-ph]].

[173] J. March-Russell, J. Unwin and S. M. West, JHEP **1208**, 029 (2012) [arXiv:1203.4854 [hep-ph]].

[174] D. E. Kaplan, G. Z. Krnjaic, K. R. Rehermann and C. M. Wells, JCAP **1110**, 011 (2011) [arXiv:1105.2073 [hep-ph]].

[175] F. -Y. Cyr-Racine and K. Sigurdson, Phys. Rev. D **87**, no. 10, 103515 (2013) [arXiv:1209.5752 [astro-ph.CO]].

[176] B. Holdom, Phys. Lett. B **166**, 196 (1986).

[177] R. Foot and X. -G. He, Phys. Lett. B **267**, 509 (1991).

[178] B. Batell, M. Pospelov and A. Ritz, Phys. Rev. D **79**, 115008 (2009) [arXiv:0903.0363 [hep-ph]].

[179] L. Hulthén, Ark. Mat. Atron. Fys. **28 A** (1942).

[180] L. Hulthén, Ark. Mat. Atron. Fys. **29 B** (1942).

[181] J. M. Cline, Z. Liu, G. Moore and W. Xue, Phys. Rev. D **89**, 043514 (2014) [arXiv:1311.6468 [hep-ph]].

[182] M. Cirelli, P. Panci, G. Servant and G. Zaharijas, JCAP **1203**, 015 (2012) [arXiv:1110.3809 [hep-ph]].

[183] G. Mangano and P. D. Serpico, Phys. Lett. B **701**, 296 (2011) [arXiv:1103.1261 [astro-ph.CO]].

[184] J. Jaeckel, M. Jankowiak and M. Spannowsky, Phys. Dark Univ. **2**, 111 (2013) [arXiv:1212.3620 [hep-ph]].

[185] M. Goodsell, J. Jaeckel, J. Redondo and A. Ringwald, JHEP **0911**, 027 (2009) [arXiv:0909.0515 [hep-ph]].

[186] H. An, M. Pospelov and J. Pradler, Phys. Rev. Lett. **111**, 041302 (2013) [arXiv:1304.3461 [hep-ph]].

[187] J. M. Cline, Z. Liu and W. Xue, Phys. Rev. D **85**, 101302 (2012) [arXiv:1201.4858 [hep-ph]].

[188] J. M. Cline, Z. Liu, G. Moore and W. Xue, arXiv:1312.3325 [hep-ph].

[189] R. Laha and E. Braaten, arXiv:1311.6386 [hep-ph].

[190] R. Foot, Phys. Rev. D **69**, 036001 (2004) [hep-ph/0308254].

[191] R. Foot, Phys. Rev. D **82**, 095001 (2010) [arXiv:1008.0685 [hep-ph]].

[192] R. Foot, Phys. Lett. B **728**, 45 (2014) [arXiv:1305.4316 [astro-ph.CO]].

[193] N. Fornengo, P. Panci and M. Regis, Phys. Rev. D **84**, 115002 (2011) [arXiv:1108.4661 [hep-ph]].

[194] M. Pospelov and A. Ritz, Phys. Lett. B **671**, 391 (2009) [arXiv:0810.1502 [hep-ph]].

[195] R. Foot, Phys. Lett. B **718**, 745 (2013) [arXiv:1208.6022 [astro-ph.CO]].

[196] F.-Y. Cyr-Racine, R. de Putter, A. Raccanelli and K. Sigurdson, arXiv:1310.3278 [astro-ph.CO].

[197] S. A. Khrapak, A. V. Ivlev, G. E. Morfill, S. K. Zhdanov, and H. M. Thomas, Transactions on Plasma Science **32** 555 (2004).

[198] S. A. Khrapak, A. V. Ivlev, G. E. Morfill, and S. K. Zhdanov, Phys. Rev. Lett. **90** 225002 (2003).

[199] S. A. Khrapak, A. V. Ivlev, and G. E. Morfill, Phys. Rev. E **70** 056405 (2004).

[200] D. A. Buote, T. E. Jeltema, C. R. Canizares and G. P. Garmire, Astrophys. J. **577**, 183 (2002) [astro-ph/0205469].

[201] G. Battaglia, A. Helmi and M. Breddels, New Astron. Rev. **57**, 52 (2013) [arXiv:1305.5965 [astro-ph.CO]].

[202] S. R. Coleman, Nucl. Phys. B **262**, 263 (1985) [Erratum-ibid. B **269**, 744 (1986)].

[203] T. D. Lee and Y. Pang, Phys. Rept. **221**, 251 (1992).

[204] G. Rosen, J. Math. Phys. **9**, 996 (1968);

[205] T. D. Lee and G. C. Wick, Phys. Rev. D **9**, 2291 (1974);

[206] R. Friedberg, T. D. Lee and A. Sirlin, Phys. Rev. D **13**, 2739 (1976).

[207] A. Kusenko, Phys. Lett. B **405**, 108 (1997) [hep-ph/9704273].

[208] K. Enqvist and J. McDonald, Phys. Lett. B **425**, 309 (1998) [hep-ph/9711514].

[209] S. Kasuya and M. Kawasaki, Phys. Rev. D **61**, 041301 (2000) [hep-ph/9909509].

[210] S. Kasuya and M. Kawasaki, Phys. Rev. D **62**, 023512 (2000) [hep-ph/0002285].

[211] S. Kasuya and M. Kawasaki, Phys. Rev. D **64**, 123515 (2001) [hep-ph/0106119].

[212] M. Dine and A. Kusenko, Rev. Mod. Phys. **76**, 1 (2003) [hep-ph/0303065].

[213] K. Griest and E. W. Kolb, Phys. Rev. D **40**, 3231 (1989);

[214] J. A. Frieman, G. B. Gelmini, M. Gleiser and E. W. Kolb, Phys. Rev. Lett. **60**, 2101 (1988).

[215] A. Kusenko, M. E. Shaposhnikov, P. G. Tinyakov and I. I. Tkachev, Phys. Lett. B **423**, 104 (1998) [arXiv:hep-ph/9801212].

[216] A. Kusenko, L. Loveridge and M. Shaposhnikov, Phys. Rev. D **72**, 025015 (2005) [arXiv:hep-ph/0405044];

[217] A. Kusenko, L. C. Loveridge and M. Shaposhnikov, JCAP **0508**, 011 (2005) [arXiv:astro-ph/0507225].

[218] A. Kusenko, Phys. Lett. B **406**, 26 (1997) [hep-ph/9705361].

[219] J. Barnard, JHEP **1101**, 101 (2011) [arXiv:1011.4944 [hep-ph]].

[220] M. Postma, Phys. Rev. D **65**, 085035 (2002) [hep-ph/0110199].

[221] D. Metaxas, Phys. Rev. D **63**, 083507 (2001) [hep-ph/0009225].

[222] M. Gleiser and A. F. Heckler, Phys. Rev. Lett. **76**, 180 (1996) [hep-ph/9509347].

[223] A. Kusenko, M. E. Shaposhnikov and P. G. Tinyakov, Pisma Zh. Eksp. Teor. Fiz. **67**, 229 (1998) [JETP Lett. **67**, 247 (1998)] [hep-th/9801041].

[224] A. Kusenko, Phys. Lett. B **404**, 285 (1997) [hep-th/9704073].

[225] D. Spector, Phys. Lett. B **194**, 103 (1987).

[226] A. D. Linde, Nucl. Phys. B **216**, 421 (1983) [Erratum-ibid. B **223**, 544 (1983)].

[227] N. Graham, Phys. Lett. B **513**, 112 (2001) [hep-th/0105009].

[228] J. A. Frieman, A. V. Olinto, M. Gleiser and C. Alcock, Phys. Rev. D **40**, 3241 (1989).

[229] R. Banerjee and K. Jedamzik, Phys. Lett. B **484**, 278 (2000) [hep-ph/0005031].