Summer 9-1-2014

Phase-Space Distributions of Galactic Dark Matter Halos and Implications for Detection

Daniel Robert Hunter
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS
Department of Physics

Dissertation Examination Committee:
Francesc Ferrer, Chair
James H. Buckley
Ramanath Cowsik
Ricardo Flores
Viktor Gruev
Michael C. Ogilvie

Phase-Space Distributions of Galactic Dark Matter Halos and Implications for Detection

by

Daniel R. Hunter

A dissertation presented to the
Graduate School of Arts & Sciences
of Washington University in St. Louis in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2014
St. Louis, Missouri

Contents

List of Figures

List of Tables

Acknowledgements

First and foremost I thank my advisor Francesc Ferrer for his invaluable guidance and patience. I am also grateful to my co-advisors James H. Buckley and Michael C. Ogilvie, and my committee members Ramanath Cowsik, Viktor Gruev, and Ricardo A. Flores.

I am thankful to my fellow students at Washington University with whom I have learned and struggled over the years and from whom I've been fortunate to receive support and sanity. In particular I thank Laleh Sadeghian, Saeed Mirshekari, Adam J. Archibald, Thomas Crockett (with Ivy Love), Tyler Harmon, Leandro Medina de Oliveira, Ryan X. Charles, Ryan P. Murphy, Benjamin P. Burch, Dimitris Manolidis, Patrick G. Johnson, John Flavin, Sarah E. Thibadeau, S. Brett McArthur, Daniel Flanagan, Sina Mossahebi, and Lauren C. Tran.

I dedicate this work to my families, both intrinsic and discovered. My parents Stanley D. Hunter and Michele M. Hunter created the foundations of my creativity and love of knowledge, for which I am eternally indebted. I adore and admire them, and I am overjoyed to make them proud. It has been a marvel to watch my brother Thomas E. Hunter grow and develop alongside me in astonishing ways and directions. It seems that the more different our paths become the closer we get, and this is just the beginning.

Finally, I acknowledge the friends I have found over the years that have become as family to me. It has been a tremendous joy and privilege to be a part of their lives and to resonate so deeply with them. Their inspiration and delightful challenge to my imagination has made me the person I am today. I especially praise the following wonderful persons:

Jason A. Smith, Elliot J. Peeples, L. Paige Johnson, Michael Bobrow, Brooke M. Baldwin, Steven Krehbiel, Stubert Smith, everyone else at Camp SUDS, Rejoice Miene, Tomko Miene, the Reiner family, Eugene "Zhenya" Libster, Shannon Tracy, and Charles Madonna.

— DRCDH

"Because I can, because no one can stop me, because it makes up for things I lost, because I'm addicted to bad ideas, and all the beauty in this world."

— The World/Inferno Friendship Society

1 Introduction

In physics we constantly attempt to align theory with observation. Given some reasonable assumptions, we compare predictions of quantities to measurements of the same quantities. The more precisely the results match, the greater confidence we can put in our understanding. The other possibility is the predictions and measurements do *not* match. At least one of these must be wrong. In some sense, measurement also depends on theory: we can only be so sure of how well a measurement represents any particular physical quality or action. Often we have such control over an experimental set-up that this is manifested only as statistical and systematic uncertainty. This typically makes the fault of disagreement between reproducable measurement and theory fall upon theory. However, in *astrophysics*, we cannot design the systems to be measured, only the probes for measurement. So when physical theory disagrees with an astronomical observation of some system, ignoring the possibility of a bad instrument, there are two possible culprits: the theory describing how parts of the system interact (*i.e.* the dynamics), or the theory describing *what quality* of the system has been measured.

This difficulty has created one of the most important open problems in modern astrophysics, *the missing mass problem*. As a brief example, which will be returned to in greater detail, consider a spiral galaxy, made up of stars, gas, and plasma. It is natural to suppose these constituents interact with one another through Newtonian gravity, meaning that there is a specific relationship between the velocity curve, which can be measured via the Doppler shift, and the contained mass, which is measured by accumulating the abundances of the constituents. We would expect these measurements to generally agree, but they do not. Along the lines of what is described above, one of two conclusions can be drawn: spiral galaxies do not exactly follow Newtonian gravity, or the measurements are incomplete. More precisely, the second of these possible conclusions is that galaxies have some component that dominates their total mass and is non-luminous. This new component is the standard explanation and is called *dark matter*. The other possibility, that Newtonian gravity fails in some way as an effective theory for galactic-scale dynamics, has also been seriously considered [89, 90, 59], but the application of this paradigm

to cosmology is incomplete, and it has problems predicting the power spectrum of the Cosmic Microwave Background. We will assume in this work that Newtonian dynamics and General Relativity are correct: unseen dark matter does indeed exist.

Accepting that dark matter exists, the next question is: what is it? At the time when astrophysical discrepancies were first found[1] by Fritz Zwicky in 1933 [141], it was known that a large portion of the luminous mass in galaxies was made up not of stars but diffuse gas and dust. We now call this the interstellar medium. Astronomical observations had not yet covered the entire electromagnetic spectrum at the time, so it was natural to expect the missing mass to be filled in by the interstellar medium at yet-unseen wavelengths. Modern observations have covered this spectral ground from radio to gamma-ray: atomic hydrogen is traced by its 21 cm line, molecular gas is identified by spectral lines from radio to infrared, ionized gas (plasma) emits thermal x-rays, and interstellar dust is a source of infrared blackbody radiation. Observations have enriched the map of galaxies and clusters of galaxies, but they have not entirely revealed the missing mass. Not satisfied with diffuse material, the existence of massive compact halo objects (MACHOs) was considered around the 1980s. A population of such objects would be non-luminous, evading observations as yet, but could, however, be determined by gravitational micro-lensing: the brightness of background stars would be occasionally enhanced by lensing over time. After lengthy monitoring of the galaxy, however, it seems that MACHOs account for at most only several percent of the total mass [125].

By cosmological arguments, the majority of matter in the universe must be non-baryonic and cold, and we identify this cosmological component as the galactic dark matter as well. We must turn to physics beyond the Standard Model. Specific candidates will not be discussed at length, but the generic weakly-interacting massive particle (WIMP) will be assumed. This class of particle candidates is well motivated by the thermal relic density calculation and by the possibility of supersymmetry. WIMP dark matter allows for the possibility of non-gravitational particle detection. Detection methods can be either direct, looking for signals from WIMP-nucleon interactions in terrestrial

[1] Oort found discrepancies in the dynamics of the solar system in 1932 that he attributed to something like dark matter, but these were quickly settled.

instruments, or indirect, looking for signatures of self-annihilation where the WIMP density is highest. Such a detection of particle dark matter has not yet been confirmed as successful. It is hoped that this work contributes toward a discovery.

In order to make specific predictions for direct and indirect detection, we need to model the structures formed by dark matter. Especially, we need to understand the distribution of dark matter in and around the Milky Way, where we are most likely to make a detection. Signatures of WIMPs in indirect detection generally scale with either their mass density or mass density squared. We can predict the distribution of mass from Newtonian simulations, which gives us an idea of where to look and the luminous extent of dark matter structures in the galaxy. We are also interested in the velocity distribution of WIMPs for two reasons: the velocity distribution of WIMPs near the earth has a large impact on direct detection, and it can also have a significant effect on annihilation rate of WIMPs, which is important for indirect detection.

These two distributions, mass and velocity, are not independent. It is reasonable to assume that galactic-scale dark matter structures are statistically stationary in the sense that their phase-space distribution is time-independent.[2] This creates a link between the mass distribution of a structure and the velocity distribution of particles within that structure. Taking advantage of this relationship is the basis of this entire work. In particular, we investigate the impact of a complicated augmentation of the annihilation rate, Sommerfeld enhancement, which requires the non-trivial calculation of the velocity distribution of WIMPs everywhere in the galaxy. We also consider the effect of anisotropy in the velocity distribution of WIMPs throughout the galaxy on direct detection prospects.

The focus of this work is phenomenology: how do different models affect predictions of WIMP detection? The many theories of WIMPs, such as the many possible supersymmetric theories, are not investigated here. This work may be considered as a treatise on the necessity of methods of calculation: in what cases are careful and difficult numerical computations necessary to find an accurate and precise prediction? In Section 2, we give an overview of modern physical cosmology, in particular the most important and

[2]The time scale of terrestrial observations is much less than dynamical time scales.

convincing evidences for non-baryonic dark matter: Big Bang Nucleosynthesis and the Cosmic Microwave Background. In Section 3 we give a somewhat-historical overview of the evidence for galactic dark matter, as well as mass density models of galactic dark matter structures and models of the baryonic components of the Milky Way. In Section 4 we briefly discuss various particle candidates for dark matter, motivate WIMPs, and discuss the thermal relic density calculation in detail. In Sections 5 and 6 we review the basic calculations involved in indirect and direct detection, as well as the local positron excess, which has been a major motivator for the possible particle physics of Sommerfeld enhancement, which is also described in detail. Sections 7 and 8 include the main results of this work: we perform the most careful to date calculation of the velocity distribution of WIMPs in the Milky Way and consider the impact on indirect detection prospects; we also derive the anisotropy profile of dark matter in the Milky Way, incorporate the result into a new empirical model of the velocity distribution, and calculate the effect on direct detection prospects. We conclude in Section 9 with some thoughts on these efforts and the context in the larger scheme of things.

All inline division is such that $a/2c = a/(2c)$. A subscript 0 generally denotes the present time or pure s-wave interaction. Effort has been made to avoid using natural units, or at least use them consistently, but wherever constant factors are accidentally neglected, $\hbar = c = k_{\mathcal{B}} = 1$.

2 Cosmology and Dark Matter

Here we review the basic concepts from physical cosmology that will be necessary in future discussions. In particular we are interested in the interplay between matter and radiation densities and expansion, the thermodynamics of these components, and the time-dependence of thermodynamic quantities. Two principles set the foundation for these derivations: general relativity and statistical homogeneity and isotropy on large scales.

Homogeneity and isotropy imply the generic space-time metric attributed to Friedmann, Lemaître, Robertson, and Walker (the FLRW metric). This naturally introduces the concept of cosmic expansion, first predicted by Lemaître in 1927 and observationally confirmed by Edwin P. Hubble in 1929.[3] The expansion of the universe is one of the three pillars of the Big Bang Model of Cosmology, the others being Big Bang Nucleosynthesis (Section 2.1) and the Cosmic Microwave Background (Section 2.2). Using the FLRW metric and assuming that the content of the universe is a perfect fluid, consistent with homogeneity and isotropy, the Einstein tensor equations of general relativity can be solved, producing the Friedmann-Lemaître equations:

$$\frac{\dot{a}^2}{a^2} = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda}{3}, \tag{1}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) + \frac{\Lambda}{3}, \tag{2}$$

where ρ is the density and P is the pressure of all components in the universe. The first of these is often simply called the Friedmann equation, while the second is called the acceleration equation. The quantity a is called the "scale factor": it tracks a length scale in an expanding universe, in the sense that if a physical distance between two points divided by the scale factor, d/a, is constant, then the two points are said to be "comoving". Similarly, we can speak of comoving volumes. We can define the "Hubble

[3]The usual story is that Hubble discovered the expansion of the universe, but there is actually some contention over the exact history. Hubble used redshift measurements by Vesto M. Slipher to determine distances, so there is some major credit due there. See [97] and references within.

parameter" that measures the rate of expansion:

$$H \equiv \frac{\dot{a}}{a}. \tag{3}$$

Two comoving points of physical distance r have a physical relative velocity of Hr in an expanding universe. The dimensionless quantity k takes the value of $+1$, 0, or -1 and determines the curvature of physical space: spherical (closed), flat, or hyperbolic (open), respectively. Observations are very compatable with flat or nearly flat space, and it is a puzzle why a non-flat space should appear so close to flat today, so the usual assumption is that space really is flat and always has been ($k = 0$). We will make this assumption from here on. The quantity $\Lambda = 8\pi G\rho_\Lambda$ is a cosmological constant attributed to the energy density of the vacuum. For the purposes of this work, Λ will be important only when discussing measurements of the Hubble parameter from supernovæobservations.

Equations 1 and 2 also combine to give the continuity equation

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} + 3\frac{\dot{a}}{a}(\rho + P) = 0. \tag{4}$$

Ignoring for a moment the cosmological constant (or absorbing it into ρ) and pressure P, there is a simple Newtonian interpretation of these equations. Consider some matter symmetrically distributed in a sphere with outgoing velocity proportional to the distance from the sphere's center. The matter's kinetic energy driving it outward, diluting it, competes with its potential energy, gravitationally drawing it back in, compressing it. If the total energy is negative, then gravity will eventually win and the matter will collapse. If the total energy is larger than zero, the matter will escape to infinity with kinetic energy to spare. If it is precisely zero, the matter will "barely" escape to infinity. These three cases correspond to closed, open, and flat space. A full relativistic derivation introduces the pressure contribution.

The continuity equation 4 applies to all components separately, assuming they do not interact. It is usually solved for a particular component of the universe by specifying an equation of state that relates ρ and P, normally of the form $P = w\rho$, where w is

a constant "equation of state parameter". Two extreme cases are of importance: non-relativistic matter or "dust" has zero pressure so $w = 0$; ultra-relativistic matter such as radiation has $w = 1/3$.[4] Using whatever equation of state, the continuity equation gives the density as a function of the scale parameter:

$$\rho_M \propto a^{-3}, \tag{5}$$

$$\rho_R \propto a^{-4}. \tag{6}$$

Obviously, the term in equation 1 contributed by curvature is proportional to a^{-2}, and the contribution from the cosmological constant is indeed constant. If we assume that the observed expansion has been monotonic always, these relations have an important implication: regardless of the composition of the universe today, as long as the universe is old enough, there was a time before which radiation dominated the total energy density, so $\rho \approx \rho_R$, and after which matter dominated the total energy density, so $\rho \approx \rho_M$.

In most situations, we will be concerned with cosmology during one of these two "eras", so it is appropriate to assume the dynamics of the universe are dominated by one or the other. Now we use the rest of the information in equations 1 and 2 to find the time-dependence of a and H. First suppose the universe is made up entirely of relativistic matter. In this case, $\rho \propto a^{-4}$, and the Friedmann equation implies

$$a(t) \propto t^{1/2}, \tag{7a}$$

$$H(t) = 1/2t, \tag{7b}$$

$$\rho(t) = \frac{3}{32\pi G t^2}. \tag{7c}$$

Next suppose the universe is dominated non-relativistic matter. In this case, $\rho \propto a^{-3}$,

[4]This is shown explicitly in Appendix A.

and the Friedmann equation implies

$$a(t) \propto t^{2/3}, \tag{8a}$$

$$H(t) = 2/3t, \tag{8b}$$

$$\rho(t) = \frac{1}{6\pi G t^2}. \tag{8c}$$

Expressing the total density ρ explicitly as a sum of components, we can rewrite the Friedmann equation as

$$\frac{3H^2}{8\pi G} = \rho_M + \rho_R + \rho_\Lambda - \frac{3}{8\pi G}\frac{kc^2}{a^2}. \tag{9}$$

If the spacial curvature is zero, then the total density is equivalent to the left-hand-side, which we define as the critical density:

$$\rho_{\text{crit}}(t) \equiv \frac{3H^2(t)}{8\pi G}. \tag{10}$$

If we know the total density of the universe ρ and the rate of expansion H, we then know whether space is curved and how (if we do not assume space is flat *a priori*). If $\rho = \rho_{\text{crit}}$ then space is flat; if ρ is greater than or less than ρ_{crit} then space is closed or open, respectively. This is in analogy with the Newtonian interpretation of the Friedmann equation described above.

Generally, we can write the total energy density in units of the critical density *today*:

$$\rho = \rho_{\text{crit},0}\left[\Omega_\Lambda + \Omega_M\left(\frac{a_0}{a}\right)^3 + \Omega_R\left(\frac{a_0}{a}\right)^4\right], \tag{11}$$

where $\rho_{\text{crit},0}$ is the critical density today,

$$\rho_{\text{crit},0} = \frac{3H_0^2}{8\pi G}, \tag{12}$$

and we have introduced the standard fractions, defined as the fraction of the critical

density made up of whatever component today,

$$\Omega_\Lambda \equiv \rho_{\Lambda,0}/\rho_{\text{crit},0}, \tag{13a}$$

$$\Omega_M \equiv \rho_{M,0}/\rho_{\text{crit},0}, \tag{13b}$$

$$\Omega_R \equiv \rho_{R,0}/\rho_{\text{crit},0}. \tag{13c}$$

Then we have

$$\Omega_\Lambda + \Omega_M + \Omega_R = 1. \tag{14}$$

In Appendix A we review derivations of the number, energy, and entropy densities, and the pressure of relativistic and non-relativistic bosons and fermions. The important results we will need later are the energy and entropy densities of all relativistic species

$$\rho_R(T) = \frac{\pi^2}{30}\mathcal{N}T^4, \tag{15}$$

$$s_R(T) = \frac{2\pi^2}{45}\mathcal{N}T^3. \tag{16}$$

Combining equation 15 with equation 7c we find a relationship between changes in time and temperature during the era of radiation domination:

$$\frac{\mathrm{d}t}{\mathrm{d}T} = -\sqrt{\frac{45}{4\pi^3 G\mathcal{N}}}T^{-3}. \tag{17}$$

An important quantity is the baryon-to-photon ratio

$$\eta \equiv n_B/n_\gamma. \tag{18}$$

We will see later that the number density of cosmological photons is easily measured. While η is still a theoretically-important quantity, it was also a main observational target before the CMB spectrum was fully observed.

Our calculations later will be during either eras of radiation domination or matter domination. The point in history when these two components are equally balanced is

approximated by

$$\rho_R = \rho_M,$$
$$\Omega_R(1+z_{\rm eq})^4 = \Omega_M(1+z_{\rm eq})^3, \tag{19}$$

which, knowing from the CMB temperature that the photon density is $\Omega_\gamma h^2 = 2.47\times10^{-5}$, the total radiation density (photons and neutrinos) is $\Omega_R h^2 = 4.12\times10^{-5}$. We also know from the CMB anisotropies that $\Omega_M h^2 = 0.14$ [3], so we estimate that radiation-matter equality occured at around a redshift of $z_{\rm eq} \simeq 3400$. The photon temperature today is $T_{\gamma,0} = 2.7255\,\mathrm{K}$ and $T_\gamma \propto 1+z$, so the photon temperature at radiation-matter equality was roughly $10^4\,\mathrm{K}$. Studies of interactions that occur at energies corresponding to temperatures well above this value can assume equation 7; at those well below can assume equation 8. On the other hand, we will also see that the density of *baryonic* matter is only $\Omega_B h^2 = 0.022$, so the redshift and photon temperature at radiation-baryon equality were $z_{\gamma,B} \simeq 530$ and $T_{\gamma,B} \simeq 1500\,\mathrm{K}$, respectively.

2.1 Big Bang Nucleosynthesis

The term "nucleosynthesis" refers to the creation of atomic nuclei, combinations of bound protons and neutrons, the two of which being generally referred to as nucleons. Essentially all of the baryonic matter in the universe consists of atomic nuclei, the most widespread of which is hydrogen, making up about three quarters of the baryonic matter; helium being the other quarter. Cosmologically, there are only trace amounts of heavier nuclei such as lithium, carbon, and so on, though these are certainly important in their own right. *Stellar* nucleosynthesis is specifically the process of creating these heavier elements in stars via hydrogen fusion and the subsequent burning of helium. The important product of hydrogen fusion is stable helium nuclei: the temperature of a typical star is high enough that the hydrogen nuclei, which are lone protons, have enough energy to tunnel through the Coulomb barrier and create the unstable isotope ^{2}He.[5] This begins the *proton-proton*

[5]The probability of this is given by the Gamow-Sommerfeld factor, first derived by George Gamow in 1928.

(p-p) chain reaction.[6] This ^{2}He nucleus may emit a positron via beta-plus decay, resulting in a ^{2}H nucleus, which is called a deuteron. This first step of the hydrogen-fusion process may be written as

$$^1\mathrm{H} + {}^1\mathrm{H} \to {}^2\mathrm{H} + e^+ + \nu_e + \gamma, \tag{20}$$

where e^+ is a positron, an anti-electron, and ν_e is an electron neutrino. The energy of the photons emitted in this reaction each have 0.421 MeV of energy, minus whatever kinetic energy the positron and neutrino carry; furthermore, the positrons created annihilate with electrons to create pairs of photons, each with at least $m_e c^2 = 0.511\,\mathrm{MeV}$ of energy (plus whatever kinetic energy the electron and positron had). These photons and those from subsequent reactions are important for creating an outward pressure in a star to prevent gravitational collapse. The next step is for deuterons to combine with another proton to create the helium isotope ^{3}He:

$$^1\mathrm{H} + {}^2\mathrm{H} \to {}^3\mathrm{He} + \gamma. \tag{21}$$

With a source of ^{3}He coming from proton-proton fusion, the dominant reaction to create ^{4}He at these temperatures is the fusion of two ^{3}He nuclei:

$$^3\mathrm{He} + {}^3\mathrm{He} \to {}^4\mathrm{He} + {}^1\mathrm{H} + {}^1\mathrm{H}. \tag{22}$$

Subdominant reactions that create helium from deuterons involve heavier nuclei, lithium and beryllium. These are important, however, because they involve weak interactions and so create neutrinos (with energy of ~ 0.1 to 10 MeV or more). Nuclei of ^{4}He, also called "alpha particles", in supply, creation of heavier nuclei is started by the triple-alpha process: two helium nuclei combine to create the unstable isotrope beryllium-8 that, before it decays, combines with another helium nuclei to finally create the important element carbon-12. Heaver nuclei are created by successive fusions of helium nuclei: oxygen-16, neon-20, *etc.*

[6] This is actually the dominant process in stars with central temperature less than about 1.7×10^7 K. Hotter stars follow the *carbon-nitrogen-oxygen cycle* [109], which we do not need to discuss for the purposes of illustrating the differences between stellar and big bang nucleosynthesis.

This process [27, 73] explains the origin of "metals", elements heavier than helium, but there is a puzzle here. The rate at which deuterons are burned is larger than that at which they are created, so the abundance of deuterons in stars should always be vanishingly small *unless* there was an initial abundance at the time a star may have been formed. In recent astrophysical history, stars are the only objects that, with the ingredients available, just protons, can reach the temperatures necessary to create deuterons. Another puzzle is the helium abundance of the universe today: about 25% of baryonic matter is helium, which is created in stars, but such a large abundance of helium, if it were created *only* in stars, would imply a much higher stellar luminosity than is seen in the sky. The stars are not bright enough to be responsible for all of the helium we see today. We must look to the early universe, when the temperatures were also high, for an initial abundance of deuterons and helium. It turns out that the deuterium and helium abundances today are an excellent probe of the overall density of nucleons when the universe was young, just minutes after the big bang.

The important difference between stars and the early universe is what ingredients are available: stellar nucleosynthesis begins with solely protons. There are no free neutrons, which decay with a lifetime of about 880 seconds [22]. This is far less than the age of the universe: ignoring any intervening processes, after 13 billion years only one out of every $100^{10^{14}}$ neutrons in an initial population would be left over. Much earlier, however, protons and neutrons were kept in chemical equilibrium via weak interactions. When the universe cooled enough, weak interactions become inefficient and two important things then happen: the neutron abundance is subject to decay (since the neutron-to-proton interactions cannot be driven), and nuclear interactions take over, just as in stars, so deuterons and helium nuclei can be formed. Different from stars, however, there are free neutrons available and the density is much lower than in stars.[7] Only two-body interactions can compete against expansion at this density, so helium nuclei must be formed via deuterons, and this goes for heavier nuclei as well. Also, while deuterons are easily created by the abundant protons and neutrons, their own abundance remains

[7]The density of the universe at this time was approximately that of air on Earth.

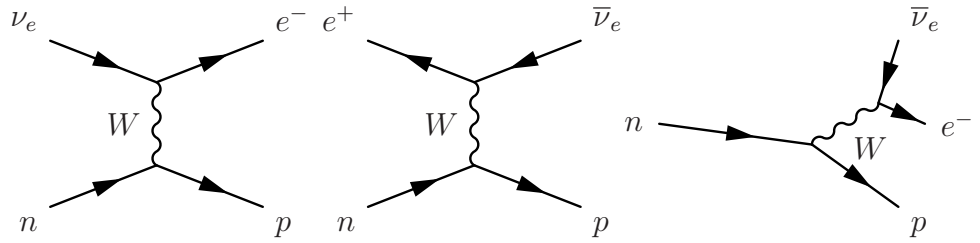

Figure 1: The primary weak interactions that govern the proton and neutron abundances while $T \gtrsim 10^{10}\,\mathrm{K}$. Once these become inefficient, nuclear interactions dominate.

small for some time, since their binding energy is so small. These two facts, that heavy nuclei must be made via deuterons and that the binding energy of deuterons is small, together cause what is called the "deuterium bottleneck". When this bottleneck stops and deuterons become abundant, helium nuclei can finally form. All of the free neutrons become bound in helium nuclei and are thus barred from decay. Since the neutrons are free to decay while the bottleneck is in effect, *the abundance of neutrons depends on when the bottleneck stops*. Observations of the deuterium/helium abundances today are a probe for the cosmological baryon density, because as we will see in more detail below that when the bottleneck stops *depends on the density of nucleons*.

To summarize, the important points are the chemical equilibrium of protons and neutrons, neutron decay, and the deuterium bottleneck. We will now review each of these concepts in detail.

Above temperatures of about $1\,\mathrm{MeV}/k_{\mathcal{B}} \approx 10^{10}\,\mathrm{K}$, the most important reactions are the following six:

$$n + \nu_e \longleftrightarrow p + e^-, \tag{23}$$

$$n + e^+ \longleftrightarrow p + \overline{\nu}_e, \tag{24}$$

$$n \longleftrightarrow p + e^- + \overline{\nu}_e. \tag{25}$$

The tree-level Feynman diagrams for these interactions are shown in Figure 1. In order to calculate the abundances of neutrons and protons when nucleosynthesis starts, we need to know how these interactions change the relative abundances with time. This means we must calculate the cross-sections and then the interaction rates of all six processes. The derivation is straightforward, but we defer the details to Appendix B.

The final result for the total rate of neutron-to-proton conversion is

$$\Gamma_{n\to p} = \frac{2G_F^2}{\pi^3}\frac{m_p^2}{m_n^2}\int_{-\infty}^{\infty}\frac{\mathrm{d}\mathcal{E}\mathcal{E}^2}{1+\exp\{\mathcal{E}/T_\nu\}}\frac{(\mathcal{E}+Q)^2\sqrt{1-m_e^2/(\mathcal{E}+Q)^2}}{1+\exp\{-(\mathcal{E}+Q)/T\}}, \tag{26}$$

where $Q = m_n - m_p$, T and T_ν are the photon and neutrino temperatures, respectively, and the integration leaves out the interval where the square root is imaginary.

If electrons and positrons have not yet annihilated, then $T = T_\nu$ and

$$\Gamma_{p\to n} = \exp\{-Q/T\}\,\Gamma_{n\to p}. \tag{27}$$

Knowing these six important interaction rates, we can finally write the differential equation describing how the number densities of neutrons n_n and protons n_p change with time. We have

$$\begin{aligned}\frac{\mathrm{d}n_n}{\mathrm{d}t} &= -\Gamma_{n\to p}\,n_n + \Gamma_{p\to n}\,n_p \\ &= -\Gamma_{n\to p}\,n_n\left(1+\exp\{-Q/T\}\,\frac{n_p}{n_n}\right).\end{aligned} \tag{28}$$

In equilibrium, the instantaneous rate of change of the neutron density (and thus also the proton density) is zero. We can solve for this equilibrium density easily:

$$\left.\frac{\mathrm{d}n_n}{\mathrm{d}t}\right|_{\mathrm{eq}} = 0 = 1+\exp\{-Q/T\}\,\frac{n_p^{\mathrm{eq}}}{n_n^{\mathrm{eq}}}. \tag{29}$$

This just the Boltzmann factor again. It is convenient to introduce the relative abundances of neutrons and protons, which are defined as the portion of all nucleons by number ($i = n, p$):

$$X_i \equiv \frac{n_i}{n_n + n_p}. \tag{30}$$

Now the equilibrium neutron abundance has the simple form

$$X_n^{\mathrm{eq}} = \frac{1}{1+\exp\{Q/T\}}. \tag{31}$$

We are neglecting any difference in chemical potential between different nucleons. Because of this, when the temperature is much larger than the mass difference Q, protons and neutrons are essentially indistinguishable, so they are evenly balanced in number and $X_n^{\rm eq} = 1/2$. As the temperature drops, however, the mass difference becomes important and it becomes more and more "appealing" to favor the lighter protons.[8] The differential equation for the *actual* neutron abundance is

$$\frac{{\rm d}X_n}{{\rm d}t} = \Gamma_{n\to p}\left(1 + e^{-Q/T}\right)\left(X_n^{\rm eq} - X_n\right). \tag{32}$$

It is useful to switch to the independent variable $x \equiv Q/T$. At the temperatures of interest, radiation dominates, so we may use equation 17. Then equation 32 becomes

$$\frac{{\rm d}X_n}{{\rm d}x} = \sqrt{\frac{45}{4\pi^3\mathcal{N}}}\frac{M_{\rm Pl}}{Q^2}\Gamma_{n\to p}\, x\left(1 + e^{-x}\right)\left(X_n^{\rm eq} - X_n\right). \tag{33}$$

This differential equation can be numerically integrated from some early time at $x_{\rm eq} \ll 1$ with the initial condition $X_n(x_{\rm eq}) = X_n^{\rm eq}(x_{\rm eq})$. Such a calculation is plotted in Figure 2. Note that the actual neutron ratio begins to deviate significantly from its equilibrium value at a temperature of around $10^{10}\,{\rm K}$.

This calculation has been done ignoring the creation of complex nuclei, meaning that it certainly becomes incorrect at a temperature of around $10^9\,{\rm K}$, when deuterons can build up a substantial population, stopping neutrons from decaying. We now want to calculate *when* this happens and the deuterium bottleneck "opens". Knowing the neutron abundance at that time, we can finally determine the abundance of various isotopes.

First let us consider the efficiency of deuteron-production. The production rate *per*

[8] As another way to understand this, consider the Helmholtz free energy $F = U - TS$, where U is the energy, T is the temperature, and S is the entropy of the system of protons and neutrons. The system tends to minimize the free energy. When the temperature is large, it is more advantageous to maximize the entropy. This is accomplished by having a even mix of protons and neutrons, maximizing the disorder. On the other hand, if the temperature is low, it is more advantageous to minimize the energy U, which is accomplished by preferring less-massive protons over neutrons.

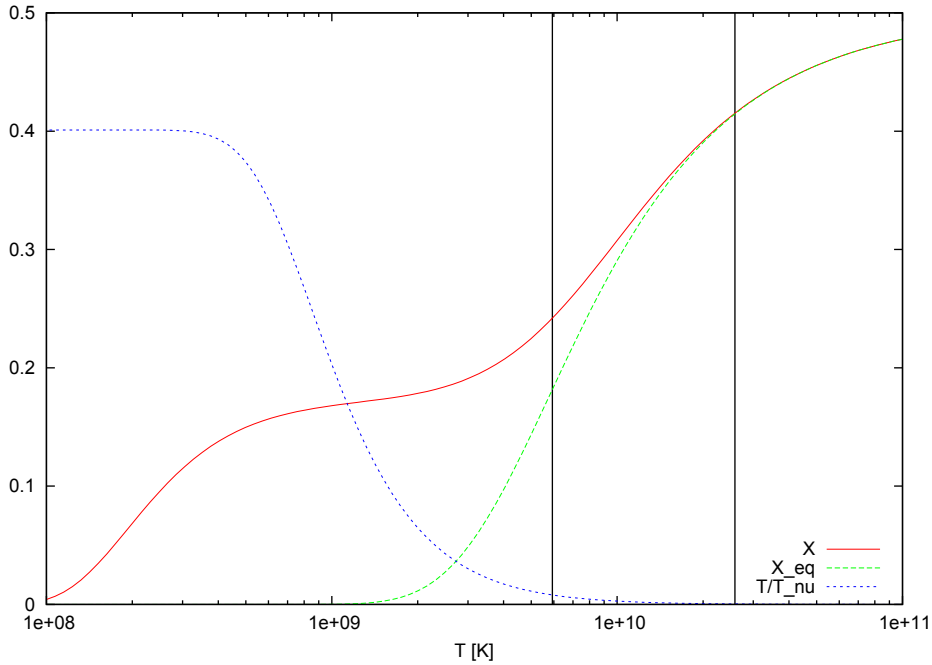

Figure 2: The neutron-to-nucleon ratio X_n versus temperature. Also shown is the equilibrium ratio $X_n^{\rm eq}$ and the ratio of photon and neutrino temperatures $(T - T_\nu)/T_\nu$. The vertical lines mark the temperatures approximately corresponding to electron/positron annihilation (0.511 MeV) and the binding energy of a deuteron (2.23 MeV).

free neutron is

$$\begin{aligned}\frac{{\rm d}\Gamma_{^2{\rm H}}}{{\rm d}N_n} &= \left(4.55\times10^{-20}\,{\rm cm}^3/{\rm s}\right) n_p \\ &= \left(2.52\times10^{4}\,{\rm s}^{-1}\right)\left(\frac{T}{10^{10}\,{\rm K}}\right)^3 X_p\Omega_B h^2. \end{aligned} \tag{34}$$

This is quite large even though the universe is young at the time of interest. During radiation-domination, $t \approx \sqrt{45/16\pi^3\mathcal{N}}M_{\rm Pl}T^{-2}$, so ignoring other processes, the number of deuterons created "per neutron" at time t would be

$$\frac{{\rm d}N_{^2{\rm H}}}{{\rm d}N_n} = t\cdot\frac{{\rm d}\Gamma_{^2{\rm H}}}{{\rm d}N_n} \approx 990\left(\frac{T}{10^{10}\,{\rm K}}\right)X_p\left(\frac{\Omega_B h^2}{0.022}\right). \tag{35}$$

This result cannot be taken literally, but it does tell us that deuteron production is very efficient, provided there are protons and neutrons available. Thus, as far as the deuterons are concerned they are in local equilibrium with protons and neutrons, so their abundance is well-approximated by its equilibrium value. We can find this equilibrium

value in relation to the proton and neutron abundances. The connection is made via the chemical potentials: $\mu_p + \mu_n = \mu_{^2\mathrm{H}}$.[9] Generally, the equilibrium number density of a particular kind of nuclei is

$$n = g\left(\frac{mT}{2\pi}\right)^{3/2} \exp\left\{-(m-\mu)/T\right\}, \tag{36}$$

where g is the number of possible spin states, m is the mass, and μ is the chemical potential. We can remove the (unknown) chemical potentials by writing[10]

$$\frac{n_{^2\mathrm{H}}}{n_p n_n} = \frac{3}{4}\left(\frac{2\pi}{T}\frac{m_{^2\mathrm{H}}}{m_p m_n}\right)^{3/2} e^{B_{^2\mathrm{H}}/T}, \tag{37}$$

where $B_{^2\mathrm{H}} = m_{^2\mathrm{H}} - m_p - m_n = 2.23\,\mathrm{MeV}$ is the binding energy of a deuteron. Defining the relative abundance of deuterons as $X_{^2\mathrm{H}} \equiv n_{^2\mathrm{H}}/n_N$, we have

$$\begin{aligned} X_{^2\mathrm{H}} &= \frac{3}{4}\left(\frac{2\pi}{T}\frac{m_{^2\mathrm{H}}}{m_p m_n}\right)^{3/2} e^{B_{^2\mathrm{H}}/T} n_N X_n X_p \\ &= \frac{9}{32\pi}\frac{(H_0/h)^2}{G m_N}\left(\frac{2\pi}{T}\frac{m_{^2\mathrm{H}}}{m_N^2}\right)^{3/2}\left(\frac{T}{T_{\gamma,0}}\right)^3 e^{B_{^2\mathrm{H}}/T} X_n X_p \Omega_B h^2 \\ &= 2.78\times 10^{-26}\left(\frac{T}{T_{\gamma,0}}\right)^{3/2} e^{B_{^2\mathrm{H}}/T} X_n X_p \Omega_B h^2 \\ &= 1.36\times 10^{-13}\left(\frac{T}{10^{10}\,\mathrm{K}}\right)^{3/2} e^{B_{^2\mathrm{H}}/T} X_n X_p \frac{\Omega_B h^2}{0.022} \end{aligned} \tag{38}$$

Clearly, deuterons were quite rare until the temperature dropped below about $10^9\,\mathrm{K} \approx 0.086\,\mathrm{MeV}$. Another way to see this is to consider the number density of photons with energy of at least $B_{^2\mathrm{H}}$, which are those that can dismantle a deuteron. Recall that the

[9] We have assumed that the chemical potentials are negligibly small elsewhere, but this relation is true regardless.

[10] A deuteron has three spin states.

number density of the photons is

$$\begin{aligned} n_\gamma &= \int_0^\infty \frac{\mathrm{d}\epsilon\, \epsilon^2}{\pi^2} \frac{1}{\exp\{\epsilon/T\} - 1} \\ &= \frac{2\zeta(3)}{\pi^2} T^3 \\ &= n_N/\eta. \end{aligned} \tag{39}$$

For temperatures below $10^{10}\,\mathrm{K}$, the partial number density $n_\gamma(\epsilon \geq B_{2_\mathrm{H}})$ is well approximated by the Wien approximation, so we may write

$$\begin{aligned} n_\gamma(\epsilon \geq B_{2_\mathrm{H}}) &\simeq \int_{B_{2_\mathrm{H}}}^\infty \mathrm{d}\epsilon\, \epsilon^2 \exp\{-\epsilon/T\}/\pi^2 \\ &= \frac{T^3}{\pi^2} \left(\frac{B_{2_\mathrm{H}}}{T} \left(\frac{B_{2_\mathrm{H}}}{T} + 2 \right) + 2 \right) e^{-B_{2_\mathrm{H}}/T} \\ &= \frac{n_\gamma^\mathrm{tot}}{2\zeta(3)} \left(\frac{B_{2_\mathrm{H}}}{T} \left(\frac{B_{2_\mathrm{H}}}{T} + 2 \right) + 2 \right) e^{-B_{2_\mathrm{H}}/T}. \end{aligned} \tag{40}$$

Then the ratio of "deuteron-capable" photons to deuterons is

$$\begin{aligned} \frac{n_\gamma(\epsilon \geq B_{2_\mathrm{H}})}{n_{2_\mathrm{H}}} &= \frac{(\eta X_{2_\mathrm{H}})^{-1}}{2\zeta(3)} \left[\frac{B_{2_\mathrm{H}}}{T} \left(\frac{B_{2_\mathrm{H}}}{T} + 2 \right) + 2 \right] e^{-B_{2_\mathrm{H}}/T} \\ &\simeq \frac{2.4 \times 10^{10}}{\eta_{10} X_{2_\mathrm{H}}} \left(\frac{B_{2_\mathrm{H}}}{T} \right)^2 e^{-B_{2_\mathrm{H}}/T} \\ &= \frac{1.6 \times 10^9}{\eta_{10} X_{2_\mathrm{H}}} \left(\frac{10^{10}\,\mathrm{K}}{T} \right)^2 \exp\left\{ -2.6 \times 10^{10}\,\mathrm{K}/T \right\}. \end{aligned} \tag{41}$$

The temperature had to drop well below the temperature corresponding to the binding energy of helium $28.3\,\mathrm{MeV}/k_\mathcal{B} = 3.3 \times 10^{11}\,\mathrm{K}$. This means that the helium abundance will reach its equilibrium value very quickly once the deuterium bottleneck opens.

We now understand when the deuterium abundance becomes significant, so we may proceed to the next step. The two reactions that combine deuterons into heavier nuclei are

$$^2\mathrm{H} + {}^2\mathrm{H} \to {}^3\mathrm{H} + p \tag{42}$$

$$^2\mathrm{H} + {}^2\mathrm{H} \to {}^3\mathrm{He} + n. \tag{43}$$

The cross-sections of these processes at the temperatures of interest have been experimentally measured; we use, for the respective reactions,

$$\langle \sigma v \rangle_1 \simeq 1.8 \times 10^{-17}\,\mathrm{cm}^3/\mathrm{s}, \tag{44}$$

$$\langle \sigma v \rangle_2 \simeq 1.6 \times 10^{-17}\,\mathrm{cm}^3/\mathrm{s}. \tag{45}$$

Via these processes, the total rate at which deuterons interact is

$$\begin{aligned}
\Gamma &= \langle \sigma v \rangle_{\mathrm{tot}} n_N X_{^2\mathrm{H}} \\
&= \left(3.4 \times 10^{-17}\,\mathrm{cm}^3/\mathrm{s}\right) \frac{\rho_{\mathrm{crit},0}}{m_N} \left(\frac{T}{T_{\gamma,0}}\right)^3 X_{^2\mathrm{H}}\,\Omega_B \\
&= \left(3.8 \times 10^{-22}\,\mathrm{s}^{-1}\right) \left(\frac{T}{T_{\gamma,0}}\right)^3 X_{^2\mathrm{H}}\,\Omega_B h^2 \\
&= \left(1.9 \times 10^{7}\,\mathrm{s}^{-1}\right) \left(\frac{T}{10^{10}\,\mathrm{K}}\right)^3 X_{^2\mathrm{H}}\,\Omega_B h^2
\end{aligned} \tag{46}$$

Once these processes are comparable to the expansion rate and become efficient, nucleosynthesis begins, locking in the neutron abundance. This occurs after electrons and positrons annihilate, so the effective number of particles species is $\mathcal{N} = 3.363$ and the expansion rate is

$$\begin{aligned}
H &= 1/2t \\
&= \left(0.28\,\mathrm{s}^{-1}\right) \left(\frac{T}{10^{10}\,\mathrm{K}}\right)^2
\end{aligned} \tag{47}$$

Equating the two rates in equations 46 and 47 gives

$$\frac{\Omega_B h^2}{0.022} \simeq \frac{6.7 \times 10^{-6}}{X_{^2\mathrm{H}}} \left(\frac{T}{10^9\,\mathrm{K}}\right)^{-1} \tag{48}$$

We have already estimated that nucleosynthesis begins at around $T^{\mathrm{Nuc}} = 10^9\,\mathrm{K}$, so this suggests that the deuteron abundance at the beginning of nucleosynthesis was $X_{^2\mathrm{H}}^{\mathrm{Nuc}} \approx 6.7 \times 10^{-6}$. Using equation 38 and estimating $X_n X_p \approx X_n(1 - X_n) \approx 0.14$ from Figure 2, we find that this abundance would be reached in thermal equilibrium at a temperature of

1.14×10^9 K, quite consistent with our starting estimate of 10^9 K. We can safely conclude that this is approximately when helium nuclei formed, capturing all the free neutrons. The fractional abundance of helium by weight Y_p is double the neutron abundance: approximately 25% of nucleons were then contained in helium nuclei, the rest being free protons (that will become atomic hydrogen). It is measured from observations of emission lines of hydrogen and helium in H II regions, metal-poor extragalactic clouds of partially ionized gas. Various groups generally report a value of $Y_p \approx 0.25$ [98, 102, 75, 18, 17], and the PDG recommends the value $Y_p = 0.2465 \pm 0.0097$ [22].

This calculation also predicts the deuterium/hydrogen ratio, which can be measured from ^{2}H I absorption lines in Lyman-α systems, giving $^2\mathrm{H}/^1\mathrm{H} = (2.53 \pm 0.04) \times 10^{-5}$. Note that because deuterium can only be created (without being quickly destroyed) during BBN, any observation sets a lower limit on the primoridal deuterium abundance, which sets an upper limit on η. Past strategies have involved spectroscopic observations of the interstellar medium, giving the ratio $(1.60 \pm 0.09^{+0.05}_{-0.10}) \times 10^{-5}$ [83]; the Jovian atmosphere, giving the ratio $(5 \pm 2) \times 10^{-5}$ [57]; and measurements of ^{3}He in the solar wind, thought to have been converted from deuterium before the Sun entered the main sequence, giving the ratio $(2.6 \pm 0.6 \pm 1.4) \times 10^{-5}$.

While the measurements of the deuterium/hydrogen ratio are now very precise and imply a value of η that is consistent with that by measurements of the helium abundance, measurements of the lithium/hydrogen ratio are troubling. The value measured from low-metallicity stars and dwarfs, $(1.6 \pm 0.3) \times 10^{-10}$, is lower than expected from the BBN calculation [112]. It is not clear how these low-mass stars may burn lithium, so it is difficult to extrapolate the observed ratios to the primordial values. Figure 3 shows the precise predictions of various isotopic abundances as a function of η and measurements of those abundances.

In our estimates we have preemptively assumed that $\Omega_B h^2 = 0.022$ in order to calculate the abundances. This is backward from what is actually done in practice: the abundances are measured and used to calculate $\Omega_B h^2$. The point here is just to show

that nucleosynthesis requires the cosmological density of baryons to be small.[11] This result alone does not necessitate non-baryonic dark matter.We must look to another major event in cosmic history, when gravity comes into play.

2.2 The Cosmic Microwave Background

We have seen that the expanding, cooling model of cosmology explains very well the primordial existence of light nuclei. Another major feature of such a universe is a nearly isotropic Planck distribution of light. This radiation "background" is leftover from the time of "recombination", when the temperature of the photon population became low enough that the matter in the universe, mostly electrons, free protons, and light nuclei, could combine and become neutral without being immediately ionized again. Before this time, the free-streaming length of photons was small. After, the universe became essentially transparent to photons, which continued to traverse, nearly unabated, to be seen today. Most of this background radiation is now in the microwave range, so the whole is called the Cosmic Microwave Background. It was first predicted in 1948 by George Gamow and later discovered (at a single wavelength) in 1965 by Arno A. Penzias and Robert W. Wilson [103]. The first all-sky observations came from the COBE mission in 1992 [117, 61]. Since then, the WMAP [78] and, most recently, Planck missions [2] have measured the temperature of the emission over the entire sky to great precision. These all-sky maps reveal small fluctuations in the CMB temperature with respect to direction, and this temperature anisotropy tells us a great deal in cosmology. We will briefly review how measurements are decomposed, the physical origins of the temperature fluctuations, and the cosmological parameters that are determined by modeling the creation of the fluctuations.

We can express the temperature observed from a particular direction $\hat{n}$ as

$$T(\hat{n}) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\hat{n}), \tag{49}$$

[11] Strictly speaking, the calculation regarding Big Bang Nucleosynthesis *alone* only constrains the baryon-to-photon ratio η. With a basic measurement of the CMB, however, we know the photon density and thus the baryon density.

where $Y_{\ell m}(\hat{n}) = Y_{\ell m}(\theta, \phi)$ are the spherical harmonics and $a_{\ell m}$ are coefficients satisfying $a^*_{\ell m} = a_{\ell -m}$ (because T is real). Equation 49 does not assume anything about the observed temperatures, but the coefficients $a_{\ell m}$ correspond to the temperature fluctuations seen at a single place in the universe (the Earth). A cosmological theory can not (and perhaps *should* not) predict this. From a cosmological standpoint, we would be interested in averages over all locations in the universe, which we denote in this section as $\langle . \rangle$. We will refer to this kind of average as a "cosmological average". If the coefficients $a_{\ell m}$ measured at a single location (such as the Earth) are randomly drawn from a distribution, then a cosmological theory would predict the distribution for each $a_{\ell m}$ and observations would provide a *single sample* from each distribution. An alternative way to think of this is that the randomness is not from the arbitrary location but rather the history of the particular universe we inhabit. Even from just one location, sampling from many iterations of the universe would give us more information about the distributions of $a_{\ell m}$, to which we could compare cosmological predictions.[12]

Unfortunately, we have only a single vantage point from which to observe and only one universe, so we cannot sample more than one point from each of the distributions of $a_{\ell m}$. The expected error of any observed $a_{\ell m}$ compared to $\langle a_{\ell m} \rangle$ is, of course, the variance of the distribution of that $a_{\ell m}$. We can calculate this variance from the predicted distribution, but we cannot measure it as we have only a single measurement of each $a_{\ell m}$. In this context, this is called "cosmic variance" and is unavoidable. There is, however, a way to dampen its effects so to allow for accurate and precise predictions. We assume in theory that the universe is homogeneous and isotropic. Then the CMB temperature fluctuations should, on cosmological average, be rotationally invariant. The *angular* average of $T(\hat{n})$ is simply the average observed temperature $T_{\gamma,0}$, and $\langle T(\hat{n}) \rangle$ would be the cosmological average photon temperature $\langle T_{\gamma,0} \rangle$. More information is contained in the two-point correlation function. Rotational invariance requires that the cosmological

[12]These two distributions — one encapsulating the randomness with location in one universe, the other the randomness at *one* location with different universes — are equivalent by the ergodic theorem [137].

average of the product of two $a_{\ell m}$ is

$$\langle a_{\ell m} a_{\ell' m'} \rangle = \delta_{\ell \ell'} \delta_{m -m'} C_\ell. \tag{50}$$

The cosmological two-point correlation function is then

$$\begin{aligned} \langle T\left(\hat{n}\right) T\left(\hat{n}'\right) \rangle &= \sum_{\ell m} C_\ell \, Y_{\ell m}(\hat{n}) Y_{\ell -m}(\hat{n}') \\ &= \sum_{\ell} C_\ell \left(\frac{2\ell + 1}{4\pi} \right) P_\ell \left(\hat{n} \cdot \hat{n}' \right), \end{aligned} \tag{51}$$

where P_ℓ are the Legendre polynomials. Inverting this equation we find

$$C_\ell = \frac{1}{4\pi} \int \mathrm{d}\hat{n} \, \mathrm{d}\hat{n}' \, P_\ell(\hat{n} \cdot \hat{n}') \langle T(\hat{n}) T(\hat{n}') \rangle. \tag{52}$$

While we cannot actually measure the cosmological average $\langle T(\hat{n}) T(\hat{n}') \rangle$, we *can* measure the angular average of $T(\hat{n}) T(\hat{n}')$ from our particular location, which is equivalent to an average over m:

$$\begin{aligned} C_\ell^{\mathrm{obs}} &= \frac{1}{2\ell + 1} \sum_m a_{\ell m} a_{\ell -m} \\ &= \frac{1}{4\pi} \int \mathrm{d}\hat{n} \, \mathrm{d}\hat{n}' \, P_\ell(\hat{n} \cdot \hat{n}') T(\hat{n}) T(\hat{n}'). \end{aligned} \tag{53}$$

We see that C_ℓ is the cosmological average $\langle C_\ell^{\mathrm{obs}} \rangle$. The question is: how confident can we be that the measured values C_ℓ^{obs} accurately represent the theoretical values C_ℓ? This uncertainty is cosmic variance. The cosmological average of the squared fractional difference between equation 52 and 53 quantifies the effect of cosmic variance.[13] If the random temperature fluctuations are Gaussian, then we have[14]

$$\left\langle \left(\frac{C_\ell - C_\ell^{\mathrm{obs}}}{C_\ell} \right)^2 \right\rangle = \frac{2}{2\ell + 1}. \tag{54}$$

[13] This is the cosmological average of the fractional error that observers throughout the universe would have. Different observers would likely badly disagree over a measurement such as an $a_{\ell m}$ or C_ℓ^{obs} if cosmic variance is significant.

[14] Assuming that $a_{\ell m}$ follow Gaussian statistics.

This means that the uncertainty due to cosmic variance in the measurement of C_ℓ^{obs}, as an approximation of C_ℓ, drops with increasing ℓ. Measurements of higher-order C_ℓ^{obs} do not suffer much from cosmic variance. We can understand this by the following. The distribution of C_ℓ^{obs} is completely determined by the distributions of $a_{\ell m}$ for $m \in [-\ell, \ell]$, and we have used $2\ell + 1$ independent measurements to determine C_ℓ^{obs} in equation 53. The number of samples of C_ℓ thus increases with ℓ, and the estimated average C_ℓ^{obs} becomes more accurate with larger ℓ. All of this depends on the starting assumption of cosmological homoegeity and isotropy. The quantity C_ℓ as a function of ℓ is called the "power spectrum" of the CMB, and it is usually presented as $\ell(\ell+1)C_\ell/2\pi$.

The first two components of the spherical harmonic decomposition are easy to understand. The first, the monopole, is simply the average temperature $T_{\gamma,0} = 2.7255 \pm 0.0006\,\text{K}$. The second, the dipole, is due to the Doppler shift from the motion of the solar system. It's amplitude is $3.355 \pm 0.008\,\text{mK}$ (note that this is *milli-Kelvins*) [70]. This is a frame-dependent quantity, so there is indeed a rest frame with respect to the CMB. The implied velocity for the solar system's barycenter is $369.0 \pm 0.9\,\text{km/s}$ [70].

Higher-order multipoles are the result from inhomogeneities in the density at the time of recombination. Most of the matter with which photons interact is hydrogen nuclei, free protons. Hydrogen has an ionization energy of about 13.6 eV, so we would expect that once most of the photons were below this energy, hydrogen could become neutral and reliably stay that way. This energy corresponds to a temperature of about $1.6 \times 10^5\,\text{K}$. Because of the high-energy tail in the Planck distribution of photons, however, recombination actually occurs at around a temperature of $0.2\,\text{eV}/k_{\mathcal{B}} \approx 2300\,\text{K}$. After recombination, radiation and matter proceeded to freely expand with separate cosmological dynamics, so the photon temperature followed $T_\gamma \propto a^{-1}$ after recombination, as it did during radiation domination. Recombination happens *after* radiation-matter equality, however, so there was a period of time during which the photon temperature varied as $T_\gamma \propto a^{-2}$. The duration of this period at a particular location depends on the local matter density, which thus affects the local photon temperature. These fluctuations are called "intrinsic fluctuations" and map the "initial conditions" at recombination. Futhermore, at the time

of decoupling from matter, photons are influenced by the gravitational potential of the matter. Any fluctuation in the matter density will result in gravitational redshift of the local photons. The combination of the intrinsic fluctuations and the gravitational redshift is called the Sachs-Wolfe effect [110]. The difference in temperature from the average is proportional to the gravitational potential due to the matter perturbation at the time of recombination:

$$\frac{\Delta T\left(\hat{n}\right)}{T\left(\hat{n}\right)} = \frac{1}{3}\phi_{\text{pert}}\left(\hat{n}\, r_{\text{recom}}\right). \tag{55}$$

The Sachs-Wolfe effect is the dominant influence on multipoles with $3 \leq \ell \lesssim 100$.

Multipoles with $100 \lesssim \ell \lesssim 1000$ reflect the acoustic oscillations that photons and baryonic matter underwent just before recombination. Small positive perturbations in the matter density would gravitationally be driven to collapse, while photon-interactions provided an outward pressure. The amplitude of these perturbations were $\mathcal{O}\left(10^{-5}\right)$, so the perturbation Fourier modes behaved linearly and independently from each other. The oscillation frequency depended on the sound speed, which depended on the density and pressure. When recombination occured, some oscillating modes happened to be at their maximum or minimum, while some happened to be at zero. These conditions were frozen after recombination and are reflected in a series of peaks in the CMB power spectrum. The odd-numbered peaks correspond to the modes that were at a maximum when recombination occurred; even-numbered peaks correspond to modes that were at a minimum. The troughs correspond to the modes that were at zero. The value of ℓ at the first peak corresponds to the Hubble radius, *i.e.* the size of the universe, at the time of decoupling.

For most purposes, recombination can be assumed to be instantaneous. It was not exactly so, however, and the multipoles corresponding to length scales (at the time) smaller than the time over which recombination occured, $\ell \gtrsim 1000$, are subject to damping. This can be thought of as diffusion in the matter-photon fluid as the "coupling" goes to zero. This is called "Silk damping" [116].

All of the features of the CMB power spectrum we have described are illustrated in Figure 4. Recent results from the Planck collaboration are also shown. The red curve is

the prediction of the ΛCDM model with the best-fit cosmological parameters, which we will now discuss. The overall normalization of the curve corresponds to the total matter in the universe: the more mass, the more powerful any mode will be. The difference in height between the first two peaks is due to the partial viscosity of the matter and photons. It corresponds to the portion of matter that interacts with photons (*i.e.* the portion that is not dark matter).

At the very least, the standard model of cosmology has eight parameters. Those already described are the Hubble parameter $h = H_0/\left(100\,\mathrm{km/s/Mpc}\right)$, the baryon density $\Omega_B h^2$, the cold dark matter density $\Omega_C h^2$, and the total density $\Omega_{\mathrm{tot}} = \rho_0/\rho_{\mathrm{crit},0}$. The spectral index n and the amplitude A together determine the power-law that describes the initial conditions of the density perturbations $\Delta^2 \simeq A\left(k/k_0\right)^{n-1}$; r is the ratio of tensor to scalar perturbations (recently measured for the first time by the BICEP-2 collaboration [4]); and τ is the integrated optical depth due to Thompson scattering between photons and electrons. Generalizations such as a non-constant $n = n(k)$ and dark energy equation of state parameter $w \neq -1$ are also possible.

What is important for this work are the baryon and cold dark matter densities. The most precise results using CMB observations alone are [3]

$$\Omega_B h^2 = 0.02207 \pm 0.00033, \tag{56}$$

$$\Omega_C h^2 = 0.1196 \pm 0.0031. \tag{57}$$

First, we see that there is excellent agreement with the calculation of $\Omega_B h^2$ from observations of isotopic abundances, as described in Section 2.1.[15] We now also have a calculation of the total *matter* density. The difference of these two is the density of non-baryonic matter $\Omega_C h^2$, which we identify as that of cold dark matter. There is over five times as much by mass as there is baryonic matter!

[15]This is impressive, considering that the time of BBN and the creation of the CMB are separate by 380,000 years.

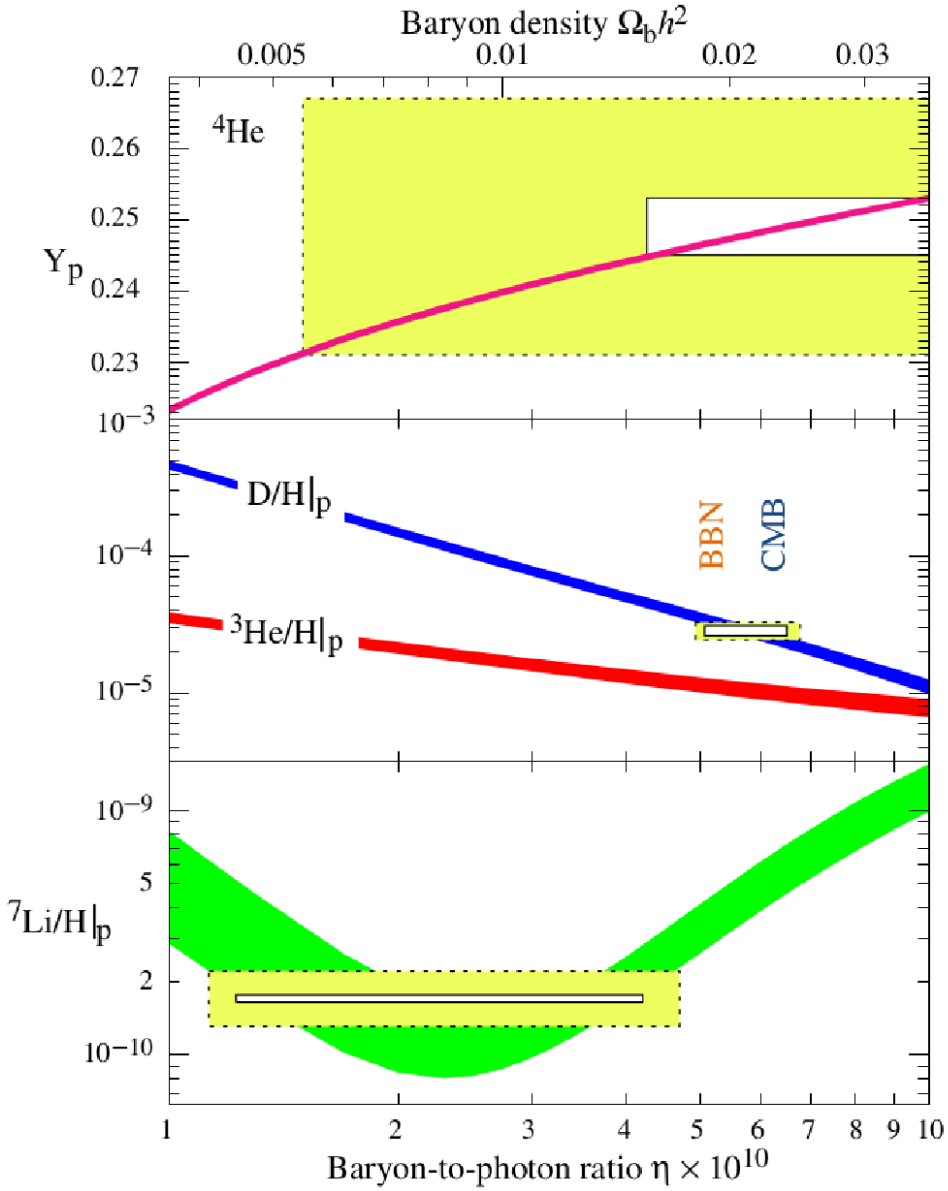

Figure 3: The abundances of ^{4}He, ^{2}H, ^{3}He, and ^{7}Li as predicted by BBN. The bands show the 95% CL range. Boxes indicate the observed light element abundances (smaller boxes: $\pm 2\sigma$ statistical errors; larger boxes: $\pm 2\sigma$ statistical *and* systematic errors). The narrow vertical band indicates the CMB measure of the cosmic baryon density, while the wider band indicates the BBN concordance range (both at 95% CL) [22].

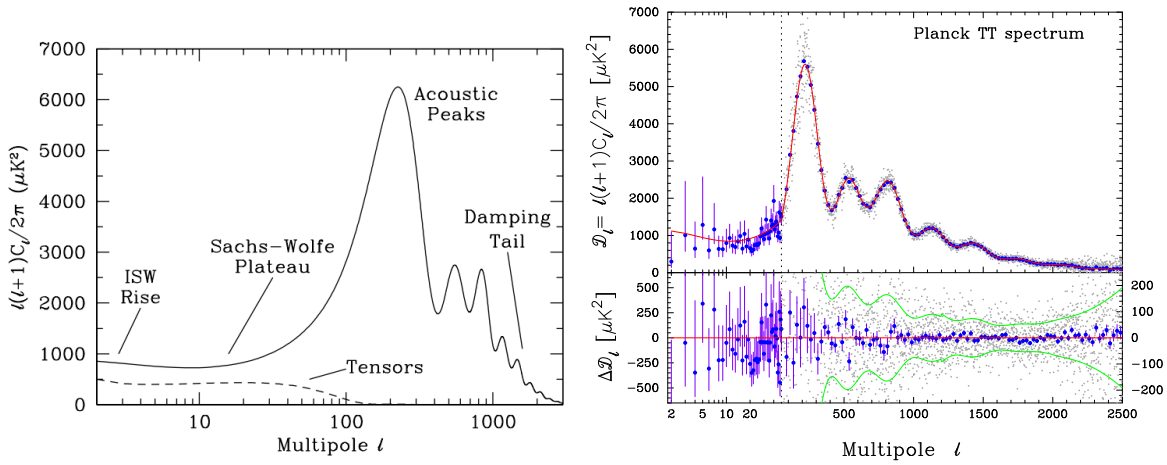

Figure 4: The Cosmic Microwave Background power spectrum. *Left:* a theoretical power spectrum calculated from `CAMB` using a standard ΛCDM model. *Right:* results from the Planck collaboration [3]. Note the enlarged bands of uncertainty from cosmic variance at small ℓ. The uncertainty at large ℓ is from foreground contamination.

3 Galactic Dark Matter

Although the strongest modern evidence for the existence of non-baryonic dark matter comes from cosmological arguments and surveys, early evidence came from observations of clusters of galaxies and, somewhat later, studies of individual galaxies. Regardless of history, all efforts to detect particle dark matter require precise understanding of the distribution of dark matter on galactic scales. The basic concept is that every galaxy, including the Milky Way, is embedded in a larger, nearly-spherical "halo" of dark matter particles. We will review the evidence from virialized clusters of galaxies, arguments about the stability of galaxies, and galactic rotation curves. We will then discuss in Section 3.1 how dark matter halos may be modeled.

One of the earliest realizations of dark matter came from using the virial theorem to measure the mass of clusters of galaxies [141]. This mass was compared to the mass implied by the luminosity, assuming a typical mass-to-light ratio. The *virial theorem* is

$$2T + \Phi = 0, \tag{58}$$

where T and Φ are the total kinetic energy and total gravitational potential energy of a system with total mass M:

$$T = M\langle v^2\rangle/2, \tag{59}$$

$$\Phi = -GM^2\langle 1/r\rangle/2, \tag{60}$$

where the velocity v and position r are measured relative to the center of mass of the system. The virial theorem is valid for systems that have reached a state of statistical equilibrium. This allows us to measure the total mass of a spherical distribution of galaxies from observations:

$$M = \frac{2\langle v^2\rangle}{G\langle 1/r\rangle}. \tag{61}$$

The root-mean-squared velocity $\sqrt{\langle v^2\rangle}$ is measured by the Doppler shift of the X-ray spectrum; $\langle 1/r\rangle$ is found from the angular separation of the individual galaxies. Results

of studies along these lines are usually expressed as a mass-to-light ratio M/L. Typical modern values[16] are 200 to $400h\,\mathcal{M}_\odot/L_\odot$ [40, 41]. This is much larger than the mass-to-light ratios for individual galaxies, $\sim 10h\,\mathcal{M}_\odot/L_\odot$ [29]. We can use the average mass-to-light ratio of clusters of galaxies to estimate the total mass density of the universe from the total luminosity density of the universe $\mathcal{L}$. The total mass density is

$$\rho_M = \mathcal{L}\left(\frac{M}{L}\right)_{\text{clusters}}, \tag{62}$$

In terms of the fraction of the critical density today, this gives us

$$\Omega_M = \frac{(M/L)_{\text{clusters}}}{\rho_{\text{crit},0}/\mathcal{L}}, \tag{63}$$

which is independent of the Hubble constant. With a value of $\mathcal{L} = 2 \times 10^8 h\,L_\odot \text{Mpc}^{-3}$, we find $\Omega_M = 0.15$.

In the late-1960s, Richard Miller, Kevin Prendergast, and, independently, Frank Hohl, wondered about the spiral structure of the Milky Way and other spiral galaxies. Obviously, something was keeping them from simply gravitationally collapsing into a lump of stars. Presumably, the structure was supported by rotation, like the solar system. Such a system is "cold", as opposed to pressure-supported "hot" systems, the particles in which have random orbits. The groups utilized the burgeoning computing power (and techniques that will be described in Section 3.1) to try to simulate the formation of spiral structure as waves in the density of a rotating disk of particles [91, 92, 71]. What they found was surprising: a rotationally-supported disk of the size of the Milky Way is unstable: it rapidly evolved into a pressure-supported axisymmetric form. Particle orbits changed from uniform and circular to elliptical and at random angles to each other. Jeremiah Ostriker pointed out that if the kinetic energy of a spheroidal system is at least 28% in the form of rotational motion, then the system is unstable and tends to increase its moment of inertia by becoming prolate. The initial conditions of the simulations by

[16]The results scale as H_0 because the angular diameter distance scales as H_0^{-1} and the luminosity L scales as H_0^{-2}.

Miller, Prendergast, and Hohl certainly fell in this regime, so the apparent instability was correct. Meanwhile, the kinetic energy of the Milky Way is almost entirely rotational: our galaxy should be very unstable![17] As a solution, Ostriker suggested that there is another component of the galaxy: a non-rotating spheroid that holds a large portion of the total mass. Then the galaxy as a whole would truly be pressure-supported. This halo would have to be dark, since there is no luminous component of the right shape and sufficient size. Ostriker worked with James Peebles in 1973 to simulate such a composite system [100]. They confirmed the instability found by Miller, Prendergast, and Hohl, and then found that the new spheroidal component did indeed provide the necessary stability for the system. The disk component that we really see kept its own rotational structure. For the purposes of stability, this new dark halo needed to extend only to about the position of the Sun and have at least 50% of the total galactic mass. Other observations from this decade, however, would show that we need much more.

Galaxies, in particular rotating disk galaxies, are non-relativistic systems, so it is reasonable to expect its particles, stars and plasma, to follow Newtonian dynamics. The Newtonian gravitational force on a constituent "tracer" of mass due to the rest of the system creates the centripedal acceleration that drives its circular motion:

$$g_N = V_{\rm rot}^2/r. \tag{64}$$

If we can consider the system's mass as smoothly distributed (*i.e.* approximate it as a fluid) then the gravitational force is just given by the enclosed mass:

$$g_N = GM/r^2. \tag{65}$$

Equating these two gives a relationship between the rotational velocity and location of a

[17] As a quick illustration of this, the Sun has a rotational velocity of about 200 km/s and additional random velocity of about 40 km/s. We see that $T_{\rm rot}/\left(T_{\rm rot}+T_{\rm rand}\right) = 200^2/(200^2+40^2) \approx 0.96$.

tracer and the mass within that tracer's position:

$$M = V_{\rm rot}^2 r/G, \tag{66a}$$

$$V_{\rm rot}^2 = \sqrt{GM/r}. \tag{66b}$$

One might expect these equations to hold for galaxies just as well as they do for the solar system. If we have tracers that are outside the majority of the mass (the planets being outside the Sun is an exaggerated example), then their rotational velocity should simply fall as $1/\sqrt{r}$. This is a Keplerian rotation curve, and it is not generally seen in galaxies! Neutral hydrogen is an excellent mass tracer of these galaxies, as it extends well beyond the stellar edge of spiral galaxies. The 21-cm line from hyperfine-splitting in the hydrogen atom and the subsequent Doppler-shift of this radio emission allows for the measurement of the velocity of the gas [108, 32, 19, 52]. Figure 5 is a modern example of these observations [20]. The measured velocity curve of stars and gas in the dwarf spiral galaxy NGC 6503 is flat at large radii. The dashed line denotes the velocity curved expected from the mass contributions of stars alone, which exhibits the Keplerian behavior beyond the stellar edge of about 5 kpc, a typical radius of dwarf spiral galaxies. Assuming Newtonian dynamics is indeed correct, there must be an additional component aside from stars and gas that has a large mass and extends well beyond the stellar disk. A dark halo provides a solution. The dot-dashed line in Figure 5 is the contribution from a halo model.

3.1 Halo Models

We have established that DM structures originate from density fluctuations that gravitationally collapse and grow with time as the universe expands. Linear perturbation theory gives us an idea of the early history of these objects, but we cannot continue this analysis through the non-linear regime to today. Furthermore, the development of DM structures is complicated by the presence of baryons, which are not an insignificant gravitational influence, and baryon-photon interactions. Effects include gas dynamics,

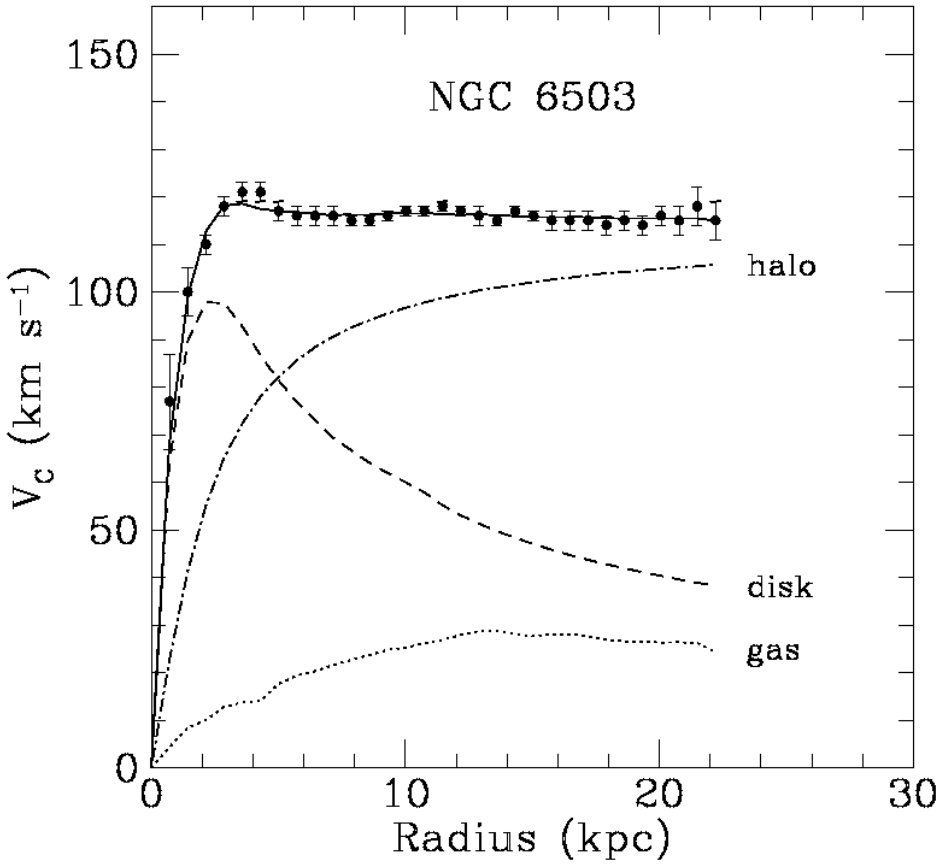

Figure 5: The measured velocity curve of the dwarf spiral NGC 6503, compared to the contributions from the stellar and gaseous components, based on luminous distribution. Adding the contribution from a theoretical DM halo provides a good fit to the flat measured curve. From [20].

radiative cooling, photoionization, recombination, and radiative transfer. If and when an indirect signature of DM is discovered, we will be able to measure its distribution in galactic systems. In the meantime, to predict the precise galactic and intergalactic distribution of DM we must turn to numerical computation. The primary approach to this problem is *N-body simulation*, the history of which goes back to Erik Holmberg in 1941 [72]. The incredible steady increase in computer power, illustrated by Moore's law, has made these efforts more and more precise, reliable, and illuminating (see Figure 6). These simulations have two types: collisional, for which individual particles are tracked as a Hamiltonian system, and collisionless, for which a system of particles is approximated as a continuous fluid following the Boltzmann equation (eq. 142). The basic algorithm

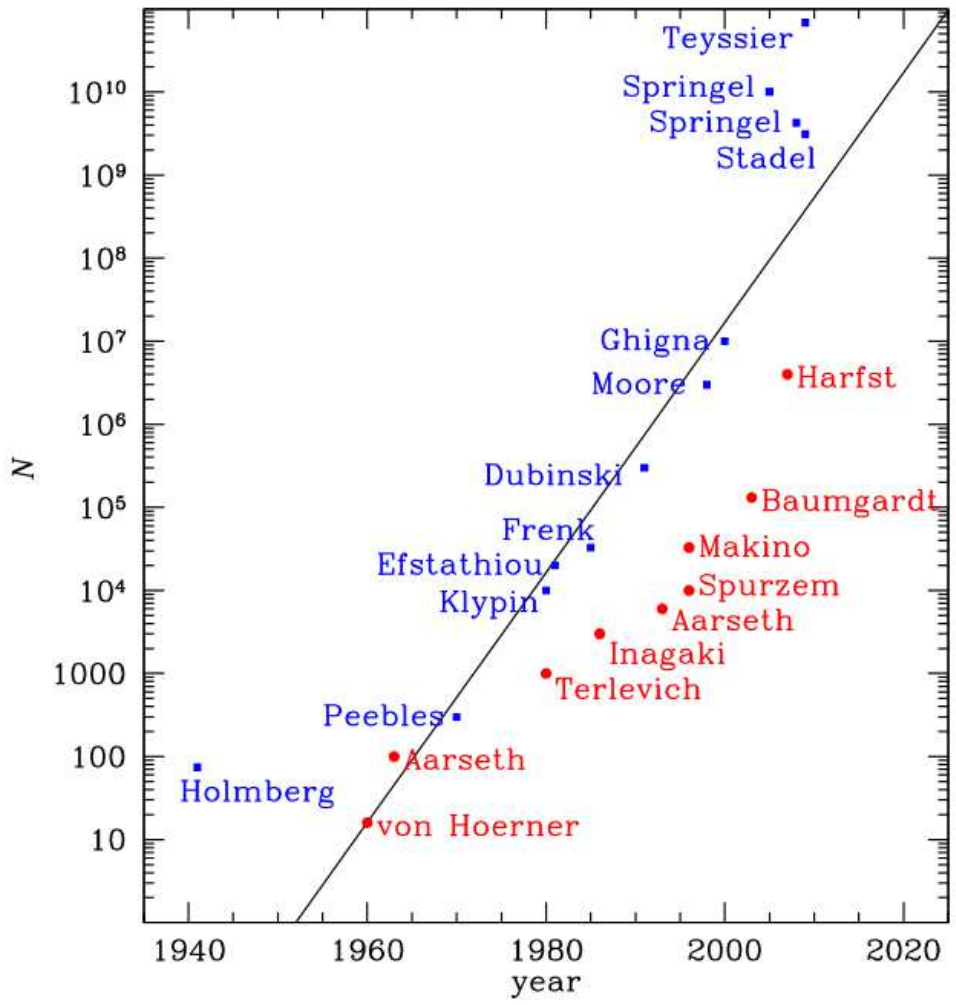

Figure 6: Complexity of N-body simulations versus time. This is an example of Moore's law. From [54].

is the same, however. Given the state of a gravitational system, either a discrete list of particle locations and velocities or samples of the phase-space distribution, calculate the gravitational force on each particle or small volume of phase-space, then evolve the system by some small, discrete time step. One then has a new state of the system and continues the process, presumably until it converges to statistical equilibrium. This is called a *Poisson solver*.

The reliability of N-body simulations is essentially quantized by their length- and mass-scales. In modern simulations of cluster-scale structure and larger, the size of the "imaginary box" is hundreds to thousands of Mpc to a side, with as many as 100 billion DM "particles", each with mass of $\approx 10^6\,\mathcal{M}_\odot$. Galactic-scale simulations are a few Mpc to a side. A force softening length scale is necessary to avoid the divergence that occurs

when two particles are very close to each other. Any results below this length scale are unreliable: in cluster-scale simulations this is around 1 to 10 kpc; in galactic-scale simulations this is around 10 pc.

Different groups have found somewhat different results, but they generally find that all simulated halos follow a *universal profile*, regardless of mass and epoch [95, 28, 118, 33, 14]. The mass distribution or density profile of any particular halo is

$$\rho(r) = \rho_{\rm s} g(r/r_{\rm s}), \tag{67}$$

where $\rho_{\rm s}$ and $r_{\rm s}$ are a scale density and scale radius, both of which may vary from halo to halo. The dimensionless function g, a function of the dimensionless radius $x \equiv r/r_{\rm s}$, is universal: it is the same for all halos. Two-power models are common:

$$g_{\rm 2-pow}(x) = x^{-\gamma_0}(1+x^{\alpha})^{-(\gamma_\infty-\gamma_0)/\alpha}. \tag{68}$$

Choosing the exponents $(\alpha, \gamma_0, \gamma_\infty)$ sets the profile model: the negative log-log slope of the density changes at around $x = 1$ from γ_0 at smaller radii to γ_∞ at larger radii. The "sharpness" of this transition is determined by the parameter α. An outer slope of $\gamma_\infty = 3$ is generally found in halo simulations, but there is some contention over the inner slope γ_0. Navarro, Frenk, and White (NFW) found an inner slope of $\gamma_0 = 1$, and this profile is very commonly used for halos of many scales (including this work), but other studies favor the Moore profile with $\alpha = \gamma_\infty = 1.5$ [93].

Another model we will use for the Milky Way halo is the Einasto profile. This profile was first used to model the luminosity of early-type galaxies and the gaseous components of clusters of galaxies, but has also been found to be a good fit to DM halos [39]. Its functional form is

$$g_{\rm Ein}(x) = \exp\left\{-\frac{2}{a}\left(x^a - 1\right)\right\}, \tag{69}$$

where taking $a = 0.17$ profiles a good fit to galactic- and cluster-sized halos in simulations [96, 67].

The halo models mentioned so far are "cusped" profiles: the density goes to infinity at the center. Of course, this does not mean the mass is infinite, and it is presumed that these models break down at small radii (remember that N-body simulations, from which we measure these profiles, have a limited resolution). Furthermore, the baryonic component of galaxies dominates near their centers, and baryons have only very recently been included in precise simulations [79]. There is also the influence of the central black hole [65, 130, 111]. If DM self-annihilates, then there is a maximum central density $\rho_{\rm core} = m_\chi/\langle\sigma v\rangle t_{\rm halo}$, where $t_{\rm halo}$ is the age of the halo.

While cusped profiles work well in larger galaxies such as the Milky Way, they are at odds with dynamical measurements of dwarf spheroidals [134, 11]. A better fit to these systems is found using "cored" profiles. The density in cored profiles becomes constant inside some "core radius". In particular, the Burkert profile [37] is effective in modeling the DM halo of dwarf spheroidals:

$$g_{\rm Bur}(x) = (1+x)^{-1}\left(1+x^2\right)^{-1}. \tag{70}$$

With a model of the mass distribution, the gravitational potential of a halo is calculable in principle, though this can be difficult. An alternative method is to first assume a potential model. This is good for making constraints on halo qualities using astronomical observations of tracer objects [35, 36].

In principle, the velocity distribution of particles in N-body simulations can be measured as well, but this is difficult because of poor resolution[18] and because there is no simple, universal model for the velocity distribution like there is for the mass distribution. In some simplified cases such as that of an isolated halo the velocity distribution and anisotropy profile is measured [132], but it is difficult to model these measurements. The empirical two-parameter model proposed in [88] provides a good fit to the velocity distribution at different locations of various simulated halos, but there is much scatter in the parameters, so no universal model is evident. For modeling the velocity distribution

[18] One would be interested in the velocity distribution at a particular location in the halo, perhaps that corresponding to the location of the Sun, meaning that particles inside a small volume must be sampled, unlike how samples of the contained mass is done.

at a particular location in a given halo, however, it has been successfully utilized [79, 74]. We generalize this model in Section 8.1.

An interesting alternative quantity to measure in simulated halos is the "pseudo-phase-space density" (PPSD), defined as

$$Q \equiv \frac{\rho}{\sigma_{\rm r}^3}, \tag{71}$$

where $\sigma_{\rm r} \equiv \sqrt{\langle v_{\rm r}^2 \rangle}$ is the radial velocity dispersion, which in general varies with position. In many simulations, the PPSD is found to vary in halos as a power-law over a large range of radii [123, 131, 87, 132]. A negative slope that is slightly shallower than 2 is typically found (also see [53]). Note that a value of exactly 2 would correspond to the case of an isothermal sphere, which has $\rho \propto r^{-2}$ and constant $\sigma_{\rm r}$. In Section 8.1, we will explore the impact of assuming such a power-law for the PPSD.

3.2 The Milky Way

Here we specify the models and parameters used later to describe the Milky Way (MW) baryonic and dark matter components. In Section 7.4 we use a NFW profile and an Einasto profile. A concentration of $c = 10$ and virial radius of $r_{\rm vir} = 200\,{\rm kpc}$ is chosen for both profile possibilities. The total mass of the MW halo is then approximately $10^{12}\,\mathcal{M}_\odot$.

We also model the baryonic bulge and disk of the MW in Section 7.4. Following [121], we take a spherically-symmetric Hernquist potential for the bulge:

$$\Phi_{\rm bulge}(r) = -\frac{GM_{\rm bulge}}{r + c_0}, \tag{72}$$

where $c_0 \sim 0.6\,{\rm kpc}$ and $M_{\rm bulge} = 1.5 \times 10^{10}\,\mathcal{M}_\odot$. We model the disk by a spherical distribution that approximates the mass and circular velocity of the exponential disk:

$$\Phi_{\rm disk}(r) = -\frac{GM_{\rm disk}}{r}\left(1 - e^{-r/b_{\rm disk}}\right), \tag{73}$$

where $b_{\rm disk} \sim 4\,{\rm kpc}$ and $M_{\rm disk} = 5 \times 10^{10}\,\mathcal{M}_\odot$. Of course, the galactic disk is not spherical,

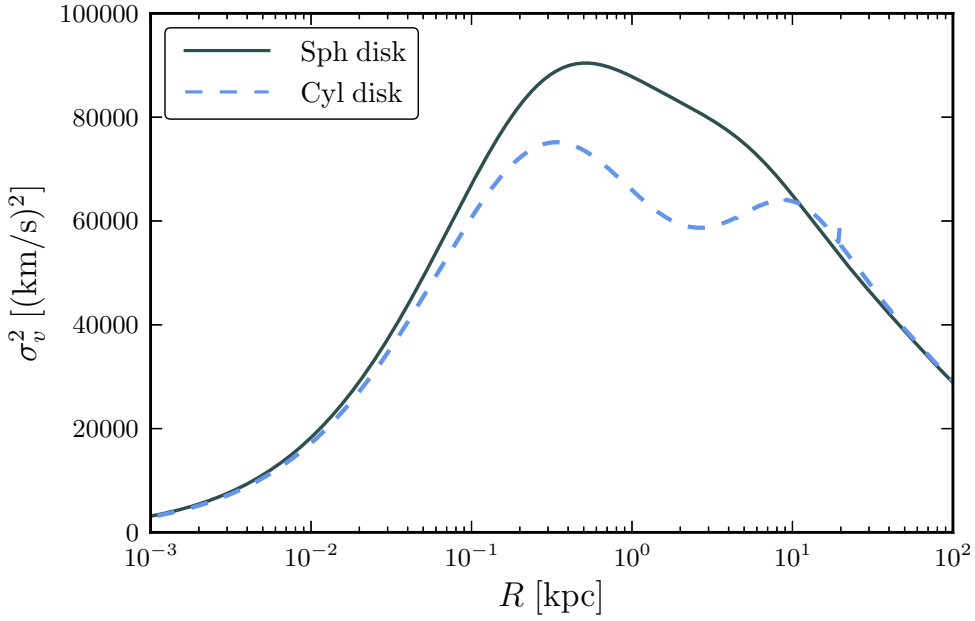

Figure 7: Velocity variance profiles of the spherical disk model in equation 73 and a more appropriate exponential disk with the same mass function.

	lower 95%	lower 68%	mean	upper 68%	upper 95%
$M_{\rm vir}$ [$10^{12}\,\mathcal{M}_\odot$]	1.23	1.33	1.49	1.64	1.86
c	13.93	16.59	19.70	22.90	24.6
$\rho_\odot$ [GeV/cm^3]	0.338	0.365	0.389	0.414	0.435
$r_\odot$ [kpc]	7.67	8.00	8.28	8.55	8.81
$\sigma_{\rm tot,\odot}$ [km/s]	276.7	281.7	287.0	292.2	297.2
$v_{\rm esc}$ [km/s]	528.5	539.7	550.7	561.7	573.3

Table 1: Assumed ranges for the halo parameters, solar radius, local total velocity dispersion, and local escape speed. Taken from Table 3 of [42] and from Table 1 of [43].

but the distribution in equation 73 contains the same amount of mass interior to a radius r as an exponential disk. We further check the validity of this simplified model by calculating the velocity variance profile of each from the Jeans equation. These are shown in Figure 7; the velocity dispersion differs with error no more than $\sim 15\%$.

In Section 8.1 we are interested in a particular derived result, as opposed to just the example of a method in Section 7.4, so we use a range of parameters that may describe the MW halo. These are given in Table 1.

4 Candidates

We see that cosmology and astrophysics requires a new massive, non-baryonic component. So far we have an idea of the role of dark matter in these contexts: we know what it *does*. We do not, however, have a firm understanding of what it *is* in the sense of particle physics. The situation is similar, perhaps, to postulating the existence of electric currents and observing their influence on charged matter, but not knowing the properties of electrons. Here we briefly discuss some particle candidates and then focus on the broad class of WIMPs and the supersymmetric neutralino.

The standard model of particle physics does not offer any viable dark matter particle candidate: there is no non-baryonic, massive, neutral particle that alone can fill the cosmological and astrophysical roles of dark matter. However, it comes close: standard model neutrinos are non-baryonic and neutral, and observations of solar emission suggests that they are actually massive (implying that the SM is incomplete). For these reasons (and that they are known to exist), neutrinos seemed a good possibility. We now know, however, that neutrinos can constitute only a small fraction of the total non-baryonic matter in the universe. As discussed in Section 2, the cosmological neutrino population expanded freely after decoupling and, since they are so weakly-interacting, have approximately kept the same number density. Using $T_{\nu,0} = 1.945\,\mathrm{K}$ and equation 227 for fermions, the number density of each family of neutrino is $112\,\mathrm{cm}^{-3}$ ($g = 2$ since each neutrino and anti-neutrino has one spin state), so the total mass density of neutrinos today is

$$\rho_{\nu,0} = \left(336\,\mathrm{cm}^{-3}\right) \sum_i m_i, \tag{74}$$

where we sum over the masses of each family of neutrino $i = \{\nu_e, \nu_\mu, \nu_\tau\}$. The fraction of the total density of the universe today is thus

$$\Omega_\nu h^2 = 0.016 \sum_i \frac{m_i}{\mathrm{eV}}. \tag{75}$$

This is a cosmological bound on the neutrino masses [49]. If neutrinos were to constitute all the dark matter, the total mass of all three types must be about $7.5\,\mathrm{eV}$.

Since cosmic neutrinos are relativistic until very recently, their free-streaming length is large ($\sim 40\,\mathrm{Mpc}$), so smaller-scale structure would be erased, resulting in "top-down" structure formation [30]. This is contrary to observations of early galaxies [31]. Thus, astrophysical observations constrain the total neutrino mass to no more than about $0.2 - 0.3\,\mathrm{eV}$ [124, 105, 140], meaning that $\Omega_\nu h^2 < 0.0048$, clearly ruling out the possibility that neutrinos are the main dark matter component.[19]

It is possible that there are additional species of neutrinos that mix with the standard model neutrinos but otherwise do not have weak interactions. These "sterile" neutrinos could be produced by standard model neutrinos via a small lepton asymmetry, and they would have a non-thermal spectrum. Because they would be non-thermal, sterile neutrinos evade the problems of standard model neutrinos described above [114]. Sterile neutrinos would be "warm" dark matter and are a viable candidate, though structure formation requires the mass of such new particles to be at least about $3\,\mathrm{keV}$.

Axions are hypothetical spinless particles that were introduced as part of a proposed solution to the strong CP problem in particle physics [136, 138]. Despite its unrelated origin, the axion is also a dark matter candidate [115]. Laboratory searches and observations of supernovae and the sun require the axion mass be very small, $\lesssim 0.01\,\mathrm{eV}$, and they must interact with standard model particles very weakly [107]. Similar to sterile neutrinos, this means axions would be a non-thermal relic. Unlike sterile neutrinos, however, and despite their small mass, axions would be non-relativistic and could satisfy bottom-up structure formation.

There are many other less well-studied hypotheses of particle dark matter (see [26, 122]). All candidates essentially fall under one of two categories: thermal or non-thermal, the former meaning that the particle population was in equilibrium at some point in cosmic history. Standard model neutrinos are thermal, while sterile neutrinos and axions are non-thermal. Candidates can further be classified by whether they were non-relativistic (cold) or relativistic (hot) at the time of decoupling. We will focus on the broad case of Weakly-Interacting Massive Particles (WIMPs), which are *cold thermal relics*. This focus

[19] Meanwhile, observations of solar and atmospheric neutrinos give a *lower* limit on the total neutrino mass of about $0.06\,\mathrm{eV}$ [22]. Also see [82].

on cold thermal relics is motivated by the needs of structure formation, the possibility of supersymmetry, which offers a viable dark matter candidate among other appealing features described in Section 4.1, and by the relic density calculation in Section 4.2.

4.1 Supersymmetric Dark Matter

All of the established symmetries in particle physics link bosons to bosons and fermions to fermions. *Supersymmetry* would link bosons to fermions and vice-versa. The framework was first developed totally independent from astrophysics. Rather, it was initially attractive for solving the hierarchy problem of particle physics and for unifying the gauge couplings at some large energy scale. It also suggests the unification of particle physics and gravity at the Planck energy. Supersymmetry first implies that every boson/fermion particle species has at least one associated fermion/boson species. Supersymmetry cannot be an exact symmetry, because then the new implied particles would have mass equal to that of their associated Standard Model particles, and we do not observe these "superpartners". The symmetry being broken, however, means that the masses of these new particles may be quite large and thus difficult to create in collider experiments. This symmetry-breaking should be soft and be associated with TeV-scale energies, but the details of this process are unknown and introduce most of the ambiguity in the theory. Still, supersymmetry offers a stable weakly-interacting massive particle that is a promising dark matter candidate.

The minimum number of new fields added to the Standard Model to make it supersymmetric is a single new particle for each SM particle, their spins being different by 1/2, and additional Higgs fields. The particle content is shown in Table 2. This is called the Minimial Supersymmetric Standard Model (MSSM). A fermionic supersymmetric particle is designated by the name of its SM counterpart, but with the postfix "-ino", generally called a bosino. A bosonic supersymmetric particle is designated by the name of its SM counterpart, but with the prefix "s-", generally called a sfermion.

Combinations of SM fields and their supersymmetric counterparts are called "supermultiplets". The gauge supermultiplet includes the gluons and their superpartner

gluinos, and the electroweak gauge bosons and their superpartner gauginos. The matter supermultiplet includes the three generations of quarks and leptons, their superpartner squarks and sleptons, and all the associated anti-particles. The Higgs supermultiplet has *two* complex Higgs doublets, the superpartner Higgsinos, and the anti-particles. The additional SM Higgs field is necessary so that both up- and down-type quarks and leptons can simultaneously have mass while being consistent with the supersymmetry. The fields of supermultiplets are then made components of "superfields". Vector superfields contain the gauge and gaugino fields; chiral superfields contain the matter and Higgs supermultiplets.

A supersymmetric Lagrangian contains Yukawa coupling and mass terms that are themselves supersymmetric. Combined with gauge invariance, gaugino fields are coupled to the fields of matter, Higgs, and the matter/Higgs superpartners. The rest of the Lagrangian is constructed by including all possible interaction terms that satisfy the usual $\mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1)$ symmetry and $B - L$ symmetry (invariance of the difference between baryon number and lepton number). Finally, soft symmetry-breaking terms are included.

We may always define the quantity $R = (-1)^{3(B-L)+2S}$, where S is spin. We see that SM particles have $R > 0$ and supersymmetric particles have $R < 0$. An additional quality of the MSSM is R-parity, which enforces the multiplicative conservation of R. This was not originally proposed as part of the MSSM, but as a way to ensure the stability of the proton in all supersymmetric models, so it is not generally essential (though the stability of the proton should be ensured somehow). It would mean, however, that supersymmetric particles can be created from SM particles only in pairs, such as in collider experiments. More important for this work, R-parity also implies that the lightest supersymmetric particle (LSP) is *stable* and an inevitable result of the decay of all heavier unstable supersymmetric particles. The LSP, if it is neutral, is a dark matter candidate offered by supersymmetry.

The gauginos and higgsinos mix, the charged fields creating two "chargino" mass states, and the neutral fields creating four "neutralino" mass states. For the purposes of

Superfield	SM particles	Spin	Superpartners	Spin
Q	$\begin{pmatrix} u_L \\ d_L \end{pmatrix}$	1/2	$\begin{pmatrix} \tilde{u_L} \\ \tilde{d_L} \end{pmatrix}$	0
U^c	$\bar{u}_R$	1/2	$\tilde{u}_R^*$	0
D^c	$\bar{d}_R$	1/2	$\tilde{d}_R^*$	0
L	$\begin{pmatrix} \nu_L \\ e_L \end{pmatrix}$	1/2	$\begin{pmatrix} \tilde{\nu_L} \\ \tilde{e_L} \end{pmatrix}$	0
E^c	$\bar{e}_R$	1/2	$\tilde{e}_R^*$	0
H_1	H_1	0	$\tilde{H}_1$	1/2
H_2	H_2	0	$\tilde{H}_2$	1/2
G^a	g	1	$\tilde{g}$	1/2
W_i	W_i	1	$\tilde{W}_i$	1/2
B	B	1	$\tilde{B}$	1/2

Table 2: Field content of the MSSM. Taken from [26].

dark matter, only the latter concern us. The lightest neutralino is often the supersymmetric dark matter candidate.

4.2 Thermal Relic Density

Consider an arbitrary particle that annihilates with its anti-particle. Even if the densities of both particles and anti-particles are equal and their chemical potentials are zero, at some point the number density becomes so low that the annihilation rate vanishes. This results in a left-over number density. In a cosmological context, if these particles were in thermal equilibrium at some early time, we call this particle a "thermal relic", which presently has some abundance or "relic density". The point at which annihilation becomes inefficient, once expansion and annihilation has diluted the number density sufficiently, is called "freeze-out". This happens roughly when the expansion rate increases to match the annihilation rate, $\Gamma \sim H$. We present first a simplified derivation to motivate some basic concepts in section 4 and then a more detailed derivation.

The interaction rate per particle/anti-particle is $n\langle\sigma v\rangle$, where $n = n(t)$ is the homogeneous number density and $\langle\sigma v\rangle$ is the thermal average of the product of the annihilation cross-section and the Møller velocity (for the purposes of this work, this velocity can be considered just the relative velocity of the annihilating particles). Without annihilation, the number of particles in a comoving volume na^3 would be conserved. With annihilation,

the rate of decrease in the number of particles in a comoving volume is $na^3 \times n\langle\sigma v\rangle$. The Boltzmann equation is then

$$\frac{\mathrm{d}n}{\mathrm{d}t} + 3Hn = -\langle\sigma v\rangle \left(n^2 - n_{\mathrm{eq}}^2\right), \tag{76}$$

where n_{eq} is the equilibrium number density, at which the rate of annihilation and creation from the background instantaneously balance. We write "instantaneously", because expansion dilutes the actual number density (due to the second term on the left-hand-side) and because the equilibrium density is not constant in time. It depends on the temperature of the background, which changes due to expansion. For non-relativistic particles of mass m and number of degrees of freedom g, the equilibrium number density at temperature T is Maxwell-Boltzmann:

$$n_{\mathrm{eq}} = g\left(\frac{mT}{2\pi}\right)^{3/2} e^{-m/T}. \tag{77}$$

We now switch to the inverse entropy per particle $Y \equiv n/s$, so $Y_{\mathrm{eq}} = n_{\mathrm{eq}}/s$, where s is the total entropy density. We also change the independent variable from time to the inverse temperature in units of the particle mass, $x \equiv m/T$. During the radiation-dominated era, the entropy density is given by equation 233, so equation 76 becomes [24, 63]

$$\frac{\mathrm{d}Y}{\mathrm{d}x} = -\langle\sigma v\rangle \frac{s}{Hx}\left(1 + \frac{1}{3}\frac{\mathrm{d}\ln\mathcal{N}}{\mathrm{d}\ln T}\right)\left(Y^2 - Y_{\mathrm{eq}}^2\right). \tag{78}$$

This will be the starting place for the more-detailed calculation later. Neglecting the logarithmic derivative of $\mathcal{N}$ and changing variables again to $\Delta \equiv Y - Y_{\mathrm{eq}}$ we have

$$\frac{\mathrm{d}\Delta}{\mathrm{d}x} = -\frac{\mathrm{d}Y_{\mathrm{eq}}}{\mathrm{d}x} - B(x)\Delta\left(2Y_{\mathrm{eq}} - \Delta\right), \tag{79}$$

where we have introduced the dimensionless function

$$B(x) \equiv \sqrt{\frac{\pi\mathcal{N}}{45}}\frac{mM_{\mathrm{Pl}}}{x^2}\langle\sigma v\rangle. \tag{80}$$

The function $\mathcal{N}(T)$ is not constant and has been calculated by a number of groups [63, 80].[20] For this simple calculation, however, we will take it as constant $\mathcal{N} = 3.363$, which is the value today. We know that the equilibrium density $Y_{\rm eq}$ is negligible for $x \gg x_* \approx 20$, so we have

$$\frac{\mathrm{d}\Delta}{\mathrm{d}x} \approx -B(x)\Delta^2. \tag{81}$$

Assuming, for now, that $\langle\sigma v\rangle$ is constant, this can easily be integrated from $x = x_*$ to today at $x = x_0 \approx \infty$. Using $\Delta(x_f) \gg \Delta(x_0)$, we find

$$Y_0 = \sqrt{\frac{45}{\pi\mathcal{N}}}\frac{x_f}{M_{\rm Pl}m_\chi}\frac{1}{\langle\sigma v\rangle}, \tag{82}$$

which tells us the contribution of this relic particle to the total density of the universe:

$$\begin{aligned}\Omega_\chi h^2 &= \rho_\chi/\rho_{\rm crit,0}\\ &= m_\chi s_0 Y_0/\rho_{\rm crit,0}\\ &\approx \frac{3\times10^{-27}\,\mathrm{cm^3 s^{-1}}}{\langle\sigma v\rangle}.\end{aligned} \tag{83}$$

To arrive at equation 83 we have made several approximations and assumptions that are very inaccurate. There are two important points to be learned from this calculation, however: the density today does not depend on the particle mass, except through logarithmic corrections that have been neglected, and that a interaction rate per particle of about $3\times10^{-26}\,\mathrm{cm^3 s^{-1}}$ gives the correct density expected of cosmological dark matter. This interaction rate is weak-scale:

$$G_{\rm F}^2 m_\chi^2 \approx 1.6\times10^{-25}\,\mathrm{cm^3 s^{-1}}\left(\frac{m_\chi}{100\,\mathrm{GeV}}\right)^2. \tag{84}$$

This motivates the generic kind of dark matter candidate called a Weakly-Interacting Massive Particle (WIMP).

Now let us return to equation 78 and find the relic density $\Omega_\chi h^2$ more accurately. In

[20] We have neglected that, in the presence of relativistic particles other than photons, the number of degrees of freedom associated with the energy density and with the entropy density are distinct. For the range of WIMP masses we are interested in, however, this distinction is negligible. See [63, 120].

principle, this equation can be numerically integrated, with the initial condition $Y = Y_{\rm eq}$ at some $x \lesssim 1$. This is a stiff differential equation, however, so it is very difficult to work with directly. Fortunately, there are two separate periods of time in which equation 78 simplifies significantly. We will solve the simplified equation in the first period and connect that with the second to find a final result [120]. Quantities will be designated with a $*$ at this cross-over between regimes, when $x = x_*$. In the earlier period, when $x < x_*$, the true number density closely follows the equilibrium number density, so $\mathrm{d}\Delta/\mathrm{d}x \ll \mathrm{d}Y_{\rm eq}/\mathrm{d}x$. We define the cross-over inverse temperature x_* by $\Delta(x_*) \equiv \mathcal{O}(1)$, where $Y/Y_{\rm eq} \equiv 1 + \Delta$ as before. In other words, we define the moment of cross-over as when the fractional deviation of Y from its equilibrium value becomes significant. Using this we replace Y with Δ and write

$$\frac{\mathrm{d}\ln Y_{\rm eq}}{\mathrm{d}\ln x} + \frac{\mathrm{d}\ln(1+\Delta)}{\mathrm{d}\ln x} = -\langle\sigma v\rangle \frac{sY_{\rm eq}}{H}\left(1 + \frac{1}{3}\frac{\mathrm{d}\ln\mathcal{N}}{\mathrm{d}\ln T}\right)\left(1 + \Delta - \frac{1}{1+\Delta}\right). \tag{85}$$

Recasting equation 77 as

$$Y_{\rm eq}(x) = \frac{45}{2\pi^2}\frac{g}{\mathcal{N}}\left(\frac{x}{2\pi}\right)^{3/2} e^{-x}, \tag{86}$$

we find a differential equation for $1 + \Delta$, with which we can find x_*:

$$\frac{\mathrm{d}\ln(1+\Delta)}{\mathrm{d}\ln x} = -\frac{3}{2} + x + \frac{\mathrm{d}\ln\mathcal{N}}{\mathrm{d}\ln T} - \langle\sigma v\rangle\frac{sY_{\rm eq}}{H}\left(1 + \frac{1}{3}\frac{\mathrm{d}\ln\mathcal{N}}{\mathrm{d}\ln T}\right)\left(1 + \Delta - \frac{1}{1+\Delta}\right). \tag{87}$$

This equation is also difficult to solve numerically, but for $x < x_*$ we may further approximate $\mathrm{d}\Delta/\mathrm{d}x \ll 1$ and, also neglecting the logarithmic derivaties, write[21]

$$x_* + \log(x_* - 1.5) - 0.5\log x_* = 20.5 + \log\left(\langle\sigma v\rangle_*/10^{-26}\,\mathrm{cm}^3/\mathrm{s}\right) + \log m - 0.5\log\mathcal{N}_*. \tag{88}$$

This equation is easily solved for x_*. For the simple case of pure s-wave annihilation, we plot the values of x_* found using the cross-section with which a WIMP of mass m gives the correct relic density. This is shown in Figure 8. We also plot the square root number

[21] This result comes from choosing $\Delta(x_*) = \left(\sqrt{5} - 1\right)/2$. The exact choice has negligible effect on the results.

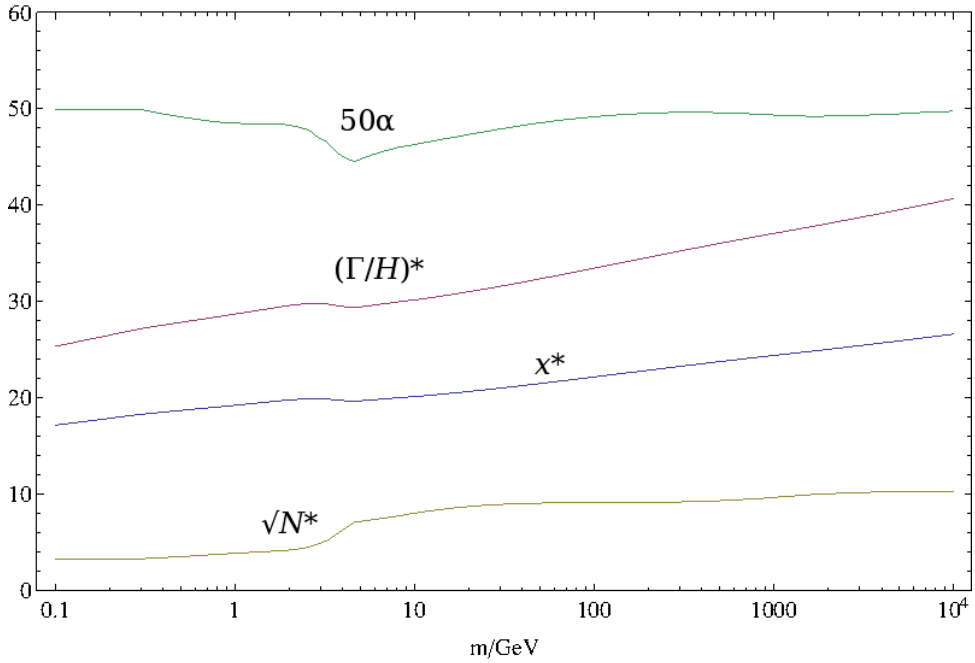

Figure 8: Solutions for $(\Gamma/H)_*$, x_*, and $\sqrt{\mathcal{N}_*}$, in the case of constant $\langle\sigma v\rangle$, that give the correct dark matter relic density $\Omega_\chi h^2 = 0.11$.

of relativistic particle species $\sqrt{\mathcal{N}_*}$ and the annihilation rate divided by the expansion rate $(\Gamma/H)_*$. This ratio continues to drop with increasing x, and it is important to note that annihilation is still efficient over expansion significantly after this cross-over point. In other words, x_* is *not* the freeze-out inverse temperature as has been erroneously assumed in some literature.

After the cross-over time, the deviation of Y from its equilibrium value quickly becomes large, so in the late-time regime $Y_{\rm eq}$ can be neglected. Equation 78 can then be integrated from x_* to around the time of galaxy formation (and WIMP-reheating) at $x_0 \approx m_\chi/20T_{\gamma,0}$ to find $Y_0 = Y(x_0)$. We have

$$\int_{Y_*}^{Y_0} \frac{\mathrm{d}Y}{Y^2} = -\int_{x_*}^{x_0} \mathrm{d}x \langle\sigma v\rangle \frac{s}{Hx}\left(1 + \frac{1}{3}\frac{\mathrm{d}\ln\mathcal{N}}{\mathrm{d}\ln T}\right). \tag{89}$$

If the annihilation is pure s-wave, then we may express the solution for Y_0 as

$$\frac{Y_0}{Y_*} = \frac{1}{1 + \alpha\,(\Gamma/H)_*}, \tag{90}$$

where

$$\alpha \equiv \int_{T_0}^{T_*} \frac{\mathrm{d}T}{T_*} \sqrt{\frac{\mathcal{N}}{\mathcal{N}_*}} \left(1 + \frac{1}{3} \frac{\mathrm{d}\ln\mathcal{N}}{\mathrm{d}\ln T}\right). \tag{91}$$

In Figure 8 we also plot α, multiplied by 50 for clarity, for the case of pure s-wave annihilation. The relic density today is $\Omega_\chi = m_\chi Y(x_0) s_0/\rho_{\mathrm{crit},0}$, so we find

$$\begin{aligned} \Omega_\chi h^2 &= \frac{8\pi G}{3} \left(100\,\mathrm{km/s/Mpc}\right)^{-2} m_\chi Y_0 s_0 \\ &= \frac{9.92 \times 10^{-28}}{\langle\sigma v\rangle} \frac{x_*}{\sqrt{\mathcal{N}_*}} \frac{\left(\Gamma/H\right)_*}{1 + \alpha\left(\Gamma/H\right)_*}. \end{aligned} \tag{92}$$

Knowing the WIMP abundance today, we can calculate the necessary interaction cross-section of a WIMP of given mass.

The interaction rate per particle density $\langle\sigma v\rangle$ is not constant in general. The kinetic average $\langle\sigma v\rangle$ depends on the velocity distribution of the WIMP population, which is a function of temperature (or x). The WIMP distribution is Maxwell-Boltzmann, so we have

$$\langle\sigma v\rangle(x_\chi) = \frac{x_\chi^{3/2}}{2\sqrt{\pi}} \int \mathrm{d}v\, \sigma(v) v\, e^{-x_\chi v^2/4}, \tag{93}$$

where the integral is over all velocities and the cross-section $\sigma(v)$ is a function of velocity determined by the particle physics. Pure s-wave annihilation makes $\langle\sigma v\rangle$ truly a constant; p-wave annihilation was considered in [38]. We will consider s-wave interactions with Sommerfeld enhancement. In this case, the interaction rate per particle density is

$$\langle\sigma v\rangle = \langle\sigma v\rangle_0 \mathcal{S}\left(x_\chi\right), \tag{94}$$

where $\langle\sigma v\rangle_0$ is the constant s-wave rate and $\mathcal{S}\left(x_\chi\right)$ is the Sommerfeld enhancement at the WIMP inverse-temperature x_χ.[22] See Section 5.2 for details about this function. The above calculation changes somewhat due to these new physics. In the early period, Sommerfeld enhancement is negligible, so x_* and $\left(\Gamma/H\right)_*$ do not change significantly, but it becomes important in the later period. The main change is that α in equations 90

[22]Before thermal decoupling, the WIMP temperature follows the photon temperature, so $x_\chi = x$. Afterward, however, the WIMP temperature falls as $x_\chi \propto x^2$.

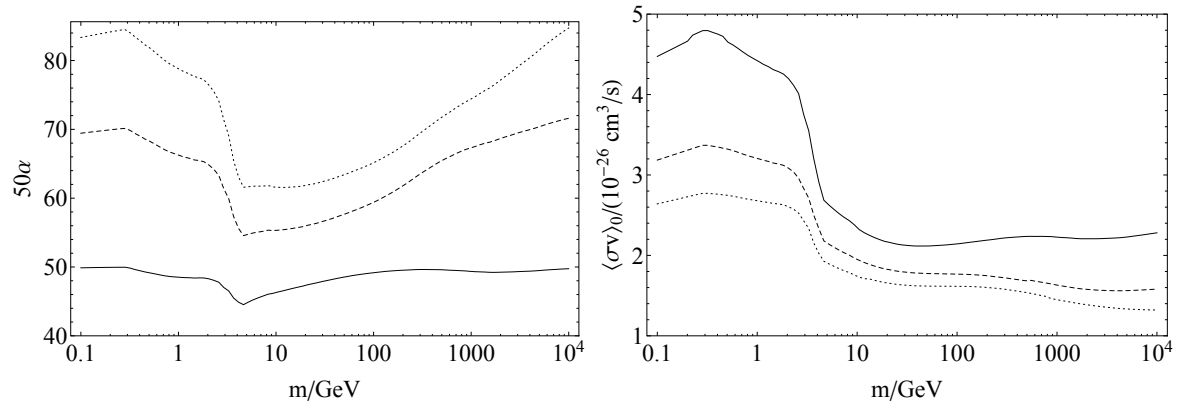

Figure 9: Solutions for 50α or $50\alpha_{\mathcal{S}}$ (left) and $\langle\sigma v\rangle_0$ (right) in the cases of pure s-wave (solid) interactions and Sommerfeld enhanced s-wave interactions far from resonance (dashed) and near resonance (dotted). The Sommerfeld enhancement takes the coupling to be 10^{-2} and, for the dotted curve, the mass ratio to be 1.4×10^{-3} (near the third resonance).

and 92 is replaced by

$$\alpha_{\mathcal{S}} \equiv \int_{T_0}^{T_*} \frac{\mathrm{d}T}{T_*} \frac{\mathcal{S}}{\mathcal{S}_*} \sqrt{\frac{\mathcal{N}}{\mathcal{N}_*}} \left(1 + \frac{1}{3} \frac{\mathrm{d}\ln\mathcal{N}}{\mathrm{d}\ln T}\right). \tag{95}$$

In Figure 9 we compare the value of α found in the pure s-wave case with $\alpha_{\mathcal{S}}$ in two cases of Sommerfeld enhancement, as well as the s-wave cross-section times velocity $\langle\sigma v\rangle_0$ necessary to get the correct relic density.

To generalize, consider some particle physics that determines $\sigma(v)$ and is specified by a set of parameters $\{\epsilon_i\}$. Using equation 93 we can calculate the interaction rate per particle density as a function of inverse temperature x. Following the algorithm above, we can calculate the relic density today and compare the result to the experimental value. If the calculated value is significantly larger, that set of parameters of the particle physics may be ruled out. If the calculated value is significantly *lower* than the experimental value, we cannot rule out the case, because there may be more than one dark matter component.[23] Alternatively, if the cross-section $\sigma(v)$ has some constant factor independent of the parameters $\{\epsilon_i\}$, say the s-wave cross-section σ_0, one can perform the same calculation and find the value of σ_0 that gives the correct relic density. Then the cross-section $\sigma(v)$ is totally specified and can be used in other calculations such as those

[23] After all, baryonic matter has more than one particle component, so why not dark matter also?

described in Section 5.

5 Indirect Detection

Generically, a WIMP undergoes self-annihilation, producing gamma-rays, neutrinos, or other Standard Model particles. Detecting these annihilation products is the basis for the indirect detection of DM [26].

Consider a self-annihilating particle. The annihilation rate is

$$\Gamma(v) = \sigma\Phi, \tag{96}$$

where σ is the annihilation cross-section and Φ is the flux of incident particles as seen by the target particle, defined as the product of the number density n and relative velocity v. In general, the relative velocity between two particles of mass m with four-momenta $p_i = (E_i, \mathbf{p}_i), i = 1, 2$ is

$$v = \frac{\sqrt{(p_1 \cdot p_2)^2 - m^4}}{E_1 E_2}. \tag{97}$$

The cross-section may be a function of the relative velocity, so the annihilation rate on the left side of equation 96 is really the annihilation rate of particles with relative velocity v. To get the total annihilation rate, we must know the relative velocity distribution of the annihilating particles. We average over velocities to get

$$\Gamma = \langle\sigma\Phi\rangle, \tag{98}$$

where $\langle . \rangle$ denotes a kinematic average.

Indirect detection of DM is concerned with the final spectrum of Standard Model particles resulting from such annihilation. Suppose DM particles χ annihilate into some final states consisting of Standard Model particles. In a volume $\mathrm{d}V$ that contains DM particles, the energy-differential rate at which particles of type j are created is

$$\frac{\mathrm{d}^2\Gamma_j}{\mathrm{d}E\,\mathrm{d}V} = n_\chi \sum_i BR_i \frac{\mathrm{d}N_{ij}}{\mathrm{d}E} \langle\sigma_i \Phi_\chi\rangle, \tag{99}$$

where the sum is over all possible reactions, and BR_i is the branching fraction for reaction

i with cross-section σ_i, which produces $\mathrm{d}N_{ij}$ particles of type j with energy between E and $E + \mathrm{d}E$. The DM flux Φ_χ (in this frame) is $n_\chi v$, so we have

$$\frac{\mathrm{d}^2\Gamma_j}{\mathrm{d}E\,\mathrm{d}V} = n_\chi^2 \sum_i BR_i \frac{\mathrm{d}N_{ij}}{\mathrm{d}E}\langle\sigma_i v\rangle, \tag{100}$$

where we now have the thermal average of the cross section times relative velocity. In general, this quantity can be complicated and very model-dependent. The cross section can vary rapidly with energy/velocity near resonances or when a new annihilation channel opens [63]. Furthermore, at early times DM particles had relativistic speeds. This is important to consider when studying the cosmological history of DM, but for the purposes of detection today, we may assume that the DM particles are non-relativistic and that the cross section can be expanded in powers of v^2:

$$\sigma v \simeq a + bv^2 + \mathcal{O}(v^4), \tag{101}$$

where a and b are constants associated with s- and p-wave interactions, respectively [119].

In this work we will discuss two cases in particular: the production of gamma-rays from DM annihilation and the synchrotron radiation from $e^\pm$ created by DM annihilation. Starting with the first case, assuming the annihilation cross section does not depend on the energy of the daughter particles, the production rate per volume of photons of any energy is

$$\frac{\mathrm{d}\Gamma_\gamma}{\mathrm{d}V} = \frac{N_\gamma}{m_\chi^2}\langle\sigma v\rangle\rho_\chi^2, \tag{102}$$

where we have integrated over energy and summed over all channels that result in photons, N_γ being the average number produced per annihilation. The relevant quantity for observation is the photon flux, the rate of photons striking an area incident from a solid angle:

$$\Phi_\gamma = \frac{N_\gamma}{4\pi m_\chi^2}\int \mathrm{d}l\,\mathrm{d}\Omega\langle\sigma v\rangle\rho_\chi^2, \tag{103}$$

where the integral is over line-of-sight and solid angle. In general, the integrand in equation 103 depends on the radius r from the center of the halo producing photons.

If we are measuring the photon flux from the galactic center on Earth, this radius is a function of l and ψ, the angle from the galactic center. In this case we have

$$\Phi_{\gamma,\mathrm{gc}} = \frac{N_\gamma}{2m_\chi^2} \int_0^\infty \mathrm{d}l \int_0^{\Delta\psi} \mathrm{d}\psi \cos(\psi) \, \langle\sigma v\rangle[r(l,\psi)] \, \rho_\chi^2[r(l,\psi)], \tag{104}$$

where $\Delta\psi$ is the opening angle of whatever detector and we have written the spacial-dependence on l and ψ through r explicitly.

If, instead of the galactic center, we are interested in dwarf spheroidals, the point spread function of the detector may contain the entire object, in which case we may simply integrate over the entire halo:

$$\Phi_{\gamma,\mathrm{dSph}} = \frac{N_\gamma}{m_\chi^2} \frac{\Delta\Omega}{4\pi} \int \mathrm{d}r \, \langle\sigma v\rangle(r) \, \rho_\chi^2(r). \tag{105}$$

The factor ρ_χ^2 is trivial, given a halo model, but the factor $\langle\sigma v\rangle$ can be more difficult to handle. The simplest thing to do it assume it is constant: there are only pure s-wave interactions. This approach is a straight-forward way to set upper limits on $\langle\sigma v\rangle$ for any given halo model, using gamma-ray data from, say, the Fermi-LAT instrument. In this case, equation 103 becomes

$$\Phi_\gamma = \frac{N_\gamma \langle\sigma v\rangle}{4\pi m_\chi^2} J(\Delta\Omega), \tag{106}$$

where $\Delta\Omega$ represents the patch of sky under observation and the so-called J-factor is

$$J(\Delta\Omega) = \int \mathrm{d}l \, \mathrm{d}\Omega \, \rho_\chi^2. \tag{107}$$

This formulation is appealing because the J-factor contains all of the astrophysical information about DM, while the rest of the right-hand side of equation 106 depends only on the particle physics. A thorough review of applications of this calculation to the galactic center and dwarf spheroidals is given in [58].

After taking $\langle\sigma v\rangle$ as constant, the next simplest approach is to consider p-wave interactions, the rate of which depend on the velocity dispersion squared, $\sigma^2 \equiv \langle v^2 \rangle$. Of course, one can take this as constant in space, but it is straightforward to calculate it

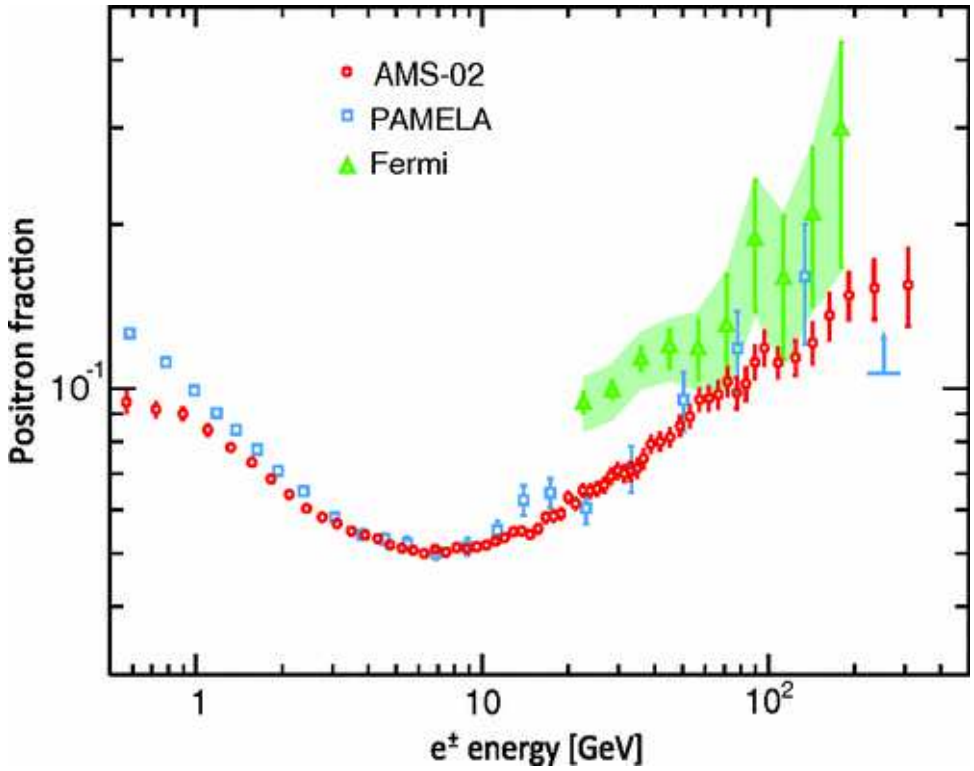

Figure 10: Measurements of the positron fraction $e^+/(e^+ + e^-)$ by the AMS-02, PAMELA, and Fermi-LAT experiments.

using the Jeans equation. This is discussed in Section 7.1.

5.1 The Positron Excess

The Payload for Antimatter Matter Exploration and Light-nuclei Astrophysics (PAMELA) reported a sharp increase in the positron fraction of electron-positron flux at energies of around 10 to 100 GeV [6], which was later confirmed by the AMS-02 experiment [7] (see Figure 10). This is contrary to the usual model of positron production from high-energy cosmic rays propogating in the interstellar medium [94]. There are two possibilities: models of cosmic-ray propogation must be refined (*e.g.* see [104, 46, 45, 47, 48]), or there is an additional source of positrons. Annihilating DM that results in electron-positron pairs would provide this source.

Interpreting the positron excess as a result of DM annihilation introduces difficulties with other observational constraints and the standard relic density calculation (see equation 83). The anti-proton flux has no unexpected behavior, so a new source of positrons cannot also be a source of anti-protons [5, 55]. Furthermore, the production of

hadrons in general must be small, since the production of π^0s from DM annihilation is well-constrained by gamma-ray observations of the galactic center, and the production of electrons and positrons would create synchrotron emission [25, 21, 44]. The other difficulty is that the interaction rate required to create the observed excess is much larger than that expected for a thermal relic. A "boost factor" of $\mathcal{O}(100)$ is needed. Both of these issues can be solved by postulating a new force in the dark sector: DM particles experience a force mediated by a new boson ϕ. Then the annihilation channel $\chi\chi \to \phi\phi$ may be the dominant one. The new ϕ particle may decay to Standard Model particles. If it is light enough, decays to hadrons are kinematically forbidden (π^0s and $p/\overline{p}$). Furthermore, a new force between DM particles can adequately boost the annihilation cross-section by a mechanism called the Sommerfeld Effect, which is described in detail in Section 5.2.

If we assume there is an attractive force between dark matter particles that provides the necessary boost to explain the postron excess, then there is a possibility of contradiction. Any boost to the annihilation rate of local DM must also apply to DM everywhere else. In particular the annihilation rate of DM at the galactic center is also boosted, so the flux of photons from these annihilations receives the same boost. It is possible that the boost from the Sommerfeld effect, while explaining the local positron flux, causes the galactic center emission to exceed observations. This possibility is investigated in detail in Section 7.4.

5.2 The Sommerfeld Effect

The Sommerfeld Effect is the change in the interaction cross-section of a target object or particle (located at the origin) due to a circular potential. If the potential grants an attractive force, then we call this the Sommerfeld Enhancement, since the cross-section is increased. We will briefly describe a classical analogue and then proceed with a quantum mechanical derivation.

Consider a point particle approaching a hard sphere of radius R from infinity with velocity v and impact parameter b. The scattering cross-section is simply the cross-

sectional area of the target sphere $\sigma_0 = \pi R^2$. If $b > R$ then no scatting occurs. Now suppose that the incident particle and the target sphere are gravitationally attracted. Then the scattering cross-section increases because the incident particle is accelerated toward the target sphere. Simply using conservation of energy and angular momentum, the modified scattering cross-section is found to be

$$\sigma = \sigma_0 \left(1 + \frac{v_{\text{esc}}^2}{v^2}\right), \tag{108}$$

where $v_{\text{esc}}^2 = 2GM/R$ is the escape velocity of the target sphere. The quantity in parentheses is the classical *Sommerfeld factor* and, in this case of an attractive force, can vary from 1 to ∞. The smaller the incident velocity, the greater the modification.

The Sommerfeld enhancement is the quantum analogue of the above classical effect. We have a circular potential $V(r)$ centered at the origin and a non-relativistic incident particle. If the incident particle were free (no potential), its wave function would simply be

$$\psi_0(\mathbf{r}) = e^{ikz}, \tag{109}$$

where the subscript 0 denotes the lack of the scattering potential $V(r)$, k is the wave-number or momentum of the particle, and, without loss of generality, we have taken the particle as moving along the z-axis. As in the classical analogue, we say that a scattering has occurred if the incident particle "hits" the target or enters the volume that it encompasses. For the most part, we will assume that the interaction is point-like: scattering occurs if the incident particle is found at the origin. The likelihood that scattering occurs, *i.e.* the cross-section, is thus proportional to the likelihood that the incident particle is found at the origin:

$$\sigma \propto \int \mathrm{d}^3 r \, \delta(r) |\psi(r)|^2 = |\psi(0)|^2. \tag{110}$$

Similarly, in the case without the potential, the unmodified cross-section is $\sigma_0 \propto |\psi_0(0)|^2 = 1$. With the potential, the modification to the cross-section, *i.e.* the Sommerfeld factor,

is

$$S = \frac{|\psi(0)|^2}{|\psi_0(0)|^2} = |\psi(0)|^2. \tag{111}$$

To find this, we must solve the Schrödinger equation.

An axisymmetric wave function subject to a circular potential may be written as

$$\psi(r,\theta) = \sum_{l=0}^{\infty} \frac{i^l e^{\delta_l}\,(2l+1)}{k} P_l\left(\cos\left(\theta\right)\right) R_l(r), \tag{112}$$

where δ_l is the phase shift associated with the partial wave solution labeled by l, $P_l(x)$ is the lth Legendre polynomial, and $R_l(r)$ is the radial wave equation for the lth solution. Since the Legendre polynomials are orthogonal, equation 111 becomes

$$S = \sum_{l=0}^{\infty} \left[(2l+1)\frac{R_l(r)}{k}\right]^2. \tag{113}$$

The equation for the radial wave equations $R_l(r)$ is

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R_l\left(r\right)}{\mathrm{d}r}\right) - \frac{l\left(l+1\right)}{r^2}R_l\left(r\right) + 2M\left(E - V(r)\right)R_l(r) = 0, \tag{114}$$

where $E = k^2/(2M)$ is the energy of the incident particle. For our purposes, we are interested in interactions described by an attractive Yukawa potential

$$V(r) = -\frac{\alpha}{r}e^{-m_\phi r}, \tag{115}$$

with coupling constant α and force-carrier mass m_ϕ. Recall that we are ultimately interested only in the wave function at $r = 0$. In the limit $r \to 0$, equation 114 is dominated by the first two terms[24]

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}R_l\left(r\right)}{\mathrm{d}r}\right) - \frac{l\left(l+1\right)}{r^2}R_l\left(r\right) = 0, \tag{116}$$

which, for $l \neq 0$, has solutions $R_{l\neq 0}(r) \propto r^l$, which vanish at the origin. Since we only

[24] As long as $V(r)$ does not diverge faster than $1/r$.

need the value of the wave function at the origin, we may ignore the $l \neq 0$ partial wave functions and focus on the $l = 0$ case. The Sommerfeld factor is then

$$S = \left| \frac{R_0(0)}{k} \right|^2 . \tag{117}$$

We recast equation 114 in terms of the function $\chi(r) = rR_0(r)$ and the dimensionless parameters

$$\rho \equiv kr, \tag{118}$$

$$\epsilon_v \equiv \frac{k/M}{\alpha} = \frac{v}{\alpha}, \tag{119}$$

$$\epsilon_\phi \equiv \frac{m_\phi/M}{\alpha}, \tag{120}$$

so we have

$$\frac{\mathrm{d}^2\chi}{\mathrm{d}\rho^2} + \left(1 + \frac{2}{\epsilon_v \rho} e^{-\epsilon_\phi \rho/\epsilon_v}\right) \chi(\rho) = 0. \tag{121}$$

This equation cannot be solved analytically in general. However, in the case of a massless force-mediator, an analytic solution exists. Taking $m_\phi = 0$, we have

$$\frac{\mathrm{d}^2\chi}{\mathrm{d}\rho^2} + \left(1 + \frac{2}{\epsilon_v \rho}\right) \chi(\rho) = 0. \tag{122}$$

This is the Coulomb wave equation[25] and is solved by the regular and irregular Coulomb wave functions, the first of which is physically permissible, so

$$\chi(\rho) = F_0(\eta, \rho) = \sqrt{\frac{2\pi/\epsilon_v}{1 - e^{-2\pi/\epsilon_v}}} \rho e^{i\rho} M(1 - i/\epsilon_v, 2, -2i\rho), \tag{123}$$

where $M(\alpha, \beta, \gamma)$ is the confluent hypergeometric function. Using $\chi(r) = rR(r)$ in equation 113, we find that the Sommerfeld factor in the case of a massless force-mediator is

$$S_C = \left| \frac{\chi(\rho)}{\rho} \right|^2_{\rho \to 0} = \frac{2\pi/\epsilon_v}{1 - e^{-2\pi/\epsilon_v}}. \tag{124}$$

[25]Equation 122 can be solved analytically even if the $l(l+1)/r^2$ term was not dropped in equation 114.

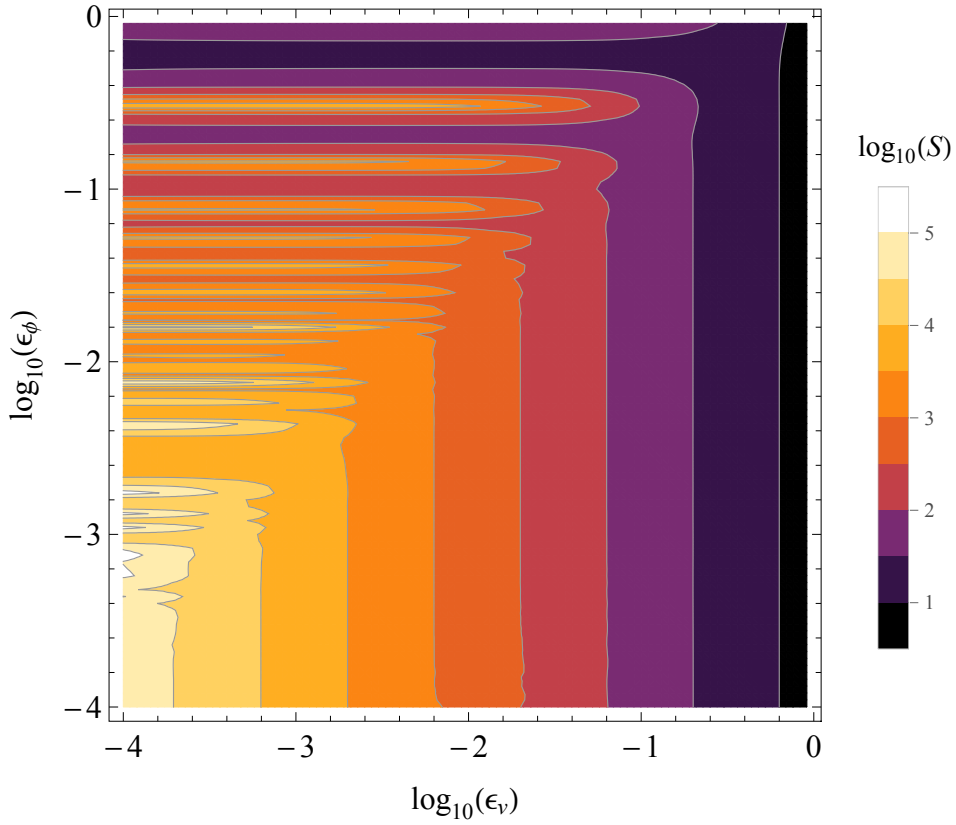

Figure 11: Sommerfeld factor for a Yuwaka potential.

Let us now return to the case of a general Yukawa potential, equation 121. Calculating this requires numerical computation. We know that in the limit $r \to 0$, $\chi(r) \sim r$, since $R_0(r) \sim$ constant. This grants an alternative expression for the Sommerfeld factor:

$$S = |\chi'(0)|^2 . \tag{125}$$

In the opposite limit, $r \to \infty$, we expect $\chi(r) = rR_0(r) \sim \sin(kr)$, a sinesoidal wave with amplitude unity. This is all the information necessary to calculate S. Details of the algorithm are given in Appendix C. The Sommerfeld factor S is plotted in Figure 11. Note the resonance pattern that arises due to the force-mediator being massive, which allows for the brief formation of bound states.

An alternative potential is the Hulthen potential

$$V_H(r) = -\alpha m_\phi \frac{e^{-m_\phi r}}{1 - e^{-m_\phi r}}, \tag{126}$$

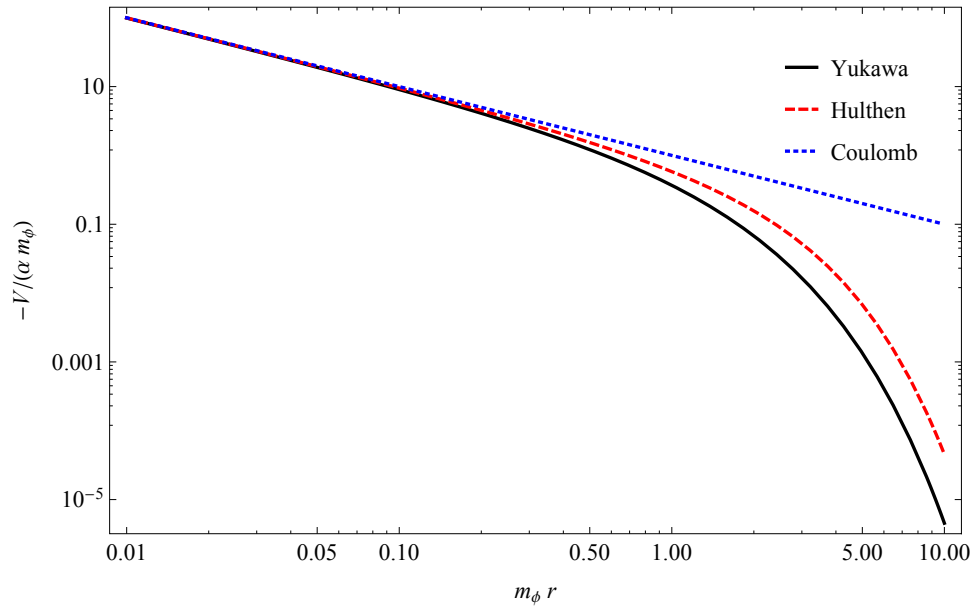

Figure 12: The three potential functions in units of αm_ϕ as a function of radius in units of $1/m_\phi$.

which behaves as a Coulomb potential $-\alpha/r$ for small r and, like the Yukawa potential, exponentially decays for large r. The three potentials so far discussed, Yukawa, Hulthen, and Coulomb, are compared in Figure 12.

The advantage of approximating the Yukawa potential with the Hulthen potential is that the Schrödinger equation is then analytically solvable. The result is

$$S_H(v) = \frac{A \sinh B}{\cosh B - \cos\sqrt{2AB - B^2}}, \tag{127}$$

where $A = 2\pi/\epsilon_v$ and $B = 12\epsilon_v/\pi\epsilon_\phi$. To demonstrate the validity of using this result, Figure 13 plots the Sommerfeld factors S and S_H as a function of the mediator-WIMP mass-ratio for two different velocities. Most important is that the structure from mass resonance is also exhibited when using the Hulthen potential. From equation 127 we see that mass resonances occur in the low velocity limit ($\epsilon_v \ll \epsilon_\phi$) at

$$\epsilon_{\phi,\text{res}} = \frac{12}{\pi^2 n^2} \quad \text{for} \quad n \in \mathbb{Z}_{>0}, \tag{128}$$

which are drawn in Figure 13

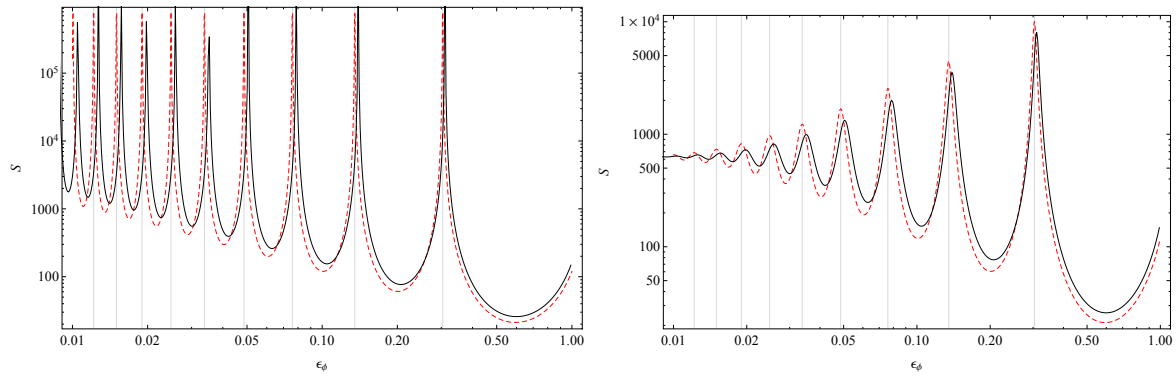

Figure 13: Comparison of the Sommerfeld factor S for a Yukawa potential (solid black) and a Hulthen potential (dashed red). Left: $\epsilon_v = 10^{-4}$. Right: $\epsilon_v = 10^{-2}$.

We will generally refer to $S = S(v)$ as the Sommerfeld factor, which is a function of the relative velocity between two annihilating particles. If we have a population of particles with some probability distribution of relative velocities $f\left(v; \{\alpha_i\}\right)$ that is specified by parameters $\{\alpha_i\}$ (usually the temperature or position in a system), then the *Sommerfeld enhancement* of the annihilation rate is

$$\begin{aligned} \mathcal{S}\left(\{\alpha_i\}\right) &\equiv \langle \sigma v \rangle / \langle \sigma v \rangle_0 \\ &= \frac{\int \mathrm{d}v\, \sigma(v) v\, S(v)\, f\left(v; \{\alpha_i\}\right)}{\int \mathrm{d}v\, \sigma(v) v\, f\left(v; \{\alpha_i\}\right)}. \end{aligned} \tag{129}$$

In the case of only s-wave annihilation, this simplifies to

$$\mathcal{S}\left(\{\alpha_i\}\right) = \int \mathrm{d}v\, S(v)\, f\left(v; \{\alpha_i\}\right). \tag{130}$$

6 Direct Detection

We have determined that the Milky Way galaxy is contained in a much larger dark matter halo. Dynamical observations of the galaxy give an idea of the local DM density $\rho_\odot \approx 0.3$ - $0.4\,\mathrm{GeV/cm^3}$ [42] [**Find some more references**]. It must therefore be that the Earth sweeps through a significant amount of DM particles as it, and the solar system, move through the galaxy. If the DM particles are WIMPs, then they will occasionally interact with other matter. In particular, we can in principle observe the interaction of passing WIMPs with nuclei: we can record the *recoil energy* imparted upon nuclei in a detector. This is the basic idea of direct detection [66, 135, 56].

The most important ingredients in predicting the rate of WIMP-nucleon interactions in a detector are the velocity distribution of local DM particles in the lab frame $f_{\mathrm{lab}}(\mathbf{v};t)$, normalized to $\rho_\odot$ here, and the WIMP-nucleon cross-section. The interaction rate between nucleons and WIMPs with velocity $\mathbf{v}$ per detector mass is [76]

$$\mathrm{d}R_{\mathbf{v}} = \frac{\rho_\odot}{m_\chi m_N} f_{\mathrm{lab}}\left(\mathbf{v};t\right) \frac{\mathrm{d}\sigma}{\mathrm{d}\left|\mathbf{q}\right|^2} \mathrm{d}\left|\mathbf{q}\right|^2 \left|\mathbf{v}\right| \mathrm{d}^3 v, \tag{131}$$

where $\left|\mathbf{q}\right|^2$ is the square momentum transfer and $\mathrm{d}\sigma/\,\mathrm{d}\left|\mathbf{q}\right|^2$ is the differential cross-section. Switching to the energy transfer $Q = \left|\mathbf{q}\right|^2/2m_N$ and integrating over velocity, we have the total energy-differential interaction rate per detector mass

$$\frac{\mathrm{d}R}{\mathrm{d}Q} = \frac{\rho_\odot}{m_\chi m_N} \int_{|\mathbf{v}|\geq v_{\mathrm{th}}} \mathrm{d}^3 v \ \left|\mathbf{v}\right| f_{\mathrm{lab}}\left(\mathbf{v};t\right) \frac{\mathrm{d}\sigma}{\mathrm{d}Q}. \tag{132}$$

We will focus on elastic scattering. Then the minimum or threshold velocity a WIMP needs in order to contribute to the differential rate with energy transfer Q is

$$v_{\mathrm{th}} = \sqrt{\frac{Qm_\chi}{2m_{\mathrm{r}}^2}}, \tag{133}$$

where $m_{\mathrm{r}} = m_\chi m_N/\left(m_\chi + m_N\right)$ is the reduced mass. Usually the focus is on spin-independent interactions, since the spin-dependent interaction rate is typically much

smaller (but note, *e.g.*, [10]). Then the differential cross-section is

$$\frac{\mathrm{d}\sigma}{\mathrm{d}Q} = \frac{\sigma_0 m_N}{2m_\mathrm{r}^2 \left|\mathbf{v}\right|^2} F^2\left(Q\right), \tag{134}$$

where σ_0 is the WIMP-nucleon cross-section at zero momentum transfer and $F\left(Q\right)$ is the form factor describing the structure of the nucleus used in the detector in question. Inserting equation 134 into equation 132, we have

$$\begin{aligned}\frac{\mathrm{d}R}{\mathrm{d}Q} &= \frac{\sigma_0}{2m_\chi m_\mathrm{r}^2} \int_{|\mathbf{v}|\geq v_\mathrm{th}} \mathrm{d}^3 v \, \frac{\rho_\odot f_\mathrm{lab}\left(\mathbf{v};t\right)}{|\mathbf{v}|} \\ &= \frac{\sigma_0}{2m_\chi m_\mathrm{r}^2} \int_{v_\mathrm{th}}^{\infty} \mathrm{d}v \, \mathrm{d}\Omega \, \rho_\odot \left|\mathbf{v}\right| f_\mathrm{lab}\left(\mathbf{v};t\right). \end{aligned} \tag{135}$$

All of the astrophysics of DM are contained in the integral of equation 135, while all of the particle physics is outside. From an astrophysical standpoint, therefore, the integrand of equation 135 is of the main interest. We denote this function of velocity as

$$g\left(\left|\mathbf{v}\right|,t\right) \equiv \int \mathrm{d}\Omega \, \rho_\odot \left|\mathbf{v}\right| f_\mathrm{lab}\left(\mathbf{v};t\right). \tag{136}$$

Note that we have written the lab-frame WIMP velocity distribution with explicit time-dependence. This is because the boost from the halo frame to the lab frame changes as the Earth orbits the Sun. Generally, it is the *combination* of the Earth's velocity with respect to the Sun and the Sun's velocity with respect to the galaxy and halo. Unless the local DM velocity distribution is anisotropic (see Section 8), a boost of constant magnitude has no effect and f_lab would indeed be independent of time. However, the Earth's orbit causes an oscillation in the net velocity boost. Thus, the interaction rate depends on the time of year. The velocity boost to the lab frame may be considered as the aggregate of three motions: the rotational motion of the Sun around the galactic center (defining the "local standard of rest"), the pecular motion of the sun, which is an instance of the relatively small random velocities in the Milky Way spiral galaxy, and the orbit of the Earth around the Sun. These combine to give the net velocity boost to the

lab frame, which is roughly

$$\mathbf{v}_{\rm lab} = \left[\Theta_0 + (15.4\,{\rm km/s})\cos\left(2\pi\left(t - t_1\right)\right)\right]\hat{\phi}, \tag{137}$$

where $\Theta_0 \simeq \sqrt{2/3}\sigma_v \approx 250\,{\rm km/s}$, and $t_1 = 0.419$ is the beginning of June in years.

There are various experimental setups for detecting WIMP-nucleon interactions. Experiments such as LUX [9], XENON [15], and ZEPLIN [81] consist of a large volume of a nobel gas, often xenon. When a WIMP interacts with a nucleus, imparting 1 to 100 keV of energy, the atom is both excited and ionized. Applying an electric field causes the free electrons to drift, eventually scintillating. These processes create two separate signals, and the time delay between the two distinguishes the incident particle. Alternatively, experiments such as CDMS [8], Edelweiss [16], CRESST [13], CoGeNT [1], and DAMA/LIBRA [23] use cryogenically-cooled crystal materials. Interactions with incident particles create scintillation photons and cause ionization, and vibrational phonons are also detectable.

Given an experimental setup, there are two basic ways of interpreting data. One can attempt to filter out background events from unwanted particles such as cosmic rays (experiments are typically located underground to protect against these), radiation from the surrounding earth, and contamination of the target material. Whatever is left, if significant, might be interpreted as a WIMP signal. Alternatively, one may neglect accounting for background radiation and look for a modulation in the total signal. If the rate of background interactions is constant, then any modulation in the total rate must be from the earth sweeping through the DM halo at varying speed, as described above. This has been the strategy of the CoGeNT and DAMA/LIBRA collaborations, which have reported a modulation, interpreted as a WIMP signature, for some time. Meanwhile, the CRESST collaboration also reports an excess of events compatible with a WIMP signal. On the other hand, collaborations such as XENON, Super-CDMS[26], ZEPLIN, Edelweiss, and, especially and most recently, LUX have all found no significant signal attributable to WIMPs. These collaborations put maximum limits on the possible WIMP mass and

[26]CDMS did find a possible signal, but this was later excluded by Super-CDMS.

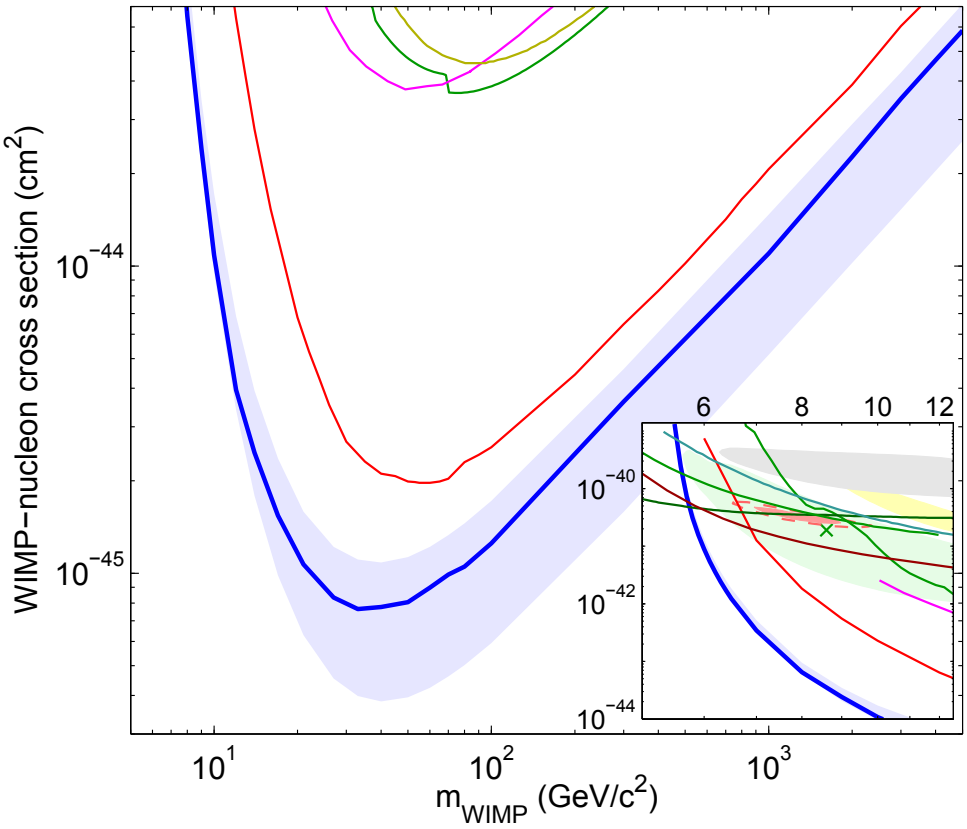

Figure 14: 90% confidence limits on the spin-independent elastic WIMP-nucleon cross-section and WIMP mass from LUX (blue and red), XENON-100 (orange), CDMS II (green), ZEPLIN-III (magenta), and Edelweiss II (yellow). Also, in the inset, the 90% allowed regions from CRESST (yellow), CDMS II silicon detectors (green), CoGeNT (red), and DAMA/LIBRA (grey). From [9].

WIMP-nucleon cross-section. Indeed, there is much disagreement between various direct detection experiments, including between positive results. Figure 14 shows the limits from experiments with null results and the allowed regions from experiments with positive results.

7 The Phase-Space Distribution and Implications for Detection

The phase-space distribution or *distribution function* (DF) contains the full information for describing a system of particles statistically. We define the function f so that $f(\mathbf{x}, \mathbf{v}, t)\,\mathrm{d}^3x\,\mathrm{d}^3v$ is the probability that at time t a *particular* particle is in the phase-space volume $\mathrm{d}^3x\,\mathrm{d}^3v$ around the position $(\mathbf{x}, \mathbf{v})$. Since any particle must be *somewhere*, this function is normalized so that

$$\int \mathrm{d}^3x\,\mathrm{d}^3v f(\mathbf{x}, \mathbf{v}, t) = 1. \tag{138}$$

The value of the DF at any location in phase-space is invariant with the coordinate system [29], so we may use arbitrary canonical phase-space coordinates $\mathbf{w} = (\mathbf{x}, \mathbf{v})$. At time passes, probability must be conserved in the same sense that mass is conserved in fluid flow. Thus we have the continuity equation

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial \mathbf{q}} \cdot (f\dot{\mathbf{q}}) + \frac{\partial}{\partial \mathbf{p}} (f\dot{\mathbf{p}}) = 0. \tag{139}$$

Using Hamilton's equations, this becomes

$$\frac{\partial f}{\partial t} + \dot{\mathbf{q}} \cdot \frac{\partial f}{\partial \mathbf{q}} + \dot{\mathbf{p}} \cdot \frac{\partial f}{\partial \mathbf{p}} = 0. \tag{140}$$

This is the *collisionless Boltzmann equation*. An alternative way to express it is to define the total derivative $\mathrm{d}f/\mathrm{d}t$ as the left-hand side of equation 140. It is the rate of change in the local probability density as seen by an observer following a particle in the system. We can then simply write

$$\frac{\mathrm{d}f}{\mathrm{d}t} = 0, \tag{141}$$

which means that the flow of "probability fluid" through phase-space is incompressible. In other words, the probability density around a particle is constant as that particle moves (though the density around *different* particles may be very different).

In Cartesian coordinates the Hamiltonian is $H = |\mathbf{v}|^2/2 + \Phi(\mathbf{x})$ and equation 140 becomes

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} - \frac{\partial \Phi}{\partial \mathbf{x}} \cdot \frac{\partial f}{\partial \mathbf{v}} = 0. \tag{142}$$

From here on we focus on systems that are in statistical equilibrium, so the DF is not a function of time: $f = f(\mathbf{x}, \mathbf{v})$. Jeans theorem tells us that we may always take the DF of a system to be a function of no more than three independent isolating integrals [29]. In a stationary gravitational potential $\Phi(\mathbf{x})$, an integral of motion is a function $I(\mathbf{x}, \mathbf{v})$ that is invariant along any orbit in the system. Mathematically, the function I satisfies equation 141. In this work we will discuss two integrals of motion: the Hamiltonian or energy H and the angular momentum L.

The Hamiltonian H is an integral of motion: the energy of a particle in a collisionless system is constant. A DF $f = f(H)$ that is a function of energy only is called *ergodic*: it is constant along energy hyper-surfaces in phase-space. It immediately follows that the mean velocity vanishes everywhere, because f is an odd function of $\mathbf{v}$:

$$\overline{\mathbf{v}}(\mathbf{x}) = \frac{\int \mathrm{d}^3 v \, \mathbf{v} f(|\mathbf{v}|^2/2 + \Phi)}{\int \mathrm{d}^3 v f(|\mathbf{v}|^2/2 + \Phi)} = 0. \tag{143}$$

Similarly, the velocity dispersion tensor is isotropic everywhere:

$$\sigma_{ij}^2(\mathbf{x}) = \overline{v_i v_j} = \frac{1}{3}\sigma^2 \delta_{ij}. \tag{144}$$

An isotropic system is thus one that is described by an ergodic DF. In the context of this work, the terms "isotropic" and "ergodic" may be used interchangably. Systems with spherical symmetry are the easiest to describe and study. Though DM halos in simulations are not spherical but rather axisymmetric, it is useful to model them as spherical for the purposes of detection prospects.

If the system is spherically symmetric, so $\Phi = \Phi(r)$, then the distribution of particles depends on the angular momentum vector $\mathbf{L}$ only through the magnitude $L = |\mathbf{L}|$. In spherical coordinates, v_r and $v_\mathrm{t} = \sqrt{v_\theta^2 + v_\phi^2}$ are the radial and tangential components of

$\mathbf{v}$; they are parallel and perpendicular, respectively, to the radial direction. Then $L = rv_\mathrm{t}$ and $H = (v_\mathrm{r}^2 + v_\mathrm{t}^2)/2 + \Phi(r)$. The mean values of v_r and v_t again vanish, but isotropy is broken: $\sigma_\mathrm{r}(r) \neq \sigma_\theta(r) = \sigma_\phi(r)$. Anisotropic systems will be studied in greater detail in Section 8.

It is worth pointing out that many of the usual concepts and techniques from statistical mechanics cannot be applied to systems such as those we are discussing here [29, 101, 86, 77]. Firstly, in a gravitational system, *energy is not an extensive quantity*: the energy of the whole is not equal to the sum of the energies of its parts. To illustrate, in a system with constant number density n of particles with mass m, the potential energy of any particular particle at radius R is $-Gm^2n\int \mathrm{d}^3x/r = -2\pi Gm^2nR^2$. The potential energy per particle is greater for more distant particles, so the total energy is not extensive. Non-extensive statistical mechanics, which generalizes the Boltzmann-Gibbs entropy, is discussed in [126, 128, 127]. Secondly, for bound systems of more than two particles, hyper-energy surfaces in phase-space are unbounded. This means that *the microcanonical ensemble cannot be defined*. Furthermore, since the energy, mass, and entropy of a gravitational system cannot be simultaneously finite, *the canonical ensemble cannot be defined* (non-extensive statistical mechanics may avoid this issue). Indeed, the phase-space *mass* densities calculated in Section 7.2 cannot be normalized: the mass is infinite. Calculating a phase-space *probability* density is an alternative, and it can be normalized to unity, but the problem resurfaces when calculating observables because the number of particles is infinite. However, since we are just interested in using phase-space distributions to empirically calculate observables and they are calculated from models that are measured anyway (not derived), we will continue despite these caveats, but they should be kept in mind.

7.1 The Jeans Equation and Applications

Recall the Boltzmann equation (142), now written using the Einstein summation convention:

$$\frac{\partial f}{\partial t} + v_i\frac{\partial f}{\partial x_i} - \frac{\partial \Phi}{\partial x_i}\frac{\partial f}{\partial v_i} = 0. \tag{145}$$

Multiplying this by v_j and integrating over all velocities gives

$$\frac{\partial}{\partial t}\int \mathrm{d}^3 v f v_j + \int \mathrm{d}^3 v v_i v_j \frac{\partial f}{\partial x_i} - \frac{\partial \Phi}{\partial x_i}\int \mathrm{d}^3 v v_j \frac{\partial f}{\partial v_i} = 0, \tag{146}$$

where we have used the facts that $\mathbf{v}$ does not depend on time and that $\partial\Phi/\partial x_i$ does not depend on $\mathbf{v}$. Since there are no particles with infinite velocity, we can apply the divergence theorem to the last term to find

$$\frac{\partial(\rho\overline{v_i})}{\partial t} + \frac{\partial(\rho\overline{v_i v_j})}{\partial x_j} + \rho\frac{\partial\Phi}{\partial x_i} = 0. \tag{147}$$

Subtracting from this $\overline{v_i}$ times the continuity equation, we finally obtain the *Jeans equation*:

$$\rho\frac{\partial\overline{v_i}}{\partial t} + \rho\overline{v_j}\frac{\partial\overline{v_i}}{\partial x_j} = -\rho\frac{\partial\Phi}{\partial x_i} - \frac{\partial(\rho\sigma_{ij}^2)}{\partial x_j} \tag{148}$$

This is essentially the Euler equation but with the fluid velocity replaced by the average particle velocity; the last term acts as a pressure force $-\nabla p$. Indeed, the object $-\rho\sigma_{ij}^2$ is a stress tensor that provides a pressure, which is anisotropic in general.

The Jeans equation in spherical coordinates is

$$\frac{\mathrm{d}(\rho\overline{v_\mathrm{r}^2})}{\mathrm{d}r} + 2\frac{\beta}{r}\rho\overline{v_\mathrm{r}^2} = -\rho\frac{\mathrm{d}\Phi}{\mathrm{d}r}, \tag{149}$$

where we have used the anisotropy parameter $\beta \equiv 1 - \sigma_t^2/2\sigma_\mathrm{r}^2$. This form will be vital in Section 8. If the anisotropy parameter is constant, we can multiply equation 149 by $r^{2\beta}$ and integrate to find

$$\overline{v_\mathrm{r}^2}(r) = \frac{1}{r^{2\beta}\rho(r)}\int_r^\infty \mathrm{d}r' r'^{2\beta}\rho(r')\frac{\mathrm{d}\Phi}{\mathrm{d}r'}, \tag{150}$$

which simplifies further for an isotropic system:

$$\overline{v_\mathrm{r}^2}(r) = \frac{1}{\rho(r)}\int_r^\infty \mathrm{d}r'\rho(r')\frac{\mathrm{d}\Phi}{\mathrm{d}r'}. \tag{151}$$

This formula has a simple physical interpretation: multiplying both sides by $\rho\Delta V$ shows that the left-hand side is the total kinetic energy of particles in a small volume ΔV at

radius r, while the right-hand side is the work done to move a mass $\rho(r)\Delta V$ from infinity to radius r.

Recall that the cross section for p-wave interactions is proportional to the second moment of the relative velocity, the relative velocity dispersion squared: $\sigma^2 \equiv \langle v^2 \rangle$. If *individual* particles follow a velocity distribution that is Maxwell-Boltzmann with dispersion σ_1, then the relative velocity dispersion is

$$\sigma = \sqrt{2}\sigma_1. \tag{152}$$

As mentioned in Section 5, the radial velocity dispersion of individual particles can be calculated directly from the density profile and potential using the Jeans equation. We do not need to know the functional form of the velocity distribution. The photon flux from p-wave interactions in an isotropic halo is, using equations 101, 103, and 152,

$$\Phi_{\gamma,p} = \frac{6bN_\gamma}{4\pi m_\chi^2} \int \mathrm{d}l \, \mathrm{d}\Omega \, \sigma_{\mathrm{r},1}^2[r(l,\psi)]\rho_\chi^2[r(l,\psi)], \tag{153}$$

and $\sigma_{\mathrm{r},1}^2$ is easily calculated from equation 151. Typically, because DM particles are very non-relativistic, the contribution from p-wave interactions is subdominant to velocity-independent s-wave interactions. However, it is possible there is some mechanism that prohibits s-wave annhilations.

If a system is spherical and isotropic, higher velocity moments (kurtosis, etc.) can be calculated from further equations analogous to the Jeans equation. This procedure can be continued to arbitrarily high velocity moments. These will be exact in the sense that they would agree with the calculations using the full phase-space distribution f. However, information is lost from integrating the Boltzmann equation. With just velocity moments, there is no way to ensure that f is everywhere non-negative.

We will see in Section 7.4 that the relative velocity distribution of DM particles is close to Maxwell-Boltzmann except near the center of halos, which is a small portion of the total volume. Therefore, unless an observation of the galactic center is in question, assuming a Maxwell-Boltzmann velocity distribution is a reasonable approximation. Since we can

easily calculate the single-particle velocity dispersion from the Jeans equation and thus also the relative velocity dispersion, we can calculate the thermal average of an arbitrary cross section. This strategy was used in considering Sommerfeld enhancement of an annihilation signal in the galaxy and dwarf spheroidals [106].

7.2 Eddington's Equation for Ergodic Systems

The phase-space density of a DM halo has often been taken as simply the product of the mass density (Section 3.1) and a Maxwell-Boltzmann velocity distribution:[27]

$$f = \rho(r) \times f_{\rm MB}(v). \tag{154}$$

This does not generally solve the Maxwell-Boltzmann equation (140). A Maxwell-Boltzmann velocity distribution is correct only for the singular isothermal sphere:

$$\rho_{SIS}(x) \propto x^{-2}. \tag{155}$$

If the velocity distribution cannot be approximated by Maxwell-Boltzmann (or any given function) and either the interaction rate per particle density $\langle \sigma v \rangle$ is not simply a linear combination of velocity moments or the Jeans equations cannot be closed, such as in the case of anisotropy, then the DF must be calculated. Given a spherical system with mass distribution $\rho(r)$ and potential $\Phi(r)$, we can find an ergodic DF that is self-consistent and a solution to the Boltzmann equation.[28]

It is convenient to introduce the **relative potential** $\Psi \equiv -\Phi + \Phi_0$ and **relative energy** $\mathcal{E} \equiv -H + \Phi_0 = \Psi - v^2/2$. The constant Φ_0 is chosen so that $f = 0$ for $\mathcal{E} \leq 0$. In other words, $\Psi = 0$ at the "edge" of the system. If the system extends to infinity, then $\Phi_0 = 0$ and Ψ is simply the binding energy of the system. In this work, Φ_0 will generally be zero, but there are cases when a non-zero value is appropriate.[29] Requiring

[27] We are now considering the DF as a phase-space distribution of *mass*, not probability. This does not change the points made earlier in Section 7.

[28] Any isolated, finite system with an ergodic DF must be spherical [29], but the converse is not necessarily true. A spherical system may have a non-ergodic DF (see Section 8).

[29] For example, if the halo profile is truncated at some finite radius $r_{\rm trunc}$, with no particles beyond

that $f(\mathcal{E} \leq 0) = 0$ also means that only particles that are bound are considered apart of the system, so the maximum velocity is the escape velocity $v_{\rm esc}(r) = \sqrt{2\Psi(r)}$.

To restate the problem precisely, we have a system with known mass distribution $\rho(r)$, and we wish to derive a DF that depends on the phase-space coordinates only through the energy: $f = f(\mathcal{E})$. We begin by noting that the mass distribution is recovered from the DF by marginalizing over velocity. Since the DF is isotropic, it depends only on the magnitude of the velocity. We can thus write

$$\rho(r) = 4\pi \int_0^{v_{\rm esc}} \mathrm{d}v\, v^2 f(\Psi - v^2/2). \tag{156}$$

We can change variables of integration to $\mathcal{E}$, and, since $\Psi(r)$ is a monotonic function, we can consider ρ as a function of Ψ instead of r. We have

$$\frac{1}{\sqrt{8}\pi}\rho(\Psi) = 2\int_0^{\Psi} \mathrm{d}\mathcal{E} f(\mathcal{E})\sqrt{\Psi - \mathcal{E}}. \tag{157}$$

The integrand vanishes at the limits of integration, so differentiating both sides with respect to Ψ gives

$$\frac{1}{\sqrt{8}\pi}\frac{\mathrm{d}\rho}{\mathrm{d}\Psi} = \int_0^{\Psi} \mathrm{d}\mathcal{E}\frac{f(\mathcal{E})}{\sqrt{\Psi - \mathcal{E}}}. \tag{158}$$

This is an Abel integral equation, which can be inverted. The solution for $f(\mathcal{E})$ may be written as

$$f(\mathcal{E}) = \frac{1}{\sqrt{8}\pi}\frac{\mathrm{d}}{\mathrm{d}\mathcal{E}}\int_0^{\mathcal{E}} \frac{\mathrm{d}\Psi}{\sqrt{\mathcal{E} - \Psi}}\frac{\mathrm{d}\rho}{\mathrm{d}\Psi}. \tag{159}$$

A physical requirement is that $f(\mathcal{E}) \geq 0$, there is no negative mass density anywhere. Equation 159 thus implies that a spherical mass distribution $\rho(r)$ and relative potential $\Psi(r)$ can be recovered from an ergodic DF if and only if the quantity

$$\int_0^{\mathcal{E}} \frac{\mathrm{d}\Psi}{\sqrt{\mathcal{E} - \Psi}}\frac{\mathrm{d}\rho}{\mathrm{d}\Psi} \tag{160}$$

is an increasing function of $\mathcal{E}$ for $\mathcal{E} \geq 0$. Otherwise, there would be some position and

this radius, then $\Phi_0 = \Phi(r_{\rm trunc})$.

velocity that has negative density. This is an important check when deriving DFs. A more useful way to express the solution to equation 158 is

$$f(\mathcal{E}) = \frac{1}{\sqrt{8}\pi}\left[\int_0^{\mathcal{E}} \frac{\mathrm{d}\Psi}{\sqrt{\mathcal{E}-\Psi}}\frac{\mathrm{d}^2\rho}{\mathrm{d}\Psi^2} + \frac{1}{\sqrt{\mathcal{E}}}\left(\frac{\mathrm{d}\rho}{\mathrm{d}\Psi}\right)_{\Psi=0}\right]. \tag{161}$$

This is *Eddington's formula* and will be used extensively in this work. Typically, the second term on the right-hand side vanishes, so for our purposes we have

$$f(\mathcal{E}) = \frac{1}{\sqrt{8}\pi}\int_0^{\mathcal{E}} \frac{\mathrm{d}\Psi}{\sqrt{\mathcal{E}-\Psi}}\frac{\mathrm{d}^2\rho}{\mathrm{d}\Psi^2}. \tag{162}$$

It is not always easy or possible to invert $\Psi(r)$ in order to express ρ as a function of Ψ.[30] However, we can always write the second derivative in equation 162 in terms of derivatives with respect to r. We have

$$\frac{\mathrm{d}^2\rho}{\mathrm{d}\Psi^2} = \left(\frac{\mathrm{d}\Psi}{\mathrm{d}r}\right)^{-2}\left(\frac{\mathrm{d}^2\rho}{\mathrm{d}r^2} - \left(\frac{\mathrm{d}\Psi}{\mathrm{d}r}\right)^{-1}\frac{\mathrm{d}^2\Psi}{\mathrm{d}r^2}\frac{\mathrm{d}\rho}{\mathrm{d}r}\right). \tag{163}$$

Then Eddington's formula becomes

$$f(\mathcal{E}) = \frac{1}{\sqrt{8}\pi}\int_0^{\mathcal{E}} \frac{\mathrm{d}\Psi}{\sqrt{\mathcal{E}-\Psi}}\left(\frac{\mathrm{d}\Psi}{\mathrm{d}r}\right)^{-2}\left(\frac{\mathrm{d}^2\rho}{\mathrm{d}r^2} - \left(\frac{\mathrm{d}\Psi}{\mathrm{d}r}\right)^{-1}\frac{\mathrm{d}^2\Psi}{\mathrm{d}r^2}\frac{\mathrm{d}\rho}{\mathrm{d}r}\right). \tag{164}$$

Or, changing the variable of integration,

$$f(\mathcal{E}) = \frac{1}{\sqrt{8}\pi}\int_{r(\Psi=\mathcal{E})}^{\infty} \frac{\mathrm{d}r}{\sqrt{\mathcal{E}-\Psi}}\left(\frac{\mathrm{d}\Psi}{\mathrm{d}r}\right)^{-1}\left(\frac{\mathrm{d}^2\rho}{\mathrm{d}r^2} - \left(\frac{\mathrm{d}\Psi}{\mathrm{d}r}\right)^{-1}\frac{\mathrm{d}^2\Psi}{\mathrm{d}r^2}\frac{\mathrm{d}\rho}{\mathrm{d}r}\right). \tag{165}$$

These forms are useful for numerical computation. We find that integrating over the potential Ψ results in much better accuracy than integrating over radius.

Consider the Milky Way halo. Using the baryonic disk and bulge models from Section 3.2, equations 72 and 73, the total gravitational potential experienced by dark matter

[30] For example, inverting the Jaffe and Hernquist potentials is easy. It is difficult, though doable, in the case of NFW (see Appendix H).

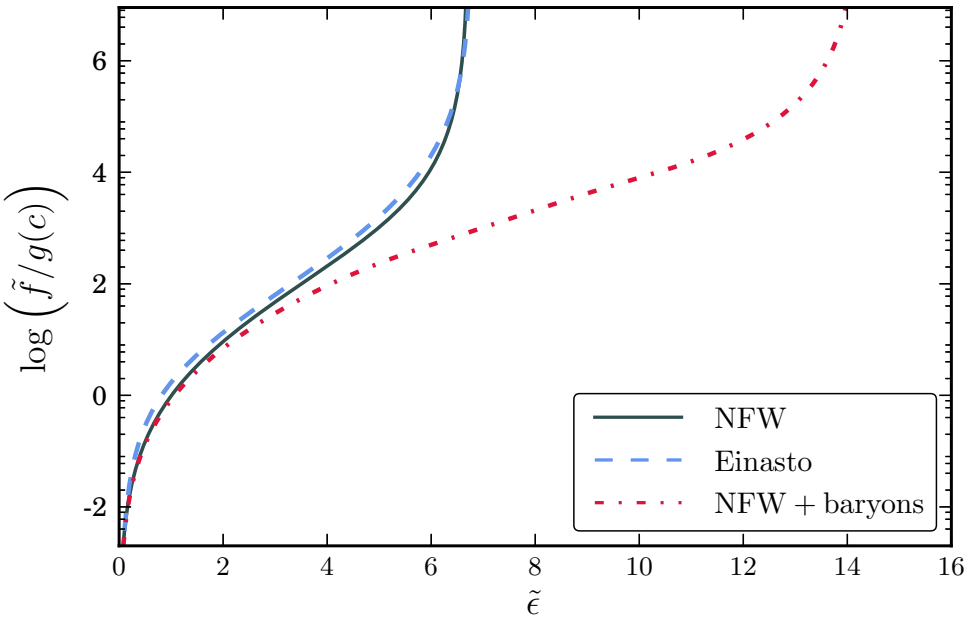

Figure 15: Phase-space distribution function of the dark matter in the Galaxy, equation 162, assuming an NFW profile, with and without baryonic components, or a dark matter only Einasto profile. See Appendix D for the definitions of the axes.

particles in the halo is

$$\Phi(r) = \Phi_{\rm bulge}(r) + \Phi_{\rm disk}(r) + \Phi_{\rm halo}(r). \tag{166}$$

We use both a NFW profile and an Einasto profile for the Milky Way halo, so $\Phi_{\rm halo}$ is derived from either equation 68 or 69. We can easily take derivatives of $\Psi = -\Phi$ and $\rho = \rho_\chi$; then equation 164 can be numerically computed. Figure 15 plots the DFs of a Milky Way-like halo for the cases of NFW with and without baryons and Einasto. The DFs vanish at zero energy because there are no bound particles infinitely far away; they diverge at a finite energy because at $r = 0$ the potential is finite and the density is infinite[31].

We do the same for a typical dwarf spheroidal, modeled after Draco. In this situation we neglect baryons and use a NFW profile and a cored Burkert profile. The DFs for these cases are plotted in Figure 16.

[31] A Jaffe profile, on the other hand, has infinite potential energy at $r = 0$, so the DF has a semi-infinite domain and diverges at infinity.

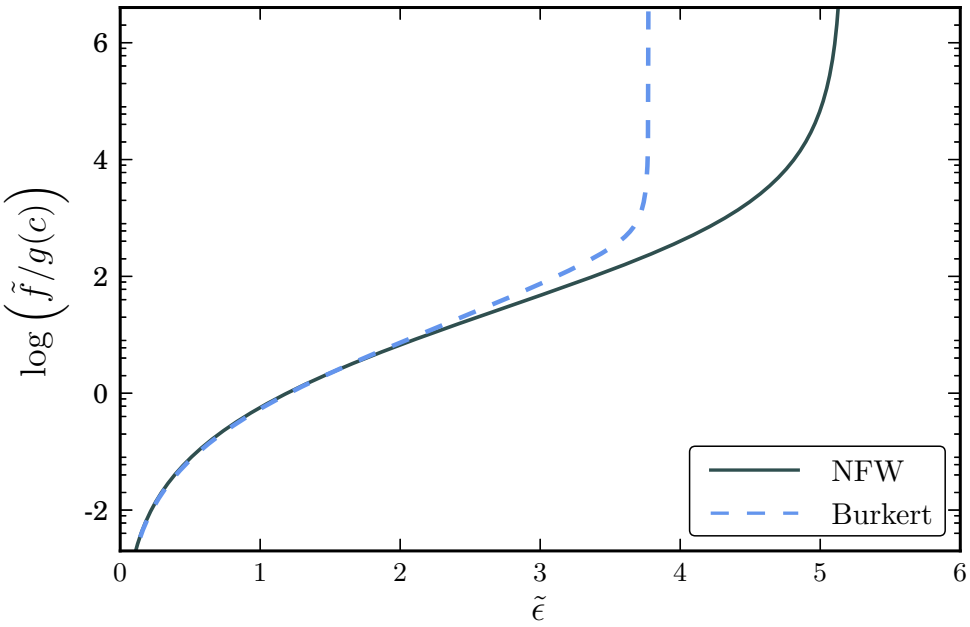

Figure 16: Phase-space distribution function of the dark matter in a dSph with an NFW or a cored Burkert profile. See Appendix D for the definitions of the axes.

7.3 Velocity Distribution

If the phase-space density $f(\mathbf{x}, \mathbf{v})$ is known, then the single-particle velocity distribution at a given position $\mathbf{x}$ is simply

$$P_1(\mathbf{v}; \mathbf{x}) = \frac{f(\mathbf{x}, \mathbf{v})}{\rho(\mathbf{x})}, \tag{167}$$

which is normalized so that $\int \mathrm{d}^3 v P_1(\mathbf{v}; \mathbf{x}) = 1$. If the system is spherical and ergodic, we have

$$P_1(v; r) = \frac{f(\Psi(r) - v^2/2)}{\rho(r)}. \tag{168}$$

The single-particle velocity distribution in Milky Way-like NFW and Einasto halos at a few radii are plotted in Figure 17. These do not include the effects of baryonic components. Figure 18 plots the NFW velocity distribution taking the baryonic disk and bulge into account. For a typical dwarf spheroidal, the cases of NFW and Burkert are plotted in Figure 19. For every curve in these plots, a Maxwell-Boltzmann distribution with the same dispersion is shown for comparison. In the DM-only plots, it is clear that the Maxwell-Boltzmann distribution is a good approximation at larger radii. We also see that the true velocity distribution deviates from Maxwell-Boltzmann at smaller radii: the

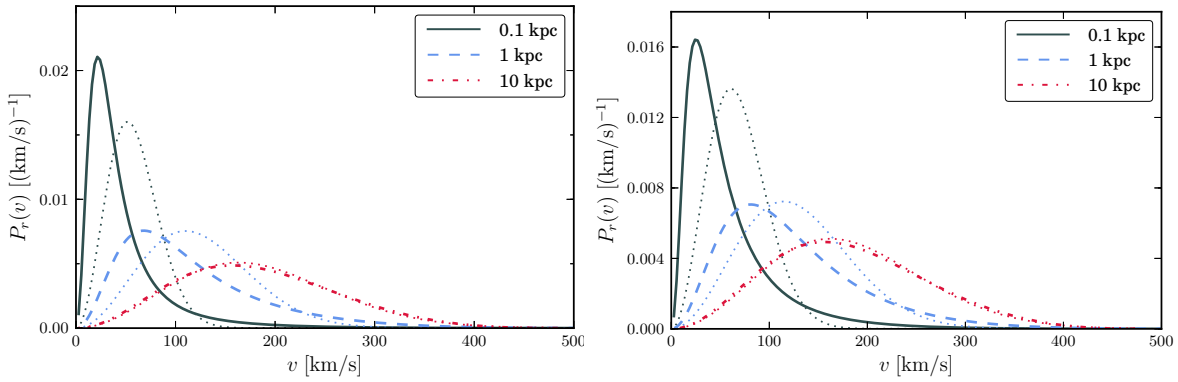

Figure 17: Single-particle velocity distribution $P_1(v;r)$ (see equation 168) for a NFW halo (left) and Einasto halo (right), both *without* baryonic components. The dotted lines show a Maxwell-Boltzmann distribution with the same velocity dispersion.

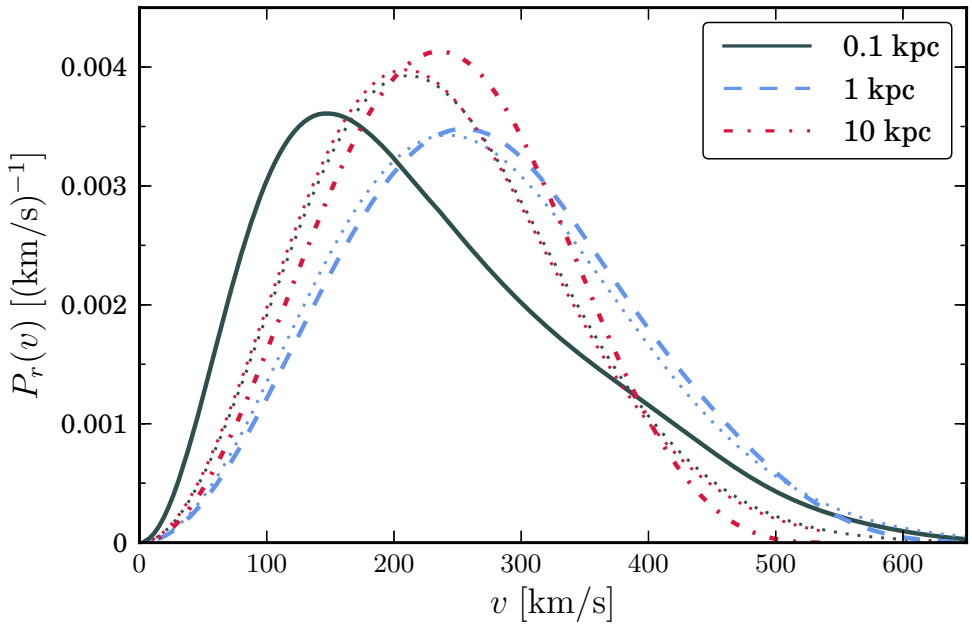

Figure 18: Single-particle velocity distribution $P_1(v;r)$ (see equation 168) for a NFW halo *with* baryonic components.

peak becomes sharper and located at a smaller velocity. The situation is less clear when the baryonic disk and bulge are included, but the deviation from Maxwell-Boltzmann still occurs at smaller radii, albeit less severely.

To calculate the annihilation rate in general, we need the *relative* velocity distribution $P(v_{\rm rel};r)$. Essentially, we need to convolve the single-particle velocity distribution in equation 167 with itself. In the center-of-momentum frame of a two-particle system, there are two relevant quantities: the relative velocity $\mathbf{v}_{\rm rel} = \mathbf{v}_1 - \mathbf{v}_2$ and the center-of-momentum velocity $\mathbf{v}_{\rm cm} = (\mathbf{v}_1 + \mathbf{v}_2)/2$. The two-particle and single-particle velocity

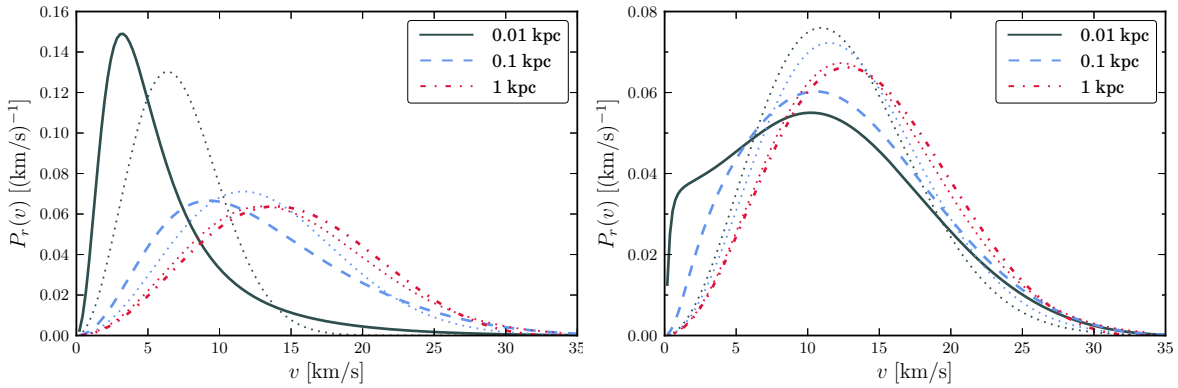

Figure 19: Single-particle velocity distribution for a dSph satellite-sized NFW halo (left) and Burkert halo (right).

distributions are related in general as (suppressing the distribution "parameter" $\mathbf{x}$)

$$P(\mathbf{v}_{\rm rel}, \mathbf{v}_{\rm cm})\, \mathrm{d}^3 v_{\rm rel}\, \mathrm{d}^3 v_{\rm cm} = P_1(\mathbf{v}_1) P_1(\mathbf{v}_2)\, \mathrm{d}^3 v_1\, \mathrm{d}^3 v_2. \tag{169}$$

The determinant of the Jacobian matrix is unity, so we have

$$P(\mathbf{v}_{\rm rel}, \mathbf{v}_{\rm cm}) = P_1(\mathbf{v}_{\rm rel}/2 + \mathbf{v}_{\rm cm}) P_1(\mathbf{v}_{\rm rel}/2 - \mathbf{v}_{\rm cm}). \tag{170}$$

Marginalizing over the center-of-momentum velocity and the direction of the relative velocity gives the relative velocity distribution. For an ergodic system, with P_1 given by equation 168, we have

$$P(v_{\rm rel}) = 8\pi^2 v_{\rm rel}^2 \int_0^{v_{\rm esc}} \mathrm{d}v_{\rm cm} v_{\rm cm}^2 \int_{-1}^{1} \mathrm{d}z\, P_1(v_+) P_1(v_-), \tag{171}$$

where $v_\pm^2 \equiv v_{\rm cm}^2 + v_{\rm rel}^2/4 \pm v_{\rm cm} v_{\rm rel} z$ and z is the cosine of the angle between the relative and center-of-momentum velocity vectors. This function is normalized so that

$$\int_0^{2v_{\rm esc}} \mathrm{d}v_{\rm rel}\, P(v_{\rm rel}) = 1. \tag{172}$$

We are now free to calculate the interaction rate per particle density with arbitrary cross section:

$$\langle \sigma v \rangle = \int_0^{2v_{\rm esc}} \mathrm{d}v_{\rm rel}\, \sigma v_{\rm rel}\, P(v_{\rm rel}). \tag{173}$$

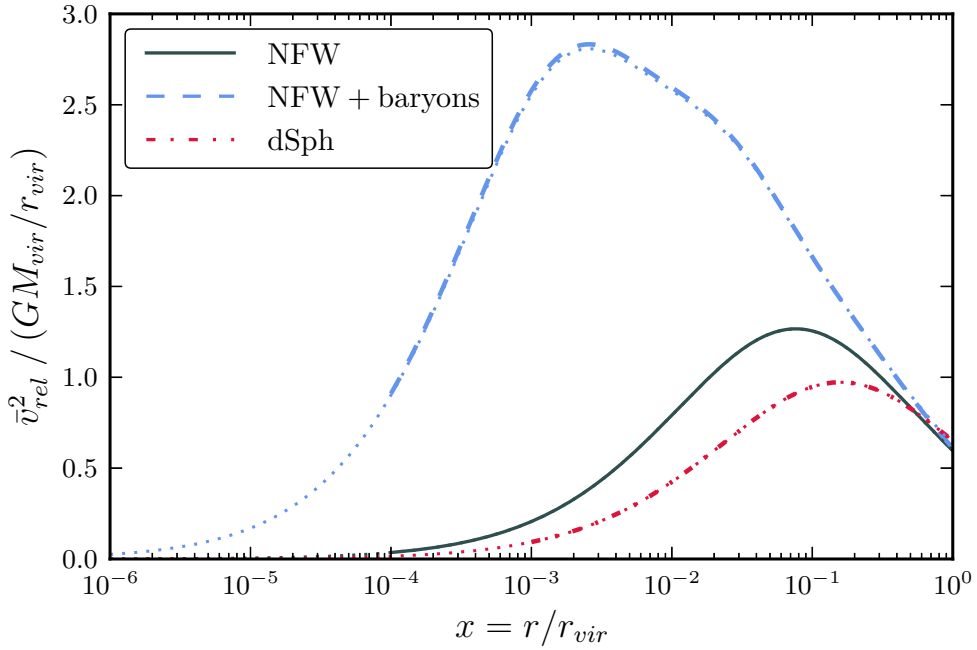

Figure 20: One dimensional relative velocity dispersion. The thick lines show the result of using the exact velocity distribution, while the thinner dotted lines use a Jeans analysis and assume that the single-particle distribution is Maxwell-Boltzmann. Note that in physical units, the line corresponding to the dSph would be lowered by a factor of $M_{\rm vir}^{\rm dSph} r_{\rm vir}^{\rm gal} / M_{\rm vir}^{\rm gal} r_{\rm vir}^{\rm dSph} \sim 1.5 \times 10^3$.

It is interesting to note that the relative velocity *dispersion* seems to follow the same relationship with the single-particle dispersion that is true for Maxwell-Boltzmann distributions. This is shown in Figure 20. This serves as a useful check of numerical computation.

Figure 21 plots the relative velocity distribution $P(v_{\rm rel})$ at various radii in a NFW halo, without and with baryonic components. The trends in the relative velocity distribution are very similar to those seen in the single-particle velocity distribution. The distribution becomes more sharply peaked at smaller values as the radius decreases.

7.4 Galactic Signature

We have developed an elaborate and complicated formalism in Sections 7.2 and 7.3 that is necessary when the Maxwell-Boltzmann distribution is a poor approximation *and* the non-relativistic expansion in equation 101 is invalid (or at least not the whole story). We will now scrutinize the necessity of such a detailed calculation. The question is: in what situations does using a MB distribution instead of the correct phase-space dis-

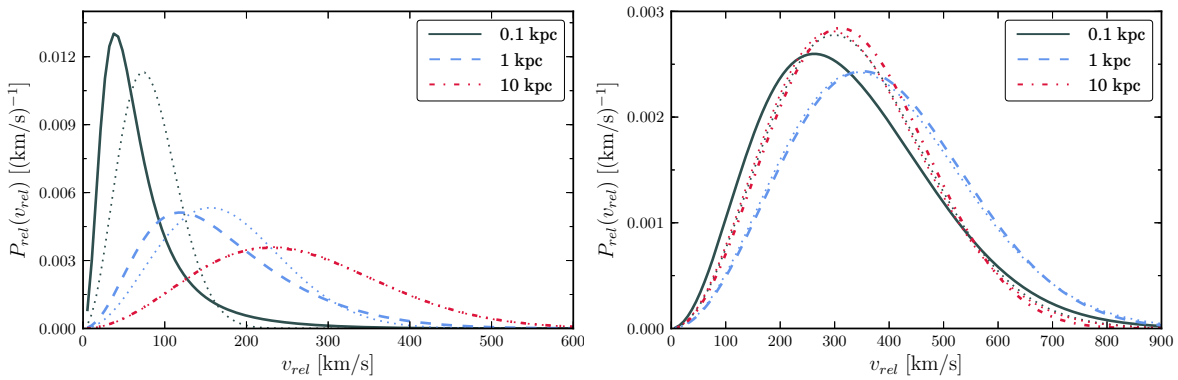

Figure 21: Relative velocity distribution at various radii in an NFW halo without (left) and with (right) the effects of baryonic components. The dotted lines show a Maxwell-Boltzmann distribution with the same velocity dispersion.

tribution result in a significantly-inaccurate calculation of observables? We assume an ergodic (spherically symmetric and isotropic) distribution function for simplicity. The aim is not so much to make predictions, but rather to test the reliability of simplifying assumptions [60].

The Sommerfeld effect motivated in Section 5.1 and described in Section 5.2 is an interesting example of $\langle \sigma v \rangle$ not simply being a linear combination of velocity moments. Instead, we have a complicated function of $v_{\rm rel}$ that must be averaged using the relative velocity distribution. Of course, the Maxwell-Boltzmann distribution can be used as an approximation, in which case calculating $\mathcal{S}(r) = \langle S(v_{\rm rel}; r) \rangle$ is a straightforward numerical calculation [106]. But we have seen that MB can be a poor approximation near the center of a halo. Is this break-down of the MB approximation very important in the context of indirect detection of dark matter? We define the quantity

$$\mathcal{F} \equiv J / J_{\rm MB}, \tag{174}$$

where J is the generalized J-factor (compare with equation 107)

$$J(\Delta\Omega) = \int \mathrm{d}l \, \mathrm{d}\Omega \, \mathcal{S} \rho_\chi^2, \tag{175}$$

which is calculated using the exact relative velocity distribution, and $J_{\rm MB}$ is the same calculation but using a Maxwell-Boltzmann distribution with dispersion 220 km/s. This

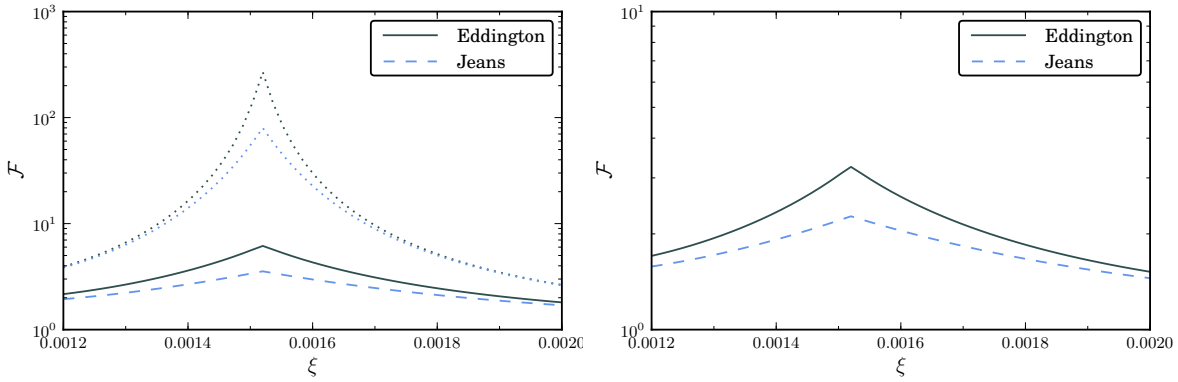

Figure 22: $\mathcal{F}$ calculated for the galactic center using a NFW profile (left) with (solid) and without (dotted) baryonic components and an Einasto profile (right) with baryonic components.

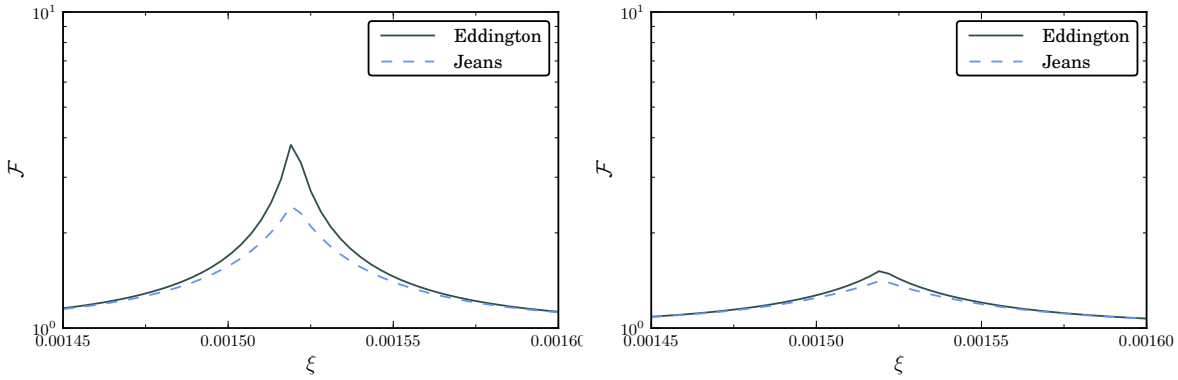

Figure 23: $\mathcal{F}$ calculated for the Draco dwarf spheroidal using a NFW profile (left) and a Burkert profile (right).

ratio quantizes the effect deviation from MB has on the observed flux. The larger the value, the more severe the deviation and a poorer the approximation. Figure 22 plots this for the case of the galactic center, using a Sommerfeld model with $\alpha = 10^{-2}$ and variable $\xi = m_\phi/m_\chi$. Figure 23 is for the case of Draco.

We are also interested in the possibility of observing secondary emission from electrons and positrons that are annihilation products interacting with the magnetic field around the galactic center. Instead of a line-of-sight integral, *i.e.* the J-factor, we calculate the volume integral

$$K = \int_0^{r_B} \mathrm{d}r' r'^2 \mathcal{S} \rho_\chi^2, \tag{176}$$

where $r_B \simeq 1\,\mathrm{kpc}$ is the radius within which the magnetic field strength can create appreciable emission. Analogous to equation 174, we quantize the adequacy of the Maxwell-

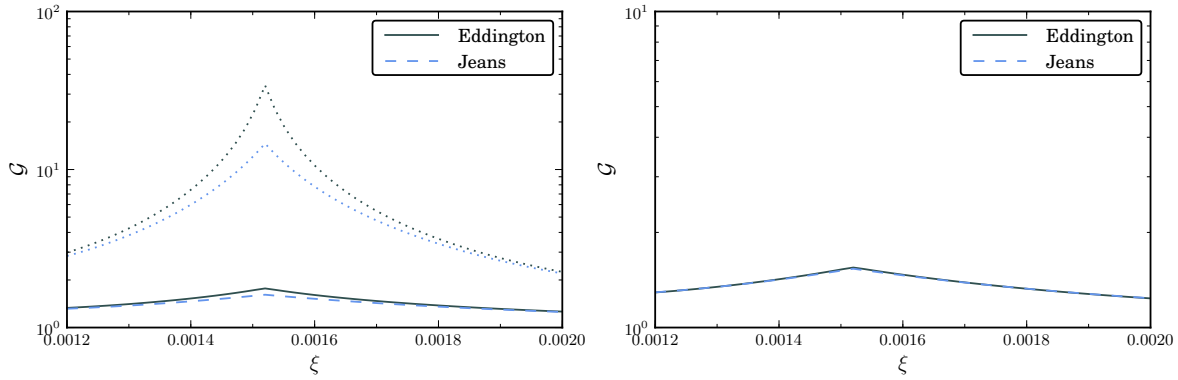

Figure 24: $\mathcal{G}$ calculated using a NFW profile (right) with (solid) and without (dotted) baryonic components and an Einasto profile (right) with baryonic components.

Boltzmann approximation with

$$\mathcal{G} \equiv K/K_{\mathrm{MB}}. \tag{177}$$

As before, we plot $\mathcal{G}$ in Figure 24 for a NFW profile and an Einasto profile, using the same Sommerfeld model with $\alpha = 10^{-2}$.

The most dramatic effect from using the DF is in the case of the galactic center without accounting for baryons. Additional boosts to the J-factor of as much as two orders of magnitude are possible. This would mean that any constaint on a combination of DM halo and Sommerfeld models using observations should actually be much more severe than previously thought. Including baryons, however, which gravitationally dominate at the galactic center, lessens this additional boost from using the DF. Still, we see that an additional boost factor of about 2 to 6 is possible. This does not depend much on the halo model used (NFW or Einasto). The reason for this additional boost over the Maxwell-Boltzmann approximation is that the true (relative) velocity distribution peaks more sharply and at a smaller velocity than the Maxwell-Boltzmann distribution (with the same dispersion) at small radii. Thus, the Sommerfeld enhancement at these radii are greater and contribute more to the overall boost.

On the other hand, other observations that get large contributions far from a halo's center are not as sensitive to the deviation from Maxwell-Boltzmann. This is because the deviation from MB occurs only at small radii, which contitutes a small portion of the entire halo. So the larger the portion of a halo away from the center contributing

to an observation, the smaller the importance of deviation from MB in predicting that observation. For example, compare the dotted Eddington curves in the left plots in Figures 22 and 24: the dotted Eddington curve in Figure 24 is smaller relative to the respective solid curve because it comes from a volume integral as opposed to a line-of-sight integral. Including baryons also reduces the Eddington curve to nearly unity. As for the case of Draco, even though the line-of-sight volume includes a large portion of the halo, since baryons are much less dominant in these systems, a cusped halo grants significant additional boosts. However, a cored profile, which seems more realistic for these systems, has very little deviation from MB, so the boost is small.

In all these plots we have also shown the boost over the SHM from using a MB distribution with variable velocity dispersion (compare with [106]). Compared to the calculation of the DF and relative velocity distribution, computing the velocity dispersion from the Jeans equation is easy. We conclude this study by suggesting that the variability of the velocity dispersion (and thus also velocity-dependent cross-sections) should always be taken into account. However, except in extreme cases of observations of the galactic center *and* Sommerfeld resonance, the full DF calculation can be neglected without too much inaccuracy.

7.5 Apparent Inner Density Slope

In Section 7.4 we showed how the *total* flux from Sommerfeld-enhanced WIMP annihilation near the galactic center can be subject to large boosts. Here we investigate how the "apparent slope" of the inner halo profile can be steepened from Sommerfeld enhancement. The enhancement increases with decreasing relative velocity, and the relative velocity distribution becomes peaked at smaller values closer to the center of the halo. Thus, we expect that the integrand in equation 103, $\langle\sigma v\rangle\rho_\chi^2 \propto \mathcal{S}(r)\rho_\chi^2$, would have a steeper slope than the square-density ρ_χ^2 alone. Motivation for this study comes from tentative observations of γ-ray excesses near the galactic center (the latest example being [51]). These flux measurements are typically fit to a power-law, and it is simplest to assume that this slope is that of the WIMP density profile. If there is Sommerfeld

enhancement, however, then there is some degeneracy between the *actual* density slope and the particle physics behind the enhancement (*i.e.* the values of the coupling α and mass ratio $\xi = m_\phi/m_\chi$).

In order to get a grasp of this ambiguity, we consider the generalized NFW profile (with $x \equiv r/r_\mathrm{s}$)

$$g_\chi(x) = x^{-\gamma}\left(1+x\right)^{-(3-\gamma)}, \tag{178}$$

and we call γ the "inner slope", which we vary from 0.8 to 1.7. For a given value of γ, we numerically compute the dimensionless potential

$$\tilde{\psi}\left(x\right) = x^{-1}\int_0^x \mathrm{d}x' x'^2 g_\chi + \int_x^\infty \mathrm{d}x' x' g_\chi. \tag{179}$$

We then compute the relative velocity distribution by way of equation 164 and 171 at about ten different radii from $x = 10^{-1}$ to 10^{-3}. Specifying the particle physics parameters, we can use this to calculate the Sommerfeld enhancement at those radii. At these radii, $\tilde{\rho}_\chi$ is essentially a simple power-law with slope γ. We fit a power-law with slope γ_app to $\sqrt{\mathcal{S}}\tilde{\rho}_\chi$ in this range of radii. To restore units to the potential, we use $\psi/\tilde{\psi} = 4\pi G r_\mathrm{s}^2\rho_\mathrm{s}$; the "velocity unit" is thus $v/\tilde{v} = \sqrt{4\pi G r_\mathrm{s}^2\rho_\mathrm{s}}$. Referring to Section 5.2, we define $\tilde{\alpha}$ and $\tilde{\xi}$ such that

$$\epsilon_v \equiv \frac{v/c}{\alpha} = \frac{\tilde{v}}{\tilde{\alpha}}, \tag{180}$$

$$\epsilon_\phi \equiv \frac{m_\phi/m_\chi}{\alpha} = \frac{\tilde{\xi}}{\tilde{\alpha}}. \tag{181}$$

In other words,

$$\tilde{\alpha}/\alpha = c/\sqrt{4\pi G r_\mathrm{s}^2\rho_\mathrm{s}} \tag{182}$$

$$\approx 645\frac{r_\mathrm{s}}{20\,\mathrm{kpc}}\left(\frac{\rho_\mathrm{s}}{10^7\,\mathcal{M}_\odot/\mathrm{kpc}^3}\right)^{1/2}. \tag{183}$$

We use the parameters $\tilde{\alpha}$ and $\tilde{\xi}$ because we then do not need to specify the halo parameters r_s and ρ_s when performing the computation.

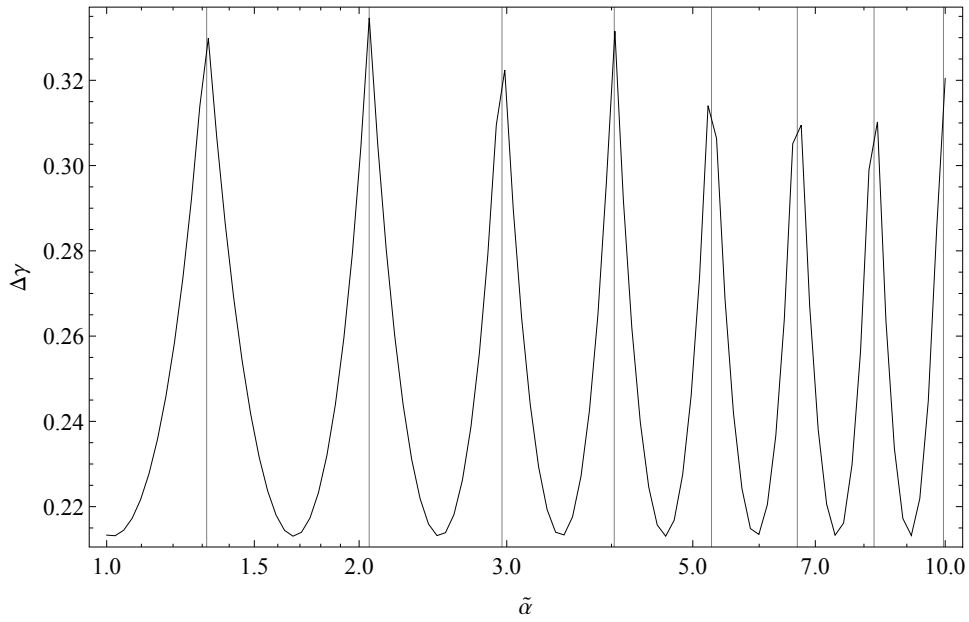

Figure 25: The change in inner slope $\Delta\gamma$ of $\sqrt{\mathcal{S}}\rho_\chi$ from ρ_χ as a function of $\tilde{\alpha}$, with $\gamma = 1$ and $\tilde{\xi} = 0.1$.

To calculate the Sommerfeld enhancement $\mathcal{S}(x)$, we use the Hulthen Sommerfeld factor $S(v_{\rm rel})$ given by equation 127. This is because the mass resonances are known to be at $\epsilon_\phi = 12/\pi^2 n^2$ or, rewritten in a more suggestive form, with the parameters used here,

$$\left.\sqrt{12\tilde{\alpha}/\tilde{\xi}}\right|_{\rm res} = \pi n \ \text{ for } \ n \in \mathbb{Z}_{>0}. \tag{184}$$

We proceed as follows. For some chosen values of γ and $\tilde{\xi}$ we find $\gamma_{\rm app}$ as a function of $\tilde{\alpha}$ over a range of $\tilde{\alpha}$ corresponding to $\epsilon_\phi \in [10^{-2}, 10^{-1}]$. An example of this is shown in Figure 25, in which we show the change in apparent slope $\Delta\gamma = \gamma_{\rm app} - \gamma$. Note that the resonance pattern from the Sommerfeld factor $S(v_{\rm rel})$ is manifest. Considering the magnitude and range of $\Delta\gamma$, this resonance pattern motivates the semi-empirical model

$$\Delta\gamma = A \log\left[12\csc^2\left(\sqrt{12\tilde{\alpha}/\tilde{\xi}}\right)\right]^p. \tag{185}$$

We can fit this model, parameterized by A and p, as a function of $\tilde{\alpha}$ to the numerical results. While this model can be a good fit to results such as that shown in Figure 25 over most of the range of $\tilde{\alpha}$, it has trouble near the resonances. The Sommerfeld models

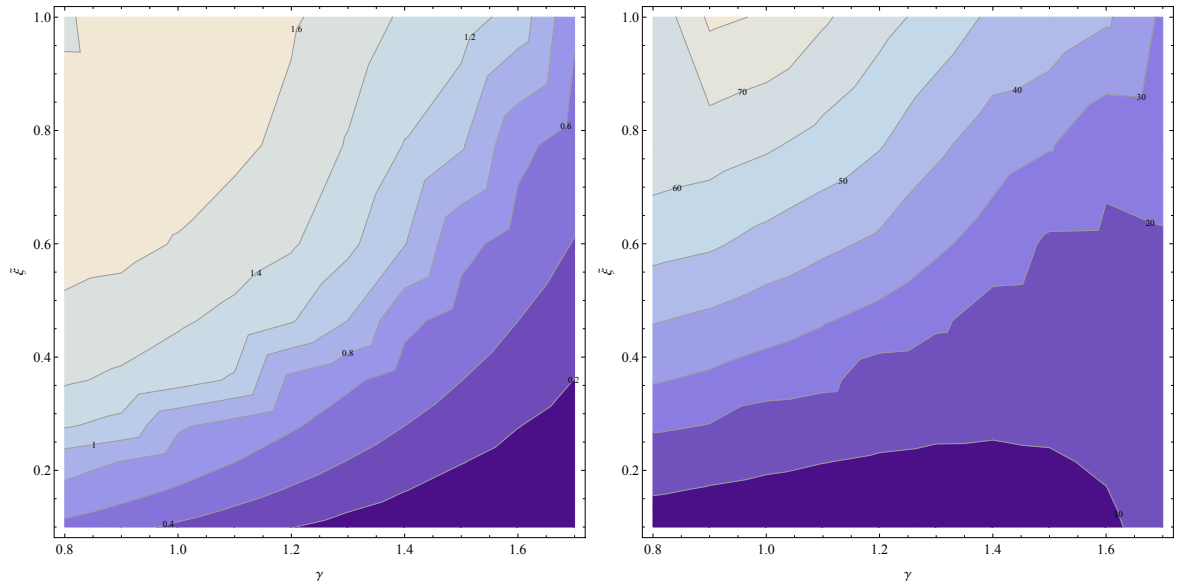

Figure 26: Best-fit values for the parameters p (left) and A (right) in equation 185.

that make the fit difficult, however, are also those that create the largest boosts to the flux amplitude [60], so it seems that, for the purposes of identifying a WIMP signal at the galactic center, accounting for the apparent inner slope is most important with Sommerfeld models *away* from resonance. Thus, in order to obtain a reliable fit to $\Delta\gamma$-$\tilde{\alpha}$ points, calculated with whatever pair of values of γ and $\tilde{\xi}$, we remove large-$\Delta\gamma$ points from the fitting algorithm until the fractional error of the model relative to the included points is no more than 0.1. Over the range of γ and $\tilde{\xi}$ of interest, no more than 20% of points are ever excluded from fits. Figure 26 plots the best-fit values of p and A as a function of $\tilde{\xi}$ for many values of γ. Figure 27 shows the minimum $\Delta\gamma$ possible as a function of γ and $\tilde{\xi}$.

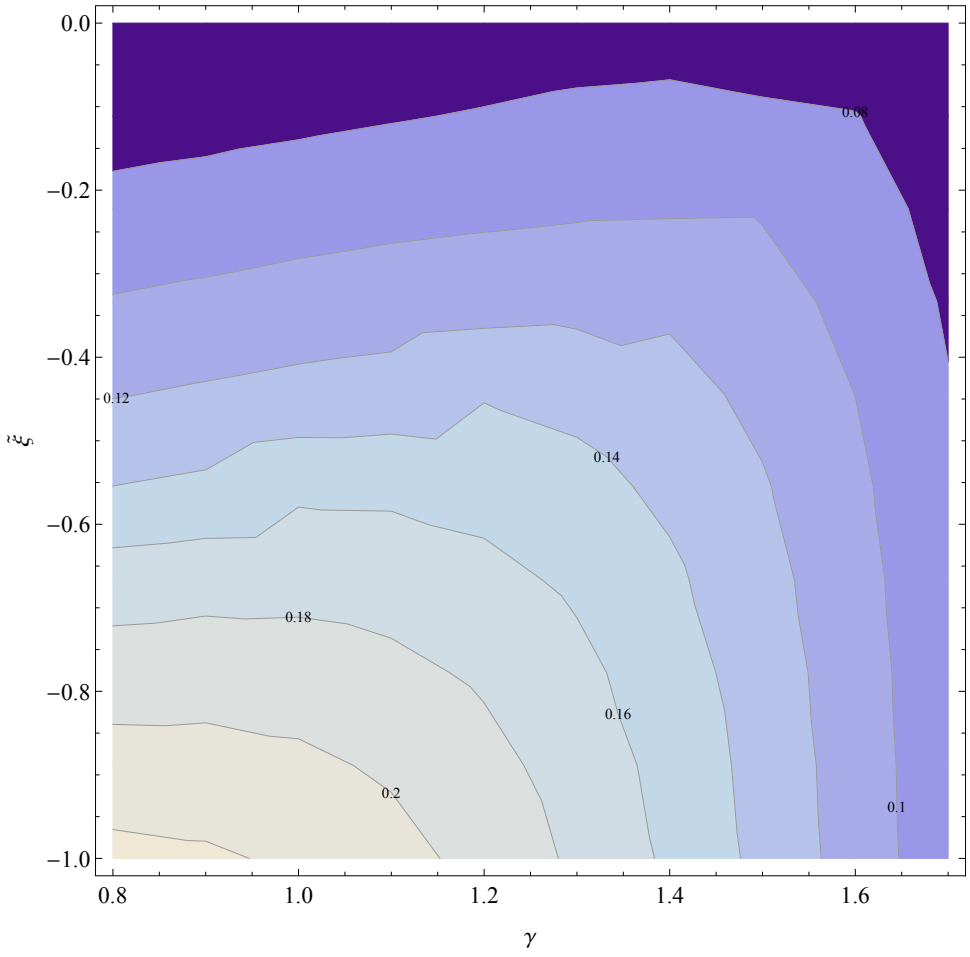

Figure 27: Contour plot of the minimum $\Delta\gamma$ possible as a function of γ and $\tilde{\xi}$.

8 Anisotropy and Implications for Detection

So far we have only discussed systems that are described by an ergodic distribution function. An ergodic distribution function depends only on the energy of a particle, which depends only on the *magnitude* of the particle's velocity. The *direction* in which any particle is moving relative to its location in the system is irrelevant. Thus the velocity distribution at any location in the system is *isotropic*. This implies that the dispersions of the velocity components are always equal. Using the components $(v_{\rm r}, v_\theta, v_\phi)$, which are the velocity magnitudes along the spherical unit vectors $(\hat{r}, \hat{\theta}, \hat{\phi})$, it is clear that

$$\sigma_{\rm r}^2 = \int {\rm d}v_{\rm r} v_{\rm r}^2 \int {\rm d}v_\theta \int {\rm d}v_\phi f\left(\Psi - \left[v_{\rm r}^2 + v_\theta^2 + v_\phi^2\right]/2\right) \tag{186}$$

and the other dispersions σ_θ and σ_ϕ have the same form, so all three are equal, and the total dispersion is simply given by the sum $\sigma^2 = 3\sigma_{\rm r}^2$.

Of course, systems with velocity anisotropy do exist (the solar system or spiral galaxies for example). The standard way of quantifying the deviation from isotropy (the "amount" of anisotropy) is the anisotropy parameter

$$\beta \equiv 1 - \frac{\sigma_\theta^2 + \sigma_\phi^2}{2\sigma_{\rm r}^2}, \tag{187}$$

which may be a position-dependent quantity. It is zero if the velocity distribution (at whatever location in a system) is isotropic. Values of β larger than zero correspond to "radial bias": the velocity of particles are likely to be pointed more closely along the radial direction $\hat{r}$. Note that β cannot be larger than one; if $\beta = 1$ then all particle orbits are exactly radial. Values of β smaller than zero correspond to "tangential bias": the velocity of particles are likely to be pointed more closely along the $(\hat{\theta}, \hat{\phi})$ plane. If all particle orbits are exactly circular, then $\sigma_{\rm r} = 0$ and $\beta = -\infty$. Often the dispersions σ_θ and σ_ϕ are assumed to be equal, so the anisotropy parameter is then witten as

$$\beta \equiv 1 - \frac{\sigma_{\rm t}^2}{2\sigma_{\rm r}^2}, \tag{188}$$

where $\sigma_{\rm t}$ is the tangential velocity dispersion.[32]

In this section we explore the ways accounting for the possibility of anisotropy can impact models and predictions. In Section 8.1 we use the anisotropic Jeans equation to derive the anisotropy profile $\beta(r)$ for a halo described by a NFW density profile *and* a power-law pseudo-phase-space density. In Section 8.2 we review some ways to derive the distribution function of an anisotropic system with a given anisotropy profile.

8.1 Jeans Analysis of the Anisotropy Profile

Here we attempt to derive the anisotropy profile $\beta(x)$ for a Milky Way-like halo that is described by a density profile and a power-law pseudo-phase-space density (PPSD). We use the NFW profile; general expressions and some details specific to NFW are deferred to Appendix E. Solving for the contained mass $M(x)$ gives us the scale density $\rho_{\rm s}$ in terms of the virial mass $M_{\rm vir} \equiv M(c)$ and concentration c.

Following Taylor and Navarro [123] (also see [53]), we take the PPSD to be a power-law with negative slope α:[33]

$$\frac{\rho}{\sigma_{\rm r}^3} = \frac{\rho_s}{\sigma_{\rm r,s}^3} x^{-\alpha}. \tag{189}$$

The radial velocity dispersion is now known (see equation 269), and its value at the scale radius $\sigma_{\rm r,s}$ may be set by assuming a local radial velocity dispersion $\sigma_{\rm r,\odot}$.

From the differential Jeans equation given in equation 149, we can solve for the anisotropy parameter (compare with [139, 113]),

$$\beta(x) = \frac{5}{6}\gamma(x) - \frac{\alpha}{3} - \frac{GM(x)}{2xr_{\rm s}\sigma_{\rm r}^2(x)}, \tag{190}$$

where we have defined the negative log-log slope of the density $\gamma(x) \equiv -{\rm d}\log(\rho)/{\rm d}\log(x)$. We know the contained mass $M(x)$, and we know the radial velocity dispersion $\sigma_{\rm r}$ from

[32] We define the tangential velocity dispersion such that $\sigma_{\rm t}^2 = \sigma_\theta^2 + \sigma_\phi^2 = 2\sigma_\theta^2$.

[33] The power of three in equation 189 was allowed to vary in [53], but it may be inappropriate to refer to such a quantity as a pseudo-phase-space density since the units no longer match.

the mass density and PPSD, so we have

$$\beta(x) = \frac{5}{6}\gamma\,(x) - \frac{\alpha}{3} - \Sigma^{-2} f(x;\alpha), \tag{191}$$

where $f(x;\alpha)$ is a somewhat complicated function of x, with α its sole parameter (*i.e.* f does not depend on the halo parameters or $\sigma_{\rm r,s}$: see equation 270 for the full expression). The quantity Σ is a dimensionless measure of the radial velocity dispersion at the scale radius, defined as

$$\Sigma^2 \equiv \frac{\sigma_{\rm r,s}^2}{4\pi G r_{\rm s}^2 \rho_{\rm s}/3}. \tag{192}$$

The denominator in equation 192 is the circular velocity squared at the edge of a spherical mass of radius $r_{\rm s}$ and constant density $\rho_{\rm s}$.

To summarize the necessary ingredients that go into equation 191, we need the PPSD slope α, the halo parameters $M_{\rm vir}$, $r_{\rm s}$, and c (one of which may be determined by the local halo density $\rho_\odot$), the local radial velocity dispersion $\sigma_{\rm r,\odot}$, and the local (solar) radius $r_\odot$.

It is shown in Appendix F that the anisotropy parameter for a NFW profile with PPSD slope $\alpha \approx 2$ has asymptotic limits

$$\beta(x) \to \begin{cases} (5-2\alpha)/6 & \text{for } x \to 0 \\ (15-2\alpha)/6 & \text{for } x \to \infty \end{cases} \tag{193}$$

This is acceptable in the small-x limit, where $\beta \to 1/6$ for, as an example, $\alpha = 2$. However, in the large-x limit, with the same value of α, $\beta \to 11/6$, which is greater than one, implying an imaginary velocity dispersion. Requiring that $\beta \leq 1$ as $x \to \infty$ would imply $\alpha \geq 9/2$, which is a far steeper slope than seen in simulations. This unphysical behavior in β may naively suggest that we cannot have a physical model that simultaneously exhibits an NFW density profile and power-law PPSD, but really this requirement for physical-ness is too restrictive. We do not expect the models or assumption of equilibrium (via the Jeans equation) to hold beyond around the virial radius. Requiring that these models are consistent and physical only up to just before they

are expected to break down is, however, reasonable and still has consequences elsewhere in a halo. Thus, let us just require that the anisotropy parameter is no greater than one everywhere *within the virial radius*.

Mathematically, we require

$$\forall x \leq c : \beta(x) \leq 1. \tag{194}$$

We can effectively satisfy this for our purposes by requiring that $\beta(c) \leq 1$. This gives a maximum value for Σ (equation 271) that depends only on the concentration c (by way of the virial mass) and PPSD log-slope α. For reasonable values of c and α, this upper limit is of order one. From the definition of Σ in equation 192, this immediately gives an upper bound on $\sigma_{\mathrm{r,s}}$ (equation 272) and thus also on $\sigma_{\mathrm{r},\odot}$ (equation 273) in terms of the halo parameters and $r_\odot$. Once $\sigma_{\mathrm{r},\odot}$ is set, the anisotropy profile $\beta(x)$ is totally specified, including the local anisotropy parameter $\beta_\odot = \beta(x_\odot)$. The *total* velocity dispersion profile σ_{tot} is then also given, using the relation $\sigma_{\mathrm{tot}}^2 = (3 - 2\beta)\sigma_{\mathrm{r}}^2$. We find that $\sigma_{\mathrm{tot},\odot}$ depends monotonically on the choice of $\sigma_{\mathrm{r},\odot}$, so we finally have an upper bound on $\sigma_{\mathrm{tot},\odot}$ (equation 274).

To get an idea of what this upper bound on $\sigma_{\mathrm{tot},\odot}$ is and its uncertainty, we use the values in Table 1 in Section 3. With these ranges of parameters, we plot the upper limit of $\sigma_{\mathrm{tot},\odot}$ versus the PPSD slope α in Figure 28. The dark, solid line uses the mean values in Table 1, while the upper and lower dashed lines take the extreme values of $\sigma_{\mathrm{tot},\odot}$ allowed by the 68% confidence intervals in Table 1. In other words, the band in Figure 28 includes all combinations of parameters within the 68% confidence intervals. Also shown is the mean value and 68% and 95% confidence intervals for $\sigma_{\mathrm{r},\odot}$ in [43].

Actually choosing a value for $\sigma_{\mathrm{r},\odot}$ (or $\sigma_{\mathrm{tot},\odot}$) determines the anisotropy profile, but this quantity is also uncertain. We use the results for $\sigma_{\mathrm{tot},\odot}$ from [43] and then take $\sigma_{\mathrm{r},\odot}^2 = \sigma_{\mathrm{tot},\odot}^2/3$, which is used to find the anisotropy profile in equation 190. Note that the factor 1/3 corresponds to the isotropic case. As we will see, we find only radial bias at the solar radius. Given the same value of $\sigma_{\mathrm{tot},\odot}$, radial bias implies a larger value of

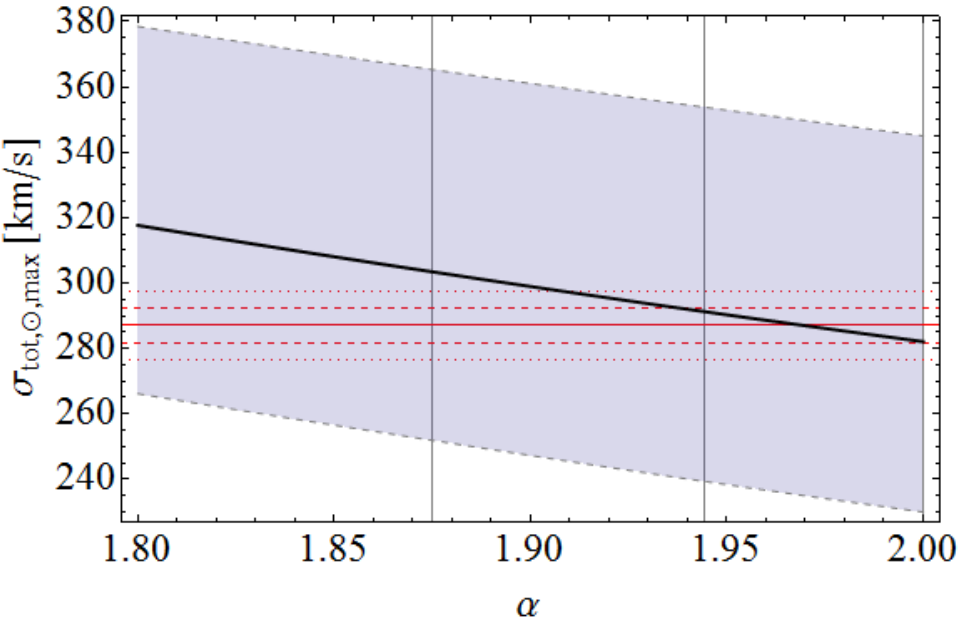

Figure 28: Maximum value of the local total velocity dispersion $\sigma_{\mathrm{tot},\odot}$. The spread of the band reflects the uncertainty in the halo parameters. The red solid line marks the mean value, while the dashed and dotted lines mark the 68% and 95% confidence intervals (see Table 1).

$\sigma_{\mathrm{r},\odot}$, which in turn gives a greater local radial bias[34]. So as far as predicting departure from isotropy, this is a conservative approximation.

We plot the anisotropy profile for fiducial values $\alpha = \{2, 35/18, 15/8\}$ in Figures 29, 30, and 31. The solid curves take the mean values in Table 1 while the dashed curves mark the area within which all parameters are within their 68% confidence interval given in Table 1. The vertical lines mark the 68% lower and upper limits of $x_\odot = r_\odot/r_\mathrm{s}$ and c. For example, if we assume a PPSD slope of 35/18 (Figure 30), we might expect a local anisotropy parameter of at least about 0.2 and no more than about 0.4.

Generally, the profile is slightly radially biased near the center, reaches a minimum at around a tenth the scale radius, and rises to a (local) maximum of around 0.4 to 0.6 before the virial radius. We see in all cases that for $x \to 0$ the anisotropy parameter rises slowly to the value in equation 277, which is independent of the halo parameters. See [69, 12] for discussion of central anisotropy. Here we do not presume that either assumed model, of the mass distribution or PPSD, necessarily stays valid at very small

[34]Successive adapting of the relation between $\sigma_{\mathrm{r},\odot}$ and $\sigma_{\mathrm{tot},\odot}$ would, of course, converge to the correct "trial value" for $\beta_\odot$.

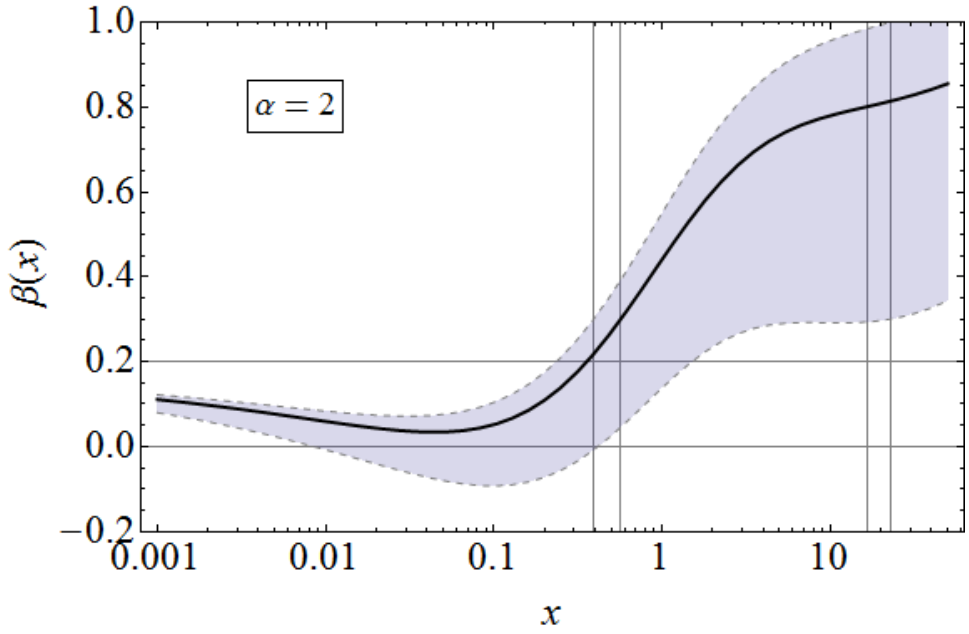

Figure 29: The anisotropy profile for $\alpha = 2$, corresponding to the isothermal case.

or very large radii. See [84] for an investigation of the break-down of the PPSD power law.

Recently, close attention has been paid to the form of the velocity distribution used to calculate predictions for indirect and direct DM detection. In some cases the functional form can make a significant difference. Especially, the assumed velocity distribution influences the interpretation of results from direct detection experiments [129, 133, 43]. Here on we focus on the local distribution and suppress the subscript $\odot$. We introduce a new, anisotropic generalization of the model proposed by Mao, et al. [88]:

$$f(\mathbf{v}) \propto \exp\left\{-\sqrt{\frac{v_{\rm r}^2}{v_{\rm r,0}^2}+\frac{v_{\rm t}^2}{v_{\rm t,0}^2}}\right\}\left(v_{\rm esc}^2-v^2\right)^p, \tag{195}$$

where $v_{\rm r} = v\cos\eta$ and $v_{\rm t} = v\sin\eta$ are the radial and tangential velocity components and η is the angle from the radial direction. The parameters $v_{\rm r,0}$ and $v_{\rm t,0}$ are not dispersions but just velocity scales. The exponent p characterizes the high-velocity tail. The function is normalized so that $\int \mathrm{d}\mathbf{v} f(\mathbf{v}) = 1$. We choose this distribution because of its recent success in modeling the Eris and ErisDark simulations (see Figure 3 in [79]). For consis-

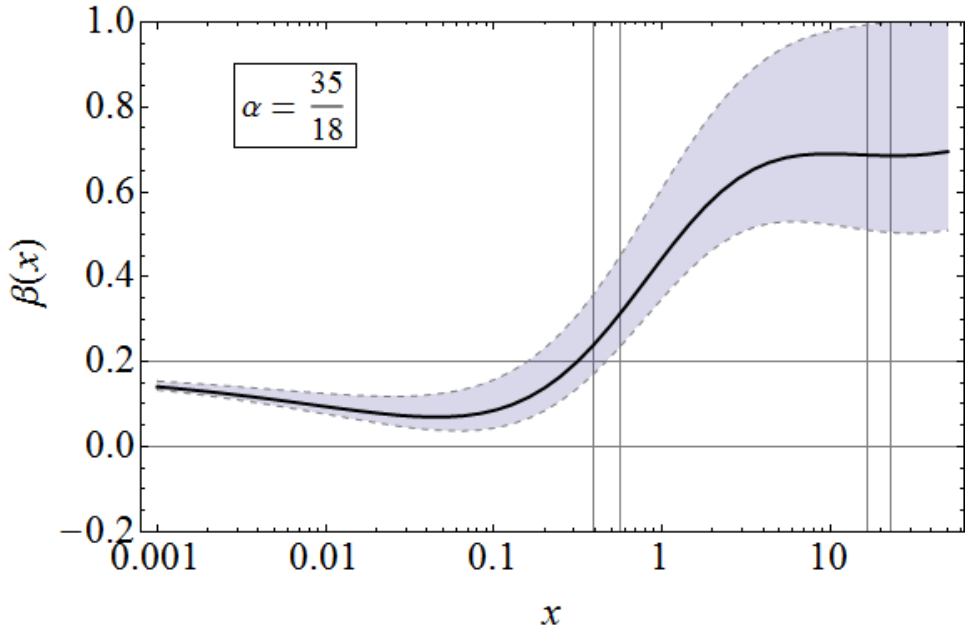

Figure 30: The anisotropy profile for $\alpha = 35/18$, the critical value discussed in [53].

tency with that study we take $p = 1.5$, which was used to model the ErisDark results.[35] The escape speed $v_{\rm esc}$ is given by the combined gravitational potential of both the DM halo and any other matter, and we use the mean value in Table 1. The total dispersion $\sigma_{\rm tot}$ and the anisotropy parameter β are then determined by the parameters $v_{\rm r,0}$ and $v_{\rm t,0}$. We require that the total dispersion equals the mean value in Table 1 and solve for $v_{\rm r,0}$ and $v_{\rm t,0}$ such that the desired anisotropy parameter is generated. Of course, the original, isotropic distribution is recovered when $v_{\rm r,0} = v_{\rm t,0}$. See Appendix G for details on the selection of values for $v_{\rm r,0}$ and $v_{\rm t,0}$.

We have checked that the uncertainties in the values of $\sigma_{\rm tot}$ and $v_{\rm esc}$ have a small impact on the following calculations. More importantly, the uncertainties affect both the isotropic and anisotropic cases equally once β has been chosen. So for the purposes of investigating the importance of modeling deviation from isotropy, we show only results using the mean values in Table 1.

We use the function in equation 195 to model the local velocity distribution with the intention of understanding the impact that anisotropy can have on direct detection. For

[35] We do not take the Eris parameter $p = 2.7$ for two reasons: we have not considered baryonic effects on the PPSD profile, and because such a steep cut-off makes it difficult to achieve anisotropy greater than $\beta \approx 1.0$ with the model in equation 195.

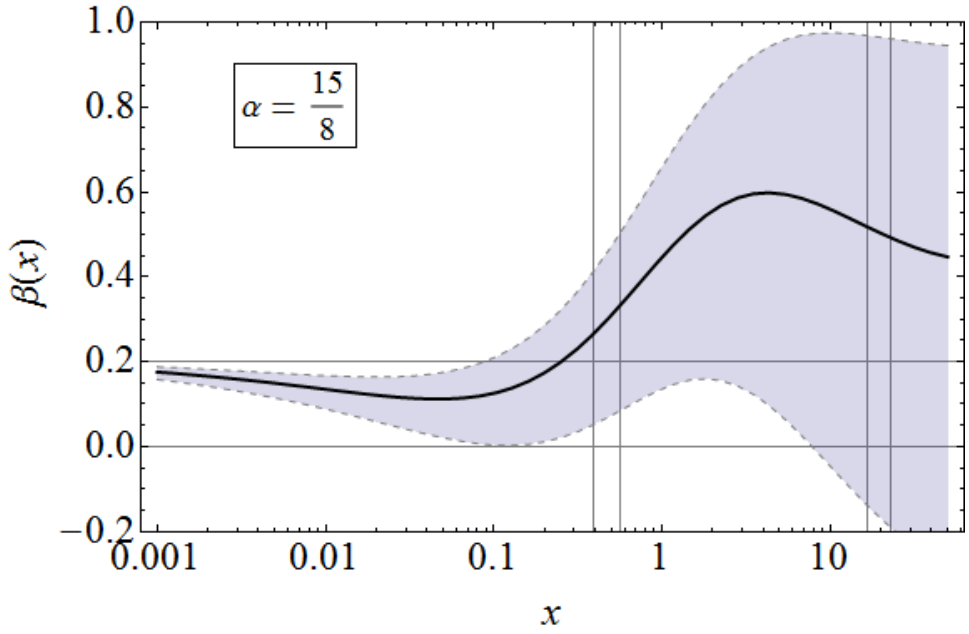

Figure 31: The anisotropy profile for $\alpha = 15/8$, the value found in [123].

the purposes of this work, we assume a conservative value of 0.2 for the anisotropy parameter β. It is straight-forward to calculate the function $g(v,t)$ defined in equation 136. In Figure 32 we plot this function for June and December; for the isotropic case and the anisotropic case. The differential detection rate is found by specifying a velocity threshold $v_{\rm th}$ for DM particles in the detector frame:

$$\frac{\mathrm{d}R}{\mathrm{d}Q} \propto G\left(v_{\rm th}, t\right) \equiv \int_{v \geq v_{\rm th}} \mathrm{d}v \, g\left(v, t\right). \tag{196}$$

The velocity threshold is determined by the specifics of any particular experiment and the DM particle mass, and we leave it free (see equation 133). Figure 33 plots the function G, averaged between June and December, for the isotropic and anisotropic cases, with the fractional difference

$$\Delta G = \left(G_{\rm Ani} - G_{\rm Iso}\right)/G_{\rm Iso}. \tag{197}$$

We also consider the modulation amplitude of the signal, defined here as half the difference between the rate in June and the rate in December:

$$A\left(v_{\rm th}\right) = \left|G\left(v_{\rm th}, t_{\rm June}\right) - G\left(v_{\rm th}, t_{\rm Dec}\right)\right|/2. \tag{198}$$

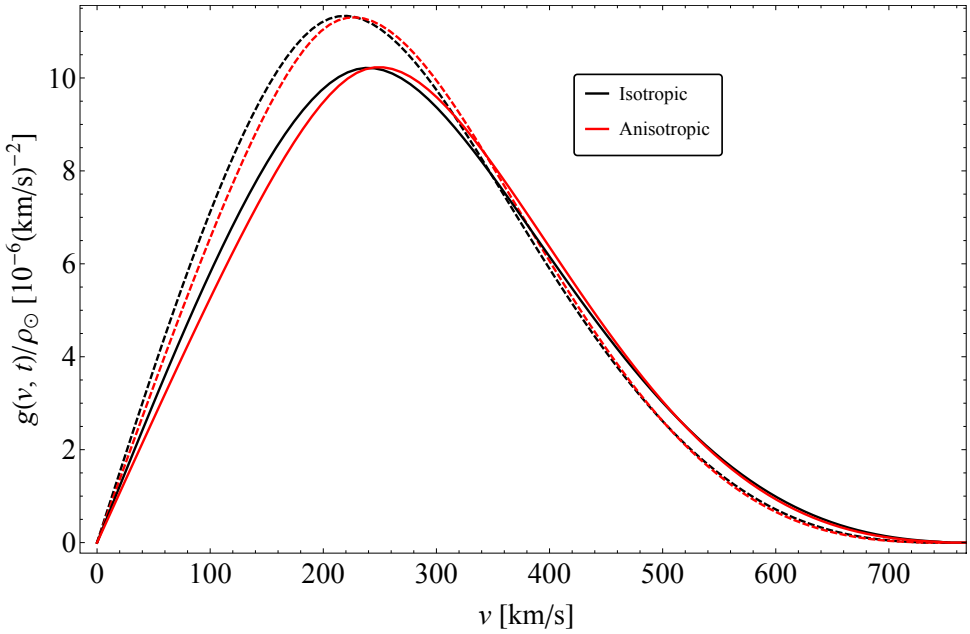

Figure 32: The function $g\,(v,t)$ defined in equation 136. Solid lines are calculated in June; dashed lines are calculated in December.

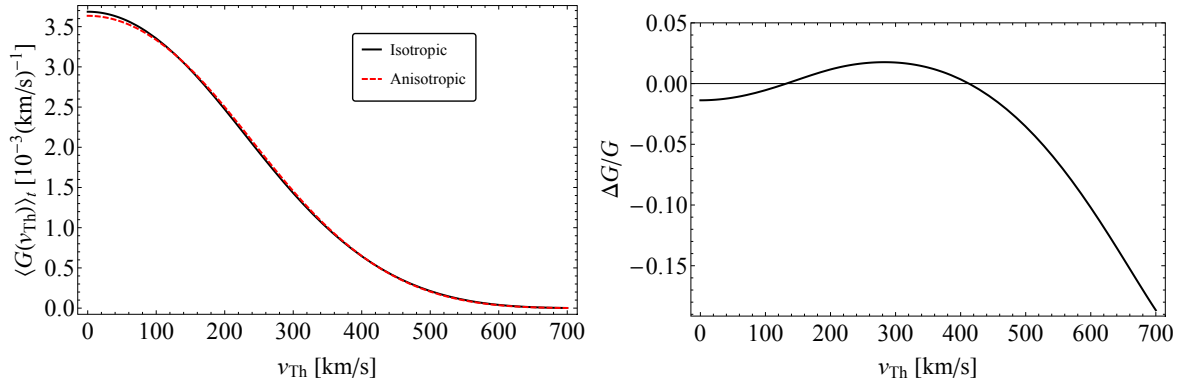

Figure 33: The time-averaged function $\langle G\,(v_{\mathrm{th}},t)\rangle_t$ as a function of the velocity threshold (see equation 196) and the fractional difference between the isotropic to anisotropic cases.

This is plotted in Figure 34 for the isotropic and anisotropic cases, with the fractional difference, analogous to equation 197.

Combining models of the mass distribution and pseudo-phase-space density, the Jeans equation gives us a particular anisotropy profile. We have plotted this profile for a few representative values of the PPSD slope and for a spread of parameters that may describe the Milky Way halo. These profile shapes are consistent with those shown in Figure 1 of [139], although those results exhibit less anisotropy overall. The anisotropy profiles found in [34] are also similar but were derived from models of the phase-space distribution. The difference in methods strengthens both their results and these.

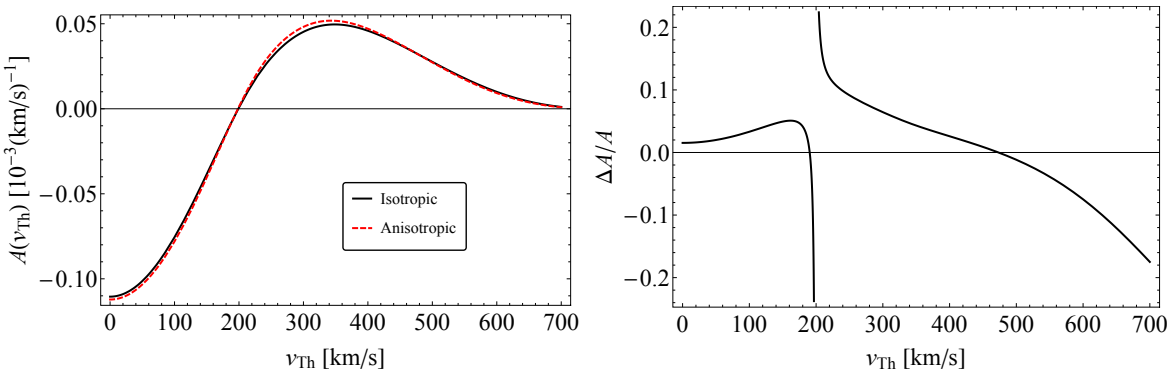

Figure 34: Signal modulation amplitude as a function of the velocity threshold and the fractional difference between the isotropic to anisotropic cases.

We have used an anisotropic modification to the model proposed by Mao et al. [88], which was also used to model the Eris simulation. We find that assuming a local anisotropy of approximately 0.2 is reasonable and conservative. In the Eris simulation, a comparable amount of radially biased anisotropy was found at the location corresponding to the solar radius (this is roughly seen by measuring the half-maximum width of the radial and azimuthal distributions in Figure 2 of [79]). On the other hand, the results of [42] favor a local *tangential* bias, though the small local radial bias found in this work and others already mentioned is approximately within their 95% confidence interval.

Different direct detection collaborations have found contradictory results (*e.g.* see [64]). One of the goals in studying the local velocity distribution is to alleviate these discrepancies. Since different experiments can have different threshold velocities, Figure 33 suggests that the difference between observed signals can vary by several percent due to the effect of local anisotropy. This may seem small, but it is comparable to the uncertainty introduced by considering different density profiles [43]. The modification to the modulation amplitude can be even more significant and is sensitive to the value of the velocity threshold, but the signal itself is smallest where the modification is greatest.

In principle, a detector that can give information about the direction of a detected WIMP's velocity would allow us to *measure* the local anisotropy. This is difficult, as it would require an individual WIMP to interact multiple times inside the detector or require a low detector density so the recoiled particle can be tracked. Once a discovery is confirmed, however, it may be viable to consider such an experiment (see [62]).

We note that the anisotropy at radii beyond about the scale radius is sensitive to the shape of the PPSD profile and to the other parameters, and it can also be quite large. However, it seems unlikely that this grants a viable observational effect, since the density is so low there and substructure would dominate any emission.

The most novel result of this work is the constraint on the velocity dispersion profile. Requiring the anisotropy parameter to be physical (*i.e.* no greater than one) inside the virial radius implies a maximum value for the local total velocity dispersion of about 300 km/s or so. Typical assumed values for the local velocity dispersion (such as in the Standard Halo Model, 220 km/s) do not seem to be in great danger, but this consistency check should be remembered in future model-building.

8.2 Anisotropic Distribution Functions

An ergodic distribution function is a function only of energy. As discussed in Section 7, systems described by ergodic DFs are spherical and isotropic. It is clear from N-body simulations, however, that DM halos are not spherically symmetric or isotropic except at their very centers. More accurately, DM halos are triaxial and have increasing radial anisotropy in the velocity distribution going outward [132]. How can we include velocity anisotropy in our models in a self-consistent way? Eddington's equation 162 is specific to ergodic DFs, but there are analogous equations for anisotropic models that we will now discuss.[36]

The simplest anisotropic DF is one with constant anisotropy, $\beta(r) = \beta \neq 0$. The DF is of the form

$$f(\mathcal{E}, L) = L^{-2\beta} f_1(\mathcal{E}), \tag{199}$$

where $L = r v_t = r v \sin\eta$ is the angular momentum. It is straightforward to check by calculating velocity moments that this form confirms the definition of β in equation 187.

[36] Analogous equations also exist for axisymmetric distributions [85].

Marginalizing over velocity to recover the mass distribution $\rho(r)$ gives the relation

$$\rho = \frac{4\pi I_\beta}{r^{2\beta}} \int_0^\infty \mathrm{d}v \, v^{2-2\beta} f_1\left(\mathcal{E}\right), \tag{200}$$

$$I_\beta \equiv \frac{\sqrt{\pi}}{2} \frac{(-\beta)!}{\left(\frac{1}{2}-\beta\right)!}, \tag{201}$$

Rearranging terms and changing the variable of integration, we have

$$\frac{2^{\beta-1/2}}{4\pi I_\beta} r^{2\beta}\rho = \int_0^\Psi \mathrm{d}\mathcal{E} \frac{f_1\left(\mathcal{E}\right)}{\left(\Psi-\mathcal{E}\right)^{\beta-1/2}}. \tag{202}$$

This is an Abel integral equation for $\frac{1}{2} < \beta < \frac{3}{2}$ and can be immediately inverted. If $\beta < \frac{1}{2}$, derivatives may be taken to make it Abel:

$$\frac{2^{\beta-1/2}}{4\pi I_\beta} \frac{\mathrm{d}^n}{\mathrm{d}\Psi^n}\left[r^{2\beta}\rho\right] = \prod_{m=1}^{n}\left(\frac{3}{2}-\beta-m\right) \int_0^\Psi \mathrm{d}\mathcal{E} \frac{f_1\left(\mathcal{E}\right)}{\left(\Psi-\mathcal{E}\right)^{\beta-1/2+n}}, \tag{203}$$

where n is the smallest integer that satisfies $\frac{1}{2}\left(1-2n\right) < \beta$. Now, for $\frac{1}{2}-n < \beta < \frac{3}{2}-n$, this can be inverted to give

$$f_1\left(\mathcal{E}\right) = \frac{2^{\beta-1/2}}{4\pi^2 I_\beta} \prod_{m=1}^{n}\left(\frac{3}{2}-\beta-m\right)^{-1} \sin\left(\pi\left(\beta-\frac{1}{2}+n\right)\right) \cdot \frac{\mathrm{d}}{\mathrm{d}\mathcal{E}} \int_0^\mathcal{E} \frac{\mathrm{d}\Psi}{\left(\mathcal{E}-\Psi\right)^{\frac{3}{2}-\beta-n}} \frac{\mathrm{d}^n}{\mathrm{d}\Psi^n}\left[r^{2\beta}\rho\right] \tag{204}$$

or

$$f_1\left(\mathcal{E}\right) = \frac{2^{\beta-1/2}}{4\pi^2 I_\beta} \prod_{m=1}^{n}\left(\frac{3}{2}-\beta-m\right)^{-1} \sin\left(\pi\left(n+\beta-\frac{1}{2}\right)\right) \cdot \left[\int_0^\mathcal{E} \frac{\mathrm{d}\Psi}{\left(\mathcal{E}-\Psi\right)^{\frac{3}{2}-\beta-n}} \frac{\mathrm{d}^{n+1}}{\mathrm{d}\Psi^{n+1}}\left[r^{2\beta}\rho\right] + \frac{1}{\mathcal{E}^{\frac{3}{2}-\beta-n}} \frac{\mathrm{d}^n}{\mathrm{d}\Psi^n}\left[r^{2\beta}\rho\right]\Big|_{\Psi=0}\right] \tag{205}$$

It is easy to check that this reduces to the usual isotropic Eddington equation for $\beta = 0$ and $n = 1$.

Equation 203 cannot be an Abel integral equation for half-integer β, i.e. $\beta = \frac{1}{2} - n$

with $n \in \mathbb{Z}_{\geq 0}$. However, this equation simplifies tremendously in these cases:

$$\frac{2^{-1-n}}{\pi^{3/2}\,\Gamma\left(\frac{1}{2}+n\right)}\frac{\mathrm{d}^n}{\mathrm{d}\Psi^n}\left[r^{1-2n}\rho\right] = \int_0^\Psi \mathrm{d}\mathcal{E}\, f_1\left(\mathcal{E}\right), \tag{206}$$

The energy-dependent part of the distribution function is then simply

$$f_1\left(\mathcal{E}\right) = \frac{2^{-1-n}}{\pi^{3/2}\,\Gamma\left(\frac{1}{2}+n\right)}\,\frac{\mathrm{d}^{n+1}}{\mathrm{d}\Psi^{n+1}}\left[r^{1-2n}\rho\right]\Big|_{\Psi=\mathcal{E}} \tag{207}$$

Osipkov-Merritt models instead take the distribution function as [99]

$$f\left(\mathcal{E}, L\right) = f\left(Q\right), \tag{208}$$

where $Q = \mathcal{E} - L^2/2r_\mathrm{a}^2$ and r_a is the *anisotropy radius*. The anisotropy parameter is now *variable* from 0 to 1 as

$$\beta\left(r\right) = \frac{r^2}{r_\mathrm{a}^2 + r^2}. \tag{209}$$

The value of r_a is the radius around which the anisotropy changes in the system. Note that in the limit of $r_\mathrm{a} \to \infty$ the Osipkov-Merritt model reduces to that of an isotropic system. The equation for Osipkov-Merritt models analogous to equation 157 and 202 is

$$\frac{1}{\sqrt{8}\pi}\left(1+\frac{r^2}{r_\mathrm{a}^2}\right)\rho = 2\int_0^\Psi \mathrm{d}Q\, f\left(Q\right)\sqrt{\Psi - Q}. \tag{210}$$

This can be made Abel by taking the derivative with respect to Ψ and can then be inverted, similar to the isotropic case. The result analogous to the Eddington equation is

$$f(Q) = \frac{1}{\sqrt{8}\pi}\frac{\mathrm{d}}{\mathrm{d}Q}\int_0^Q \frac{\mathrm{d}\Psi}{\sqrt{Q-\Psi}}\frac{\mathrm{d}}{\mathrm{d}\Psi}\left[\left(1+\frac{r^2}{r_\mathrm{a}^2}\right)\rho\right]. \tag{211}$$

Note that we cannot make the integral in equation 210 trivial by choosing parameters as we could to find equation 207. This is a limitation of the Osipkov-Merritt model, but we can restore the capability by the following.

We now generalize to Cuddeford models, which take the distribution function as

$$f\left(\mathcal{E}, L\right) = L^{-2\beta_0} f_1\left(Q\right), \tag{212}$$

where Q is defined as before and β_0 is a constant [50]. The anisotropy now varies from β_0 to 1 as

$$\beta\left(r\right) = \frac{\beta_0\, r_\mathrm{a}^2 + r^2}{r_\mathrm{a}^2 + r^2}. \tag{213}$$

Note that in the limit of $r_\mathrm{a} \to \infty$ the Cuddeford model also reduces to that of constant anisotropy, and in the limit of $\beta_0 \to 0$ it reduces to the Osipkov-Merritt model. The equation analogous to equation 202 is

$$\frac{2^{\beta_0 - 1/2}}{4\pi I_{\beta_0}} r^{2\beta_0} \left(1 + \frac{r^2}{r_\mathrm{a}^2}\right)^{1-\beta_0} \rho = \int_0^\Psi \mathrm{d}Q \frac{f_1\left(Q\right)}{\left(\Psi - Q\right)^{\beta_0 - 1/2}}. \tag{214}$$

Restricting to half-integer values of β_0, we have, similar to equation 207,

$$f_1\left(Q\right) = \frac{2^{-1-n}}{\pi^{3/2}\,\Gamma\left(\frac{1}{2} + n\right)} \frac{\mathrm{d}^{n+1}}{\mathrm{d}\Psi^{n+1}} \left[r^{1-2n} \left(1 + \frac{r^2}{r_\mathrm{a}^2}\right)^{1/2+n} \rho \right] \Bigg|_{\Psi = \mathcal{E}} \tag{215}$$

As in Section 7, we can trivially obtain the (anisotropic) single-particle velocity distribution from the phase-space distribution function. In the general case that $f = f(\mathcal{E}, L)$, we have, analogous to equation 168,

$$P_1(v_\mathrm{r}, v_\mathrm{t}; r) = \frac{f\left(\Psi(r) - \left(v_\mathrm{r}^2 + v_\mathrm{t}^2\right)/2, r v_\mathrm{t}\right)}{\rho(r)}. \tag{216}$$

The calculation of the relative velocity distribution is complicated by anisotropy. Because of the additional variable in P_1, the integral analogous to that in equation 171 becomes a *four* dimensional integral:

$$P(v_\mathrm{rel}) = 8\pi v_\mathrm{rel}^2 \int_0^\pi \mathrm{d}\eta \sin\eta \int_0^\pi \mathrm{d}\Delta\eta \int \mathrm{d}v_{\mathrm{cm,r}} \int \mathrm{d}v_{\mathrm{cm,t}} v_{\mathrm{cm,t}} P_1\left(v_{1,\mathrm{r}}, v_{1,\mathrm{t}}\right)\, P_1\left(v_{2,\mathrm{r}}, v_{2,\mathrm{t}}\right), \tag{217}$$

where the radial and tangential components of the individual particles are given by

$$v_{(1,2),\mathrm{r}} = |v_{\mathrm{cm,r}} \pm v_{\mathrm{rel}} \cos(\eta)/2| \tag{218}$$

$$v^2_{(1,2),\mathrm{t}} = v^2_{\mathrm{cm,t}} + (v_{\mathrm{rel}} \sin(\eta)/2)^2 \pm v_{\mathrm{cm,t}} v_{\mathrm{rel}} \sin(\eta) \cos(\Delta\eta). \tag{219}$$

9 Conclusions

The open problem of dark matter is one of the biggest in modern astrophysics. The full seriousness of the problem was realized from calculations of big bang nucleosynthesis and the cosmic microwave background power spectrum, discussed in Sections 2.1 and 2.2. The baryonic material with which we are most familiar is actually a very small fraction of the total mass and energy in the universe. The history of dark matter-related observations goes back further, including the dynamics of galaxies and clusters of galaxies, discussed in Section 3.

The investigation of particle dark matter theory has been broad and deep. The identification of dark matter's particle properties is the next big step toward a full understanding and could serve as a jumping off point for further physics beyond the Standard Model such as Supersymmetry (see Section 4.1). A great amount of imagination and carefulness has gone into making predictions for indirect and direct detection (see Sections 5 and 6). In Section 7 we performed the most detailed calculation to date of the total flux and flux slope from galactic dark matter that annihilates with a complicated velocity-dependent cross-section. A novel result about the local dark matter velocity dispersion was found from a Jeans analysis of anisotropic galactic halos in Section 8.1, and a generalization of a successful empirical model was introduced.

Experimental efforts have been impressive, but are so far inconclusive. Considering this null result, some theoretical results may seem strangely precise. The necessity of the detailed computation in Section 7 is very significant only for particular cases and only given the assumption of a Sommerfeld enhancement. The 2% change due to anisotropy found in Section 8.1 seems unlikely to settle the disgreements between direct detection experiments, especially now that LUX is so dominating. Such results may seem desperate, but they are not wasted. Once a discovery is actually made, a tremendous amount of relevant work could already be done in accidental preparation. We will thus likely learn much detail about galactic dark matter very quickly once a detection is confirmed, thanks to efforts such as this work.

A Statistics and Thermodynamics in Cosmology

The universe is not an equilibrium system, but in many cases particle interactions are so rapid that we can assume *local* equilibrium. The test of this is to compare the rate of the interactions of interest against the expansion rate. The rate at which incident particles with density n and relative velocity v interact with target particles with cross-section σ is generically

$$\Gamma = \sigma n v. \tag{220}$$

Comparing this with the expansion rate, if the ratio

$$\Gamma/H \tag{221}$$

is much greater than unity we can assume that equilibrium is reached. Another way to think of this is to define a collision time $t_c \simeq 1/\Gamma$ and compare this against the cosmic time $t_H \simeq 1/H$. A system is in equilibrium if many collisions happen during a cosmic time.

Particles are either bosons or fermions, and in equilibrium they follow the respective statistics, the well-known Bose-Einstein and Fermi-Dirac distributions:

$$f(\epsilon) = \frac{1}{\exp\left\{(\epsilon - \mu)/k_{\mathcal{B}}T\right\} \mp 1}, \tag{222}$$

where the sign is $-$ for bosons and $+$ for fermions. This is the fraction of states with energy ϵ that are filled. To get a number density, the number of particles in a physical volume, we need to calculate the number of states per energy. This is the phase-space volume element, calculated in Appendix B. The number density, energy density, pressure,

and entropy density are in general

$$n(T) = \frac{1}{(\hbar c)^3}\frac{g}{2\pi^2}\int_m^\infty \frac{\varepsilon\sqrt{\varepsilon^2 - m^2}\,\mathrm{d}\varepsilon}{\exp\{\varepsilon/k_{\mathcal{B}}T\} \mp 1}, \tag{223}$$

$$\rho(T) = \frac{1}{(\hbar c)^3}\frac{g}{2\pi^2}\int_m^\infty \frac{\varepsilon^2\sqrt{\varepsilon^2 - m^2}\,\mathrm{d}\varepsilon}{\exp\{\varepsilon/k_{\mathcal{B}}T\} \mp 1}, \tag{224}$$

$$P(T) = \frac{\rho(T)c^2}{3} - \frac{1}{\hbar^3 c}\frac{m^2 g}{6\pi^2}\int_m^\infty \frac{\sqrt{\varepsilon^2 - m^2}\,\mathrm{d}\varepsilon}{\exp\{\varepsilon/k_{\mathcal{B}}T\} \mp 1}, \tag{225}$$

$$s(T) = \frac{\rho(T) + P(T) - \mu n(T)}{T}. \tag{226}$$

For very relativistic or massless particles, we set $m = 0$ and, if the chemical potential is negligible, we can evaluate the integrals for the cases of bosons and fermions. The number density is

$$n(T) = \begin{cases} \frac{\zeta(3)}{\pi^2} g T^3 & \text{bosons} \\ \frac{3\zeta(3)}{4\pi^2} g T^3 & \text{fermions.} \end{cases} \tag{227}$$

The energy density is

$$\rho(T) = \begin{cases} \frac{\pi^2}{30} g T^4 & \text{bosons} \\ \frac{7\pi^2}{240} g T^4 & \text{fermions.} \end{cases} \tag{228}$$

The pressure is simply

$$P(T) = \frac{\rho c^2}{3}. \tag{229}$$

The entropy density is (again, neglecting chemical potential)

$$s = (\rho + P)/T \tag{230}$$

$$= \frac{4\rho}{3T}. \tag{231}$$

The constant multiplicative difference of 7/8 in the energy densities of bosons and fermions is important. It means that the contributions to the total relativistic energy density, pressure, and entropy density are the same for each degree of freedom of any particle species

at temperature T, *except* for a factor of 7/8 for fermions. We may write

$$\rho(T) = \frac{\pi^2}{30}\mathcal{N}T^4, \tag{232}$$

$$s(T) = \frac{2\pi^2}{45}\mathcal{N}T^3, \tag{233}$$

where we have defined the effective number of degrees of freedom

$$\mathcal{N} \equiv \sum_B g_B + \frac{7}{8}\sum_F g_F, \tag{234}$$

with the summation being over the numbers of degrees of freedom of bosonic species and of fermionic species.

In the case of very non-relativistic particles, if we have $(m-\mu)/T \gg 1$ then spin-statistics are unimportant and the number density is

$$n(T) \simeq \frac{1}{(\hbar c)^3} g \left(\frac{m k_{\mathcal{B}} T}{2\pi}\right)^{3/2} \exp\left\{-\left(m-\mu\right)/k_{\mathcal{B}}T\right\}. \tag{235}$$

The energy density is simply the mass density

$$\rho(T) = mn(T), \tag{236}$$

and the pressure is negligible. Neglecting chemical potential, the entropy density is easily found from equation 226:

$$s(T) = \frac{mn(T)}{T}. \tag{237}$$

B Neutron → Proton Rate

Calculating the amplitude for these diagrams is straightforward, using the Feynman rules for fermions and the weak vertex factor.[37] For the first process we find

$$\mathcal{M}_{n\nu} = -\frac{g_w^2}{8}\left[\overline{u}(p)\gamma^\mu(1-\gamma^5)u(n)\right] W_{\mu\nu}\left[\overline{u}(e)\gamma^\nu(1-\gamma^5)u(\nu)\right], \tag{238}$$

[37] The weak coupling is a "vector minus axial vector" coupling that violates parity maximally.

where $g_w = \sqrt{4\pi\alpha_w}$ is the weak coupling constant. The propagator for a W boson with momentum q is, in general,

$$W_{\mu\nu} = \frac{-i\left(g_{\mu\nu} - q_\mu q_\nu / M_W^2\right)}{q^2 - M_W^2}, \tag{239}$$

but the typical momentum transfer of the interactions we are discussing is small in comparison to the W boson mass $M_W = 80.385\,\mathrm{GeV}$[22], so we may simplify this to

$$\mathcal{M}_{n\nu} = -\frac{ig_w^2}{8M_W^2}\left[\overline{u}(p)\gamma^\mu(1-\gamma^5)u(n)\right]\left[\overline{u}(e)\gamma_\mu(1-\gamma^5)u(\nu)\right]. \tag{240}$$

As is usual, we are not interested in the spins of particles, so we sum over possible spins. Applying "Casimir's trick" [68] after squaring the amplitudes, we find

$$\sum_{\{s_i\}} |\mathcal{M}_{n\nu}|^2 = \frac{G_F^2}{2}\mathrm{Tr}\left[\gamma^\mu\left(\not{p}_n + m_n\right)\gamma^\nu\left(\not{p}_p + m_p\right)\right] \times \mathrm{Tr}\left[\gamma_\mu \not{p}_\nu \gamma_\nu\left(\not{p}_e + m_e\right)\right] \tag{241}$$

Working out the traces and averaging over the initial spins, we find

$$\langle|\mathcal{M}_{n\nu}|^2\rangle = 64G_F^2(p_n \cdot p_\nu)\,(p_p \cdot p_e). \tag{242}$$

The amplitude for the other two processes illustrated in Figure 1 have very similar results. We can write a general amplitude-squared for all three neutron-to-proton interactions as

$$\langle|\mathcal{M}|^2\rangle = \frac{1}{2s_1}\mathcal{A}^2, \tag{243}$$

where s_1 is the number of spin states the incoming lepton can take (or, if there is no incoming lepton as in the case of neutron decay, $s_1 = 1$). In the first and third processes, the neutron has two spin states (and the neutrino has one), so $2s_1$ is two total; in the second process the incoming positron also has two spin states, bringing the total to four. For convenience we have defined

$$\mathcal{A}^2 \equiv 128G_F^2(p_n \cdot p_\nu)\,(p_p \cdot p_e). \tag{244}$$

To find the rates at which these three processes occur, we must find the cross-section in the case of the first two and the decay rate in the case of the third. First let us find the cross-section of the scattering processes. The differential cross-section is given by the Lorentz-invariant expression

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left(\frac{1}{8\pi}\right)^2 \frac{S\mathcal{A}^2/2s_1}{(p_n + p_1)^2} \sqrt{\frac{(p_p \cdot p_2)^2 - m_p^2 m_2^2}{(p_n \cdot p_1)^2 - m_n^2 m_1^2}}. \tag{245}$$

Here 1 and 2 designate the incoming and outgoing leptons, respectively. The statistical factor S is necessary because there are many free (anti-)neutrinos and, if they have not yet annihilated, many free electrons and positrons. These fill up the fermionic states available for the *outgoing* leptons in the interactions in question. Thus, the cross-section is reduced by the factor $(1 - f_2)$, where f_2 is the Fermi-Dirac distribution for the outgoing lepton.[38]

We calculate the cross-section in the neutron's rest-frame, so $(p_n + p_1)^2 \simeq m_n^2$ and $(p_n \cdot p_1)^2 - m_n^2 m_1^2 \simeq m_n^2 \left|\mathbf{p}_1\right|^2$. The daughter proton is non-relativistic, so we have similar relations between it and the outgoing lepton. Using these and integrating over solid angle, the $n + l_1 \to p + l_2$ cross-section is

$$\sigma = \frac{1}{16\pi}(1 - f_2)\frac{\mathcal{A}^2}{2s_1}\frac{1}{m_n^2}\frac{m_p|\mathbf{p}_2|}{m_n|\mathbf{p}_1|} \tag{246}$$

$$\sigma = \frac{1}{16\pi}(1 - f_2)\frac{2(g_w^2/M_W)^4\epsilon_1\epsilon_2}{s_1}\frac{m_p^2}{m_n^2}\frac{|\mathbf{p}_2|}{|\mathbf{p}_1|} \tag{247}$$

$$\sigma = \frac{1}{16\pi}(1 - f_2)\frac{2(g_w^2/M_W)^4\epsilon_2^2}{s_1}\frac{m_p^2}{m_n^2}\frac{v_2}{v_1} \tag{248}$$

We have a population of neutrons and a incoming flux of leptons with speed $v_1 = |\mathbf{p}_1|/\epsilon_1$. The rate at which neutrons and leptons with energy between ϵ_1 and $\epsilon_1 + \mathrm{d}\epsilon_1$ interact to create a proton and a lepton (of a different kind) with energy between ϵ_2 and $\epsilon_2 + \mathrm{d}\epsilon_2$ is

$$\mathrm{d}\Gamma_{n1}(\epsilon_1, \epsilon_2) = n_1(\epsilon_1; T_1)\,\sigma(\epsilon_1, \epsilon_2)\frac{|\mathbf{p}_1|}{\epsilon_1}\delta(\epsilon_2 - \epsilon_1 - Q)\,\mathrm{d}\epsilon_1\,\mathrm{d}\epsilon_2, \tag{249}$$

[38] Of course, if these interactions were occuring in otherwise empty space, then $f \to 0$, $S \to 1$, and the daughter particles would be free to take any state whatsoever.

where $n_1(\epsilon_1; T_1)\,\mathrm{d}\epsilon_1$ is the number density of incident leptons with energy between ϵ_1 and $\epsilon_1 + \mathrm{d}\epsilon_1$ at temperature T_1, and $Q = m_n - m_p$. This is the product of the Fermi-Dirac distribution and the phase-space volume element. For any energy ϵ_1 there are s_1 possible state(s) for a particle, and the probability that any particular state is filled is given by the Fermi-Dirac distribution, so we have

$$n_1(\epsilon_1; T_1)\,\mathrm{d}\epsilon_1 = s_1 f_1 \int_{|\mathbf{p}_1|}^{|\mathbf{p}_1|+\mathrm{d}|\mathbf{p}_1|} \mathrm{d}^3p/(2\pi\hbar)^3 \tag{250}$$

$$\simeq s_1 f_1 \frac{|\mathbf{p}_1|^2\,\mathrm{d}|\mathbf{p}_1|}{2\pi^2\hbar^3} \tag{251}$$

$$= s_1 f_1 \frac{|\mathbf{p}_1|\epsilon_1\,\mathrm{d}\epsilon_1}{2\pi^2\hbar^3} \tag{252}$$

Inserting this into equation 249 gives

$$\mathrm{d}\Gamma_{n1}(\epsilon_1, \epsilon_2) = \frac{1}{2\pi^2\hbar^3} s_1 f_1\, \sigma(\epsilon_1, \epsilon_2)\, |\mathbf{p}_1|^2\, \delta(\epsilon_2 - \epsilon_1 - Q)\,\mathrm{d}\epsilon_1\,\mathrm{d}\epsilon_2. \tag{253}$$

Note that with the approximations made so far,

$$\mathcal{A}^2 = 128 G_F^2 m_p m_n \epsilon_\nu \epsilon_e \tag{254}$$

$$= 128 G_F^2 m_p m_n \epsilon_1 \epsilon_2. \tag{255}$$

Finally, inserting this and equation 246 into equation 253 gives

$$\mathrm{d}\Gamma_{n1}(\epsilon_1, \epsilon_2) = \frac{2}{G_F^2\pi^3\hbar^3} \frac{m_p^2}{m_n^2} f_1\,(1 - f_2)\,\epsilon_1\epsilon_2 |\mathbf{p}_1||\mathbf{p}_2| \delta(\epsilon_2 - \epsilon_1 - Q)\,\mathrm{d}\epsilon_1\,\mathrm{d}\epsilon_2 \tag{256}$$

Utilizing energy conservation via the delta-function, we integrate over the lepton energies to find the total rate:

$$\Gamma_{n1}^{\mathrm{tot}} = \frac{2G_F^2}{\pi^3\hbar^3} \frac{m_p^2}{m_n^2} \int_{m_1}^{\infty} \frac{\mathrm{d}\epsilon_1 \epsilon_1^2 \sqrt{1 - m_1^2/\epsilon_1^2}}{1 + \exp\{\epsilon_1/T_1\}} \frac{(\epsilon_1 + Q)^2 \sqrt{1 - m_2^2/(\epsilon_1 + Q)^2}}{1 + \exp\{-(\epsilon_1 + Q)/T_2\}} \tag{257}$$

The dynamics beind all three processes, neutron-neutrino scattering, neutron-positron scattering, and neutron-decay, are the same. What is different is the kinematics, the

relation between the lepton energies and Q. We can write the sum of the interaction rates that "convert" neutrons into protons as $\Gamma_{n\to p} = \Gamma_{n\nu} + \Gamma_{n\bar{e}} + \Gamma_{n\,\text{decay}}$ and express this as the integral in equation 257 with expanded limits to generalize the kinematics:

$$\Gamma_{n\to p} = \frac{2G_F^2}{\pi^3}\frac{m_p^2}{m_n^2}\int_{-\infty}^{\infty}\frac{\mathrm{d}\mathcal{E}\mathcal{E}^2}{1+\exp\{\mathcal{E}/T_\nu\}}\frac{(\mathcal{E}+Q)^2\sqrt{1-m_e^2/(\mathcal{E}+Q)^2}}{1+\exp\{-(\mathcal{E}+Q)/T\}}, \tag{258}$$

where the integration leaves out the interval where the square root is imaginary.

C Sommerfeld Factor Computation

We want to numerically solve the differential equation

$$\frac{\mathrm{d}^2\chi}{\mathrm{d}\rho^2} + \left(1 + \frac{2\eta}{\rho}e^{-\epsilon_\phi\eta\rho}\right)\chi(\rho) = 0 \tag{259}$$

with the initial conditions $\chi(\rho_0) = C\rho_0$ and $\chi'(\rho_0) = C$, where C is a constant and $\rho_0 = \mathcal{O}(10^{-5})$ (this is to avoid the divergence). The correct choice of C squared is in fact the Sommerfeld factor S. To find the correct choice, we take an arbitrary "trial" value for C (that may as well be unity) and then evolve the function χ from $\rho = \rho_0$ to $\rho = \rho_{\text{sin}}$ that is large enough so that the function exhibits sinesoidal behavior as $\sin(\rho+\delta_l)$. Typically, $\rho_{\text{sin}} \approx 100$. The amplitude squared of $\chi(\rho)$ is measured by calculating $A^2 = \chi^2(\rho_{\text{sin}}) + \chi^2(\rho_{\text{sin}} - \pi/2)$, which should be unity. The normalization of $\chi(\rho)$ is adjusted by $1/A$, so the Sommerfeld factor is

$$S = C^2/A^2. \tag{260}$$

We used the GNU Science Library to perform the computation, splitting the original equation into two first-order equations:

$$\chi' = \Phi, \tag{261}$$

$$\Phi' = -\left(1 + \frac{2\eta}{\rho}e^{-\epsilon_\phi\eta\rho}\right)\chi. \tag{262}$$

D Dimensionless Distribution Functions in Plots

We describe the dimensionless scaling used in Section 7.2 to illustrate the derived distribution functions. Note that this scaling scheme is not else elsewhere in this work, just to compare DFs of different profiles. The dimensionless radius and density are defined in this case as [60]

$$x \equiv r/r_{\rm vir}, \tag{263}$$

$$\tilde{\rho}\left(x\right) = \rho\left(xr_{\rm vir}\right)/\rho_{\rm s}, \tag{264}$$

where $r_{\rm vir}$ is the virial radius and $\rho_{\rm s}$ is the scale density. For example, the NFW model has

$$\tilde{\rho}\left(x\right) = \left[cx\left(1+cx\right)^2\right]^{-1}, \tag{265}$$

where c is the concentration. The quantity $g(c)$ is defined for a general profile by

$$\begin{aligned} g(c) &\equiv \int_0^1 {\rm d}x x^2 \tilde{\rho}(x;c) \\ &= M_{\rm vir}/4\pi r_{\rm s}^3 \rho_{\rm s}. \end{aligned} \tag{266}$$

The dimensionless energy $\tilde{\mathcal{E}}$ is given in terms of $\mathcal{E} \equiv \Psi - v^2/2$ by

$$\mathcal{E} = \frac{GM_{\rm vir}}{r_{\rm vir}}\,\tilde{\mathcal{E}}, \tag{267}$$

where G is the gravitational constant and $M_{\rm vir}$ is the virial mass.

E Details of Jeans Analysis

From the PPSD power-law in equation 189 and the general density profile in equation 68, we have the radial velocity dispersion

$$\sigma_{\rm r}^2(x) = \sigma_{\rm r,s}^2\left[x^{-\gamma_0+\alpha}\left(\frac{2}{1+x}\right)^{\gamma_\infty-\gamma_0}\right]^{2/3}. \tag{268}$$

For the NFW profile this is

$$\sigma_{\rm r}^2(x) = \sigma_{\rm r,s}^2 \left(\frac{4x^{-1+\alpha}}{(1+x)^2} \right)^{2/3}. \tag{269}$$

The expression for the anisotropy parameter $\beta\,(x)$ in equation 190 is general. Specific to the case of the NFW profile, it is

$$\beta(x) = \frac{5+15x}{6+6x} - \frac{\alpha}{3} - \Sigma^{-2} \cdot 3x^{-(2\alpha+1)/3} \left(\frac{1+x}{2} \right)^{1/3} \left[-x + (1+x)\log(1+x) \right]. \tag{270}$$

with $\Sigma^2 \equiv \sigma_{\rm r,s}^2/(4\pi G r_{\rm s}^2 \rho_{\rm s}/3)$. The upper limit on Σ in the case of a NFW profile is

$$\Sigma^2 \leq \Sigma_{\rm max}^2 \equiv \frac{3^2 \cdot 2^{2/3}(1+c)^{4/3}\left[-c + (1+c)\log(1+c) \right]}{c^{(1+2\alpha)/3}\left[9c - 2\alpha(1+c) - 1 \right]}. \tag{271}$$

This translates to the upper limits on $\sigma_{\rm r,s}$ and $\sigma_{\rm r,\odot}$:

$$\sigma_{\rm r,s}^2 \leq \sigma_{\rm r,s,max}^2 \equiv (4\pi G r_{\rm s}^2 \rho_{\rm s}/3)\Sigma_{\rm max}^2\,(\alpha, c)\,, \tag{272}$$

$$\sigma_{\rm r,\odot}^2 \leq \sigma_{\rm r,\odot,max}^2 \equiv \left(\frac{4\,x_\odot^{-1+\alpha}}{(1+x_\odot)^2} \right)^{2/3} \sigma_{\rm r,s,max}^2\,(\alpha, M_{\rm vir}, r_{\rm s}, c)\,. \tag{273}$$

Finally, because $\sigma_{\rm tot,\odot}$ increases monotonically with $\sigma_{\rm r,\odot}$, its upper limit is

$$\sigma_{\rm tot,\odot}^2 \leq (3 - 2\beta_\odot)\sigma_{\rm r,\odot,max}^2, \tag{274}$$

which depends on α, $M_{\rm vir}$, $r_{\rm s}$, c, and $r_\odot$.

F Asymptotic Behavior of the Anisotropy Profile

We split the function for the NFW anisotropy parameter in equation 270 into two parts, so $\beta(x) = A(x) + B(x)$, with

$$A(x) = \frac{5 + 15x}{6 + 6x} - \frac{\alpha}{3}, \tag{275}$$

$$B(x) = -\Sigma^{-2} \cdot 3x^{-(2\alpha+1)/3} \left(\frac{1+x}{2}\right)^{1/3} \left[-x + (1+x)\log(1+x)\right]. \tag{276}$$

The first part has simple asymptotic limits

$$A(x) \to \begin{cases} (5 - 2\alpha)/6 & \text{for } x \to 0 \\ (15 - 2\alpha)/6 & \text{for } x \to \infty \end{cases} \tag{277}$$

while the second is more complicated. In the limit $x \to 0$, we have

$$B(x) \to \begin{cases} -\infty & \text{if } \alpha > 5/2 \\ -2^{-4/3} \cdot 3 \times \Sigma^{-2} & \text{if } \alpha = 5/2 \\ 0 & \text{if } \alpha < 5/2 \end{cases} \tag{278}$$

and in the limit $x \to \infty$, we have

$$B(x) \to \begin{cases} 0 & \text{if } \alpha > 3/2 \\ -\infty & \text{if } \alpha \leq 3/2 \end{cases} \tag{279}$$

As long as $3/2 < \alpha < 5/2$, the extreme values of $\beta(x)$ are determined solely by α.

G Details of the Anisotropic Mao *et al.* Distribution

We use the velocity distribution in equation 195 to model the local velocity distribution, with $p = 1.5$ from [79] and with $v_{\rm esc} = 550.7\,{\rm km/s}$ from [43]. The choice of parameters $v_{\rm r,0}$ and $v_{\rm t,0}$ determine the velocity dispersion and anisotropy parameter. Figure 35 plots

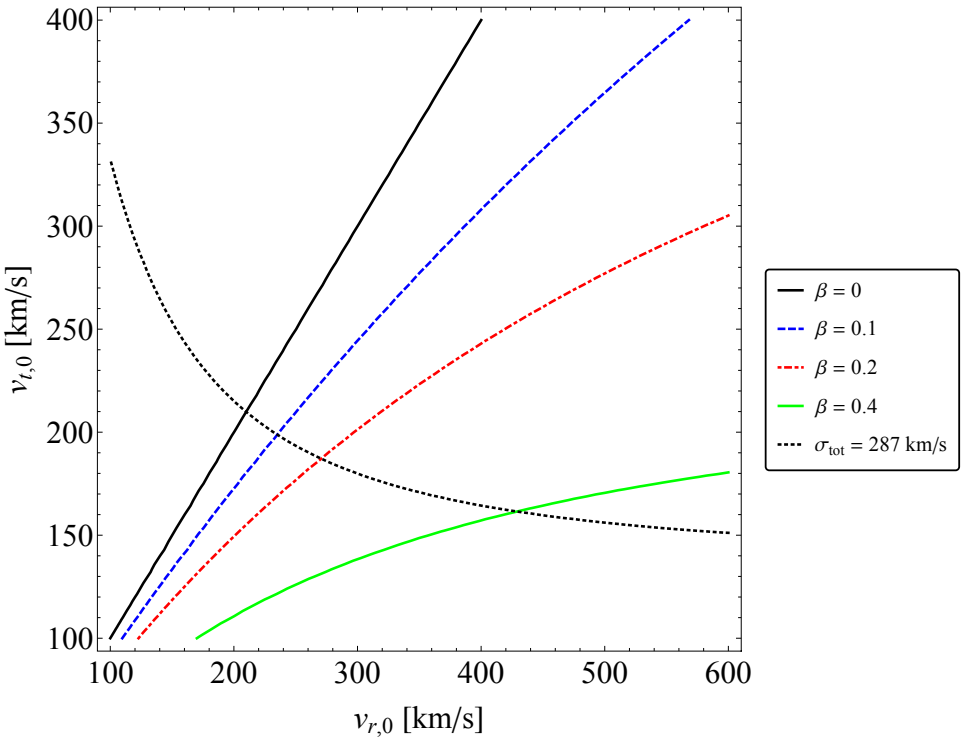

Figure 35: Contours of the parameters $v_{r,0}$ and $v_{t,0}$ that give the specified values of the anisotropy parameter or total velocity dispersion.

contours that give the specified value of β or σ_{tot}. In this work we choose $\sigma_{tot} = 287\,\mathrm{km/s}$ as the mean value [43]. For the isotropic case, this implies $v_{r,0} = v_{t,0} = 209.8\,\mathrm{km/s}$; for the anisotropic case, with $\beta = 0.2$, this implies $v_{r,0} = 270.4\,\mathrm{km/s}$ and $v_{t,0} = 187.1\,\mathrm{km/s}$.

H Anisotropic NFW Distribution Functions

The dimensionless density profile proposed by Navarro, Frenk, and White and its potential energy function are

$$\rho(x) = x^{-1}(1+x)^{-2}, \tag{280}$$

$$\Psi(x) = \frac{\log(1+x)}{x}. \tag{281}$$

We may invert the latter equation as follows:

$$\Psi = \log(1+x)/x, \tag{282}$$

$$-\Psi - \Psi x = -\Psi - \log(1+x), \tag{283}$$

$$-\Psi = (-\Psi - \log(1+x))(1+x)^{-1}, \tag{284}$$

$$-\Psi e^{-\Psi} = (-\Psi - \log(1+x))e^{-\Psi - \log(1+x)}. \tag{285}$$

Here we recognize the form of the definition of the Lambert W function. Since Ψ varies from zero to one, the quantity $-\Psi - \log(1+x)$ varies from negative infinity to negative one. This means that the W function we use takes its values on the lower branch:

$$-\Psi - \log(1+x) = W_{-1}(-\Psi e^{-\Psi}), \tag{286}$$

$$\log(1+x) = -\Psi - W_{-1}(-\Psi e^{-\Psi}), \tag{287}$$

$$\Psi x = -\Psi - W_{-1}(-\Psi e^{-\Psi}), \tag{288}$$

$$x = \frac{-\Psi - W_{-1}(-\Psi e^{-\Psi})}{\Psi}. \tag{289}$$

It is convenient to define the function

$$A(\Psi) = W_{-1}(-\Psi e^{-\Psi}), \tag{290}$$

which has the following derivative:

$$\frac{dA}{d\Psi} = W'_{-1}(-\Psi e^{-\Psi}) \frac{d}{d\Psi}\left[-\Psi e^{-\Psi}\right] \tag{291}$$

$$= \frac{A}{1+A}\frac{1-\Psi}{\Psi}. \tag{292}$$

Using what we have found, the density can be written in terms of the potential energy:

$$\rho(\Psi) = \left[\left(-1 - \frac{A}{\Psi}\right)\left(\frac{A}{\Psi}\right)^2\right]^{-1} \tag{293}$$

$$= \frac{\Psi^3}{A^2(-\Psi - A)} \tag{294}$$

In the cases of constant anisotropy with β equal to positive or negative one-half, the distribution function is easy to derive using $\rho(\Psi)$:

$$f_{\beta=\frac{1}{2}}(\mathcal{E}, L) = \frac{1}{L}\frac{1}{\pi^2}\frac{\mathcal{E}}{A^2}\frac{\mathcal{E}+A}{1+A}, \tag{295}$$

$$f_{\beta=-\frac{1}{2}}(\mathcal{E}, L) = \frac{L}{\pi^2}\frac{\mathcal{E}}{A^2}\frac{\mathcal{E}^2\,(\mathcal{E}+2A^2)}{(1+A)(\mathcal{E}+A)^2}. \tag{296}$$

References

[1] C. Aalseth et al. Results from a Search for Light-Mass Dark Matter with a P-type Point Contact Germanium Detector. *Phys.Rev.Lett.*, 106:131301, 2011.

[2] P. Ade et al. Planck 2013 results. I. Overview of products and scientific results. 2013.

[3] P. Ade et al. Planck 2013 results. XVI. Cosmological parameters. 2013.

[4] P. Ade et al. BICEP2 I: Detection Of B-mode Polarization at Degree Angular Scales. 2014.

[5] O. Adriani, G. Barbarino, G. Bazilevskaya, R. Bellotti, M. Boezio, et al. A new measurement of the antiproton-to-proton flux ratio up to 100 GeV in the cosmic radiation. *Phys.Rev.Lett.*, 102:051101, 2009.

[6] O. Adriani et al. An anomalous positron abundance in cosmic rays with energies 1.5-100 GeV. *Nature*, 458:607–609, 2009.

[7] M. Aguilar et al. First Result from the Alpha Magnetic Spectrometer on the International Space Station: Precision Measurement of the Positron Fraction in Primary Cosmic Rays of 0.5350 GeV. *Phys.Rev.Lett.*, 110(14):141102, 2013.

[8] Z. Ahmed et al. Dark Matter Search Results from the CDMS II Experiment. *Science*, 327:1619–1621, 2010.

[9] D. Akerib et al. First results from the LUX dark matter experiment at the Sanford Underground Research Facility. 2013.

[10] D. Albornoz Vasquez, G. Belanger, J. Billard, and F. Mayet. Probing neutralino dark matter in the MSSM & the NMSSM with directional detection. *Phys.Rev.*, D85:055023, 2012.

[11] N. Amorisco, A. Agnello, and N. Evans. The core size of the Fornax dwarf Spheroidal. 2012.

[12] J. H. An and N. W. Evans. A theorem on central velocity dispersions. *Astrophys.J.*, 701:1500–1505, 2009.

[13] G. Angloher, M. Bauer, I. Bavykina, A. Bento, C. Bucci, et al. Results from 730 kg days of the CRESST-II Dark Matter Search. *Eur.Phys.J.*, C72:1971, 2012.

[14] R. Angulo, V. Springel, S. White, A. Jenkins, C. Baugh, et al. Scaling relations for galaxy clusters in the Millennium-XXL simulation. 2012.

[15] E. Aprile et al. Analysis of the XENON100 Dark Matter Search Data. *Astropart.Phys.*, 54:11–24, 2014.

[16] E. Armengaud et al. Final results of the EDELWEISS-II WIMP search using a 4-kg array of cryogenic germanium detectors with interleaved electrodes. *Phys.Lett.*, B702:329–335, 2011.

[17] E. Aver, K. A. Olive, R. Porter, and E. D. Skillman. The primordial helium abundance from updated emissivities. *JCAP*, 1311:017, 2013.

[18] E. Aver, K. A. Olive, and E. D. Skillman. An MCMC determination of the primordial helium abundance. *JCAP*, 1204:004, 2012.

[19] K. Begeman. H I rotation curves of spiral galaxies. I - NGC 3198. *Astron.Astrophys.*, 223:47–60, 1989.

[20] K. Begeman, A. Broeils, and R. Sanders. Extended rotation curves of spiral galaxies: Dark haloes and modified dynamics. *Mon.Not.Roy.Astron.Soc.*, 249:523, 1991.

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

[22] J. Beringer et al. Review of Particle Physics (RPP). *Phys.Rev.*, D86:010001, 2012.

[23] R. Bernabei et al. New results from DAMA/LIBRA. *Eur.Phys.J.*, C67:39–49, 2010.

[24] J. Bernstein, L. S. Brown, and G. Feinberg. The Cosmological Heavy Neutrino Problem Revisited. *Phys.Rev.*, D32:3261, 1985.

[25] G. Bertone, M. Cirelli, A. Strumia, and M. Taoso. Gamma-ray and radio tests of the e+e- excess from DM annihilations. *JCAP*, 0903:009, 2009.

[26] G. Bertone, D. Hooper, and J. Silk. Particle dark matter: Evidence, candidates and constraints. *Phys.Rept.*, 405:279–390, 2005.

[27] H. Bethe. Energy production in stars. *Phys.Rev.*, 55:434–456, 1939.

[28] P. Bett, V. Eke, C. S. Frenk, A. Jenkins, J. Helly, et al. The spin and shape of dark matter haloes in the Millennium simulation of a lambda-CDM universe. *Mon.Not.Roy.Astron.Soc.*, 376:215–232, 2007.

[29] J. Binney and S. Tremaine. *Galactic Dynamics: Second Edition*. Princeton University Press, 2008.

[30] J. Bond, G. Efstathiou, and J. Silk. Massive Neutrinos and the Large Scale Structure of the Universe. *Phys.Rev.Lett.*, 45:1980–1984, 1980.

[31] J. Bond and A. Szalay. The Collisionless Damping of Density Fluctuations in an Expanding Universe. *Astrophys.J.*, 274:443–468, 1983.

[32] A. Bosma. 21-cm line studies of spiral galaxies. 2. The distribution and kinematics of neutral hydrogen in spiral galaxies of various morphological types. *Astron.J.*, 86:1825, 1981.

[33] M. Boylan-Kolchin, V. Springel, S. D. White, A. Jenkins, and G. Lemson. Resolving Cosmic Structure Formation with the Millennium-II Simulation. *Mon.Not.Roy.Astron.Soc.*, 398:1150, 2009.

[34] N. Bozorgnia, R. Catena, and T. Schwetz. Anisotropic dark matter distribution functions and impact on WIMP direct detection. 2013.

[35] B. Burch and R. Cowsik. Properties of Galactic Dark Matter: Constraints from Astronomical Observations. *Astrophys.J.*, 779:35, 2013.

[36] B. P. Burch. *Galactic Dark Matter*. PhD thesis, Washington University in St. Louis, 2013.

[37] A. Burkert. The Structure of dark matter halos in dwarf galaxies. *IAU Symp.*, 171:175, 1996.

[38] S. Campbell and B. Dutta. Effects of P-wave Annihilation on the Angular Power Spectrum of Extragalactic Gamma-rays from Dark Matter Annihilation. *Phys.Rev.*, D84:075004, 2011.

[39] V. F. Cardone, E. Piedipalumbo, and C. Tortora. Spherical galaxy models with power law logarithmic slope. *Mon.Not.Roy.Astron.Soc.*, 358:1325–1336, 2005.

[40] R. Carlberg, H. Yee, E. Ellingson, R. Abraham, P. Gravel, et al. Galaxy cluster virial masses and Omega. *Astrophys.J.*, 462:32, 1996.

[41] R. G. Carlberg, H. Yee, and E. Ellingson. The Average mass and light profiles of galaxy clusters. *Astrophys.J.*, 478:462, 1997.

[42] R. Catena and P. Ullio. A novel determination of the local dark matter density. *JCAP*, 1008:004, 2010.

[43] R. Catena and P. Ullio. The local dark matter phase-space density and impact on WIMP direct detection. *JCAP*, 1205:005, 2012.

[44] M. Cirelli, P. Panci, and P. D. Serpico. Diffuse gamma ray constraints on annihilating or decaying Dark Matter after Fermi. *Nucl.Phys.*, B840:284–303, 2010.

[45] R. Cowsik and B. Burch. Analysis of the Positron Fraction and the Spectrum of the Electronic Component in Cosmic Rays. 2009.

[46] R. Cowsik and B. Burch. On the Positron Fraction and the Spectrum of the Electronic Component in Cosmic Rays. 2009.

[47] R. Cowsik and B. Burch. Positron fraction in cosmic rays and models of cosmic-ray propagation. *Phys.Rev.*, D82:023009, 2010.

[48] R. Cowsik, B. Burch, and T. Madziwa-Nussinov. The origin of the spectral intensities of cosmic-ray positrons. *Astrophys.J.*, 786:124, 2014.

[49] R. Cowsik and J. McClelland. An Upper Limit on the Neutrino Rest Mass. *Phys.Rev.Lett.*, 29:669–670, 1972.

[50] P. Cuddeford. An analytic inversion for anisotropic spherical galaxies. MNRAS, 253:414–426, Dec. 1991.

[51] T. Daylan, D. P. Finkbeiner, D. Hooper, T. Linden, S. K. N. Portillo, et al. The Characterization of the Gamma-Ray Signal from the Central Milky Way: A Compelling Case for Annihilating Dark Matter. 2014.

[52] W. de Blok, S. S. McGaugh, A. Bosma, and V. C. Rubin. Mass density profiles of LSB galaxies. *Astrophys.J.*, 552:L23–L26, 2001.

[53] W. Dehnen and D. McLaughlin. Dynamical insight into dark-matter haloes. *Mon.Not.Roy.Astron.Soc.*, 363:1057–1068, 2005.

[54] W. Dehnen and J. Read. N-body simulations of gravitational dynamics. *Eur.Phys.J.Plus*, 126:55, 2011.

[55] F. Donato, D. Maurin, P. Brun, T. Delahaye, and P. Salati. Constraints on WIMP Dark Matter from the High Energy PAMELA $\bar{p}/p$ data. *Phys.Rev.Lett.*, 102:071301, 2009.

[56] A. Drukier, K. Freese, and D. Spergel. Detecting Cold Dark Matter Candidates. *Phys.Rev.*, D33:3495–3508, 1986.

[57] N. et al. *Science*, 272:846–849, 1996.

[58] N. Evans, F. Ferrer, and S. Sarkar. A 'Baedecker' for the dark matter annihilation signal. *Phys.Rev.*, D69:123501, 2004.

[59] B. Famaey and S. McGaugh. Modified Newtonian Dynamics (MOND): Observational Phenomenology and Relativistic Extensions. *Living Rev.Rel.*, 15:10, 2012.

[60] F. Ferrer and D. R. Hunter. The impact of the phase-space density on the indirect detection of dark matter. *JCAP*, 1309:005, 2013.

[61] D. Fixsen, E. Cheng, J. Gales, J. C. Mather, R. Shafer, et al. The Cosmic Microwave Background spectrum from the full COBE FIRAS data set. *Astrophys.J.*, 473:576, 1996.

[62] P. Gondolo. Recoil momentum spectrum in directional dark matter detectors. *Phys.Rev.*, D66:103513, 2002.

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

[64] P. Gondolo and G. B. Gelmini. Halo independent comparison of direct dark matter detection data. *JCAP*, 1212:015, 2012.

[65] P. Gondolo and J. Silk. Dark matter annihilation at the galactic center. *Phys.Rev.Lett.*, 83:1719–1722, 1999.

[66] M. W. Goodman and E. Witten. Detectability of Certain Dark Matter Candidates. *Phys.Rev.*, D31:3059, 1985.

[67] A. W. Graham, D. Merritt, B. Moore, J. Diemand, and B. Terzic. Empirical models for Dark Matter Halos. I. Nonparametric Construction of Density Profiles and Comparison with Parametric Models. *Astron.J.*, 132:2685–2700, 2006.

[68] D. Griffiths. Introduction to elementary particles. 2008.

[69] S. H. Hansen and B. Moore. A Universal density slope - velocity anisotropy relation for relaxed structures. *New Astron.*, 11:333, 2006.

[70] G. Hinshaw et al. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Data Processing, Sky Maps, and Basic Results. *Astrophys.J.Suppl.*, 180:225–245, 2009.

[71] F. Hohl. Numerical Experiments with a Disk of Stars. ApJ, 168:343, Sept. 1971.

[72] E. Holmberg. *Astroph.J.*, 94:385, 1941.

[73] F. Hoyle. On Nuclear Reactions Occuring in Very Hot STARS.I. the Synthesis of Elements from Carbon to Nickel. ApJS, 1:121, Sept. 1954.

[74] D. R. Hunter. Derivation of the anisotropy profile, constraints on the local velocity dispersion, and implications for direct detection. *JCAP*, 1402:023, 2013.

[75] Y. I. Izotov, T. Thuan, and G. Stasinska. The primordial abundance of He-4: A self-consistent empirical analysis of systematic effects in a large sample of low-metallicity HII regions. *Astrophys.J.*, 662:15–38, 2007.

[76] G. Jungman, M. Kamionkowski, and K. Griest. Supersymmetric dark matter. *Phys.Rept.*, 267:195–373, 1996.

[77] J. Katz. Thermodynamics and selfgravitating systems. *Found.Phys.*, 33:223–269, 2003.

[78] E. Komatsu et al. Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *Astrophys.J.Suppl.*, 192:18, 2011.

[79] M. Kuhlen, A. Pillepich, J. Guedes, and P. Madau. The Distribution of Dark Matter in the Milky Way's Disk. 2013.

[80] M. Laine and Y. Schroder. Quark mass thresholds in QCD thermodynamics. *Phys.Rev.*, D73:085009, 2006.

[81] V. Lebedenko, H. Araujo, E. Barnes, A. Bewick, R. Cashmore, et al. Result from the First Science Run of the ZEPLIN-III Dark Matter Search Experiment. *Phys.Rev.*, D80:052010, 2009.

[82] B. W. Lee and S. Weinberg. Cosmological Lower Bound on Heavy Neutrino Masses. *Phys.Rev.Lett.*, 39:165–168, 1977.

[83] J. L. Linsky, A. Brown, K. Gayley, A. Diplas, B. D. Savage, et al. Goddard high-resolution spectrograph observations of the local interstellar medium and the deuterium/hydrogen ratio along the line of sight toward Capella. *Astrophys.J.*, 402:694–709, 1993.

[84] A. D. Ludlow, J. F. Navarro, M. Boylan-Kolchin, V. Springel, A. Jenkins, et al. The density and pseudo-phase-space density profiles of cold dark matter haloes. *Mon.Not.Roy.Astron.Soc.*, 415:3895–3902, 2011.

[85] D. Lynden-Bell. Stellar dynamics: Exact solution of the self-gravitation equation. MNRAS, 123:447, 1962.

[86] D. Lynden-Bell. Negative specific heat in astronomy, physics and chemistry. *Physica*, A263:293–304, 1999.

[87] C.-P. Ma, P. Chang, and J. Zhang. Is the Radial Profile of the Phase-Space Density of Dark Matter Halos a Power-Law? 2009.

[88] Y.-Y. Mao, L. E. Strigari, R. H. Wechsler, H.-Y. Wu, and O. Hahn. Halo-to-Halo Similarity and Scatter in the Velocity Distribution of Dark Matter. *Astrophys.J.*, 764:35, 2013.

[89] M. Milgrom. A modification of the Newtonian dynamics: implications for galaxy systems. *Astrophys.J.*, 270:384–389, 1983.

[90] M. Milgrom. MOND: A pedagogical review. *Acta Phys.Polon.*, B32:3613, 2001.

[91] R. H. Miller and K. H. Prendergast. Stellar Dynamics in a Discrete Phase Space. ApJ, 151:699, Feb. 1968.

[92] R. H. Miller, K. H. Prendergast, and W. J. Quirk. Numerical Experiments on Spiral Structure. ApJ, 161:903, Sept. 1970.

[93] B. Moore, T. R. Quinn, F. Governato, J. Stadel, and G. Lake. Cold collapse and the core catastrophe. *Mon.Not.Roy.Astron.Soc.*, 310:1147–1152, 1999.

[94] I. Moskalenko and A. Strong. Production and propagation of cosmic ray positrons and electrons. *Astrophys.J.*, 493:694–707, 1998.

[95] J. F. Navarro, C. S. Frenk, and S. D. White. The Structure of cold dark matter halos. *Astrophys.J.*, 462:563–575, 1996.

[96] J. F. Navarro, E. Hayashi, C. Power, A. Jenkins, C. S. Frenk, et al. The Inner structure of Lambda-CDM halos 3: Universality and asymptotic slopes. *Mon.Not.Roy.Astron.Soc.*, 349:1039, 2004.

[97] H. Nussbaumer and L. Bieri. Who discovered the expanding universe? 2011.

[98] K. A. Olive and E. D. Skillman. A Realistic determination of the error on the primordial helium abundance: Steps toward non-parametric nebular helium abundances. *Astrophys.J.*, 617:29, 2004.

[99] L. P. Osipkov. Spherical systems of gravitating bodies with an ellipsoidal velocity distribution. *Soviet Astronomy Letters*, 5:42–44, 1979.

[100] J. P. Ostriker and P. J. E. Peebles. A Numerical Study of the Stability of Flattened Galaxies: or, can Cold Galaxies Survive? ApJ, 186:467–480, Dec. 1973.

[101] T. Padmanabhan. Statistical Mechanics of Gravitating Systems. *Phys.Rept.*, 188:285, 1990.

[102] M. Peimbert, V. Luridiana, and A. Peimbert. Revised Primordial Helium Abundance Based on New Atomic Data. *Astrophys.J.*, 666:636–646, 2007.

[103] A. A. Penzias and R. W. Wilson. A Measurement of excess antenna temperature at 4080-Mc/s. *Astrophys.J.*, 142:419–421, 1965.

[104] S. Profumo. Dissecting cosmic-ray electron-positron data with Occam's Razor: the role of known Pulsars. *Central Eur.J.Phys.*, 10:1–31, 2011.

[105] S. Riemer-Sorensen, C. Blake, D. Parkinson, T. M. Davis, S. Brough, et al. The WiggleZ Dark Energy Survey: Cosmological neutrino mass constraint from blue high-redshift galaxies. *Phys.Rev.*, D85:081101, 2012.

[106] B. Robertson and A. Zentner. Dark Matter Annihilation Rates with Velocity-Dependent Annihilation Cross Sections. *Phys.Rev.*, D79:083525, 2009.

[107] L. Rosenberg and K. van Bibber. Searches for invisible axions. *Phys.Rept.*, 325:1–39, 2000.

[108] V. Rubin, N. Thonnard, and J. Ford, W.K. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/. *Astrophys.J.*, 238:471, 1980.

[109] M. S. Longair. *High Energy Astrophysics*. Cambridge University Press, 1992.

[110] R. Sachs and A. Wolfe. Perturbations of a cosmological model and angular variations of the microwave background. *Astrophys.J.*, 147:73–90, 1967.

[111] L. Sadeghian, F. Ferrer, and C. M. Will. Dark matter distributions around massive black holes: A general relativistic analysis. 2013.

[112] L. Sbordone, P. Bonifacio, E. Caffau, H.-G. Ludwig, N. Behara, et al. The metal-poor end of the Spite plateau. 1: Stellar parameters, metallicities and lithium abundances. *Astron.Astrophys.*, 522:A26, 2010.

[113] K. B. Schmidt, S. H. Hansen, J. H. An, L. L. Williams, and A. V. Macci'o. Dark Matter Angular Momentum Profile from the Jeans Equation. *Astrophys.J.*, 694:893–901, 2009.

[114] X.-D. Shi and G. M. Fuller. A New dark matter candidate: Nonthermal sterile neutrinos. *Phys.Rev.Lett.*, 82:2832–2835, 1999.

[115] P. Sikivie. Axion Cosmology. *Lect.Notes Phys.*, 741:19–50, 2008.

[116] J. Silk. Cosmic black body radiation and galaxy formation. *Astrophys.J.*, 151:459–471, 1968.

[117] G. F. Smoot, C. Bennett, A. Kogut, E. Wright, J. Aymon, et al. Structure in the COBE differential microwave radiometer first year maps. *Astrophys.J.*, 396:L1–L5, 1992.

[118] V. Springel, J. Wang, M. Vogelsberger, A. Ludlow, A. Jenkins, et al. The Aquarius Project: the subhalos of galactic halos. *Mon.Not.Roy.Astron.Soc.*, 391:1685–1711, 2008.

[119] M. Srednicki, R. Watkins, and K. A. Olive. Calculations of Relic Densities in the Early Universe. *Nucl.Phys.*, B310:693, 1988.

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

[121] L. E. Strigari and R. Trotta. Reconstructing WIMP Properties in Direct Detection Experiments Including Galactic Dark Matter Distribution Uncertainties. *JCAP*, 0911:019, 2009.

[122] M. Taoso, G. Bertone, and A. Masiero. Dark Matter Candidates: A Ten-Point Test. *JCAP*, 0803:022, 2008.

[123] J. E. Taylor and J. F. Navarro. The Phase - space density profiles of cold dark matter halos. *Astrophys.J.*, 563:483–488, 2001.

[124] S. A. Thomas, F. B. Abdalla, and O. Lahav. Upper Bound of 0.28eV on the Neutrino Masses from the Largest Photometric Redshift Survey. *Phys.Rev.Lett.*, 105:031301, 2010.

[125] P. Tisserand et al. Limits on the Macho Content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds. *Astron.Astrophys.*, 469:387–404, 2007.

[126] C. Tsallis. Possible Generalization of Boltzmann-Gibbs Statistics. *J.Statist.Phys.*, 52:479–487, 1988.

[127] C. Tsallis. *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World*. Springer, 2009.

[128] C. Tsallis, F. Baldovin, R. Cerbino, and P. Pierobon. Introduction to Nonextensive Statistical Mechanics and Thermodynamics. *eprint arXiv:cond-mat/0309093*, Sept. 2003.

[129] P. Ullio and M. Kamionkowski. Velocity distributions and annual modulation signatures of weakly interacting massive particles. *JHEP*, 0103:049, 2001.

[130] P. Ullio, H. Zhao, and M. Kamionkowski. A Dark matter spike at the galactic center? *Phys.Rev.*, D64:043504, 2001.

[131] E. Van Hese, M. Baes, and H. Dejonghe. The dynamical structure of dark matter halos with universal properties. *Astrophys.J.*, 690:1280–1291, 2009.

[132] E. Visbal, A. Loeb, and L. Hernquist. A Simple Model for the Density Profiles of Isolated Dark Matter Halos. 2012.

[133] M. Vogelsberger, A. Helmi, V. Springel, S. D. White, J. Wang, et al. Phase-space structure in the local dark matter distribution and its signature in direct detection experiments. *Mon.Not.Roy.Astron.Soc.*, 395:797–811, 2009.

[134] M. G. Walker and J. Penarrubia. A Method for Measuring (Slopes of) the Mass Profiles of Dwarf Spheroidal Galaxies. *Astrophys.J.*, 742:20, 2011.

[135] I. Wasserman. Possibility of Detecting Heavy Neutral Fermions in the Galaxy. *Phys.Rev.*, D33:2071–2078, 1986.

[136] S. Weinberg. A New Light Boson? *Phys.Rev.Lett.*, 40:223–226, 1978.

[137] S. Weinberg. Cosmology. 2008.

[138] F. Wilczek. Problem of Strong p and t Invariance in the Presence of Instantons. *Phys.Rev.Lett.*, 40:279–282, 1978.

[139] A. Zait, Y. Hoffman, and I. Shlosman. Dark Matter Halos: Velocity Anisotropy – Density Slope Relation. 2007.

[140] G.-B. Zhao, S. Saito, W. J. Percival, A. J. Ross, F. Montesano, et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: weighing the neutrino mass using the galaxy power spectrum of the CMASS sample. *Mon.Not.Roy.Astron.Soc.*, 436:2038–2053, 2013.

[141] F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helv.Phys.Acta*, 6:110–127, 1933.